Personality traits and the marriage market

Arnaud Dupuy, Alfred Galichon

Personality Traits and the Marriage Market

Arnaud Dupuy

Centre for Population, Poverty and Public Policy Studies/INSTEAD, Maastricht School of Management, and Institute for the Study of Labor (IZA)

Alfred Galichon

Sciences Po, Institute for the Study of Labor (IZA), and Centre for Economic Policy Research

Which and how many attributes are relevant for the sorting of agents in a matching market? This paper addresses these questions by constructing indices of mutual attractiveness that aggregate information about agents' attributes. The first k indices for agents on each side of the market provide the best approximation of the matching surplus by a k-dimensional model. The methodology is applied on a unique Dutch household survey containing information about education, height, body mass index, health, attitude toward risk, and personality traits of spouses.

I. Introduction

Marriage, understood in a broad sense, is probably one of the most important factors for happiness (see, e.g., Stutzer and Frey 2006; Zimmermann and Easterlin 2006). It also plays an important role in the gen-

We thank five anonymous referees, the editor (Phil Reny), as well as Raicho Bojilov, Odran Bonnet, Xavier Gabaix, Jim Heckman, Zuzanna Kosowska-Stamirowska, Jean-Marc Robin, Marko Terviö, Bertrand Verheyden, Simon Weber, and seminar participants at the University of Chicago, Université de Montréal, Paris 1 Panthéon—Sorbonne, University of Alicante, Tilburg University, Sciences Po, Harvard—Massachusetts Institute of Technology, Université de Lausanne, and the 2013 European Economic Association meeting for their comments and Lex Borghans for useful discussions about the DNB data. Jinxin He provided excellent research assistance. Dupuy warmly thanks ROA at Maastricht University, where part of this paper was written. Galichon's research has received funding from the European Research Council under the European Union's Seventh Framework Programme

[*Journal of Political Economy,* 2014, vol. 122, no. 6]

eration of welfare and its redistribution across individuals. An in-depth understanding of marriage patterns is therefore of crucial importance for the study of a wide range of economic issues. A growing body of the economic literature studies the determinants of marriage, seen as a competitive matching market, both empirically and theoretically. This literature draws insights from the seminal model of the marriage market developed by Becker (1973). At the heart of Becker's theory lies a two-sided assignment model with transferable utility in which agents on both sides of the market (men and women) are characterized by a set of attributes only partly observed by the researcher. Each agent aims at matching with a member of the opposite sex so as to maximize his or her own payoff. This model is particularly interesting since under certain conditions, one can identify and estimate features of agents' preferences. A central question in this market is which and how many attributes are relevant for the sorting of agents?

A large body of literature has focused on the identification and estimation of preferences in the marriage market and in other matching markets;[1] however, it has been constrained by some methodological limitations regarding the quantitative methods available to identify and estimate features of the joint utility function. In the current state of the art, no estimation tool can handle sorting on multiple continuous attributes in a convenient manner. Until recently, most empirical literature assumed that sorting occurs on a single continuous dimension, which is a single index aggregating the various attributes of the agents. The choice of this approach was strongly influenced by Becker's seminal model of "positive assortative mating," which is essentially single-dimensional. Because of this limitation, empirical studies to date have therefore either focused on one attribute at a time, hence ignoring the effect of other attributes on sorting (see, e.g., Charles et al. 2013), or assumed that all observed attributes matter but only through a single index of mutual attractiveness (see, e.g., Wong 2003; Anderberg 2004; Chiappori et al. 2012). More recently, however, a new vein of the literature initiated by Choo and Siow (2006), and pursued by Chiappori et al. (2010), Fox (2010, 2011), and Galichon and Salanié (2010, 2013), among others, has

(FP7/2007–13)/European Research Council grant agreement 313699, and from FiME, Laboratoire de Finance des Marchés de l'Energie. Data are provided as supplementary material online.

[1] For the marriage markets, see, among others, Becker (1991), Wong (2003), Anderberg (2004), Choo and Siow (2006), Chiappori and Oreffice (2008), Chiappori, Salanié, and Weiss (2010), Hitsch, Hortacsu, and Ariely (2010), Oreffice and Quintana-Domeque (2010), Bruze (2011), Chiappori, Oreffice, and Quintana-Domeque (2012), Charles, Hurst, and Killewald (2013), Echenique et al. (2013), Jacquemet and Robin (2013), and Browning, Chiappori, and Weiss (2014). For other markets, see, e.g., Gabaix and Landier (2008), Terviö (2008), and Fox (2010, 2011). A nice survey is Graham (2011).

built on discrete-choice theory and is therefore restricted to the case of discrete characteristics. It seems fair to assess that a standard procedure for the estimation of continuous multivariate matching models is still needed, in spite of recent attention on the matter.[2] Another limitation of the current empirical literature is related to the set of observable attributes available in the data. Most studies solely have access to data on education and earnings, and only a few observe other dimensions such as anthropometric measures captured by height and body mass index (BMI) or self-assessed measures of health (Oreffice and Quintana-Domeque [2010] and Chiappori et al. [2012] are notable exceptions).

In the present paper, we contribute to the literature on three accounts. First, on the modeling front, we extend (i) the Choo and Siow matching model to account for possibly continuous multivariate attributes and (ii) Galichon and Salanié's (2010, 2013) surplus estimator of the Choo and Siow model to the continuous case.[3] Extending Choo and Siow's model to continuous regressors is an important problem, which has been left open so far. Indeed, many attributes that appear in empirical studies on the marriage market are intrinsically continuous: income, wealth, height, BMI, and, as our paper illustrates, psychometric attributes such as personality traits. Even if measuring necessarily involves discretization, it remains desirable to have models that treat attributes as continuous. Using the Choo and Siow model directly on the discretized attributes to perform inference is problematic since changing the level of discretization of the data will imply modifying the assumptions of the model. To solve this problem, we make use of a continuous version of the logit choice framework, pioneered by Cosslett (1988) and Dagsvik (1994), which relies on extreme value stochastic processes. This ensures that our assumptions do not depend on the level of discretization of the data.

Second, on the data analysis front, we introduce a new technique, which we call "saliency analysis," to determine the most relevant dimensions on which sorting occurs in a matching market. The starting point of this analysis requires inferring the strength of complementarities between men's and women's attributes. Using our structural model, we evaluate the intensity of assortativeness (positive or negative) between any pair of attributes, and we call the resulting matrix "the affinity matrix." Saliency analysis consists in analyzing the affinity matrix by means of a singular value decomposition. This allows one to derive "indices of mutual attractiveness," such that the joint utility of matching is a sum of mutually

[2] Recently, two papers have studied markets in which sorting occurs on more than one dimension. Coles and Francesconi (2011) and Chiappori et al. (2012) study sorting on a single continuous index and a binary variable. Nesheim (2012) focuses on the identification of multivariate hedonic models without heterogeneity and based on the observation of the price.

[3] Neither of these two papers allows for continuous observable characteristics.

exclusive pairwise interaction terms. The first k indices (for males and females) provide a convenient approximation of the joint utility by a model in which attributes are vectors of only k dimensions. As a consequence, one can perform inference on the number of dimensions that are required to explain the equilibrium sorting by testing how many singular values differ from zero.

Third, on the empirical front, we make use of a data set that allows us to observe a wide range of attributes of both spouses. The set of attributes we observe in the data includes socioeconomic variables such as education, anthropometric measures such as height and BMI, a measure of self-assessed health, as well as psychometric attributes such as risk aversion and the "big five" personality traits well known in psychology: conscientiousness, extraversion, agreeableness, emotional stability, and autonomy. This paper is, to the extent of our knowledge, the first attempt to evaluate the importance of personality traits in the sorting of men and women in the marriage market. We will show that although education explains 28 percent of a couple's observable joint utility, personality traits explain another 17 percent and different personality traits matter differently for men and for women. Our results relate to the literature showing the importance of personality traits in making economic decisions (e.g., Borghans et al. 2008). Bowles, Gintis, and Osborne (2001) and Mueller and Plug (2006), among others, have shown the importance of personality traits for earnings inequality. Closer to our focus, Lundberg (2012) studies the impact of personality traits on the odds in and out of a relationship (marriage and divorce) and finds empirical evidence that personality traits significantly affect the extensive margin in the marriage market. In particular, conscientiousness increases the probability of marriage at the age of 35 for men and extraversion increases the odds of marriage at the age of 35 for women. In the present work, we study the intensive margin, that is, to whom conscientious men and extraverted women are the most attractive. We show among other things that conscientious men have preferences for conscientious women whereas extraverted women have preferences for autonomous and less agreeable men.

The rest of the paper is organized as follows. Section II presents an important extension of the model of Choo and Siow to continuously distributed observables. Section III deals with parametric estimation of the joint utility function in this setting. Section IV presents a methodology for deriving indices of mutual attractiveness that determine the principal dimensions on which sorting occurs. The problem of inferring the number of dimensions on which sorting occurs is dealt with in Section V. Section VI presents the data used for our empirical estimation, and Section VII discusses the results. Section VIII presents conclusions.

II. The Continuous Choo and Siow Model

A. *The Becker-Shapley-Shubik Model of Marriage*

The setting is a one-to-one, bipartite matching model with transferable utility. Men and women are characterized by vectors of attributes, respectively denoted $x \in \mathcal{X} = \mathbb{R}^{d_x}$ for men and $y \in \mathcal{Y} = \mathbb{R}^{d_y}$ for women. Matched men and women are by definition in equal number; we let P and Q be the respective probability distributions of their attributes. Throughout the paper, P and Q are treated as exogenous, except in Appendix D, where we show that incorporating singles leaves the analysis unchanged while allowing us to identify reservation utilities. The distributions P and Q are assumed to have densities with respect to the Lebesgue measure denoted, respectively, f and g.[4] Without loss of generality, it is assumed throughout that P and Q are centered distributions, that is, $\mathbb{E}_P[X] = \mathbb{E}_Q[Y] = 0$.

A *matching* is the probability density $\pi(x, y)$ of occurrence of a couple with characteristics (x, y) from the matched population. Quite obviously, this imposes that the marginals of π should be P and Q. Write $\pi \in \mathcal{M}(P, Q)$, where

$$\mathcal{M}(P, Q) = \left\{ \pi : \pi(x, y) \geq 0, \int_{\mathcal{Y}} \pi(x, y)\, dy = f(x), \right.$$
$$\left. \text{and} \int_{\mathcal{X}} \pi(x, y)\, dx = g(y) \right\}.$$

Let $\Phi(x, y)$ be the joint utility generated when a man x and a woman y match, which is shared endogenously between them. Let $\Phi(x, \varnothing)$ and $\Phi(\varnothing, y)$ be the utility of man x and woman y, respectively, if they remain single; in Appendix D, we shall show that $\Phi(x, \varnothing)$ and $\Phi(\varnothing, y)$ are identified if and only if the populations of singles are observed but that the identification of $\Phi(x, y)$ is not impeded if singles are not observed.[5] In the rest of the paper we shall assume that only the matched population is observed, so we will not focus on $\Phi(x, \varnothing)$ and $\Phi(\varnothing, y)$; as a result, the matching *surplus* $\Phi(x, y) - \Phi(x, \varnothing) - \Phi(\varnothing, y)$ will not be identified. Shapley and Shubik (1972) have shown that the equilibrium matching π maximizes the total utility

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

[4] While we present the case with continuous distributions for x and y, our framework easily extends to the case in which some dimensions of x and y are discrete.

[5] In App. D, this will be shown to be a consequence of the independence of irrelevant alternatives property of the logit model.

Optimality condition (1) leads to very strong restrictions on (X, Y), which are rarely met in practice.[6] We need to incorporate some amount of unobserved heterogeneity in the model.

B. Adding Heterogeneities

Bringing the model to the data requires the additional step of acknowledging that sorting might also occur on attributes that are unobserved to the econometrician. In the case in which men's and women's attributes are *discrete*, Choo and Siow (2006) introduced unobservable heterogeneities into the matching problem by considering that if a man m of attributes $x_m = x$ and a woman w of attributes $y_w = y$ match, they create a joint utility $\Phi(x, y) + \varepsilon_m(y) + \eta_w(x)$, where $\varepsilon_m(y)$ and $\eta_w(x)$ are unobserved random "sympathy shocks" drawn by individuals. Assuming that $(\varepsilon_m(y))_y$ and $(\eta_w(x))_x$ have independent and identically distributed (i.i.d.) centered Gumbel (extreme value type I) distributions with scaling parameter $\sigma/2$, Choo and Siow have shown that the joint utility $\Phi(x, y)$ can be split into $\Phi(x, y) = U(x, y) + V(x, y)$ such that the utility of man m matching with a woman of type y is given by

$$U(x, y) + \varepsilon_m(y),$$

with a similar expression for the utility of woman w. An important implication of this setting is that at equilibrium, agents are indifferent between partners with the same observable attributes: the matching utility of man m at equilibrium depends only on the observable attributes of that woman. As a consequence, each agent in the market solves a discrete-choice problem.

In the Choo and Siow model, partners are assumed to have i.i.d. Gumbel sympathy shocks for the discrete attributes of the opposite side of the market. However, in many applied settings, these attributes are continuous random vectors, and even though the data that the analyst handles are obviously discretized, there is a strong need for a continuous framework. To illustrate, we shall take a setting in which only the height of the partners is relevant and assume that the precision of the measure is poor; say it is rounded to the nearest foot. A direct implication of the Choo and Siow assumptions is that individuals' sympathy shocks are perfectly correlated within a foot bracket and perfectly independent across feet. Suppose instead that height is measured at the nearest inch. Choo and Siow's assumptions would now imply that individuals' sympathy shocks are per-

[6] A basic result in the theory of optimal transportation (Brenier's theorem) implies that when $\Phi(x, y) = x'Ay$, the optimal matching is characterized by $(AY)_i = \partial V(X)/\partial x_i$, where V is some convex function. Hence as soon as A is invertible, the matching is *pure*, in the sense that no two men of the same type may marry women of different types. This is obviously never observed in the data.

fectly correlated within an inch bracket and perfectly independent across inches, which of course comes at odds with the previous assumptions. So, while it is of course always possible to apply the Choo and Siow setting to the discretized data, this implicitly leads to ad hoc assumptions that depend on the level of discretization of the available data.

In the present paper, we shall present an application in which x and y measure height, BMI, and various personality traits, which have a continuous multivariate distribution. Hence we need to model the random processes for $\varepsilon_m(x)$ and $\eta_w(y)$ accordingly. A legitimate candidate in the wake of Choo and Siow's approach is the continuous logit model. Although very natural and particularly tractable, this setting has been surprisingly little used in economic modeling, with some notable exceptions. McFadden (1976) initiated the literature of continuous logit models by extending the definition of independence of irrelevant alternatives beyond finite sets. Ben-Akiva and Watanatada (1981) and Ben-Akiva, Litinas, and Tsunekawa (1985) define continuous logit models by taking the limits of the discrete-choice probabilities, with applications in particular to the context of spatial choice models. Cosslett (1988) and Dagsvik (1994) have independently suggested using max-stable processes to model continuous choice. We base our approach on their insights.

Assume that each man m of type $x_m = x$ knows only a random subset of the total population of women we will call "acquaintances" and that man m considers potential partners only from his set of acquaintances. These acquaintances are indexed by $k \in \mathbb{N}$, and their observable attributes are represented by y_k^m.[7] Each of these acquaintances is associated with a random "sympathy shock" ε_k^m, which enters additively into the man's utility, so that the utility of a man m who marries a woman with attributes y_k^m can be written as

$$U(x, y_k^m) + \frac{\sigma}{2}\varepsilon_k^m, \tag{2}$$

where $U(x, y)$ is the "systematic" (in Choo and Siow's term) part of the utility obtained by man x matching with a woman with attributes y, whose existence and characterization will be provided in theorem 1 below. Note that in contrast to the original setting of Choo and Siow described above, men do not have access to the whole population of women, but only to their randomly selected set of acquaintances, which is a subset of the whole population.

We have yet to specify the distribution of y_k^m and ε_k^m. Following Cosslett and Dagsvik's idea, we assume that $\{(y_k^m, \varepsilon_k^m),\ k \in \mathbb{N}\}$ are the points of a

[7] As explained in App. A, it will result from the distributional assumptions that each man draws an infinite but countable number of acquaintances almost surely, so that these can be indexed by the set of integers.

Poisson process on $\mathcal{Y} \times \mathbb{R}$ of intensity $dy \times e^{-\varepsilon} d\varepsilon$. This means that (i) the probability that man m has an acquaintance whose observable attributes are in a small set of infinitesimal size dy around y and with sympathy shock in a set of infinitesimal size $d\varepsilon$ around ε is equal to $e^{-\varepsilon} d\varepsilon dy$, and (ii) letting S and S' be two disjoint subsets of $\mathcal{Y} \times \mathbb{R}$, the events "$m$ has an acquaintance in S" and "m has an acquaintance in S'" are independent. According to the standard theory of Poisson point processes, this implies that, for S a subset of $\mathcal{Y} \times \mathbb{R}$, the probability that man m has no acquaintance in set S is $\exp(-\int_S e^{-\varepsilon} dy d\varepsilon)$. In Appendix A we show that this yields a continuous version of the multinomial logit choice model. As a result, the probability distribution of man m choosing a woman with attributes y is given by its density of probability

$$\pi_{Y|X}(y|x) = \frac{\exp[U(x,\ y)/(\sigma/2)]}{\displaystyle\int_{\mathcal{Y}} \exp[U(x,\ y')/(\sigma/2)]\,dy'}, \tag{3}$$

which is clearly the extension of the logit formalism to the continuous choice setting. Similarly, the utility of a woman w with attributes $y_w = y$ who marries a man with attributes x is

$$V(x_l^w,\ y) + \frac{\sigma}{2}\eta_l^w, \tag{4}$$

where $V(x, y)$ is the systematic part of the utility, and $\{(x_l^w,\ \eta_l^w),\ l \in \mathbb{N}\}$ are the points of a Poisson process on $\mathcal{X} \times \mathbb{R}$ of intensity $dx \times e^{-\eta} d\eta$, so that the probability distribution of woman w choosing a man with attributes x is given by its density of probability

$$\pi_{X|Y}(x|y) = \frac{\exp[V(x,\ y)/(\sigma/2)]}{\displaystyle\int_{\mathcal{X}} \exp[V(x',\ y)/(\sigma/2)]\,dx'}. \tag{5}$$

The continuous logit framework inherits the structural assumptions of the discrete multinomial logit model. In particular, the *independence* property, which implies that the sympathy shock for women whose attributes are in a small set around y is perfectly uncorrelated with the sympathy shock for women whose attributes are in a small set around $y' \neq y$. Hence the logit framework (continuous or discrete) does not allow for a systematic sympathy shock, that is, correlated sympathy shocks across observables. In the example in which the attribute of interest is height, it may be desirable to accommodate a random sympathy shock for height (some men prefer taller women, some prefer shorter women, regardless of their own observable attributes). We conjecture, however, that if the amount of variation of the unobserved heterogeneity is small, the misspecification of the sympathy shocks has only a minor impact on the market outcome and the identification of the joint utility.

Taking the logarithm of equations (3) and (5), respectively, yields $U - a = (\sigma/2)\log\pi$ and $V - b = (\sigma/2)\log\pi$, where

$$a(x) = \frac{\sigma}{2}\log\int_{\mathcal{Y}} \frac{e^{[U(x,y')/(\sigma/2)]}}{f(x)}\,dy', \quad b(y) = \frac{\sigma}{2}\log\int_{\mathcal{X}} \frac{e^{[V(x',y)/(\sigma/2)]}}{g(y)}\,dx', \tag{6}$$

and since $\Phi = U + V$, one obtains by summation

$$\log\pi(x, y) = \frac{\Phi(x, y) - a(x) - b(y)}{\sigma}. \tag{7}$$

We formalize this result in theorem 1, which extends Galichon and Salanié (2010) to the continuous case.

THEOREM 1. Under the assumptions stated above, the following hold:

i. The equilibrium matching π maximizes the social gain

$$\max_{\pi\in\mathcal{M}(P,Q)} \int\int_{\mathcal{X}\times\mathcal{Y}} \Phi(x, y)\pi(x, y)\,dxdy - \sigma\int\int_{\mathcal{X}\times\mathcal{Y}} \log\pi(x, y)\pi(x, y)\,dxdy. \tag{8}$$

ii. In equilibrium, for any $x\in\mathcal{X}$, $y\in\mathcal{Y}$,

$$\pi(x, y) = \exp\left[\frac{\Phi(x, y) - a(x) - b(y)}{\sigma}\right], \tag{9}$$

where the potentials $a(x)$ and $b(y)$ are determined such that $\pi\in\mathcal{M}(P, Q)$. They exist and are uniquely determined up to a constant.

iii. A man m with attributes x who marries a woman k^* from his set of acquaintances obtains utility

$$U(x, y^m_{k^*}) + \frac{\sigma}{2}\varepsilon^m_{k^*} = \max_k\left[U(x, y^m_k) + \frac{\sigma}{2}\varepsilon^m_k\right], \tag{10}$$

where

$$U(x, y) = \frac{\Phi(x, y) + a(x) - b(y)}{2}. \tag{11}$$

Similarly, a woman w with attributes y who marries man l^* from her set of acquaintances obtains utility

$$V(x^w_{l*}, y) + \frac{\sigma}{2}\eta^w_{l*} = \max_l \left[V(x^w_l, y) + \frac{\sigma}{2}\eta^w_l \right], \tag{12}$$

where

$$V(x, y) = \frac{\Phi(x, y) - a(x) + b(y)}{2}. \tag{13}$$

As in Galichon and Salanié (2010, 2013) and, independently, Decker et al. (2013), part i of this result expresses the fact that the equilibrium matching reflects a trade-off between sorting on the observed attributes (which tends to maximize the term $\int \Phi(x, y)\pi(x, y)dxdy$) and sorting on the unobserved attributes (which in turn tends to maximize the entropic term $\int \log \pi(x, y)\pi(x, y)dxdy$). The second term will therefore pull the solution toward the random matching, where partners are randomly assigned; the parameter σ, which captures the intensity of the unobserved heterogeneity, measures the intensity of this trade-off. The smaller the σ (i.e., the less unobserved heterogeneity in the model), the closer the solution will be to the solution without heterogeneity. On the contrary, the higher the σ, the larger the probabilistic independence between the observed attributes of men and women. As an illustration, we consider the simple toy example below, in which this phenomenon is explicit.

Example 1. When P and Q are the standard univariate Gaussian distribution $\mathcal{N}(0, 1)$ and $\Phi(x, y) = -\frac{1}{2}(x - y)^2$, the equilibrium matching π is such that $\pi_{Y|X}(y|x) = \mathcal{N}(tx, 1 - t^2)$, where $t = \sqrt{(\sigma^2/4) + 1} - (\sigma/2)$. Hence, $\sigma = 0$ implies $t = 1$, in which case $Y = X$ (sorting predominates and we have positive assortative matching), while $\sigma \to \infty$ implies $t \to 0$, in the limit of which Y becomes independent from X (unobserved heterogeneity predominates; there is no sorting on observables). Closed-form formulas can also be provided in the multivariate case in which P and Q are Gaussian and Φ is quadratic; see Bojilov and Galichon (2013).

Part ii of theorem 1 is an expression of the first-order optimality conditions. The program is an infinite-dimensional linear programming problem in which $a(x)$ and $b(y)$ are the Lagrange multipliers corresponding to the constraints $\int \pi(x, y)dy = f(x)$ and $\int \pi(x, y)dx = g(y)$, respectively. Equation (9), or more precisely its logarithmic transform equation (7), will be the basis of our estimation strategy. Together with the constraint $\pi \in \mathcal{M}(P, Q)$, this equation provides a nonlinear system of equations in $a(\cdot)$ and $b(\cdot)$. In the applied mathematical literature it is known as the *Schrödinger-Bernstein equation*, or more commonly as the *Schrödinger problem.* Existence and uniqueness (up to a constant) are well studied under very general conditions on P and Q; see, for instance, Rüschendorf and Thomsen (1993) and references therein. An efficient algorithm for the numerical determination of the solution based on a fixed-point idea is studied in Rüschendorf (1995). For completeness, it is further explained in Appendix C.

Part iii of theorem 1 explains how the joint utility is shared at equilibrium. Unsurprisingly, just as in Choo and Siow (2006) and the ensuing literature, individuals do not transfer their sympathy shock at equilibrium, which is expressed by (10) and (12). Expressions (11) and (13) provide the formulas for the systematic part of the utility. As previously noted, $a(x)$ and $b(y)$ are the Lagrange multipliers of the scarcity constraints of men's observable attributes x and women's attributes y. Hence a higher $a(x)$ shall imply a higher relative scarcity for x and therefore a greater prospect for utility extraction.

Identification.—From an identification perspective, note that equations (3) and (5) imply that the observation of $\pi(x, y)$ leads to identification of $U(x, y)$ up to an additive term $c(x)$ and, similarly, $V(x, y)$ up to an additive term $d(y)$ by

$$U(x, y) = \frac{\sigma}{2}[\log\pi_{Y|X}(y|x) + c(x)],$$
$$V(x, y) = \frac{\sigma}{2}[\log\pi_{X|Y}(x|y) + d(y)],$$

and thus

$$\Phi(x, y) = \frac{\sigma}{2}[\log \pi_{Y|X}(y|x) + \log\pi_{X|Y}(x|y) + c(x) + d(y)].$$

As a result, $\Phi(x, y)$ is identified up to a separatively additive function since we restrict our attention to the matched population.[8] Since $\Phi(x, y)$ yields the same equilibrium matching as $\Phi(x, y) + c(x) + d(y)$, the identified quantity is actually the cross-derivative $\partial^2\Phi(x, y)/\partial x\partial y$, while neither $\partial\Phi(x, y)/\partial x$ nor $\partial\Phi(x, y)/\partial y$ can be identified, nor can their signs be identified. To illustrate, assume that there is only one dimension—education. It may be that men and women who are more educated generate more utility, which we call "absolute attractiveness," and which translates into $\partial\Phi(x, y)/\partial x \geq 0$ and $\partial\Phi(x, y)/\partial y \geq 0$. However, this is not identifiable in our model because models in which the joint utility is $\Phi(x, y)$ are observationally indistinguishable from models in which the joint utility is $\Phi(x, y) + c(x) + d(y)$, and the terms c and d might be strongly negatively correlated with education. Instead, the present framework allows us to determine whether education is mutually attractive in the sense that $\partial^2\Phi(x, y)/\partial x\partial y \geq 0$, meaning not only that highly educated men and women attract each other but also that lower-educated men and women attract each other. Hence, our model allows us to measure the strength

[8] Appendix D explains how these results are extended when singles are observed.

of *mutual attractiveness* (or assortativeness) on various dimensions, but not *absolute attractiveness.*

III. Parametric Estimation

A. *Specification of the Matching Utility*

In this section, we shall specify a parametric form for the joint utility function, the estimation of which shall be discussed in the next section. While Choo and Siow's estimator is fully nonparametric, the fact that the variables under study are continuous reinforces the need for a parametric estimator. For the purpose of this discussion, we shall look back at the illustrative example from the introduction in which only both partners' heights are observed. The Choo and Siow analysis provides a nonparametric estimator for the joint utility $\Phi(x, y)$ of a match between a man of height x and a woman of height y. If heights were to be rounded to the nearest inch and individuals' heights in inches ranged, say, from 50 to 90, then the dimension of vector $\Phi(x, y)$ would be $40 \times 40 = 1{,}600$. Note that this would worsen significantly if several characteristics were observed. But even in the single-dimensional case, there would be a serious missing data problem, since the odds that one would observe data for every pair of heights are virtually zero. Moreover, even if one were lucky enough to obtain the full nonparametric estimator of $\Phi(x, y)$, one would have to heavily process this information before being able to draw any stylized conclusion. This simple example highlights the need for a parametric estimation when considering continuous variables.

Throughout the rest of the paper, we shall assume a quadratic parameterization of Φ: for A a $d_x \times d_y$ matrix, we take

$$\Phi_A(x, y) = x'Ay,$$

where we call matrix A the *affinity matrix.* One has

$$A_{ij} = \frac{\partial^2 \Phi(x, y)}{\partial x_i \partial y_j}.$$

The parameter A_{ij} accounts for the strength of mutual attractiveness (which can be positive or negative) between dimensions x_i and y_j. It measures how the (marginal) gain in joint utility from increasing the man's ith attribute evolves as the woman's jth attribute increases. It captures the intensity of the complementarity/substitutability between attribute x_i of man x and attribute y_j of woman y in the joint utility.

Two comments about this parametric choice are noteworthy. First, this parametric choice is arguably the simplest one that captures nontrivial complementarities between any pair of attributes. If $A_{ij} > 0$, x_i and y_j are complements, and (all things else being equal) high x_i tend to match with high y_j. It reflects positive assortative matching across men's ith attribute and women's jth attribute. For instance, the level of education of one of the partners may be complementary with the risk aversion of the other partner. On the contrary, if $A_{ij} < 0$ and x_i and y_j are substitutes, there is negative assortative matching between x_i and y_j. Note that attributes x and y should not be interpreted as an absolute quality (where a greater value of x_i, the ith dimension of x, would be more socially desirable than a smaller value of x_i). In fact, the model is observationally indistinguishable from a model in which x is changed into $-x$ and y is changed into $-y$.

Second, this quadratic setting in which Φ is bilinear in x and y is less restrictive than it seems and can be extended to the case in which the various observed attributes have nonlinear contributions to the joint utility.[9] For instance, it may be plausible that extraverted men are indifferent about the education and the height of their wives, but that if a woman is tall, then men prefer her with a higher education. Our setting can easily be extended to incorporate such nonlinear effects. We assume no restrictions on the attributes that enter x and y, so that the observables can be enriched by the addition of nonlinear functions of them, that is, adding x_i^2, x_i^3, and so forth and $x_i x_j$ as observable attributes for men and similarly for women. This will allow $\Phi(x, y)$ to be any polynomial function of x and y. Thus, our setting can easily incorporate any utility function that is a polynomial expression of the observable attributes.

B. *Inference*

We turn to the estimation of the affinity matrix A. The technique we apply here was introduced by Galichon and Salanié (2010); we discuss this extension to the continuous case. By taking the cross-derivative of equation (7), one has

$$\frac{A_{ij}}{\sigma} = \frac{\partial^2 \log \pi(x, y)}{\partial x_i \partial y_j}. \tag{14}$$

A seemingly natural procedure would consist in estimating π nonparametrically and obtaining A from the cross-derivatives with respect to x_i and y_j. While feasible, this procedure faces a number of issues both in theory and in practice. First, it requires a nonparametric estimation of the second derivatives of the log likelihood, which is quite challenging: the "curse of dimensionality" would fully apply.[10] Second, since equa-

[9] We thank a referee for pointing this out.

[10] In our application, both x and y have 10 dimensions, so (x, y) is of dimension 20.

tion (14) is valid at any point (x, y), this equation is an overidentifying restriction to the estimation of A. The right-hand side of (14) depends on (x, y), while the left-hand side does not. One may certainly take some averaging of the right-hand side of equation (14), but it is not quite obvious how to weigh each point optimally, and it would only partially offset the problems stemming from the curse of dimensionality. As a result, this procedure will be statistically inefficient.

Instead, we prefer to resort to a moment matching procedure, which is relatively simple while achieving asymptotic statistical efficiency as shown in theorem 2 below. Let us provide intuition for this method. Each value of the matrix A yields an equilibrium matching distribution, which we denote $\pi^A(x, y)$. As argued in Appendix C, π^A can be computed efficiently using a fixed-point method. Recall that we have assumed that the distributions of X and Y have zero mean and introduce the *cross-covariance matrix*

$$\Sigma_{XY} = (\mathbb{E}[X_i Y_j])_{ij} = \mathbb{E}[XY'], \tag{15}$$

which is observed in the data. The idea is to look for the value of A such that, for all i and j, the covariances predicted by the model match the covariances observed in the data, that is,

$$\mathbb{E}_{\pi^A}[X_i Y_j] = \mathbb{E}[X_i Y_j]. \tag{16}$$

This yields a map $A \to (\mathbb{E}_{\pi^A}[X_i Y_j])_{ij}$ that is invertible. The inversion of this map (in order to estimate A) can be formulated as a convex optimization problem, thus making it easy to solve numerically. To see this, we shall recall that the equilibrium matching π maximizes the social gain

$$\mathcal{W}_\sigma(A) := \max_{\pi \in \mathcal{M}(P,Q)} \mathbb{E}_\pi[X'AY] - \sigma \mathbb{E}_\pi[\ln \pi(X, Y)], \tag{17}$$

and we see that models with parameters (A, σ) and models with parameters $(A/\sigma, 1)$ are observationally equivalent, which translates mathematically into positive homogeneity $\mathcal{W}_\sigma(A) = \sigma \mathcal{W}_1(A/\sigma)$. By the envelope theorem, the predicted covariance between X_i and Y_j coincides with the partial derivative of $\mathcal{W}_\sigma$ with respect to A_{ij}, that is,

$$\mathbb{E}_{\pi^A}[X_i Y_j] = \frac{\partial \mathcal{W}_\sigma}{\partial A_{ij}}(A) = \frac{\partial \mathcal{W}_1}{\partial A_{ij}}(A/\sigma), \tag{18}$$

which implies that, upon normalization $\sigma = 1$, the map $A \to (\mathbb{E}_{\pi^A}[X_i Y_j])_{ij}$ is invertible since $\mathcal{W}_1$ is strictly convex (see lemma 3). This led Galichon and Salanié (2013) to conclude, in a setting with discrete observable attributes, that $B = A/\sigma$ is identified as a solution to the following convex optimization program:

$$\min_{B\in\mathcal{M}_{d_X d_Y}(\mathbb{R})}\left\{\mathcal{W}_1(B)-\sum_{ij}B_{ij}\Sigma_{XY}^{ij}\right\},\tag{19}$$

whose first-order conditions are precisely $\partial\mathcal{W}_1(B)/\partial B_{ij}=\Sigma_{XY}^{ij}$, that is, $\mathbb{E}_{\pi^B}[X_i Y_j]=\Sigma_{XY}^{ij}$. In the present setting with continuous observable attributes, things work in an identical manner. Since the model is scale invariant, only A/σ is identified, and we normalize A so that $\|A\|=1$, where $\|A\|=(\sum_{ij}A_{ij}^2)^{1/2}$. Then A and σ are obtained by $A=B/\|B\|$ and $\sigma=1/\|B\|$. Let us denote A^{XY} the (unique) solution to this problem, which will be our estimator of the affinity matrix A. Affinity matrix A^{XY} is "dual" to cross-covariance matrix Σ_{XY} in the sense that there is a one-to-one correspondence between them by equation (18). However, the former has a structural interpretation: it measures the strength of the interactions between pairs of attributes.

At this point, it is worth commenting on the relevance of the structural approach. Indeed, it does not suffice to just look at the variance-covariance matrix inside matches to infer the sign of complementarities, as illustrated on the following example. Imagine two observed characteristics, where the first dimension is education and the second dimension is risk aversion. Suppose that we observe positive correlation in partners' educations and in partners' risk aversions (i.e., $\Sigma_{11}>0$ and $\Sigma_{22}>0$). One might naively infer that there is positive complementarity both in education and in risk aversion (i.e., $A_{11}>0$ and $A_{22}>0$). However, this is not necessarily the case; there could actually be negative complementarity in risk aversion ($A_{22}<0$), but positive association between individuals' education and risk aversion, if positive complementarity in education ($A_{11}>0$) dominates the negative complementarity in risk aversion, thus leading to positive correlation in risk aversions inside matches. The structural approach allows us to avoid this misinterpretation by allowing us to control for the marginal distributions (e.g., control for the fact that there is positive association between individuals' education and risk aversion).

Once the affinity matrix A^{XY} has been estimated, two questions arise. First, what is the rank of A^{XY}? This question is of importance since one would like to know the number of dimensions of x and y on which sorting occurs. Second, how can we construct "indices of mutual attractiveness" such that each pair of indices for men and women explains a mutually exclusive part of the matching utility? Many studies resort to a technique called "canonical correlation," which essentially relies on a singular value decomposition of Σ^{XY}. In Dupuy and Galichon (forthcoming), we argue that this technique is not well suited for studying assortative matching and that the resulting procedure is inconsistent. Instead, in Section IV, we propose a method we call "saliency analysis" in order to accurately answer these two questions. This method is essentially based on the singular value decomposition of the affinity matrix A^{XY} (instead of Σ^{XY} as in canonical correlation). Testing the rank of the affinity matrix is equivalent to testing the number of (potentially multiple) singular values different from zero. Performing this decomposition allows one to construct the indices of mutual attractiveness that each explain a separate share of the joint utility.

IV. Saliency Analysis

In this section we set out to determine the rank of the affinity matrix A^{XY} and the principal dimensions in which it operates. For this, we introduce and describe a novel technique we call *saliency analysis*, which is similar in spirit to canonical correlation but does not suffer the pitfalls of the latter. Instead of performing a singular value decomposition of the (renormalized) cross-covariance matrix Σ_{XY}, we shall perform a singular value decomposition of the affinity matrix A^{XY}, properly renormalized. This idea is similar in spirit to the proposal of Heckman (2007), who interprets the assignment matrix as a sum of Cobb-Douglas technologies using a singular value decomposition in order to refine bounds on wages.

Recall that we have defined the cross-covariance matrix $\Sigma_{XY} = \mathbb{E}_\pi[XY']$, and let us introduce S_X and S_Y the diagonal matrices whose diagonal terms are, respectively, the variances of the X_i and the Y_j, that is,

$$S_X = \operatorname{diag}[\operatorname{Var}(X_i),\ i = 1, \ldots, d_x],$$
$$S_Y = \operatorname{diag}[\operatorname{Var}(Y_j),\ j = 1, \ldots, d_y].$$

We shall work with the rescaled attributes $S_X^{-1/2}X$ and $S_Y^{-1/2}Y$, whose entries each have unit variance. By lemma 1 in Appendix B, the affinity matrix between the rescaled attributes $S_X^{-1/2}X$ and $S_Y^{-1/2}Y$ is

$$\Theta = S_X^{1/2} A^{XY} S_Y^{1/2},$$

for which a singular value decomposition of Θ yields

$$\Theta = U'\Lambda V,$$

where Λ is a diagonal matrix with nonincreasing elements $(\lambda_1, \ldots, \lambda_d)$, $d = \min(d_x, d_y)$, and U and V are orthogonal matrices. Define the vectors of *indices of mutual attractiveness*

$$\tilde{X} = US_X^{-1/2}X \quad \text{and} \quad \tilde{Y} = VS_Y^{-1/2}Y,$$

where each index is a weighted sum of the observed attributes. Let $A^{\tilde{X}\tilde{Y}}$ be the affinity matrix on the rescaled vectors of characteristics $\tilde{X}$ and $\tilde{Y}$. From lemma 1, it follows that $A^{\tilde{X}\tilde{Y}} = \Lambda$, and as a result,

$$\Phi_A(x, y) = \sum_{i=1}^{d_x} \sum_{j=1}^{d_y} A_{ij} x_i y_j = \sum_{i=1}^{d} \lambda_i \tilde{x}_i \tilde{y}_i.$$

Hence, the new indices $\tilde{x}$ and $\tilde{y}$ are such that $\tilde{x}_i$ and $\tilde{y}_j$ are complements for $i = j$, and neither complements nor substitutes if $i \neq j$. In other words, there is positive assortative matching between $\tilde{x}_i$ and $\tilde{y}_j$ for $i = j$ and no assortativeness for $i \neq j$. This justifies the choice of terminology: $\tilde{x}_i$ and $\tilde{y}_i$ are "mutually attractive" because they are complementary with each other and only with each other. All things being equal, a man with a higher $\tilde{x}_i$ tends to match with a woman with a higher $\tilde{y}_i$.

The weights of each index of mutual attractiveness constructed by saliency analysis are given by the associated row of $US_X^{-1/2}$ for men and $VS_Y^{-1/2}$ for women. The value $\lambda_i/(\sum_i \lambda_i)$ indicates the share of the observable matching utility of couples explained by the ith pair of indices. The fact that U and V are orthogonal implies strong restrictions on how $\tilde{x}$ and $\tilde{y}$ are obtained from $S_X^{-1/2}x$ and $S_Y^{-1/2}y$. In particular, this mapping preserves distances between points; that is, the distance between $\tilde{x}$ and $\tilde{x}'$ is equal to the distance between $S_X^{-1/2}x$ and $S_X^{-1/2}x'$.

We observe that in contrast to canonical correlation analysis, a convenient feature of saliency analysis is that the results do not change when the attributes are measured using different measurement units, as expressed in lemma 2. For instance, if the partners' heights are measured in feet rather than in meters, the outcome of saliency analysis does not change.

For illustrative purposes, we give a stylized example of how saliency analysis operates in a simple two-dimensional situation.

EXAMPLE 2. Assume that there are two dimensions on each side of the market and that $S_X = S_Y = Id$, so that $\Theta = A$. Suppose that

$$A = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix}. \tag{20}$$

Then the singular value decomposition of A is $A = U'\Lambda V$, where

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \Lambda = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } V = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The economic interpretation of this simple example is the following: if the joint utility is given by $\Phi(x, y) = 4x_1y_2 - x_2y_1$, then the indices of mutual attractiveness should be given by $\tilde{x}_1 = x_1$, $\tilde{x}_2 = x_2$ and $\tilde{y}_1 = y_2$, $\tilde{y}_2 = -y_1$. One has $\Phi(\tilde{x}, \tilde{y}) = 4\tilde{x}_1\tilde{y}_1 + \tilde{x}_2\tilde{y}_2$. The vectors $\tilde{x} = (x_1, x_2)$ and $\tilde{y} = (y_2, -y_1)$ can be interpreted as indices of mutual attractiveness, meaning that there is sorting between $\tilde{x}_1 = x_1$ and $\tilde{y}_1 = y_2$ and between $\tilde{x}_2 = x_2$

and $\tilde{y}_2 = -y_1$. If one was willing to approximate the model by a one-dimensional sorting model, then saliency analysis advocates to keep $\tilde{x}_1 = x_1$ as a proxy for the attributes of men and $\tilde{y}_1 = y_2$ as a proxy for the attributes of women. In this case the joint utility is approximated by $\Phi(\tilde{x}, \tilde{y}) = 4\tilde{x}_1\tilde{y}_1$.

Example 2 is the occasion to compare singular value decomposition to another matrix decomposition, the eigenvalue decomposition, which economists may be more familiar with. Eigenvalue decomposition consists in writing, whenever possible, a square matrix M as $M = R\Lambda R^{-1}$, with Λ diagonal and R invertible. In the context of saliency analysis, this decomposition cannot be performed on A as A is not necessarily a square matrix; further, as soon as A is not symmetric, this decomposition does not necessarily exist. In particular, when A is given by (20), it does not exist since A has no real eigenvalue.[11]

As example 2 also illustrates, the observation of vector Λ will allow one to draw conclusions about the multivariate nature of the sorting and on the number of dimensions on which the sorting occurs. In particular, testing for multidimensional sorting versus unidimensional sorting is equivalent to testing whether at least two singular values λ_i are significantly larger than zero, as we shall elaborate in the next section.

V. Inferring the Number of Sorting Dimensions[12]

Assume that a finite sample of size n is observed. For the sake of readability, dependence in n of the estimators will be dropped from the notation. The vector of mutual attraction weights estimated on the sample is denoted $\hat{\Lambda}$, while the vector of mutual attraction weights in the population is denoted Λ. Similarly, $\hat{A}$ is the estimator of A (which was denoted A^{XY} in Sec. III.B, where the construction of that estimator is described). Let $\hat{S}_X$, $\hat{S}_Y$, and $\hat{\Sigma}_{XY}$ be the sample estimators of S_X, S_Y, and Σ_{XY}, respectively. For a given quantity M, we shall denote

$$\delta M = \hat{M} - M, \tag{21}$$

the difference between the estimator of $\hat{M}$ and M.

Consider two important matrices associated with the large sample properties of the model. The Fisher information matrix is defined by

[11] However, the singular values of A can be interpreted as eigenvalues of a larger matrix. Indeed, letting H be the constant Hessian matrix of the map 2Φ, then H is a symmetric matrix of size $(d_x + d_y)$ written blockwise with two zero blocks on its diagonal and A and A' as off-diagonal blocks, and the eigenvalues of H are plus and minus the singular values of A (see Horn and Johnson 1991, 135).

[12] In a discussion with one of the authors, Jim Heckman suggested the intuition of the approach proposed in this paper to test for multidimensional sorting.

$$\mathbb{F}_{kl}^{ij} = \mathbb{E}_{\pi}\left[\frac{\partial \log \pi(X,Y)}{\partial A_{ij}} \frac{\partial \log \pi(X,Y)}{\partial A_{kl}}\right], \tag{22}$$

where we note that the lines of $\mathbb{F}$ are indexed by pairs of integers (ij), just as the columns of $\mathbb{F}$ are indexed by pairs kl. (The reason is that the parameter to be estimated, A_{ij}, is not a vector but a matrix.) Hence $\mathbb{F}$ is a "doubly-indexed matrix," which we shall denote using a "blackboard" font. Some basic formalism on doubly-indexed matrices is recalled in Appendix B, Section D.

Our next result expresses the asymptotic distribution of the estimators of A, S_X, and S_Y. It will be the main building block for testing the rank of the affinity matrix (and the number of sorting dimensions).

THEOREM 2. The following convergence holds in distribution for $n \to +\infty$:

$$n^{1/2}(\delta A,\ \delta S_X,\ \delta S_Y) \Rightarrow \mathcal{N}\left(0,\ \begin{pmatrix} \mathbb{F}^{-1} & 0 & 0 \\ 0 & \mathbb{K}_{\mathcal{X}\mathcal{X}} & \mathbb{K}_{\mathcal{X}\mathcal{Y}} \\ 0 & \mathbb{K}'_{\mathcal{X}\mathcal{Y}} & \mathbb{K}_{\mathcal{Y}\mathcal{Y}} \end{pmatrix}\right),$$

where $\mathbb{F}$ has been defined in (22), $\mathbb{K}_{XY}$ is defined by

$$(\mathbb{K}_{XY})_{ij}^{kl} = 1_{\{i=j,k=l\}}\mathrm{Cov}_{\pi}(X_i X_j,\ Y_k Y_l),$$

and we define similarly $\mathbb{K}_{XX}$ and $\mathbb{K}_{YY}$ by

$$(\mathbb{K}_{XX})_{ij}^{kl} = 1_{\{i=j,k=l\}}\mathrm{Cov}_{\pi}(X_i X_j,\ X_k X_l),$$

$$(\mathbb{K}_{YY})_{ij}^{kl} = 1_{\{i=j,k=l\}}\mathrm{Cov}_{\pi}(Y_i Y_j,\ Y_k Y_l).$$

Note that, as shown in lemma 5 in Appendix B, $\mathbb{F}$ can be evaluated numerically as the Hessian matrix of $\mathcal{W}_1$. Theorem 2 implies in particular that the asymptotic variance-covariance matrix of our estimator of A is the inverse of the Fisher information matrix. As a result, our estimator attains asymptotic statistical efficiency.

Now denoting $\hat{\Theta} = \hat{S}_X \hat{A} \hat{S}_Y$ the estimated counterpart of Θ, whose singular value decomposition is denoted $\hat{\Theta} = \hat{U}' \hat{\Lambda} \hat{V}$, we show in Appendix B that $\hat{\Theta}$ is asymptotically normal, and we give an expression for its asymptotic variance-covariance matrix in lemma 6. We use this asymptotic result to test the rank of the affinity matrix Λ. Testing the rank of a matrix is an important issue with a distinguished tradition in econometrics (see, e.g., Robin and Smith [2000] and references therein). Here, we use results from Kleibergen and Paap (2006). One wishes to test the null hypothesis H_0: the rank of the affinity matrix is equal to $p = 1, 2, \ldots,$

$d-1$. 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 $\hat{\Lambda}_1$, $\hat{U}'_{11}$, and $\hat{V}_{11}$ are $p\times p$ square matrices. Define

$$\hat{T}_p=(\hat{U}'_{22}\hat{U}_{22})^{-1/2}\hat{U}'_{22}\hat{\Lambda}_2\hat{V}_{22}(\hat{V}'_{22}\hat{V}_{22})^{-1/2},$$

$$\hat{A}_{p\perp}=\begin{pmatrix}\hat{U}'_{21}\\ \hat{U}'_{22}\end{pmatrix}(\hat{U}'_{22})^{-1}(\hat{U}'_{22}\hat{U}_{22})^{1/2},$$

$$\hat{B}_{p\perp}=(\hat{V}'_{22}\hat{V}_{22})^{1/2}\hat{V}_{22}^{-1}(\hat{V}_{21}\hat{V}_{22}),$$

so that our next result provides a test for the number of sorting dimensions.

THEOREM 3. Under the null hypothesis that the rank of the affinity matrix is p, the quantity $n^{1/2}\hat{T}_p$ is asymptotically normally distributed, and the expression of its variance-covariance matrix Ω_p is given in Appendix B, formula (B18). As a result, the test statistic

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

converges under the null hypothesis to a $\chi^2((d_x-p)(d_y-p))$ random variable.

VI. The Data

A. *The Data Set*

In this paper, we use the 1993–2002 waves of the DNB Household Survey (DHS) to estimate preferences in the marriage market. For a thorough description of the setup and the quality of these data, we refer the reader to Nyhus (1996). This data set is a representative panel of the Dutch population with respect to region, political preference, housing, income, degree of urbanization, and age of the head of the household, among others. The DHS data were collected via online terminal sessions, and each participating family was provided with a computer and a modem if necessary. The panel contains, on average, about 2,200 households in each wave.

These data include three main features that are particularly attractive for our purposes. First, within each household, all persons aged 16 or over were interviewed. This implies that the data contain detailed information not only about the head of the household but about all individuals in the household. In particular, the data identify "spouses" and "permanent partners" of the head of each household. This information reveals the nature of the relationship between the various individuals of each household and allows us to reconstruct "couples."

Second, this data set contains very detailed information about individuals. This rich set of information includes sociodemographic variables such as birth year and education, as well as variables about the anthropometry of respondents (height and weight), a self-assessed measure of health, and, above all, information about personality traits and risk attitude, which are included in the 1993–2002 waves.

Finally, as for most panel data, the DHS data suffer from attrition problems. The attrition of households is, on average, 25 percent each year (cf. Das and van Soest [1999], among others). To remedy attrition, refreshment samples were drawn each year such that, over the period 1993–2002, about 7,700 distinct households were interviewed at least once. Since the methodology implemented in this paper relies essentially on the availability of a cross section of households, attrition and its remedy is in fact an asset of these data as it allows us to have access to a rather large pool of potential couples.

Note that our methodology could be applied to other panel data sets (such as the German Socio-economic Panel [GSOEP], e.g.) that also include supplementary questionnaires enabling one to construct measures of personality traits and risks aversion together with sociodemographic and morphological variables. However, the main asset of the DNB data set is that it allows us to measure all relevant variables in a single wave, whereas in GSOEP, one would have to use the panel structure to match measures of BMI (from the 2008 and 2009 waves) and measures of personality traits (from the 2006 and 2007 waves).

B. Variables

Educational attainment is measured from the respondent's reported highest level of education achieved. The respondents could choose among 13 categories (seven in the later waves), ranging from primary to university education. The reduction to seven categories in the later waves implies that only three broad educational categories can be consistently constructed. We coded responses as follows: (1) lower (kindergarten, primary, elementary, secondary) education, (2) intermediate (secondary, preuniversity, vocational) education, and (3) higher (university) education.

The respondents were also asked about their height and weight. The answers to these questions allowed us to calculate the BMI of each respondent as the weight in kilograms divided by the square of the height measured in meters. The respondents were also asked to report their general health. The phrasing of the question was "How do you rate your general health condition on a scale from 1, excellent, to 5, poor?" We make use of the panel structure to deal partly with nonresponses on socioeconomic and health variables. When missing values for height, weight, education, year of birth, and so forth were encountered, values reported in adjacent years were imputed. We defined our measure of health by subtracting the answer to this question from 6.

In Appendix E, Section B, we recall the methodology of Nyhus and Webley (2001), which we followed in order to construct five factors of personality traits. These factors were labeled as follows:

- emotional stability: a high score indicates that the person is less likely to interpret ordinary situations as threatening and minor frustrations as hopelessly difficult;
- extraversion (outgoing): a high score indicates that the person is more likely to need attention and social interaction;
- conscientiousness (meticulous): a high score indicates that the person is more likely to be self-disciplined and to plan his or her actions;
- agreeableness (flexibility): a high score indicates that the person is more likely to be pleasant with others and go out of his way to help others;
- autonomy (tough-mindedness): a high score indicates that the person is more likely to be direct, rough, and dominant.

C. *Couples*

Our definition of a couple is a man and a woman living in the same household and reporting being either head of the household, spouse of the head, or a permanent partner (not married) of the head.[13] To construct our data set of couples, we first pool all the selected waves (1993–2002). We then keep only those respondents who report being head of the household, spouse of the head, or permanent (not married) partner of the head. This sample contains roughly 13,000 men and women and identifies about 7,700 unique households. We then split this sample and create two data sets, one containing women and one containing men. Each data set identifies about 6,500 different men and women. We then merge the men data set with the women data set using the household identifier. We identify 5,500 unique couples while roughly 1,250 men and 1,250 women remain unmatched.

[13] Note that using the subsample of legally married couples does not affect the three main results of our analysis mentioned in the abstract. These results are available from the authors on request.

Given the aim of our main analysis, we further restrict our sample to relatively newly formed couples. In the absence of information about when couples actually formed, following the literature (see, e.g., Chiappori et al. 2012), we select only couples whose wives are younger than 40 years old.

Table 1 reports the number of identified young couples and the number of young couples for whom we have complete information on the various dimensions. For nearly all couples we have information on both spouses' educational attainment. However, out of 2,897 couples, only 1,595 provide complete information on education, height, health, and BMI. We lose another 337 couples for whom personality traits are not fully observed. Another 100 couples are lost when attitude toward risk is additionally taken into consideration. Our working data set therefore contains 1,158 young couples.

Table 2 presents summary statistics for men and women. On average, in our sample, men are 3 years older than women,[14] are slightly more educated, are taller by 13 centimeters, have a BMI of 1 kilogram per square meter higher, are less conscientious (meticulous), are less extraverted but more emotionally stable, and are more risk averse. On average, in our sample, men and women have similar (good) health and a comparable degree of agreeableness and autonomy.

Oreffice and Quintana-Domeque (2010) estimate features of the (observed) matching function between men and women in the marriage market using the Panel Study of Income Dynamics data for the United States. Their strategy consists in regressing each attribute of men on all attributes of women and vice versa. This procedure can easily be replicated with our data in an attempt to compare features of the matching function in both data sets (US vs. Dutch marriage market). Interestingly enough, we find results very similar to those obtained by Oreffice and Quintana-Domeque. For instance, these authors find that an additional unit in the husband's BMI is associated with a 0.4 additional unit in the wife's BMI. Using our sample, our estimate is also significant and of similar magnitude even after controlling for personality traits (i.e., 0.25). Furthermore, they find that an additional inch in the husband's height is associated with an additional 0.12 inch in the wife's height. Here, too, our estimate is significant and of similar magnitude (0.15). Yet, Oreffice and Quintana-Domeque find that richer men (higher-educated men in our case) tend to be married with wives with a lower BMI (an increase of 10 percent in the husband's earnings is associated with a decrease of

[14] Note that the mean age at first marriage is relatively high in the Netherlands (30.1 and 32.8 for women and men, respectively; source: UN Economic Commission for Europe, 2010 Statistical Database) compared to the United States (26.1 and 28.2), which is reflected in the relatively high mean age of men and women in our sample of couples.

TABLE 1
Number of Identified Young Couples and Number of Young Couples with Complete Information for Various Subsets of Variables

	Observations
Identified couples	2,897
Couples with complete information on:	
Education	2,883
The above + health, height, and BMI[a]	1,595
The above + personality traits (big 5)	1,258
The above + measure of risk aversion	1,158

Source.—DNB; own calculation.
Note.—We have excluded all individuals taller than 210 centimeters or shorter than 145 centimeters and all individuals lighter than 40 kilograms; no one is heavier than 200 kilograms in our data. These exclusions represent less than 1 percent of the sample of adults in the source data. 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.

TABLE 2
Sample of Young Couples with Complete Information: Summary Statistics by Gender ($N = 1{,}158$)

	Husbands		Wives	
	Mean	Standard Error	Mean	Standard Error
Age	35.52	6.01	32.78	4.84
Educational level	2.01	.57	1.87	.57
Height	182.33	7.20	169.35	6.41
BMI	24.53	2.94	23.44	3.83
Health	3.21	.66	3.11	.69
Conscientiousness	−.25	.64	.01	.68
Extraversion	−.12	.69	.16	.60
Agreeableness	−.06	.65	−.04	.64
Emotional stability	.17	.57	−.19	.53
Autonomy	.00	.67	−.01	.69
Risk aversion	.06	.68	−.12	.88

0.21 points in his wife's BMI). In our sample, we find that higher-educated men (interpreting education as permanent income) are matched with women with a lower BMI; that is, a man with one additional level of education is matched with a woman whose BMI is 0.56 units lower.

VII. Empirical Results

We apply the saliency analysis, outlined in the previous section, on our sample of couples. The procedure requires us first to estimate the affinity matrix A. This is done by applying the technique presented in

Section III. The estimation results are reported in table 3. It is important to note that the estimates reported in the table are obtained using standardized attributes rather than the original ones. The main advantage of using standardized attributes is that the magnitude of the coefficients is directly comparable across attributes, allowing a direct interpretation in terms of comparative statics.

The estimates of the affinity matrix reveal four important and remarkable features.

1. *On-diagonal:* Education is the single most important attribute in the marriage market. The largest coefficient of the affinity matrix is indeed observed on the diagonal for education. This coefficient is more than twice as large as the second-largest coefficient obtained on the diagonal for the variable BMI. Loosely speaking, this means that increasing the education of both spouses by 1 standard deviation increases the couple's joint utility by 0.56 units. To achieve a similar increase in utility, the BMI of both spouses should be increased by 1.63 standard deviations each.

2. *Off-diagonal:* The table clearly indicates the importance of cross-gender interactions between the various attributes as many off-diagonal coefficients of the affinity matrix are significantly different from zero. This implies that important trade-offs take place between the various attributes. For instance, men's emotional stability interacts positively with women's conscientiousness (i.e., 0.21). Stated otherwise, this means that increasing the husband's emotional stability increases the joint utility of couples whose wives are relatively conscientious. Other examples are noticeable: men's autonomy interacts negatively with women's conscientiousness (i.e., −0.11) but positively with women's extraversion (i.e., 0.11). Conversely, men's agreeableness interacts positively with women's conscientiousness (i.e., 0.13) but negatively with women's extraversion (i.e., −0.14).

3. *Asymmetry:* The affinity matrix is not symmetric, indicating that preferences for attributes are not similar for men and women. For instance, increasing a wife's conscientiousness by one standard deviation increases the joint utility of couples with more agreeable men relatively more (significant coefficient of magnitude 0.13) while increasing the husband's conscientiousness by one standard deviation has the same impact on a couple's joint utility, indifferently of how agreeable his wife is.

4. *Personality traits:* Personality traits matter for preferences, not only directly (terms on the diagonal are significant for conscientiousness and risk aversion and of respective magnitude, 0.14 and 0.11) but mainly indirectly through their interactions with other attributes. For instance, the single most important interaction between observable attributes of men and women is found between the emotional stability of husbands and the conscientiousness of women (i.e., 0.21), a magnitude that matches

TABLE 3

Estimates of the Affinity Matrix: Quadratic Specification ($N = 1{,}158$)

	Wives									
Husbands	Education	Height	BMI	Health	Conscientiousness	Extraversion	Agreeableness	Emotional Stability	Autonomy	Risk Aversion
Education	.56*	.02	−.08	.02	−.04	−.01	−.03	−.04	.05	−.02
Height	.01	.18*	.04	−.01	−.04	.05	.02	.02	.02	.02
BMI	−.05	.05	.21*	.01	.06	.00	−.04	.04	−.01	−.01
Health	−.07	.00	−.06	.14*	−.04	.05	−.04	.04	.02	.00
Conscientiousness	−.06	−.03	.07	.00	.14*	.07	.04	.06	−.02	−.01
Extraversion	.01	−.02	.05	.02	−.06	.02	−.02	−.01	−.03	−.05
Agreeableness	.00	.01	−.08	.02	.13*	−.14*	.02	.11	−.09	−.04
Emotional stability	.03	.00	.12*	.04	.21*	.05	−.03	−.04	.08	.01
Autonomy	.02	.00	.00	.01	−.11*	.11*	−.04	.03	−.09	.01
Risk aversion	.00	.02	−.03	.02	.01	−.01	−.01	−.05	.05	.11*

* Significant at the 5 percent level.

with the direct effect of BMI. Also, personality traits interact not only with other personality traits but also with anthropometry. The emotional stability of men interacts positively with women's BMI (i.e., 0.12).

Using the estimated affinity matrix, we then proceed to the saliency analysis as introduced in the previous section. This enables us (i) to test whether sorting is unidimensional, that is, occurs on a single index, and (ii) to construct pairs of indices of mutual attractiveness for men and women.

We first test the dimensionality of the sorting in the marriage market. For $p = 1$, that is, testing against the null hypothesis that sorting occurs on a single index, we find that $n\hat{T}_1'\hat{\Omega}_1^{-1}\hat{T}_1 = 273.45$, which is significant at the 1 percent level. This implies that sorting in the marriage market does not occur on a single index as has been assumed in most of the previous literature. In fact, our test statistic never becomes insignificant. Even for $p = 9$, we have $n\hat{T}_9'\hat{\Omega}_9^{-1}\hat{T}_9 = 13.62$, which is still significant at the 1 percent level. This suggests that the affinity matrix has full rank and that sorting occurs on at least 10 observed indices. Our results therefore clearly highlight that sorting in the marriage market is multidimensional and individuals face important trade-offs between the attributes of their spouses.

Each pair of indices derived from saliency analysis explains a mutually exclusive part of the total observable matching utility of couples. The share explained by each of our 10 indices is reported in table 4. The table shows that the share of the first eight pairs of indices is significantly different from zero at the 1 percent level.

As for the principal component analysis, the labeling of each dimension is subjective and becomes increasingly difficult to interpret as one considers more dimensions. Table 5 therefore contains only the three pairs of indices explaining most of the joint utility. Together these three pairs of indices explain about 60 percent of the total matching utility. The first pair, indexed I1, explains about 28 percent of the joint utility. These indices load heavily (weight ≥ 0.5) on education, and the weights on education are of similar magnitude for men and women. This confirms that education plays the most important role in sorting in the marriage market. However, the second pair of indices, which explains another 17 percent of the joint utility, loads heavily on personality traits (i.e., emotional stability for men and conscientiousness for women). Personality traits play a strong role in the sorting in the marriage market. Interestingly enough, while conscientiousness matters only for the attractiveness of women, emotional stability matters only for the attractiveness of men. The third pair explains another 14 percent of the joint utility and loads on BMI and extraversion for women and agreeableness

TABLE 4

Share of Observed Joint Utility Explained

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
Share of joint utility explained	27.98***	16.60***	14.20***	10.07***	9.18***	8.51***	6.24***	4.14***	2.09	.99
Standard deviation of shares	2.25	1.55	1.59	1.54	1.68	1.48	2.58	1.91	2.26	1.03

Note.—The column headings I1–I10 indicate the 10 indices created by the singular value decomposition of the affinity matrix.

*** Significant at the 1 percent level.

TABLE 5
INDICES OF ATTRACTIVENESS

	I1		I2		I3	
ATTRIBUTES	Men	Women	Men	Women	Men	Women
Education	.97[a]	.96[a]	.15	.21	−.01	−.02
Height	.02	.04	.02	.05	−.39	−.27
BMI	−.16	−.19	.41	.51	−.35	−.56[a]
Health	−.08	.02	−.20	.02	−.04	−.02
Conscientiousness	−.17	−.14	.37	.82[a]	.04	.39
Extraversion	.01	−.02	−.08	−.02	−.17	−.59[a]
Agreeableness	−.02	−.05	.16	.00	.75[a]	.17
Emotional stability	−.01	−.09	.71[a]	.01	−.15	.22
Autonomy	.05	.08	−.30	.18	−.33	−.17
Risk aversion	.02	−.02	−.00	−.02	−.06	−.11
Cumulative share	27.98		44.58		58.78	

NOTE.—The column headings I1–I3 are the respective indices.
[a] Indicates coefficients that are larger than 0.5.

for men. This result corroborates Chiappori et al.'s (2012) finding that BMI is important for sorting in the marriage market.

VIII. Summary and Discussion

This paper has introduced a novel technique to test for the dimensionality of sorting in the marriage market and derive indices of mutual attractiveness, namely, saliency analysis. This technique is grounded in the structural equilibrium model of Choo and Siow (2006), which we have extended to the continuous case in this paper. Indices of mutual attractiveness derived in saliency analysis, in contrast to canonical correlation, for instance, have a structural interpretation and are therefore informative about agents' preferences.

Saliency analysis has been performed on a data set of Dutch households containing information about education, height, BMI, health, attitude toward risk, and five personality traits of both spouses. The empirical results of this paper reveal two important features of the marriage market. First, our results clearly show that sorting occurs on multiple indices rather than just on a single one, as assumed in most of the current literature. This implies that individuals face important trade-offs between the attributes of their potential spouse. For instance, in the data set we studied, more conscientious men prefer more conscientious women (0.14), but more autonomous men prefer less conscientious women (−0.11). Hence, women face a trade-off between being attractive for more conscientious men and being attractive for more autonomous men. Similarly, more conscientious women prefer more agreeable men (0.13), but more extraverted women prefer less agreeable men (−0.14). Men therefore face a trade-off between being attractive for more conscientious women and being attractive for more extraverted women.

Second, personality traits and attitude toward risk matter for the sorting of spouses in the marriage market. In fact, although education explains the largest share (28 percent) of the observable joint utility of spouses, personality traits explain a rather large share too (17 percent). Interestingly enough, different traits matter differently for men and women. For instance, women find emotionally stable men more attractive. Yet, men prefer conscientious women but are indifferent about the emotional stability of women.

The analysis presented in this paper opens up interesting possibilities for further research. In particular, our analysis could be applied to other markets besides the marriage one, such as the market for chief executive officers. A recent literature led by Bertrand and Schoar (2003), Falato, Li, and Milbourn (2012), and Custodio, Ferreira, and Matos (2013) acknowledges the multidimensionality of CEOs' talent but assumes that sorting occurs on a single index. Our setting can then be used to extend the seminal contributions of Gabaix and Landier (2008) and Terviö (2008), who calibrate a single-dimensional multiplicative sorting model in order to explain CEO compensation. An important difference in the CEO compensation literature is that transfers (i.e., salaries) are typically observed, in contrast to the case of the marriage market considered in the present paper. The observation of the transfers has interesting consequences for identification. Assume that x is a CEO's vector of characteristics (say track record, education, political inclinations, cultural affinities) and y is a vector of a firm's characteristics. Let $\alpha(x, y)$ be the nonpecuniary utility of CEO x working with firm y, and let $\gamma(x, y)$ be the productivity (in monetary units) of CEO x if hired by firm y. In the case in which transfers are unobserved, only the joint utility $\Phi = \alpha + \gamma$ is identified. However, in the case in which transfers are observed, it is possible to identify α and γ separately. Hence when CEO compensation data are available, the results of the present paper can be easily extended to identify simultaneously the CEO's productivity and his or her nonpecuniary utility for working with a given firm.

Finally, we observe that the Poisson process approach that appears in the framework of this paper may provide the "missing link" between search models and matching with unobserved heterogeneity. Indeed, Poisson processes are central to search models, and the fact that they also play an important role in our model suggests that they may provide an interesting connection. The key difference, of course, comes from the fact that in search models, agents are faced with an optimal stopping problem: agents cannot know what their opportunities will be in advance, and they cannot retain offers, while in our framework they are fully aware of all their opportunities from the start. While we briefly elaborate on the formal connection in Appendix A, we leave a full exploration of the matter for future work.

Appendix A

Continuous Logit Formalism

In this paragraph, we expound the main ideas of Cosslett (1988) and Dagsvik (1994), who show how to obtain a continuous version of the multinomial logit model. Assume that $\{(y_k^m, \varepsilon_k^m), k \in \mathbb{N}\}$ are the points of a Poisson point process on $\mathcal{Y} \times \mathbb{R}$ of intensity $dy \times e^{-\varepsilon} d\varepsilon$. We recall that this implies that for S a subset of $\mathcal{Y} \times \mathbb{R}$, the probability that man m has no acquaintance in set S is $\exp(-\int_S e^{-\varepsilon} dy d\varepsilon)$. From (2), man m chooses woman k among his acquaintances such that his utility is maximized; that is, man m solves

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

Letting Z be the value of this maximum, one has for any $c \in \mathbb{R}$

$$\Pr(Z \leq c) = \Pr(U(x, y_k^m) + \varepsilon_k^m \leq c \;\forall\, k),$$

which is exactly the probability that the Poisson point process (y_k, ε_k^m) has no point in $\{(y, \varepsilon) : U(x, y) + \varepsilon > c\}$; thus

$$\begin{aligned}
\log \Pr(Z \leq c) &= -\int\int_{\mathcal{Y}\times\mathbb{R}} 1(U(x, y) + \varepsilon > c)\, dy e^{-\varepsilon} d\varepsilon \\
&= -\int_{\mathcal{Y}}\int_{c-U(x,y)} e^{-\varepsilon} d\varepsilon dy \\
&= -\int_{\mathcal{Y}} e^{-c+U(x,y)} dy \\
&= -\exp\left[-c + \log\int_{\mathcal{Y}} \exp U(x, y)\, dy\right].
\end{aligned}$$

Hence Z is a $(\log \int_{\mathcal{Y}} \exp U(x, y) dy, 1)$ Gumbel. In particular,

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

and the choice probabilities are given by their density with respect to the Lebesgue measure

$$\pi(y|x) = \exp[U(x, y)] / \left[\int_{\mathcal{Y}} \exp U(x, y')\, dy'\right].$$

The same logic also implies that $\{\varepsilon_k : k \in \mathbb{N}\}$ has a Gumbel distribution. Indeed, the probability that this Poisson point process has no element in the set $\{\varepsilon : \varepsilon > c\}$ is equal to

$$\exp\left(-\int_c^{+\infty} e^{-\varepsilon}d\varepsilon\right) = \exp[-\exp(-c)],$$

which is equivalent to saying that $\Pr(\max_{k\in\mathbb{N}}\varepsilon_k \leq c) = \exp[-\exp(-c)]$. Finally, note that a similar argument would show that m has almost surely an infinite, though countable, number of acquaintances, as announced.

Note that an interesting connection remains to be explored with the search literature (Shimer and Smith 2000; Atakan 2006). Assume that each man m draws a Poisson sample of acquaintances (y_k^m, ε_k^m), where y_k^m is the type of partner of index k, and ε_k^m is now the time at which this acquaintance is met. Assume that it has been agreed that if x matches with y, x will receive utility $U(x, y)$ out of the joint utility $\Phi(x, y)$. In the spirit of Atakan (2006), assume that unmatched agents pay a utility cost equal to σ per unit of time while unmatched, such that if x matches with y at time ε, his utility is $U(x, y) - \sigma\varepsilon$. If agents could perfectly foresee their opportunities (i.e., know the full sample (y_k^m, ε_k^m) in advance), they would choose opportunity k so as to maximize the quantity $U(x^m, y_k^m) - \sigma\varepsilon_k^m$ exactly as in the present paper. The difference, of course, comes from the fact that in search models, agents are faced with an optimal stopping problem: agents cannot know what their opportunities will be in advance, and they cannot retain offers. At each time t they know only what opportunities have already been received up to time t, and they do not know about the set of k's such that $\varepsilon_k^m > t$. This is an optimal stopping problem with a Poisson process, well studied in probability theory and operations research following seminal work by Elfving (1967). The basic idea is as follows: there exists a function $\psi : \mathbb{R} \to \mathbb{R}$ such that the partner chosen by m is the first partner (in terms of meeting time) such that $U(x^m, y_k^m)$ exceeds $\psi(\varepsilon_k^m)$. The function ψ can be characterized as a solution to an ordinary differential equation and, in some cases, can be expressed analytically.

Appendix B

Proofs

A. *Proof of Theorem 1*

Part i: The first part of the argument extends Galichon and Salanié (2010) to the continuous case; the argument is decomposed in four steps, which are now briefly commented on. In step 1, we shall show that the expression of the social welfare is given by

$$\begin{aligned} &\min_{U,V} \int_{\mathcal{X}} G_x(U(x,\cdot))f(x)dx + \int_{\mathcal{Y}} H_y(V(\cdot,\ y))g(y)dy \\ &\text{subject to } U(x,y) + V(x,y) \geq \Phi(x,y), \end{aligned} \tag{B1}$$

where $U(x, y)$ (respectively, $V(x, y)$) is the share of the systematic joint utility going to man x (woman y), and $G_x(U)$ ($H_y(V)$) is the ex ante indirect utility of a man of type x (a woman of type y), namely,

$$\begin{aligned} G_x(U(x,\cdot)) &= \mathbb{E}\left[\max_k \left\{ U(x, y_k^m) + \frac{\sigma}{2}\varepsilon_k^m \right\}\right], \\ H_y(V(\cdot, y)) &= \mathbb{E}\left[\max_l \left\{ U(x_l^w, y) + \frac{\sigma}{2}\varepsilon_l^w \right\}\right]. \end{aligned} \tag{B2}$$

Welfare expression (B1) has a straightforward interpretation in terms of equilibrium. The constraint $U + V \geq \Phi$ is a stability condition, and the minimization of the sum of the individual ex ante indirect utility function expresses the absence of rents.

In step 2, we shall express the dual of variational problem (B1) as

$$\mathcal{W} = \sup_{\pi \in \mathcal{M}(P,Q)} \int \Phi d\pi - \mathcal{I}(\pi),$$

where

$$\begin{aligned} \mathcal{I}(\pi) = \sup_U &\left[\int_{\mathcal{X}\times\mathcal{Y}} U(x,y)\, d\pi(x,y) - \int_{\mathcal{X}} G_x(U(x,\cdot))\, dP(x) \right] \\ &+ \sup_V \left[\int_{\mathcal{X}\times\mathcal{Y}} V(x,y)\, d\pi(x,y) - \int_{\mathcal{Y}} H_y(V(\cdot,y))\, dQ(y) \right]. \end{aligned}$$

In step 3, we shall show that under the distributional assumptions made on the heterogeneities, the expression of $\mathcal{I}$ is given by

$$\mathcal{I}(\pi) = \sigma \int\int_{\mathcal{X}\times\mathcal{Y}} \log \frac{\pi(x,y)}{\sqrt{f(x)g(y)}} \pi(x,y)\, dxdy.$$

In step 4, we shall show that as a result, the social welfare, which is the value of variational problem (B1), can be expressed up to irrelevant constants as

$$\max_{\pi \in \mathcal{M}(P,Q)} \int\int_{\mathcal{X}\times\mathcal{Y}} \Phi(x,y)\pi(x,y)\, dxdy - \sigma \int\int_{\mathcal{X}\times\mathcal{Y}} \log \pi(x,y)\pi(x,y)\, dxdy, \tag{B3}$$

which will establish part i.

Step 1: Introduce $\varepsilon_m(\cdot)$ a stochastic process on $\mathcal{Y}$ defined by

$$\varepsilon_m(y) = \frac{\sigma}{2}\max_k \{\varepsilon_k^m : y_k = y\}$$

if the set $\{k : y_k = y\}$ is nonempty, and $\varepsilon_m(y) = -\infty$ otherwise. Similarly, introduce $\eta_w(x)$ a stochastic process on $\mathcal{X}$ defined by

$$\eta_w(x) = \frac{\sigma}{2}\max_l \{\eta_l^w : x_l = x\}$$

if the set $\{l : x_l = x\}$ is nonempty, and $\eta_w(x) = -\infty$ otherwise. By the results of Shapley and Shubik (1972), extended to the continuous case by Gretsky, Ostroy, and Zame (1992), the equilibrium matching solves the dual transportation problem, which expresses the social welfare

$$\mathcal{W} = \inf_{u_m + v_w \geq \Phi(x_m, y_m) + \varepsilon_m(y) + \eta_w(x)} \int u_m dm + \int v_w dw. \tag{B4}$$

Now, the constraint can be rewritten as

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

where U and V have been defined as

$$U(x, y) = \inf_m [u_m - \varepsilon_m(y)],$$
$$V(x, y) = \inf_w [v_w - \eta_w(x)],$$

which implies that u_m and v_w can be expressed in $U(x, y)$ and $V(x, y)$ by

$$\begin{aligned} u_m &= \sup_{y \in \mathcal{Y}} [U(x, y) + \varepsilon_m(y)], \\ v_w &= \sup_{x \in \mathcal{X}} [V(x, y) + \eta_w(x)]. \end{aligned} \tag{B5}$$

Therefore, replacing u_m and v_w by their expression in U and V, (B4) is rewritten as (B1), with G_x and H_y given by (B2).

Step 2: Rewrite (B1) as a saddlepoint problem

$$\begin{aligned} \mathcal{W} = \inf_{U,V} \sup_{\pi} \Bigg[\iint_{\mathcal{X} \times \mathcal{Y}} \Phi d\pi + \int_{\mathcal{X}} G(U(x, \cdot)) dP(x) \\ - \iint_{\mathcal{X} \times \mathcal{Y}} U d\pi + \int_{\mathcal{Y}} H(V(\cdot, y)) dQ(y) - \iint_{\mathcal{X} \times \mathcal{Y}} V d\pi \Bigg] \end{aligned}$$

or, in other words,

$$\mathcal{W} = \sup_{\pi} \int \Phi d\pi - \mathcal{I}(\pi),$$

where

$$\begin{aligned} \mathcal{I}(\pi) = \sup_U \left[\iint_{\mathcal{X} \times \mathcal{Y}} U d\pi - \int_{\mathcal{X}} G_x(U(x, \cdot)) dP(x) \right] \\ + \sup_V \left[\iint_{\mathcal{X} \times \mathcal{Y}} V d\pi - \int_{\mathcal{Y}} H_y(V(\cdot, y)) dQ(y) \right]. \end{aligned}$$

Step 3: From the derivation in Appendix A, we get that

$$G_x(U(x, \cdot)) = \frac{\sigma}{2} \log \int_{\mathcal{Y}} \exp \frac{U(x, y)}{\sigma/2} dy$$

and

$$H_y(V(\cdot, y)) = \frac{\sigma}{2}\log\int_{\mathcal{X}} \exp\frac{U(x,y)}{\sigma/2}\,dx.$$

Now, in order to get an expression for $\mathcal{I}(\pi)$, it remains to compute

$$\sup_{U(x,y)} \int\int_{\mathcal{X}\times\mathcal{Y}} U(x,y)\pi(x,y)\,dxdy - \int G_x(U(x,\cdot))f(x)\,dx \tag{B6}$$

and the similar expression on the other side of the market.

By first-order conditions,

$$\pi(x,y) = \frac{f(x)\exp\dfrac{U(x,y)}{\sigma/2}}{\displaystyle\int_{\mathcal{Y}} \exp\frac{U(x,y)}{\sigma/2}\,dy},$$

which implies that the value of the problem is infinite unless $\int \pi(x,y)\,dy = f(x)$, in which case it is

$$(\sigma/2)\int\int_{\mathcal{X}\times\mathcal{Y}} \pi(x,y)\log\frac{\pi(x,y)}{f(x)}\,dxdy,$$

which is the value of (B6). A symmetric expression is obtained for the other side of the market. Finally, $\mathcal{I}(\pi)$ obtains as

$$\mathcal{I}(\pi) = \sigma\int\int_{\mathcal{X}\times\mathcal{Y}} \log\frac{\pi(x,y)}{\sqrt{f(x)g(y)}}\,\pi(x,y)\,dxdy$$

if $\pi \in \mathcal{M}(P, Q)$, while $\mathcal{I}(\pi) = +\infty$ otherwise.

Step 4: One has

$$\begin{aligned}\mathcal{I}(\pi) = \sigma&\int\int_{\mathcal{X}\times\mathcal{Y}} \log\pi(x,y)\pi(x,y)\,dxdy\\ &- (\sigma/2)\int_{\mathcal{X}} \log f(x)f(x)\,dx - (\sigma/2)\int_{\mathcal{Y}} \log g(x)g(x)\,dx.\end{aligned}$$

The last two terms do not depend on the particular matching $\pi \in \mathcal{M}(P, Q)$ and thus are irrelevant in the expression of the social welfare, which establishes (B3) and part i.

Part ii: Letting

$$a(x) = \frac{-\sigma}{2}\log\frac{f(x)}{\displaystyle\int_{\mathcal{Y}} \exp U(x,y)\,dy},$$

$$b(y) = \frac{-\sigma}{2}\log\frac{g(y)}{\displaystyle\int_{\mathcal{X}} \exp V(x,y)\,dx},$$

one has

$$\log \pi(x, y) = \frac{U(x, y) - a(x)}{\sigma/2},$$
$$\log \pi(x, y) = \frac{V(x, y) - b(y)}{\sigma/2},$$

and by summation,

$$\pi(x, y) = \exp\left[\frac{\Phi(x, y) - a(x) - b(y)}{\sigma}\right].$$

Part iii: One has

$$U(x, y) = \frac{\sigma \log \pi(x, y)}{2} + a(x) = \frac{\Phi(x, y) + a(x) - b(y)}{2},$$

and similarly,

$$V(x, y) = \frac{\Phi(x, y) - a(x) + b(y)}{2}.$$

By (B5), one sees that if man m of type x marries a woman of type x, he gets utility

$$u_m = \sup_{y' \in \mathcal{Y}} [U(x, y') + \varepsilon_m(y')] = U(x, y) + \varepsilon_m(y).$$

QED

B. Useful Lemmas

We state several useful lemmas that are useful in Sections IV and V and in the proof of theorem 2. First, we need a formula that expresses the affinity matrix of the rescaled attributes as a function of the affinity matrix between X and Y. This is given in the following lemma.

LEMMA 1. For M and N two invertible matrices, one has

$$A^{MX,NY} = (M')^{-1} A^{XY} N^{-1}. \tag{B7}$$

This result should be compared with the expression of the cross-covariance matrix between MX and NY, namely, $\Sigma_{MX,NY} = M \Sigma_{XY} N'$. A quick dimensionality check is coherent, as the unit of A^{XY} is the inverse of the product of the units of X and Y, while the unit of Σ_{XY} is the product of the units of X and Y.

Proof of lemma 1. Recall that every affinity matrix A^{XY} is characterized by the fact that

$$\frac{\partial \mathcal{W}^{P,Q}}{\partial A_{ij}} (A^{XY}) = \Sigma_{XY}^{ij}. \tag{B8}$$

Let P_M (respectively, Q_N) be the distribution of MX (NY). We therefore have

$$\begin{aligned}\frac{\partial \mathcal{W}^{P_M,Q_N}}{\partial A_{ij}}(A^{MX,NY}) &= \Sigma^{ij}_{MX,NY} = M\Sigma^{ij}_{XY}N' \\ &= M\frac{\partial \mathcal{W}^{P,Q}}{\partial A_{ij}}(A^{XY})N',\end{aligned} \tag{B9}$$

where the second equality follows by definition and the third by using (B8). A simple calculation shows that

$$\mathcal{W}^{P_M,Q_N}(A^{MX,NY}) = \mathcal{W}^{P,Q}(M'A^{MX,NY}N).$$

Taking the derivative with respect to A yields

$$\frac{\partial \mathcal{W}^{P_M,Q_N}}{\partial A}(A^{MX,NY}) = M\frac{\partial \mathcal{W}^{P,Q}}{\partial A}(M'A^{MX,NY}N)N'. \tag{B10}$$

And, by comparing (B9) and (B10), one gets

$$\frac{\partial \mathcal{W}^{P,Q}}{\partial A}(M'A^{MX,NY}N) = \frac{\partial \mathcal{W}^{P,Q}}{\partial A}(A^{XY}).$$

From the strict convexity of $\mathcal{W}^{P,Q}$, we therefore have $M'A^{MX,NY}N = A^{XY}$, and given that M and N are invertible, it follows that

$$A^{MX,NY} = (M')^{-1}A^{XY}N^{-1}.$$

QED

As a consequence of lemma 1, we are able to state that the results of saliency analysis are invariant with respect to a (linear) change in the measurement units.

Lemma 2. For ζ and ξ_j two vectors of positive scalars, let

$$\hat{X}_i = \zeta_i X_i \quad \text{and} \quad \hat{Y}_j = \xi_j Y_j$$

be the values of partners' attributes measured under different measurement units. Then the outcome of saliency analysis under the new measurement units coincides with the outcome under the former.

Proof. Saliency analysis consists in determining the singular value decomposition of $\Theta = \sigma_X A^{X,Y}\sigma_Y$ under the old units and of $\hat{\Theta} = \sigma_{\hat{X}}A^{\hat{X},\hat{Y}}\sigma_{\hat{Y}}$ under the new units. Letting $D_\zeta = \operatorname{diag}(\zeta_i)$ and $D_\xi = \operatorname{diag}(\xi_j)$, one has

$$A^{\hat{X},\hat{Y}} = D_\zeta^{-1}A^{X,Y}D_\xi^{-1},\ \sigma_{\hat{X}} = \sigma_X D_\zeta,\ \text{and}\ \sigma_{\hat{Y}} = D_\xi\sigma_Y;$$

thus $\hat{\Theta} = \Theta$. QED

The next lemma shows that $\mathcal{W}_1(A)$ is strictly convex.

Lemma 3. The map $A \to \mathcal{W}_1(A)$ is strictly convex.

Proof. Consider two matrices A and $\tilde{A}$. Let π be the matching associated with $\Phi(x,y) = x'Ay$ and $\tilde{\pi}$ be the matching associated with $\Phi(x,y) = x'\tilde{A}y$ (uniqueness of π and $\tilde{\pi}$ follows from the uniqueness of the solution to the Schrödinger problem; see Rüschendorf and Thomsen 1993, theorem 3). Then convexity of $\mathcal{W}_1$ implies

$$\mathcal{W}_1(\tilde{A}) \geq \mathcal{W}_1(A) + \langle \nabla \mathcal{W}_1(A),\ \tilde{A} - A \rangle, \tag{B11}$$

where, by the envelope theorem, $\nabla \mathcal{W}_1(A) = \mathbb{E}_\pi[XY']$. In order to show strict convexity, we need to show that equality in (B11) implies $A = A'$. Assume that (B11) holds as an equality. One has

$$\begin{aligned}\mathcal{W}_1(\tilde{A}) &= \mathcal{W}_1(A) + \langle \nabla \mathcal{W}_1(A), \tilde{A} - A \rangle \\ &= \mathbb{E}_\pi[X'AY] - \mathbb{E}_\pi[\ln \pi(X,Y)] + \mathbb{E}_\pi[X\tilde{A}Y'] - \mathbb{E}_\pi[X'AY] \\ &= \mathbb{E}_\pi[X\tilde{A}Y'] - \mathbb{E}_\pi[\ln \pi(X,Y)].\end{aligned}$$

This implies that π is optimal for the matching problem associated with $\Phi(x,y) = x'\tilde{A}y$. Again by the uniqueness of the solution to the Schrödinger problem mentioned above, it follows that $\pi = \tilde{\pi}$ and hence that

$$A = \partial^2 \ln \pi(x,y)/\partial x \partial y = \partial^2 \ln \tilde{\pi}(x,y)/\partial x \partial y = \tilde{A}.$$

QED

The following lemma allows us to characterize the conditional expectations of the gradient of the log likelihood.

LEMMA 4. Let $\pi_A \in \mathcal{M}(P,Q)$ be the equilibrium matching computed for joint utility function Φ_A. Then

$$\mathbb{E}\left[\frac{\partial \log \pi_A}{\partial A_{ij}} X = x\right] = \mathbb{E}\left[\frac{\partial \log \pi_A}{\partial A_{ij}} Y = y\right] = 0, \tag{B12}$$

and

$$\begin{aligned}\sigma \mathbb{E}_\pi\left[\frac{\partial \log \pi_A(X,Y)}{\partial A_{ij}} \frac{\partial \log \pi_A(X,Y)}{\partial A_{kl}}\right] &= \mathbb{E}_\pi\left[\frac{\partial \log \pi_A(X,Y)}{\partial A_{ij}} x_k y_l\right] \\ &= \mathbb{E}_\pi\left[x_i y_j \frac{\partial \log \pi_A(X,Y)}{\partial A_{kl}}\right].\end{aligned} \tag{B13}$$

Proof of lemma 4. It follows from equation (7) that

$$\sigma \frac{\partial \log \pi}{\partial A_{ij}}(x,y) = x_i y_j - \frac{\partial a}{\partial A_{ij}}(x) - \frac{\partial b}{\partial A_{ij}}(y).$$

But by differentiation of $\int_y \pi(x,y)\,dy = f(x)$ with respect to A_{ij}, one gets

$$\int_y \frac{\partial \log}{\partial A_{ij}} \pi(x,y) \pi(x,y)\,dy = 0;$$

thus

$$\mathbb{E}_{\pi}\left[\frac{\partial \log}{\partial A_{ij}}\pi(X, Y)|X = x\right] = 0,$$

which proves (B12). Then (B13) follows directly. QED

The final lemma in this section shows that the Hessian of $\mathcal{W}$ coincides with the Fisher information matrix $\mathbb{F}$.

LEMMA 5. The Hessian of $\mathcal{W}_1$ is given by

$$\frac{\partial^2 \mathcal{W}_1}{\partial A_{ij}\partial A_{kl}} = \mathbb{F}^{ij}_{kl},$$

where the expression of $\mathbb{F}$ is given by (22).

Proof of lemma 5. By the envelope theorem,

$$\frac{\partial \mathcal{W}_1}{\partial A_{ij}} = \int x_i y_j \pi_A(x, y)\, dxdy.$$

Thus,

$$\begin{aligned}\frac{\partial^2 \mathcal{W}_1}{\partial A_{ij}\partial A_{kl}} &= \int x_i y_j \frac{\partial \log \pi_A(x, y)}{\partial A_{kl}} \pi_A(x, y)\, dxdy \\ &= \mathbb{F}^{ij}_{kl},\end{aligned}$$

where the second equality follows from lemma 4. QED

C. *Proof of Theorem 2*

The proof of theorem 2 relies on the auxiliary results derived in the previous paragraph. In the sequel, we assume $\sigma = 1$; by positive homogeneity, this is without loss of generality.

Let

$$\hat{\pi}(x, y) = \frac{1}{n}\sum_{k=1}^{n} \delta(x - X_k)\delta(y - Y_k)$$

be the distribution of the empirical sample under observation, and π_A is the equilibrium matching computed for matching utility function Φ_A (we shall drop the subscript A when there is no ambiguity). Recall that the (population) affinity matrix A and its sample estimator $\hat{A}$ are respectively characterized by

$$\frac{\partial \mathcal{W}_1(A)}{\partial A_{ij}} = \Sigma^{ij}_{XY} \quad \text{and} \quad \frac{\partial \mathcal{W}_1(\hat{A})}{\partial A_{ij}} = \hat{\Sigma}^{ij}_{XY}.$$

By the delta method, we get

$$(\mathbb{F} \cdot \delta A)^{ij} = \int \frac{\partial \log \pi_A}{\partial A_{ij}}(\hat{\pi} - \pi)\, dxdy + o_D(n^{-1/2}), \tag{B14}$$

where $\mathbb{F}$ is the Hessian of $\mathcal{W}_1$ at A, whose expression is

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

where $\pi \in \mathcal{M}(P, Q)$ is the equilibrium matching computed for the joint utility function Φ_A. Further,

$$(\delta S_X)^{ij} = 1_{\{i=j\}} \int x_i x_j (\hat{\pi} - \pi) dxdy + o_D(n^{-1/2}),$$
$$(\delta S_Y)^{kl} = 1_{\{k=l\}} \int y_i y_j d\pi (\hat{\pi} - \pi) dxdy + o_D(n^{-1/2}).$$

Hence

$$\mathbb{E}[(\mathbb{F} \cdot \delta A)_{ij} (\delta S_X)_{kl}] = \mathrm{Cov}\left(\frac{\partial \log \pi}{\partial A_{ij}}, X_k X_l\right) 1_{\{k=l\}} = 0,$$

where we have used (B12), and, similarly, $\mathbb{E}[(\delta A)_{ij} (\delta S_Y)_{kl}] = 0$. This proves the asymptotic independence between δA and $(\delta S_X, \delta S_Y)$. The conclusion follows by noting that the asymptotic variance-covariance matrix of δA is $\mathbb{F}^{-1}$ and that of $(\delta S_X, \delta S_Y)$ is

$$\begin{pmatrix} \mathbb{K}_{XX} & \mathbb{K}_{XY} \\ \mathbb{K}'_{XY} & \mathbb{K}_{YY} \end{pmatrix}.$$

QED

D. Proof of Theorem 3

In order to give asymptotic distributions of matrix estimators, it is convenient to represent matrices as vectors, an operation that is called *vectorization* in matrix algebra. Linear operators acting on these vectorized matrices will therefore be called *doubly-indexed matrices*, for which we shall use the blackboard notation to distinguish them from simply-indexed matrices. If $\mathbb{R}$ is a doubly-indexed matrix, its general term will be denoted $\mathbb{R}^{ij}_{kl}$, where ij indexes the lines and kl indexes the columns of $\mathbb{R}$. Then $\mathbb{R} \cdot M$ will denote the (simple) matrix N such that $N^{ij} = \sum_{kl} \mathbb{R}^{ij}_{kl} M^{kl}$. We recall the definition of the Kronecker product: for two matrices A and B, $A \otimes B$ is the doubly-indexed matrix $\mathbb{R}$ such that

$$\mathbb{R}^{ij}_{kl} = A_{ik} B_{jl}.$$

LEMMA 6. The following convergence holds in distribution for $n \to +\infty$:

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

where

$$\mathbb{V} = \mathbb{T}_{XY} \mathbb{F}^{-1} \mathbb{T}'_{XY} + \mathbb{T}_X \mathbb{K}_{XX} \mathbb{T}'_X + \mathbb{T}_Y \mathbb{K}_{YY} \mathbb{T}'_Y + \mathbb{T}_X \mathbb{K}_{XY} \mathbb{T}'_Y + \mathbb{T}_Y \mathbb{K}'_{XY} \mathbb{T}_X.$$

Proof of lemma 6. As

$$\delta S_X^{1/2} = (I \otimes S_X^{1/2} + S_X^{1/2} \otimes I)^{-1} \delta S_X,$$

one has

$$\begin{aligned} \delta\Theta &= (S_Y^{1/2} \otimes S_X^{1/2}) \delta A + (S_Y^{1/2} A' \otimes I) \delta S_X^{1/2} + (I \otimes S_X^{1/2} A) \delta S_Y^{1/2} \\ &= \mathbb{T}_{XY} \delta A + \mathbb{T}_X \delta S_X + \mathbb{T}_Y \delta S_Y, \end{aligned}$$

where

$$\mathbb{T}_X = (S_Y^{1/2} A' \otimes I)(S_X^{1/2} \otimes I + I \otimes S_X^{1/2})^{-1}, \tag{B15}$$

$$\mathbb{T}_{XY} = S_Y^{1/2} \otimes S_X^{1/2}, \tag{B16}$$

and

$$\mathbb{T}_Y = (I \otimes S_X^{1/2} A)(S_Y^{1/2} \otimes I + I \otimes S_Y^{1/2})^{-1}. \tag{B17}$$

QED

The proof of theorem 3 follows as an easy consequence.

Proof of theorem 3. Let

$$\Omega_p = (B_{p\perp} \otimes A'_{p\perp}) \mathbb{V} (B_{p\perp} \otimes A'_{p\perp})'. \tag{B18}$$

By Kleibergen and Paap's theorem 1, the convergence

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

holds for $n \to +\infty$, and theorem 3 follows. QED

Appendix C

Computation

Let a and b be the solutions of equation (7), and introduce

$$\tilde{a}(x) = \exp[-a(x)/\sigma] \quad \text{and} \quad \tilde{b}(y) = \exp[-b(y)/\sigma],$$

so equation (7) can be rewritten as

$$\pi(x, y) = \tilde{a}(x) \tilde{b}(y) K(x, y), \tag{C1}$$

where $K(x, \ y) = \exp[\Phi(x, \ y)/\sigma]$, and the system of equations formed by the constraints on the marginals can be rewritten as

$$\tilde{a}(x) = f(x) \left[\int_y \tilde{b}(y) K(x, y) dy \right]^{-1} \tag{C2}$$

and

$$\tilde{b}(y) = g(y)\left[\int_{\mathcal{X}} \tilde{a}(x)K(x,y)\,dx\right]^{-1}. \tag{C3}$$

Note that by (C3), $\tilde{b}$ can be expressed as a function of $\tilde{a}$. Then $\tilde{a}$ is rewritten as a fixed-point equation $\tilde{a} = F(\tilde{a})$, where F is given by

$$F(\tilde{a})(x) = f(x)\left\{\int_{\mathcal{Y}}\left[g(y)\int_{\mathcal{X}} \tilde{a}(x')K(x',y)\,dx'\right]^{-1} K(x,y)\,dy\right\}^{-1}.$$

The iterative projection fitting procedure consists in starting with some proper choice of $\tilde{a}_0(x)$ that ensures integrability of $x \to \tilde{a}(x)K(x,y)$ and iteratively applying $\tilde{a}_{k+1} = F(\tilde{a}_k)$. Details and proof of convergence are provided in Rüschendorf (1995); convergence is very quick in practice.

Appendix D

Incorporating Singles

Throughout this appendix, the symbol $\varnothing$ stands for singlehood; this enlarges the sets of marital choices of men and women, which we denote $\mathcal{X}_0 = \mathcal{X} \cup \{\varnothing\}$ and $\mathcal{Y}_0 = \mathcal{Y} \cup \{\varnothing\}$. Let $\bar{f}(x)$ be the density of the mass of men of type x, $f_0(x)$ be the density of the mass of single men of type x, and, as in the rest of the paper, $f(x)$ be the density of the mass of matched men of type x, so that $\bar{f}(x) = f_0(x) + f(x)$. Introduce similar notation on the other side of the market: $\bar{g}(y) = g_0(y) + g(y)$, and note that the total mass of men and women no longer needs to coincide; that is, in general one has

$$\int_{\mathcal{X}} \bar{f}(x)\,dx \neq \int_{\mathcal{Y}} \bar{g}(y)\,dy.$$

The set of acquaintances of man m is now expanded to include singlehood: $\{(y_k^m, \varepsilon_k^m),\ k \in \mathbb{N}\}$ are now the points of a Poisson process on $\mathcal{Y}_0 \times \mathbb{R}$ of intensity $\lambda_0 \times e^{-\varepsilon}d\varepsilon$, where for $B \subseteq \mathcal{Y}_0$,

$$\lambda_0(S) = 1\{\varnothing \in B\} + \lambda(B \setminus \{\varnothing\}),$$

where λ is the Lebesgue measure on $\mathcal{Y}$. As in Appendix A, the utility of a man m matching with acquaintance k is determined at equilibrium by $U(x, y_k^m) + (\sigma/2)\varepsilon_k^m$, but y_k^m can now take value $\varnothing$, in which case $U(x, \varnothing) = \Phi(x, \varnothing)$. The indirect utility of man m is thus given by

$$Z = \max_k \left\{U(x, y_k^m) + \frac{\sigma}{2}\varepsilon_k^m\right\},$$

and one has

$$\begin{aligned}\log \Pr(Z \le c) &= -\int\int_{\mathcal{Y}_0\times\mathbb{R}} 1\left(U(x,y) + \frac{\sigma}{2}\varepsilon > c\right) d\lambda_0(y) e^{-\varepsilon} d\varepsilon \\ &= -\exp\left\{-c + \log\left[\exp\frac{\Phi(x,\varnothing)}{\sigma/2} + \int_{\mathcal{Y}} \exp\frac{U(x,y)}{\sigma/2}\,dy\right]\right\},\end{aligned}$$

so that

$$\frac{f_0(x)}{\bar{f}(x)} = \frac{\exp\dfrac{\Phi(x,\varnothing)}{\sigma/2}}{\exp\dfrac{\Phi(x,\varnothing)}{\sigma/2} + \displaystyle\int_{\mathcal{Y}} \exp\frac{U(x,y)}{\sigma/2}\,dy},$$

$$\frac{g_0(y)}{\bar{g}(y)} = \frac{\exp\dfrac{\Phi(\varnothing,y)}{\sigma/2}}{\exp\dfrac{\Phi(\varnothing,y)}{\sigma/2} + \displaystyle\int_{\mathcal{X}} \exp\frac{V(x,y)}{\sigma/2}\,dx},$$

while

$$\pi(y|x) = \frac{\exp\dfrac{U(x,y)}{\sigma/2}}{\displaystyle\int_{\mathcal{Y}} \exp\frac{U(x,y')}{\sigma/2}\,dy'},$$

$$\pi(x|y) = \frac{\exp\dfrac{V(x,y)}{\sigma/2}}{\displaystyle\int_{\mathcal{X}} \exp\frac{V(x',y)}{\sigma/2}\,dx'}.$$

Hence we see that the observation of π identifies $U(x, y)$ up to an additive term $c(x)$ and $V(x, y)$ up to an additive term $d(y)$; hence U and V are identified by

$$\begin{aligned}U(x,y) &= \sigma/2[\log\pi(y|x) + c(x)],\\ V(x,y) &= \sigma/2[\log\pi(x|y) + d(y)],\end{aligned}$$

and

$$\Phi(x,y) = \frac{\sigma}{2}[\log\pi(y|x) + \log\pi(x|y) + c(x) + d(y)],$$

where $c(x)$ and $d(y)$ are undetermined. This is precisely the identification achieved in Section II.B. The crucial conclusion is that the observation of singles does not change anything in the identification of U and V. This is a consequence of the independence of irrelevant alternatives of the logit model: indeed, the incentive for remaining single does not affect the odds ratios of the choices of the partners' types. As a result, the distributions of matched men and women $f(x)$ and $g(y)$ may be treated as exogenous.

Once U and V have been identified, one has

$$\frac{f_0(x)}{\bar f(x)} = \frac{\exp\dfrac{\Phi(x,\varnothing)}{\sigma/2}}{\exp\dfrac{\Phi(x,\varnothing)}{\sigma/2} + \exp c(x)},$$

$$\frac{g_0(y)}{\bar g(y)} = \frac{\exp\dfrac{\Phi(\varnothing,y)}{\sigma/2}}{\exp\dfrac{\Phi(\varnothing,y)}{\sigma/2} + \exp d(y)}.$$

Hence by inversion,

$$\Phi(x,\varnothing) = \frac{\sigma}{2}\left[\log\frac{f_0(x)}{\bar f(x) - f_0(x)} + c(x)\right],$$

$$\Phi(\varnothing,y) = \frac{\sigma}{2}\left[\log\frac{g_0}{\bar g(y) - g_0(y)} + d(y)\right],$$

which implies that the observation of single individuals allows one to identify the reservation utilities. As a result, the utility surplus from matching $\Phi(x,y) - \Phi(x,\varnothing) - \Phi(\varnothing,y)$ is identified in the data by

$$\log\left\{\frac{\pi(y|x)[\bar f(x) - f_0(x)]}{f_0(x)}\,\frac{\pi(x|y)[\bar g(y) - g_0(y)]}{g_0(y)}\right\}, \tag{D1}$$

and the ex ante expected utility surpluses of men of type x and women of type y are given just as in Choo and Siow by

$$u(x) = \log\frac{\bar f(x)}{f_0(x)} \quad \text{and} \quad v(y) = \log\frac{\bar g(y)}{g_0(y)}. \tag{D2}$$

These formulas are the continuous extensions of the formulas given in Choo and Siow (2006), where the surplus from matching is identified by $\log[\mu_{xy}^2/(\mu_{x0}\mu_{0y})]$, where μ_{x0} and μ_{0y} are, respectively, the number of single men and women of type x and y, respectively, and μ_{xy} is the number of xy pairs.

Appendix E

Further Details on the Data Set

A. Questionnaire about Personality and Attitudes[15]

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

[15] The website http://www.centerdata.nl/en/TopMenu/Databank/DHS_data/Codeboeken/ provides a link to the complete description of the questionnaire.

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 toward things/oriented toward 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 or offended/sensitive, easily hurt or offended
TEG9: trusting, credulous/suspicious
TEG10: oriented toward 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 toward 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 1 2 3 4 5 6 7 totally agree

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.

B. *Construction of the "Big Five" Personality Factors*

The DHS panel contains three lists of items that would allow one to assess a respondent's personality traits.

1. The first list contains 150 items and refers to the Five-Factor Personality Inventory (FFPI) measure, developed by Hendriks et al. (1999). This list was included in a supplement to the 1996 wave.
2. The second list refers to the 16 Personality Adjective (16PA) scale developed by Brandstätter (1988) and was included in the module "Economic and Psychological Concepts" from 1993 until 2002.
3. From 2003 on, the panel replaced the 16PA scale by the International Personality Item Pool (IPIP) developed by Goldberg (1999). The 10-item list

version of the IPIP scale is used except for the 2005 wave, where the 50-item list was implemented.

Of the three scales, the 16PA scale covers the largest sample of individuals. For that reason, the 16PA scale was chosen to measure personality traits. This scale offers the respondents the opportunity to locate themselves on 16 personality dimensions. Each dimension is represented by two bipolar scales so that the full scale contains 32 items. Nyhus and Webley (2001) show that this scale distinguishes five factors.[16] They labeled these factors as emotional stability, extraversion, conscientiousness, agreeableness, and autonomy. Of the 32 items associated with the 16PA measure, the first half was asked in 1993, 1995, and each year between 1997 and 2002 while the other half was asked in 1994 and 1996 only. Constructing the full scale, therefore, requires losing all respondents but those who responded in two successive years between 1993 and 1996. To avoid throwing out too many observations, we constructed the five dimensions using only those 16 items included in the 1993, 1995, and 1997–2002 waves. Since answers given to the same item by the same person in different waves are strongly correlated (see Nyhus and Webley 2001), we simply collapse the data by individual using the person's median answer to each item.

We have constructed our five factors by adding the (standardized) items identified by Nyhus and Webley (2001) for the respective scales. In other words, "emotional stability" is constructed using the following items:

- "oriented toward reality"/"dreamer,"
- "happy with myself"/"doubtful,"
- "need to be supported"/"independent,"
- "well balanced"/"quick tempered,"
- "slow thinker"/"quick thinker," and
- "easily worried"/"not easily worried."

"Agreeableness" is constructed using the following items:

- "creature of habit"/"open to changes,"
- "slow thinker"/"quick thinker," and
- "quiet, calm"/"vivid, vivacious."

"Autonomy" is constructed using

- "direct, straightforward"/"diplomatic,"
- "quiet, calm"/"vivid, vivacious," and
- "shy"/"dominant."

[16] Using the 1996 wave that contains both the FFPI module and the 16PA module, Nyhus and Webley (2001) checked the correlation between the five factors identified by the 16PA scale and the (big) five factors identified by the FFPI. The correlation is generally high though not perfect. This suggests that both sets of factors assess slightly different aspects of the latent factors. We followed Nyhus and Webley and use a slightly less general wording for the various dimensions identified from the 16PA scale.

"Extraversion" is based on

"oriented toward things"/"toward people,"
"flexible"/"stubborn," and
"trusting, credulous"/"suspicious."

"Conscientiousness" is constructed using

"little self-control"/"disciplined,"
"carefree"/"meticulous," and
"not easily hurt"/"easily hurt, sensitive."

As a robustness check, we constructed the full scale using the 1993, 1994, 1995 and 1996 waves. We followed Nyhus and Webley (2001) and constructed the five factors using Principal Component Analysis and varimax rotation on the five main factors. The correlation between each of the factors we constructed using only 16 items and the corresponding factor using the full scale varies between .42 for agreeableness and .76 for emotional stability.

References

Anderberg, D. 2004. "Getting Hitched: The Equilibrium Marriage Market Behaviour of a British Cohort." Discussion Papers in Economics, Univ. London, Royal Holloway.
Atakan, A. 2006. "Assortative Matching with Explicit Search Costs." *Econometrica* 74 (3): 667–80.
Becker, G. S. 1973. "A Theory of Marriage: Part I." *J.P.E.* 81 (4): 813–46.
———. 1991. *A Treatise on the Family.* Cambridge, MA: Harvard Univ. Press.
Ben-Akiva, M., N. Litinas, and K. Tsunekawa. 1985. "Continuous Spatial Choice: The Continuous Logit Model and Distributions of Trips and Urban Densities." *Transportation Res., Part A General* 19:83–206.
Ben-Akiva, M., and T. Watanatada. 1981. *Application of a Continuous Spatial Choice Logit Model in Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, MA: MIT Press.
Bertrand, M., and A. Schoar. 2003. "Managing with Style: The Effect of Managers on Firm Policies." *Q.J.E.* 118 (4): 1169–1208.
Bojilov, R., and A. Galichon. 2013. "Closed-Form Formulas for Multivariate Matching." Manuscript, Soc. Sci. Res. Network.
Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel. 2008. "The Economics and Psychology of Personality Traits." *J. Human Resources* 43 (4): 972–1059.
Bowles, S., H. Gintis, and M. Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *J. Econ. Literature* 39 (4): 1137–76.
Brandstätter, H. 1988. "Sixteen Personality Adjective Scales as a Substitute for the 16PF in Experiments and Field Studies." *Zeitschrift fur Experimentelle und Angewandte Psychologie* 35:370–90.
Browning, M., P.-A. Chiappori, and Y. Weiss. 2014. *Economics of the Family.* Cambridge: Cambridge Univ. Press.

Bruze, G. 2011. "Marriage Choices of Movie Stars: Does Spouse's Education Matter?" *J. Human Capital* 5 (1): 1–28.

Charles, K. K., E. Hurst, and A. Killewald. 2013. "Marital Sorting and Parental Wealth." *Demography* 50 (1): 51–70.

Chiappori, P.-A., and S. Oreffice. 2008. "Birth Control and Female Empowerment: An Equilibrium Analysis." *J.P.E.* 116 (1): 113–40.

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque. 2012. "Fatter Attraction: Anthropometric and Socieconomic Matching on the Marriage Market." *J.P.E.* 120 (4): 659–95.

Chiappori, P.-A., B. Salanié, and Y. Weiss. 2010. "Assortative Matching on the Marriage Market: A Structural Investigation." Working paper, Columbia Univ.

Choo, E., and A. Siow. 2006. "Who Marries Whom and Why." *J.P.E.* 114 (1): 175–201.

Coles, M., and M. Francesconi. 2011. "On the Emergence of Toyboys: The Timing of Marriage with Aging and Uncertain Careers." *Internat. Econ. Rev.* 52 (3): 825–53.

Cosslett, S. 1988. "Extreme-Value Stochastic Processes: A Model of Random Utility Maximization for a Continuous Choice Set." Technical report, Ohio State Univ.

Custodio, C., M. Ferreira, and P. P. Matos. 2013. "Generalists versus Specialists: Lifetime Work Experience and Chief Executive Officer Pay." *J. Financial Econ.* 108 (2): 471–92.

Dagsvik, J. 1994. "Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes." *Econometrica* 62:1179–1205.

Das, M., and A. van Soest. 1999. "A Panel Data Model for Subjective Information on Household Income Growth." *J. Econ. Behavior and Org.* 40 (4): 409–26.

Decker, C., E. Lieb, R. McCann, and B. Stephens. 2013. "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market." *J. Econ. Theory* 148:778–92.

Dupuy, A., and A. Galichon. Forthcoming. "Canonical Correlation and Assortative Matching: A Remark." *Ann. Econ. and Statis.*

Echenique, F., S. Lee, B. Yenmez, and M. Shum. 2013. "The Revealed Preference Theory of Stable and Extremal Stable Matchings." *Econometrica* 81 (1): 153–71.

Elfving, G. 1967. "A Persistency Problem Connected with a Point Process." *J. Appl. Probability* 4:77–89.

Falato, A., D. Li, and T. T. Milbourn. 2012. "Which Skills Matter in the Market for CEOs? Evidence from Pay for CEO Credentials." Working paper, Fed. Reserve Board, Washington, DC.

Fox, J. 2010. "Identification in Matching Games." *Quantitative Econ.* 1:203–54.

———. 2011. "Estimating Matching Games with Transfers." Working paper, Univ. Michigan.

Gabaix, X., and A. Landier. 2008. "Why Has CEO Pay Increased So Much?" *Q.J.E.* 123:49–100.

Galichon, A., and B. Salanié. 2010. "Matching with Trade-offs: Revealed Preferences over Competing Characteristics." Discussion Paper no. 7858, Centre Econ. Policy Res., London.

———. 2013. "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models." Working paper, Soc. Sci. Res. Network.

Goldberg, L. R. 1999. "A Broad-Bandwidth, Public-Domain, Personality Inventory Measuring the Lower-Level Facets of Several Five-Factor Models." In *Personality Psychology in Europe*, vol. 7, edited by I. Mervielde, I. Deary, F. De Fruyt, and F. Ostendorf, 7–28. Tilburg, Neth.: Tilburg Univ. Press.

Graham, B. 2011. "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers." In *Handbook of Social Economics*, edited by J. Benhabib, A. Bisin, and M. Jackson. Amsterdam: Elsevier.

Gretsky, N., J. Ostroy, and W. Zame. 1992. "The Nonatomic Assignment Model." *Econ. Theory* 2 (1): 103–27.

Heckman, J. J. 2007. Notes on Koopmans and Beckmann's "Assignment Problems and the Location of Economic Activities." Lecture notes, Univ. Chicago.

Hendriks, A. A. J., W. K. B. Hofstee, B. De Raad, and A. Angleitner. 1999. "The Five-Factor Personality Inventory (FFPI)." *Personality and Individual Differences* 27:307–25.

Hitsch, G. J., A. Hortacsu, and D. Ariely. 2010. "Matching and Sorting in Online Dating." *A.E.R.* 100 (1): 130–63.

Horn, R., and C. Johnson. 1991. *Topics in Matrix Analysis*. Cambridge: Cambridge Univ. Press.

Jacquemet, N., and J.-M. Robin. 2013. "Marriage with Labor Supply." Working paper, Sciences Po.

Kleibergen, F., and R. Paap. 2006. "Generalized Reduced Rank Tests Using the Singular Value Decomposition." *J. Econometrics* 133:97–126.

Lundberg, S. 2012. "Personality and Marital Surplus." *IZA J. Labor Econ.* 1:3.

McFadden, D. 1976. "The Mathematical Theory of Demand Models." In *Behavioral Travel-Demand Models*, edited by P. Stopher and A. Meyburg, 305–14. Lexington, MA: Heath.

Mueller, G., and E. Plug. 2006. "Estimating the Effect of Personality on Male and Female Earnings." *Indus. and Labor Relations Rev.* 60 (1): 3–22.

Nesheim, L. 2012. "Identification in Multidimensional Hedonic Models." Working paper, Univ. Coll. London.

Nyhus, E. K. 1996. "The VSB-CentER Savings Project: Data Collection Methods, Questionnaires and Sampling Procedures." Progress Report. VSB-CentER Savings Project, Tilburg Univ.

Nyhus, E. K., and P. Webley. 2001. "The Role of Personality in Household Saving and Borrowing Behaviour." *European J. Personality* 15:S85–S103.

Oreffice, S., and C. Quintana-Domeque. 2010. "Anthropometry and Socioeconomics among Couples: Evidence in the United States." *Econ. and Human Biology* 8 (3): 373–84.

Robin, J.-M., and R. J. Smith. 2000. "Tests of Rank." *Econometric Theory* 16:151–75.

Rüschendorf, L. 1995. "Convergence of the Iterative Proportional Fitting Procedure." *Ann. Statis.* 23:1160–74.

Rüschendorf, L., and W. Thomsen. 1993. "Note on the Schrödinger Equation and I-Projections." *Statis. and Probability Letters* 17:369–75.

Shapley, L., and M. Shubik. 1972. "The Assignment Game I: The Core." *Internat. J. Game Theory* 1:111–30.

Shimer, R., and L. Smith. 2000. "Assortative Matching and Search." *Econometrica* 68:343–69.

Stutzer, A., and B. Frey. 2006. "Does Marriage Make People Happy, or Do Happy People Get Married?" *J. Socio-economics* 35 (2): 326–47.

Terviö, M. 2008. "The Difference That CEOs Make: An Assignment Model Approach." *A.E.R.* 98:642–68.

Wong, L. Y. 2003. "Structural Estimation of Marriage Models." *J. Labor Econ.* 21 (3): 699–728.

Zimmermann, A., and R. Easterlin. 2006. "Happily Ever After? Cohabitation, Marriage, Divorce, and Happiness in Germany." *Population and Development Rev.* 32 (3): 511–28.