CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 613

Analyzing Matching Patterns in Marriage: Theory and Application to Italian Data

Pierre-André Chiappori, Edoardo Ciscato, and Carla Guerriero

May 2021

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 613

Analyzing Matching Patterns in Marriage: Theory and Application to Italian Data

Pierre-André Chiappori[†], Edoardo Ciscato[‡], and Carla Guerriero[§]

Abstract
Social scientists have long been interested in marital homogamy and its relationship with inequality. However, measuring homogamy is not straightforward, particularly when one is interested in assessing marital sorting based on multiple traits. In this paper, we argue that Separate Extreme Value (SEV) models not only generate a matching function with several desirable theoretical properties, but they are also suited for the study of multidimensional sorting. Specifically, we show how a small number of factors can be identified that capture most of the explained variance in matching patterns. We then use rich small-scale survey data to examine sorting among parents of children attending schools in Naples. Our findings show that homogamy is pervasive; not only do men and women sort by age, education, height, and physical characteristics, but they also look for partners that share similar health-related behavioral traits and risk attitude. We also show that marital patterns are well explained by a low number of dimensions, the most important being age and human capital. Moreover, children of parents with a high human capital endowment perform better at school, although they report lower levels of subjective well-being and perceived quality of relationship with their mothers.

Keywords: Homogamy, Matching, Intergenerational Inequality.

Acknowledgements: Funding from Programma Star UNINA Linea 1 Junior Principal Investigator Grant 2016 (Compagnia di San Paolo and Istituto Banco di Napoli) is gratefully acknowledged. This paper would have never been possible without the attention and enthusiasm of the children, their parents, and the precious help of the school teachers. A special thank you to the DISES-UNINA administrative staff. A development version of the R package used to produce the quantitative findings of this paper can be found on the following GitHub website: https://github.com/edoardociscato/affinitymatrix.

[††] Columbia University.

[‡] KU Leuven.

[§] Università di Napoli Federico II and CSEF.

Table of Contents

1 Introduction

Since the pioneering work of Becker (1973), a large number of studies have analyzed matching patterns in the marriage market. From a social sciences perspective, a major motivation to study partner choice, among several reasons, is the relationship between marital patterns and inequality. Assortative matching has a direct impact on inequality *within generations*; if individuals have a large propensity to marry their own likes, initial inequalities in individual endowments tend to be amplified at the household level (Fernández and Rogerson, 2001; Greenwood et al., 2003, 2014). More importantly, recent studies have emphasized the potential impact of assortative matching on social reproduction and *intergenerational* inequality. Individuals with a high level of human capital tend to match assortatively; these households tend to invest heavily in their children's human capital, and their investments tend to be particularly productive. Given the complementarities involved in the human capital production function, such trends amplify initial inequalities for the next generation, generating what can be termed an "inequality spiral".[1]

While numerous studies in economics, sociology, and demography have investigated marital patterns,[2] they mostly concentrate on a specific trait, such as income or human capital.[3] However, the real-life process of marital matching is obviously much more complex, and involves a host of other characteristics: age, race, religion, but also tastes and preferences, cultural background, physical attractiveness, etc. From a methodological viewpoint, whether the empirical strategy adopted can account for this diversity and, more importantly, whether it can disentangle the respective impacts of multiple traits, especially when the latter appear to be correlated, is an important question.[4]

The present paper is based on the so-called Separable Extreme Value (SEV) model, which has become dominant in the empirical analysis of matching models. The SEV approach uses a frictionless matching framework with Transferable Util-

[1] See for instance Del Boca et al. (2014), Chiappori et al. (2017), Chiappori et al. (2017) and Chiappori et al. (2020).

[2] A non-exhaustive list includes Schoen (1981); Qian and Preston (1993); Blackwell and Lichter (2004); Schwartz and Mare (2005); Bouchet-Valat (2014); Schwartz and Han (2014); Gonalons-Pons and Schwartz (2017).

[3] An obvious exception is Dupuy and Galichon (2014), which is discussed later on.

[4] For a general presentation of multidimensional matching, see Chiappori et al. (2010).

ity, in which unobserved heterogeneity is captured through an additive, separable random term and the latter is typically assumed to follow a type 1 extreme value distribution. Our focus is on its empirical implementation in a multidimensional context. We describe how a simple extension of the basic SEV model, borrowed from Dupuy and Galichon (2014), allows one to estimate the interactions between the various traits characterizing the spouses. We also show how one can estimate more restricted models in which these numerous traits only matter through a small number of unknown "factors" (or "indices").

Next, we apply our methodology to a rich Italian dataset, a survey of parents and their children aged 6 to 19 attending schools in the Campania region around Naples, contributing to the literature on matching and human capital formation with a unique dataset of 276 families, comprising information on parents, children, and families as a whole. The data collected on parents include sociodemographic variables (age and education), anthropometric characteristics (height and weight), health-related behavior (e.g., smoking and sports activity), household-level characteristics (e.g., number of children and the time spent by the mother at home), and an incentivized question relating to the relative healthiness of eating habits. The survey also collects parental psychometric information on risk behavior, with a focus on health and recreational risk.

The data on children includes anthropometric characteristics, measures of risk preferences, and two incentivized questions, one on attitude and taste for healthy food, and one on the ability to defer gratification. In addition, the survey collects information on children's educational outcomes, physical and mental health, attitude towards healthy lifestyle and eating habits, the amount of time spent in front of screens, the amount of time spent with them by parents, the type of relationship with their parents, and the children's degree of altruism. This information is particularly interesting, since it relates to what can be considered as *outcomes* of parental investment. We show how our empirical analysis, and in particular the identification of the main factors that drive matching patterns, can be formally related to these outcomes, and how, conversely, these outcomes help providing an interpretation of the factors.

Our conclusions can be summarized as follows. Marital patterns are charac-

terized by a surprisingly large degree of homogamy, that relates not only to age, education, and anthropometric measures, but also to *all* behavioral traits under consideration, including those related to health and risk attitudes. We find that sorting can essentially be reduced to a low-dimensional process. Indeed, a model based on a small number of factors, three for the most general version of our model, constitutes a very good approximation of the marriage market. In particular, we find that the main sorting factor reflects the market segmentation in different age cohorts, while the second most important dimension can be interpreted as capturing sorting based on human capital. Not only educated women tend to marry educated men, but sorting also brings together individuals that are health-conscious, e.g., that smoke less and have a preference for healthy food.

Finally, we analyze the consequences of these assortative matching patterns in terms of child outcomes. Positive assortative mating may play an important role in children's direct socialization; parents who want to transmit their characteristics to their children may be more likely to choose a partner with a similar bundle of traits (Bisin and Verdier, 2000, 2001; Bisin et al., 2004). We find that parents with a high level of human capital have children that perform better at school and have healthier eating habits. Interestingly, however, these same children report lower levels of happiness and a worse relationship with their mothers, suggesting that parental investments come with cost in terms of children's welfare.

2 Analyzing Marital Patterns: Theory

2.1 The Matching Model

In this section, we review the static, one-to-one, bipartite matching framework with transferable utility of Dupuy and Galichon (2014). We consider two populations, men and women, each defined by a set of characteristics X and Y, whose distributions are assumed to have densities f and g, respectively. Each woman is described by a vector of characteristics x of length m, each man by a vector y of length n. The matching can then be described by a *measure* μ on the product space $X \times Y$, and $\mu(x, y)$ denotes the probability that a woman of type x is matched with a man of type y. However, μ is constrained by the fact that its marginals must coincide

with $f(x)$ and $g(y)$; in other words, the matching is *feasible* if and only if

$$\mu \in M(f,g) \coloneqq \left\{ \mu : \begin{array}{c} \mu(x,y) \geq 0 \quad \forall x,y, \\ \int_Y \mu(x,y)dy = f(x) \quad \forall x, \\ \int_X \mu(x,y)dx = g(y) \quad \forall y \end{array} \right\}. \tag{1}$$

It is assumed that each woman i with type x only considers K men when choosing a partner; each potential partner k is characterized by a type y_k and an additively separable taste shock ε_k, so that i) the utility for i of choosing k as a husband is $U(x,y_k)+\varepsilon_k$, and ii) $\{(y_k,\varepsilon_k), k \in \{1,...K\}\}$ are the points of a Poisson process on $Y \times \mathbb{R}$ of intensity $dy \times e^{-\varepsilon}d\varepsilon$. Intuitively, $U(x,y)$ is the systematic utility generated, by woman of type x marrying a husband of type y, whereas the vector ε represents i's idiosyncratic preferences for different potential partners. Eventually, i chooses her husband by maximizing her utility:

$$\max_k \left(U(x,y_k)+\varepsilon_k\right). \tag{2}$$

This implies that a man j will be chosen with probability

$$\Pr\left(\varepsilon_j - \varepsilon_k \geq U(x,y_k) - U(x,y_j), \forall k \in (1,...K)\right) = \frac{\exp\left(U(x,y_j)\right)}{\int_Y \exp\left(U(x,y)\right)dy}, \tag{3}$$

where the equality follows from the distributional assumption of the taste shocks. Men choose a wife in an analogous way, and their utility of choosing a woman l of type y_l corresponds to $V(x,y_l)+\eta_l$.

The market is at equilibrium when the matching is stable, i.e., when there are is no pair of individuals who would both prefer being matched together rather than with their current partners. Hence, the optimal matching must satisfy (3) for every agent on both sides of the market. Combining supply and demand conditions, one obtains

$$\mu(x,y) = \exp\left(\Phi(x,y) - a(x) - b(y)\right) \tag{4}$$

where Φ represents the *systematic match surplus*, while a and b are determined so that the equilibrium matching is feasible. More formally, these quantities are defined

as

$$\Phi(x,y) \coloneqq U(x,y) + V(x,y) \tag{5}$$

$$a(x) \coloneqq f(x)^{-1} \int_Y \exp\left(U(x,y)\right) dy \tag{6}$$

$$b(y) \coloneqq g(y)^{-1} \int_X \exp\left(V(x,y)\right) dx. \tag{7}$$

An important property of this choice model is the Independence from Irrelevant Alternatives (IIA), i.e., the fact that the *relative* probabilities of any set of possible choices do not depend on the presence of other ("irrelevant") alternatives. This implies that estimating the probability of choosing a partner of type y instead of type y', *conditional on marriage*, will give the same conclusions as analyzing *unconditional* choices (i.e., also taking singlehood as a possible choice).

2.2 Parameterization and Estimation

The surplus function is conveniently parameterized as follows,

$$\Phi\left(x,y\right) = x'Ay = \sum_{k,l} A_{kl} x_k y_l, \tag{8}$$

so that every element of the $m \times n$ *affinity matrix* A tells us whether two traits x_k and y_l are complements ($A_{kl} > 0$) or substitutes ($A_{kl} < 0$). In our context, $m = n$ since the various traits are observed for both spouses. In particular, diagonal terms $A_{ii}, i = 1, n$ are directly informative about the super- ($A_{ii} > 0$) or sub- ($A_{ii} < 0$) modularity of the surplus with respect to characteristic i.

The affinity matrix can be estimated using a maximum likelihood estimator. With a sample of N couples, the likelihood function corresponds to

$$\mathcal{L}(A) = \frac{1}{N} \sum_{i=1}^{N} \mu^A(x_i, y_i), \tag{9}$$

which means that the equilibrium matching μ^A must be computed for every choice of A. In practice, the computation of μ^A consists in finding two functions a and b so that μ^A is feasible.

When the dimension of X and Y is high relative to the sample size N, Dupuy

et al. (2019) suggest using a penalized likelihood method in order to allow for A to be sparse and increase the precision of the parameter estimates. A way of introducing sparsity is to impose a restriction to the rank of A. In practice, this can be achieved by generalizing (9) with

$$\mathcal{L}(A) = \frac{1}{N}\sum_{i=1}^{N} \mu^A(x_i, y_i) + \tau \left\|A\right\|_* , \tag{10}$$

where $\left\|A\right\|_*$ represents the nuclear norm of A, defined as the sum of its singular values, whereas τ is the Lagrangian multiplier associated with the rank constraint. In practice, τ is chosen through cross validation.

2.3 Dimensionality and Matching Factors

The biquadratic specification offers an important advantage; one can rewrite Φ as the linear combination of independent factors, each capturing a different dimension of assortativeness. This is insightful for multiple reasons. Firstly, it allows one to infer the number of dimensions of assortativeness. For instance, we can test the hypothesis that attractiveness is well summarized by a single index (or a small number of indices) subsuming numerous observable traits. Secondly, when multiple dimensions of assortativeness matter, it quantifies their relative importance. Lastly, it allows one to describe the role played by the observables x and y in each dimension of assortativeness.

In practice, the affinity matrix can be decomposed through a Singular Value Decomposition (SVD) as[5]

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

where Λ is a diagonal matrix whose positive nonincreasing elements $(\lambda_1, \ldots, \lambda_K)$, with $K = \min\{m, n\}$, capture the relative importance of each sorting dimension, while the columns of U and V are loading vectors that describe the nature of each dimension. In other words, we can define the *factors* (or *indices*) as $\tilde{x} = Ux$ and

[5]Importantly, it is convenient to work with demeaned and rescaled data. More precisely, each observable characteristic x_k must be demeaned so that the sample mean $\hat{\mu}_k$ is zero; and all characteristics x must be rescaled so that the diagonal elements of the sample covariance $\hat{\Sigma}$ are one (i.e., $\hat{\sigma}_k = 1$ for every k). A similar transformation must be applied to y.

$\tilde{y} = Vy$, and rewrite the surplus as

$$x'Ay = \tilde{x}'\Lambda\tilde{y} = \sum_{k=1}^{K} \lambda_k \tilde{x}_k \tilde{y}_k,$$

where each k term $\lambda_k \tilde{x}_k \tilde{y}_k$ represents the surplus contribution of an independent dimension of assortativeness.

In our empirical analysis, we can perform SVD on the estimated affinity matrix $\hat{A}$ to obtain estimates of U, V, and Λ. In this way, we can discuss the relative importance and nature of the different dimensions of assortativeness. Confidence intervals for U, V, and Λ can be obtained with bootstrap techniques.

However, what is unknown is how many relevant dimensions of assortativeness do we observe as well as how many elements of Λ are positive and significant. Dupuy and Galichon (2014) outline a method to answer these questions and develop a test of joint significance of the estimated Λ. In summary, the method consists of testing the rank of the estimated affinity matrix $\hat{A}$. The null hypothesis is a restriction on the rank of Z, i.e., $rank(A) = k$; if it is rejected, then the number of positive diagonal elements of Λ will be higher than k, which will lead us to conclude that the number of relevant dimensions of assortativeness is higher than k.

3 Data

The survey data used in this paper contain information about the preferences, beliefs, and actions of both parents and their children. Compared with previous studies on parents, this rich dataset fleshes out the separate role of each spouse and their characteristics in the matching function, making it possible to assess the effects of matching on children's outcomes. This study is part of a large research project, CHILDROLE, that explores the role of children as decision makers within the family. The data were collected in five schools in the Metropolitan City of Naples, Italy's third-largest city, from February to April 2019. The metropolitan city is one of the most densely populated areas in Europe. Naples and the surrounding towns are marked by sharp income and cultural differences, and offer a good setting for collecting representative sample data on Italian households. Families were recruited through schools that agreed to participate and classes were randomly selected to take

part in the study. Five schools (three elementary, one middle and one high school) agreed to take part. The schools, all public, are located in different districts of the city and nearby towns, with a good socio-economic mix. The parents were surveyed by face-to-face interviews, and pencil-and-paper questionnaire. To avoid reciprocal influence, fathers and mothers were asked to complete the questionnaire in separate rooms. Children were interviewed individually in class using the same questionnaire format. Younger children were helped to fill in the questionnaire during one-to-one interviews.

3.1 Matching Variables

The study collected information about parents' demographics, preferences, beliefs, and individual actions. Respondents were asked to report their date and place of birth.[6] Educational attainment is the self-reported highest level achieved: 1) Primary, 2) Middle School, 3) High School, or 4) University. Parents' self-reported height and weight are used to measure Body Mass Index (BMI). Respondents were also asked about their health-related actions, measured by three variables: smoking, physical exercise, and propensity for healthy diet. Smoking and physical exercise are measured by multiple-choice questions with three possible answers: "Never" (coded as 0),"Seldom" (1), and "Often" (2). Propensity for healthy dietary choices was assessed through an incentivized question, with each parent choosing snack to consume after completing the questionnaire.[7] A generic question, "Do you worry about your own health?", also investigated respondents' concern for their own health.

Next, information was collected on parents' measures of domain-specific, risk-taking preferences. Weber et al. (2002) originally developed and tested an individual measure of risk-taking in judgment and decision-making in different conventional domains: financial, ethical, health/safety, social, and recreational (Weber et al., 2002; Blais and Weber, 2006). We used only the questions on the health and safety, and the recreational domains. The questions were modified with the help of a child psychologist, in order to explore how risk preferences and individual actions

[6]Missing answers were supplied from school registry.

[7]Parents were shown three different types of snack before the interview and were asked to select one to have immediately after the interview. The choices were: a banana, a Parmesan bar, and a chocolate muffin. According to the Centers for Disease Control, the three correspond to different degrees of healthiness: respectively, very healthy, healthy and unhealthy.

vary among household members, including children and youths. After extensive piloting, additional questions applicable to both parents and children were included in the final questionnaire (see Appendix A). For further details on the questions, see Guerriero et al. (2018). Risk-taking preferences and actions were measured with eight statements. Three explored risk preferences and actions in the health and safety domain (e.g. "I wear sunscreen to avoid sunburns"). The remaining five statements referred to preferences for recreational risk (e.g. "I would go on safari in the jungle").

3.2 Outcome Variables

As part of the CHILDROLE data project, we also collected information about objective outcomes, such as children's academic performance and anthropometric measures, and subjective indicators, such as children's subjective wellbeing and perceived quality of the relationship with their parents.

The first set of outcome variables relate to children's cognitive competence. From the class register, we observe grades in Italian and math, and whether the child failed the academic year. Additionally, we acquired information on children's nationally standardized OCSE-PISA test results in Italian and math collected by the National Institute for the Educational Evaluation of Instruction and Training (INVALSI). In the child psychology domain, a child's ability to defer gratification, especially at younger ages, has been found to be an important predictor of cognitive and social competence later in life (Mischel et al., 1989; Castillo et al., 2019; Watts et al., 2018). Hence, we measure children's discount rate using a simple incentivized experiment assessing whether they prefer to receive one snack today rather than two snacks tomorrow.[8]

The second set of variables relate to children's physical and mental health. Height and weight were directly measured by the interviewers. Using parents' answers, we also have information on children's weight at birth, which has been found to be positively associated with better longer-run outcomes such as IQ, education,

[8]Children were asked whether they wanted to wait one day to receive two (instead of one) of their most preferred snack. Children could choose between three alternatives: a banana, a Parmesan bar, and a chocolate muffin. The snacks were left in the class to show that they would be readily available the day after, to reduce potential distrust about actually receiving the rewards, and because the visibility of the reward itself has been associated with the ability to defer gratification (Mischel et al., 1989; Watts et al., 2018).

and earnings (Black et al., 2007). Cognitive and affective evaluations of one's life were measured using three different questions (Bradshaw, 2015; Nima et al., 2020). The first question investigates how children evaluate their life as a whole with the generic question: "Are you satisfied with your own life?"[9] The affective component of subjective well-being assesses emotions that people experience in daily life. After extensive piloting, we selected two relevant questions that were easily understood by all age groups in order to measure this affective component: "How often do you feel tired?" and "How often do you feel happy?"

The third set of variables relate to attitudes towards a healthy lifestyle in children. A study by Case and Paxson (2002) documented how parental behavior, such as smoking, use of child-seat and belt, and regular bedtime, are important predictors of children's health. Fisher et al. (2002) found that daughters of parents who consumed more fruit and vegetables had higher fruit and vegetable intake, as well as a higher micronutrient and a lower fat intake. Studies in the public health domain also suggest that there is substantial parent-to-child transmission of dietary behaviors (Vepsäläinen et al., 2018). We investigate attitudes towards healthy lifestyle in children in our dataset and try to uncover what they like, what they think to be healthy, and whether they can restrain themselves from eating unhealthy snacks when given the option. For this purpose, we ask them to indicate how much they like three different snacks (a banana, a Parmesan bar, and chocolate) and which one they would like to eat as a reward after the survey. In addition, we ask children whether they worry about their health, how often they eat vegetables and fruit, practice sports, smoke, and consume soft drinks. Finally, we ask them how much time they spend every day using a tablet, smartphone, and/or watching TV.

The fourth set of variables in our study measure the amount of time children spend with their parents and the type of relationship they have together. Del Boca et al. (2014) find that both parents' inputs are important for children's cognitive development and that parental time inputs, especially active time, are generally more productive for cognitive development than financial expenditure on "child goods" (such as tutoring or toys). We asked parents how much time they spend outside

[9]During the pilot phase we checked children's comprehension of the question and their capacity to transform their own evaluations to a 5-Likert scale following González-Carrasco et al. (2015).

the home every day and we asked children how much time they spend with their parents. In order to investigate the relationship between the child and their parnet we asked children whether they get along with their mother and father. Finally, we asked the children how they perceive the role their parents play within the family.[10]

The fifth set of variables measures children's financial autonomy. The ability to deal with money is an essential skill that people must acquire to successfully function in society (McCormick, 2009; Suiter and Meszaros, 2005). During childhood and adolescence, the availability and use of money is also a powerful measure of parents' non-paternalistic altruism and of their lack of control over children's choices (Barnet-Verzat and Wolff, 2002). Nevertheless, little is known on how much pocket money young children receive from their parents. In our study, we collect information about the frequency and amount of pocket money allowances.

The sixth set of variables considered in this study measure are children's risk preferences and actions in the health and safety domain, using the same questions posed to parents. Parents can mold their children's preferences in order to align them with their own. A study conducted by Dohmen et al. (2012) on a large sample of German citizens shows that willingness to take risks and to trust others are transmitted from parents to children. In a study on Bangladeshi households, Chowdhury et al. (2018) investigate how the risk, time, and social preferences of children are influenced by both parents. The study finds that both mother's and father's preferences are significantly related to the child's in all three domains.

Finally, we measured children's degree of altruism. We asked children whether they would try to help a classmate in trouble to assess their generic altruism, and whether they would give money to a classmate that needs to buy a snack to assess their non-paternalistic altruism (Guerriero et al., 2018).

3.3 Summary Statistics

Table 1 reports the number of households participating in the study: 632 children were contacted. For 332 of them, at least one parent participated in the survey, and for 276 of them, both parents participated in the survey. The latter constitute the core of our sample, since information on both parents is necessary for our empirical

[10]Children were asked who works and who takes care of the home in order to identify if tasks are equally shared or parents specialize (Amato and Booth, 1995; Kaufman, 2000).

Table 1: Frequency of children by parents' survey participation

	Father participates		
Mother participates	No	Yes	Total
No	326	8	334
Yes	48	276	324
Total	374	284	658

Notes: Participation is coded as the parent reporting at least some basic information.

analysis. Table 2 reports summary statistics on parents' education and age. We have complete information on age, educational attainment, height, and BMI for 259 of 276 participant couples. Henceforth, this group of 259 couples will be termed Sample 1. Table 3 reports summary statistics on parents' health and recreational risk behavior. 205 of 276 participant couples had both parents complete the behavioral questionnaire. This group of 205 couples will be termed Sample 2. In Table 2, we see that men are on average 43 years old and women 40 years old. Women are also slightly better educated. The average man in our sample is overweight according to WHO standards (BMI greater than 25.0), while women are on average in the healthy range (BMI between 18.5 and 24.9). Table 3 shows summary statistics on risk behavior. Women smoke less but are also less active. Preferences for healthy snacks and concerns about own health are on average similar across genders. Gender differences are more pronounced for specific health and recreational risk behaviors. Women are more likely to use sunscreen at the beach, more scared of riding a fast moped, and less interested in doing extreme sports.

4 Results

4.1 Reduced Affinity Matrix

We start with a small-size version of our model, in which we only consider matching patterns on four socio-demographic characteristics: age, education, height, and BMI. The corresponding matrix is displayed in Table 4. All diagonal coefficients are positive and significant, implying that, for all characteristics, homogamy increases the surplus generated by the match. Not surprisingly, the largest and most significant association relates to age, a feature that essentially reflects the presence

Table 2: Summary statistics - parents

	N	Mean	St Dev	10th P	90th P
Mother's education	323	2.9	0.8	2.0	4.0
Mother's age	322	39.7	6.8	31.0	49.0
Mother's height (cm)	314	163.6	5.6	158.0	170.0
Mother's BMI	308	24.3	3.9	20.4	29.3
Father's education	281	2.7	0.8	2.0	4.0
Father's age	283	43.1	7.3	35.0	53.0
Father's height (cm)	279	175.3	6.6	168.0	183.0
Father's BMI	274	26.8	3.4	23.0	31.1

Notes: the table reports number of observations, mean, standard deviation, 10th and 90th percentile of a variable. Education is coded as a four-category variable. Individuals that reported implausibly low anthropometric measures were excluded (height below 130cm or weight below 40kg).

of several cohorts among parents. Strongly significant is homogamy on education, which was expected, BMI, and height. Additional patterns emerge; for instance, more educated men tend to have older and thinner wives (while the opposite is not significant); and more educated wives tend to have taller husbands.

The factor decomposition is given in Table 5. The last line indicates that the first factor (Index 1) weighs three times more than the three other factors combined, and mostly reflects age differences. In other words, parents in our sample belong to different "cohorts" (defined by year of birth), and people tend to marry a spouse from a cohort that is close to their own. It should be noted that, in our sample, age is positively correlated with education, reflecting the well-known fact that more educated people tend to both marry and have children later. This explains the positive ans significant impact of education on the first factor. The second factor (Index 2) is particularly interesting. It singles out individuals who are more educated, as well as taller and, at least for women, thinner. In other words, matching patterns, while primarily driven by age, also capture a mix of education and physical appearance, possibly reflecting various dimensions of social status. Figures 1 and 2 provide a graphical visualization of the first two matching dimensions and we will return to them in the next sections for further interpretation.[11]

[11] In Appendix C, we also plot the marginal distributions of the first and second indices, as well as their joint distribution. This helps visualize the strength of sorting on different dimensions. By definition, sorting is stronger on the first dimension.

Table 3: Summary statistics - parents

	N	Mean	St Dev
Mother smokes	310	0.6	0.8
Mother does sports	310	0.8	0.8
Mother chooses healthy snacks	320	0.8	0.8
Mother wears sunscreen	309	1.4	0.7
Mother washes hands	310	1.8	0.4
Mother worries about her own health	311	1.5	0.6
Mother would go on safari	308	0.8	0.8
Mother fears speed	295	1.3	0.7
Mother likes holidays in known places	307	1.1	0.7
Mother would do extreme sports	311	0.2	0.5
Mother crosses carefully	311	1.9	0.3
Father smokes	269	0.8	0.9
Father does sports	267	0.9	0.7
Father chooses healthy snacks	277	0.8	0.8
Father wears sunscreen	270	1.0	0.8
Father washes hands	266	1.8	0.5
Father worries about his own health	269	1.5	0.6
Father would go on safari	267	0.8	0.8
Father fears speed	260	0.7	0.7
Father likes holidays in known places	265	1.1	0.7
Father would do extreme sports	269	0.4	0.6
Father crosses carefully	268	1.8	0.5

Notes: The table reports number of observations, mean and standard deviation. Possible answers are (0) never, (1) sometimes, (2) often. The only exception is the question about healthy snacks, for which possible answers are (0) chocolate, (1) Parmesan bar, (2) banana. See Table 8 for more details.

Table 4: Estimated affinity matrix (Sample 1)

Husband \ Wife	Education	Age	Height	BMI
Education	0.74	0.62	0.14	-0.24
	(0.11)	(0.17)	(0.09)	(0.10)
Age	0.24	3.30	0.25	0.07
	(0.16)	(0.33)	(0.13)	(0.13)
Height	0.25	0.37	0.16	-0.07
	(0.09)	(0.14)	(0.07)	(0.07)
BMI	0.10	0.20	0.01	0.28
	(0.08)	(0.13)	(0.07)	(0.07)

Notes: 259 couples. Standard errors in parentheses. The estimator $\hat{A}$ is asymptotically normal.

Table 5: Saliency analysis (Sample 1)

	Men		Women	
	Index 1	Index 2	Index 1	Index 2
Education	0.21	0.92	0.12	0.91
	(0.02)	(0.03)	(0.02)	(0.02)
Age	0.97	−0.23	0.99	−0.12
	(0.01)	(0.02)	(0.00)	(0.02)
Height	0.12	0.30	0.08	0.15
	(0.03)	(0.08)	(0.03)	(0.07)
BMI	0.06	−0.05	0.01	−0.36
	(0.02)	(0.05)	(0.02)	(0.05)
Index share	0.75	0.17	0.75	0.17
	(0.07)	(0.02)	(0.07)	(0.02)

Notes: The table reports men's and women's singular vectors, V and U respectively, and singular values, $diag(\Lambda)$, from the singular value decomposition of $\hat{A} = U'\Lambda V$. We report standard errors in parentheses; they are obtained with 1,000 bootstrap replications (Milan and Whittaker, 1995). In the last line, each value of $diag(\Lambda)$ can be interpreted as the relative importance of each sorting dimension.

4.2 Global Affinity Matrix

Here, we look at the global (15×15) affinity matrix. A first and very striking feature is the high level of homogamy that prevails within the population. *Each* of the 15 diagonal coefficients in Table 6 is positive, indicating positive assortativeness along that specific dimension. The probability of getting such a pattern under random matching would be less than .01%. Moreover, all but one are statistically significant at 5%, and most are actually significant at 1%. This is even more remarkable given the relatively small sample size of only 205 couples.[12]

[12]If the spouses become more similar in terms of health-related behavior and risk attitude *after* the marriage, our estimates of the diagonal coefficients would be upward biased. If we had information on the length of the relationship, we would be able to run a robustness check with newlyweds only. However, there is already evidence in the literature that these kind of traits do not significantly alter after marriage: Dohmen et al. (2012) find that correlation patterns between spouses' risk attitude and trust do not change with the length of the relationship.

Table 6: Estimated affinity matrix (Sample 2)

Husband \ Wife	Education	Age	Height	BMI	Smokes	Does sports	Chooses healthy snacks	Wears sun-screen	Washes hands	Worries about health	Would go on safari	Fears speed	Likes holidays in usual places	Would do extreme sports	Careful when crossing
Education	0.29	0.32	0.10	-0.11	-0.04	0.06	0.12	0.04	0.01	-0.03	0.04	0.00	-0.12	0.04	-0.07
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Age	0.17	0.95	0.10	0.05	0.06	0.04	0.03	0.17	0.02	-0.01	-0.07	0.03	-0.13	0.04	-0.08
	(0.05)	(0.06)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
Height	0.10	0.04	0.04	-0.07	0.01	-0.00	0.01	-0.01	0.02	-0.04	0.02	-0.03	-0.02	0.01	-0.01
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)	(0.03)
BMI	-0.02	0.05	-0.01	0.11	-0.06	-0.02	-0.01	0.04	0.02	0.01	-0.04	0.03	-0.03	0.00	-0.04
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Smokes	-0.10	-0.00	-0.03	-0.02	0.14	-0.03	-0.10	-0.03	-0.01	-0.03	0.01	-0.03	-0.02	-0.03	0.07
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Does sports	-0.06	0.01	0.02	-0.05	-0.03	0.08	-0.06	-0.05	0.01	0.07	0.05	0.02	0.03	0.00	0.07
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
Chooses healthy snacks	0.11	0.02	-0.01	-0.02	-0.02	-0.00	0.25	0.06	0.02	-0.02	0.02	0.05	0.06	0.01	-0.05
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Wears sunscreen	0.08	-0.04	0.00	-0.00	-0.07	-0.01	0.16	0.07	-0.02	0.01	-0.06	-0.01	0.04	-0.00	-0.06
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Washes hands	-0.09	0.03	0.02	0.01	-0.00	-0.01	0.01	0.02	0.10	0.04	-0.02	0.00	0.02	-0.02	0.08
	(0.05)	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Worries about health	-0.05	0.02	0.02	0.00	-0.05	0.04	-0.02	0.01	-0.00	0.07	-0.03	-0.02	-0.01	-0.01	0.03
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Would go on safari	0.03	-0.07	0.01	-0.08	-0.01	0.05	0.01	-0.08	0.04	0.00	0.17	0.05	-0.02	0.01	0.04
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
Fears speed	-0.04	0.05	-0.03	0.05	0.02	-0.00	0.04	0.02	-0.01	0.00	0.01	0.05	0.03	0.00	-0.01
	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)
Likes holidays in usual places	0.01	0.02	-0.05	0.05	-0.05	0.01	0.06	-0.01	-0.06	-0.02	0.04	0.07	0.16	0.04	-0.08
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
Would do extreme sports	0.06	0.08	0.02	-0.03	-0.01	0.01	0.06	0.02	0.01	-0.01	0.01	0.01	0.01	0.01	-0.02
	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)
Careful when crossing	-0.16	-0.01	0.00	-0.01	0.05	0.01	-0.04	-0.01	0.07	0.06	-0.01	-0.00	0.01	-0.04	0.13
	(0.04)	(0.05)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)	(0.03)

Notes: 205 couples. Standard errors in parentheses. The estimator $\hat{A}$ is asymptotically normal.

Several aspects revealed in the reduced matrix are still visible here - for instance, a positive (and significant) interaction between his education and her age, or between her education and his height, as well as a negative interaction between his education and her BMI. Others are less expected. Wives of more educated men are more likely to choose healthy snacks and to favor vacations at an unknown place; husbands of more educated women are also more likely to eat healthy food, but are less careful when crossing a street.[13]

The factor decomposition enriches the conclusions drawn from the reduced matrix. Factor loadings are reported in Table 7. Index 1, which accounts for about a third of the total systematic surplus by itself, essentially recaptures the cohort pattern observed on the reduced matrix. As can be seen in the first and fourth column of Table 7, age plays a dominant role in the first sorting dimension, while the role of education follows from the observed positive correlation between education and both age at marriage and age at first birth. More surprisingly, the factor is also associated with a preference for extreme sports among men, and a dislike for known holiday destinations among women.

Index 2 is of a different nature. It singles out individuals who are younger, more educated, taller (men) or thinner (women) and more health-conscious (they are more likely to eat healthy food and wear sunscreen, and less likely to smoke or experience health problems). All in all, Index appears to capture various dimensions of parents' *human capital* (education health), which in turn appears to be highly correlated with social status; parents with a high Index 2 live in a more residential neighborhood (hence, the reduced requirement to pay attention when crossing the street). In addition, some traits are more idiosyncratic, such as individuals with a high Index 2 being fond of holidays in known destinations. Along this second dimension of matching, some characteristics are gender-specific; for instance, among individuals with a high Index 2, men are less likely to do sports, although they have a preference for extreme sports, and to wash their hands, possibly because they are

[13]Street crossing behavior also reflects the safety of the neighborhood where the family lives. Individuals who live in safer places, with more pedestrian areas and less dangerous crossing points, are less careful when crossing the street. Only 71% of parents from students at the centrally located Umberto I high school report being very careful when crossing the street, compared to 87% in other schools. Parents of students at Umberto I high school are by far the most educated in the sample, and yet street crossing behavior is uncorrelated with education, which suggests that it is primarily driven by the neighborhood where the family lives.

Table 7: Saliency analysis (Sample 2)

	Men			Women		
	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3
Education	0.39	0.39	0.47	0.27	0.59	0.31
	(0.02)	(0.04)	(0.06)	(0.02)	(0.04)	(0.06)
Age	0.90	−0.27	−0.13	0.91	−0.27	−0.05
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.03)
Height	0.07	0.14	0.28	0.12	0.00	0.22
	(0.03)	(0.06)	(0.07)	(0.03)	(0.06)	(0.08)
BMI	0.05	−0.06	−0.29	0.01	−0.14	−0.53
	(0.02)	(0.05)	(0.07)	(0.03)	(0.05)	(0.07)
Smokes	−0.05	−0.33	0.14	0.02	−0.25	0.13
	(0.02)	(0.05)	(0.06)	(0.02)	(0.05)	(0.06)
Does sports	−0.03	−0.16	0.20	0.05	0.01	0.18
	(0.03)	(0.05)	(0.07)	(0.03)	(0.05)	(0.07)
Chooses healthy snacks	0.07	0.48	−0.20	0.10	0.58	−0.25
	(0.02)	(0.05)	(0.06)	(0.02)	(0.04)	(0.05)
Wears sunscreen	0.02	0.37	−0.27	0.17	0.05	−0.30
	(0.03)	(0.05)	(0.07)	(0.03)	(0.06)	(0.08)
Washes hands	−0.00	−0.18	−0.09	0.02	−0.08	0.11
	(0.03)	(0.05)	(0.06)	(0.03)	(0.05)	(0.07)
Worries about health	0.00	−0.12	−0.04	−0.03	−0.12	−0.04
	(0.03)	(0.05)	(0.07)	(0.03)	(0.06)	(0.07)
Would go on safari	−0.06	0.12	0.47	−0.05	0.12	0.39
	(0.02)	(0.04)	(0.05)	(0.03)	(0.04)	(0.05)
Fears speed	0.03	−0.02	−0.22	0.03	0.08	−0.11
	(0.03)	(0.06)	(0.08)	(0.03)	(0.05)	(0.07)
Likes holidays in known places	0.01	0.22	−0.38	−0.14	0.11	−0.39
	(0.02)	(0.04)	(0.05)	(0.02)	(0.04)	(0.05)
Would do extreme sports	0.09	0.13	0.03	0.05	0.09	0.01
	(0.03)	(0.06)	(0.08)	(0.03)	(0.06)	(0.09)
Careful when crossing	−0.06	−0.35	0.02	−0.11	−0.32	0.22
	(0.03)	(0.05)	(0.07)	(0.03)	(0.05)	(0.07)
Index share	0.38	0.18	0.12	0.38	0.18	0.12
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)

Notes: The table reports men's and women's singular vectors, V and U respectively, and singular values, $diag(\Lambda)$, from the singular value decomposition of $\hat{A} = U'\Lambda V$. We report standard errors in parentheses; they are obtained with 1,000 bootstrap replications (Milan and Whittaker, 1995). In the last line, each value of $diag(\Lambda)$ can be interpreted as the relative importance of each sorting dimension.

less likely to have a manual occupation, while women appear to like safaris.

Finally, the third factor emphasizes yet other traits, several of which are related to physical shape. Individuals with a high Index 3 are younger, more educated, but also taller, thinner, and more likely to do sport. They dislike usual holidays and prefer safaris. They are also less likely to choose healthy snacks and wear sunscreen.

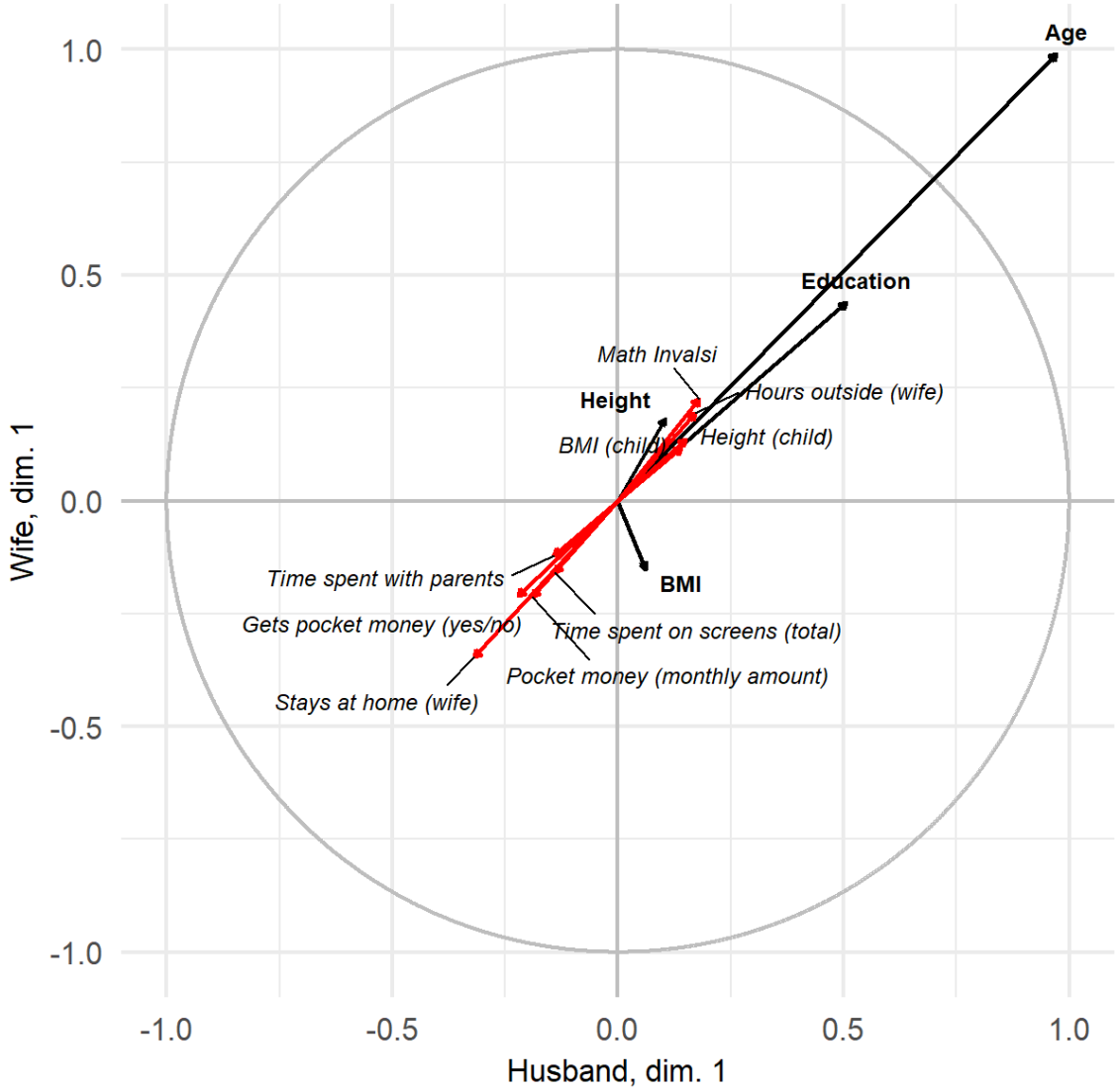
Together, these three factors explain more than the remaining 13 factors combined. A formal test, presented in Table 12, suggests that these factors are in fact sufficient to fully summarize the matching affinity matrix; the null hypothesis that its rank is equal to 3 is not rejected.

4.3 Outcomes

An interesting aspect of our data is that they include what could be considered as "outcome" variables, i.e., indicators reflecting choices made within the household and their consequences. A standard example is labor supply behavior. While most married men are active on the labor market, women may or may not participate, and these decisions appear to be related to matching patterns. Another type of outcome regards the impact of matching on children. In our data, we observe both objective outcomes, such as grades and academic performance, and more subjective indicators, such as the child's subjective well-being and the perceived quality of the relationship with their parents.[14] A list of all outcome variables is provided in Appendix A.

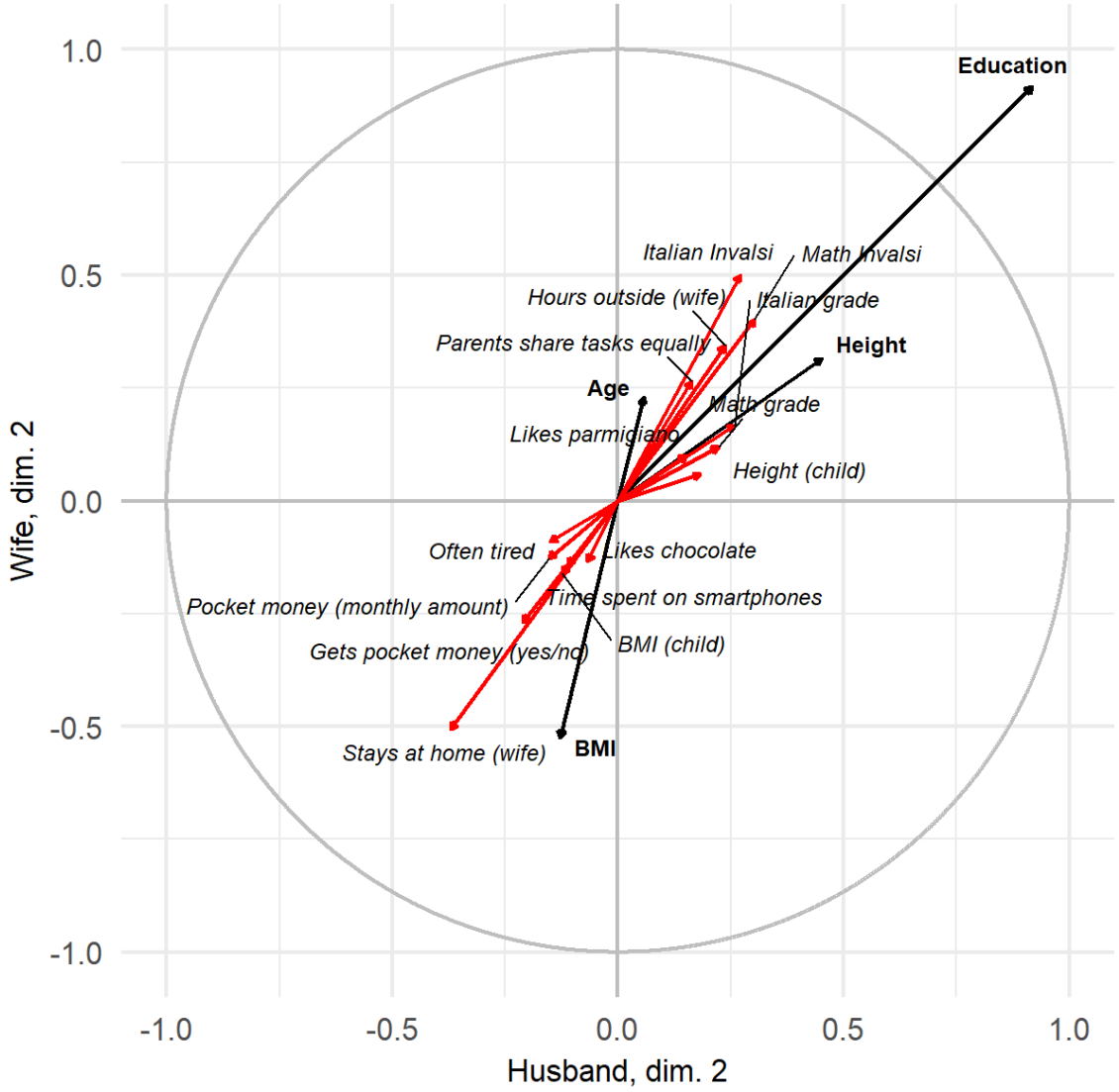
[14]The latter variables exhibit a clear age pattern, older children (i.e., teenagers) typically exhibiting less satisfaction and more difficult parental connections. For that reason, we often "clean up" these variable from their age-related component. In practice, we regress the variable under consideration on the child's age, and we use the residual of this regression as our new indicator. More details are provided in Table 9.

Figure 1: Correlation between Index 1, matching variables, and outcomes (Sample 1)



Notes: We plot correlation rates of both matching variables and outcome variables with the husband's Index 1 (x-axis) and the wife's (y-axis). The matching variables include all the parents' background variables used in the main estimation. The outcomes include additional variables that are excluded from the main estimation, e.g., the number of children, the child's grades, and the wife's labor supply. In order to improve the readability of the graph, we only plot those variables whose correlation rate is significantly different from zero at the 5% level in a two-tailed test.

Figure 2: Correlation between Index 2, matching variables, and outcomes (Sample 1)



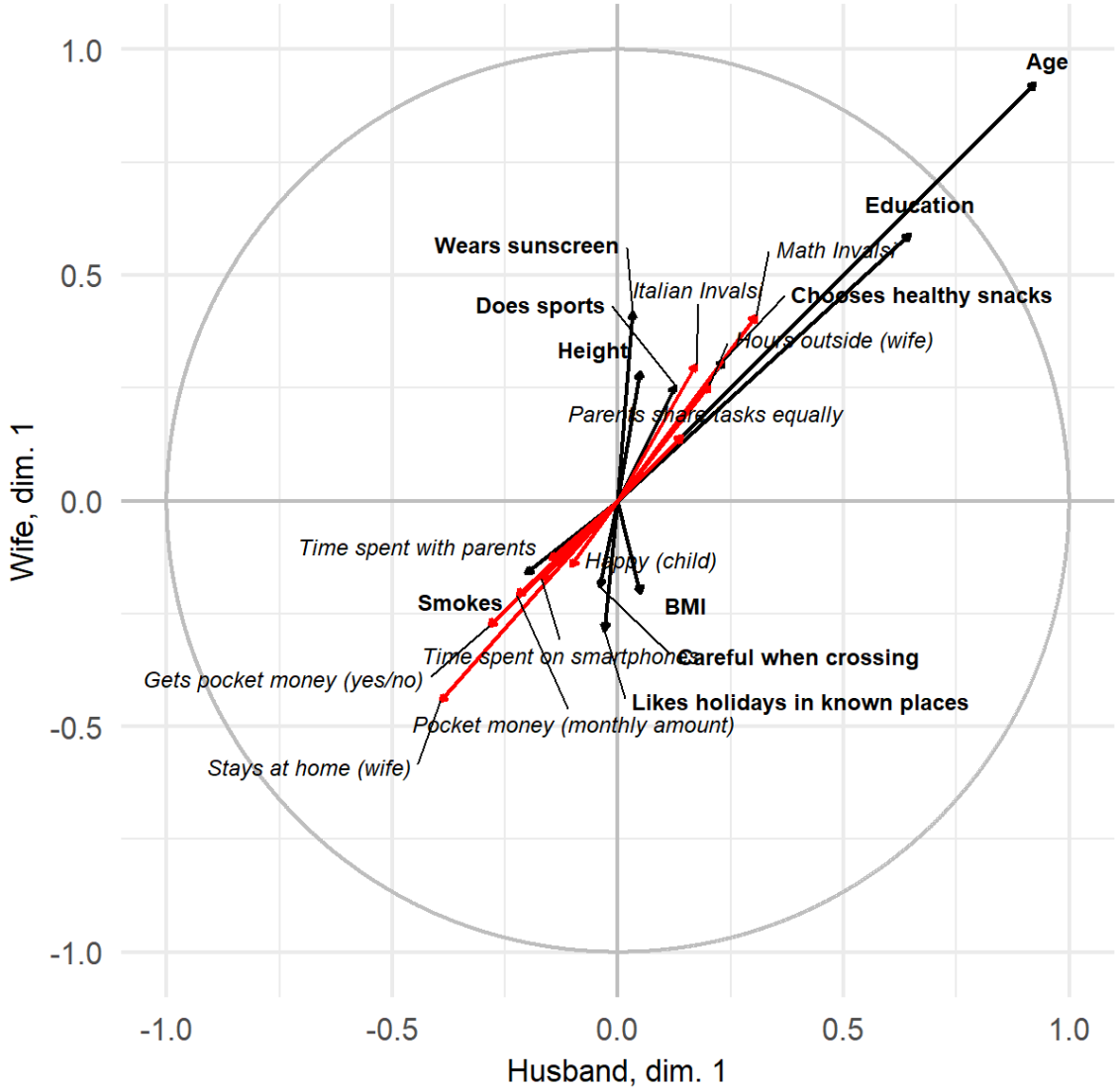
Notes: We plot correlation rates of both matching variables and outcome variables with the husband's Index 2 (x-axis) and the wife's (y-axis). See notes to Figure 1.

Figure 1 is based on the reduced affinity matrix; it plots the composition of the wife's and husband's Index 1 together with selected output variables. In order to improve the readability of the graph, we only keep variables whose correlation with the corresponding factor is significant at the 5% level.[15] Parents with a high Index 1 are older, which correlates with children reporting lower levels of subjective well-being. This correlation, however, appears to be spurious, and mostly reflecting the child's age; once controlling for age, the correlation disappears. Similarly, women with a a high Index 1 are more likely to work outside home, but again this correlation is likely to partly reflect the older age of their children. All in all, the cohort component of the matching patterns does not seem to be strongly correlated with any output, except for children being less likely to receive pocket money and to spend time on screens.

Things are however quite different with the second factor, which we interpreted as being mostly driven by parents' human capital. In Figure 2, we see that women with a high Index 2 are more likely to work outside home; interestingly, these couples are also more likely to share domestic tasks equally. More importantly, high levels of Index 2 are strongly correlated with better educational outcomes, as measured by children's grades, including both nationally standardized INVALSI test results (math and Italian) and class grades (always math and Italian, standardized at class level). The children, just like their parents, are also taller than their peers. Finally, they receive less pocket money and spend less time on smartphones. The fact that they are possibly less financially autonomous and more monitored suggests their parents are likely to adopt markedly distinct, and possibly more intensive, parenting practices (Doepke and Zilibotti, 2017).

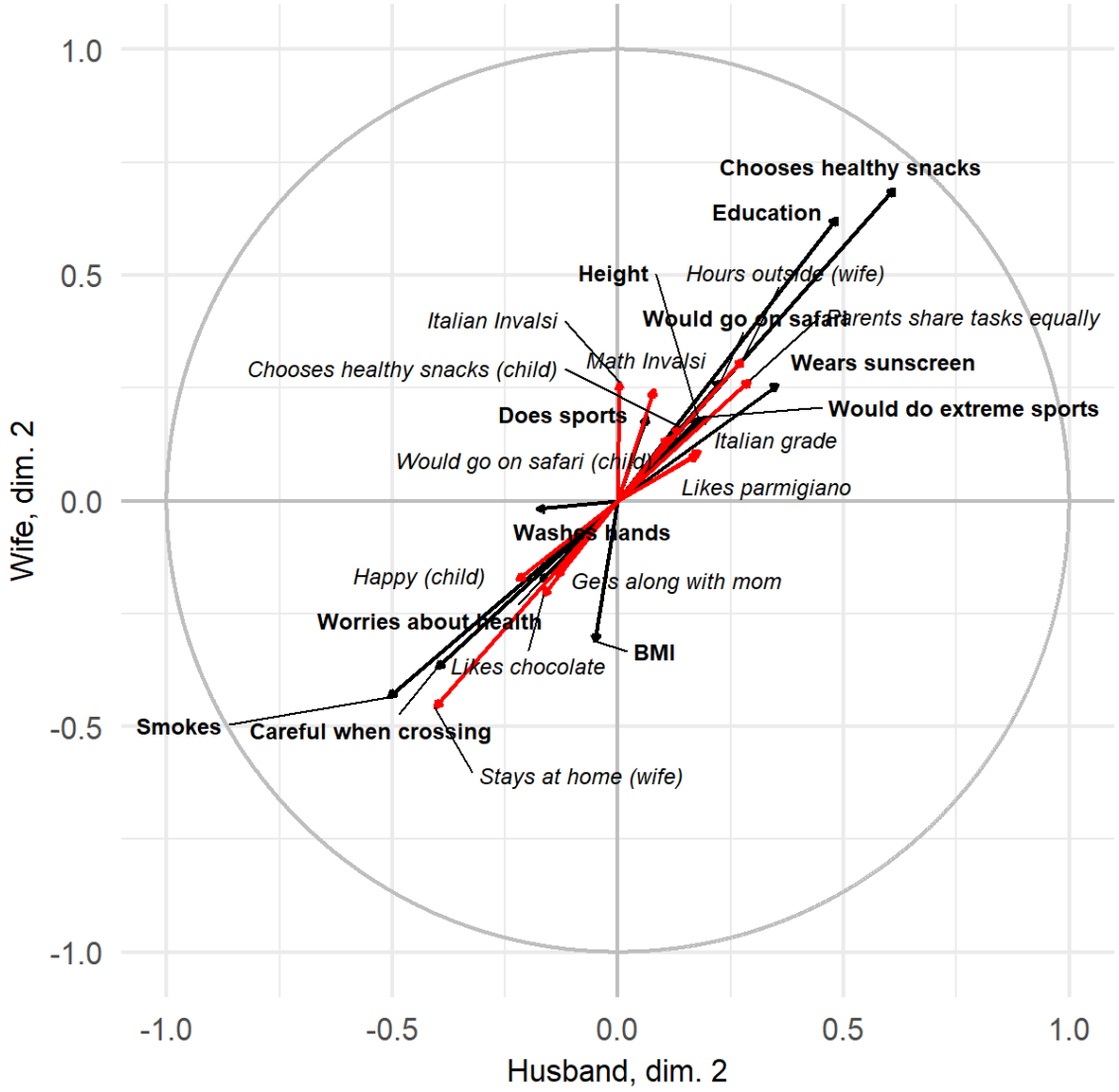
[15] Tables 13 and 14 in Appendix B report all correlation rates between the indices and the outcome variables.

Figure 3: Correlation between Index 1, matching variables, and outcomes (Sample 2)



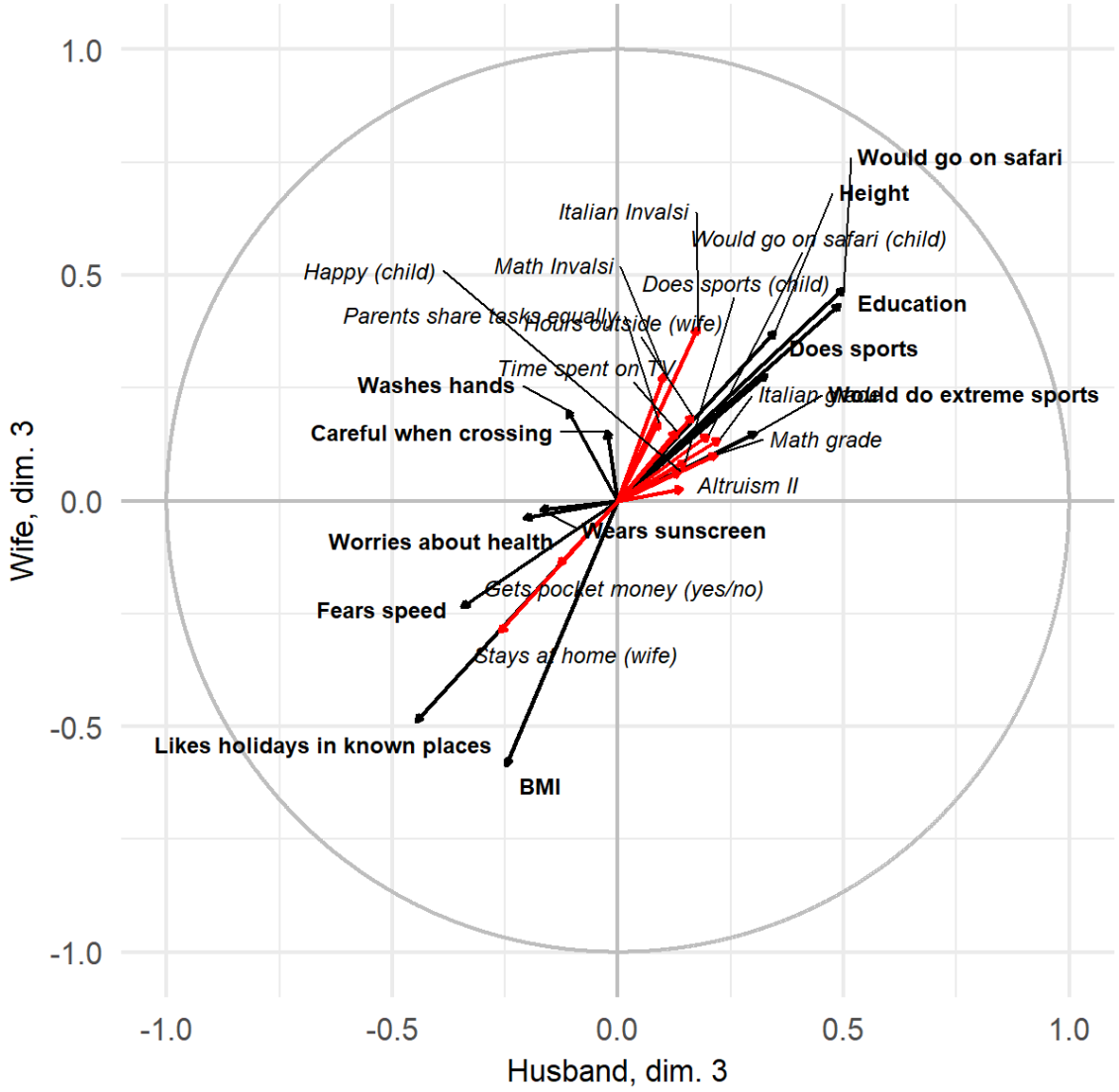
Notes: We plot correlation rates of both matching variables and outcome variables with the husband's Index 1 (x-axis) and the wife's (y-axis). See notes to Figure 1.

Figure 4: Correlation between Index 2, matching variables, and outcomes (Sample 2)



Notes: We plot correlation rates of both matching variables and outcome variables with the husband's Index 2 (x-axis) and the wife's (y-axis). See notes to Figure 1.

Figure 5: Correlation between Index 3, matching variables, and outcomes (Sample 2)



Notes: We plot correlation rates of both matching variables and outcome variables with the husband's Index 3 (x-axis) and the wife's (y-axis). See notes to Figure 1.

These conclusions are enriched when considering the global affinity matrix (Figures 3, 4, and 5). As before, Index 1 is only correlated with women's participation to the labor market, though with the same caveats as before, less pocket money for the kids, and a better performance in nationally standardized math tests. Again, children of parents with a higher Index 2 have significantly better grades than their peers, and appear to be more health-conscious, as they are more likely to choose healthy food and like Parmesan, while less likely to like chocolate. Their mothers are more likely to spend time outside home and to share domestic tasks equally with their husband. However, Index 2 is also strongly correlated with children reporting to be less happy and less likely to get along with their mothers. A possible interpretation is that educated parents' investment in their children's human capital, while fruitful in terms of health-related habits and academic performance, comes at a price, since children appears to resent the corresponding pressure.[16] This is in stark contrast with couples exhibiting a high level of Index 3. Their children also have better grades and less pocket money, but do sports more often and would like to go on safaris. Last but not least, they appear to be overall happier than their (nerdier) peers.

What is remarkable here is that the factors are recovered *exclusively from matching patterns*; neither parental investment nor any output variable are used for their estimation. Yet, parents' human capital, the most important factor driving assortativeness after parental age/cohort, appears to be strongly correlated with both children's human capital achievements (positively) and well-being (negatively). This strongly suggests not only that future investments in children's human capital are an explicit part of individuals' marital strategies, but also that these aspects are crucially important, since the corresponding factor dominates all other patterns.

5 Conclusion

A large body of literature exists across economics, sociology, and demography that has studied homogamy and measured it using data on marital patterns. In this

[16] This notion had already been mentioned in the literature (Heinrich, 2014; Dinisman and Ben-Arieh, 2016). One explanation is that the parenting style of educated parents is more oriented towards altruism than paternalism: this means that educated parents pressure children to perform well at school, which results in a welfare loss for the child (Doepke and Zilibotti, 2017).

paper, we argue that Separable Extreme Value (SEV) models can be used to study multidimensional sorting and can easily handle numerous discrete classes and continuous variables. We show that the SEV approach can generate rich empirical findings by estimating a multidimensional and parametric model, borrowed from Dupuy and Galichon (2014), with data from a survey of parents of children attending schools in Campania, a region of Southern Italy. We show that marital patterns are characterized by a high level of homogamy; not only do men and women sort based on demographic and socioeconomic traits such as age, BMI, height, and education, but they also look for partners that share similar health-related behavioral traits and risk attitudes. Our estimates are also insightful about the number and nature of the sorting dimensions that can rationalize the marital patterns observed in the data. We find that a relatively low number of sorting dimensions, three in our sample of couples, are sufficient to summarize an individual's attractiveness on marriage markets. While the first dimension of sorting mainly captures market segmentation across age cohorts, the second dimension describes sorting on human capital, so that educated and health-conscious women tend to marry men with similar traits. Finally, when we look at family outcomes, we find that children of parents with a high level of human capital perform better at school, but this comes at a cost, as they are also more likely to report lower levels of subjective well-being and a worse relationship with their mother.

A Variable Description

Table 8 contains the survey questions that were used to measure different behavioral traits. In Table 9, a complete list of all variables that were used in section 4.3 as "outcome" variables are reported. Note that some outcome variables may never show up in our plots because they are found to be uncorrelated with the estimated indices of attractiveness.

Table 8: Matching variables

Matching variables		
Variable	Question	Possible answers
Chooses healthy snacks	*At the end of the experiment, we will give you a snack. Which one do your prefer?*	1=chocolate, 2=Parmesan bar, 3=banana
Smokes	*I smoke*	0=Never, 1=Sometimes, 2=Often
Does sports	*I do sports*	0=Never, 1=Sometimes, 2=Often
Wears sunscreen	*I wear sunscreen to avoid sunburns*	0=Never, 1=Sometimes, 2=Often
Washes hands	*I wash my hands before eating*	0=Never, 1=Sometimes, 2=Often
Worries about health	*I worry about my health*	0=Never, 1=Sometimes, 2=Often
Would go on safari	*I would go on a safari in the jungle*	0=Never, 1=Sometimes, 2=Often
Fears speed	*I am scared of mopeds riding fast*	0=Never, 1=Sometimes, 2=Often
Likes holidays in known places	*I like going on holidays in places I know because it is safer*	0=Never, 1=Sometimes, 2=Often
Would do extreme sports	*I would do extreme sports*	0=Never, 1=Sometimes, 2=Often
Careful when crossing	*I am very careful when crossing the street*	0=Never, 1=Sometimes, 2=Often

Notes: The table reports the translated text of the questions that were asked in the written questionnaire.

Table 9: Outcome variables

Outcome variables	
Variable	Description
Number of children	Children report family composition, including number and sex of siblings.
Stays at home (wife)	Dummy variable. Answers "housewife" when asked about profession.
Hours outside (wife)	Mothers are asked: *If you work, how many hours do you usually spend outside the home?* Possible answers: Does not work, 3-5 hours, 6-8 hours, ¿8 hours.
Share equal tasks	Dummy variable. Answers "Both parents work and share chores" when asked about gender roles in their family.
Height*	Measured by interviewers.
BMI	Measured by interviewers.
Birth weight	Reported by mothers.
Failed the school year	From class register.
Italian grade	From class register. It ranges from 0 to 10. We use deviation from class mean.
Math grade	From class register. It ranges from 0 to 10. We use deviation from class mean.
Italian Invalsi grade	Standardized national test. It ranges from 0 to 300.
Math Invalsi grade	Standardized national test. It ranges from 0 to 300.
Patience*	Dummy variable. Chooses to wait one day to have two snacks instead of one. See Section 4.3.
Subjective well-being*	Children are asked: *How happy are you about your life?* Possible answers are: 1=Very sad, ... 5=Very happy.

Outcome variables	
Variable	Description
Happy*	Children are asked: *Are you happy?* Possible answers are: 0=Never, 1=Sometimes, 2=Often.
Often tired*	Self-reported. Possible answers: 0=Never, 1=Sometimes, 2=Often.
Gets along with mum*	Children are asked: *Do you get along with your mum?* Possible answers are: 0=Never, 1=Sometimes, 2=Often.
Gets along with dad*	Children are asked: *Do you get along with your dad?* Possible answers are: 0=Never, 1=Sometimes, 2=Often.
Likes chocolate*	Children are asked if they like chocolate. Possible answers range from 1=Not at all, to 5=Very much.
Likes Parmesan*	Children are asked if they like Parmesan. Possible answers range from 1=Not at all, to 5=Very much.
Likes bananas*	Children are asked if they like bananas. Possible answers range from 1=Not at all, to 5=Very much.
Chooses healthy snacks*	Children are asked to choose a snack to eat after the interview. Possible choices are: 1=chocolate, 2=Parmesan bar, 3=banana.
Gets pocket money*	1=Children report receiving at least 5€per month, 0=Otherwise.
Pocket money*	Children's self-reported amount of pocket money per month. Trimmed at 450€.
Number of daily snacks*	Self-reported.
Drinks sugary drinks*	1=Reports drinking sodas or other sugary drinks when snacking, 0=Otherwise.
Time spent on tablets*	Self-reported. Possible answers: ¡1 hour per day, 1-3 hours per day, ¿3 hours per day.

Outcome variables	
Variable	Description
Time spent on smartphones*	Self-reported. Possible answers: same as "time spent on tablets".
Time spent on TV*	Self-reported. Possible answers: same as "time spent on tablets".
Time spent on screens*	Sum of time spent on tablets, smartphones, and TV.
Time spent with parents*	Self-reported. Possible answers: 0=Seldom, 1=Often, 2=Very often.
Eats vegetables*	Measures vegetables *and* fruit consumption. Self-reported. Possible answers: 0=Never, 1=Sometimes, 2=Often.
Does sports*	Children are asked the same questions about health behavior and risk attitudes as their parents. See Table 8 for details.
Smokes*	See Table 8.
Wears sunscreen*	See Table 8.
Washes hands*	See Table 8.
Worries about health*	See Table 8.
Would go on safari*	See Table 8.
Likes holidays in known places*	See Table 8.
Would do extreme sports*	See Table 8.
Careful when crossing*	See Table 8.
Altruism I*	Children are asked: *If your classmate is in trouble, do you try to help him/her?* Possible answers are: 0=Never, 1=Sometimes, 2=Often.

Outcome variables	
Variable	Description
Altruism II*	Children are asked: *Would you give money to your classmate if he/she has no money to buy a snack?* Possible answers are: 0=Never, 1=Sometimes, 2=Often.

Notes: The table clarifies how outcome variables were measured. Some variables were built straight from the class register or direct measurement, others from the answers to the survey. All variables marked with * correspond to the residuals obtained by regressing the raw variable on the child's age.

B Additional Tables

Table 10: Frequency of children by parents' presence at home

	Father is present		
Mother is present	No	Yes	Total
No	0	3	3
Yes	16	613	629
Total	16	616	632

Notes: A parent is present if he/she participates to the survey and/or is reported as living at home by the child.

Table 11: Rank test for $\hat{A}$ (Sample 1)

H_0: $rk(A) = k$	$k = 1$	$k = 2$	$k = 3$
χ^2	70.34	17.35	1.91
df	9	4	1
Rejected?	Yes	Yes	No

Notes: Each column reports the statistic resulting from testing the null hypothesis that the rank of $\hat{A}$ is equal to k. We report whether the null hypothesis was rejected at the 1% level. These tests lead us to conclude that sorting occurs on at least 3 orthogonal dimensions.

Table 12: Rank test for $\hat{A}$ (Sample 2)

H_0: $rk(A) = k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$
χ^2	544.54	240.17	127.39	69.78	28.13	11.85	3.44	0.17	0.00
df	196	169	144	121	100	81	64	49	36
Rejected?	Yes	Yes	No	No	No	No	No	No	No

Notes: Each column reports the statistic resulting from testing the null hypothesis that the rank of $\hat{A}$ is equal to k. We report whether the null hypothesis was rejected at the 1% level. These tests lead us to conclude that sorting occurs on at least 3 orthogonal dimensions.

Table 13: Correlation of matching indices with household outcomes (Sample 1)

	Men		Women	
	Index 1	Index 2	Index 1	Index 2
Number of children	−0.00	−0.02	0.08	0.04
	(0.99)	(0.75)	(0.19)	(0.53)

	Men		Women	
	Index 1	Index 2	Index 1	Index 2
Hours outside (wife)	0.17	0.24	0.19	0.34
	(0.01)	(0.00)	(0.00)	(0.00)
Stays at home (wife)	−0.32	−0.37	−0.35	−0.51
	(0.00)	(0.00)	(0.00)	(0.00)
Parents share tasks equally	0.11	0.16	0.12	0.27
	(0.07)	(0.01)	(0.06)	(0.00)
Height (child)	0.15	0.18	0.14	0.06
	(0.02)	(0.00)	(0.03)	(0.35)
BMI (child)	0.14	−0.12	0.12	−0.16
	(0.03)	(0.05)	(0.06)	(0.01)
Failed the school year	−0.04	0.06	−0.00	0.04
	(0.53)	(0.35)	(0.95)	(0.56)
Italian grade	0.00	0.26	0.05	0.17
	(0.97)	(0.00)	(0.45)	(0.01)
Math grade	0.04	0.22	0.04	0.12
	(0.51)	(0.00)	(0.49)	(0.06)
Italian Invalsi	0.12	0.27	0.11	0.50
	(0.27)	(0.01)	(0.30)	(0.00)
Math Invalsi	0.18	0.30	0.23	0.40
	(0.09)	(0.00)	(0.03)	(0.00)
Patience (child)	−0.08	0.05	−0.05	0.04
	(0.19)	(0.38)	(0.41)	(0.47)
Subjective well-being (child)	0.00	−0.03	−0.02	−0.06
	(0.95)	(0.63)	(0.75)	(0.39)
Happy (child)	−0.06	−0.06	−0.09	−0.03
	(0.35)	(0.36)	(0.14)	(0.61)
Gets along with mom	0.02	−0.08	0.00	−0.07
	(0.81)	(0.21)	(0.96)	(0.26)
Gets along with dad	0.02	0.01	0.02	−0.08
	(0.78)	(0.85)	(0.73)	(0.23)
Likes chocolate	−0.07	−0.07	−0.03	−0.13
	(0.25)	(0.29)	(0.67)	(0.03)

	Men		Women	
	Index 1	Index 2	Index 1	Index 2
Likes Parmesan	−0.02	0.15	0.03	0.10
	(0.75)	(0.02)	(0.68)	(0.12)
Likes bananas	−0.03	−0.08	−0.04	−0.06
	(0.64)	(0.21)	(0.49)	(0.33)
Chooses healthy snacks (child)	−0.08	−0.01	0.00	0.09
	(0.21)	(0.85)	(0.98)	(0.13)
Gets pocket money (yes/no)	−0.22	−0.21	−0.21	−0.27
	(0.00)	(0.00)	(0.00)	(0.00)
Pocket money (monthly amount)	−0.19	−0.15	−0.21	−0.13
	(0.00)	(0.02)	(0.00)	(0.06)
Number of daily snacks	−0.06	0.02	−0.09	−0.01
	(0.31)	(0.81)	(0.16)	(0.82)
Drinks sugary drinks	−0.03	−0.09	−0.05	0.02
	(0.61)	(0.16)	(0.40)	(0.72)
Time spent on tablets	−0.09	−0.01	−0.10	−0.01
	(0.14)	(0.92)	(0.13)	(0.92)
Time spent on smartphones	−0.10	−0.11	−0.09	−0.14
	(0.11)	(0.07)	(0.16)	(0.02)
Time spent on TV	−0.06	−0.01	−0.09	−0.02
	(0.38)	(0.90)	(0.15)	(0.77)
Time spent on screens (total)	−0.14	−0.07	−0.16	−0.08
	(0.03)	(0.29)	(0.01)	(0.20)
Time spent with parents	−0.14	−0.00	−0.12	−0.07
	(0.02)	(1.00)	(0.05)	(0.29)
Eats vegetables	0.07	0.09	0.07	0.00
	(0.24)	(0.13)	(0.30)	(0.97)
Often tired	−0.01	−0.15	0.01	−0.09
	(0.85)	(0.02)	(0.85)	(0.15)
Smokes (child)	0.00	0.04	−0.02	0.05
	(0.97)	(0.53)	(0.76)	(0.40)
Likes sports (child)	−0.05	0.12	−0.02	0.06
	(0.44)	(0.06)	(0.73)	(0.36)

	Men		Women	
	Index 1	Index 2	Index 1	Index 2
Worries about own health (child)	−0.04	0.00	−0.04	−0.00
	(0.55)	(0.97)	(0.54)	(0.98)
Wears sunscreen (child)	0.03	−0.00	0.02	0.06
	(0.61)	(0.95)	(0.72)	(0.30)
Washes hands (child)	0.03	−0.01	0.01	0.10
	(0.61)	(0.89)	(0.83)	(0.12)
Would go on safari (child)	0.05	0.12	0.04	0.09
	(0.40)	(0.05)	(0.48)	(0.13)
Likes usual holidays (child)	−0.05	−0.03	−0.02	−0.05
	(0.40)	(0.68)	(0.72)	(0.43)
Would do extreme sports (child)	0.09	0.04	0.06	0.07
	(0.14)	(0.56)	(0.36)	(0.25)
Careful when crossing (child)	0.04	−0.01	0.03	−0.01
	(0.54)	(0.89)	(0.60)	(0.82)
Altruism I	0.02	0.07	0.00	0.08
	(0.72)	(0.30)	(0.94)	(0.19)
Altruism II	−0.00	0.11	−0.01	0.01
	(0.94)	(0.08)	(0.86)	(0.86)

Notes: 259 couples. Correlations of matching indices with household outcomes. P-values of the two-tailed significance test in parentheses.

Table 14: Correlation of matching indices with household outcomes (Sample 2)

	Men			Women		
	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3
Number of children	0.02	0.02	−0.04	0.11	0.08	0.05
	(0.74)	(0.78)	(0.54)	(0.13)	(0.26)	(0.46)
Hours outside (wife)	0.20	0.28	0.17	0.26	0.31	0.19
	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.01)
Stays at home (wife)	−0.39	−0.41	−0.26	−0.44	−0.46	−0.29
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Parents share tasks equally	0.14	0.29	0.09	0.14	0.27	0.17

	Men			Women		
	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3
	(0.04)	(0.00)	(0.20)	(0.04)	(0.00)	(0.01)
Height (child)	0.12	0.05	0.01	0.10	0.03	0.03
	(0.09)	(0.46)	(0.84)	(0.17)	(0.68)	(0.68)
BMI (child)	0.13	−0.12	−0.09	0.11	−0.08	−0.09
	(0.06)	(0.09)	(0.21)	(0.12)	(0.26)	(0.18)
Failed the school year	−0.03	0.03	−0.01	0.04	0.02	0.06
	(0.68)	(0.67)	(0.89)	(0.57)	(0.74)	(0.37)
Italian grade	0.03	0.18	0.23	0.05	0.11	0.14
	(0.68)	(0.01)	(0.00)	(0.51)	(0.12)	(0.05)
Math grade	0.07	0.13	0.22	0.04	0.05	0.10
	(0.32)	(0.06)	(0.00)	(0.54)	(0.52)	(0.14)
Italian Invalsi	0.17	0.00	0.18	0.30	0.26	0.39
	(0.17)	(0.98)	(0.16)	(0.01)	(0.04)	(0.00)
Math Invalsi	0.31	0.08	0.10	0.41	0.25	0.28
	(0.01)	(0.52)	(0.42)	(0.00)	(0.05)	(0.02)
Patience (child)	−0.10	−0.03	0.06	−0.08	−0.11	0.02
	(0.15)	(0.66)	(0.36)	(0.23)	(0.12)	(0.80)
Subjective well-being (child)	−0.03	−0.03	0.07	−0.04	−0.05	−0.04
	(0.63)	(0.65)	(0.30)	(0.59)	(0.50)	(0.54)
Happy (child)	−0.10	−0.22	0.14	−0.15	−0.18	0.07
	(0.14)	(0.00)	(0.05)	(0.04)	(0.01)	(0.35)
Gets along with mom	−0.05	−0.14	−0.01	−0.08	−0.17	−0.02
	(0.45)	(0.05)	(0.87)	(0.25)	(0.02)	(0.79)
Gets along with dad	0.04	−0.07	0.04	0.02	−0.08	0.06
	(0.58)	(0.33)	(0.58)	(0.79)	(0.23)	(0.37)
Likes chocolate	−0.12	−0.17	0.03	−0.11	−0.21	−0.02
	(0.10)	(0.02)	(0.64)	(0.12)	(0.00)	(0.79)
Likes Parmesan	−0.01	0.18	0.04	0.02	0.10	0.06
	(0.90)	(0.01)	(0.54)	(0.77)	(0.16)	(0.43)
Likes bananas	−0.05	0.05	−0.03	−0.04	0.08	−0.01
	(0.45)	(0.46)	(0.71)	(0.56)	(0.29)	(0.89)
Chooses healthy snacks (child)	−0.00	0.14	0.01	0.04	0.16	−0.01

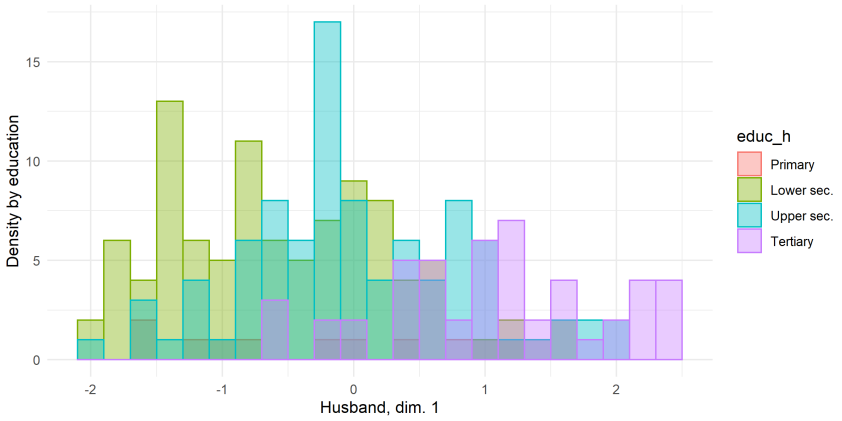
	Men			Women		
	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3
	(0.98)	(0.05)	(0.86)	(0.59)	(0.02)	(0.92)
Gets pocket money (yes/no)	−0.29	−0.14	−0.13	−0.28	−0.12	−0.14
	(0.00)	(0.05)	(0.06)	(0.00)	(0.08)	(0.04)
Pocket money (monthly amount)	−0.22	−0.07	−0.05	−0.21	−0.02	−0.03
	(0.00)	(0.39)	(0.51)	(0.00)	(0.81)	(0.74)
Number of daily snacks	−0.07	−0.03	0.01	−0.07	−0.02	−0.02
	(0.29)	(0.70)	(0.90)	(0.35)	(0.73)	(0.80)
Drinks sugary drinks	−0.01	−0.00	0.00	−0.04	0.10	−0.02
	(0.84)	(0.97)	(0.95)	(0.53)	(0.15)	(0.80)
Time spent on tablets	0.00	0.08	0.02	0.01	0.11	−0.05
	(0.96)	(0.25)	(0.74)	(0.90)	(0.13)	(0.51)
Time spent on smartphones	−0.17	−0.07	−0.05	−0.17	−0.07	−0.13
	(0.02)	(0.30)	(0.50)	(0.01)	(0.31)	(0.06)
Time spent on TV	−0.11	−0.07	0.13	−0.11	0.01	0.16
	(0.13)	(0.30)	(0.07)	(0.12)	(0.87)	(0.03)
Time spent on screens (total)	−0.13	0.00	0.04	−0.14	0.04	−0.02
	(0.07)	(0.99)	(0.61)	(0.05)	(0.55)	(0.78)
Time spent with parents	−0.15	−0.10	0.07	−0.13	−0.08	0.01
	(0.03)	(0.15)	(0.35)	(0.06)	(0.24)	(0.92)
Eats vegetables	0.09	−0.01	0.09	0.08	0.04	−0.07
	(0.21)	(0.87)	(0.21)	(0.27)	(0.55)	(0.33)
Often tired	−0.03	0.01	−0.09	−0.01	0.01	−0.04
	(0.64)	(0.90)	(0.21)	(0.84)	(0.91)	(0.54)
Smokes (child)	0.02	0.11	−0.09	0.02	0.13	0.03
	(0.76)	(0.13)	(0.19)	(0.74)	(0.07)	(0.72)
Likes sports (child)	−0.07	0.01	0.15	−0.07	−0.03	0.09
	(0.29)	(0.89)	(0.03)	(0.30)	(0.71)	(0.22)
Worries about own health (child)	−0.03	0.01	0.08	−0.04	0.00	−0.00
	(0.68)	(0.88)	(0.28)	(0.55)	(0.96)	(0.96)
Wears sunscreen (child)	0.06	0.08	−0.03	0.03	0.04	−0.06
	(0.39)	(0.25)	(0.66)	(0.68)	(0.52)	(0.43)
Washes hands (child)	0.02	−0.04	−0.02	0.01	−0.01	0.08

	Men			Women		
	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3
	(0.75)	(0.62)	(0.74)	(0.85)	(0.88)	(0.23)
Would go on safari (child)	0.03	0.11	0.20	0.03	0.14	0.15
	(0.67)	(0.11)	(0.00)	(0.71)	(0.05)	(0.04)
Likes usual holidays (child)	−0.07	−0.05	0.02	−0.07	−0.06	−0.08
	(0.33)	(0.46)	(0.73)	(0.34)	(0.43)	(0.23)
Would do extreme sports (child)	0.07	0.05	0.06	0.07	0.05	0.09
	(0.31)	(0.49)	(0.38)	(0.32)	(0.45)	(0.18)
Careful when crossing (child)	0.12	0.00	−0.03	0.08	−0.07	−0.01
	(0.09)	(0.98)	(0.63)	(0.26)	(0.32)	(0.92)
Altruism I	0.06	0.09	0.06	0.03	0.08	0.04
	(0.37)	(0.20)	(0.40)	(0.71)	(0.25)	(0.54)
Altruism II	0.04	0.11	0.14	0.03	0.05	0.03
	(0.53)	(0.12)	(0.04)	(0.71)	(0.51)	(0.71)

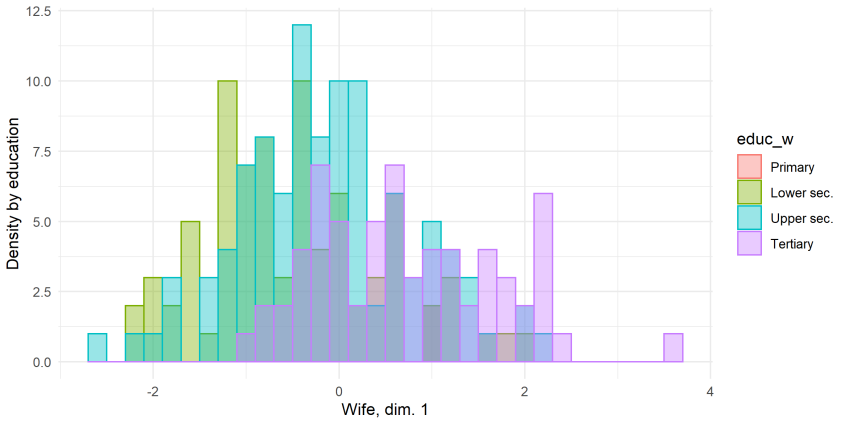
Notes: 205 couples. Correlations of matching indices with household outcomes. P-values of the two-tailed significance test in parentheses.

C Additional Figures

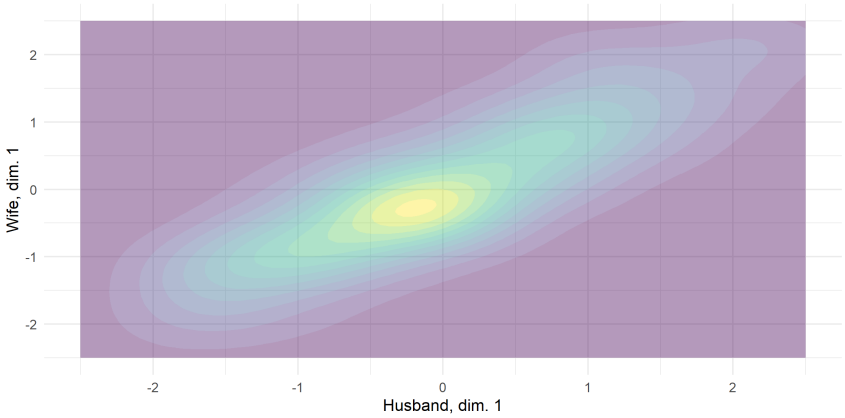
Figure 6: Distribution of Index 1 by education and gender (Sample 1)



(a) Men



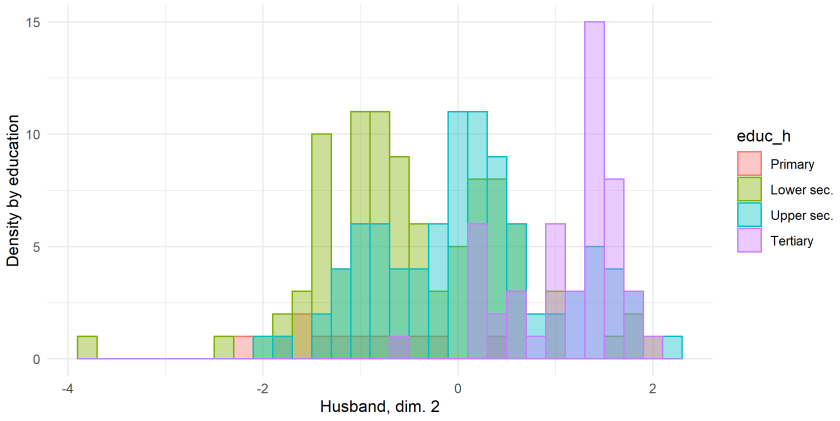
(b) Women



(c) Joint

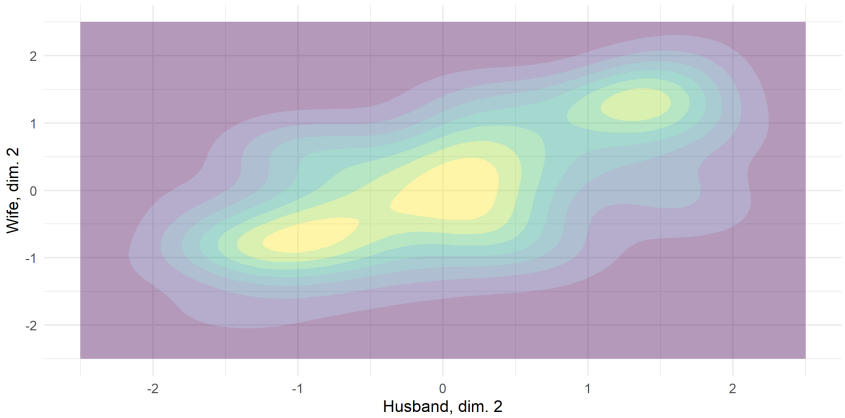
Notes: We plot the distribution of the men's Index 1 by their education (a), the distribution of women's Index 1 by their education (b), and the joint distribution of men's Index 1 and women's Index 1 (c).

Figure 7: Distribution of Index 2 by education and gender (Sample 1)



(a) Men

(b) Women



(c) Joint

Notes: We plot the distribution of the men's Index 2 by their education (a), the distribution of women's Index 2 by their education (b), and the joint distribution of men's Index 2 and women's Index 2 (c).

References

Paul R. Amato and Alan Booth. Changes in gender role attitudes and perceived marital quality. *American Sociological Review*, 60(1):58–66, 1995.

Christine Barnet-Verzat and François-Charles Wolff. Motives for pocket money allowance and family incentives. *Journal of Economic Psychology*, 23(3):339–366, 2002.

Gary S Becker. A theory of marriage: Part i. *Journal of Political Economy*, 81(4): 813–846, 1973.

Alberto Bisin and Thierry Verdier. "beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3):955–988, 2000.

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319, 2001.

Alberto Bisin, Giorgio Topa, and Thierry Verdier. Religious intermarriage and socialization in the united states. *Journal of political Economy*, 112(3):615–664, 2004.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439, 2007.

Debra L Blackwell and Daniel T Lichter. Homogamy among dating, cohabiting, and married couples. *The Sociological Quarterly*, 45(4):719–737, 2004.

Ann-Renée Blais and Elke U Weber. A domain-specific risk-taking (dospert) scale for adult populations. *Judgment and Decision Making*, 1(1):33, 2006.

Milan Bouchet-Valat. Changes in educational, social class and social class of origin homogamy in france (1969-2011). *Revue Française de Sociologie*, 55(3):459–505, 2014.

Jonathan Bradshaw. Subjective well-being and social policy: Can nations make their children happier? *Child Indicators Research*, 8(1):227–241, 2015.

Anne Case and Christina Paxson. Parental behavior and child health. *Health Affairs*, 21(2):164–178, 2002.

Marco Castillo, Jeffrey L Jordan, and Ragan Petrie. Discount rates of children and high school graduation. *Economic Journal*, 129(619):1153–1181, 2019.

Pierre-André Chiappori, Robert J McCann, and Lars P Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Economic Theory*, 42(2):317–354, 2010.

Pierre-André Chiappori, Bernard Salanié, and Yoram Weiss. Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–67, 2017.

Pierre-André Chiappori, Monica Costa Dias, and Costas Meghir. Changes in assortative matching: Theory and evidence for the us. Technical report, National Bureau of Economic Research, 2020.

Shyamal Chowdhury, Matthias Sutter, and Klaus Zimmermann. Evaluating intergenerational persistence of economic preferences: A large scale experiment with families in bangladesh. Technical report, 2018. Working paper.

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household choices and child development. *Review of Economic Studies*, 81(1):137–185, 2014.

Tamar Dinisman and Asher Ben-Arieh. The characteristics of children's subjective well-being. *Social Indicators Research*, 126(2):555–569, 2016.

Matthias Doepke and Fabrizio Zilibotti. Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5):1331–1371, 2017.

Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2):645–677, 2012.

Arnaud Dupuy and Alfred Galichon. Personality traits and the marriage market. *Journal of Political Economy*, 122(6):1271–1319, 2014.

Arnaud Dupuy, Alfred Galichon, and Yifei Sun. Estimating matching affinity matrices under low-rank constraints. *Information and Inference: A Journal of the IMA*, 8(4):677–689, 2019.

Raquel Fernández and Richard Rogerson. Sorting and long-run inequality. *Quarterly Journal of Economics*, 116(4):1305–1341, 2001.

Jennifer Orlet Fisher, Diane C Mitchell, Helen Smiciklas-Wright, and Leann Lipps Birch. Parental influences on young girls' fruit and vegetable, micronutrient, and fat intakes. *Journal of the American Dietetic Association*, 102(1):58–64, 2002.

Pilar Gonalons-Pons and Christine R Schwartz. Trends in economic homogamy: changes in assortative mating or the division of labor in marriage? *Demography*, 54(3):985–1005, 2017.

Mònica González-Carrasco, Sara Malo, Ferran Casas, Gemma Crous, Mireia Baena, and Dolors Navarro. The assessment of subjective well-being in young children: Strengths and challenges. *Journal of Social Research & Policy*, 6(2), 2015.

Jeremy Greenwood, Nezih Guner, and John A Knowles. More on marriage, fertility, and the distribution of income. *International Economic Review*, 44(3):827–862, 2003.

Jeremy Greenwood, Nezih Guner, Georgi Kocharkov, and Cezar Santos. Marry your like: Assortative mating and income inequality. *American Economic Review*, 104 (5):348–53, 2014.

Carla Guerriero, John Cairns, Fabrizio Bianchi, and Liliana Cori. Are children rational decision makers when they are asked to value their own health? a contingent valuation study conducted with children and their parents. *Health Economics*, 27 (2):e55–e68, 2018.

Carolyn J Heinrich. Parents' employment and children's wellbeing. *The Future of Children*, 24(1):121–146, 2014.

Gayle Kaufman. Do gender role attitudes matter? family formation and dissolution among traditional and egalitarian men and women. *Journal of Family Issues*, 21 (1):128–144, 2000.

Martha Henn McCormick. The effectiveness of youth financial education: A review of the literature. *Journal of Financial Counseling and Planning Volume*, 20(1): 71, 2009.

Luis Milan and Joe Whittaker. Application of the parametric bootstrap to models that incorporate a singular value decomposition. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 44(1):31–49, 1995.

Walter Mischel, Yuichi Shoda, and Monica I Rodriguez. Delay of gratification in children. *Science*, 244(4907):933–938, 1989.

AA Nima, KM Cloninger, BN Persson, S Sikström, and D Garcia. Validation of subjective well-being measures using item response theory. *Frontiers in psychology*, 10:3036, 2020.

Zhenchao Qian and Samuel H Preston. Changes in american marriage, 1972 to 1987: Availability and forces of attraction by age and education. *American Sociological Review*, 58(4):482–495, 1993.

Robert Schoen. The harmonic mean as the basis of a realistic two-sex marriage model. *Demography*, 18(2):201–216, 1981.

Christine R Schwartz and Hongyun Han. The reversal of the gender gap in education and trends in marital dissolution. *American Sociological Review*, 79(4):605–629, 2014.

Christine R Schwartz and Robert D Mare. Trends in educational assortative marriage from 1940 to 2003. *Demography*, 42(4):621–646, 2005.

Mary Suiter and B Meszaros. Teaching about saving and investing in the elementary and middle school grades. *Social Education*, 69(2):92–95, 2005.

Henna Vepsäläinen, Jaakko Nevalainen, Mikael Fogelholm, Liisa Korkalo, Eva Roos, Carola Ray, and Maijaliisa Erkkola. Like parent, like child? dietary resemblance in families. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1):1–11, 2018.

Tyler W Watts, Greg J Duncan, and Haonan Quan. Revisiting the marshmallow test: A conceptual replication investigating links between early delay of gratification and later outcomes. *Psychological science*, 29(7):1159–1177, 2018.

Elke U Weber, Ann-Renee Blais, and Nancy E Betz. A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, 15(4):263–290, 2002.