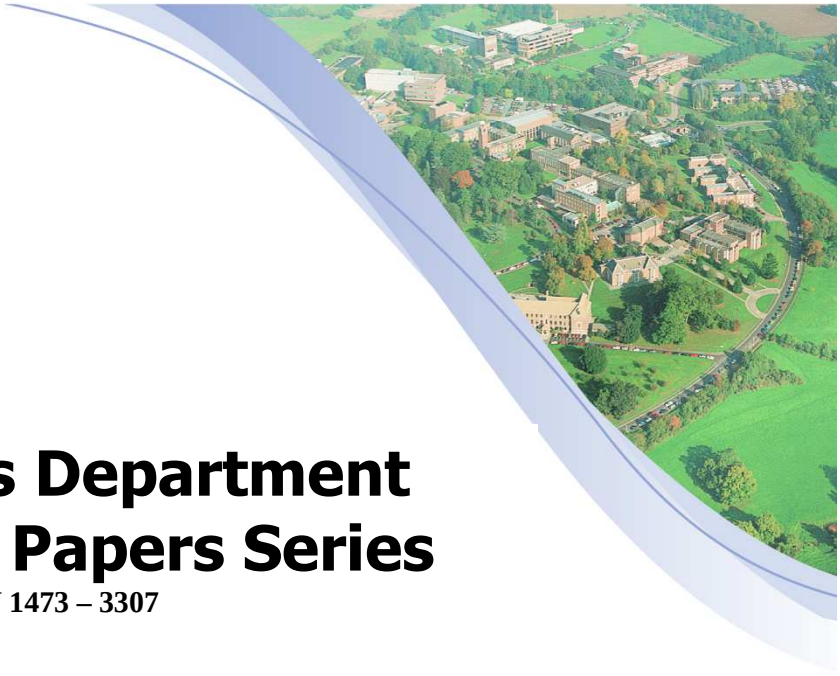

Economics Department Discussion Papers Series

ISSN 1473 – 3307

Who Marries Whom? The Role of Identity, Cognitive and Noncognitive Skills in Marriage

Annalisa Marini

Paper number 19/04

URL Repec page: http://ideas.repec.org/s/exe/wpaper.html

Who Marries Whom? The Role of Identity, Cognitive and Noncognitive Skills in Marriage

Annalisa Marini*

May 25, 2019

Abstract

I estimate a structural model of marriage sorting on a representative sample of British individuals. The paper first investigates the importance of numerical skills in the selection of the partner and the role of identity for marriage matching on a British sample. The findings show that identity is among the most important attributes, together with education and physical characteristics, in marriage sorting. Cognitive skills are both direct and indirect determinants of marriage matching. Personality traits are also relevant in the choice of the partner: conscientiousness and openness to experience play, in addition to risk propensity, a direct and an indirect role, while agreeableness, extraversion and neuroticism matter only indirectly. Interesting findings, robust to both alternative specifications and a sensitivity analysis, and heterogeneous preferences between males and females emerge from the analysis.

JEL: J12, D9, C01

Keywords: Marriage, Identity, Cognitive and Noncognitive Skills

*Marini: University of Exeter, Streatham Court, Rennes Drive, EX4 4PU, Exeter, UK, a.marini@exeter.ac.uk. First draft: September 2017. I am very grateful to Sonia Oreffice, Climent Quintana-Domeque and Jo Silvester for useful suggestions and feedback. All the remaining errors are mine.

"*Wives and Oxes of Your Own Places*", Italian Proverb

"*Better to Marry a Neighbor than a Stranger*", Uruguayan Proverb

"*Birds of a Feather Flock Together*", English Proverb

1 Introduction

Marriage is one of the most traditional forms of social interactions and it is the most common way partners use to transmit their values, norms and identity in a family and across generations. Indeed, married couples act as a group in several day-life situations and identity, norms, values and culture play often a role in many decisions they make. For instance, when parents socialize their children they are often motivated by a sort of 'paternalistic altruism', namely, they put effort in socializing their kids (so they are altruistic), and often this altruism is paternalistic, that is, parents evaluate their offsprings' actions from their own point of view and preferences and they want to transmit to the kids their own values and norms (Bisin and Verdier, 2000, 2001).

Yet, what determines the formation of couples and their identity? When choosing whom they want to get married to, individuals take into account that they are likely, in the future, to make decisions about themselves and their offsprings jointly with their partner. The outcome of this process is a sort of 'we thinking', which implies the creation of a group pride and the importance, for the couple, of being esteemed (Akerlof, 2016). Thus, understanding what drives marriage sorting is crucial to explain both the formation of a family as a club good (see e.g., Carvalho, 2016, for a def-

inition) and the dynamic evolution of identity, values and norms across generations.

This paper shows that identity is a crucial determinant of assortative mating and marriage. It is well recognized by the literature (Bisin and Verdier, 2000) that direct socialization, that is, socialization from parents to children, is more efficient when a family is homogamous (i.e., a family whose parents share the same cultural traits) than when it is heterogamous (i.e., a family where parents have different cultural traits) because it is easier for parents with similar values and identity to have common perspectives about what is better for the family and the children.

Drawing on this literature, in this paper I estimate a model of marriage sorting where the decision of couples to form a marriage partnership is based on multiple attributes, one of which is a measure of ethnolinguistic fractionalization. To the best of my knowledge, this paper is the first to empirically assess the impact of identity in marriage matching on a sample of British individuals. In more details, I test for the importance of ethnic identity for marriage sorting and the determination of the joint utility of married couples. To represent identity, I decide to focus on ethnic heterogeneity defined as in Desmet, Ortuño-Ortín and Wacziarg (2017). They develop an index of ethnic heterogeneity for 76 countries. The index measures the average probability (by country) that a person meets or is matched with another person from a different ethnic group, and it is meant to measure social antagonism: it ranges from 0 to 1, where 1 indicates extreme social antagonism. They use this index, among others, to investigate the effect of ethnic and cultural heterogeneity on various political economy outcomes.

Figure 1: Ethnolinguistic fractionalization index (Desmet, Ortuño-Ortín and Wacziarg, 2017)

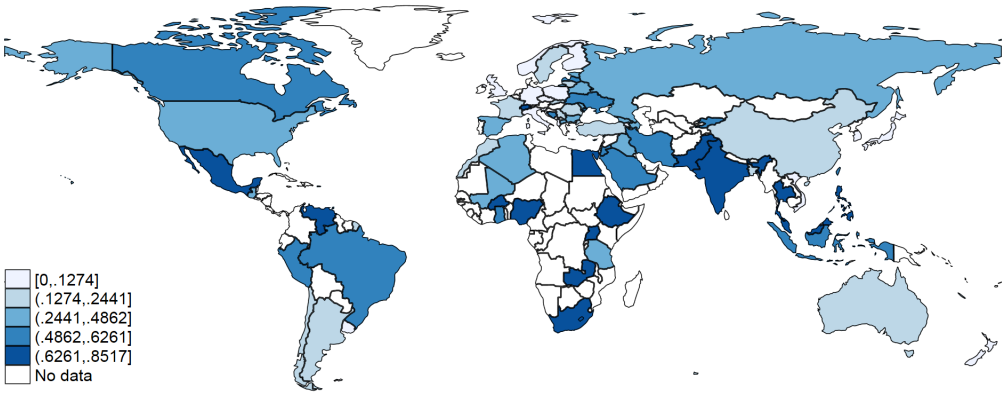

Figure 1 shows the values of this index. As it is noticeable from the Figure, the index is particularly low for some countries, such as Finland, Italy, Norway, North Korea, low for others such as China, Australia, Argentina and Chile; it is higher for North America and some countries of Central and Latin America, such as Mexico and Venezuela, and it is also high in many of the African countries and some countries in Europe (e.g., Switzerland) and in Asia (e.g., India and Pakistan).

I chose to use ethnolinguistic fractionalization (ELF) to study the importance of identity in marriage sorting for specific reasons. First, the literature on marriage has recognized the crucial importance of ethnic and religious traits to explain assortative mating. As a matter of fact, parents ground their family on, and transmit to their offsprings, well defined and specific traits, which they inherited themselves from their parents and forebears. Thus, in presence of ethnic heterogeneity it is comprehensible that individuals may want to get married to a partner that shares their own values and norms. This may in part motivate and can be linked to the absence, even in

heterogeneous societies such as the United States or the United Kingdom, of a 'melting pot', where individuals with different ethnic and religious backgrounds would eventually assimilate into a society (e.g., Bisin and Verdier, 2000), and the persistence of social and residential segregation of minorities and generations of immigrants. The tendency to have more homogamous marriages (Bisin and Verdier, 2000) is sometimes stronger for migrants, who tend to be strongly tied to and to emphasize even more their cultural inheritance (see among others Cutler, Glaeser and Vigdor, 2008). Second, the authors who developed the index assumed that ELF could be a source of social antagonism and be linked to (and cause) civil conflicts or other conflicts for resource allocation within countries. In this paper I argue that ELF is important also for the determination of happiness and joint utility of a couple. However, while Desmet, Ortuño-Ortín and Wacziarg (2017) show that ELF may be detrimental for macroeconomic conditions of a country, this may not be the case when we investigate marriage decisions made by couples. Instead, ELF could eventually be beneficial to the well-being of a couple.

In order to conduct the analysis I use the years 2009-2011 of the Understanding Society data, which is a survey representative of the British population. I use this data set because it is particularly suitable to address the question of the present work. Indeed, the United Kingdom has an ethnically heterogeneous population, so the survey can easily be used to investigate the impact of ELF on individual marriage sorting. Furthermore, in the data set we have information not only on ethnicity of the respondent, but also on ethnic origins of parents and grandparents, which I use in order

to derive inherited ethnic heterogeneity. In addition, this data set contains information on education and demographics, which are generally used in the marriage literature, as well as on additional variables, such as cognitive and noncognitive skills, physical characteristics, such as BMI and Height, which are relevant to mating, and a variable for risk propensity. This allows me to build on the most recent advances of the marriage matching literature (e.g., Oreffice and Quintana-Domeque, 2010; Chiappori, Oreffice and Quintana-Domeque, 2012; Dupuy and Galichon, 2014) in that I let matching of couples be based on multiple attributes of various nature.

The paper contributes to the literature as follows. First, to the best of my knowledge, this is the first work to investigate the impact of ELF on marriage matching. Indeed, studies of sorting in marriage have often focused on the role of attributes such as education, income, physical and more recently genetic and psychological characteristics to explain assortative mating in marriage (Chiappori, Salanie and Weiss, 2010; Chiappori, Oreffice and Quintana-Domeque, 2012; Dupuy and Galichon, 2014). Despite identity and ELF are recognized to be important determinants of various socioeconomic outcomes (Akerlof and Kranton, 2000; Bisin and Verdier, 2001; Bénabou and Tirole, 2006; Guiso et al., 2008; Desmet, Ortuño-Ortín and Wacziarg, 2017) and are considered among the most salient characteristics in the determination of marriage matching in the theoretical literature, their relevance has not been empirically tested yet on a nationally representative sample. As a matter of fact, while some recent work tests for the importance of family values and gender norms (Bertrand, Kamenica and Pan, 2015; Goussé, Jaquement and Robin, 2015*b*, 2017) for marriage re-

lated outcomes, none of these contributions investigates the importance of ethnic identity. As far as I know Ciscato, Galichon and Goussé (2018) and Adda, Pinotti and Tura (2019) are the other two papers providing evidence on the role of identity in the marriage market. However, these two papers work out results that either refer to a specific subsample (i.e. Ciscato et al. (2018) conduct the analysis on a sample of Californian individuals) or are dependent on national laws (i.e., Adda et al., 2019). In addition, they do not use the measure of ELF worked out by Desmet et al. (2017) I use here. Thus, this work is different from the existing contributions because, on the one hand, it provides more general results on the role of identity in marriage; on the other hand, while values and gender norms may vary across families and may differ within a country or a same ethnic group, ethnolinguistic fractionalization captures individual cultural and ethnic identity and its inheritance. This is the first main empirical contribution of this paper. Second, in addition to considering the role of identity, education and personality traits in assortative mating, I also consider the importance of cognitive skills in the formation of a partnership. While the literature has often considered education, income, and more recently noncognitive skills, I first show that numerical skills are also important to investigate who marries whom and that their consideration to study marriage sorting may provide further interesting insights about the choice of the partner. This contribution is the other main empirical innovation of the paper. Finally, my work can be considered an extension of previous work. Indeed, I estimate a structural model similar to the model in Dupuy and Galichon (2014) to investigate which attributes are important determinants of marriage sorting.

By using British data, the paper extends the results of the literature to a British sample; in so doing, I can compare the results obtained here to the findings in Dupuy and Galichon (e.g., 2014) and provide evidence of external validity of the technique they propose. Also, it innovates with respect to Dupuy and Galichon (2014) by adding the two empirical contributions mentioned above (i.e., the use of ELF and numerical skills).

The results show that ELF is one of the most important attributes to explain marriage matching. Second, cognitive skills also matter for the matching process and its inclusion sheds further light on the formation of couples. Third, in accord with the previous literature (e.g., Dupuy and Galichon, 2014), education is the most important attribute in the determination of the joint utility of couples, followed by physical characteristics and ELF. Fourth, personality traits are good predictors of positive assortative matching of couples; in particular, openness to experience and conscientiousness of both partners are directly relevant to explain the joint utility of the couple; in addition, these and the other personality traits determine matching preferences indirectly, through the interaction with other attributes. Fifth, risk propensity plays both a direct and an indirect role in the determination of who marries whom. Sixth, in accord with the marriage literature, the results show that individuals choose their partner based on physical characteristics too. Seventh, the results support findings of the behavioral literature suggesting the presence of heterogeneous preferences across males and females in decision-making and delineate very interesting patterns. Eighth, findings are consistent across various specifications; also, they show that ELF is slightly more relevant when the ELF index inherited

from the mother is used than when I use the one inherited from the father (which is in line with the findings of the literature, see for instance Ljunge, 2014) and that identity inherited from the father is also partially indirectly relevant. Ninth, although I do not use exactly the same attributes used by Dupuy and Galichon (2014) in their work, I show my findings align to theirs and for the British sample are sometimes more significant, providing some interesting insights. Finally, the sensitivity analysis shows that results are robust to the changing of starting values.

The rest of the paper is structured as follows. The next section presents the related literature. Section 3 describes the data and the empirical methodology. Section 4 reports the results, alternative specifications and the sensitivity analysis. In section 5 I present a discussion and concluding remarks.

2 Related Literature

This work relates to various fields of economics. First, drawing on the most recent advances of the literature, it contributes to the assortative mating and marriage sorting literature. This literature dates back to Becker (1973, 1974); since this pioneering work, several contributions have been made to the literature and progresses in identification were crucial to allow the estimation of marriage models without imposing too much structure. Choo and Siow (2006) propose and estimate a static model of marriage with transferable utility and spillovers that minimizes a priori assumptions on preferences of individuals and estimate the model on a sample of US couples from the 1970 and 1980; Choo and Siow (2007) extend this model to

a dynamic framework. They found that gains in marriage diminished for younger generations and that the effect of some policies, such as for instance abortion, played a role in this fall. Also, most of the literature has focused so far on the importance of a single characteristic to explain matching patterns (Chiappori and Oreffice, 2008; Charles, Hurst and Killewald, 2013); other work allowed multiple characteristics to count in marriage sorting, but only by means of a single index (Anderberg, 2004; Chiappori, Oreffice and Quintana-Domeque, 2012). Only recently, marital matching started to be explained using multiple attributes. Some articles limited the study of marriage sorting to the use of discrete characteristics (Chiappori, Salanie and Weiss, 2010; Goussé, Jaquement and Robin, 2015*a*). The work by Choo and Siow (2006) cited earlier is another example. Very recently Dupuy and Galichon (2014) extend the previous literature in order to allow the use of continuous attributes to determine marital sorting. They also introduce the concept of 'saliency analysis' to determine which of these attributes count in the matching process. This empirical framework allows the researcher to investigate not only which attributes count most in the sorting process, but also the extent to which they are complement or substitutes and the presence of heterogeneities between husbands and wives. In the present paper I adopt the framework they develop to investigate which attributes drive marriage matching in the United Kingdom.

In addition, I link the marriage sorting literature to the identity economics literature by testing the importance of ethnic heterogeneity for marriage sorting. Indeed, it is recognized that people of different background, ethnicity and religion may have heterogeneous preferences for mar-

riage (e.g., Choo and Siow, 2006). In particular the role of identity may give rise to segregation in marriage where some ethnicites may prefer homogamous marriages (Bisin and Verdier, 2000). Thus, this paper contributes to the literature by first testing the importance of ethnic heterogeneity for marriage on a representative sample of British individuals. While several aspects of identity may matter for marriage sorting, I use this index because ethnicity is one of the traits that is likely to lead to homogamous marriage (Bisin and Verdier, 2000); also, while cultural heterogeneity may be relevant, it has been recently showed that ethnic identity can predict cultural values (Desmet, Ortuño-Ortín and Wacziarg, 2017). In this sense, the ELF index I use is a more general index than others capturing for instance cultural heterogeneity.

Some of the most recent contributions of the marriage literature are now investigating the role of values and norms in marriage outcomes (Bertrand, Kamenica and Pan, 2015; Goussé, Jaquement and Robin, 2015*b*, 2017). Goussé, Jaquement and Robin (2017), for instance, construct a Family Value Index (FVI) that captures family attitudes and used it, together with other attributes, such as gender wage inequality, heterogeneous preferences of males and females and home production technologies, to explain gender division of labor. Using the BHPS (1991-2008), they build a search-and-matching-and-bargaining model where individual utility is determined by a market good, leisure and a good produced within the household using the non-market time of the components of the family. Their results show that there exist homophily in education; also, the matrix shows that while progressive individuals are more open to marry a partner disregarding her/his

family values, traditional individuals are more keen to marry a partner with traditionalist values. The estimation of the model parameters also reveals, not surprisingly, that traditionalist females prefer to spend more of their time for leisure and home production, while traditionalist males are less keen to spend their time in home production. Furthermore, their counterfactual simulations show that if all individuals had progressive values there would be more singles and females would significantly increase their labor supply and if in addition females had the same preferences as males, this would result in a further increase in female labor supply. Bertrand, Kamenica and Pan (2015) analyze the impact that gender identity norms have in marriage. Using the Panel Study of Income Dynamics (PSID), the Survey of Income and Programme Participation (SIPP), and the Census and the American Community Survey, they study causes and consequences of relative income within households. They show that more conservative/traditionalist views of gender norms have lead to the decline in marriage rates over the last three decades; also, they show that females earning more than their partner may still be perceived as a problem and that females with higher earning potential than their partners may be willing to either not participate in the labor force or to accept lower wages in order not to get gender role reversal, but then an unhappy marriage and divorce may become more likely.

My contribution is different from this work investigating the role of gender norms and family values in the determination of marriage matching and family decisions because it investigates the role of ethnic identity rather than gender identity or family values. In so doing, the paper aligns to the other two contemporaneous studies that investigate the impact of

identity in marriage sorting, namely, Ciscato, Galichon and Goussé (2018) and Adda, Pinotti and Tura (2019). As a matter of fact, Ciscato et al. (2018), using a sample of individuals living in California during the period 2008-2012, estimate a structural model of marriage sorting of same-sex couples and compare the results to heterosexual couples. They model marriage sorting as a function of multiple attributes among which ethnicity; their findings suggest that regarding identity for various reasons different-sex couples show stronger preferences for homogamy with respect to same-sex couples. Instead, Adda et al. (2019) estimate a structural model of marriage between natives and migrants in Italy. In Italy marrying a native gives access to legal status acquisition and to citizenship; by exploiting the enlargement of the European Union to Eastern European countries, they set up a natural experiment to disentangle between marriage to acquire legal status and marriage for cultural distance. They showed, among other findings, that after the Eastern European countries joined the European Union, migrants started to get married more with partners belonging to their own communities and less with natives. Also, cultural affinity vary depending on area of origin of migrants and it is partially correlated with genetic distance. My paper differs from these two contributions because it estimates a model of marriage sorting on a representative sample of British individuals. So, as in their studies I provide evidence on the role of identity in marriage sorting. However, my results are neither specific to a sub-national area (e.g., California) as in Ciscato et al. (2018), nor aimed at investigating the effect of European and state legislation on marriages between natives and migrants. Thus, the results in this paper are more general and can be

generalized (by eventually studying the impact of ELF on other representative samples) and used to investigate (dis)similarities in the role of ELF and more broadly in matching complementarities across countries. Furthermore, they do not use the ELF index and while (in line with Dupuy and Galichon, 2014) I let sorting be based on a comprehensive set of attributes that includes demographics, cognitive and noncognitive skills, Ciscato et al. (2018) and Adda et al. (2019) use a more limited set of attributes.

Finally, as in Dupuy and Galichon (2014) I let sorting of couples depend on a series of attributes: in addition to ELF (which they do not include among the attributes), I let marriage sorting be a function of cognitive and noncognitive skills. Although the presence of the highest level of education attained can in part capture cognitive abilities and income, I consider important to explicitly account for the presence of cognitive skills: indeed, they are recognized to be good predictors of outcomes over the life cycle also because they can well explain behavior in presence of uncertainty (see e.g., Borghans et al., 2008). Thus, to the best of my knowledge I first use a variable for numeric ability to study how cognitive skills affect marriage sorting. By doing this, I also align to the literature that considers the importance of both cognitive and noncognitive skills to explain individual behavior over the lifetime (e.g., Heckman, Stixrud and Urzua, 2006; Borghans et al., 2008; Cunha and Heckman, 2008; Cunha, Heckman and Schennach, 2010). In line with the literature, I use, in addition to cognitive skills, personality traits and a measure for risk propensity to capture noncognitive skills and risk preferences of individuals.

3 Model and Data

3.1 Empirical Framework

The model I estimate in the next section grounds on the model in Dupuy and Galichon (2014), who extend the Choo and Siow (2006) marriage model, where marriage matching is based on discrete attributes, to include continuous attributes. Thus, as in Dupuy and Galichon (2014) I make use of the continuous logit model (see for instance McFadden, 1976; Ben-Akiva and Watanatada, 1981; Ben-Akiva, Litinas and Tsunekawa, 1985; Cosslett, 1988; Dagsvik, 1994).

In the sample we have males and females, denoted respectively with m and w, searching for a partner. The search leads to a one-to-one bipartite matching model with transferable utility. Thus, after couples are matched the number of males and females in the sample is the same. Each man has a series of attributes $y \in Y = R^{dy}$ and each female has attributes $x \in X = R^{dx}$. I assume that males and females look for a partner by searching in the set of their acquaintances, indexed respectively by $l \in M$ and $k \in N$.

A male and a female will match when the matching, namely, the probability density that a couple with certain attributes is formed, maximizes the total utility of the couple. Formally, the utility of a man m of type y matching with a woman w of type x is $U\left(y, x_l^m\right) + \frac{\sigma}{2}\varepsilon_l^m$ and similarly the utility of a woman w of type x matching with a man m of type y is $V\left(x, y_k^w\right) + \frac{\sigma}{2}\eta_k^w$, where $U(\cdot)$ and $V(\cdot)$ are the utilities based on the observable attributes of the potential partner, σ is a parameter that measures dispersion of un-

observed heterogeneity, and ε_l^m and η_k^w are the two sympathy shocks for respectively candidate husbands and candidate wives.

In equilibrium matching maximizes social gains. The maximization problem is as follows:

$$\max_{\pi \in M(P,Q)} \int\int_{YxX} \phi(y,x)\pi(y,x)dydx - \sigma \int\int_{YxX} \log \pi(y,x)\pi(y,x)dydx \qquad (1)$$

namely, the maximization of the utility function subject to the probability the matching occurs. In this optimization problem, $\phi(y,x)$ is the joint utility, $\pi(y,x)$ is the probability distribution that a couple with characteristics (y,x) is formed, σ is the parameter that measures the dispersion of unobserved heterogeneity and P and Q are the probability distributions of the attributes of males and females. The first term of the maximization captures sorting based on observable attributes and the second term of the maximization represents sorting based on unobservables. The two probability distributions for males and females choosing a partner with attributes x and y are, respectively:

$$\pi_{X|Y}(x|y) = \frac{\exp[U(y,x)/(\sigma/2)]}{\int\limits_x \exp[U(y,x')/(\sigma/2)]dx'} \qquad (2)$$

and

$$\pi_{Y|X}(y|x) = \frac{\exp[U(x,y)/(\sigma/2)]}{\int\limits_y \exp[U(x,y')/(\sigma/2)]dy'}, \qquad (3)$$

which both are continuous logit. Taking logarithms of the probability density functions and exploiting the fact that the joint utility is the sum of the

utilities of the two partners, the equilibrium solution is:

$$\pi\left(x,y\right)=\exp\left[\frac{\phi\left(x,y\right)-\left(\sigma/2\right)\left(a(y)-b(x)\right)}{\sigma}\right] \tag{4}$$

where $a(y)=\log\int_{X}\frac{\exp[U(y,x')/(\sigma/2)]}{f(y)}dx'$ and $b(x)=\log\int_{Y}\frac{\exp[V(x,y')/(\sigma/2)]}{g(x)}dy'$.

By substitution, the two equilibrium utilities for husbands and wives are, respectively:

$$U(y,x)=\frac{\phi(x,y)+(a(y)-b(x))}{2} \tag{5}$$

$$V(x,y)=\frac{\phi(x,y)+(-a(y)+b(x))}{2}. \tag{6}$$

Let $\{(x_l^m,\varepsilon_l^m),l\in M\}$ be Poisson distributed with intensity $dx\times\varepsilon^{-\varepsilon}d\varepsilon$. This guarantees independence across disjoint subsets (i.e., the independence of irrelevant alternatives) and it yields to the continuous version of the multinomial logit (see Dupuy and Galichion (2014) for details). This assumption allows to rule out the presence of a systematic sympathy shock (i.e., correlated sympathy shock across women observables). The same is true for men observables holding the same assumption for their attributes and shocks. It could be appropriate to accommodate a random sympathy shock for attributes, because of partner preferences; however, if most of the matching is determined through observables, the impact of unobservables in marriage sorting can be considered negligible.

In the optimization process, σ, the parameter for the dispersion of unobserved heterogeneity, measures the extent to which the matching is determined by the observable attributes or by the unobserved heterogeneity.

In particular, a small σ implies that most of the matching is determined by observed attributes of candidate husbands and wives, a high value of this parameter determines that matching occurs in large part based on unobserved heterogeneity (i.e., the matching is independent on the observable attributes). So if matching is largely determined by the observable attributes of the couple the role played by unobserved heterogeneity is small, that is, the solution is far from a random-matching type of solution and we can assume that the independence of irrelevant alternative is met.

Dupuy and Galichon (2014) explain that the utility function, $\phi(y, x)$, which can be estimated also based on a distributional assumption on unobservables, can be written as $\phi_A(x, y) = y'Ax$, where A is the affinity matrix whose elements are defined as

$$A_{ij} = \frac{\delta^2 \phi(x, y)}{\delta x_i \delta y_j}, \tag{7}$$

Identification can be reached up to a separable additive function because only the cross-derivative $\partial^2 \phi\left(x, y\right) / \partial x \partial y$, whose elements constitute the affinity matrix, A_{ij}, can be identified, while we cannot identify the first derivatives with respect to x and y.

This limitation renders impossible to identify *absolute attractiveness*; however, the cross-derivative allows the researcher to identify *mutual attractiveness*; in particular, when the elements of the matrix are positive, there is positive assortative matching (i.e., complementarity) between the attributes of the husband and the wife, while when the element of the affinity matrix is negative there is negative assortative matching (i.e., the attributes of husband and wife are substitutes). The properties of the affin-

ity matrix are interesting also because they enable the researcher to assess the presence of heterogeneities between candidate husbands and wives in the matching process.

Finally, the structural approach, by controlling for marginal distributions controls for the possible presence of misleading results due to correlations across variables. Once the matrix has been estimated it is possible to conduct a saliency analysis to both determine the rank of the affinity matrix and to construct indices of mutual attractiveness among couples (see Dupuy and Galichon (2014) for additional details on both the methodology and identification).

3.2 Data

The main data set used in the analysis is the British Household Understanding Society data set (BHUS). This data set is a representative survey of the British population and it is the continuation of the British Household Panel Study (BHPS). The BHUS survey is particularly suitable for the idea I test in this paper because it has a series of variables that allow the researcher to investigate which attributes can predict marriage sorting. Indeed, the BHUS data set is one of the national surveys conducted in various countries (other examples are the American Panel Study of Income Dynamics (PSID), the German SocioEconomic Panel (GSOEP) and the DNB Household Survey (DHS), this last one used by Dupuy and Galichon (2014)) that not only include demographics of the interviewed individuals, but also have records of physical characteristics, such as for instance Height and BMI, and personality traits, that can be crucial to understand individual behavior,

such as who marries whom.

The structural model of marriage sorting I estimate in this paper includes thus a series of variables, among which demographics, cognitive and noncognitive skills, physical characteristics that can matter in the determination of attractiveness between husband and wife; by including the ethnolinguistic measure I can test whether and the extent to which identity matters in the choice of a partner. The list of attributes I use to study marriage sorting is similar to the ones used by Dupuy and Galichon (2014) and it is as follows: education, height, Body Mass Index (BMI), the Big 5 personality Traits -Conscienciousness, Extraversion, Agreableeness, Neuroticism, Openness to Experience-, Risk Propensity, a variable capturing cognitive skills -Numeric Ability-, to which I add the index for ethnolinguistic heterogeneity (ELF).

As mentioned above, the variable for ethnolinguistic fractionalization is taken from Desmet, Ortuño-Ortín and Wacziarg (2017): it measures the average probability, by country, that a person meets or is matched with another person from a different ethnic group. Desmet, Ortuño-Ortín and Wacziarg (2017) constructed this index to measure social antagonism, which takes extreme values when the index is equal to 1. However, while social antagonism may be detrimental to macroeconomic outcomes such as provision of public goods and civil conflicts (Eatserly and Levine, 1997; Alesina et al., 2003; Desmet, Ortuño-Ortín and Wacziarg, 2017), this may not be the case for assortative mating. On the one hand, we are studying assortative mating in a social context where scarcity of resources is much less pronounced than in other (notably developing) countries, and where the ability to absorb im-

migrants and facilitate integration and assimilation may offset the negative effects that ethnolinguistic fractionalization can generate from a macroeconomic perspective. On the other hand, both the migration and marriage literature recognize that social interactions of immigrants and marriage are more likely to occur with people of the same ethnicity and country of origin and that the perpetuation of such behavior across multiple generations may prevent the creation of a melting pot and assimilation of immigrants (Bisin and Verdier, 2000; Cutler, Glaeser and Vigdor, 2008). If this was the case, ELF, as proxy for identity, is more likely to play a positive role in assortative mating of couples than it does in macroeconomics, because homophily would be at work for ELF sorting.

In the BHUS data set it is possible to identify country of origin of parents and grandparents of each respondent. So I assign a particular value of ELF to each individual using the information on family origins. Since the values of the mother are considered to be the most likely to be transmitted to the offsprings (e.g., Ljunge, 2014), I first assign the ELF value corresponding to the country of origin of the mother, then if this information is missing, I assign an ELF value to the individual origin using, in order, the information on the father, the information on the mother's mother and mother's father country of origin. I will also alternatively use only ELF of mother and father to test the importance of ELF inherited from each parent.

The variable I used for education is the maximum level of education attained by each individual ($MaxEdu$). I prefer to use this variable rather than the years of education because it is more indicative of the individual achievements obtained. It ranges from 1 to 3, where 1 indicates the in-

dividual has almost some degree of compulsory education, 2 indicates the individual has some intermediate level of education (i.e., at most a high school degree) and 3 indicates that the individual has a higher education degree or any further higher education. This attribute is generally used in the marriage literature because it is more likely that individuals with similar education level find each other more attractive. While some work in the marriage market uses the income of the two partners to explain matching (Oreffice and Quintana-Domeque, 2010), I follow Dupuy and Galichon (2014) and use education instead.

In addition to the maximum level of education attained, I also use a measure for cognitive skills. There are various measures for cognitive skills in the BHUS data set; I use numeric ability (*Numeric*) because it tests skills in problem-solving that individuals can face in everyday life; it is also related to financial matters: these characteristics make of numeric ability a good attribute to study sorting in marriage. This variable takes values from 1 to 5 and it is a count of the number of problems the respondent has been able to solve. Thus, during the analysis I refer to education as the attribute capturing the tendency of couples with the same education and eventually income to attract themselves; while numeric ability should be taken more as an attribute capturing organizational and problem-solving abilities of the two partners. However, since cognitive skills and maximum education are closely related concepts, I will show that the results hold also when including one of these two variables at a time. The Big 5 inventory and risk taking capture personality characteristics that are often used by the literature to explain behavioral outcomes that cannot be explained by

cognitive skills alone.

Since physical attractiveness is one of the attributes that could also count for the choice of a partner, following the marriage literature I also consider height of the individuals (*Height*, expressed in centimeters) and the value of the Body Mass Index (BMI). This index captures the extent to which an individual is underweight (a $BMI < 18.5$), normal weight ($18.5 < BMI < 24.99$), overweight ($25 < BMI < 29.99$) or obese ($BMI > 30$). In the regression I will use the exact value of BMI (continuos measure) for each individual.[1]

I use the 'Big 5' personality traits as measures of noncognitive skills because they are the personality traits that economists generally borrow from the psychology literature to assess the impact of noncognitive skills on individual behavior (Borghans et al., 2008). These personality traits are grouped in five traits, namely, Conscienciousness, Extraversion, Agreableeness, Neuroticism (or Emotional Stability, depending on how the question is formulated in each survey), and Openness to Experience. These personality traits are, in the BHUS survey as in other surveys, obtained by means of factor analysis using a larger set of questions on personality. They range from 1 to 7, where 1 corresponds to the lowest value and 7 to the highest value of each trait. Conscientiousness (*Consc*) captures the ability to self-discipline, to stay focused of a person and it captures the ability to comply with rules and to plan in advance. Openness to experience (*Open*) measures

[1]While most of the variables used are fixed (e.g., ELF) or available in a single wave (e.g., noncognitive skills), others may have different values over time (i.e., education and BMI). For these attributes I use either the individual average or the maximum value across the available values. This does not affect the results because they are computed using values from 1 to at most 3 consecutive years.

the extent to which a person is open to something new, the need for intellectual stimulation, intelligence and imagination. Extraversion (*Extrav*) catches the degree to which a person interacts with others, the tendency to be involved in social activity and warmth of a person. Neuroticism (here *Neurot*, but it is sometimes called Emotional Stability if the coding has a positive rather than negative connotation) measures anxiety, depression and how well individuals can control their emotions under stress. Finally, Agreeableness (*Agree*) measures the tendency to trust others, altruism, cooperation, and the ability of a person to have harmonious and balanced relations with other individuals. In line with the previous work (Borghans et al., 2008; Dupuy and Galichon, 2014) I also control for propensity to take risk of each person. Thus, I include in the analysis a variable called *Risk* that takes values from 1 to 10 measuring the propensity to take risk of each respondent, where 1 corresponds to the lowest value for propensity and 10 to the highest.

In order to conduct the analysis, I need information for all the variables used for both males and females. I use the first three waves of the BHUS, corresponding to the years 2009-2011; then, I keep the couples for which I have full information about the attributes, and I preserve the first year for each couple. I did not impose age limits when deriving the sample, except that I only consider individuals aged at least 18, so the age range of selected partners is 18-88.[2] Thus, I am left with 2,374 couples, that

[2]I could have run the analysis by including only individuals who were born during and after the 1970s because they could have been exposed to a different, more open, marriage market compared to older individuals. I could also have excluded individuals who are not at their first marriage experience, because it can be questioned that the reasons to marry a second time could be influenced by other factors. However, the sample size would have been limited and this sample reduction may have altered the quality of the

is, 4,748 individuals (however, the number of couples varies slightly with different specifications). This is the sample used along the paper.

4 Results

4.1 Descriptive Statistics and The Affinity Matrix

Table 1 presents descriptive statistics for the characteristics of males and females.

Table 1: Descriptive Statistics for Matched Males and Females

		Males		Females		Range		p-values
	Observations	Mean	Standard Deviation	Mean	Standard Deviation	Minimum	Maximum	
Age	2,374	52.29	14.77	49.94	14.50	19	88	0.000
MaxEdu	2,374	2.18	0.77	2.15	0.78	1	3	0.001
Consc	2,374	5.48	1.04	5.70	1.03	1	7	0.000
Neurot	2,374	3.16	1.36	3.80	1.39	1	7	0.000
Open	2,374	4.66	1.24	4.46	1.31	1	7	0.000
Extrav	2,374	4.42	1.32	4.73	1.30	1	7	0.000
Agree	2,374	5.41	1.03	5.80	0.94	1	7	0.000
Numeric	2,374	4.03	0.96	3.57	0.99	0	5	0.000
Risk	2,374	5.50	2.53	4.73	2.41	0	10	0.000
BMI	2,374	28.40	5.16	27.83	5.56	18.2^m 15.2^w	163.2^m 55.8^w	0.000
Height	2,374	175.18	7.08	161.96	6.26	84.2^m 141.4^w	202.2^m 187.5^w	0.000

Notes: Descriptive statistics for matched males and females. Superscripts m and w indicate ranges for males and females respectively.
Source: BHUS, years 2009, 2010, 2011

The table shows that males are, on average, older than females and females are slightly less educated than males. Females are more conscientious than males, but also more neurotic (less emotionally stable). Males are slightly more open to experience than females, while females are more extraverted than males and also on average more agreeable. As far as cognitive skills are concerned, numeric ability is on average higher for males than for females, and this in part reflects the results of the existing litera-

results. Although I reckon ideally these factors should be taken into account, they should not affect much the sorting based on ELF.

ture according to which males are generally better in maths than females; however, the difference is not high and this supports the result that in more equal societies, such as the United Kingdom, such differences tend to disappear (see Kimura, 2000; Halpern et al., 2007; Guiso et al., 2008, for the literature on gender differences in mathematics). The propensity to take risk is higher for males than for females, and this result also aligns to the existing literature (e.g., Borghans et al., 2009). Also, physical characteristics reflect the usual differences between males and females, since males have higher values of BMI and are overall taller than females. Furthermore, the BMI index indicates that both males and females are on average slightly overweight. Indeed, both males and females have average values above 25 and below 30, which are respectively the thresholds for normal weight and overweight.

Figure 2 shows the map of ELF present in the BHUS sample, based on the country of origin of the respondents's parents and grandparents. The

Figure 2: Ethnolinguistic fractionalization index (Desmet, Ortuño-Ortín and Wacziarg, 2017) in the BHUS sample

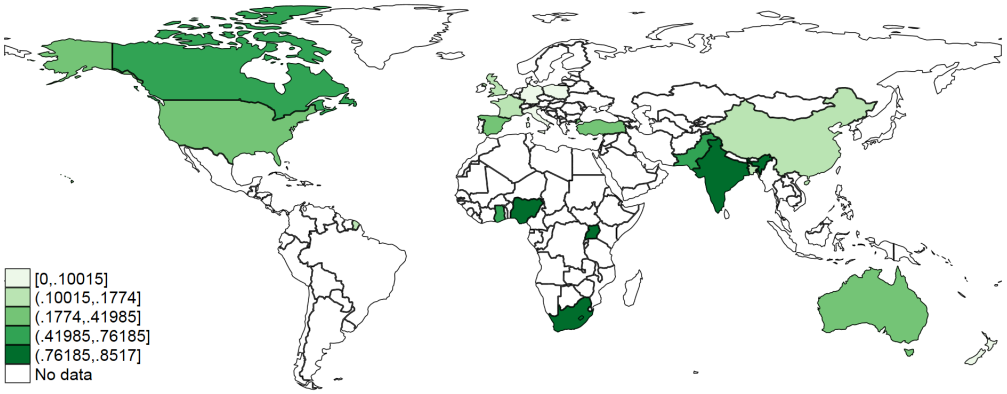

Figure shows that the British sample is comprehensive of individuals with

heterogeneous ELF background, thus, as pointed out earlier in the paper, it is suitable to study the impact of ELF on the selection of the partner and on the determination of joint utility of the couple.

Table 2 reports the estimates of the affinity matrix. The estimation results refer to standardized attributes, as in Dupuy and Galichon (2014); this renders the estimates comparable across attributes. Looking at the on-diagonal estimates, we can notice that ELF is one of the highest coefficients: increasing ELF of both spouses by 1 standard deviation increases the couple's joint utility by 0.16 units. Education has the highest largest coefficient, 0.38, and it is statistically significant too. This last result, indicating the presence of homophily in education, is in line with previous findings (Goussé, Jaquement and Robin, 2017; Chiappori, Salanie and Weiss, 2017); Height and BMI, with a on-diagonal entry of respectively 0.26 and 0.21, are the second and third most important attributes to determine the couple's joint utility, in support to the existing findings (Dupuy and Galichon, 2014). In addition, increasing risk of both spouses by one standard deviation increases the couple's joint utility by 0.12 units. When we look at the noncognitive skills, we notice that increasing openness and conscientiousness by 1 standard deviation also increases the couples' joint utility by 0.12 units for openness and by 0.07 units for conscientiousness: thus, we can infer that there is complementarity in openness to experience and conscientiousness of males and females in the determination of joint utility. Finally, there exists complementarity also in cognitive skills of males and females, since increasing cognitive skills of candidate partners by 1 standard deviation also increases the joint utility of the couple (by 0.15 units).

These results show that ELF is a crucial attribute in marriage sorting and it contributes to positive assortative matching and marriage decisions. It is only less important in magnitude than the maximum level of education and physical characteristics, which is comprehensible since both these sets of attributes are found to be the main determinants of marriage sorting. In accord with the previous literature (Oreffice and Quintana-Domeque, 2010; Dupuy and Galichon, 2014), the findings show that the highest level of education attained, followed by physical characteristics, such as BMI and Height, by ELF, risk and personality traits, such as conscientiousness and openness to experience, and by numerical skills have a direct influence on matching of partners. Some personality traits do not seem to play a direct significant role in positive assortative mating.

Table 2: Estimates of the Affinity Matrix using ELF of family

						Wives					
Husbands	**ELF**	**Agree**	**Neurot**	**Extrav**	**Consc**	**Open**	**Numeric**	**BMI**	**Height**	**Risk**	**MaxEdu**
ELF	0.16***	0.03	-0.03	-0.07	-0.02	0.03	-0.01	-0.03	-0.04	0.05	0.10
	[3.460]	[0.859]	[-0.893]	[-1.499]	[-0.359]	[0.574]	[-0.183]	[-0.375]	[-0.644]	[0.993]	[1.355]
Agree	-0.03	0.02	-0.06**	0.00	-0.04*	0.02	-0.01	0.01	-0.01	-0.07***	-0.00
	[-0.485]	[0.920]	[-2.171]	[0.186]	[-1.870]	[0.683]	[-0.371]	[0.268]	[-0.376]	[-2.999]	[-0.130]
Neurot	-0.01	0.03	-0.01	-0.06**	-0.04*	-0.04*	-0.00	0.01	-0.02	0.05**	0.05*
	[-0.138]	[1.369]	[-0.495]	[-2.263]	[-1.856]	[-1.747]	[-0.171]	[0.309]	[-0.899]	[2.292]	[1.686]
Extrav	-0.01	0.04*	-0.07***	-0.01	0.02	-0.06***	-0.03	-0.01	-0.01	-0.01	-0.06**
	[-0.239]	[1.830]	[-2.708]	[-0.330]	[0.897]	[-2.410]	[-1.027]	[-0.418]	[-0.525]	[-0.390]	[-1.995]
Consc	-0.03	0.02	-0.03	-0.03	0.07***	-0.03	0.02	0.05	0.05*	-0.00	-0.01
	[-0.428]	[0.688]	[-1.151]	[-0.970]	[2.760]	[-0.951]	[0.763]	[1.351]	[1.654]	[-0.056]	[-0.196]
Open	0.06	0.00	0.01	-0.04	-0.06***	0.12***	0.02	-0.07*	0.01	0.05**	0.06*
	[1.092]	[0.145]	[0.462]	[-1.392]	[-2.538]	[4.235]	[0.676]	[-1.919]	[0.487]	[2.101]	[1.690]
Numeric	-0.05	-0.04	-0.02	-0.02	-0.00	-0.04	0.15***	-0.07*	-0.00	0.02	0.06**
	[-0.881]	[-1.534]	[-0.622]	[-0.658]	[-0.079]	[-1.473]	[6.024]	[-1.909]	[-0.154]	[0.838]	[2.031]
BMI	-0.03	0.01	-0.02	-0.03	0.02	-0.02	-0.05***	0.21***	0.02	-0.00	-0.02
	[-0.535]	[0.583]	[-0.879]	[-1.145]	[0.645]	[-1.073]	[-2.514]	[5.041]	[0.597]	[-0.096]	[-0.648]
Height	-0.03	0.01	0.02	0.02	-0.01	0.04	-0.05*	0.04	0.26***	0.02	0.14***
	[-0.472]	[0.232]	[0.717]	[0.632]	[-0.389]	[1.546]	[-1.869]	[1.181]	[8.254]	[0.750]	[4.440]
Risk	0.01	-0.02	0.00	-0.03	-0.01	-0.02	0.03	-0.03	0.06**	0.12***	0.02
	[0.097]	[-0.873]	[0.146]	[-1.196]	[-0.574]	[-0.578]	[1.339]	[-0.818]	[2.276]	[4.162]	[0.567]
MaxEdu	0.06	-0.04	0.04	0.07**	-0.02	0.10***	0.08***	-0.12***	0.02	-0.00	0.38***
	[0.794]	[-1.229]	[1.391]	[2.256]	[-0.635]	[3.258]	[2.419]	[-2.848]	[0.491]	[-0.163]	[9.970]

Notes: *** indicates significance at the 1% level, ** at the 5%, and * at the 10% level. t-statistics are in parentheses.
Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

Nonetheless, all the attributes considered so far, in addition to a di-

rect impact on the joint utility of a couple (i.e., matching of couples by attribute), are likely to interact also with other attributes of the respective partner. So they can indirectly contribute to marriage sorting and to the joint utility of the couple. This can be checked by looking at the significance of the off-diagonal coefficients. The off-diagonal elements for ELF show that while ELF is directly very important to explain mutual attractiveness, it is not indirectly relevant.

Personality traits are also relevant to marriage sorting through their interaction with the other attributes. Males' agreeableness, for instance, negatively correlates with females' neuroticism, their conscientiousness, and risk propensity; this result indicates that having partners who are more emotionally stable, not too rational or self-disciplined and who are not too risk lovers either increases the utility of couples whose husband is relatively more friendly and compassionate. Having a wife who is agreeable increases the utility of a couple whose husband has comparatively higher extraversion.

The results for neuroticism show that increasing extraversion, conscientiosness and openness to experience of wives reduces the joint utility of couples whose husband is relatively less emotionally stable; instead, increasing risk propensity and education of wives for this group of husbands has a beneficial effect on the the joint utility of the couple. The results for wives show that increasing both agreeableness and extraversion of husbands has a negative effect on the joint utility of a couple whose wife is relatively less emotionally stable. In addition, increasing neuroticism, openness to experience and education of wives whose husband is more extraverted by 1 standard deviation reduces the joint utility of the couple by, respectively,

0.07, 0.06 and 0.06 units, while increasing their agreableeness increases the joint utility by 0.5. Neuroticism has a similar impact when we look at the interactions between extraversion of females and their interaction with the attributes of husbands; instead, the maximum level of education has a positive effect. Having a tall wife increases the joint utility of a couple whose husband is relatively more conscientious; agreeableness, neuroticism and openness to experience of husbands have a negative impact on the determination of indirect joint utility of a couple whose wife is more conscientious. For the last trait of the 'Big 5' we find that increasing by 1 standard deviation neuroticism and extraversion of husbands whose wives are relatively more open to experience reduces the joint utility of the couple by, respectively, 0.04 and 0.06, and increasing husband's education increases it by 0.10. Increasing conscientiousness and BMI of wives by 1 standard deviation has a negative impact on the determination of the joint utility of the couple whose husbands are relatively more open to experiences, and increasing risk propensity and education increases it.

The off-diagonal entries of the physical attributes are also significant. BMI of husbands is negatively and significantly correlated with numerical skills of females and BMI of wives is also negatively correlated with openness to experience, numerical skills and education of husbands. Height of husband is positively correlated with maximum education and negatively correlated with numeric ability of the wife. Height of wives is positively correlated with risk propensity and conscientiousness of husbands.

Finally, the maximum level of education of wives positively correlates with husbands' neuroticism, their height, openness to experience and his

numerical skills and negatively correlates with his extraversion; while increasing extraversion, openness to experience and numerical skills of wives increases the joint utility of couples whose husband has a relatively higher education, a higher BMI reduces it.

All in all, the results suggest that most of the attributes contribute to the determination of joint utility of a couple also indirectly.

4.2 Saliency Analysis: The Role of Attributes in the Explanation of the Joint Utility

Saliency analysis is a method proposed by Dupuy and Galichon (2014), which allows to determine the rank of the affinity matrix and the number of attributes on which the marriage sorting decision of couples is based. The method consists in performing a singular value decomposition of the affinity matrix, which for rescaled attributes can be written as:

$$\Theta = S_Y^{1/2} A^{YX} S_X^{1/2} = U' \Lambda V$$

where S_X and S_Y are the diagonal matrices whose terms are, respectively, the variances of the X and Y attributes, $S_X^{-1/2}X$ and $S_Y^{-1/2}Y$ are the rescaled attributes whose entries have unit variance, $\tilde{X} = VS_X^{-1/2}X$ and $\tilde{Y} = US_Y^{-1/2}Y$ are the indices of mutual attractiveness and Λ is a diagonal matrix with nonincreasing elements $(\lambda_1, .., \lambda_d), d = min(d_y, d_x)$ and U and V are orthogonal matrices. As explained in Dupuy and Galichon (2014), the joint utility of a couple can then be written as:

$$\phi_A(y, x) = \Sigma_{i=1}^{d_y} \Sigma_{j=1}^{d_x} A_{ij} y_i x_j = \Sigma_{i=1}^{d} \lambda_i \tilde{y}_i \tilde{x}_i$$

which indicates that the attributes x and y are 'mutually attractive' when $i = j$ (i.e., there is presence of positive assortative matching), which implies that an individual with a higher value for one of the attributes ($\tilde{y}$) tends to match to a partner with higher values of $\tilde{x}$, other things equal. This method is similar to canonical correlation but it has some advantages, one of which is the comparability of the results disregarding the units of the analysis (see Dupuy and Galichon, 2014, for details and proofs).

As it has been previously shown (Dupuy and Galichon, 2014; Kleibergen and Paap, 2006), it is possible to use the asymptotic properties of the affinity matrix to test the null hypothesis that the rank of the affinity matrix is χ^2 distributed with $(d_y - p)(d_x - p)$ degrees of freedom. Thus, I test for the dimensionality of the affinity matrix. The null hypothesis that marriage sorting occurs on p dimensions is rejected for the first 10 out of 11 indices. Indeed, for $p = 1$ the test statistic is equal to 491.83, which is statistically significant at the 1 per cent level, and this is indicative that the choice of the partner depends on multiple attributes. The test remains significant till for $p = 9$, which means that sorting is based at least on 10 attributes. This result aligns to the previous literature pointing out that sorting is multidimensional (Dupuy and Galichon, 2014).

Table 3 reports in Panel A the share of joint utility explained by each of the attributes, and in Panel B the principal component analysis to investigate how much is explained by the indices. Panel A shows that most of the indices are statistically significant; since the attributes explain the totality of the observed matching utility of the couples of the sample, the importance of the remaining indices is minor. Also, the observables explain

the totality of the joint utility, so we can assume that the attribute specific random sympathy shock is negligible; the first four indices explain about the 68 percent of the total utility.

Table 3: Saliency Analysis using ELF of family

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11
Panel A: Share of Joint Utility Explained											
Share of joint utility explained	27.57	17.06	12.51	10.94	9.49	6.92	6.15	4.77	2.39	1.97	0.22
Standard deviation of shares	1.20	1.27	1.23	1.31	1.28	1.38	1.36	2.17	1.48	1.30	1.67
Panel B: Indices of Attractiveness											
	I1 M	**I1 W**	**I2 M**	**I2 W**	**I3 M**	**I3 W**					
ELF	0.23	0.20	0.20	0.22	-0.69	-0.65					
Agree	-0.04	-0.08	0.00	-0.03	0.04	-0.24					
Neurot	0.04	0.10	0.02	-0.04	-0.16	0.08					
Extrav	-0.17	0.08	0.06	-0.04	-0.00	0.32					
Consc	-0.07	-0.10	-0.20	-0.07	0.10	0.16					
Open	0.25	0.27	0.12	-0.01	-0.32	-0.23					
Numeric	0.16	0.19	0.18	0.26	0.54	0.46					
BMI	-0.22	-0.33	-0.44	-0.51	-0.27	-0.26					
Height	0.28	0.16	-0.82	-0.76	0.02	0.12					
Risk	0.10	0.09	-0.08	-0.01	-0.01	-0.22					
MaxEdu	0.83	0.82	0.04	-0.17	0.12	-0.02					

Notes: In Panel A the shares of joint utility explained by the attributes and the standard deviation of each share are reported. I1-I11 indicate the indices created by the singular value decomposition of the affinity matrix. I1-I3 in Panel B indicate the respective indices for males (M) and females (W).

Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

As it is possible to notice in panel B, where I report the first three pairs of indices, the first index, which explains about 28 per cent of the total observed matching utility, loads, similarly for males and females, on the maximum level of education. Instead, the second index loads more on physical characteristics and the third index loads on ELF and cognitive and noncognitive skills of partners. This confirms the importance of identity in the formation of couples and in general remarks the importance of multiple attributes to explain marriage matching.

Thus, saliency analysis shows that, taken all together, the different types of attributes are crucial to understand marriage sorting. Not all the noncognitive skills are equally important in the determination of marriage sorting,

but this is consistent with the literature on personality traits, according to which some traits are more important than others to explain economic and social behavior of individuals (e.g., Borghans et al., 2008).

4.3 Comparing Parents' ELF

In the previous section I have used the information available for parents and grandparents to retrieve the values for ELF. However, I now present the results on the affinity matrix obtained by using either only the information on the mother or only the information on the father. This further analysis is important for a series of reasons. First, it allows me to show that also when the information on ethnicity of a single parent is used the results still hold. Second, by estimating the matrix using either the information on the mother or the information on the father, I can unfold whether identity 'inherited' from one of the two parents is more relevant than the one inherited from the other parent. Third, I can provide a comparison between these findings and the findings in the previous literature (e.g., Dupuy and Galichon, 2014; Ciscato, Galichon and Goussé, 2018; Adda, Pinotti and Tura, 2019). Also, although there exists some degree of heterogeneity in ELF, it can be questioned that variability is not enough and consequently it could alter the estimation results. For this reason I increase the variability in ELF. In the BHPS there is, in the variables for ethnic origins, the possibility to know the ethnic origin at a sub-national level when the ethnicity is British; that is, for each British ethnic origin, it is possible to know if the person or her/his forebears are from: England, Scotland, Wales or Northern Ireland. So I derive from the office of national statistics and in

particular the Census 2011 data for each of these regions the percentage of individuals with white British ethnic origin and weight the value of ELF for each percentage. Since the percentages are very similar across regions and they are very close to unity, it turns out that using the inverse of the percentages does not substantially change the values for the ELF index while increasing, at the same time, the variability of the index itself.

The estimation results for the fathers' and mothers' affinity matrix are reported, respectively, in Table 4 and 5, the share of joint utility explained by each of the attributes are presented in Table 6 and finally the results for the respective principal component analysis are reported in Table 7. For the sake of brevity, from now onwards I will not provide detailed comments of the results, but I will only briefly report the most relevant findings. Also, in Tables 4 and 5 I exclude cognitive skills to make the results more comparable with the results in Dupuy and Galichon (2014).

Table 4 shows that ELF of father is important to directly explain marriage matching and the results are in line with the ones obtained in the affinity matrix where the information on the whole family was used. ELF also contributes indirectly because increasing agreeableness of a wife increases the joint utility of the couple whose husband has a relatively high ELF. This result could mean that having a wife who is more compassionate, understanding, and willing to cooperate may benefit the couple if the husband comes from a family where the father has a relatively high ELF index, so that may have been exposed to ethnic heterogeneity or conflicts. Once again, some personality traits play both a direct and indirect role in the determination of the joint utility of the couple; the same can be said for the

Table 4: Estimates of the Affinity Matrix using ELF of Fathers

	Wives									
Husbands	**ELF**	**Agree**	**Neurot**	**Extrav**	**Consc**	**Open**	**MaxEdu**	**BMI**	**Height**	**Risk**
ELF	0.16***	0.07*	-0.03	-0.06	-0.01	0.03	0.03	-0.02	-0.02	0.05
	[2.921]	[1.878]	[-0.666]	[-1.326]	[-0.164]	[0.558]	[0.621]	[-0.284]	[-0.396]	[1.010]
Agree	-0.04	0.02	-0.06**	0.00	-0.04	0.02	-0.00	0.00	-0.01	-0.06***
	[-0.535]	[1.019]	[-2.312]	[0.019]	[-1.602]	[0.788]	[-0.026]	[0.003]	[-0.191]	[-2.521]
Neurot	-0.02	0.04*	-0.01	-0.06**	-0.06***	-0.03	0.05	0.01	-0.02	0.06***
	[-0.459]	[1.676]	[-0.298]	[-2.324]	[-2.411]	[-1.348]	[1.598]	[0.272]	[-0.662]	[2.433]
Extrav	-0.02	0.05**	-0.07***	-0.01	0.02	-0.05*	-0.07**	-0.02	-0.01	-0.01
	[-0.331]	[2.122]	[-2.769]	[-0.488]	[0.683]	[-1.817]	[-2.234]	[-0.552]	[-0.503]	[-0.456]
Consc	-0.01	0.02	-0.03	-0.03	0.06***	-0.03	-0.01	0.05	0.04	-0.01
	[-0.176]	[0.762]	[-1.158]	[-1.140]	[2.524]	[-0.967]	[-0.406]	[1.489]	[1.554]	[-0.204]
Open	0.04	0.00	0.01	-0.03	-0.05**	0.12***	0.06*	-0.08**	0.01	0.05*
	[0.798]	[0.096]	[0.432]	[-0.953]	[-2.126]	[4.299]	[1.939]	[-2.175]	[0.301]	[1.939]
MaxEdu	0.01	-0.04	0.03	0.07***	-0.01	0.10***	0.44***	-0.14***	0.02	-0.00
	[0.102]	[-1.391]	[0.938]	[2.264]	[-0.404]	[3.169]	[11.682]	[-3.450]	[0.455]	[-0.056]
BMI	-0.02	0.01	-0.02	-0.02	0.01	-0.03	-0.04	0.20***	0.02	0.00
	[-0.374]	[0.586]	[-0.872]	[-0.727]	[0.270]	[-1.379]	[-1.259]	[4.846]	[0.636]	[0.056]
Height	-0.04	0.00	0.01	0.01	-0.02	0.03	0.12***	0.04	0.26***	0.02
	[-0.788]	[0.021]	[0.557]	[0.534]	[-0.687]	[1.242]	[3.988]	[1.125]	[8.320]	[0.930]
Risk	0.03	-0.02	0.00	-0.04	-0.02	-0.02	0.03	-0.03	0.07***	0.12***
	[0.541]	[-0.721]	[0.150]	[-1.433]	[-0.715]	[-0.685]	[0.862]	[-0.752]	[2.548]	[4.285]

Notes: *** indicates significance at the 1% level, ** at the 5%, and * at the 10% level. t-statistics are in parentheses.
Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

other attributes. As expected, the maximum level of education attained plays the major role among the on diagonal direct impacts on marriage sorting.

Table 5 reports the results obtained using the information about the identity of the mother.

ELF has a direct role in sorting of partners and once again is sizeable: increasing ELF of both spouses by 1 standard deviation increases the couple's joint utility by 0.17 units. Some personality traits have both a direct and indirect effect on sorting. The maximum level of education is the most important attribute, followed by physical characteristics and ELF, on the main diagonal. In addition, the interactions between the maximum level of education and the off-diagonal attributes in the column of wives supports

Table 5: Estimates of the Affinity Matrix using ELF of Mothers

	Wives									
Husbands	**ELF**	**Agree**	**Neurot**	**Extrav**	**Consc**	**Open**	**MaxEdu**	**BMI**	**Height**	**Risk**
ELF	0.17***	0.05	-0.03	-0.07	-0.01	0.03	0.05	-0.02	-0.02	0.05
	[2.920]	[1.154]	[-0.804]	[-1.456]	[-0.300]	[0.635]	[0.890]	[-0.299]	[-0.288]	[0.959]
Agree	-0.03	0.02	-0.06**	0.00	-0.04	0.02	-0.00	0.00	-0.01	-0.07***
	[-0.461]	[0.960]	[-2.260]	[0.175]	[-1.534]	[0.665]	[-0.146]	[0.150]	[-0.257]	[-2.913]
Neurot	-0.01	0.04	-0.01	-0.06***	-0.05**	-0.04	0.04	0.01	-0.02	0.05**
	[-0.125]	[1.606]	[-0.226]	[-2.360]	[-2.118]	[-1.632]	[1.478]	[0.441]	[-0.881]	[2.287]
Extrav	-0.01	0.05**	-0.07***	-0.00	0.02	-0.06**	-0.07**	-0.01	-0.02	-0.01
	[-0.291]	[2.045]	[-2.663]	[-0.121]	[0.853]	[-2.206]	[-2.241]	[-0.431]	[-0.603]	[-0.375]
Consc	-0.01	0.02	-0.03	-0.03	0.06***	-0.02	-0.01	0.05	0.04	-0.00
	[-0.147]	[0.973]	[-1.053]	[-1.253]	[2.384]	[-0.869]	[-0.241]	[1.482]	[1.556]	[-0.026]
Open	0.05	-0.00	0.01	-0.04	-0.05**	0.12***	0.07**	-0.08**	0.01	0.05**
	[1.052]	[-0.004]	[0.201]	[-1.328]	[-2.083]	[4.314]	[2.007]	[-2.090]	[0.492]	[1.991]
MaxEdu	0.03	-0.05	0.03	0.06**	-0.01	0.10***	0.44***	-0.14***	0.02	0.00
	[0.375]	[-1.583]	[0.840]	[2.066]	[-0.370]	[3.169]	[11.791]	[-3.446]	[0.659]	[0.005]
BMI	-0.03	0.01	-0.01	-0.02	0.01	-0.03	-0.04	0.20***	0.02	-0.00
	[-0.599]	[0.574]	[-0.783]	[-0.617]	[0.399]	[-1.383]	[-1.281]	[5.046]	[0.675]	[-0.109]
Height	-0.02	0.00	0.02	0.01	-0.02	0.03	0.13***	0.04	0.26***	0.02
	[-0.360]	[0.053]	[0.830]	[0.614]	[-0.674]	[1.349]	[4.191]	[1.149]	[8.299]	[0.848]
Risk	0.01	-0.02	0.00	-0.03	-0.01	-0.02	0.03	-0.03	0.07***	0.12***
	[0.226]	[-0.626]	[0.078]	[-1.346]	[-0.569]	[-0.777]	[0.981]	[-0.757]	[2.338]	[4.307]

Notes: *** indicates significance at the 1% level, ** at the 5%, and * at the 10% level. t-statistics are in parentheses.
Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

the presence of interesting patterns across variables. As a matter of fact, we notice that increasing openness to experience of husbands increases the joint utility of a couple whose wife has relatively higher education. Said in other words, a wife with relatively better education could benefit from marrying a husband that could possibly be a source of destabilization for the couple (with openness to experience being the main cause of it). This finding is in line with the results in Tables 2 and 4.

If we compare the two affinity matrices we notice that, despite the presence of similarities, they somehow differ. In particular, the on-diagonal entry for the ELF index obtained with the information about the mother is slightly higher: increasing ELF of both partners by 1 standard deviation increases the joint utility of the couple by 0.16 when using the ELF information of fathers, but 0.17 when using the ELF information on mothers.

This result could be indicative, in accord with the previous literature, of a higher direct importance of identity and values inherited from the mother than those inherited from the father (e.g., Ljunge, 2014). However, while the ELF index is only directly relevant to explain the joint utility of the couple in Table 5, when considering the ELF of fathers (Table 4) it becomes significant also in the interaction with wives' agreeableness, maybe because one of the values transmitted by the father is that a good wife should be understanding, compassionate and cooperative. Overall, the results suggest that values inherited from either parents are likely to be good predictors when explaining the actions and interactions of the offsprings, with a slightly higher direct impact using the ELF of mothers. This finding is valid not only to explain actions of children, but also other decisions, such as marriage, indicating that values and beliefs inherited in the past may still influence lifetime decisions made at a later stage of life.

In Tables 6 and 7 I report respectively the shares and the loads relative to the two affinity matrices for fathers and mothers (Tables 4 and 5).

Table 6: Saliency Analysis using ELF of parents: Share of Joint Utility Explained

	Father's ELF									
	I1	**I2**	**I3**	**I4**	**I5**	**I6**	**I7**	**I8**	**I9**	**I10**
Share of joint utility explained	30.47	17.59	13.40	9.75	7.74	7.38	6.98	3.13	2.36	1.21
Standard deviation of shares	1.29	1.37	1.33	1.42	1.39	1.49	1.72	2.29	1.61	1.41
	Mother's ELF									
Share of joint utility explained	31.07	17.30	13.44	9.88	8.08	7.26	6.04	3.35	2.42	1.15
Standard deviation of shares	1.32	1.36	1.32	1.40	1.37	1.48	1.71	2.31	1.59	1.40

Notes: The shares of joint utility explained by the attributes and the standard deviation of each share are reported. I1-I10 indicate the indices created by the singular value decomposition of the affinity matrix for males (M) and females (W).

Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

The attributes explain the totality of the joint utility, and the first four

indices explain around the 71 per cent (in both cases) of it.

Table 7 shows that the first index loads on the maximum level of education attained, disregarding the use of mother or father information on ELF; the second index loads again mainly on physical attributes (and partially on personality traits) and finally the third index loads on ELF, personality traits and risk propensity. In conclusion, the results using the information on each of the two parents' ELF are overall in line with the initial findings.

Finally, since the attributes used in my analysis are similar to the ones used by Dupuy and Galichon (2014), I can compare the results obtained in the two studies. I then compare table 5, which is the one using the ELF of mothers, to Table 3 in Dupuy and Galichon (2014). As far as the on-diagonal estimates are concerned, we find that both the maximum level of education attained and physical attributes are the ones with the highest estimates, as in Dupuy and Galichon (2014). The point estimates of physical attributes are also similar in size to the ones obtained by Dupuy and Galichon (2014), while education has a smaller (but still sizeable) impact for the British sample. Conscientiousness and risk propensity also remain significant when using the British sample, and the estimated parameters are also similar in magnitude. As in Dupuy and Galichon (2014), I also find that some personality traits do not have a direct impact in marriage sorting, however, conscientiousness still maintains its significance using the BHUS; they do not use openness to experience, which is the other noncognitive skill that has a significant direct impact in my matrix, so this finding cannot be compared. The interactions across attributes (i.e., the indirect impact) are more significant in the British sample. So, overall the results are similar to

Table 7: Indices of Attractiveness for Males and Females using ELF of parents

	Father's ELF						Mother's ELF					
	I1 M	**I1 W**	**I2 M**	**I2 W**	**I3 M**	**I3 W**	**I1 M**	**I1 W**	**I2 M**	**I2 W**	**I3 M**	**I3 W**
ELF	0.07	0.05	0.21	0.23	0.68	0.58	0.13	0.11	0.21	0.23	-0.69	-0.57
Agree	-0.01	-0.08	0.02	0.02	-0.18	0.23	-0.03	-0.09	0.02	0.00	0.19	-0.20
Neurot	0.04	0.09	-0.01	-0.02	0.20	-0.04	0.03	0.08	0.02	-0.05	-0.22	0.07
Extrav	-0.16	0.11	0.07	0.00	0.02	-0.40	-0.16	0.08	0.07	-0.03	-0.01	0.46
Consc	-0.08	-0.08	-0.22	-0.03	0.01	-0.15	-0.06	-0.07	-0.22	-0.05	-0.06	0.15
Open	0.23	0.26	0.15	0.09	0.36	0.10	0.24	0.26	0.15	0.08	-0.38	-0.09
MaxEdu	0.88	0.87	0.12	-0.07	-0.26	-0.12	0.87	0.86	0.10	-0.09	0.31	0.15
BMI	-0.21	-0.34	-0.39	-0.50	-0.05	-0.10	-0.21	-0.33	-0.44	-0.54	0.04	0.04
Height	0.28	0.16	-0.83	-0.82	0.12	0.20	0.28	0.18	-0.81	-0.80	-0.14	-0.21
Risk	0.09	0.07	-0.17	-0.08	0.50	0.60	0.10	0.08	-0.15	-0.06	-0.42	-0.56

Notes: I1-I3 indicate the respective indices for males (M) and females (W).

Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

those obtained in Dupuy and Galichon (2014) although the estimated parameters cannot be directly compared because we include slightly different attributes; also, the interactions that determine the significance of the indirect effect of attributes on the joint utility of a couple are more significant for the British sample and they unfold very interesting patterns.

Furthermore, although I cannot compare the results on identity with Dupuy and Galichon (2014) because they did not include this attribute in their study, I can compare my results to the existing literature that does it. My results cannot be compared with the work in Adda et al. (2019), given the specificity of their analysis. Instead, as in Ciscato, Galichon and Goussé (2018) I find evidence of homophily in identity, but identity is more important in Ciscato et al. (2018) than in my analysis, since the gain in the joint utility of the couple is much higher than the one I find here. However, their results could be specific to California; also, while they let marriage sorting be based only on a few attributes, I include a larger set of attributes in the analysis. Thus, the comparison between the two studies cannot be exhaustive. I also dropped cognitive and noncognitive skills and re-estimate the matrix to let sorting be based on a similar set of attributes as the one used in Ciscato et al. (2018), but my on-diagonal entries are still smaller than the ones they obtained on the different-sex couples; the estimation results, dropped for space reason, are available upon requested to the interested reader. Finally, as aforementioned, the results of these two studies cannot be generalized because they are obtained on very specific samples. Consequently, the divergence of my results with the findings in these two studies may be due to the specificity of their samples or the choice

of a more restricted set of attributes they used.

4.4 Assessing the Impact of Cognitive and Noncognitive Skills

In this subsection I present the estimate of the affinity matrix using the mother information for the ELF index, and I include both cognitive and noncognitive skills (I exclude the maximum level of education). This result permits to get insights about the impact of both cognitive and noncognitive skills in marriage sorting. In so doing I follow the literature assessing the impact of these two sets of skills on lifetime outcomes (e.g., Heckman, Stixrud and Urzua, 2006; Borghans et al., 2008, 2009). Table 8 reports the estimation results.

Table 8: Estimates of the Affinity Matrix: Mother's ELF, Cognitive and Noncognitive Skills

						Wives				
Husbands	**ELF**	**Agree**	**Neurot**	**Extrav**	**Consc**	**Open**	**Numeric**	**BMI**	**Height**	**Risk**
ELF	0.17***	0.05	-0.03	-0.06	-0.02	0.03	0.01	-0.02	-0.01	0.05
	[2.734]	[1.174]	[-0.791]	[-1.404]	[-0.351]	[0.757]	[0.133]	[-0.302]	[-0.177]	[1.059]
Agree	-0.03	0.02	-0.06***	0.00	-0.04	0.01	-0.01	0.01	-0.01	-0.07***
	[-0.553]	[0.969]	[-2.363]	[0.107]	[-1.502]	[0.579]	[-0.314]	[0.177]	[-0.282]	[-2.926]
Neurot	-0.01	0.03	-0.00	-0.06**	-0.05**	-0.03	0.02	-0.00	-0.02	0.06***
	[-0.124]	[1.121]	[-0.189]	[-2.266]	[-2.271]	[-1.294]	[0.783]	[-0.022]	[-0.666]	[2.583]
Extrav	-0.02	0.06**	-0.07***	-0.01	0.03	-0.07***	-0.05**	-0.01	-0.02	-0.01
	[-0.478]	[2.311]	[-2.893]	[-0.368]	[1.066]	[-2.637]	[-2.164]	[-0.202]	[-0.806]	[-0.594]
Consc	-0.02	0.03	-0.03	-0.03	0.06***	-0.03	0.02	0.05	0.04	-0.00
	[-0.292]	[0.987]	[-1.063]	[-1.313]	[2.371]	[-1.055]	[0.594]	[1.559]	[1.463]	[-0.105]
Open	0.07	-0.01	0.02	-0.02	-0.06***	0.15***	0.06**	-0.10***	0.02	0.06**
	[1.556]	[-0.495]	[0.584]	[-0.861]	[-2.404]	[5.196]	[2.205]	[-2.668]	[0.882]	[2.281]
Numeric	-0.05	-0.06**	-0.00	0.01	-0.02	0.01	0.20***	-0.10***	0.01	0.03
	[-0.888]	[-2.379]	[-0.191]	[0.223]	[-0.857]	[0.289]	[8.139]	[-3.034]	[0.354]	[1.310]
BMI	-0.03	0.02	-0.02	-0.02	0.02	-0.03	-0.08***	0.22***	0.02	-0.01
	[-0.676]	[0.810]	[-1.205]	[-1.008]	[0.665]	[-1.766]	[-3.607]	[5.232]	[0.622]	[-0.222]
Height	-0.01	-0.01	0.02	0.02	-0.02	0.06**	0.00	0.02	0.27***	0.03
	[-0.211]	[-0.538]	[0.980]	[0.909]	[-0.965]	[2.316]	[0.159]	[0.684]	[8.825]	[1.365]
Risk	0.02	-0.02	0.01	-0.03	-0.02	-0.01	0.05**	-0.04	0.07***	0.13***
	[0.366]	[-0.891]	[0.313]	[-1.084]	[-0.781]	[-0.474]	[2.213]	[-1.046]	[2.491]	[4.445]

Notes: *** indicates significance at the 1% level, ** at the 5%, and * at the 10% level. t-statistics are in parentheses.
Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

Once again, for the sake of brevity I do not provide detailed comments. Numeric ability is the attribute with the highest direct impact on the joint utility of the couple, together with physical characteristics, and plays both a direct and indirect role in marriage sorting. ELF is the third most important attribute in the determination of the joint utility of a couple and once again it is only directly relevant. In line with previous findings and the literature, some personality traits contribute both directly and indirectly, while others only have an indirect influence on matching. All the other attributes, with the exception of ELF, impact both directly and indirectly the utility of the couple. Finally, findings on cognitive skills are interesting: they have a direct impact on sorting, indicating that partners with similar numerical and organizational skills mutually attract each other; but they also have an indirect impact, confirmed through the interaction with openness to experience, extraversion and risk of husbands, which respectively increases, reduces and increases the joint utility of a couple whose wife is relatively more endowed of numerical skills. This finding, in particular the positive contribution of the interaction with openness to experience and risk of husband, suggests the inclination of 'relatively more rational' women to be attracted by men with opposite characteristics (if numerical skills can be linked to organization and the ability to deal with problem solving, openness to experience and risk propensity may act as destabilizer and be the source of problems). This result aligns to the findings obtained in the previous tables.

4.5 Sensitivity Analysis

Finally, as last check I report the results using as starting values to compute the affinity matrix of Table 5 (Mothers) the estimated parameters obtained in the original article by Dupuy and Galichon (2014), when applicable; where not possible, I left zeros as starting values.

Table 9: Estimates of the Affinity Matrix using the information in Dupuy and Galichon (2014)

					Wives					
Husbands	**ELF**	**Agree**	**Neurot**	**Extrav**	**Consc**	**Open**	**MaxEdu**	**BMI**	**Height**	**Risk**
ELF	0.17***	0.05	-0.03	-0.07	-0.01	0.03	0.05	-0.02	-0.02	0.05
	[2.919]	[1.154]	[-0.804]	[-1.456]	[-0.300]	[0.635]	[0.890]	[-0.299]	[-0.288]	[0.960]
Agree	-0.03	0.02	-0.06**	0.00	-0.04	0.02	-0.00	0.00	-0.01	-0.07***
	[-0.461]	[0.960]	[-2.260]	[0.175]	[-1.534]	[0.665]	[-0.146]	[0.150]	[-0.257]	[-2.913]
Neurot	-0.01	0.04	-0.01	-0.06***	-0.05**	-0.04	0.04	0.01	-0.02	0.05**
	[-0.125]	[1.606]	[-0.226]	[-2.360]	[-2.118]	[-1.632]	[1.478]	[0.441]	[-0.881]	[2.286]
Extrav	-0.01	0.05**	-0.07***	-0.00	0.02	-0.06**	-0.07**	-0.01	-0.02	-0.01
	[-0.291]	[2.045]	[-2.663]	[-0.121]	[0.853]	[-2.206]	[-2.241]	[-0.431]	[-0.603]	[-0.376]
Consc	-0.01	0.02	-0.03	-0.03	0.06***	-0.02	-0.01	0.05	0.04	-0.00
	[-0.147]	[0.973]	[-1.053]	[-1.253]	[2.384]	[-0.869]	[-0.241]	[1.482]	[1.556]	[-0.026]
Open	0.05	-0.00	0.01	-0.04	-0.05**	0.12***	0.07**	-0.08**	0.01	0.05**
	[1.052]	[-0.004]	[0.201]	[-1.328]	[-2.083]	[4.314]	[2.007]	[-2.090]	[0.492]	[1.991]
MaxEdu	0.03	-0.05	0.03	0.06**	-0.01	0.10***	0.44***	-0.14***	0.02	0.00
	[0.375]	[-1.583]	[0.840]	[2.066]	[-0.370]	[3.169]	[11.791]	[-3.446]	[0.659]	[0.005]
BMI	-0.03	0.01	-0.01	-0.02	0.01	-0.03	-0.04	0.20***	0.02	-0.00
	[-0.599]	[0.574]	[-0.783]	[-0.617]	[0.399]	[-1.383]	[-1.281]	[5.046]	[0.675]	[-0.109]
Height	-0.02	0.00	0.02	0.01	-0.02	0.03	0.13***	0.04	0.26***	0.02
	[-0.360]	[0.053]	[0.830]	[0.614]	[-0.674]	[1.349]	[4.191]	[1.149]	[8.299]	[0.848]
Risk	0.01	-0.02	0.00	-0.03	-0.01	-0.02	0.03	-0.03	0.07***	0.12***
	[0.226]	[-0.626]	[0.078]	[-1.346]	[-0.569]	[-0.777]	[0.981]	[-0.757]	[2.338]	[4.307]

Notes: *** indicates significance at the 1% level, ** at the 5%, and * at the 10% level. t-statistics are in parentheses.
Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

This final check is useful to show that the results are not sensitive to the change of starting values. The findings, presented in Table 9, are substantially unchanged if we compare them to the results obtained in Table 5, so we can conclude that the results are not sensitive to the use of alternative starting values.

The shares and the loadings are reported in Table 10.

Table 10: Saliency Analysis using the information in Dupuy and Galichon (2014)

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
	Panel A: Share of Joint Utility Explained									
Share of joint utility explained	31.07	17.30	13.44	9.88	8.08	7.26	6.04	3.35	2.42	1.15
Standard deviation of shares	1.32	1.36	1.32	1.40	1.37	1.48	1.71	2.31	1.59	1.40
	Panel B: Indices of Attractiveness									
	I1 M	**I1 W**	**I2 M**	**I2 W**	**I3 M**	**I3 W**				
ELF	0.13	0.11	0.21	0.23	-0.69	-0.57				
Agree	-0.03	-0.09	0.02	0.00	0.19	-0.20				
Neurot	0.03	0.08	0.02	-0.05	-0.22	0.07				
Extrav	-0.16	0.08	0.07	-0.03	-0.01	0.46				
Consc	-0.06	-0.07	-0.22	-0.05	-0.06	0.15				
Open	0.24	0.26	0.15	0.08	-0.38	-0.09				
MaxEdu	0.87	0.86	0.10	-0.09	0.31	0.15				
BMI	-0.21	-0.33	-0.44	-0.54	0.04	0.04				
Height	0.28	0.18	-0.81	-0.80	-0.14	-0.21				
Risk	0.10	0.08	-0.15	-0.06	-0.42	-0.56				

Notes: In Panel A the shares of joint utility explained by the attributes and the standard deviation of each share are reported. I1-I10 indicate the indices created by the singular value decomposition of the affinity matrix. I1-I3 in panel B indicate the respective indices for males (M) and females (W).

Source: BHUS, years 2009, 2010, 2011 and Desmet, Ortuño-Ortín and Wacziarg (2017).

Again, the shares reported in Panel A show that the indices altogether explain the totality of the joint utility; looking at the indices of attractiveness presented in Panel B, the first indices load on maximum education, the second ones load more on physical characteristics and the third ones load on ELF and noncognitive skills, so they also reflect the previous findings. Thus, we can conclude that neither the affinity matrix nor saliency analysis are affected by the change of starting values.

5 Discussion and Conclusions

5.1 Discussion

Some interesting features emerge from the empirical analysis. First, males and females base their marriage decision on different characteristics; the

lack of symmetry in the affinity matrix indicates that there exists a certain degree of heterogeneity, as pointed out by the literature (e.g., Heckman, Stixrud and Urzua, 2006; Borghans et al., 2008, 2009) in decision making across gender, and that, when choosing a partner, the determinants of such choice are different for the candidates husbands and wives (e.g., Dupuy and Galichon, 2014). Thus, for instance, while risk of females is positively correlated with neuroticism, openness to experience, and risk propensity of males, risk of males is positively correlated with risk of females, their height, and sometimes their cognitive skills; this finding suggests that, while females who are more likely to take risk look for (and benefit from) husbands that are less emotionally stable, more imaginative, intelligent and open to experience, males prone to take risk are more likely to search for and benefit from wives that have high cognitive skills. Said in other words, while females tend to look for males with similar characteristics to their own one, males are attracted by females who are also risk lovers (showing some complementarity in this attribute), but may prefer to start a family with (or, better, the couple benefits more from) wives whose characteristics counterbalance and compensate their own attributes. Again, while wives with relatively high cognitive skills or maximum level of education seem to benefit from marrying husbands with characteristics that can be seen as destabilizers (opennes to experience, risk propensity, and neuroticism) this pattern is not well delineated for husbands. While I do not intend to be exhaustive on this finding here, it suggests that wives and husbands benefit from different characteristics of the partner.

Then, some interactions between skills and physical characteristics should

be remarked: numerical skills are negatively correlated with BMI of the partner, and this finding may be indicative of the awareness of people with relatively high cognitive skills of the negative consequences of being overweight or obese.

Furthermore, the analysis of complement and substitute attributes is also relevant to understand the 'natural' formation of couples. Sorting based on ELF explains the presence of homophily in identity and the absence of a melting pot. Results on ELF also indicate that identity transmitted from each of the two parents may be slightly different. Looking at personality traits the results suggest that while some similarity is needed in the matching for certain psychological traits, other traits can only be substitutes. For example, while we expect that individuals that are similarly conscientious or open to experience are more likely to form a couple, we would expect that if one of the partners is less emotionally stable the other should be more conscientious in order to reach and maintain a steady equilibrium. A 'natural' degree of complementarity in physical characteristics is another outcome: we expect that individuals with similar characteristics are attractive to each other.

In sum, while some attributes show that similar people attract each other ('wives and oxes of your own places', 'better to marry a neighbor than a stranger', 'birds of a feather flock together'), they also show that some degree of dissimilarity or difference in the attributes of the couple is needed to reach a balance in the couple ('opposites attract') and that sorting is heterogeneous in gender.

Finally, the similarity between the results I obtain here on a British sam-

ple and the findings obtained by Dupuy and Galichon (2014) on a Dutch sample encourage cross-country comparisons; thus, while the results presented here differ from the only other two studies assessing the importance of identity in marriage matching, the release of new evidence obtained on national representative samples could shed light on the relevance of identity on marriage sorting in multicultural societies.

5.2 Conclusions

Marriage is one of the simplest forms of social interactions individuals experience over their own life, but it is also one of the most important means they use to strengthen and transmit their values, norms, and identity; besides, it is related to an individual and a couple's well-being and happiness (Stutzer and Frey, 2006; Zimmermann and Easterlin, 2006; Dupuy and Galichon, 2014). Thus, understanding what determines marriage sorting is crucial to explain individual behaviors, the ability to manage wealth and to decide on investments for both partners and their children (Chiappori, Salanie and Weiss, 2017) as well as the intergenerational transmission of values, norms, and habits (which are driven by identity and may relate to other outcomes such as homotetic preferences for marriage and social and racial segregation) and the transmission of cognitive and noncognitive skills (Grönqvist, Ökert and Vlachos, 2017).

In this paper I estimate a structural model of marriage sorting using the BHUS, a representative data set of British individuals, which allows me to study matching in marriage across couples living in the United Kingdom.

The results show that ELF is among the attributes that most contribute

to the determination of marriage sorting and to the explanation of the observed joint utility of the couple. This finding confirms the importance of identity to explain marriage matching, in accord with the literature supporting the prevalence of homogamous marriages (Bisin and Verdier, 2000; Bisin et al., 2016).

It also aligns to the 'identity economics' (Akerlof and Kranton, 2000, 2010; Akerlof, 2016; Carvalho, 2016) literature and suggests that individuals prefer to socialize and form a group (a family) with other individuals similar to them: while the results cannot convey if higher or lower degrees of ethnolinguistic fractionalization are better (recall, we cannot identify absolute attractiveness), they suggest that individuals with similar values of ethnolinguistic fractionalization *mutually prefer* each other. This evidence is important because the formation of a group tends to reinforce, due to spillovers between partners and the generation of a group pride (Carvalho, 2016), norms, values, and identity itself; these features of resilience and homophily may explain the rise of phenomena such as the absence of a 'melting pot' in heterogeneous societies (Cutler, Glaeser and Vigdor, 2008) and the radicalization and stereotyping of values and cultural traits (Carvalho, 2016). Furthermore, it shows that while ELF may have negative consequences for macroeconomic outcomes, such as provision of public goods and civil conflicts, it reinforces a family identity and contributes to the joint utility and happiness of a couple. In addition, ELF, although relevant, is less important than the maximum level of education/cognitive skills and physical characteristics; this result is consistent with the reality of marriage sorting. Finally, another interesting finding on ELF is that by

comparing affinity matrices computed using the ELF of the mother and of the father it emerges that while the identity inherited from the mother is slightly more relevant in positive assortative matching and in the determination of the joint utility of a couple, the identity inherited from the father suggests that having a wife who is more agreeable, willing to compromise and to cooperate is beneficial for husbands with relatively higher levels of ELF.

Other findings emerge from the study. First, in accord with the previous literature (e.g., Goussé, Jaquement and Robin, 2015*b*; Dupuy and Galichon, 2014), education plays the most important role in marriage sorting. Second, personality traits are also relevant to understand who marries whom: conscientiousness and openness to experience are both directly and indirectly relevant for marriage sorting, and extraversion, neuroticism and agreeableness are relevant when we look at their interaction with the other attributes, i.e., they play an indirect role. This result is consistent with the previous literature (e.g., Borghans et al., 2008) on the 'Big 5' personality traits suggesting some of these traits are more relevant than others in the prediction of individual behavior. Third, risk propensity also enters marriage sorting decisions and has both a direct and indirect role in the determination of a partnership. Fourth, cognitive skills are important, and they show that while numerical skills of both partners are directly relevant to explain the joint utility of a couple, organizational skills of wives are more likely to indirectly significantly affect the marriage sorting decisions and the joint utility of a couple than the organizational skills of candidate husbands, suggesting the presence of role differences in the choice and formation of

couples. Fifth, in line with this result, the asymmetry of the affinity matrix shows that preferences differ for males and females. The findings, in accord with the previous literature on marriage (Dupuy and Galichon, 2014) and the behavioral literature (e.g., Borghans et al., 2008, 2009), show that while females tend to be attracted and prefer males with similar characteristics to their own, males generally prefer to choose wives whose characteristics counterbalance their own ones. Also, I estimate a model similar to the one in Dupuy and Galichon (2014) on a data set other than the one they used in their analysis and I provide both evidence of external validity of the technique they developed and a sensitivity analysis.

All in all this paper shows that identity, physical attributes, cognitive and noncognitive skills all play a role and should be considered when studying assortative mating and marriage sorting; also, it is in line with the recent literature (Akerlof and Kranton, 2010; Akerlof, 2016; Carvalho, 2016) pointing out that more attention should be devoted to the importance of identity to explain not only individual behavior and decisions, but also social interactions, the formation of groups and the transmission of values and norms across generations. While this study refers to marriage sorting, these attributes may also explain marriage dynamics as well as lifetime outcomes also of children through their intergenerational transmission; this is left for future work.

A Appendix

Table A1 reports the list of variables, their definitions and sources.

Table A1: List of Variables, Definitions and Source

Variables	Definition	Source
Age	age of respondent (r henceforth)	BHUS, `'w'_dvage`
ELF	index for ethnolinguistic fractionalization. It ranges from 0 to 1, with 1 indicating an extreme rate of ethnolinguistic fractionalization.	Desmet, Ortuño-Ortín and Wacziarg (2017)
Agree	indicator ranging from 1 to 7 (where 7 is the highest) that measures r agreeableness	BHUS, `c_big5a_dv`
Neurot	indicator ranging from 1 to 7 (where 7 is the highest) that measures r neuroticism	BHUS, `c_big5n_dv`
Extrav	indicator ranging from 1 to 7 (where 7 is the highest) that measures r extraversion	BHUS, `c_big5e_dv`
Consc	indicator ranging from 1 to 7 (where 7 is the highest) that measures r conscientiousness	BHUS, `c_big5c_dv`
Open	indicator ranging from 1 to 7 (where 7 is the highest) that measures r openness to experience	BHUS, `c_big5o_dv`
Numeric	indicator ranging from 1 to 5 (where 5 is the highest) that measures r numeric ability. It is the count of correct answers to questions asked to test numeric ability. It tests skills in everyday problem-solving and financial abilities.	BHUS, `c_cgna_dv`
BMI	it measures r's body mass index. The value for each individual is used (continue measure). In general, BMI indicates if the r is underweight (<18.5), normal weight ($18.5 < BMI < 24.99$), overweight ($25 < BMI < 29.99$), or obese ($BMI >= 30$)	BHUS, `bmival`
Height	variable taking the value of the r height in centimeters	BHUS, `'w'_height`
Risk	indicator measuring the propensity to risk of the r. It ranges from 0 to 10 (where 10 is the highest).	BHUS, `a_scriska`
MaxEdu	indicator measuring the highest level of education attained by the r. 1 means the r has at most a compulsory level of education, 2 indicates the r has an intermediate level of education (at most a high school education degree), and 3 indicates the r has a higher education or further education degree.	BHUS, `'w'_hiqual_dv`

References

Adda, J., P. Pinotti, and G. Tura. 2019. "There's More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations." mimeo.

Akerlof, G. A., and R. E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Akerlof, G. A., and R. E. Kranton. 2010. *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton University Press.

Akerlof, R. 2016. "'We Thinking' and Its Consequences." *American Economic Review*, 106(5): 415–419.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth*, 8(2): 155–194.

Anderberg, D. 2004. "Getting Hitched: The Equilibrium Marriage Market Behavior of a British Cohort." Discussion Paper, Royal Holloway University of London.

Becker, G. S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Becker, G. S. 1974. "A Theory of Marriage: Part II." *Journal of Political Economy*, 82(2): S11–S26.

Bénabou, R., and J. Tirole. 2006. "Beliefs in a Just World and Redistributive Politics." *The Quarterly Journal of Economics*, 121(2): 699–746.

Ben-Akiva, M., and T. Watanatada. 1981. "Application of a Continuous Spatial Choice Logit Model in Structural Analysis of Discrete Data

with Econometric Applications." *Structural Analysis of Discrete Data and Econometric Applications*, Chapter 8, 320–343. MIT Press, Cambridge, MA.

Ben-Akiva, M., N. Litinas, and K. Tsunekawa. 1985. "Continuous Spatial Choice: The continuous logit model and distribution of trips and urban densities." *Transportation Research Part A: General*, 19(2): 119–154.

Bertrand, M., E. Kamenica, and J. Pan. 2015. "Gender Identity and Relative Income within Households." *The Quarterly Journal of Economics*, 130: 571–614.

Bisin, A., and T. Verdier. 2000. "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits." *The Quarterly Journal of Economics*, 115(3): 955–988.

Bisin, A., and T. Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory*, 97(2): 298–319.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou. 2016. "Bend it like Beckham: Ethnic identity and integration." *European Economic Review*, 90(C): 146–164.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel. 2008. "The Economics and Psychology of Personality Traits." *the Journal of Human Resources*, 43(4): 972–1059.

Borghans, L., B. H. Golsteyn, J. J. Heckman, and H. Meijers. 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association*, 7(2/3): 649–658.

Carvalho, J.-P. 2016. "Identity-based Organizations." *American Economic Review*, 106(5): 410–414.

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." *Journal of Political Economy*, 116(1): 113–140.

Chiappori, P.-A., B. Salanie, and Y. Weiss. 2010. "Assortative Matching on the Marriage Market: A Structural Investigation." working paper, Columbia University.

Chiappori, P.-A., B. Salanie, and Y. Weiss. 2017. "Partner Choice, Investment in Children, and the Marital College Premium." *American Economic Review*, 107(8): 2109–2167.

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque. 2012. "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market." *Journal of Political Economy*, 120(4): 659–695.

Choo, and Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.

Choo, E., and A. Siow. 2007. "Lifecycle marriage matching: Theory and evidence." Society for Economic Dynamics, 2007 meeting paper.

Ciscato, E., A. Galichon, and M. Goussé. 2018. "Like Attract Like? A Structural Comparison of Homogamy Across Same-sex and Different-sex Households." Available at SSRN: https://ssrn.com/abstract=2530724 or http://dx.doi.org/10.2139/ssrn.2530724.

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

Cunha, F., and J. J. Heckman. 2008. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Journal of Human Resources*, 43(4): 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica*, 78(3): 883–931.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor. 2008. "Is the Melting Pot Still Hot? Explaining the Resurgence of Immigrant Segregation." *Review of Economics and Statistics*, 90: 478–497.

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

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg. 2017. "Culture, Ethnicity, and Diversity." *American Economic Review*, 107(9): 2479–2513.

Dupuy, A., and A. Galichon. 2014. "Personality Traits and The Marriage Market." *Journal of Political Economy*, 122(6): 1271–1319.

Eatserly, W., and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *The Quarterly Journal of Economics*, 112(4): 1203–1250.

Goussé, M., N. Jaquement, and J.-M. Robin. 2015*a*. "Cupid's Invisible Hand: Social Surplus and Identification in Match-

ing Models." Available at SSRN: http://ssrn.com/abstract=1804623 or http://dx.doi.org/10.2139/ssrn.1804623.

Goussé, M., N. Jaquement, and J.-M. Robin. 2015*b*. "Household labour supply and the marriage market in the UK, 1991-2008." *Labour Economics*, 46(C): 131–149.

Goussé, M., N. Jaquement, and J.-M. Robin. 2017. "Marriage, Labor Supply, and Home Production." *Econometrica*, 85(6): 1863–1919.

Grönqvist, E., B. Ökert, and J. Vlachos. 2017. "The Intergenerational Transmission of Cognitive and Noncognitive Abilities." *Journal of Human Resources*, 52(4): 887–918.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales. 2008. "Culture, Gender, and Maths." *Science*, 320: 1164–1165.

Halpern, D. F., C. P. Benbow, D. C. Geary, R. C. Gur, J. S. Hyde, and M. A. Gernsbacher. 2007. "The Science of Sex Differences in Science and Mathematics." *Psychological Science in the Public Interest*, 8(1): 1–51.

Heckman, J. J., J. Stixrud, and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

Kimura, D. 2000. *Sex and Cognition.* MIT, Boston.

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

Ljunge, M. 2014. "Trust issues: evidence on the intergenerational trust transmission among children of immigrants." *Journal of Economic Be-*

havior and Organization, 106: 175–196.

McFadden. 1976. "The Mathematical Theory of Demand Models." *Behavioral Travel-Demand Models*, ed. P. Stopher and A. Meyburg, Chapter 17, 305–314. Lexington, MA.

Oreffice, S., and C. Quintana-Domeque. 2010. "Antropometry and Socio-Economics among Couples: Evidence in the United States." *Economics and Human Biology*, 8(3): 373–384.

Stutzer, A., and B. S. Frey. 2006. "Does marriage make people happy, or do happy people get married?" *The Journal of Socio-Economics*, 35(2): 326–347.

Zimmermann, A. C., and R. A. Easterlin. 2006. "Happily Ever After? Cohabitation, Marriage, Divorce, and Happiness in Germany." *Population and Development Review*, 32(3): 511–528.