Royal Holloway
University of London

# Discussion Paper Series

2004 – 12

Department of Economics
Royal Holloway College
University of London
Egham TW20 0EX

# Getting Hitched: The Equilibrium Marriage Market Behaviour of a British Cohort*

By Dan Anderberg

*Royal Holloway University of London, CEPR, CESifo and IFS*

December 8, 2003

**Abstract**

This paper structurally estimates an equilibrium model of the marriage market using data from the British Cohort Study. The results suggest that both search frictions and selective behaviour determine who marries whom and when, and that education is the most important determinant of an individual's attractiveness. Simulations of the model, however, only replicate a fraction of the correlations of partners' characteristics observed in the data, leading us to question some assumptions commonly made in the literature. Simulations also suggest that increased participation in education may have contributed to increased sorting into partnerships.

## I Introduction

How individuals sort into households is of fundamental importance to the economy in terms of inequality, intergenerational mobility and policy.[1] A substantial literature, mainly descriptive in nature, has documented the degree of sorting in the process of partnership formation (see e.g. Mare, 1991, Lam, 1988). While this literature has been very useful in mapping out on which characteristics sorting occurs, mainly education and income, it gives little insight into

*I would like to thank Richard Blundell, Martin Browning, John Ermisch, Alissa Goodman, Costas Meghir, Shannon Seitz, and seminar participants in Berlin, Essex and Copenhagen for useful comments. Errors are mine. Financial support from the ESRC is gratefully acknowledged.

[1]Recent contributions that stress the link between sorting, inequality and mobility include e.g. Kremer (1997), Fernandez, Guner and Knowles (2001), Fernandez and Rogerson (2001), Kremer and Chen (1999).

how sorting relates to the underlying distribution of characteristics. Is the degree of equilibrium sorting independent of the degree of variability in characteristics or would e.g. a reduction in the variance of wages reduce the incentives to sort? Is increased sorting associated with higher age at first marriage?

One approach to tackling these questions is to estimate a structural model of the marriage market that allow us to recover the underlying preference and technology parameters. With a structural model at hand one can consider potential equilibrium effects of various types of policies, e.g. minimum wage policies, tax policies, and education policies, on the process of partnership formation, not only in terms of if and when people marry, but also whom they marry.

Recent years have seen a number of important theoretical papers on sorting in marriage markets with search frictions, most notably Burdett and Coles (1997), Bloch and Ryder (2000) and Shimer and Smith (2000). The framework put forward in these papers is attractive both from a theoretical- and an empirical perspective. The models provide a theory of timing of marriage as well as a theory of sorting in marriage. Moreover, the models naturally allow the equilibrium degree of sorting to depend on the moments of the underlying distribution of types. This contrasts the static matching literature (Becker, 1973) which typically predict perfect sorting irrespective of the variability in variable on which sorting occurs.[2] From an empirical point of view, the search models also have a structure that make them amenable for direct estimation as particular types of duration models.

This paper takes a search model of the marriage market and applies it to a British cohort dataset. The approach I adopt is fundamentally different from that in the aforementioned empirical literature on assortative mating: rather than investigating whether there is positive sorting along one dimension or another, I assume that an individual's characteristics combine into a single-dimensional index representing that individual's input into marital production. I can then apply the theoretical model by letting the marital input index be the dimension along which individuals differ.[3] The input thus constitutes an individual's level of "attractiveness" when interacting in the marriage market. The approach rests on two fundamental assumptions:

[2] E.g. in the framework used by Becker (1991), with transferable utility and complementary characteristics, the male with the $n$th largest endowment marries the female with the $n$th largest endowment, irrespectively of the degree of variability in endowments.

[3] Recall that the matching models that are available all rest on the assumption of a unidimensional type-space.

1. All individuals interact with everyone else in the economy – the marriage market is not "segmented".

2. All individuals agree on the ranking, in terms of attractiveness, of the members of the opposite sex.

In equilibrium there is then positive assortative mating (in the appropriate stochastic sense) on attractiveness. The fundamental parameters of interest are (i) the parameters of the mapping from characteristics to marriage market attractiveness, and (ii) the degree of search frictions in the marriage market. To recover these parameters I use data which contains information about the respondent's characteristics, the age at entry into first marriage, and the characteristics of the subsequent partner. The estimation strategy is to nest the marriage market equilibrium within a maximum likelihood routine; in other words, for each trial value of the parameters I numerically solve for the marriage market equilibrium and then evaluate the probability of the observed outcome at that equilibrium.

The characteristics that I focus on in the current paper are education and wage since these are the two characteristics that have been focused on in the literature on assortative mating.[4] The parameter estimates suggest that:

- The rate at which individuals enter marriage strongly depends both on search frictions and on selective behaviour in the choice of partners.

- Education is a key determinant of marriage market attractiveness; wage on the other hand, while having a positive impact, is unlikely to substantially increase an individual's rank in the marriage market once the individual's education is fixed.

- There are important unobserved components that contribute to an individual's attractiveness; nevertheless, these components do not overshadow the importance of the measured characteristics.

After estimating the model I use the recovered parameters to simulate the marriage market equilibrium. I start with a simulation based on the distribution of characteristics observed in the data; this serves as a useful check on the adequacy of the model specification. The results

[4]Authors have typically found that there is positive assortative mating on both education and wages. The positive sorting on wages goes against Becker's original prediction based on the notion of specialization. Lam (1988) demonstrates that positive sorting on wages can obtain from the consumption of household public goods.

from this simulation are striking. First, the model somewhat fails to replicate the shape of the observed CDF of marriage dates. Second, and more importantly, the model underpredicts the correlation in partners' characteristics. In particular, the model only replicates about one fourth of the observed correlations in education and wages. I argue that this underscores the limits imposed by the two aforementioned key assumptions, both of which are in line with the received theory of marital matching. The results suggest that, in order to further explain the observed patterns of marital sorting, one would need to generalize the models to allow the marriage market to be less fully integrated or to allow individuals to rank bundles of partners' characteristics differently.

I then also consider a number of counterfactual scenarios where I simulate the model on alternative distributions of characteristics. The purpose of these simulations is to explore how e.g. the expansion of higher education and changed wage inequality may have affected marriage patterns both in terms of timing and in terms of sorting. Here I use data from an earlier cohort to construct suitable counterfactual distributions.

The literature that attempts to structurally estimate marriage market behaviour is scarce. An early paper is Keeley (1977). However, Keeley uses both the respondent's and the respondent's partner's characteristics as exogenous variables explaining the couple's time of marriage. Hence, Keeley does not model the incentives to sort. Boulier and Rosenzeig (1984) consider jointly the decision to invest in schooling and when to marry. Using data from the Philippines they show that the two decisions are linked – an individual takes into account that the education decision has an impact for her marriage prospects. The approach adopted by Boulier and Rosenzweig is however quite different from that of the current paper: they postulate the existence of a stable matching function according to which that partner's characteristics is a function of the respondent's characteristics and time of marriage and treat time of marriage as a choice variable. In contrast, we assume that individuals choose which type of partners to accept and treat time of marriage as a random variable.

The paper closest to this is Wong (2003) who also estimates an equilibrium search model of the marriage market based education and wages. This paper, however, differs from Wong's along several dimensions. First, I use a "cohort" data set – the British Cohort Study – which implies that all individuals in the sample have the same age and hence truly participate in the same marriage market. This is likely to be important since the distribution of characteristics can differ substantially across cohorts.[5] Second, I base my analysis on a model with transferable utility; this is important since, as is well-known, transferable utility (i.e. bargaining over the division of surplus) typically reduces the incentives to sort.[6] It also implies a very different algorithm for solving for the marriage market equilibrium. Third, I use simulations to understand the limitations of the model specification and to consider counterfactual scenarios.

The outline of the paper is as follows. Section II explains the approach; in particular, I outline the setup, describe the equilibrium model – a version of the model by Shimer and Smith (2000) – and the solution algorithm, detail how the model is empirically implemented, and derive the likelihood function. In Section III I describe the data and in Section IV I present the results from the estimation. Section V presents the simulations and Section VI offers a discussion. Section VII concludes.

# II The Model

## The Setup

Consider an economy with an equal number of men and women. Each individual is characterized by a (time invariant) vector of characteristics $\mathbf{z}$. The characteristics that I focus on here are education $e \in E$ and log hourly wage, $\omega \in \mathbb{R}_+$. Hence $\mathbf{z} = (e, \omega)$. These characteristics, along with some unobserved components (see below), determine an individual's marriage market attractiveness, denoted by $x$. We are interested in understanding the relation between characteristics $\mathbf{z}$ and attractiveness $x$. Let $\boldsymbol{\beta}$ and $\sigma$ be parameters of that relation and of the unobserved components. Also let the indices $m$ and $f$ represent men and women respectively.

The joint distribution of characteristics is assumed to be known (and is estimated from the data); let the joint CDF for $\mathbf{z}$ be $F(\cdot|m)$ for males and $F(\cdot|f)$ for women. In practice, since education is naturally discrete, $F(\cdot|k)$, $k = m, f$, can conveniently be expressed in terms of a marginal distribution of education and a set of conditional distributions of wages. For simplicity, I approximate each conditional distribution of wages $F(\omega|e,k)$ with a normal distribution with

[5] One can also argue that social norms regarding e.g. age at marriage are likely to have changed over time which would make it in appropriate to treat individuals from different cohorts as interacting on the same marriage market.

[6] I also use a much simpler error structure, thus reducing the number of parameters; this is valuable since the likelihood function is typically fails to be globally concave, leading me to use non-Newtonian optimization techniques.

mean $\mu(e,k)$ and standard deviation $\sigma(e,k)$.[7]

All individuals start out as single. While single, an individual randomly meets new potential partners at the fixed rate $\rho$. The choice problem facing the individual is which potential partners to agree to marry. I assume that the utility as single is zero: an agent does not appreciate his/her own attractiveness while single. When a male of type $x^i$ and an female of type $x^j$ match their marital output is $x^i x^j$ (see below for a discussion).

How this output is shared is determined through Nash bargaining, thus making utility transferable. Once a couple marry they are assumed to remain married forever. Each individual decides whom to propose marriage to in order to maximize his/her discounted lifetime utility.

The fundamental parameters to be estimated are $\boldsymbol{\beta}, \rho$ and $\sigma$. These parameters are to be estimated by maximum likelihood with equilibrium imposed on the estimation. Thus when I evaluate the likelihood of a parameter vector $(\boldsymbol{\beta}, \rho, \sigma)$ I start by numerically solving for the marriage market equilibrium; I then compute, for each respondent, the probability of the observed outcome – the date of marriage and the partner's characteristics – at that equilibrium.

Next I describe the details of the marriage market equilibrium and how it is computed. In order to make that computation speedy I discretize the type space so that I can obtain an equilibrium using standard matrix operations.

### The Marriage Market Equilibrium Model

Let there be a discrete ordered set of attractiveness types, $X \equiv \{x^1, ..., x^N\}$. The type space $X$ is the same for men and women; however, the distribution over types may differ. Let $\ell(x|m)$ and $\ell(x|f)$ denote the PDF over $X$ among singles for men and women respectively. The distribution of attractiveness types among singles will depend on the rate at which individuals marry: types that marry slowly should accumulate in pool of singles. In order to capture this, I assume that when someone marries, he or she is replaced by someone whose characteristics are drawn from the distribution of characteristics observed in the sample.[8] For a given $\beta$ and $\sigma$, let $l(x|m)$ and $l(x|f)$ denote the distribution of attractiveness in the observed sample.

A *strategy* for a male of type $x \in X$ is a time-invariant *acceptance set* $A^m(x) \subset X$, the set of

[7] An alternative would be to try to characterize each conditional distribution non-parametrically. However, assuming normality is likely to be a reasonably good approximation; moreover, the simplicity of the characterization is useful later on when I construct counterfactuals to be used in the simulation exercises.

[8] I adopt the one-for-one replacement assumption since I'm not interested in the size of the singles pool, only its composition.

female types with whom he is willing to match. Similarly, a strategy for a female of type $x \in X$ is a time-invariant *acceptance set*, $A^f(x) \subset X$, the set of types with whom she is willing to match. From the two correspondences $A^m(\cdot)$ and $A^f(\cdot)$ we obtain a matching matrix $M$ with elements $m_{ij}$, $i,j = 1,...,N$. $m_{ij}$ is equal to one if a male of type $x^i$ accepts *and* is accepted by a female of type $x^j$ and zero otherwise. Hence, for $i,j = 1,...,N$,

$$m_{ij} = \begin{cases} 1 & \text{if } x^j \in A^m(x^i) \text{ and } x^i \in A^f(x^j) \\ 0 & \text{otherwise} \end{cases}. \tag{1}$$

Note that, in the matrix $M$, each rows correspond to males and columns correspond to females. The matrix $M$ may or may not be symmetric.

In a steady state *Nash equilibrium*, each agent chooses a strategy that maximizes his/her discounted expected utility given the behaviour of everyone else. Let $S^m(x^i|x^j)$ denote the capital gain for a type-$x^i$ male who enters into marriage with type-$x^j$ female, while $S^f(x^j|x^i)$ is defined analogously as the capital gain to the female. Then the value of being a single male of type $x^i$, denoted $W^m(x^i)$, satisfies

$$rW^m(x^i) = \rho \sum_{j=1}^{N} m_{ij} S^m(x^i|x^j)\, \ell(x^j|f), \tag{2}$$

and similarly, for women of type $x^j$, the value as single, $W^f(x^j)$, satisfies

$$rW^f(x^j) = \rho \sum_{i=1}^{N} m_{ij} S^f(x^j|x^i)\, \ell(x^i|m). \tag{3}$$

Nash bargaining implies equal sharing of the surplus; hence

$$S^m(x^i|x^j) = S^f(x^j|x^i) = \frac{1}{2}\left(\frac{x^i x^j}{r} - W^m(x^i) - W^f(x^j)\right). \tag{4}$$

The match $(x^i, x^j)$ is mutually agreeable precisely when the surplus in (4) is positive,

$$m_{ij} = \begin{cases} 1 & \text{if } \frac{x^i x^j}{r} - W^m(x^i) - W^f(x^j) \geq 0 \\ 0 & \text{else} \end{cases}. \tag{5}$$

In equilibrium, a single male (female) of type $x^i \in X$ has a constant marriage hazard rate, denoted $\lambda^m(x^i)$ (alt. $\lambda^f(x^i)$) which satisfies

$$\lambda^m(x^i) \equiv \rho \sum_{j=1}^{N} m_{ij}\ell(x^j|f) \quad \text{and} \quad \lambda^f(x^i) \equiv \rho \sum_{j=1}^{N} m_{ji}\ell(x^j|m). \tag{6}$$

We can now derive the steady state distribution of types in the singles pool. Equating, for each type, the number of new entrants to the singles pool with the number leaving that pool we have that

$$\ell\left(x^i|k\right)\lambda^k\left(x^i\right)=l\left(x^i|k\right)\sum_{j=1}^{N}\ell\left(x^j|k\right)\lambda^k\left(x^j\right) \tag{7}$$

for $k=m,f$ and $i=1,...,N$.

Hence an *equilibrium* is a collection of two value functions (two length-$N$ vectors), $W^m$ and $W^f$, and a (binary) matching matrix $M$, and two pdfs (two length-$N$ vectors) of types in the singles pool, $\ell(x|m)$ and $\ell(x|f)$, that satisfy (2), (3), (5) and (7).

A word on the chosen utility specification is in order. From the data we know that there is positive assortative mating both on education and wages. Below I will model "attractiveness" as a simple linear function of education and wage (which we expect to be increasing in both arguments). Hence I choose the marital production function to exhibit sufficient complementarity that the marriage market equilibrium exhibits positive sorting on $x$; the results by Shimer and Smith (2000) tell us that the product function $x^i x^j$ will accomplish this.

### Solution Algorithm

To find an equilibrium I follow Shimer and Smith (2000) and apply an iterative process updating, in turn, the value functions $W^m$ and $W^f$, the matching matrix $M$, and the distribution of types in the singles pool, using (2), (3) and (5). Note that this is a very fast process. Updating of the value functions involves solving a linear equation system; updating $M$ simply involves checking the signs in a matrix; updating the type distributions involves solving two more linear equation systems. The speediness of the process is key since the estimation involves solving for the equilibrium at each trial value of the parameters. The process takes as fixed inputs the type space $X$, the meetings rate $\rho$, and the distributions over types, $L(\cdot|m)$ and $L(\cdot|f)$ in the sample. It then starts from some initial matching matrix $M$ (e.g. the identity matrix).

**Step (i): Updating the Value Functions** This step treats (2) and (3) as a system of $2N$ linear equations in $2N$ unknown which can be solved through matrix inversion. To see this let $\widehat{\rho}=\rho/2$ and define

$$\zeta_i^m\equiv r+\widehat{\rho}\sum_{j=1}^{N}m_{ij}\ell\left(x^j|f\right)\quad\text{and}\quad\zeta_j^f\equiv r+\widehat{\rho}\sum_{i=1}^{N}m_{ij}\ell\left(x^i|m\right). \tag{8}$$

Note that the term $\sum_{j=1}^{N} m_{ij}\ell\left(x^j|f\right)$ in $\zeta_i^m$ is simply the probability that a male of type $x^i$ will marry a randomly encountered female. Let $\zeta^k = \left(\zeta_1^k, ..., \zeta_N^k\right)'$. Also define

$$\xi_i^m \equiv \widehat{\rho}\sum_{j=1}^{N} m_{ij}\frac{x^i x^j}{r}\ell\left(x^j|f\right) \quad \text{and} \quad \xi_j^f \equiv \widehat{\rho}\sum_{i=1}^{N} m_{ij}\frac{x^i x^j}{r}\ell\left(x^i|m\right), \tag{9}$$

and let $\xi^k = \left(\xi_1^k, ..., \xi_N^k\right)'$ for $k = m, f$. Define frequency-scaled versions of $M$ for males and females as follows

$$M^m = \begin{bmatrix} m_{11}\ell\left(x^1|f\right) & ... & m_{1N}\ell\left(x^N|f\right) \\ \vdots & \ddots & \vdots \\ m_{N1}\ell\left(x^1|f\right) & ... & m_{NN}\ell\left(x^N|f\right) \end{bmatrix}, \text{ and } M^f = \begin{bmatrix} m_{11}\ell\left(x^1|m\right) & ... & m_{N1}\ell\left(x^N|m\right) \\ \vdots & \ddots & \vdots \\ m_{1N}\ell\left(x^1|m\right) & ... & m_{NN}\ell\left(x^N|m\right) \end{bmatrix}.$$

The equation system can then be written condensely in matrix form as follows

$$\begin{bmatrix} diag\left(\zeta^m\right) & \widehat{\rho}M^m \\ \widehat{\rho}M^f & diag\left(\zeta^f\right) \end{bmatrix} \begin{bmatrix} W^m \\ W^f \end{bmatrix} = \begin{bmatrix} \xi^m \\ \xi^f \end{bmatrix},$$

where $diag\left(\zeta^m\right)$ is the diagonal matrix constructed from the vector $\zeta^m$ and where $W^k = \left(W^k\left(x^1\right), .., W^k\left(x^N\right)\right)'$ for $k = m, f$. Given $M$ the updated values of $W^m$ and $W^f$ are then obtained by simple matrix inversion.

**Step (ii): Updating the Matching Matrix** Given $W^m$ and $W^m$ from step 1, step 2 involves computing the surplus matrix $S^m\left(\cdot|\cdot\right)$ and then updating $M$ by setting $m_{ij} = 1$ if $S^m\left(x^i|x^j\right) > 0$ and zero otherwise.

**Step (iii): Updating the Steady State Distribution of Types** Given $M$ one can determine the rate at which all types marry using (6), and hence I can update the distribution of types in the singles pool. This involves arranging (7) as two equation systems (one for each gender), putting these systems in matrix form and solving.[9]

If convergence is obtained, an equilibrium has been located. While no general results regarding the convergence properties of the process are available (e.g. it does not constitute a contraction), in practice it converges rapidly.

[9] Note that the equations are linearly dependent; in practice I drop one equation and replace it with the adding up condition $\sum_i \ell\left(x^i|k\right) = 1$.

### From Model to Empirics

One key aim is to investigate is how $e$ and $\omega$ combine to produce an individual's input into marital production, denoted $x$. Since this $x$ is what determines an individual's rank in the marriage market I will interchangable refer to the input $x$ as the individual's "attractivness".

A natural approach is to let $x$ be a simple linear function of the characteristics

$$x_i = \beta_0 + \beta_1 e_i + \beta_2 \omega_i + \varepsilon_i, \tag{10}$$

where $\varepsilon_i \sim N\left(0, \sigma^2\right)$ and is assumed independent of the characteristics. The error term $\varepsilon_i$ captures unobserved components of an individual's attractiveness.

### Discretization

In order to apply the above model and algorithm I generate a discrete type space from the distribution of attractiveness. It is important to note that a new discretization is made for each trial value of the parameters. The number of discrete types is set to $N = 10$.

The steps of the discretization are as follows. Given $\boldsymbol{\beta}$, $\sigma$ and the observed distributions of characteristics $F(\cdot|k)$, I first compute, using (10), the aggregate CDF for attractiveness, denoted $L(x)$ (treating $x$ as a continuous variable). Using this CDF, and following Wong (2003), I let the decile points identify a set of *boundary points*; i.e. for $i = 1, ..., 9$, I define $\overline{x}^i$ through $L\left(\overline{x}^i\right) = 0.i$ (and I set $\overline{x}^0 = -\infty$ and $\overline{x}^{10} = \infty$). The type space $X$ is then obtained by assigning as *value* $x^i$, $i = 1, ..., 10$, the expected value of $x$ conditional on $x \in \left(\overline{x}^{i-1}, \overline{x}^i\right)$. Finally, I obtain gender-specific distributions over $X$, denoted $L\left(x^i|k\right)$ above, by computing the measure of males and females that fall within each interval $\left(\overline{x}^{i-1}, \overline{x}^i\right)$, $i = 1, ..., 10$.

I also impose a normalization on the type space $X$. In particular, I set $x^1 = 1$. This is achieved by continuously adjusting $\beta_0$. Hence in the estimation, only $\beta_1$ and $\beta_2$ are considered as free parameters. This normalization has a number of advantages; it ensures that no one will be "unmarriageable" and it reduces the number of parameters to be estimated without restricting the model in terms of the total and relative value of the two characteristics $e$ and $\omega$ (as captured by the length and direction of the vector $(\beta_1, \beta_2)$).

### The Unobserved Components

The unobserved component of attractiveness $\varepsilon_i$ in (10) is an important ingredient in the model in that it allows for the possibility that an individual's true attractiveness is different from that based on his observed characteristics. It is assumed that these unobserved components are observable to all participants in the marriage market.

It is easy to see why including unobserved component is necessary. For any $\boldsymbol{\beta}$ there will always be observations where e.g. a male with very low observed attractiveness marries a female with a very high observed attractiveness or vice versa. Absent noise, such an observation would have zero probability unless the matching sets were very large; hence the model can be expected to respond to the non-inclusion of noise by e.g. setting very low meetings rate $\rho$ and by making $\boldsymbol{\beta}$ small enough that characteristics are deemed to be unimportant, which ensures that no one ever turns anyone down. Allowing unobserved components should allow us to avoid this source of bias.

An individual $i$'s true attractivenss is said to be $x^i \in X$ if and only if $x_i \in \left(\overline{x}^{i-1}, \overline{x}^i\right)$ where $x_i$ is given in (10). Given the assumed distribution for $\varepsilon_i$ I can calculate, for each individual $i$ and each possible attractiveness type $x \in X$, the probability that individual $i$'s true type is $x$. Denote this probability $\pi_i(x)$ (which depends both on the individual's characteristics and on the parameters). The fact that we need to keep track of the probability of each individual being of each type is a key factor restricting the number of types $N$ in the discretization.

### The Likelihood Function

If an individual marries, then I observe two outcome variables: (i) the age at entry into marriage, denoted $t_i$, and (ii) the characteristics of the subsequent partner, denoted $\mathbf{z}_i^p = (e_i^p, \omega_i^p)$. However, it could be that an individual still hasn't married by the end of the sampling period, $T$ – i.e. the observation is censored – in which case I only observe that $t_i > T$.

When computing the likelihood at $(\boldsymbol{\beta}, \rho, \sigma)$ I start by discretizing the type space as outlined above to obtain $X$ and the gender-specific distributions over types in $X$. Using these components I compute the marriage market equilibrium which generates the matching matrix $M$ and the equilibrium steady state distribution of types in the singles pool, $\ell(\cdot|k)$, $k = m, f$. Also, before proceeding I calculate for each individual $i$, his/her probability $\pi_i(x)$ of being of each type $x \in X$. Hence when proceeding to evaluate the probability of the individual outcomes we carry with us $X$, $\ell(\cdot|k)$, $M$, as well as $\pi_i(\cdot)$ for each individual.

In computing the probability of an individual's outcome it is convenient to condition on the unobserved true attractiveness type $x \in X$; one can then obtain the probability of the individual's observed outcome by summing over true types, weighted by the individual probability distribution over $X$.[10] There are two useful features of the marriage market equilibrium that simplify the formulation of the likelihood function once we condition on true type $x$. First, each individual has a constant exit rate from singlehood, i.e. the length of each individual's spell as single is an exponentially distributed random variable (with a hazard rate that depends on $x$). Second, the characteristics of the subsequent partner is independent of age at entry into marriage.

**Noncensored Observations** Consider a male who is observed to enter marriage after $t_i$ periods. If this male is of type $x^i \in X$, then his constant exit rate from singlehood is given by (6) (for $k = m$). Hence the exit date is exponentially distributed with parameter $\lambda^m\left(x^i\right)$ which implies that the pdf of exit dates, evaluated at $t_i$, is

$$\lambda^m\left(x^i\right)e^{-\lambda^m\left(x^i\right)t_i}. \tag{11}$$

In order to derive the probability of the partner's observed characteristics I assume that the partner's attractiveness is observed without error; the partner's type is then $x^{i^p} \in X$ if $x_i^p = \beta_0 + \beta_1 e_i^p + \beta_1 \omega_i^p \in \left(\overline{x}^{i-1}, \overline{x}^i\right)$. We then seek an expression for $\Pr\left(\mathbf{z}_i^p|x^i\right)$. But this is simply the relative frequency of the characteristics $\mathbf{z}_i^p$ among all women that a male of type $x^i$ marries

$$\frac{m_{ii^p} f\left(\mathbf{z}_i^p|f\right)}{\sum_{j=1}^N m_{ij}\ell\left(x^j|f\right)}.$$

where $i^p$ is index of the partner's type. (The indicator variable $m_{ii^p}$ sets the probability to zero if the male does not marry a female of type $x^{i,p}$).

Putting the likelihood together and integrating out the unobserved true type, the contribution to the likelihood of male $i$ who is noncensored is

$$\mathcal{L}_i^{nc} = \sum_{x^i \in X} \pi_i\left(x^i\right)\lambda^m\left(x^i\right)e^{-\lambda^m\left(x^i\right)t_i}\frac{m_{ii^p} f\left(\mathbf{z}_i^p|f\right)}{\sum_{j=1}^N m_{ij}\ell\left(x^j|f\right)},$$

where I used that the exit date and the partner's characteristics are independent events conditional on $x^i$. A similar expression holds for noncensored females.

**Censored Observation** Consider now a male who has not married by the end of the sampling period; for this individual I only observe that $t_i > T$. Again, I start by conditioning on the

[10] Suppose e.g. we observe the outcome $\left(t_i, \mathbf{z}_i^p\right)$ for an individual with characteristics $\mathbf{z}_i$. Then we use that $\Pr\left(t_i, \mathbf{z}_i^p|\mathbf{z}_i\right) = \sum_{x \in X} \Pr\left(t_i, \mathbf{z}_i^p|x\right)\pi_i\left(x|\mathbf{z}_i\right)$.

unobserved type $x^i \in X$; then since time to first marriage is exponentially distributed with parameter $\lambda^m(x^i)$, the probability of the agent being censored is $e^{-\lambda^m(x^i)T}$. Integrating out the unobserved true type, the contribution to the likelihood of male $i$ who is noncensored is

$$\mathcal{L}_i^c = \sum_{x^i \in X} \pi_i(x^i)\, e^{-\lambda^m(x^i)T}. \tag{12}$$

A similar expression holds for censored females. The overall likelihood is then obtained as the product over then individuals in the sample, for each individual using the relevant case (i.e. male/female and noncensored/censored).[11]

# III Data

I apply the framework described above to data from the British Cohort Study (BCS), a survey of all children born between 5 and 11 April 1970 with followup samples at ages 5, 10, 16, 26, 30. This type of data is well-suited for the current purpose: it allows me to follow one particular cohort passing through the marriage market from the day they enter that market until the age of thirty. The variables of interest are the respondent's characteristics (education $e_i$ and log wage $\omega_i$), the age at entry into first marriage $t_i$, and the subsequent partner's characteristics (education $e_i^p$ and log wage $\omega_i^p$).

For education $e$ I use "age at which left full time education" since this variable is recorded both for respondents and for partners. This is a naturally discrete variable taking on integer values. However, since I need to estimate the distribution of wages at each education outcome, I choose to use a somewhat coarser partition which is more closely related to qualification structure. I code $e \in E = \{1, ..., 5\}$ where $e = 1$ for leaving full time education at before 16 ("no qualifications"), $e = 2$ for leaving at 16 or 17 (completed GCSEs), $e = 3$ for leaving at 18 or 19 (completed A-levels), $e = 4$ for leaving at 20 or 21 (some higher education) and $e = 5$ for leaving at after 22 or higher (university education). One issue is whether one should measure wage at the time of marriage or at one date for all. It is reasonable to argue that individuals choose partners based on expected adult earnings capacity. Hence I choose to use predicted log

[11] Preliminary inspection of the log likelihood function suggests that it is not globally concave. To tackle this problem I optimize the likelihood using "simulated annealing" rather than a standard Newtonian method; simulated annealing is capable of overcoming problem with local maxima by randomly sampling points at some distance away from the current trial value. After finding a maximum I turn to grid search around that point to verify that the point is indeed a global optimum.

wage at age 30.[12]

I drop individuals who are self-employed, disabled, or in education at the age of thirty. I also drop all observations where information about the respondent's characteristics were missing. This leaves me with 9233 individuals, 4648 of whom have married by the age of thirty. I cannot drop individuals for whom I lack information about the characteristics of the partner: doing so would increase the fraction of censored individuals in the sample, immediately generating a bias. I deal with this problem in several ways. If the individual is divorced and information about the ex-partner is missing, then I use information on any current partner as a proxy. 86 individuals failed to report age at marriage, and 540 individuals failed to report partner's education. For these individuals I impute values by drawing randomly a date of marriage from the observed CDF of marriage dates and/or an education level for the partner from the observed distributions pertaining to the respondent's gender/education category. Descriptive statistics are given in Table 1. The observed correlation in partner's education is 0.509 and that observed correlation in (log) wage is 0.187.

A small number of individuals marry at low ages, between 15 and 18. For reason that will be clear later on, I assume that the individuals enter the marriage market only at age 18 and I recode these early marriages as occurring in the first year in the marriage market.

## IV Estimates

The point estimates of the parameters are presented (with standard errors) in Table 2. The estimates suggest education and wage are both important determinants of marriage market attractiveness, that an individual meets a potential partner about once every seven years, and that there is substantial unobserved components of attractiveness. The rest of this section will illustrate these results in a number of ways.

Since the approach nests the marriage market equilibrium, the equilibrium matching matrix $M$ is recovered in the estimation. The matching matrix is illustrated in Figure 1 where the meetings that are converted into marriages are highlighted. E.g. a male of type $x^3$ will marry

[12] The predicted wage is obtained from estimating a standard wage equation with correction for selection. Details of the regression (which included gender, marital status, children, and dummies for education, ethnicity, social class and occupations) are available on request. The code used for cleaning the wage variable was based on original code by Lorraine Dearden and Alissa Goodman of the IFS, kindly made available through the Data Archive.

any female of type $x^2$ through $x^6$. The matrix in this case is not symmetric. The figure indicates that there is a fair amount of assortative mating based on attractiveness.

FIG 1 HERE

Using the equilibrium matching matrix along with the gender-specific distributions over types, one can determine the probability that a typical random meeting is converted into marriage,

$$\sum_{i=1}^{N}\sum_{j=1}^{N} m_{ij}\ell\left(x^i|m\right)\ell\left(x^j|f\right) = 0.22$$

This highlights that the average time spent in the marriage market is the combined outcome of both search friction and selective choices.

Figure 2 illustrates the estimated type space $X$.

FIG 2 HERE

Since the type space is obtained by splitting the population into deciles, 10 percent of the aggregate population belongs to each type. Hence the function illustrated in Figure 2 is effectively the inverse of the aggregate CDF of attractiveness. The figure offer some perspective on the estimated size of the unobserved components $\varepsilon_i$. Consider e.g. an individual whose observed characteristics imply the predicted attractiveness 58, making him median in the population. If his error term $\varepsilon_i$ is minus one standard deviation, his rank will drop to about the 20nd percentile while if the error term is plus one standard deviation his rank increases to about the 80th percentile. Hence with nearly 70 percent probability the individual belongs to the middle six deciles in terms of attractiveness. In other words, while unobserved components of attractiveness are important, they clearly do not overshadow the importance of the measured characteristics.

Finally, Figure 3 illustrates the level curves corresponding to the types $x^2$ through to $x^9$ in the space of education and hourly wages.[13] Here is should be recalled that a one unit increase in education corresponds to about a 2 years increase in years of schooling (see above). The figure highlights the overwhelming importance of education.

FIG 3 HERE

[13]The level curves for $x^1$ and $x^{10}$ do not appear in the figure but are located further down and to the right respectively; their extreme positions reflect the size of the unobserved noise.

By construction an upwards move from one level curve to the next corresponds to (on average) a one-decile increase in rank in the marriage market.[14] In this sense the level curves have a cardinal interpretation which helps us interpret the distance between them. Note e.g. that around education $e = 2$, the level curves are relatively close. This reflects that $e = 2$ is the most common education choice; hence conditional on $e = 2$, an given increase in wage can substantially increase an individual's rank.

It is interesting to note that from Figure 1 men and women of type $x^1$ are effectively unmarriageable. However, Figure 3 also reveals that if a person is unmarriageable it is not due to his/her observed characteristics: having low education and/or wage is not sufficient (but, effectively, necessary) to be of type $x^1$.

# V Simulations

After estimating the model I now turn to some simulations. A simulation based on the observed characteristics provides a useful "specification test" of the model; in particular, a simulation will reveal if the model fails to replicate some key features of the observed outcome.

I also use simulations to predict equilibrium responses to changes in the primitives in the economy, in this case the joint distributions of characteristics. Over the last couple of decades age at first marriage has increased rapidly (See e.g. Gould and Paserman (2003) for a discussion of the US experience). Simultaneously there has been in increase in inequality and an increase in the participation in higher education.[15] It is then natural to ask whether these phenomena are related. Search models offer a natural framework for studying this question. In this framework, the option value of remaining single will depend on the variability in attractiveness of the participants in the singles market. Hence we would e.g. expect that an increase in the inequality of wages to lead to more search to more sorting. Indeed, the hypothesis that increased wage inequality leads to delayed entry into marriage has recently been tested empirically by Gould and Paserman (2003): using US data and exploiting variation in inequality over time and space, Gould and Paserman demonstrate that women living in cities with more male wage inequality

[14] For a plot of the boundary points $\overline{x}^i$, $i = 0, ..., 10$, the interpretation as a one-decile increase would have been exact.

[15] The increase in inequality has been extensively documented; see e.g. Gottschalk (1997) and Machin (1999) for overviews and discussions. On the increase in participation in education in the UK, see Blanden, Gregg and Machin (2002) who also discuss the link to inequality.

spend longer time searching for a husband.

In this section I thus treat the estimated parameter values $\boldsymbol{\beta}$, $\rho$ and $\sigma$ as fixed technology parameters and simulate the model for different joint distributions of characteristics for a large number of individuals.

### The Simulation Exercise

A simultion exercise proceeds in several steps. The first step involves choosing a number of individuals $n$ and the joint distributions of characteristics for men and women. The second step involves using these distributions to randomly assign characteristics (and a noise term $\varepsilon$) to $n/2$ men and $n/2$ women. Using the distributions of characteristics, and using $\boldsymbol{\beta}$ and $\sigma$, I determine the distribution of attractiveness $x$ which I then discretize as above and each individual's type $x^i \in X$ is determined. The third step uses the type space $X$ and the gender-specific distributions over $X$ and computes the marriage market equilibrium $M$. Using an individual's type and the equilibrium matching matrix $M$ I finally simulate a marriage date (a draw from an individual-specific exponential distribution) and a partner (first in terms of attractiveness type and then in terms of characteristics). I set $n = 250,000$, repeat the exercise ten times and take averages of the relevant outcome variables.

### A Simulation Based on the Observed Distribution of Characteristics

I start by simulating the model based on the joint distributions of education and wages for men and women observed in the BCS data. This serves as a "specification test" of the structural model that we have imposed on the data. In particular, it should help us spot if there are some restrictions imposed by the theoretical model that do not seem to be compatible with the observed outcome.

The outcome in terms of marriage dates is illustrated in Fig. 4.

FIG 4 HERE

The figure highlights a first shortcoming of the fitted model. Note that the actual CDF of age at first marriage is convex at low ages (below 24), reflecting that fewer individuals marry at e.g. 19 or 20 than at 21 or 22. In the theoretical model that was fitted to the data, on the other hand, each individual has a constant hazard rate, implying a concave CDF. Even though the hazard rates are individual-specific, the simulated aggregate CDF of marriage dates inherits the concavity property of the exponential distribution. In short, the theoretical model fails to capture the fact that few individuals marry at low ages. This was also the reason why I assumed that the individuals enter the marriage market only at age 18; if I had assumed that they enter at the age of 15 (the earliest observed marriage) the fit would have been significantly worse.

The simulated equilibrium does exhibit positive assortative mating on both education and wage. However, it substantially underpredict the degree of sorting: for either characteristic, the correlation in the simulated equilibrium is only forty percent of the correlation observed in the BCS data. See Table 4 (Simulation 1).

What can account for this underprediction? One possible explanation is that the specification of the empirical model is not adequate; e.g. it could be that the error structure is not general enough, that the type-space is too coarsely discretized, or that the functional form of the attractiveness index is wrong.[16] If, however, if the model, even after generalizations, fail to replicate the high observed correlations (for education in particular), then we must conclude that the problem likely lies with the two fundamental assumptions of the approach. In particular, the result then casts doubt on the assumptions that all men (women) rank the women (men) in the same way and that everyone is equally likely to interact with everyone else. We would then conclude that the theoretical models need to be generalized to allow for individuals to rank the members of the opposite sex differently, or to allow for the possibility that not everyone interacts with everyone else with equal probability.

**Changing the Distribution of Characteristics**

I now turn to considering changes in the underlying distribution of characteristics. For that purpose I want to construct relevant counterfactual distributions of characteristics. The BCS was preceeded by another cohort study, the National Child Development Study (NCDS), which surveys all children born in Britain between 3 and 9 March 1958, i.e. 12 years before the BCS cohort. Followups in this survey occured at ages 7, 11, 16, 23, 33 and 42. I thus obtain the characteristics of the respondents in the NCDS in the same way as I did for BCS. I predict log hourly wages (at age 33, in year 1991) using the same variables that I used to predict log wages for the BCS cohort (at age 30, in year 2000).

How has the distribution of characteristics changed from the 1950 cohort to the 1970 cohort?

[16] I have estimated the model with other specifications for the marital production function; however this has not improved the model's ability to predict sorting.

Consider first education. The BCS cohort is considerable better educated than the NCDS cohort (Table 3). In particular, we see that not only is the fraction of individuals leaving school before age 16 is larger in the 1958 cohort, but the number of individual proceeding past secondary education is also lower for both men and women.[17]

A second change concerns the relation between education and conditional average (log) wage. In fact, as is by now well-recognized, while the returns to education increased sharply in the 1980s, this trend was broken in the 1990s. Indeed, Figure 5 (Panel A) plots average log wage for each education group relative to average low wage for those with $e = 1$ ("no qualifications"). The solid lines represent the NCDS cohort and the dashed lines represents the BCS cohort; squares represent women and cirles represent men. Hence the figure shows that "returns to education" has reduced sharply for women and has reduced somewhat for men.[18]

FIG 5 HERE

The third change involves within-group inequality as measured by the variability in wages within each gender/education category. Figure 5 (Panel B) depicts the standard deviations in predicted log wages for each gender/education category. (Again, the solid lines represent the NCDS cohort and the dashed lines represents the BCS cohort; squares represent women and cirles represent men.) The figure suggests that wages have become more variable at high education levels but somewhat less variable for those with low levels of education.

I simulate the model on three counterfactual distributions of characteristics. In particular, I use the distribution of characteristics in BCS as starting point; I then change one component at a time: the distribution of education, the "returns to education", and the within-group inequality.

**Changing the Distribution of Education** Did the increased participation in education affect the degree of sorting in the marriage market? To get a handle on this I simulate the model again; I leave the conditional wage distributions be unchanged, but I replace the distribution of

[17] In terms of coding of education $e = 1, ..., 5$ one can easily verify that the variance of education is higher in the BCS cohort than in the NCDS cohort.

[18] Since my primary interest is in the distribution of characteristics (education and log wage) the figure is constructed from the "raw" conditional average predicted log wage; nevertheless, the decrease in the returns to education suggested by the figure has been confirmed e.g. by Dearden, Fiorini and Goodman (2001) who control for various background variables.

education with that observed in the 1958 cohort. The result is reported in Table 4 (Simulation 2). Since the correlations obtained in this case is lower, the simulation suggests that the expansion of education may have contributed to higher correlations in partners' characteristics, both in terms of education and wages.

**Changing the Returns to Education** What effect do returns to education have on sorting into marriage? To investigate this I leave average wages for men and women unchanged, but I change the slopes of the mean (log) wage functions to match the slopes observed in the NCDS data (i.e. in terms of figure 5 I replace the dashed lines with the dotted lines). The result is reported in Table 4 (Simulation 3).

Since the simulated correlations are slightly larger than in the baseline simulation, the result suggests that the decreasing returns to education may have contributed to decreasing the degree of sorting. It is interesting to note that the result comes through for the wage correlation in particular; this is exactly what one would expect given that sorting occurs primarily on education. Suppose e.g. that everyone always married exactly according to rank in terms of education. In that case a reduction in the returns to education would have no effect on the correlation in partners' education (remember that we are not changing the distribution of education), but a negative effect on the correlation in wages (since the conditional wage distributions of different educational groups will "move closer").

**Changing Within-Group Inequality** Finally, I simulate the model this time replacing the standard deviation of wages for each group by that observed in the NCDS data. Since the change in within-group inequality has not been unambiguous, we should not expect any major effects. Indeed, the simulated effects on the correlations of partners' characteristics are negligable, both for education and wages. See Table 4 (Simulation 4). It is important to keep in mind that sorting occurs more heavily on education than on wages. It is then quite conceivable for an increased wage variance to lead to lower correlation in wages simply as a result of the fact that all individuals' wages become more scattered.

I have so far only reported the results in terms of correlations of characteristics. One might also conjecture that changes in the distribution of characteristics would also change the timing of entry into marriage. However, I do not find any such effects: no counterfactual distribution changed the simulated CDF of marriage dates by more than one or two percentage points at any age. In fact, the model predicts that the increase in participation in education should have slightly *increased* the fraction of individuals married at age 30; this happens simply since highly educated agents are also highly marriageable. Thus the model does not offer any insights into why there has been a significant increase in the age at first marriage over the last couple of decades.

# VI Discussion and Extensions

The above model is a first attempt to estimate a structural search model of the marriage market on British cohort data. While the results have been illuminating, they also present a number of problems to be solved. First, it is noteable that the simulations can only replicate a fraction of the observed correlations in partner's characteristics. This casts doubt on the two fundamental assumptions noted in the Introduction, i.e. that all individuals rank all members of the opposite sex in the same way and that everyone is equally likely to interact with everyone else. It is e.g. highly plausible that highly educated men have a stronger preference for highly educated women. Note that this could, at least in principle be tested once the model has been estimated. One way to do this would be to find, in the data, pairs of men (say) who have different characteristics but whose expected attractiveness (i.e. they are on the same level curve in Figure 3). According to the theoretical model (and hence in the simulated equilibrium) the wifes of these men should be drawn from the same distribution. It should not be that, systematically, the higher educated male should be married to a more highly educated female.

A second problem is that the model as it stands fails to capture why so few individuals marry at low ages. Hence, the results suggests that a more plausible empirical search model of marriage markets must include some reason why some individuals become active in the marriage market at a higher age (e.g. in conjunction with leaving education). A possible third problem that I have not tackled in the current model is that the values of the education and wage may differ by gender. Hence a natural extension to the model would be allow the $\boldsymbol{\beta}$ vector (and possibly also $\sigma$) to differ by gender; work on this extension is under way.[19]

[19]Preliminary results suggest that education is relatively more important for women than for men, but that allowing for gender-specific $\boldsymbol{\beta}$ vectors will not radically improve the simulated correlations in partners' characteristics. However, the estimation so far has proved somewhat sensitive.

## VII Conclusions

This paper provides an attempt to estimate an equilibrium search model of the marriage market structurally using a particularly well-suited British data set, the British Cohort Study, which has followed all children born in Britain in one week in March 1970 until the year 2000. This data allows me to model who marries whom and when. An individual's attractiveness (his/her input into marital production) is modeled as a simple linear function of education and (log) wage. Agents meet randomly and decide which potential partners to marry and which to turn down.

The results suggest that the rate at which individuals marry strongly depends both on search frictions and on selective behaviour in the choice of partners. Eduction is found to be a somewhat more important determinant of an individual's attractiveness than wage; nevertheless wage can be an important characteristics for those whose education level is shared by many others. Marriage market attractiveness is also found to have important unobserved components.

After estimating the model, I also simulate the model. This serves two purposes. First it serves as a specification test. Here I find that the model does not replicate the high degree of sorting on characteristics that is observed in the data. I argue that this casts some doubt on two fundamental assumptions that are common in equilibrium models of the marriage market. Second, it allows me to investigate if changes in the distribution of characteristics (education and wages) may have affected equilibrium sorting and timing of marriage. Here I find that e.g. increased participation in education may have contributed to increased sorting in the marriage market.

## References

Becker, G. S. (1973), 'A theory of marriage: Part I', *Journal of Political Economy* **81**, 813–846.

Becker, G. S. (1991), *A Treatise on the Family*, 2'nd edn, Harvard University Press, Cambridge, Mass.

Blanden, J., Gregg, P. & Machin, S. (2002), 'Education and family income'. Mimeo, University College London.

Bloch, F. & Ryder, H. (2000), 'Two-sided search, marriages and matchmakers', *International Economic Review* **41**, 93–115.

Boulier, B. L. & Rosenzweig, M. R. (1984), 'Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior', *Journal of Political Economy* **92**, 712–732.

Burdett, K. & Coles, M. G. (1997), 'Marriage and class', *Quarterly Journal of Economics* **112**, 141–168.

Fernandez, R., Guner, N. & Knowles, J. (2001), 'Love and money: A theoretical and empirical analysis of household sorting and inequality'. CEPR Discussion Paper No. 3040.

Fernandez, R. & Rogerson, R. (2001), 'Soorting and long-run inequality', *Quarterly Journal of Economics* **116**, 1305–1341.

Gottschalk, P. (1997), 'Inequality, income growth, and mobility: The basic facts', *Journal of Economic Perspectives* **11**, 21–40.

Gould, E. D. & Paserman, M. D. (2003), 'Waiting for Mr. Right: Rising inequality and declining marriage rates', *Journal of Urban Economics* **53**, 257–281.

Keeley, M. (1977), 'The economics of family formation', *Economic Inquiry* **15**, 238–250.

Kremer, M. (1997), 'How much does sorting increase inequality?', *Quarterly Journal of Economics* **112**, 115–139.

Kremer, M. & Chen, D. (1999), 'Income distribution dynamics with endogenous fertility', *American Economic Review* **89**, 155–160.

Lam, D. (1988), 'Marriage markets and assortative mating with household public goods: Theoretical results and empirical implications', *Journal of Human Resources* **23**, 462–487.

Machin, S. (1999), Wage inequality in the 1970's, 1980's and 1990's, *in* P. Gregg & J. Wadsworth, eds, 'The State of Working Britain', Manchester University Press.

Mare, D. (1991), 'Five decades of educational assortative mating', *American Sociological Review* **56**, 15–32.

Shimer, R. & Smith, L. (2000), 'Assortative matching and search', *Econometrica* **68**, 343–370.

Wong, L. Y. (2003), 'Structural estimation of marriage market models', *Journal of Labor Economics* **21**, 699–727.

| | **Men** | **Women** |
|---|---|---|
| Hourly Wage: Mean | 9.20 | 6.86 |
| Hourly Wage: S.D. | 7.78 | 7.98 |
| Education: $e = 1$ | 0.018 | 0.019 |
| Education: $e = 2$ | 0.648 | 0.608 |
| Education: $e = 3$ | 0.159 | 0.210 |
| Education: $e = 4$ | 0.055 | 0.047 |
| Education: $e = 5$ | 0.120 | 0.114 |
| Sample size | 9233 | |

Table 1: Descriptive Statistics: Mean wages, standard deviation for wages, distribution of education over discrete categories and sample size.

| | Estimate | Standard Error |
|---|---|---|
| Education ($\beta_1$) | 15.93 | 0.227 |
| Log Wage ($\beta_2$) | 9.27 | 0.101 |
| Meetings Rate ($\rho$) | 0.155 | 0.0005 |
| Noise ($\sigma$) | 32.35 | 0.524 |

Table 2: Point estimates and estimated asymptotic standard errors.

| | **Men** | **Women** |
|---|---|---|
| Education: $e = 1$ | 0.021 | 0.030 |
| Education: $e = 2$ | 0.676 | 0.662 |
| Education: $e = 3$ | 0.128 | 0.158 |
| Education: $e = 4$ | 0.070 | 0.067 |
| Education: $e = 5$ | 0.104 | 0.083 |

Table 3: The Distribution of Education in the NCDS (1958) cohort.

| Simulation | Corr. in Education | Corr. in Wage |
|---|---|---|
| Simulation 1 | 0.099 | 0.047 |
| Simulation 2 | 0.089 | 0.041 |
| Simulation 3 | 0.100 | 0.052 |
| Simulation 4 | 0.097 | 0.045 |

Table 4: Simulated equilibrium correlation in partners' characteristics.

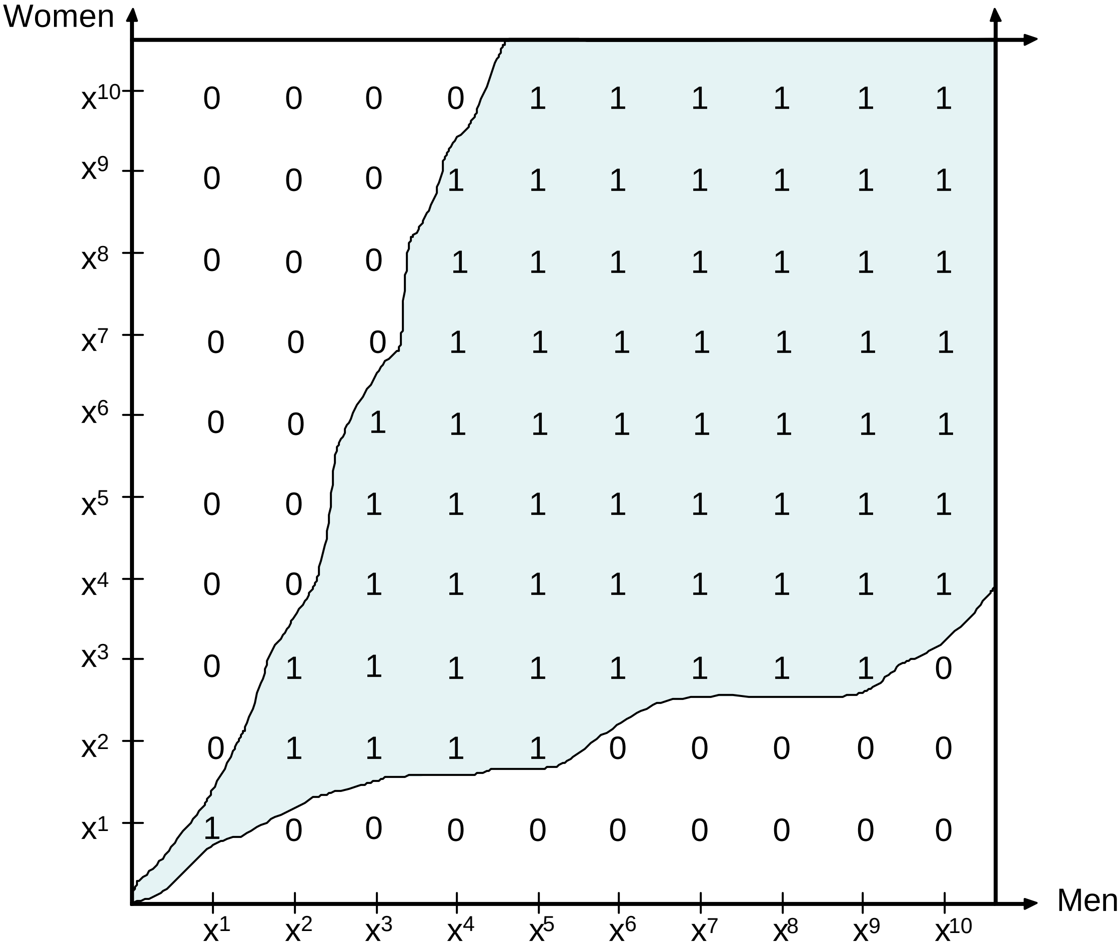


Figure 1: The equilibrium matching set.

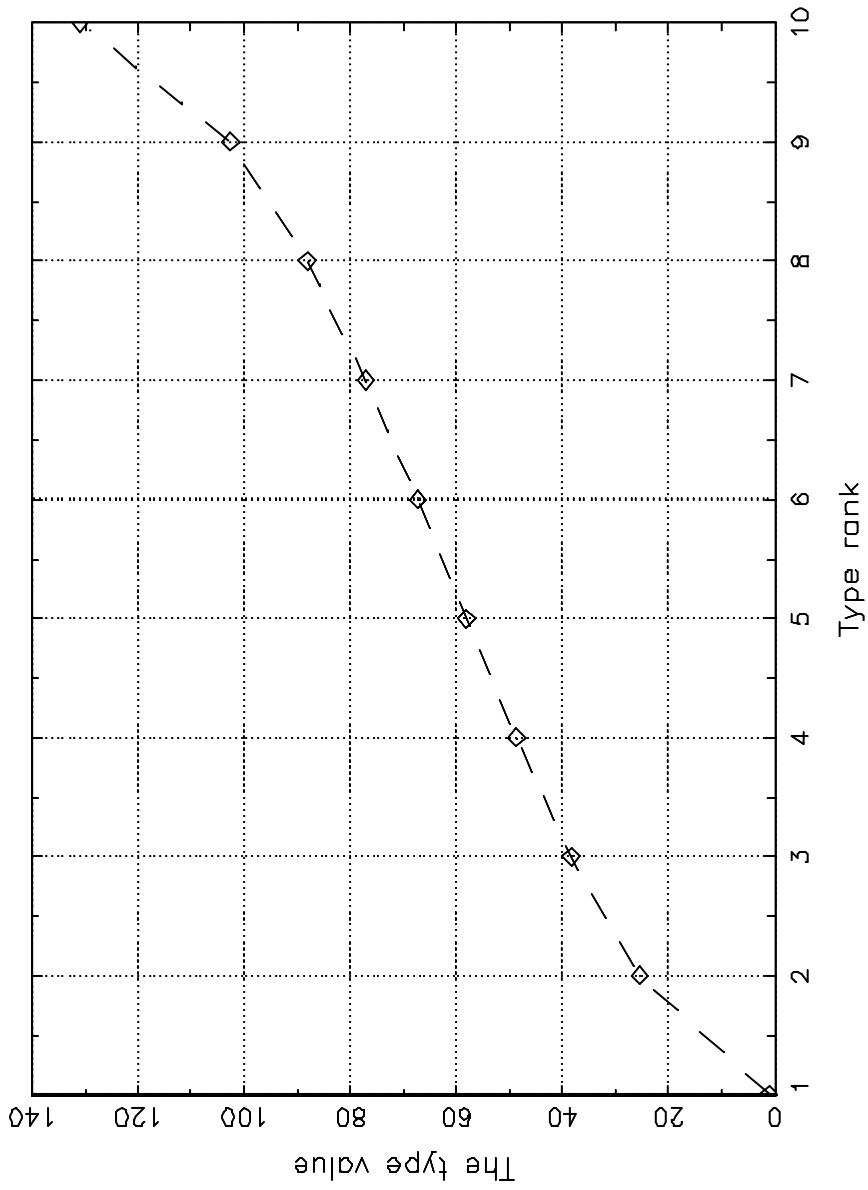


Figure 2: The estimated type space $X$.

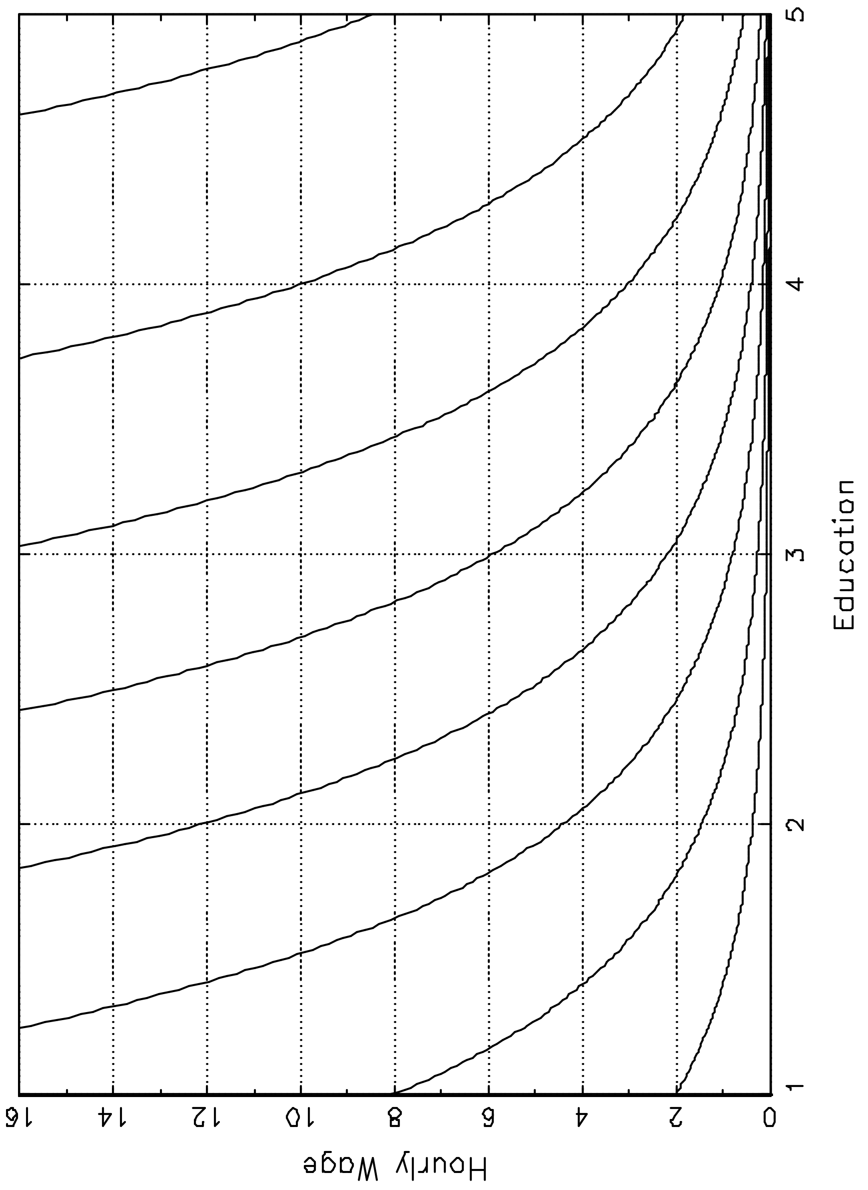


Figure 3: Estimated level curves associated with types $x^2$ through $x^9$.

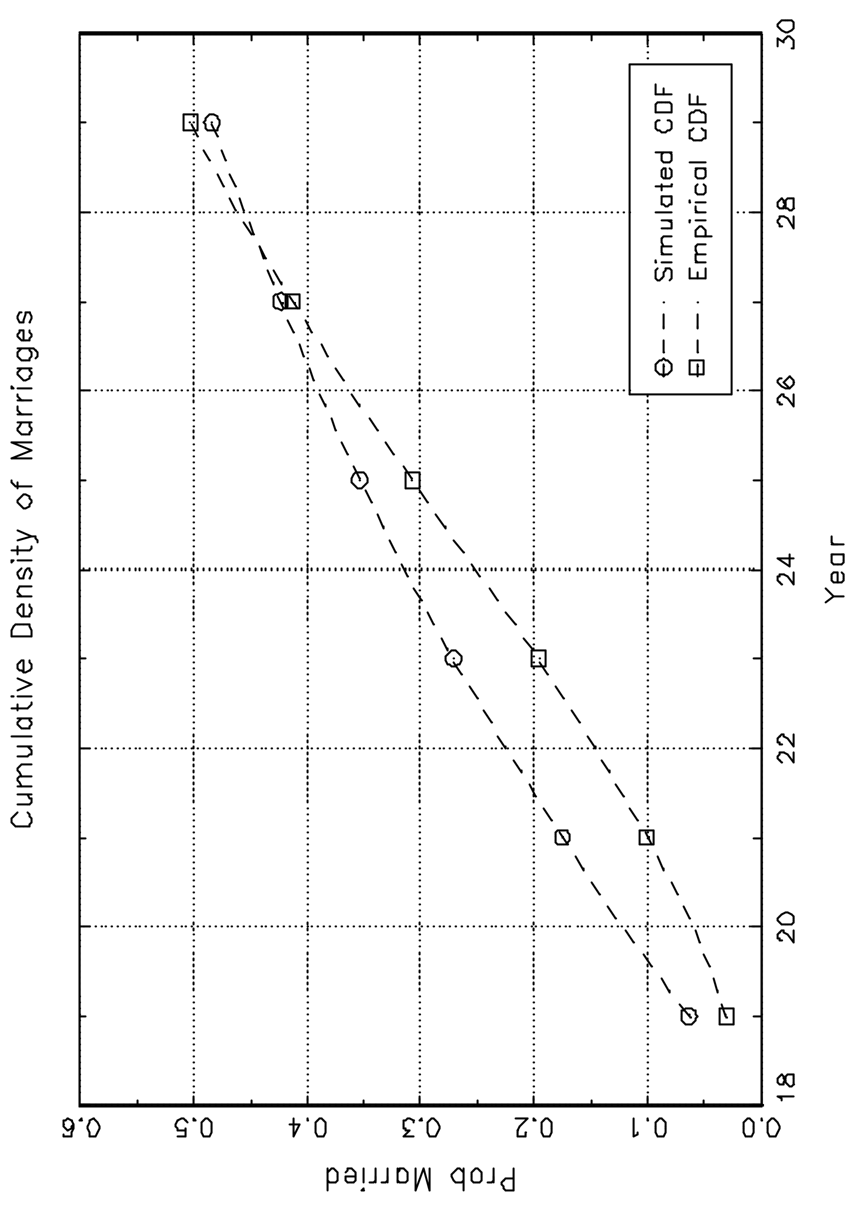


Figure 4: Simulated and actual CDF of marriage dates.

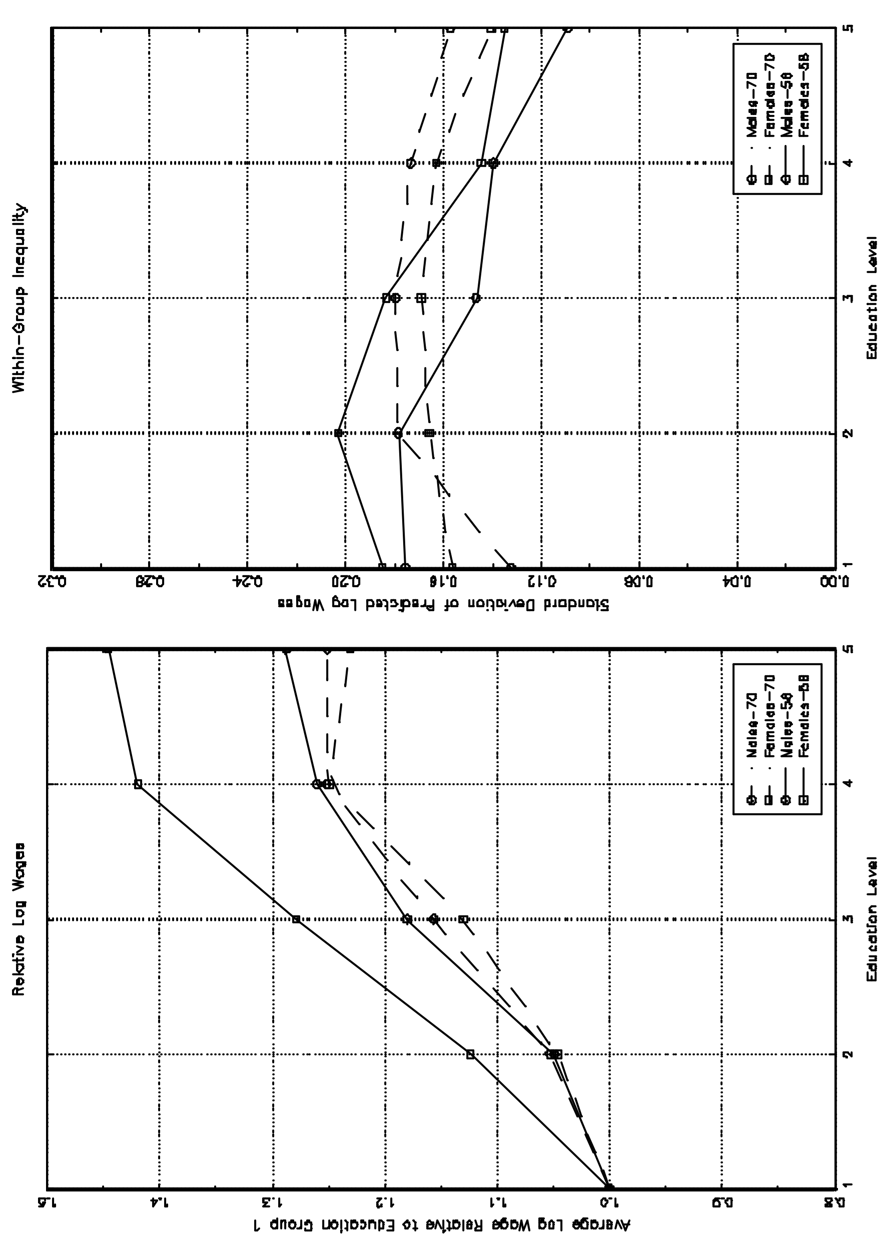


Figure 5: Relative wages and within-group wage variability in the 1958 (NCDS) cohort and the 1970 (BCS) cohort,