Personality traits and the marriage market

Alfred Galichon
Sciences Po, Paris

Joint work with Arnaud Dupuy
(CEPS/INSTEAD and Maastricht University)

University of Chicago, May 10, 2013

1 Introduction

- Goal: study the **determinants of the patterns in the marriage market**.
  Endogamy (Becker 1991, Browning, Chiappori, and Weiss, 2007):
  - *Race*; Fryer, 2007,
  - *Education, income*; Wong, 2003, Anderberg, 2004, Chiappori, Salanié and Weiss, 2010, Bruze 2011,
  - *Height, weight, BMI, look*; Hitsch, Hortaçsu and Ariely, 2010, Oreffice and Quintana-Domeque, 2010, Chiappori, Oreffice, Quintana-Domeque 2012a,
  - *Parental wealth*; Charles, Hurst and Killewald, 2011.

- Central questions: **which** and **how many** attributes are relevant for sorting? In terms of personality traits, do "**Opposites Attract**" or "**Like Attracts Like**"?
  This paper: psychological characteristics of marital partners.

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

- Univariate continuous: Becker, 1973, Chiappori, Oreffice, Quintana-Domeque 2012a, 2012b. Yet these yield (too) stark predictions: "Positive assortative matching" on a single-dimensional "ability index".

- Discrete: Wong, 2003, Anderberg, 2004, Choo and Siow, 2006, Fox, 2010, Galichon and Salanié, 2012.

- Continuous multivariate: this paper.

Our strategy:

1. Build a **structural model of multivariate matching** and **estimate surplus function**,

2. Construct **indices of attractivness** from this estimator,

3. Test for the **number of significant indices**.

2 Equilibrium on the marriage market

2.1 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. Consider a population with n_x men of type x, and m_y women of type y, such that $\sum_x n_x = \sum_y m_y$. Introduce:

- $\alpha_{xy} > 0$ utility of man x with woman y, 0 if single
- $\gamma_{xy} > 0$ utility of woman y with man x, 0 if single
- $\Phi_{xy} = \alpha_{xy} + \gamma_{xy}$ the total gains to marriage.

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

Introduce μ_{xy} the number of (x, y) pairs. Then (**Becker-Shapley-Shubik**): the market clears in order to maximize utilitarian social surplus

$$\max_{\mu \geq 0} \sum_{x,y} \mu_{xy} \Phi_{xy}$$

$$s.t. \qquad \sum_{y} \mu_{xy} = n_x \ \sum_{x} \mu_{xy} = m_y.$$

and the equilibrium payoffs u_x and v_y are determined by the dual of the former program

$$\begin{aligned} \min \qquad & \sum_{x} n_x u_x + \sum_{y} m_y u_y. \\ s.t. \qquad & u_x + v_y \geq \Phi_{xy} \end{aligned}$$

Gretsky, Ostroy and Zame (1992): extension to continuous characteristics.

2.2 Assortativeness

When partners' characteristics x and y are **continuous** and **scalar**, and when $\Phi(x, y)$ is **supermodular**, i.e.

$$\frac{\partial^2 \Phi}{\partial x \partial y} \geq 0$$

for which a leading example is

$$\Phi(x, y) = xy.$$

Then the optimal matching is such that

$$Y = T(X)$$

for some nondecreasing deterministic map $T(.)$: **positive assortative matching (PAM)**.

Sorting on an index. Positive assortative matching is observed, with varying intensity, on many dimensions of vectors of observed attributes x and y. It is sometimes postulated that there are "indices of ability" $\bar{x}$ and $\bar{y}$

$$\bar{x} = \sum_{i=1}^{d_x} \lambda_i x_i \text{ and } \bar{y} = \sum_{j=1}^{d_y} \nu_j y_j$$

on which positive assortative matching occurs

$$\Phi\left(\bar{x}, \bar{y}\right) = \bar{x}\bar{y} = \sum_{i=1}^{d_x} \sum_{j=1}^{d_y} \lambda_i \nu_j x_i y_j.$$

Then, concerns are:

- how to determine the index weights (λ_i) and $\left(\nu_j\right)$

- does a pair of single indices suffice for explaining the sorting?

2.3 On the importance of being structural

Canonical correlation is often proposed, and used, to determine the index weights (λ_i) and $\left(\nu_j\right)$. Principle: find (λ_i) and $\left(\nu_j\right)$ so as to maximize

$$\max_{\lambda,\nu} cov(\sum_{i=1}^{d_x} \lambda_i X_i, \sum_{j=1}^{d_y} \nu_j Y_j)$$

subject to normalization

$$var(\sum_{i=1}^{d_x} \lambda_i X_i) = var(\sum_{j=1}^{d_y} \nu_j Y_j) = 1.$$

Canonical correlation amounts to performing a **singular value decomposition** (SVD) of

$$\Sigma_X^{-1/2} \Sigma_{XY} \Sigma_Y^{-1/2} = U' D V$$

where
$\Sigma_{XY} = \mathbb{E}_\mu \left[XY'\right]$, $\Sigma_X = \mathbb{E}_\mu \left[XX'\right]$ and $\Sigma_Y = \mathbb{E}_\mu \left[YY'\right]$, and U and V are orthogonal, and D is diagonal.

This method is consistent in the Gaussian case, but **canonical correlation is inconsistent in general**. It is easy to build a counterexample where the "true" indices exhibit perfect PAM, yet canonical correlation does not recover the true indices.

This fact highlights the need for a structural model. **Chiappori, Oreffice, and Quintana-Domeque** (JPE 2012) solve the problem of consistent estimation of the weights. However, their model remains single-dimensional.

3 A continuous model with heterogeneities

The model of **Choo and Siow** (JPE, 2006) has become the "industry standard" of structural estimation of matching surplus. However, this model faces two limitations:

- it is based on a discrete choice (logit) setting, and only applies to the case where the partners' characteristics are **discrete**, and

- it is **nonparametric** in nature.

We now present a continuous version of the Choo-Siow model. Next, we shall introduce a convenient parameterization.

Attributes: $x \in \mathcal{X} = \mathbb{R}^{d_x}$ and $y \in \mathcal{Y} = \mathbb{R}^{d_y}$, with probability distribution $X \sim P$ and $Y \sim Q$, assumed centered $\mathbb{E}_P[X] = \mathbb{E}_Q[Y] = 0$, with finite second moments and absolutely continuous with respect to the Lebesgue measure.

The set of matchings $\mu \in \mathcal{M}(P, Q)$ is set of distributions μ of (X, Y) with $X \sim P$ and $Y \sim Q$.

Without heterogeneitites, surplus of pair (x, y) is

$$\Phi(x, y)$$

and the optimal matching without heterogeneities is

$$\max_{\mu \in \mathcal{M}(P,Q)} \mathbb{E}_\mu [\Phi(X, Y)]. \tag{1}$$

Now introduce unobserved heterogeneities in joint surplus as

$$\Phi(x, y) + \sigma_1 \varepsilon_m(y) + \sigma_2 \eta_w(x). \tag{2}$$

Discrete case. Choo and Siow (2006) assume $(\varepsilon_m(y))_y$ and $(\eta_w(x))_x$ are i.i.d. Gumbel (extreme-value type I), extended by Galichon and Salanié (2012, hereafter GS) for general distributions. Idea:

$$\Phi(x,y) = U(x,y) + V(x,y)$$

and man m and woman w solve discrete choice problems

$$\max_y (U(x_m, y) + \sigma_1 \varepsilon_m(y)) \text{ and } \max_x (V(x, y_w) + \sigma_2 \eta_w(x)).$$

Problem with continuous extension: how to make sense of a continuously infinite number of independent utility shocks?

Continuous case. Here, we use a continuous version of the logit model in order to extend Choo and Siow and GS to continuous case. See McFadden (1976), Ben-Akiva and Watanatada (1981) and Ben-Akiva, Litinas and Tsunekawa (1985), and especially Cosslett (1988) and Dagsvik (1988). The surplus will still be given by expression (S) above, but $\left(\varepsilon_m\left(y\right)\right)_y$ and $\left(\eta_w\left(x\right)\right)_x$ will be modelled as Extreme-value stochastic processes, in a sense we now explain.

Here, we follow the insights of Cosslett (1988) and Dagsvik (1988). Two steps:

- In the first step, man m draws an (exogenous) **network of "acquaintances"** y_k^m, along with a "sympathy shock" ε_k^m, where $\left(y_k^m, \varepsilon_k^m\right)$ is the enumeration of a Poisson point process on $\mathcal{Y} \times \mathbb{R}$ of intensity $dy \times e^{-\varepsilon} d\varepsilon$. Note that:
  - the set of acquaintances drawn by each man is infinite but discrete, thus $k \in \mathbb{Z}^+$
  - ε_k^m are distributed as i.i.d. Gumbel variables

- In the second step, man m chooses his **preferred partner** from his network of acquaintances, i.e. solves

$$\max_k \left\{U\left(x, y_k^m\right) + \sigma_1 \varepsilon_k^m\right\}.$$

The expected indirect utility of man m of type x is

$$\begin{aligned} G_x(U) &= \mathbb{E}\left[\max_k \{U(x, y_k^m) + \sigma_1 \varepsilon_k^m\}\right] \\ &= \sigma_1 \log(\int_{\mathcal{Y}} \exp(U(x,y)/\sigma_1) dy) \end{aligned}$$

which implies that the density of probability of this man choosing a woman of type y is

$$d\mu(y|x) = \frac{\exp(U(x,y)/\sigma_1)dy}{\int_{\mathcal{Y}} \exp(U(x,y')/\sigma_1)dy'}.$$

Similar formulas hold for the choice probabilities of the women, in particular

$$H_y(V) = \sigma_2 \log(\int_{\mathcal{X}} \exp(V(x,y)/\sigma_2)dx)$$

and

$$\mu(x|y) = \frac{\exp(V(x,y)/\sigma_2)\,dx}{\int_{\mathcal{X}} \exp(V(x',y)/\sigma_2)\,dx'}.$$

The social welfare is given by the following variational problem

$$\begin{aligned}\mathcal{W}(\Phi) = \quad & \min_{U,V} \int_{\mathcal{X}} G_x(U)\, dP(x) + \int_{\mathcal{Y}} H_y(V)\, dQ(y) \\ s.t. \quad & U(x,y) + V(x,y) \geq \Phi(x,y)\,.\end{aligned}$$

In the present setting, $\mathcal{W}(\Phi)$ is given by minimizing

$$\begin{aligned} & \sigma_1 \int_{\mathcal{X}} \log\Big(\int_{\mathcal{Y}} \exp(U(x,y)/\sigma_1) dy\Big) dP(x) \\ + \quad & \sigma_2 \int_{\mathcal{X}} \log\Big(\int_{\mathcal{X}} \exp(V(x,y)/\sigma_2) dQ(y)\Big)\Big) \end{aligned}$$

subject to

$$U(x,y) + V(x,y) \geq \Phi(x,y)\,.$$

This problem has a straightforward interpretation by duality.

Let $\sigma = \sigma_1 + \sigma_2$.

Theorem. Let $\mu \in \mathcal{M}(P, Q)$ be the distribution of partners attributes in the stable matching.

(i) μ has a density with respect to the Lebesgue measure on $\mathcal{X} \times \mathcal{Y}$, also denoted μ.

(ii) $\mu \in \mathcal{M}(P, Q)$ is solution to

$$\mathcal{W}(\Phi) = \max_{\mu \in \mathcal{M}(P,Q)} \mathbb{E}_\mu [\Phi(X, Y)] - \sigma \mathbb{E}_\mu [\ln \mu(X, Y)] . \tag{3}$$

(iii) μ is a solution of

$$\log \mu(x, y) = \frac{\Phi(x, y) + a(x) + b(y)}{\sigma}$$

where $a(x)$ and $b(y)$ are such that $\mu \in \mathcal{M}(P, Q)$.

Two remarks:

- Note that $a(x)$ and $b(y)$ are determined by

$$\begin{aligned} \exp\left(a\left(x\right)/\sigma\right)\int_{\mathcal{Y}}\exp\left(\Phi/\sigma\right)\exp\left(b\left(y\right)/\sigma\right)dy &= p\left(x\right) \\ \exp\left(b\left(y\right)/\sigma\right)\int_{\mathcal{X}}\exp\left(\Phi/\sigma\right)\exp\left(a\left(x\right)/\sigma\right)dx &= q\left(y\right) \end{aligned}$$

  and hence can be obtained by an iterative fitting procedure.

- The model is scale invariant, so w.l.o.g. one may set $\sigma = 1$ and $\Phi = \Phi/\sigma$.

4 Parametric estimation

We introduce the following quadratic parameterization of surplus:

$$\Phi\left(x, y\right) = \Phi_A\left(x, y\right) := x^{'} A y = \sum_{ij} A_{ij} x_i y_j$$

so that A_{ij} (the "Affinity Matrix") reflects complementarity (or substitutability) of characteritsics x_i and y_j. This model is arguably the simplest model of interactions between any pair of attributes.

Note that when there is PAM on a pair of ability indices $\bar{x} = \sum_i \lambda_i x_i$ and $\bar{y} = \sum_j \nu_j y_j$, this corresponds to the case where

$$A_{ij} = \lambda_i \nu_j$$

hence this is the case where A is of rank one.

Questions:

- Estimation of A?
- Rank of A? that is: how many dimensions matter for sorting?
- Can we construct indices of mutual attractiveness out of Φ_A? that is: which dimensions matter?

To estimate A, GS have proposed the following estimator, which naturally extends to the present setting. Let $\mathcal{W}(A) := \mathcal{W}(\Phi_A)$.

- By the envelope theorem

$$(\Sigma_{XY})_{ij} = \frac{\partial \mathcal{W}(A)}{\partial A_{ij}}, \tag{4}$$

  where $(\Sigma_{XY})_{ij} = \mathbb{E}_\mu \left[X_i Y_j \right]$ is the *predicted* cross-covariance matrix.

- This leads to estimator of A given by

$$\min_{A \in \mathcal{M}_{d_x d_y}(\mathbb{R})} \left\{ \mathcal{W}(A) - Tr\left(A' \hat{\Sigma}_{XY} \right) \right\} \tag{5}$$

  where $(\hat{\Sigma}_{XY})_{ij} = \mathbb{E}_{\hat{\mu}} \left[X_i Y_j \right]$ is the *observed* cross-covariance matrix.

We call A^{XY} the resulting estimator. It equates predicted and observed cross-covariance matrix.

(5) is a convex optimization problem. Need to compute $\mathcal{W}(A)$: easy given remark made above.

Asymptotics. Assume that a finite sample of size n is observed. The notation $\hat{M}$ means the finite-sample estimator of M, and

$$\delta M = \hat{M} - M. \tag{6}$$

The following result holds:

Theorem. The following convergence holds in distribution for $n \longrightarrow +\infty$:

$$n^{1/2}\left(\delta A, \delta\mathbf{\Sigma}_X, \delta\mathbf{\Sigma}_Y\right) \Longrightarrow \mathcal{N}\left(0, \begin{pmatrix} \mathbb{F}^{-1} & 0 & 0 \\ 0 & \mathbb{K}_{XX} & \mathbb{K}_{XY} \\ 0 & \mathbb{K}'_{XY} & \mathbb{K}_{YY} \end{pmatrix}\right)$$

where $\mathbb{F}$ is the Fisher Information matrix

$$\mathbb{F}^{ij}_{kl} = \mathbb{E}_\mu\left[\frac{\partial \log \mu\left(X,Y\right)}{\partial A_{ij}} \frac{\partial \log \mu\left(X,Y\right)}{\partial A_{kl}}\right],$$

and $\mathbb{K}$ is the variance-covariance matrix of the quadratic moments of (X,Y)

$$\begin{aligned} \left(\mathbb{K}_{XX}\right)^{kl}_{ij} &= cov_\mu\left(X^i X^j, X^k X^l\right), \\ \left(\mathbb{K}_{YY}\right)^{kl}_{ij} &= cov_\mu\left(Y^i Y^j, Y^k Y^l\right), \text{ and} \\ \left(\mathbb{K}_{XY}\right)^{kl}_{ij} &= cov_\mu\left(X^i Y^j, X^k Y^l\right). \end{aligned}$$

5 Saliency analysis

Let A^{XY} be the affinity matrix estimated previously. We would like to test the rank of A^{XY}, and understand which combinations of the x and the y's matter for sorting.

For this purpose, we introduce *saliency analysis*. The idea is to rewrite the surplus function

$$\Phi(x, y) = \sum_{ij} A_{ij} x_i y_j$$

as

$$\Phi(x, y) = \sum_{k=1}^{d} \lambda_k \tilde{x}_k \tilde{y}_k$$

where the $\tilde{x}_k$ and the $\tilde{y}_k$'s are indices obtained linearly from x and y called "indices of mutual attractivness". The best approximation of the market by a one dimensional model is given by taking ability indices as $\tilde{x}_1$ and $\tilde{y}_1$; the best approximation by a two-dimensional model is given by taking ability indices as $(\tilde{x}_1, \tilde{x}_2)$ and $(\tilde{y}_1, \tilde{y}_2)$; etc.

Canonical correlation analysis is based on the Singular Value Decomposition (SVD) of

$$\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1/2}$$

which is the cross-covariance matrix of dimensionless vectors $\Sigma_X^{-1/2}X$ and $\Sigma_Y^{-1/2}Y$.

In contrast, we suggest to base saliency analysis on the SVD of the affinity matrix of $\Sigma_X^{-1/2}X$ and $\Sigma_Y^{-1/2}Y$. That is

$$\Theta := \Sigma_X^{1/2}A^{XY}\Sigma_Y^{1/2} = U'\Lambda V,$$

where

- Λ is diagonal with nonincreasing elements $(\lambda_1, ..., \lambda_d)$, $d = \min(d_x, d_y)$, and

- U and V are orthogonal matrices.

Define $\tilde{X} = U\mathbf{\Sigma}_X^{-1/2} X$ and $\tilde{Y} = V\mathbf{\Sigma}_Y^{-1/2} Y$, so that

$$\Phi(X, Y) = \tilde{X}' \Lambda \tilde{Y} = \sum_{k=1}^{d} \lambda_k \tilde{X}_k \tilde{Y}_k$$

i.e. there is no cross-interaction between $\tilde{X}_k$ and $\tilde{Y}_l$ for $k \neq l$.

One may check that the affinity matrix between $\tilde{X}$ and $\tilde{Y}$ is indeed diagonal, and

$$A^{\tilde{X},\tilde{Y}} = \left(\mathbf{\Sigma}_X^{-1/2} U'\right)^{-1} A^{XY} \left(V\mathbf{\Sigma}_Y^{-1/2}\right)^{-1} = \Lambda$$

Some remarks:
- The rows of $U\mathbf{\Sigma}_X^{-1/2}$ and $V\mathbf{\Sigma}_Y^{-1/2}$ are the weights of indices of mutual attractivness for men and women.
- Quantity $\lambda_k / (\sum_k \lambda_k)$ is the share of observable surplus explained by i^{th} pair of indices.

Asymptotic distribution. In this section, one would like test hypotheses on the rank of the affinity matrix Λ.

Theorem. The following convergence holds in distribution for $n \longrightarrow +\infty$:

$$n^{1/2}\left(\hat{\Theta} - \Theta\right) \Longrightarrow \mathcal{N}\left(0, \mathbb{V}\right)$$

where

$$\begin{aligned} \mathbb{V} &= \mathbb{T}_{XY}\mathbb{F}^{-1}\mathbb{T}'_{XY} + \mathbb{T}_X\mathbb{K}_{XX}\mathbb{T}'_X + \mathbb{T}_Y\mathbb{K}_{YY}\mathbb{T}'_Y \\ &\quad + \mathbb{T}_X\mathbb{K}_{XY}\mathbb{T}'_Y + \mathbb{T}_Y\mathbb{K}'_{XY}\mathbb{T}_X \end{aligned}$$

and

$$\begin{aligned} \mathbb{T}_X &= \left(\Sigma_Y^{1/2}A' \otimes I\right)\left(\Sigma_X^{1/2} \otimes I + I \otimes \Sigma_X^{1/2}\right)^{-1} && (7) \\ \mathbb{T}_{XY} &= \Sigma_Y^{1/2} \otimes \Sigma_X^{1/2} && (8) \\ \mathbb{T}_Y &= \left(I \otimes \Sigma_X^{1/2}A\right)\left(\Sigma_Y^{1/2} \otimes I + I \otimes \Sigma_Y^{1/2}\right)^{-1} && (9) \end{aligned}$$

The previous result is used to test the rank of the affinity matrix Λ. Here, we use results from Kleibergen and Paap (2006) – see also Robin and Smith (2000). One would like to test the null hypothesis

$$H_0 : rank\,(\Lambda) = p.$$

Following Kleibergen and Paap, the singular value decomposition $\hat{\Theta} = \hat{U}'\hat{\Lambda}\hat{V}$ is written blockwise

$$\hat{\Theta} = \begin{pmatrix} \hat{U}'_{11} & \hat{U}'_{21} \\ \hat{U}'_{12} & \hat{U}'_{22} \end{pmatrix} \begin{pmatrix} \hat{\Lambda}_1 & 0 \\ 0 & \hat{\Lambda}_2 \end{pmatrix} \begin{pmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{pmatrix}$$

where the blocks are dimensioned so that $\hat{U}'_{11}$ and $\hat{V}_{11}$ are two $p \times p$ square matrices. Define

$$\begin{aligned} \hat{T}_p &= \left(\hat{U}'_{22}\hat{U}_{22}\right)^{-1/2} \hat{U}'_{22}\hat{\Lambda}_2\hat{V}_{22} \left(\hat{V}'_{22}\hat{V}_{22}\right)^{-1/2} \\ \hat{A}_{p\perp} &= \begin{pmatrix} \hat{U}'_{21} \\ \hat{U}'_{22} \end{pmatrix} \left(\hat{U}'_{22}\right)^{-1} \left(\hat{U}'_{22}\hat{U}_{22}\right)^{1/2} \\ \hat{B}_{p\perp} &= \left(\hat{V}'_{22}\hat{V}_{22}\right)^{1/2} \hat{V}_{22}^{-1} \begin{pmatrix} \hat{V}_{21} & \hat{V}_{22} \end{pmatrix}. \end{aligned}$$

We get, as a consequence of Kleibergen and Paap, Theorem 1:

Theorem. Under H_0, the following convergence holds in distribution for $n \to +\infty$:

$$n^{1/2}\hat{T}_p \Longrightarrow \mathcal{N}\left(0, \Omega_p\right)$$

where $\Omega_p = \left(B_{p\perp} \otimes A'_{p\perp}\right) \mathbb{V} \left(B_{p\perp} \otimes A'_{p\perp}\right)'$. As a result, the test-statistic

$$n\hat{T}'_p\hat{\Omega}_p^{-1}\hat{T}_p$$

converges under H_0 to a $\chi^2\left(\left(d_x - p\right)\left(d_y - p\right)\right)$ random variable.

6 Empirical Results

6.1 The data

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:

- Detailed information about all individuals in the household: allows us to reconstruct "couples".
- Rich information set: socio-demographic variables (birth year and education), morphology (height and weight), self-assessed health and information about personality traits.
- 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.

- Lower education: lower vocational training, kindergarten and primary education, continued primary education or elementary secondary education,
- Intermediate education: secondary education, junior vocational training
- Higher education: 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 from a range of 1 to 5. Higher score, better health.

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.

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 of 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).

6.2 Results: 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)

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	

Note: Bold coefficients indicates coefficients larger than 0.5.

7 Conclusion

Summary:

- Saliency analysis allows us to test
  - on how many dimensions sorting occurs
  - which dimensions are most relevant
- Saliency analysis is grounded in structural equilibrium model (continuous Choo and Siow (2006) model)
- Indices of mutual attractiveness, in contrast to Canonical Correlation for instance, have a structural interpretation and are therefore informative about agents' preferences.

Empirical findings: on the Dutch marriage market:

1. Sorting is multidimensional (at least 10 dimensions)

2. Personality traits and attitude towards risk matter: personality traits explain 19% of the surplus (education 26%). Different traits matter differently for men and women.

- Open question: can we extend saliency analysis beyond quadratic surplus?

Appendix: Questionnaire about personality and attitudes

Personality traits, the 16PA scale.

Now we would like to know how you would describe your personality. Below we have mentioned a number of personal qualities in pairs. The qualities are not always opposites. Please indicate for each pair of qualities which number would best describe your personality. If you think your personality is equally well characterized by the quality on the left as it is by the quality on the right, please choose number 4. If you really don't know, type 0 (zero). Scale: 1 2 3 4 5 6 7

TEG1: oriented towards things oriented towards people.

TEG2 slow thinker quick thinker.

TEG3: easily get worried not easily get worried.

TEG4: flexible, ready to adapt myself stubborn, persistent.

TEG5: quiet, calm vivid, vivacious.

TEG6: carefree meticulous.

TEG7 shy dominant.

TEG8: not easily hurt/offended sensitive, easily hurt/offended.

TEG9: trusting, credulous suspicious.

TEG10: oriented towards reality dreamer.

TEG11: direct, straightforward diplomatic, tactful.

TEG12: happy with myself doubts about myself.

TEG13: creature of habit open to changes.

TEG14: need to be supported independent, self-reliant.

TEG15: little self-control disciplined.

TEG16: well-balanced, stable irritable, quick-tempered.

Attitude towards risk.

The following statements concern saving and taking risks. Please indicate for each statement to what extent you agree or disagree, on the basis of your personal opinion or experience.

totally disagree totally agree: 1 2 3 4 5 6 7

SPAAR1: I think it is more important to have safe investments and guaranteed returns, than to take a risk to have a chance to get the highest possible returns.

SPAAR2: I would never consider investments in shares because I find this too risky.

SPAAR3: if I think an investment will be profitable, I am prepared to borrow money to make this investment.

SPAAR4: I want to be certain that my investments are safe.

SPAAR5: I get more and more convinced that I should take greater financial risks to improve my financial position.

SPAAR6: I am prepared to take the risk to lose money, when there is also a chance to gain money.

8 Appendix: canonical correlation is biased

Let P be the distribution of (X_1, X_2) where X_1 takes value 1 with probability $1/2$ and -1 with probability $1/2$, and X_2 is exponentially distributed with parameter 1 and independent of X_1. Let G be the c.d.f. of X_2, so that $G(z) = 1 - \exp(-z)$. Let $Q = \mathcal{U}([0,1])$. Set $\lambda_1 = \lambda_2 = 1/\sqrt{2}$, so that $\hat{X} = \frac{X_1 + X_2}{\sqrt{2}}$. Hence the optimal coupling $\left(\hat{X}, \hat{Y}\right)$ is such that $\hat{Y} = F_{\hat{X}}\left(\hat{X}\right)$ where $F_{\hat{X}}(.)$ is the c.d.f. of $\hat{X}$, which is expressed as

$$F_{\hat{X}}(x) = \frac{1}{2}\left(G\left(x\sqrt{2}+1\right) + G\left(x\sqrt{2}-1\right)\right).$$

Thus

$$\hat{Y} = \begin{cases} \frac{1}{2}\left(G(X_2) + G(X_2 - 2)\right) \text{ if } X_1 = -1 \\ \frac{1}{2}\left(G(X_2 + 2) + G(X_2)\right) \text{ if } X_1 = 1, \end{cases}$$

and a calculation shows that

$$cov\left(X_1, \hat{Y}\right) = \frac{\mathbb{E}G\left(X_2+2\right) - \mathbb{E}G\left(X_2-2\right)}{4}$$

and as

$$\begin{aligned}\mathbb{E}G\left(X_2+2\right) &= 1 - e^{-2}/2 \text{ and} \\ \mathbb{E}G\left(X_2-2\right) &= e^{-2}/2,\end{aligned}$$

we get

$$cov\left(X_1, \hat{Y}\right) = \frac{1}{4}\left(1 - e^{-2}\right). \tag{10}$$

Similarly,

$$\begin{aligned}\mathbb{E}\left[X_2\hat{Y}\right] &= \frac{1}{4}\mathbb{E}\left[X_2G\left(X_2-2\right)\right] + \frac{1}{4}\mathbb{E}\left[X_2G\left(X_2+2\right)\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[X_2G\left(X_2\right)\right]\end{aligned}$$

and using the fact that $\mathbb{E}\left[X_2G\left(X_2-2\right)\right] = 7e^{-2}/4$, that $\mathbb{E}\left[X_2G\left(X_2+2\right)\right] = 1 - e^{-2}/4$, and that

$$\mathbb{E}\left[X_2G\left(X_2\right)\right] = 3/4,$$

we get $\mathbb{E}\left[X_2\hat{Y}\right] = \left(3e^{-2}+5\right)/8$, hence, as

$$\mathbb{E}\left[X_2\right]\mathbb{E}\left[\hat{Y}\right] = 1/2,$$

one gets

$$cov\left(X_2, \hat{Y}\right) = \frac{3e^{-2}+1}{8}. \tag{11}$$

Now the Canonical Correlation estimator $(\lambda_1^c, \lambda_2^c)$ of (λ_1, λ_2) solves in this setting

$$\max_{\hat{\lambda}_1, \hat{\lambda}_2} \hat{\lambda}_1 cov\,(X_1, Y) + \hat{\lambda}_2 cov\,(X_2, Y)$$
$$s.t.\ \hat{\lambda}_1^2 + \hat{\lambda}_2^2 = 1$$

which implies

$$\frac{\lambda_2^c}{\lambda_1^c} = \frac{cov\left(X_2, \hat{Y}\right)}{cov\left(X_1, \hat{Y}\right)}.$$

Using (10) and (11), this becomes

$$\frac{\lambda_2^c}{\lambda_1^c} = \frac{3+e^2}{2e^2-2} \simeq 0.81 \neq \frac{\lambda_2}{\lambda_1} = 1.$$

Therefore the Canonical Correlation estimator is not consistent in this example.