

# On the Empirical Content of the Beckerian Marriage Model

Jianfei Cao · Xiaoxia Shi · Matthew Shum

February 18, 2018

**Abstract** This note studies the empirical content of a simple marriage matching model with transferable utility, based on Becker (1973). Under Becker’s conditions, the equilibrium matching is unique and assortative. However, this note shows that, when the researcher only observes a subset of relevant characteristics, the unique assortative matching does not uniquely determine a distribution of observed characteristics. This precludes standard approaches to point estimation of the underlying model parameters. We propose a solution to this problem, based on the idea of “random matching”.

**Keywords.** Beckerian marriage model, Assortative Matching, Indeterminacy, Random matching.

**JEL codes.** C51, C78

## 1 An empirical model based on Becker (1973)

In this note, we investigate the problem of doing empirical analysis on the classical two-sided marriage matching model of Becker (1973). We add to the Beckerian model the unobserved heterogeneity in the form of unobserved characteristics, as is reasonable in empirical applications. Consider a setting where each man (woman) is characterized by  $\tilde{X} \equiv (X, \epsilon)$  ( $\tilde{Y} \equiv (Y, \eta)$ ), where  $X$  and  $Y$  are observables and  $\epsilon$  and  $\eta$  are unobservables. Furthermore each man’s unobserved component  $\epsilon$  (resp. woman’s  $\eta$ ) is drawn from the distribution  $F_\epsilon(\cdot, \beta_\epsilon)$  (resp.  $F_\eta(\cdot, \beta_\eta)$ ), which is known up to the finite-dimensional parameters  $\beta_\epsilon$  ( $\beta_\eta$ ). The characteristics are independently drawn across individuals.

---

We thank Alejandro Robinson-Cortes for research assistance.

Jianfei Cao

University of Chicago Booth School of Business E-mail: jcao0@chicagobooth.edu

Xiaoxia Shi

University of Wisconsin at Madison E-mail: xshi@ssc.wisc.edu

Matthew Shum

California Institute of Technology E-mail: mshum@caltech.edu

Consider a transferable utility two-sided matching market, and let  $S(X, \epsilon, Y, \eta)$  denote the total surplus function that a man with characteristics  $(X, \epsilon)$  and a woman with characteristics  $(Y, \eta)$  obtain from being matched.<sup>1</sup>

**Assumption 1** (a)  $\tilde{X} \sim F_{X, \epsilon}$ , and  $\tilde{Y} \sim F_{Y, \eta}$ ,

(b) each agent's observed and unobserved characteristics affect the surplus function via a single-index; that is to say,  $S(X, \epsilon, Y, \eta, \beta_S) = S(f(X, \epsilon, \beta_S), g(Y, \eta, \beta_S))$ , for scalar-valued known mappings  $f(\cdot)$  and  $g(\cdot)$  and a parameter  $\beta_S$ ;<sup>2</sup> and

(c)  $S(u, v)$  is super-modular in the indices  $u$  and  $v$ .

In what follows, we will refer to  $U := f(X, \epsilon, \beta_S)$  and  $V := g(Y, \eta, \beta_S)$  as the quality indices of men and women, and denote generic values of these elements by  $u$  and  $v$ , respectively. We also let  $\theta = (\beta'_\epsilon, \beta'_\eta, \beta'_S)'$  and let  $\Theta$  be the space of  $\theta$  under consideration.  $\beta_\epsilon$  (or  $\beta_\eta$ ) may contain parameters describing how  $X$  and  $\epsilon$  (or  $Y$  and  $\eta$ ) are correlated.

**Equilibrium assortative matching.** Becker (1973) shows that the single index assumption and super-modularity imply matching that is associative in the indices. Specifically, for fixed values of  $\theta$ , the joint distributions  $F_{X, \epsilon}(\cdot; \beta_\epsilon)$  and  $F_{Y, \eta}(\cdot; \beta_\eta)$  imply unique marginal distributions for the quality indices  $U$  and  $V$ , which are denoted as

$$F_U(u; \beta_\epsilon, \beta_S) \text{ and } F_V(v; \beta_\eta, \beta_S). \quad (1)$$

And the single index super-modularity assumption (Assumption 1) implies assortative matching:

**Proposition 1 (Single Index Assortative Matching Rule)** *Suppose that Assumption 1 holds. Then, a man with  $U = u$  is matched with a woman with  $V = F_V^{-1}(\tau; \beta_\eta, \beta_S)$  where  $\tau = F_U(u; \beta_\epsilon, \beta_S)$ .*

The proposition is implied by the arguments in Section 1 of the Appendix of Becker (1973).

In this note, we investigate the empirical implication of this theoretical result. Specifically, we consider the empirical situation where one observes the joint distribution of the observable characteristics  $F_{X, Y}$  of married couples in the data and would like to learn about  $\theta$  based on  $F_{X, Y}$ . That is, we investigate the possibility of forming an empirical model for the observed data on  $(X, Y)$ , based on Becker's (1973) theoretical result.

**Related literature.** Our work is related to the large and growing empirical literature on the identification and structural estimation of two-sided matching models, including Choo and Siow (2006), Fox (2010), Uetake and Watanabe (2012), Menzel (2013), Echenique, Lee, Shum, and Yenmez (2013), Graham (2013), and Galichon and Salanié (2015).

Particularly, the papers of Choo and Siow (2006) and Galichon and Salanié (2015) estimate transferable utility two-sided matching models for the marriage market, also under the assumption that men and women have both observed and unobserved heterogeneity. However, in those papers, the uniqueness of

<sup>1</sup> The assumptions are similar to those of the transferable utility model studied in Chiappori et. al. (2012), although we do not assume that  $X$  and  $\epsilon$  ( $Y$  and  $\eta$ ) are separable in the single index functions.

<sup>2</sup> In Becker's original formulation, these single-indices are interpreted as the time inputs that a husband or wife contributes towards household production. See also Roth and Sotomayor (1990).

the equilibrium matching directly implies uniqueness of the joint distribution of observed characteristics  $(X, Y)$  among married couples. There are two important differences between their setting and ours. First, the unobserved heterogeneity in their model is better characterized as *preference heterogeneity*, as it allows (say) two different men to obtain different contributions to surplus from the same woman. In our model, unobserved heterogeneity takes the form of *unobserved characteristics*, so that a woman with characteristics  $(Y, \eta)$  contributes the same surplus ( $= g(Y, \eta)$ ) to all potential partners. Second, the models of Chooslow and Galichon-Salanié are not single-index models, but rather multidimensional models, so that the equilibrium matching is not assortative (indeed, in the multidimensional setting, it is not even clear what the notion of “assortative” means). Our model, in contrast, is a unidimensional model (ie. both men and women’s contributions to surplus can be summarized by a single scalar index) so that assortativeness (in surplus indices) is well-defined. Thus, our analysis shows how the assumption of assortative matching can have limited empirical implications.

## 2 Indeterminacy of the matching model

Now we show that even under the strong assumptions guaranteeing assortative matching, the equilibrium distribution of the observed covariates  $X, Y$  is not uniquely determined. This precludes standard point estimation approaches for the structural model parameter  $\theta$ .

**Theorem 1** (a) *Assumption 1 does not imply a unique joint distribution of  $X$  and  $Y$  in married couples:  $F_{X,Y}^*(x, y, \theta)$ , for each  $\theta \in \Theta$ .*

(b) *This nonuniqueness persists even if we further assume that  $F_{X,\epsilon}$  and  $F_{Y,\eta}$  are continuous distributions (w.r.t. the Lebesgue measure).*

*Proof* (a) We prove part (a) by giving a very simple counter example where  $X, \epsilon, Y, \eta$  are independent *Bernoulli*(0.5) and the index functions are additive:

$$U := X + \epsilon \quad \text{and} \quad V := Y + \eta.$$

For a supermodular surplus function  $S(U, V)$ , the optimal assortative matching is characterized by:

$$\begin{aligned} (X, \epsilon) &\iff (Y, \eta) : & (2) \\ (1, 1) &\iff (1, 1) \\ \left. \begin{array}{l} (1, 0) \\ (0, 1) \end{array} \right\} &\iff \left\{ \begin{array}{l} (1, 0) \\ (0, 1) \end{array} \right. \\ (0, 0) &\iff (0, 0) \end{aligned}$$

In an optimal match, any woman with characteristics  $(0, 1)$  or  $(1, 0)$  can match with any man with characteristics  $(0, 1)$  or  $(1, 0)$ . This leads to indeterminacy of the joint distribution of  $(X, Y)$  in the matched population. In particular, for any  $c \in [0, 1]$ , all joint distribution  $F_{X,Y}$  described by the following probability

mass function are consistent with Assumption 1:

$$f_{X,Y}(x, y) = \begin{cases} 0.25 + 0.25c & (x, y) = (1, 1) \\ 0.25(1 - c) & (x, y) = (1, 0) \\ 0.25(1 - c) & (x, y) = (0, 1) \\ 0.25 + 0.25c & (x, y) = (0, 0) \end{cases} \quad (3)$$

(b) We prove part (b) by giving another counter example, where  $(X, Y, \epsilon, \eta)$  are all mutually independent standard normal random variables, and the index functions are additive:  $U = X + \epsilon$ , and  $V = Y + \eta$ . In this case, an assortative matching will match men and women according to the equality

$$X + \epsilon = Y + \eta. \quad (4)$$

This implies a continuum of joint distributions for  $(X, Y)$ . For instance, Eq. (4) is consistent with the matching

$$\{(X, Y, \epsilon, \eta) : X = \eta, Y = \epsilon\}$$

in which case  $(X, Y) \sim N(0, I_2)$ . It is also consistent with the matching

$$\{(X, Y, \epsilon, \eta) : X = Y, \eta = \epsilon\}$$

$(X, Y) \sim N(0, [1, 1; 1, 1])$ . In fact, it is consistent with anything between these polar examples as well.

As we can see from the examples, the key to the indeterminacy of  $F_{X,Y}^*$  is that different combinations of  $X$  and  $\epsilon$  can yield the same value of  $U \equiv f(X, \epsilon, \beta_S)$ , and different combinations of  $Y$  and  $\eta$  can yield the same value of  $V \equiv g(Y, \eta, \beta_S)$ .<sup>3</sup> In other words, a man with  $U = u$  is indifferent to women with the same  $V = F_V^{-1}(F_U(u, \beta_\epsilon, \beta_S), \beta_\eta, \beta_S)$  but different combinations of  $(Y, \eta)$ ; and vice versa. Under the index assumptions, any man (woman) will be indifferent between all matching partners with a given index value, regardless of the particular values of  $Y$  ( $X$ ) and  $\epsilon$  ( $\eta$ ).

Formally, the indeterminacy problem can be described in the following way. The single index assortative matching assumption implies a unique joint distribution of  $(U, V)$ , given by the Frechet upper-bound copula<sup>4</sup>

$$F_{U,V}(u, v, \theta) = \min(F_U(u; \beta_\epsilon, \beta_S), F_V(v; \beta_\eta, \beta_S)). \quad (5)$$

Also, the joint distribution of  $(U, V)$  is related to that of  $(X, Y)$  through the following integral equation:

$$F_{X,Y}^*(x, y, \theta) = F_Y(y) \int \left[ \int F_{X,\epsilon|Y,\eta}(x, e|y, a) de \right] dF_\eta(a). \quad (6)$$

Equation (5) uniquely determines  $F_{f(X,\epsilon,\beta_S)|g(Y,\eta,\beta_S)}(u|v)$  but that is not sufficient to uniquely determine the matched  $F_{X,\epsilon|Y,\eta}(x, e|y, a)$ , causing  $F_{X,Y}^*(x, y, \theta)$  to be underdetermined.

<sup>3</sup> As such, the fact that  $\epsilon$  and  $\eta$  are unobserved has nothing to do with the underlying indeterminacy problem. Even if  $\epsilon$  and  $\eta$  were observed by the researcher, the index structure of the problem still causes the joint distribution of  $(X, Y)$  indeterminate in equilibrium.

<sup>4</sup> This is the joint distribution of  $(U, V)$  having marginal distributions  $F_U$  and  $F_V$  with maximal positive correlation between  $U$  and  $V$ ; see, eg., Joe (1997).

Indeed, the single index assumption lies behind this indeterminacy feature: in the index function, many combinations of  $(X, \epsilon)$  – the “matching set”, as it is called in the literature<sup>5</sup> – lead to the same value of  $U$  (and analogously for  $V$ ), necessarily implying a one-to-many mapping from  $F_{U,V}$  to  $F_{X,Y}$ . This problem cannot be resolved with only shape restrictions on the index functions or the marginal distributions of  $\epsilon$  and  $\eta$ .

The indeterminacy problem **pointed out in the previous section** makes standard parametric estimation of this model difficult. In particular, a typical approach to estimation would be to find a value of the parameter vector  $\theta$  to satisfy the following system of equations, which equate the equilibrium distribution of  $(X, Y)$ , depending on unknown parameters  $(\theta)$ , to the distribution of  $(X, Y)$  observed by the researcher:

$$F_{X,Y}(x, y) = F_{X,Y}^*(x, y, \theta), \quad \forall x, y. \quad (7)$$

This system of equations, linking observed to unobserved elements, fully summarizes the empirical content of the model and can be used to estimate the the parameter of interest,  $\theta$ , if the joint distribution  $F_{X,Y}(x, y)$  can be estimated from the data. However, as we pointed out in the previous section, the equilibrium distribution of  $(X, Y)$ ,  $F_{X,Y}^*(\cdot, \cdot, \theta)$ , may not be uniquely determined, which makes this system of equations ill-defined. Similarly, maximum likelihood estimation of the parameters  $\theta$  also is precluded by the indeterminacy of  $F_{X,Y}(\cdot, \cdot; \theta)$ .

Relatedly, the indeterminacy of joint distribution of observable characteristics can prevent point identification of the model parameters. To see this, consider a binary Bernoulli example where the index functions are  $U = X + \epsilon$  and  $V = Y + \eta$ .  $X$  and  $Y$  are independent *Bernoulli*(0.5), and  $\epsilon$  and  $\eta$  are independent *Bernoulli*( $\theta$ ). The equilibrium matching in this setting is assortative in the indices, and characterized by Eq. (2) above. If we assume in addition that people only match with those who have the exactly the same characteristics with themselves, the joint distribution of  $(X, Y)$  would be a pair of perfectly correlated *Bernoulli*(0.5) random variables. This joint distribution is consistent with *any*  $\theta \in [0, 1]$ , so that  $\theta$  is not point identified.

### 3 Resolving Indeterminacy via Random Matching

We propose one resolution of the indeterminacy problem using the notion of “random matching”. Roughly, random matching means that agents with non-singleton matching sets have an equal chance of being matched to any element of their matching sets. Formally, this is specified in the following assumption:<sup>6</sup>

**Assumption 2 (Random Matching)** *The joint distribution of  $(X, \epsilon, Y, \eta)$  for a married couple satisfies*

$$F_{X,\epsilon|Y,\eta}(x|y, a) = F_{X,\epsilon|g(Y,\eta)}(x, e|g(y, a)), \quad \text{and} \quad F_{Y,\eta|X,\epsilon}(y, a|x, e) = F_{Y,\eta|f(X,\epsilon)}(y, a|f(x, e)).$$

*That is, two women (men) with respectively characteristics  $(y, a)$  and  $(y', a')$  ( $(x, e)$  and  $(x', e')$ ), such that  $g(y, a) = g(y', a')$  ( $f(x, e) = f(x', e')$ ), have equal opportunities among men (women).*

<sup>5</sup> See e.g. Shimer and Smith (2000), Atakan (2006).

<sup>6</sup> In the assumption and the rest of this section, we ignore the parameter  $\theta$  for notational simplicity.

**Comments.** (a) This assumption only requires the matches to occur randomly among potential matches with the same surplus. Thus, the randomness does not affect the optimality of the matching outcome.

(b) Assumption 2 may be micro-founded as limit versions of the search and matching models of (i) Shimer and Smith (2000), as we take the agents' patience to infinity; or (ii) Atakan (2006), as the agents' search cost approaches zero. Indeed, the random arrival of potential partners in those search models provides a natural randomization mechanism, and the ex ante symmetry of the agents guarantees the "equal opportunity" part of the assumption.

(c) In the Bernoulli example in the proof of Theorem 1(a), this assumption is equivalent to imposing  $c = 0.5$ .  $\square$

The following theorem shows that random matching resolves the indeterminacy problem.

**Theorem 2** *Assumptions 1 and 2 together imply a unique joint distribution of  $X$  and  $Y$  in married couples:  $F_{X,Y}^*(x, y)$ .*

*Proof* Let the asterisk superscription denote the distribution function that involves variables from both sides of the market, which could potentially be undetermined without the random matching assumption. Under Assumption 2, the equilibrium matching can be defined by the following two conditions:

- (i)  $F_{U,V}(u, v) = \min\{F_U(u), F_V(v)\}$ ; and
- (ii)  $F_{X|Y,\eta}(x|y, a) = F_{X|g(Y,\eta)}(x|g(y, a))$  and  $F_{Y|X,\varepsilon}(y|x, e) = F_{Y|f(X,\varepsilon)}(y|f(x, e))$ .

For any  $(y, a, x, e)$  such that  $F_U(f(x, e)) > 0$  and  $F_V(g(y, a)) > 0$ , consider the derivation:

$$\begin{aligned}
F_{Y,\eta,X,\varepsilon}(y, a, x, e) &= F_{Y,\eta|X,\varepsilon}(y, a|x, e)F_{X,\varepsilon}(x, e) \\
&= F_{Y,\eta|U}(y, a|f(x, e))F_{X,\varepsilon}(x, e) \\
&= F_{Y,\eta,U}(y, a, f(x, e)) \frac{F_{X,\varepsilon}(x, e)}{F_U(f(x, e))} \\
&= F_{U|Y,\eta}(f(x, e)|y, a)F_{Y,\eta}(y, a) \frac{F_{X,\varepsilon}(x, e)}{F_U(f(x, e))} \\
&= F_{U|V}(f(x, e)|g(y, a))F_{Y,\eta}(y, a) \frac{F_{X,\varepsilon}(x, e)}{F_U(f(x, e))} \\
&= F_{U,V}(f(x, e), g(y, a)) \frac{F_{Y,\eta}(y, a)}{F_V(g(y, a))} \frac{F_{X,\varepsilon}(x, e)}{F_U(f(x, e))} \\
&= \min\{F_U(f(x, e)), F_V(g(y, a))\} \frac{F_{Y,\eta}(y, a)}{F_V(g(y, a))} \frac{F_{X,\varepsilon}(x, e)}{F_U(f(x, e))}, \tag{8}
\end{aligned}$$

where the first, third, fourth, and sixth equalities hold by the Bayes rule, the second and the fifth equalities hold by (ii), and the last equality holds by (i). The expression in the last line of the above display is fully and uniquely determined by the bivariate distribution functions  $F_{X,\varepsilon}$  and  $F_{Y,\eta}$ . Thus, the joint distribution of  $F_{Y,\eta,X,\varepsilon}$  is uniquely determined by those bivariate distributions. Finally, the theorem is proved by noting that the matched joint distribution  $F_{X,Y}^*$  can be obtained by

$$F_{X,Y}^*(x, y) = \int \int F_{Y,\eta,X,\varepsilon}(y, a, x, e) da de. \tag{9}$$

Next we illustrate Theorem 2 in the Bernoulli and the Standard Normal examples described in the proof of Theorem 1 above.

**Bernoulli Example, continued.** In this example, random matching implies that the  $(1, 0)$  and  $(0, 1)$  women match “randomly” with the  $(1, 0)$  and  $(0, 1)$  men. Formally, under Assumption 2, the matched joint distribution of  $(X, \epsilon, Y, \eta)$  satisfies

- (i)  $\Pr(U = V) = 1$  (assortative matching and identical distribution of  $U$  and  $V$ ); and
- (ii)  $\Pr(X = 1|Y = y, \eta = a) = \Pr(X = 1|Y + \eta = y + a)$  (random matching)

The two conditions uniquely determine the joint distribution of  $X$  and  $Y$  in the matched population by the following derivation. First observe that

$$\begin{aligned} \Pr(X = 1|Y = 1, \eta = 1) &= \Pr(X = 1|V = 2) = \Pr(X = 1|U = 2) \\ &= \Pr(X = 1|X = 1, \epsilon = 1) = 1 \end{aligned} \tag{10}$$

where the first equality holds by condition (ii) and the second equality holds by condition (i). Similarly,

$$\begin{aligned} \Pr(X = 1|Y = 0, \eta = 0) &= \Pr(X = 1|V = 0) = \Pr(X = 1|U = 0) \\ &= \Pr(X = 1|X = 0, \epsilon = 0) = 0 \end{aligned} \tag{11}$$

$$\begin{aligned} \Pr(X = 1|Y = 1 - \eta) &= \Pr(X = 1|V = 1) = \Pr(X = 1|U = 1) \\ &= \Pr(X = 1|X = 1, \epsilon = 0 \text{ or } X = 0, \epsilon = 1) = 0.5 \end{aligned} \tag{12}$$

(13)

Therefore,

$$\begin{aligned} \Pr(X = 1|Y = 1) &= 0.5\Pr(X = 1|Y = 1, \eta = 1) + 0.5\Pr(X = 1|Y = 1, \eta = 0) = 0.75 \\ \Pr(X = 1|Y = 0) &= 0.5\Pr(X = 0|Y = 0, \eta = 1) + 0.5\Pr(X = 1|Y = 0, \eta = 0) = 0.25 \end{aligned} \tag{14}$$

These fully and uniquely characterize the joint distribution of  $X$  and  $Y$ . ■

**Bivariate normal example, continued.** For the bivariate normal example, under Assumption 2, the matched joint distribution of  $(X, \epsilon, Y, \eta)$  satisfies

- (i)  $\Pr(U = V) = 1$ ; and
- (ii)  $f_{X|Y, \eta}(x|y, a) = f_{X|Y + \eta}(x|y + a)$ .

Now we show that in the optimal match defined by (i) and (ii), the unique distribution of observed characteristics  $(X, Y)$  is mean-zero bivariate normal with variance covariance matrix  $[1, 0.5; 0.5, 1]$ . Observe that

$$\begin{aligned} f_{X|Y, \eta}(x|y, a) &= f_{X|Y + \eta}(x|y + a) = f_{X|X + \epsilon}(x|y + a) \\ &= f_{X, X + \epsilon}(x, y + a) / f_{X + \epsilon}(y + a) \\ &= \frac{(2\pi)^{-1} \exp(-2^{-1}(2x^2 - 2x(y + a) + (y + a)^2))}{(4\pi)^{-1/2} \exp(-4^{-1}(y + a)^2)} \\ &= \pi^{-1/2} \exp(-x^2 + xy + xa - (y + a)^2/4), \end{aligned} \tag{15}$$

where the second equality holds with probability one by condition (i). Thus,

$$f_{X|Y}(x|y) = \int_{-\infty}^{\infty} f_{X|Y, \eta}(x|y, a) f_{\eta}(a) da$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \pi^{-1/2} \exp(-x^2 + xy + xa - (y+a)^2/4) (2\pi)^{-1/2} \exp(-a^2/2) da \\
&= \frac{1}{\sqrt{2\pi}} \exp(-x^2 + xy - y^2/4) \int_{-\infty}^{\infty} \exp(xa - ya/2 - 3a^2/4) da,
\end{aligned} \tag{16}$$

where the first equality holds due to the independence of  $Y$  and  $\eta$ . Therefore,

$$\begin{aligned}
f_{X,Y}(x, y) &= f_{X|Y}(x|y) f_Y(y) \\
&= \frac{1}{2\pi^{3/2}} \exp(-x^2 + xy - 3y^2/4) \int_{-\infty}^{\infty} \exp(xa - ya/2 - 3a^2/4) da \\
&= \frac{1}{\sqrt{3\pi}} \exp(-(x, y) \Sigma^{-1} (x, y)' / 2),
\end{aligned} \tag{17}$$

where  $\Sigma = [1, 0.5; 0.5, 1]$ . This shows that  $(X, Y) \sim N(0, \Sigma)$ . ■

### 3.1 Estimation strategies

After establishing the uniqueness of  $F_{X,Y}(\cdot, \cdot; \theta)$ , either a simulated likelihood or a simulated method of moments approach can be used to estimate the model parameter  $\theta$ . The  $F_{X,Y}^*(\cdot, \cdot, \theta)$  function, which often does not have closed form solution, can be simulated via the following steps.

- Step 1.** Independently draw  $N$  i.i.d. observations of  $(X, \epsilon)$  and  $(Y, \eta)$  from  $F_{X,\epsilon}(\cdot, \cdot, \beta_\epsilon)$  and  $F_{Y,\eta}(\cdot, \cdot, \beta_\eta)$  respectively for a large number  $N$ . Denote the draws by  $\{X_i, \epsilon_i\}_{i=1}^N$  and  $\{Y_i, \eta_i\}_{i=1}^N$ .
- Step 2.** Compute  $\{U_i = f(X_i, \epsilon_i, \beta_S)\}_{i=1}^N$  and  $\{V_i = g(Y_i, \eta_i, \beta_S)\}_{i=1}^N$ .
- Step 3.** Sort both  $\{(X_i, U_i)\}$  and  $\{(Y_i, V_i)\}$  in decreasing order of  $U_i$  and  $V_i$  respectively. Ensure that ties (if any) are sorted randomly. Let the sorted data be denoted  $\{\tilde{X}_i, \tilde{\epsilon}_i, \tilde{Y}_i, \tilde{\eta}_i\}_{i=1}^N$ .
- Step 4.** Then, compute the simulated  $F_{X,Y}^*(x, y; \theta)$  as follows:

$$F_{X,Y}^*(x, y; \theta) = N^{-1} \sum_{i=1}^N 1(\tilde{X}_i \leq x, \tilde{Y}_i \leq y). \tag{18}$$

Both simulated maximum likelihood and simulated method of moments are well-studied estimation procedures in the econometrics literature.<sup>7</sup> The asymptotic properties of the resulting estimators are standard and well known, and thus we omit such discussions in this note.

## 4 Simulation and Empirical Application

In this section we consider a set of Monte Carlo simulations, followed by an empirical application using census data from China. We specify  $g$  and  $f$  as follows:

$$f(X, \epsilon) = X' \epsilon, \quad g(Y, \eta) = Y' \eta, \tag{19}$$

where  $X$  is a  $d_x$  vector of observables for women, and  $Y$  is a  $d_y$  vector of observables for men, and  $\epsilon$  and  $\eta$  are respectively the random weights of those observables in the utility indices for women and men.

We carry out both Monte Carlo and empirical application using simulated GMM. Potentially there are infinite number of moment conditions to use because  $E[1(X \leq x, Y \leq y)] = F_{X,Y}^*(x, y, \theta)$  needs to hold at

<sup>7</sup> See Gourieroux and Montfort (1997), for example.



all points of  $(x, y)$ . For implementation, we use a finite number of grid points on  $(x, y)$  and use two-step GMM to estimate  $\theta$ , where the procedure in Section 3.1 is used to simulate the equilibrium matching for each candidate parameter vector  $\theta$ .

#### 4.1 Monte Carlo

For our Monte Carlo exercise, we consider three different models. For all three, we let  $X = [X_1, X_2]'$ , where  $X_1 \sim N(0, 1)$  and  $X_2$  is drawn from  $\{-1, 0, 1\}$  with equal probability. Similarly, let  $Y = [Y_1, Y_2]'$ , where  $Y_1 \sim N(1, 1)$  and  $Y_2$  is drawn from  $\{0, 1, 2, 3\}$  with equal probability. We also let all random variables be independently drawn.

**Model 1** Consider a simplest model with  $\varepsilon$  and  $\eta$  being non-random. Normalize the first entries of  $\varepsilon$  and  $\eta$  to one, with  $\varepsilon = [\varepsilon_1, \varepsilon_2]' = [1, 2]'$  and  $\eta = [\eta_1, \eta_2]' = [1, 0.5]'$ . Therefore, our parameter of interest in the model is  $\theta_0 = (\varepsilon_2, \eta_2) = (2, 0.5)$ .

**Model 2** We add randomness in this model while keeping the first entries of  $\varepsilon$  and  $\eta$  constant. Let Model 2 be exactly like Model 1 except that  $\varepsilon_2 \sim N(\mu_\varepsilon, \sigma_\varepsilon^2)$  and  $\eta_2 \sim N(\mu_\eta, \sigma_\eta^2)$ . Let our parameter of interest be  $\theta_0 = (\mu_\varepsilon, \sigma_\varepsilon^2, \mu_\eta, \sigma_\eta^2) = (2, 0.5, 0.5, 1)$ .

**Model 3**  $\varepsilon_1$  and  $\eta_1$  are also random, but the distribution is normalized to be  $N(1, 1)$ . Therefore, our parameter of interest is still  $\theta_0 = (\mu_\varepsilon, \sigma_\varepsilon^2, \mu_\eta, \sigma_\eta^2) = (2, 0.5, 0.5, 1)$ .

We draw a sample size of 10,000 and set number of simulations for simulated moments as 1,000,000. We report our results in Table 1. For comparison, we also report the result of the canonical correlation method of Dupuy and Galichon (2012). They showed that the canonical correlation method is consistent in Model 1 where there is no random coefficient and  $X$  and  $Y$  are normal. For optimization, we use the simulated annealing method and report both the results using the true value as starting value (SA-True) and using a randomly generated starting value (SA-Random).<sup>8</sup>

As we can see, our method outperforms the canonical correlation method in estimating the means of the random coefficients in all three models (with  $\mu_\eta$  in Model 3 as an exception).<sup>9</sup>

#### 4.2 Empirical Application

In this subsection, we report an empirical application using the 0.95% sample of the Chinese 2000 census. We consider *age* and *years of education* as conditioning covariates, and examine how they affect the marriage matching function over time. Both *age* and *years of education* are continuous random variables. We divide the data according to the year when a couple got married, to see whether the matching function has changed over time. For the age variable, we use either the age of the agent in 2000 (“current age”), or the age when he or she got married (“marriage age”). Since differences in the marriage distribution according to marriage vs. current age arise from marriage dissolution, through either divorce/separation or (less likely) death, the

<sup>8</sup> We also tried Nelder-Mead, but the performance is poor because it tends to get stuck in local minima.

<sup>9</sup> In a sense, this is not a fair comparison because the canonical correlation method is not shown to be consistent when the coefficients are random or when the covariates are non-normal.

Table 1: Simulated Bias and Variance of Parameter Estimators

Model 1	Canonical Correlation		SA-True		SA-Random	
	<i>Bias</i>	<i>Var</i>	<i>Bias</i>	<i>Var</i>	<i>Bias</i>	<i>Var</i>
$\varepsilon_2$	-0.1565	0.0001	0.0262	0.0670	0.0015	0.0009
$\eta_2$	0.0065	0.0000	0.0019	0.0000	0.0003	0.0000
Model 2						
$\mu_\varepsilon$	0.3388	0.0073	0.0027	0.0040	0.0045	0.0041
$\sigma_\varepsilon^2$	—	—	0.0079	0.0072	0.0128	0.0078
$\mu_\eta$	-0.0739	0.0003	-0.0036	0.0009	-0.0048	0.0009
$\sigma_\eta^2$	—	—	0.0368	0.0047	0.0343	0.0046
Model 3						
$\mu_\varepsilon$	0.5225	0.0248	0.0171	0.0187	0.0200	0.0199
$\sigma_\varepsilon^2$	—	—	0.0606	0.0489	0.0721	0.0537
$\mu_\eta$	0.0255	0.0008	0.0330	0.0026	0.0310	0.0028
$\sigma_\eta^2$	—	—	0.0359	0.0301	0.0284	0.0315

difference in the results may shed light on the nature of divorce costs, an avenue which we do not pursue here.

Table 2 gives the empirical results. The matching functions follow the form in (19). The variables  $X$  and  $Y$  are both two-dimensional, where the first dimension stands for years of schooling and the second is either current age or marriage age. The model we used is Model 2 described in the Monte Carlo subsection above.<sup>10</sup> For both models, the parameters of interest are  $\theta_0 = (\mu_\varepsilon, \sigma_\varepsilon^2, \mu_\eta, \sigma_\eta^2)$ . Since  $X'\varepsilon$  represents the quality index of a man, which is identically perceived by all women, we can interpret  $\varepsilon$  as the idiosyncratic “weight” that a man’s education has in his quality index. Thus  $\mu_\varepsilon$  represents the average weight of education, and  $\sigma_\varepsilon$  is the variation in this across all men. An analogous interpretation applies to the women’s parameters  $\mu_\eta$  and  $\sigma_\eta$ .

There are 256,663 couples that we are able to track both the marriage age and the current age in the data. We divided observations into six groups according to when they got married, and treated each group as one marriage market. We also provide estimates for the whole population to use as a benchmark (also visualized by horizontal dashed lines in Figure 1).

As shown in Table 2 as well as Figure 1, results by using current age and by marriage are qualitatively similar, so we focus on interpreting the estimates obtained by using marriage age. The estimates for  $\mu_\varepsilon$  (or  $\mu_\eta$ ) can be interpreted as the average ratio of importance of age and education in marriage market. For example, before 1951, one year of education is on average as valuable as being roughly a half-year (0.4532) older for men. We observe that the estimated average education weight for women ( $\mu_\eta$ ) has a sharp fall in the 1980s, while the parameter for men ( $\mu_\varepsilon$ ) exhibits a steady upward trend. These trends are plotted in Figure 1.

One possible explanation for these trends, which is consistent with some discussions in the public press, is that this finding represents some long-term impact of the infamous Cultural Revolution (1966-76) in

<sup>10</sup> We also estimated Model 3, although the result is harder to interpret because of the different way of normalization. The results are available upon request.

Table 2: Empirical Results

	before 1951	1951-1960	1961-1970	1971-1980	1981-1990	1991-2000	Overall
current age							
$\mu_\varepsilon$	0.3565	0.3783	0.6995	0.4098	0.2307	0.5268	0.6727
$\sigma_\varepsilon^2$	2.1416	0.5160	0.2174	0.1069	0.0538	0.1373	0.0008
$\mu_\eta$	0.2410	0.4150	0.6462	0.5490	0.7648	1.0870	0.8291
$\sigma_\eta^2$	1.1602	0.4227	0.0990	0.1067	0.0742	0.0089	0.0016
marriage age							
$\mu_\varepsilon$	0.4532	0.8132	1.0389	0.9786	0.4126	0.7083	1.0153
$\sigma_\varepsilon^2$	1.6218	1.8299	0.0678	0.3301	0.0713	0.0002	0.2145
$\mu_\eta$	0.4691	0.7212	0.9345	1.0918	1.2501	1.9554	1.4277
$\sigma_\eta^2$	1.2532	0.7660	0.4138	0.0809	0.1671	0.4050	0.1476
Observations	6683	17513	35531	55949	87611	53376	256663

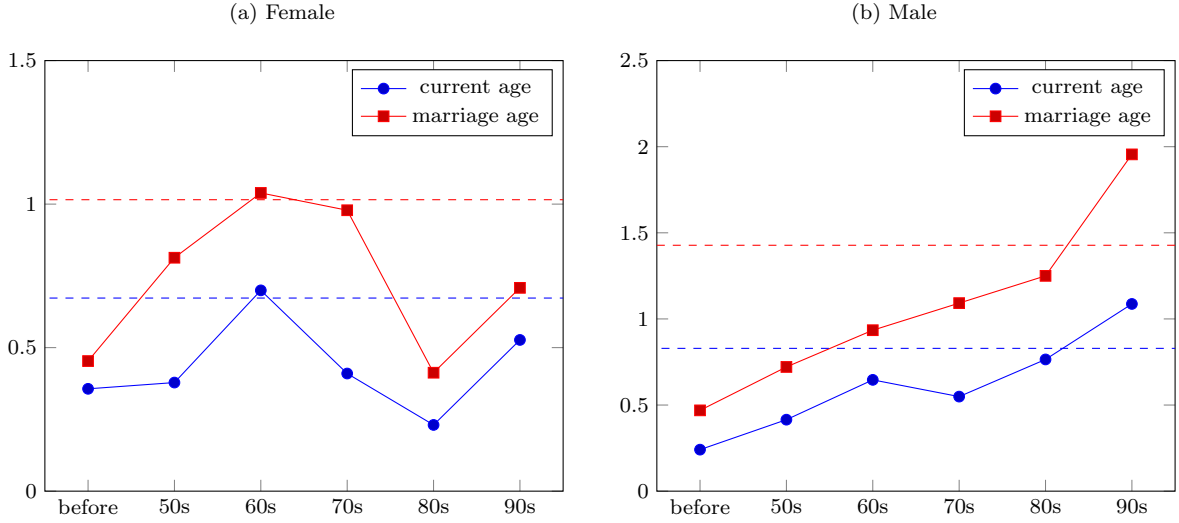


Fig. 1: Evolution of effect of education on marriage market relative to effect of age

China, in which education was de-emphasized and educated urban “elites” were demonized and exiled to the countryside. Those who became of marriageable age in the 1980s were exactly those who grew up during the cultural revolution. Nevertheless, this is only a partial explanation, as our results suggest that the value of education fell only for women but not for men. We will investigate these effects further in future research.

However, the above discussion is based on an assumption that people’s valuation towards age, or the importance of age in the marriage market, is relatively stable across time, because we normalize the weight of age in our model to be one. Because of this normalization, another valid interpretation of the increase in the valuation of men’s education is that a man’s age has become less important over time.

## 5 Concluding remark

In this note, we have focused throughout on uniqueness of the joint distribution of  $(X, Y)$ , as it makes standard parametric methods for point estimates of the model parameters  $(\beta_\epsilon, \beta_\eta, \theta)$  (via MLE or method of moments) feasible. It may be possible to (partially) estimate the model parameters using incomplete model approaches. We do not investigate that possibility here because it is not clear how to generate estimating equalities or inequalities in our setup; furthermore, there are also computational and inferential difficulties which are beyond the scope of this short note. Rather, we focus on resolving the indeterminacy issue via the additional assumption of random matching – which may be quite realistic in a real-world marriage application as it is justified by search-theoretic arguments.

## References

1. Atakan, A. E. (2006), “Assortative Matching with Explicit Search Cost,” *Econometrica*, Vol. 74, pp. 667–680.
2. Becker, G. (1973), “A Theory of Marriage, Part 1,” *Journal of Political Economy*, Vol. 81, pp. 813-846.
3. Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2012), “Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market,” *Journal of Political Economy*, Vol. 120, pp. 659-695.
4. Choo, E. and A. Siow (2006), “Who marries whom and why,” *Journal of Political Economy*, Vol. 114, pp. 175-201.
5. Dupuy, A., & A. Galichon (2015), “Canonical Correlation and Assortative Matching: a remark,” *Annals of Economics and Statistics*, no. 119/120, pp. 375-383.
6. Echenique, F., S. Lee, M. Shum, and B. Yenmez (2013), “The revealed preference theory of stable and extremal stable matchings,” *Econometrica*, Vol. 81, pp. 153-171.
7. Fox, J. (2010), “Identification in Matching Games,” *Quantitative Economics*, Vol. 1, pp. 203-254.
8. Galichon, A., and B. Salanié (2015), “Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models,” Available at SSRN: <https://ssrn.com/abstract=1804623>.
9. Gourieroux, C. and A. Monfort (1997), *Simulation-Based Econometrics Methods (Core Lectures)*, Oxford University Press, New York.
10. Graham, B. (2013), “Comparative Static and Computational Methods for an Empirical One-to-one Transferable Utility Matching Model,” in *Structural Econometric Models (Advances in Econometrics, Vol. 31)*, Emerald Press.
11. Joe, H. (1997), *Multivariate Models and Multivariate Dependence Concepts*, CRC Press.
12. Menzel, K. (2013), “Large Matching Markets as Two-sided Demand Systems,” *Econometrica* Vol. 83, pp. 897-941.
13. Roth, A. and M. Sotomayor (1990), *Two-Sided Matching*, Econometric Society Monographs, Cambridge University Press.
14. Shimer, R. and L. Smith (2000), “Assortative Matching and Search,” *Econometrica* Vol. 68, pp. 342–369.
15. Uetake, K. and Y. Watanabe (2012), “Entry by Merger: Estimates from a Two-Sided Matching Model with Externalities,” Working paper.