Econometrics of the marriage market

Alfred Galichon
"Matching Market" course
Sciences Po

April, 10 2012

Announcements

- Rescheduled class will be held next Tuesday on April 16 8am-10am.
  Location is bâtiment A (27, rue Saint-Guillaume) Room 32.

- A 90mn exam will take place during last lecture on April 17, 8am-9.30am. Only handouts distributed in class and handwritten notes will be allowed.

1 Introduction

Econometrics of the marriage market: a few socio-economic variables (education, age, income, race, wealth, sometimes biometric data...) and a lot of unobserved heterogeneity both in terms of characteristics and in taste.

One would like to understand how the market clears in order to examine sociological theories (endogamy, assortativeness, transmission of social capital...) or economic/public policy issues (eg. impact on divorce law on gender inequality).

Revealed preference problem: What do observed marriage patterns reveal about mutual preferences of partners? What is the surplus of both partners at equilibrium?

In other words, how to identify matching surpluses based on observed actual marriages?

Beyond marriage, the analysis also extends to other matching settings, such as e.g. the market of CEOs. What are the complementarities between firm and CEO's characteristics?

The framework is a matching model with transferable utility and unobserved heterogeneity. The object of interest is the joint surplus generated by a match between two partner types.

- We observe: socio-economic, sometimes biometric data on observed characteristics of both partners in matches and singles. Sometimes, transfers between them (salaries in the case of CEOs).

- We would like to: Estimate a parametric form of the surplus function, in particular in order to test cross-assortativeness or complementarities between the observable characteristics of both partners.

In most cases, we will observe a single market. Sometimes, we observe several disconnected markets where participants may be assumed to have similar characteristics distributions.

2 Equilibrium on the marriage market

Becker (JPE 1973–74): Marriage as a competitive matching market with transferable utility, with one dimensional characteristics and no heterogeneities.

Yet these yield to too stark predictions: "Positive assortative matching" on a single-dimensional "ability index".

A few years ago, Choo and Siow (JPE 2006) (hereafter, CS) have incorporated logit-type heterogeneities in Becker's model and show that in this framework, the marital surplus can be nonparametrically identified. This started a rich literature on identification and estimation in matching models.

2.1 Reminder: the Becker-Shapley-Shubik theory of marriage

Transferable utility: surplus of a pair can be split without restrictions between man and woman. Static matching, no frictions. Observable types are discrete. We recall the Becker-Shapley-Shubik setting first. Consider a population with n_x men of type x, and m_y women of type y. Introduce:

- α_{xy} utility of man x with woman y, 0 if single

- γ_{xy} utility of woman y with man x, 0 if single

Transferable utility: τ_{xy} utility transfer from x to y. Then

- $\alpha_{xy} - \tau_{xy}$ post-transfer utility of man x

- $\gamma_{xy} + \tau_{xy}$ post-transfer utility of woman y

Then the market clears in order to maximize

$$\max_{\mu \geq 0} \sum_{x,y} \mu_{xy} \Phi_{xy} : \sum_{y} \mu_{xy} \leq n_x, \; \sum_{x} \mu_{xy} \leq m_y$$

where

- $\Phi_{xy} = \alpha_{xy} + \gamma_{xy}$ is the total gains to marriage,

- μ_{xy} is the number of (x, y) pairs,

$\mu_{x0} = n_x - \sum_y \mu_{xy}$ is the number of single men of type x, and $\mu_{0y} = m_y - \sum_x \mu_{xy}$ is the number of single women of type y.

2.2 Choo and Siow's model

$|\mathcal{X}|$ groups of men of same observable characteristics, indexed by x; $|\mathcal{Y}|$ groups of women, indexed by y (education, race, income, religion...). Market participants observe everybody's full characteristics – analyst does not.

Choo and Siow: utility of a man m of group x who marries a woman of group y can be written:

$$\alpha_{xy} - \tau_{xy} + \varepsilon_{xym},$$

where $\tau_{xy} =$ utility transfer in equilibrium, and ε_{xym} is a standard type-I E.V. unobserved heterogeneity. If single, gets utility

$$0 + \varepsilon_{x0m},$$

(0 is a choice of normalization w.l.o.g.). Similarly, the utility of a woman w of group y who marries a man of group x can be written as

$$\gamma_{xy} + \tau_{xy} + \eta_{xyw},$$

and she gets utility

$$0 + \eta_{0yw}$$

is she is single.

Denote μ_{xy} the number of marriages between men of group x and women of group y; μ_{x0} the number of single men of group x; and μ_{0y} the number of single women of group y.

Problem. *The matching surpluses (α_{xy} and γ_{xy}) are not observed; only matching patterns μ_{xy} are observed. What are the restrictions on the surpluses?*

Choo and Siow proved the following result:

Theorem 0 (Choo and Siow). *In equilibrium, if there are very large numbers of men and women within each group,*

$$\exp(\frac{\Phi_{xy}}{2}) = \frac{\mu_{xy}}{\sqrt{\mu_{x0}\mu_{0y}}},$$

where Φ_{xy} *denotes the* total systematic net gains to marriage*:*

$$\Phi_{xy} = \alpha_{xy} + \gamma_{xy}.$$

Therefore marriage patterns directly identify the gains to marriage Φ *in such a model.*

As we shall show, one can in fact extend significantly Choo and Siow's model to allow for a much larger class of unobservable heterogeneities and yet get tractable models.

2.3 Assumptions

Denote n_x the number of men of group x, and m_y the number of women of group y; then

$$\forall x \geq 1,\ \ \sum_{y=0}^{|\mathcal{Y}|} \mu_{xy} = n_x \ ;\ \forall y \geq 1,\ \ \sum_{x=0}^{|\mathcal{X}|} \mu_{xy} = m_y. \tag{1}$$

For future reference, we denote $\mathcal{M}$ the set of $(|\mathcal{X}|\,|\mathcal{Y}| + |\mathcal{X}| + |\mathcal{Y}|)$ non-negative numbers (μ_{xy}) that satisfy these $(|\mathcal{X}| + |\mathcal{Y}|)$ equalities. Each element of $\mathcal{M}$ is called a "matching" as it defines a feasible set of matches (and singles).

Assumption L (Large market). *The number of individuals on the market* $N = \sum_{x=1}^{|\mathcal{X}|} n_x + \sum_{y=1}^{|\mathcal{Y}|} m_y$ *goes to infinity; and the ratios* (n_x/N) *and* (m_y/N) *are constant.*

Assumption S (Separability). *The surplus from a match between a man m of group x and a woman w of group y must decompose into*

$$\Phi_{mw} = \Phi_{xy} + \varepsilon_{xym} + \eta_{xyw},$$

where the ε and η can be normalized to have zero mean.

Assumption D (Distribution of Unobserved Variation in Surplus).

a) For any man m such that $x_m = x$, the ε_{xym} are drawn independently from a $(|\mathcal{Y}|+1)$-dimensional distribution P_x;

b) For any woman w such that $y_w = y$, the η_{xym} are drawn independently from an $(|\mathcal{X}|+1)$-dimensional distribution Q_y;

c) These draws are independent across men and women.

2.4 Discrete choice structure

Men's choice. Assume that, at equilibrium, men of group x get average utility w_y from their partner of type x. Then sum of the expected utilities of the men of group x is

$$G_x(w) = n_x \mathbb{E}_{\mathbf{P}_x}\left[\max_{y=0,\ldots,|\mathcal{Y}|}(w_y + \varepsilon_y)\right]$$

where the expectation is taken over a random vector of utility shocks $(\varepsilon_0, \ldots, \varepsilon_{|\mathcal{Y}|}) \sim \mathbf{P}_x$.

Define the *generalized entropy* G^* as the convex conjugate (Legendre-Fenchel transform) of G_x as

$$G_x^*(n_x, a_1, \ldots, a_{|\mathcal{Y}|}) = \max_{w=(w_0,\ldots,w_{|\mathcal{Y}|})}\left(\sum_{y=0}^{|\mathcal{Y}|} a_y w_y - G_x(w)\right)$$

where $a_0 = n_x - \sum_{y=1}^{|\mathcal{Y}|} a_y$. Note that the a_y/n_x are interpreted as the conditional probability that men of type

x choose partner y. At equilibrium, by the Envelope theorem,

$$a_y/n_x = \frac{\partial G_x(w)}{\partial w_y}.$$

As an important special case (Choo and Siow), when the utility shocks ε and η have the iid Type-I Extreme Value (Gumbel) distribution, one has the logit structure

$$\begin{aligned} G_x(w) &= n_x \log \left(\sum_{y=0}^{|\mathcal{Y}|} e^{w_y} \right) \\ G_x^* \left(n_x, a_1, \dots, a_{|\mathcal{Y}|} \right) &= \sum_{y=0}^{|\mathcal{Y}|} \frac{a_y}{n_x} \log \frac{a_y}{n_x} \end{aligned}$$

(where $a_0 = n_x - \sum_{y=1}^{|\mathcal{Y}|} a_y$). Hence G_x^* is a usual entropy function in this case.

Women's choice. Similarly, the sum of the expected utilities of the women of group y is

$$H_y(z) = m_y \mathbb{E}_{\mathbf{Q}_y} \left[\max_{x=0,\dots,|\mathcal{X}|} (z_x + \eta_x) \right],$$

and the associated generalized entropy is denoted

$$H_y^*(m_y, b_1, \ldots, b_{|\mathcal{X}|}).$$

At equilibrium, both men's and women's problems induce a relation between the structure of partner choice and the shares of the surplus. We need to impose feasibility in order to solve for the equilibrium.

2.5 Social surplus

The equilibrium matching μ is determined by the following result.

Theorem 1. *(**Social Surplus**) Under assumptions (L), (S) and (D), the market equilibrium maximizes the social gain*

$$\mathcal{W}(\mu) = \sum_{x,y \geq 1} \mu_{xy} \Phi_{xy} + \mathcal{E}(n, m, \mu),$$

over all feasible matchings $\mu \in \mathcal{M}$, where $\mathcal{E}$ is the generalized entropy given by

$$\mathcal{E}(n, m, \mu) = -\sum_{x=1}^{|\mathcal{X}|} G_x^* \left(n_x, \mu_{x.}\right) - \sum_{y=1}^{|\mathcal{Y}|} H_y^* \left(m_y, \mu_{.y}\right).$$

The first term $\sum_{xy} \mu_{xy} \Phi_{xy}$ in the social reflects "group preferences": if groups x and y generate more surplus when matched, then they should be matched with higher probability. In the one-dimensional example in Becker, an increasing x or y could reflect higher education. If the marital surplus is complementary in the educations of the two partners, Φ_{xy} is supermodular and this first term is maximized when matching partners with similar education levels (as far as feasibility constraints allow.)

On the other hand, the second term $\mathcal{E}(\mu)$ reflects the effect of the dispersion of individual affinities, conditional on observed characteristics: those men m in a group x that have more affinity to women of group y should be matched to women of group y. The formula for $\mathcal{W}(\mu)$ in Theorem 1 incorporates these two considerations. To take the education example again, a marriage between a man with a college degree and a woman who is a high-school dropout generated less marital surplus on average than a marriage between college graduates; but because of the dispersion of marital surplus that comes from the ε and η terms, it will be optimal to have some marriages between dissimilar partners.

Interpretation. Going back to the discrete choice problem with mean utilities w:

$$G_x(w) = n_x E_{\mathbf{P}_x}\left[\max_{y=0,\dots,|\mathcal{Y}|}(w_y + \varepsilon_y)\right],$$

first note that (again assuming differentiability for simplicity)

$$\frac{\partial G_x}{\partial w_y}(w) = n_x \Pr(y|x;w),$$

where $\Pr(y|x;w)$ denotes the probability that the maximum is achieved for a choice of partner in group y. As a consequence, the first-order conditions in $G_x^*(a)$ can be written

$$a_y = n_x \Pr(y|x;w),$$

so that any vector w that achieves the maximum in $G_x^*(n_x, a)$ is a vector of mean utilities that rationalizes the assignment of partner groups in a for men of group x. Moreover, denote $e(y|x;w)$ the conditional expectation of ε_y given that y was chosen by a man of group x when mean utilities are w. Then by construction

$$G_x(w) = n_x \sum_{y=0}^{|\mathcal{Y}|} \mathsf{Pr}(y|x;w)\left(w_y + e(y|x;w)\right);$$

and it follows that

$$G_x^*(n_x, a) = -n_x \sum_{y=0}^{|\mathcal{Y}|} \mathsf{Pr}(y|x;w) e(y|x;w).$$

Take $a = \mu_{x.}$, the observed assignment of groups of women to men of group x; and let $w^x(\mu_{x.})$ be a vector of mean utilities that rationalizes $\mu_{x.}$. Then

$$G_x^*\left(n_x, \mu_{x.}\right) = -\sum_{y=0}^{|\mathcal{Y}|} \mu_{xy} e\left(y|x; w^x(\mu_{x.})\right).$$

Summing up, the generalized entropy $\mathcal{E}(\mu)$ can be rewritten as

$$\sum_{x=0}^{|\mathcal{X}|} \sum_{y=0}^{|\mathcal{Y}|} \mu_{xy} e(y|x; w^x(\mu_{x.})) + \sum_{y=0}^{|\mathcal{Y}|} \sum_{x=0}^{|\mathcal{X}|} \mu_{xy} e(x|y; w^y(\mu_{.y})).$$

Hence the social gain from a matching $\mathcal{W}(\mu)$ is

$$\mathcal{W}(\mu) = \sum_{xy} \mu_{xy} \Phi_{xy} + \mathcal{E}(\mu).$$

2.6 Sharing the surplus

Theorem 1 has several important consequences. In particular, it yields a remarkably simple formula for the utilities participants of any type obtain in equilibrium. We state the result for men—the one for women follows with the obvious change in notation.

Theorem 2. *(***Participant Utilities***) Under assumptions (L), (S) and (D),*

a) In equilibrium, a man $m \in x$ who marries a woman of group y obtains utility

$$U_{xy} + \varepsilon_{xym}$$

where

$$U_{xy} = \frac{\partial G_x^*}{\partial \mu_{xy}} (n_x, \mu_{x\cdot})$$

can also be computed by solving the system of equations

$$\frac{\partial G_x}{\partial w_{xy}} (U_{x\cdot}) = \mu_{xy} \text{ for } y = 0, \ldots, |\mathcal{Y}| ,$$

given the normalization $U_{x0} = 0$.

b) The average expected utility of the men of group x is

$$u_x = \frac{G_x(U_{x.})}{n_x} = -\frac{\partial G_x^*}{\partial n_x}(n_x, \mu_{x.}). \tag{2}$$

Part b) of Theorem 2, in particular, makes it extremely easy to evaluate the participant utilities. The data directly yield the number of participants of this type (n_x) and their matching patterns ($\mu_{x.}$); and the specification of the distribution of unobserved heterogeneity determines the function G_x^*, allowing for the computation of u_x.

In practice, two cases will arise in empirical applications:

- Matching patterns (μ_{xy}) are observed but transfers τ_{xy} are not: typical in marriage market applications.

- Both matching patterns and transfers are observed: often the case in labour market applications.

Our results allow for handling both situations.

2.7 Consequences for identification

Remember that $\Phi_{xy} = U_{xy} + V_{xy}$; then Theorem 2 implies the following:

Theorem 3. *(***Identification***) Under assumptions (L), (S) and (D),*

a) In equilibrium, for any $x, y \geq 1$

$$\Phi_{xy} = -\frac{\partial \mathcal{E}(n, m, \mu)}{\partial \mu_{xy}} = \frac{\partial G_x^*}{\partial \mu_{xy}}(n_x, \mu_{x\cdot}) + \frac{\partial H_y^*}{\partial \mu_{xy}}\left(m_y, \mu_{\cdot y}\right); \tag{3}$$

b) Denote the systematic part of pre-transfer utilities (α, γ) *and of transfers* τ*. Then*

$$U_{xy} = \alpha_{xy} - \tau_{xy} \text{ and } V_{xy} = \gamma_{xy} + \tau_{xy}.$$

Hence:

- If transfers are not observed, then only the joint surplus (sum of the utilities of men and women $\alpha_{xy} + \gamma_{xy}$) are identified.

- If transfers are observed, individual surpluses of men and women α_{xy} and γ_{xy} are also identified.

2.8 Examples

We now study a couple of examples.

Example 1 (Heteroskedastic logit). *Assume that* ε_{ijm} *and* η_{ijw} *are type-I extreme value random variables with scaling factors* σ_i^m *and* σ_j^w *respectively. Then (focusing on men)*

$$G_i(w) = p_i \sigma_i^m \log \sum_{j=0}^{J} \exp\left(\frac{w_j}{\sigma_i^m}\right).$$

Take numbers of marriages $(a_1, \ldots, a_J)$ *for men of type* i*, and denote* $a_0 = p_i - \sum_{j=1}^{J} a_j$*. These marriage patterns can be rationalized by the mean utilities*

$$w_i^j(p_i, a) = \sigma_i^m \log \frac{a_j}{p_i} + t_i(a),$$

where $t_i(a)$ *is an arbitrary scalar function. As a result,*

$$G_i^*(p_i, a_1, \ldots, a_J) = \sigma_i^m \sum_{j=0}^{J} a_j \ln \frac{a_j}{p_i};$$

and

$$\mathcal{E}(p,q,\mu) = -\sum_{i=1}^{I} \sigma_i^m \sum_{j=0}^{J} \mu_{ij} \ln \frac{\mu_{ij}}{p_i} - \sum_{j=1}^{J} \sigma_j^w \sum_{i=0}^{I} \mu_{ij} \ln \frac{\mu_{ij}}{q_j}.$$

Hence (3) simplifies to

$$2\pi_{ij} = \left(\sigma_i^m + \sigma_j^w\right) \ln \mu_{ij} - \sigma_i^m \ln \mu_{i0} - \sigma_j^w \ln \mu_{0j}; \quad (4)$$

men of type i get an average expected utility

$$u_i = -\sigma_i^m \ln \frac{\mu_{i0}}{p_i},$$

and women of type j get an average expected utility

$$v_j = -\sigma_j^w \ln \frac{\mu_{0j}}{q_j}.$$

In the homoskedastic case, this simplifies to Choo and Siow's model.

Example 2 (Choo and Siow). *As a particular case of the above example when* $\sigma_i^m = \sigma_j^w = 1$*, we get*

$$\mathcal{E}(\mu) = -\sum_{i=1}^{I}\sum_{j=0}^{J} \mu_{ij} \ln \frac{\mu_{ij}}{p_i} - \sum_{j=1}^{J}\sum_{i=0}^{I} \mu_{ij} \ln \frac{\mu_{ij}}{q_j}.$$

which implies Choo and Siow's result.

As a more complex example of a GEV distribution, consider a nested logit.

Example 3 (Nested logit). *Suppose for instance that men of type i choose among "nests" A_l^i for $l = 1, \ldots, m_i$, and that the scale parameter is σ_{il}^m in nest l, and s_i^m overall. Then the system of equations that defines the U_{ij}:*

$$\frac{\partial G_i}{\partial w_{ij}}\left(U_{i.}\right) = \mu_{ij} \text{ for } j = 0, \ldots, J,$$

can be rewritten as

$$\begin{aligned} \frac{\mu_{ij}}{p_i} &= \frac{\left(\sum_{j' \in A_l^i} \exp\left(\frac{U_{ij'}}{\sigma_{il}^m}\right)\right)^{\sigma_{il}^m / s_i^m}}{\sum_{k=1}^{m_i} \left(\sum_{j' \in A_k^i} \exp\left(\frac{U_{ij'}}{\sigma_{ik}^m}\right)\right)^{\sigma_{ik}^m / s_i^m}} \qquad (5) \\ &\quad \times \frac{\exp\left(U_{ij}/\sigma_{il}^m\right)}{\sum_{j' \in A_l^i} \exp\left(U_{ij'}/\sigma_{il}^m\right)} \end{aligned}$$

where l is the index of the nest such that $j \in A_l^i$. There is no general closed-form expression for U_{ij}; however, note *that within a nest A_l^i,*

$$U_{ij} = \sigma_{il}^{m} \log \frac{\mu_{ij}}{p_i} + t_l^i$$

and that in (5) only the constants t_l^i remain to be determined numerically.

While the GEV framework is convenient, the mixed logit model has also become quite popular in the applied literature; it is our last example.

Example 4 (Mixed logit). *Take nonnegative numbers* α_{ik} *such that* $\sum_{k=1}^{K} \alpha_{ik} = 1$ *for each* i.. *Consider the mixture model in which for any type* i *of men, with probability* α_{ik} *the distribution* $\mathbf{P}_i$ *is iid type-I extreme value with standard error* σ_{ik}^m.

Then the U_{ij} *solve*

$$\frac{\mu_{ij}}{p_i} = \sum_{k=1}^{K} \alpha_{ik} \frac{e^{U_{ij}/\sigma_{ik}^m}}{\sum_{j'=0}^{J} e^{U_{ij'}/\sigma_{ik}^m}}.$$

In the previous examples, the generalized entropies G_x^* and H_y^*, and hence $\mathcal{E}$ can be found in closed form. These are particular instances of McFadden's Generalized Extreme Value (GEV) framework (see McFadden 1978). Consider functions $g_x : \mathbb{R}^{|\mathcal{Y}|+1} \longrightarrow \mathbb{R}$ and $h_y : \mathbb{R}^{|\mathcal{X}|+1} \longrightarrow \mathbb{R}$ such that the following four conditions hold:

- each g_x or h_y is positive homogeneous of degree one
- they go to $+\infty$ whenever any of their arguments goes to $+\infty$
- their partial derivatives of order k exist outside of 0 and have sign $(-1)^k$
- the functions defined by

$$\begin{aligned} \mathbf{P}_x\left(w_0, ..., w_{|\mathcal{Y}|}\right) &= \exp\left(-g_x\left(e^{-w_0}, ..., e^{-w_{|\mathcal{Y}|}}\right)\right) \\ \mathbf{Q}_y\left(z_0, ..., z_{|\mathcal{X}|}\right) &= \exp\left(-h_y\left(e^{-z_0}, ..., e^{-z_{|\mathcal{X}|}}\right)\right) \end{aligned}$$

  are multivariate cumulative distribution functions.

Then introducing utility shocks $\varepsilon_x \sim \mathbf{P}_x$, and $\eta_y \sim \mathbf{Q}_y$, we have by a theorem of McFadden (1978):

$$\begin{aligned}\frac{G_x(w)}{n_x} &= \mathbb{E}_{\mathbf{P}_x}\left[\max_{y=0,1,\ldots,|\mathcal{Y}|}\{w_y+\varepsilon_y\}\right] = \log g_x\left(e^{w}\right)+\gamma \\ \frac{H_y(z)}{m_y} &= \mathbb{E}_{\mathbf{Q}_y}\left[\max_{x=0,1,\ldots,|\mathcal{X}|}\{z_x+\eta_x\}\right] = \log h_y\left(e^{z}\right)+\gamma\end{aligned}$$

where γ is the Euler constant $\gamma \simeq 0.5772$.

Therefore,

$$\begin{aligned}G_x^*\left(n_x,a\right) &= \left(n_x-\sum_{y=1}^{|\mathcal{Y}|}a_y\right)w_0^x(n_x,a)+\sum_{y=1}^{|\mathcal{Y}|}a_y w_y^x\left(n_x,a\right) \\ &\quad -n_x\left(\log g_x\left(e^{w^x(n_x,a)}\right)+\gamma\right)\end{aligned}$$

where for $x=0,\ldots,|\mathcal{X}|$, the vector $w^x\left(n_x,a\right)$ solves the system

$$\left(n_x-\sum_{y=1}^{|\mathcal{Y}|}a_y, a_1,\ldots,a_{|\mathcal{Y}|}\right)=n_x\frac{\partial}{\partial w}\log g_x\left(e^{w^x}\right). \tag{6}$$

Similarly, if $\sum_{x=0}^{|\mathcal{X}|} b_x = m_y$ then

$$H_y^*(b) = \sum_{x=0}^{|\mathcal{X}|} b_x z_x^y(m_y, b) - m_y\left(\log h_y\left(e^{z^y(m_y,b)}\right) + \gamma\right) \tag{7}$$

where the vectors $z^y(m_y, b)$ solve the systems

$$\left(m_y - \sum_{x=1}^{|\mathcal{X}|} b_x, b_1, \ldots, b_{|\mathcal{X}|}\right) = m_y \frac{\partial}{\partial z} \log h_y\left(e^{z^y}\right).$$

Hence,

$$\begin{aligned} & \mathcal{E}(n, m, \mu) \\ = & \sum_{x=1}^{|\mathcal{X}|}\left(n_x \log g_x\left(e^{w^x(n_x,\mu_{x\cdot})}\right) - \sum_{y=0}^{|\mathcal{Y}|} \mu_{xy} w_y^x(n_x, \mu_{x\cdot})\right) \\ + & \sum_{y=1}^{|\mathcal{Y}|}\left(m_y \log h_y\left(e^{z^y(m_y,\mu_{\cdot y})}\right) - \sum_{x=0}^{|\mathcal{X}|} \mu_{xy} z_x^y\left(m_y, \mu_{\cdot y}\right)\right) \\ & + C \end{aligned}$$

where $C = \gamma\left(\sum_{x=1}^{|\mathcal{X}|} n_x + \sum_{y=1}^{|\mathcal{Y}|} m_y\right)$, and for $x, y \geq 1$

$$\frac{\partial \mathcal{E}}{\partial \mu_{xy}}(n, m, \mu) = -w_y^x\left(n_x, \mu_{x\cdot}\right) - z_x^y\left(m_y, \mu_{\cdot y}\right).$$

3 Parametric estimation

Linear expansion of Φ_{xy} as $\Phi^{\lambda}_{xy} = \sum_{k=1}^{K} \lambda_k \phi^k_{xy}$, where ϕ^k_{xy} are known basis functions.

Leading example: x is a vector of characteristics $x \in \mathbb{R}^d$ and also $y \in \mathbb{R}^d$. The interaction surplus is quadratic, i.e.

$$\phi^{ml}_{xy} = x^m y^l$$

and then the surplus function is given by

$$\Phi_A(x, y) = \sum_{m,l} A_{ml} x^m y^l$$

where (A_{ml}) is the *affinity matrix*. (The estimated λ_k should be identified with term A_{ml} of matrix A, and $\phi^k(x, y) = x^m y^l$).

This structural parameter has a matching interpretation as the vector of weight of interactions of the various components:

$$\frac{\partial \mathcal{W}}{\partial A_{ml}}\left(A^{XY}\right) = \Sigma^{ml}_{XY}$$

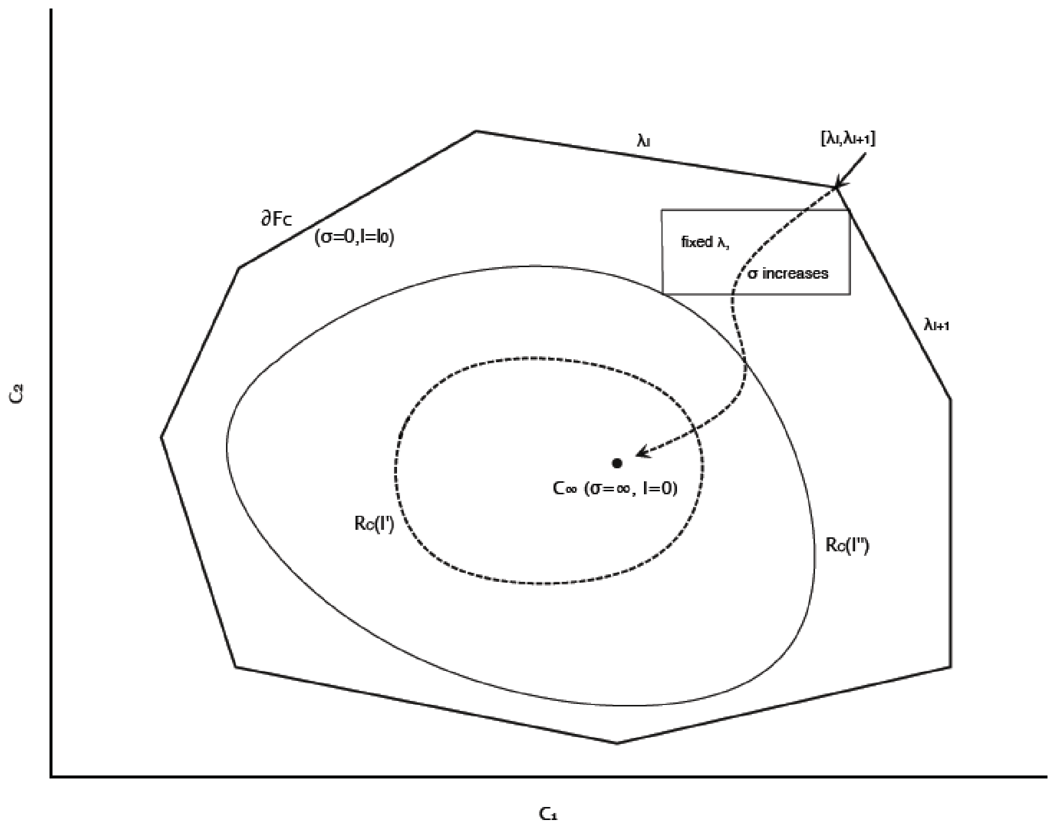
where $\mathbf{\Sigma}_{XY}^{ml}$ is the covariance between X^m and Y^l.

To simplify the exposition, we assume no singles, and we normalize the number of men and women to one. Thus μ_{xy} can be interpreted a probability of drawing a (x, y) pair from the sample. This probability has margins $n(x)$ (probability of drawing a man of type x) and $m(y)$ (probability of drawing a woman of type y). We denote $\mu \in \mathcal{M}(n, m)$.

Our goal is estimating λ. At equilibrium, we have

$$\mathcal{W}(\lambda) = \sup_{\mu \in \mathcal{M}(n,m)} \mathbb{E}_{\mu}[\Phi_{\lambda}(X, Y)] + \mathcal{E}(\pi)$$

which generates a likelihood μ^{λ}.



λl
[λl,λl+1]
∂Fc
(σ=0,l=l0)
fixed λ,
σ increases
λl+1
C2
C∞ (σ=∞, l=0)
Rc(l')
Rc(l")
C1

MLE has the following property it is the unique $\hat{\lambda}$ such that $\mu^{\hat{\lambda}}$ generates moments $\mathbb{E}_{\mu^{\lambda}}\left[\Phi^k(X,Y)\right]$ that coincide with the empirical moments:

$$\hat{C}_n^k = \mathbb{E}_n\left[\Phi^k(X,Y)\right] = \frac{1}{n}\sum_{t=1}^{n}\Phi^k(X_t,Y_t),$$

$\hat{\lambda}$ is obtained by solving the optimization program:

$$\min_{\lambda\in\mathbb{R}^k}\left(\mathcal{W}(\lambda) - \sum_{k=1}^{K}\lambda_k\hat{C}_n^k\right).$$

Estimation of $\hat{\lambda}$ requires the computation of $\mathcal{W}(\lambda)$. This is solved using a generalization of the RAS algorithm.

4 Computation of $\mathcal{W}$

We use an Iterative Projection Fitting Procedure (IPFP), a.k.a. RAS algorithm, to determine u, v. The algorithm iterates over values (u_k, v_k) so that

$$\pi_k(x,y) = p(x)\,q(y)\exp\left(\frac{\Phi(x,y) - u_k(x) - v_k(y)}{\sigma}\right)$$

converges to the solution.

Start from any initial guess of u_0 and v_0.

At step $(k+1)$ we adjust iteratively

- $v_k(y)$ into $v_{k+1}(y)$ so to fit the y-margin, that is, set v_{k+1} such that

$$e^{v_{k+1}(y)/\sigma} = \sum_x q(y)\exp\left(\frac{\Phi(x,y) - u_k(x)}{\sigma}\right)$$

- and then $u_k(x)$ into $u_{k+1}(x)$ so to fit the x-margin, that is, set u_{k+1} such that

$$e^{u_{k+1}(x)/\sigma} = \sum_y p(x) \exp\left(\frac{\Phi(x,y) - v_{k+1}(y)}{\sigma}\right)$$

Stop when sufficiently close to a fixed point. The algorithm converges to the right functions u and v.

IPFP algorithm converges *much* faster for $\sigma > 0$ than classical assignment algorithms for $\sigma = 0$. For a population of a couple of thousands, convergence in half a second vs. a few minutes.

5 Saliency analysis

5.1 Methodology

Our goal in this section is to build one (or several) index (indices) of attractiveness.

Classical example: Linear Canonical Correlation method (suggested by Becker, 1974). Advantage of the method: immediate availability in most statistical packages, simplicity of use and empirical intuition.

Drawback: purely descriptive, not rooted in a structural equilibrium model, hence it is not informative about the agents' preferences, and no clear interpretation in a matching context.

We propose an alternative approach, though related, we call saliency analysis, that is grounded in an equilibrium model.

Instead of performing a Singular Value Decomposition of the (renormalized) cross-covariance matrix Σ, we perform a Singular Value Decomposition of the renormalized affinity matrix, which is the structural matching parameter estimated in the GS framework.

As before, we assume a quadratic parametrization of Φ: for A an $n_x \times n_y$ matrix, we take

$$\Phi_A(x, y) = x' A y.$$

Let A^{XY} be parameter estimated and call it *affinity matrix* between vectors X and Y.

In the sequel we are going to normalize X and Y into

$$\begin{aligned}\bar{X} &= \Sigma_X^{-1/2} X \\ \bar{Y} &= \Sigma_Y^{-1/2} Y\end{aligned}$$

so that $\bar{X}, \bar{Y} \sim N(0, I_d)$. Note that $A^{\overline{XY}} = \Sigma_X^{1/2} A^{XY} \Sigma_Y^{1/2}$.

Singular Value Decomposition (SVD) of affinity matrix $A^{\overline{XY}}$ yields:

$$A^{\overline{XY}} = U'\Lambda V$$

where Λ is diagonal matrix with nonnegative, nonincreasing elements $(\lambda_1, ..., \lambda_d)$, and U' and V are orthogonal matrices.

Define $\tilde{X} = U\overline{X}$ and $\tilde{Y} = V\overline{Y}$ as salient vectors of attractiveness. One has:

The affinity matrix on $\tilde{X}$ and $\tilde{Y}$ is $A^{\tilde{X}\tilde{Y}} = \Lambda$.

Interpretation: $\tilde{X}^1$ and $\tilde{Y}^1$ are the best explaination of the matching structure using a one-dimension model; $\left(\tilde{X}^1, \tilde{X}^2\right)$ and $\left(\tilde{Y}^1, \tilde{Y}^2\right)$ are the best explanation of the matching structure using a two-dimension model; etc.

5.2 Application

The data source is DNB Household Survey (DHS), Waves 1993-2002. Representative panel of the Dutch population (region, political preference, housing, income, degree of urbanization, and age of the head of the household). 2000 households in each wave. Within each household, all persons aged 16 or over were interviewed.

Data particularity:

1. detailed information about all individuals in the household: allows us to reconstruct "couples".

2. rich information set: socio-demographic variables (birth year and education), morphology (height and weight), self-assessed health and information about personality traits.

3. We make use of the panel structure to deal (partly) with nonresponses on socioeconomic and health variables. When missing values for education, height, weight, education, year of birth etc. were encountered, values reported in adjacent years were imputed.

Measuring educational attainment. Respondent's reported highest level of education achieved.

1. Lower education: lower vocational training, kindergarten/primary education, continued primary education or elementary secondary education,

2. Intermediate education: secondary education, junior vocational training

3. Higher education: University: university education.

Measuring morphology and health.

- Height and weight: Body Mass Index of each respondent as the weight in Kg divided by the square of the height measured in meters.

- The respondents were also asked to report their general health. "How do your rate your general health condition on a scale from 1, excellent, to 5, poor?".

Measuring personality traits. Using multiple waves to construct the full 16PA scale of Brandstätter (1988), Nyhus and Webley (2001) showed that this scale distinguishes 5 factors. They labelled these factors as:

1. Emotional stability: a high score = less likely to interpret ordinary situations as threatening, and minor frustrations as hopelessly difficult,

2. Extraversion (outgoing): a high score = more likely to need attention and social interaction,

3. Conscientiousness (meticulous): a high score = more likely to be meticulous,

4. Agreeableness (flexibility): a high score = more likely to be pleasant with others and go out of their way to help others and,

5. Autonomy (tough-mindedness): a high score = more likely to direct, rough and dominant.

Measuring risk preference.

- Attitude toward risk: using a list of 6 items of the type "I am prepared to take the risk to lose money, when there is also a chance to gain money...... from a scale from 1, totally disagree, to 7, totally agree".

- Collapse the data by individuals using the person's median answer to each item.

- We then construct an index of risk aversion by adding the answers to the respective items.

Construction of working dataset.

- Pool all the waves selected (1993-2002).

- Keep only head of the household, spouse of the head or a permanent partner of the head: sample of roughly 13,000 men and women and identifies about 7,700 unique households.

- Create women and men datasets: each data set identifies about 6,500 different men and women.

- Create working dataset: merging men dataset to women dataset using Household id. 5,445 unique couples identified (roughly 1,250 unmatched men and women).

6 Tables

Table 1: Number of identified couples and number of couples with complete information for various subset of variables.

	N
Identified couples	5,445
Couples with complete information on:	
Education	5,409
The above + Health, Height and BMI[a]	3,214
The above + Personality traits (Big 5)	2,573
The above + measure of risk aversion	2,378

Notes: The selected sample for our analysis is the one from the last row.

a: Excluding health produces exactly the same number of couples at this stage.

Source: DNB. Own calculation.

Table 2: Sample of couples with complete information: summary statistics by gender.

	Husbands			Wives		
	N	mean	S.E.	N	mean	S.E.
Educational level	2378	2.0	0.6	2378	1.8	0.6
Height	2378	180.8	7.2	2378	168.4	6.5
BMI	2378	24.8	2.9	2378	23.9	4.1
Health	2378	4.1	0.7	2378	4.0	0.7
Conscientiousness	2378	-0.1	0.7	2378	0.1	0.7
Extraversion	2378	-0.1	0.7	2378	0.2	0.6
Agreeableness	2378	-0.1	0.6	2378	-0.1	0.6
Emotional stability	2378	0.1	0.6	2378	-0.2	0.5
Autonomy	2378	-0.0	0.7	2378	-0.0	0.7
Risk aversion	2378	0.1	0.7	2378	-0.2	1.0

Table 3: Estimates of the Affinity matrix: quadratic specification ($N = 2378$).

Wives Husbands	Education	Height.	BMI	Health	Consc.	Extra.	Agree.	Emotio.	Auto.	Risk
Education	**0.46**	0.00	**-0.06**	0.01	-0.02	0.03	-0.01	-0.03	0.04	0.01
Height	0.04	**0.21**	0.04	0.03	**-0.06**	0.03	0.02	0.00	-0.01	0.02
BMI	-0.03	0.03	**0.21**	0.01	0.03	0.00	**-0.05**	0.02	0.01	-0.02
Health	-0.02	0.02	- 0.04	**0.17**	- 0.04	0.02	- 0.01	0.01	-0.00	0.03
Conscienciousness	**-0.07**	-0.01	**0.07**	-0.00	**0.16**	0.05	0.04	0.06	0.01	0.01
Extraversion	0.00	-0.01	0.00	0.01	**-0.06**	**0.08**	-0.04	-0.01	0.02	**-0.06**
Agreeableness	0.01	0.01	-0.06	0.02	**0.10**	**-0.11**	0.00	0.07	-0.07	-0.05
Emotional	0.03	-0.01	0.04	**0.06**	**0.19**	0.04	0.01	-0.04	**0.08**	**0.05**
Autonomy	0.03	0.02	0.01	0.02	**-0.09**	**0.09**	-0.04	0.02	**-0.10**	0.03
Risk	0.03	-0.01	-0.03	-0.01	0.00	-0.02	-0.03	-0.03	**0.08**	**0.14**

Note: Bold coefficients are significant at the 5 percent level.

Table 4: Share of observed surplus explained.

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
Share of surplus explained	25.8***	18.5***	12.4***	11.0***	9.5***	7.6***	6.7***	4.8***	2.4	1.4
Standard deviation of Shares	(1.7)	(1.2)	(1.1)	(1.1)	(1.1)	(1.2)	(1.0)	(1.4)	(1.4)	(1.4)

*** significant at 1 percent

Table 5: Indices of attractiveness.

	I1		I2		I3	
Attributes	M	W	M	W	M	W
Education	**0.91**	**0.93**	0.15	0.13	-0.34	-0.32
Height	0.15	0.08	-0.13	-0.08	**0.58**	**0.60**
BMI	-0.24	-0.31	0.08	0.06	-0.15	-0.19
Health	0.12	0.13	-0.01	0.14	**0.64**	**0.64**
Conscientiousness	-0.23	-0.11	**0.58**	**0.90**	0.03	0.07
Extraversion	-0.00	0.02	-0.27	-0.06	-0.03	0.18
Agreeableness	0.08	-0.02	0.39	0.26	0.22	0.06
Emotional	0.05	-0.03	**0.63**	0.23	0.24	0.16
Autonomy	0.07	0.08	-0.17	0.09	0.21	-0.12
Risk	0.20	0.14	0.04	0.19	0.00	0.24
Cum. share	0.258		0.443		0.567	

*** significant at the 1 percent

Note: M means Men and W means women. Bold coefficients indicates coefficients larger than 0.5.

References

[CIH] Galichon, Salanié. Cupid's invisible hand: Social surplus and identification in matching models.

[PT] Dupuy, Galichon. Personnality traits and the marriage market.