DISCUSSION PAPER SERIES

The Empirical Content of Marital Surplus in Matching Models

Karina Doorley
Arnaud Dupuy
Simon Weber

The Empirical Content of Marital Surplus in Matching Models

Karina Doorley
ESRI, TCD and IZA

Arnaud Dupuy
CREA, University of Luxembourg and IZA

Simon Weber
KU Leuven

ABSTRACT

The Empirical Content of Marital Surplus in Matching Models

This note investigates the extent to which structural estimates of marital surplus are informative about subjective well-being and separation. We first estimate the marital surplus using a simple matching model of the marriage market with perfectly transferable utility and heterogeneity in tastes applied to a rich German panel dataset. We then show that these estimates of the marital surplus are negatively correlated with separation and the difference in spouses' subjective satisfaction.

Corresponding author:
Karina Doorley
The Economic and Social Research Institute
Whitaker Square
Sir John Rogerson's Quay
Dublin 2
Ireland
E-mail: karina.doorley@esri.ie

KARINA DOORLEY[‡], ARNAUD DUPUY[§], SIMON WEBER[†]

1. Introduction

In a seminal paper, Choo and Siow (2006) introduced an empirically tractable methodology to study two-sided matching models with perfectly transferable utility and heterogeneity in tastes, also known as TU matching models. The key feature of this methodology is that it recasts TU models as a set of two-sided discrete choice problems. This allows the analyst to apply tools from discrete choice random utility models to estimate preferences using a single cross-section of data. Following this breakthrough, a fast growing literature has estimated marital preferences using a wide range of attributes such as education, BMI, personality traits etc., see e.g. Choo and Siow (2006), Dupuy and Galichon (2014), Chiappori, Salanie, and Weiss (2017), Ciscato and Weber (2017), Ciscato, Galichon, and Goussé (2015), Dupuy (2018).

The marital preferences estimated by this method are, by construction, consistent with the marital patterns observed in the data: the estimated marital surplus is larger for types of couples that are more prevalent in the data once demographics have been accounted for. Realized marital choices of individuals are taken as input and the marital surplus which rationalizes these choices is inferred. Yet, little is known about whether these estimates of marital preferences provide valuable information about spouses' well-being or marital stability.

The aim of this paper is to provide a first assessment of the empirical content of marital preferences as estimated using TU matching models. Our approach consists of comparing marital surplus estimates with measures of subjective well-being and separation. To this aim, we first use the 2005 wave of the German SOcio-Economic Panel (GSOEP) and a parametric version of the Choo and Siow (2006) model to estimate couples' marital surplus. Second, we investigate the extent to which the estimated marital surplus is associated with measures of subjective satisfaction and, using later waves of data, the probability of separation.

2. The Choo and Siow Model

Consider a two-sided one-to-one matching model with transferable utility. Men and women, who aim at forming a heterosexual pair, are heterogenous and we denote $x_i \in \mathcal{X}$ ($y_j \in \mathcal{Y}$) the observed type of a (wo)man i (j). Individuals have the option to remain single which is accounted for by including 0 in individuals' choice set, i.e. (wo)men can choose a partner among set $\mathcal{Y}^0 = \mathcal{Y} \cup \{0\}$ ($\mathcal{X}^0 = \mathcal{X} \cup \{0\}$). There is a mass n_x (m_y) of (wo)men of type x (y). We assume a large market so that there is a large mass of agents of each type in the market.

A couple formed of a man i of type x and a woman j of type y derives a joint surplus[1]

$$\Phi_{xy} + \varepsilon_{iy} + \eta_{xj},$$

where ε_{iy} (η_{xj}) is the idiosyncratic preference of (wo)man i (j) for a (wo)man of type y (x), assumed to be i.i.d. distributed across individuals and following a $(0,1)-$Gumbel type I distribution.[2] Utility being transferable, the joint surplus Φ_{xy} is shared among spouses, the husband getting U_{xy}, the wife $V_{xy} = \Phi_{xy} - U_{xy}$.

Men and women maximize their utility. A (wo)man i (j) of type x (y) hence solves

$$(2.1) \qquad u_i = \max_{y \in \mathcal{Y}^0} \left(U_{xy} + \varepsilon_{iy}, \varepsilon_{i0} \right), \text{ respectively } v_j = \max_{x \in \mathcal{X}^0} \left(V_{xy} + \eta_{xj}, \eta_{0j} \right),$$

where u_i (v_j) are the payoffs of (wo)man i (j).

Given U (V), the optimal solutions to programs 2.1 generate conditional distributions of (wo)men of type x (y) marrying type y (x) (wo)men, say $\mu_{y|x}$ and $\mu_{x|y}$ respectively. It follows that:

---

[1]Without loss of generality, the systematic utility of remaining single for agents of each type is normalized to 0, which allows us to use the term surplus rather than utility for married agents.

[2]Galichon and Salanie (2017) have shown that the methodology can be extended to more general distributional assumptions.

Definition 1. An equilibrium outcome consists of a feasible matching $\mu \in \mathcal{M}(n, m)$ and payoffs (u, v) such that solutions to utility maximization (program 2.1) generate

$$\mu_{y|x} \times n_x = \mu_{xy} = \mu_{x|y} \times m_y,$$

where $\mu_{xy} \geq 0$ is the mass of couples of type (x, y) and

$$\mathcal{M}(n, m) = \left\{ \mu | \sum_{y \in \mathcal{Y}^0} \mu_{xy} = n_x \; \forall x \in \mathcal{X}, \sum_{x \in \mathcal{X}^0} \mu_{xy} = m_y \; \forall y \in \mathcal{Y} \right\}$$

is the set of feasible matchings.

A well-known result by Koopmans and Beckmann (1957) and Shapley and Shubik (1971) states that there exists an equilibrium outcome, which is i) unique in large markets (Gretsky, Ostroy, and Zame, 1992; Gretsky, Ostroy, and Zame, 1999), ii) Pareto optimal and iii) pairwise stable.

Denoting $G_x(U_{x\cdot})$ the expected indirect utility of a man of type x, one has

$$G_x(U_{x\cdot}) := \mathbb{E}\left[\max_{y \in \mathcal{Y}} (U_{xy} + \varepsilon_{iy}, \varepsilon_0) \right] = \log(1 + \sum_{y \in \mathcal{Y}} \exp(U_{xy})),$$

with a similar expression for women. It follows that an application of the Daly-Zachary-Williams theorem yields a logit structure so that one uncovers expressions of U and V in terms of the matching as

$$(2.2) \qquad U_{xy} = \log \frac{\mu_{xy}}{\mu_{x0}}, \; V_{xy} = \log \frac{\mu_{xy}}{\mu_{0y}},$$

and hence an expression of the joint surplus Φ in terms of the matching as

$$(2.3) \qquad \Phi_{xy} = 2\log \frac{\mu_{xy}}{\sqrt{\mu_{x0}\mu_{0y}}}.$$

Interestingly, equation 2.3 can easily be inverted to yield an expression for the equilibrium mass of couples of type (x, y) given the equilibrium mass of singles of

type x and y respectively and the joint surplus generated by couples (x, y) as

$$(2.4) \qquad \mu_{xy} = \sqrt{\mu_{x0}\mu_{0y}} \exp\left(\frac{\Phi_{xy}}{2}\right).$$

Plugging this expression into the feasibility constraints in Definition 1 and using a simple iterative algorithm (IPFP, see Galichon, Kominers and Weber, 2015), one can compute the unique equilibrium matching μ given $(\Phi_{xy})_{x,y}$.

3. Empirical analysis

3.1. **Data.** We use the GSOEP data which provides detailed information on a rich set of individual and household characteristics. For our analysis, the base year is set at 2005, In this wave, a rich set of individual characteristics including, for the first time, personality traits, are available. We keep all men (women) who are aged no less than 25 (23) and no more than 49 (47) in 2005.[3] These individuals are then matched to their post-2005 marriage spells and subjective well-being.

The vectors of attributes X and Y which couples are expected to match along include the following 11 variables: age, height, BMI, self-assessed health status, years of education, willingness to take risks and the big 5 measures of personality.[4]

The same individuals are asked questions about their well-being. For each couple in the 2005 wave, we therefore have, among others, each spouse's subjective satisfaction with life, satisfaction with the household's income, whether they have children, who according to each spouse has the last word in household's financial decisions, each spouse's degree of pessimism about the future and whether each spouse attends

---

[3]The age selection is standard in this literature and changes in the selection did not alter the main conclusion of our analysis.

[4]See the online appendix A for more details about the construction of these variables as well as for some descriptive statistics.

church. Moreover, using later waves, one can track 2005 spouses' marital spells and hence know whether these couples got separated between 2005 and 2016.[5]

3.2. **Empirical methodology.** The 2005 wave of the panel provides us with a representative sample of $\hat{N}$ households. Each household h can consist of either a couple, for which we observe the pair of vectors of attributes (X_h, Y_h) or a single (wo)man, for which we observe the vector X_h (resp. Y_h), where X_h and Y_h are of size $K + 1$, with the convention that the last element is a constant.

We assume the surplus function is linear in parameters and reads as the sum of three terms:

$$(3.1) \qquad \Phi_{xy}^{\lambda} = \sum_{k=1}^{K} \sum_{l=1}^{K} \lambda_{kl} \left| x^{(k)} - y^{(l)} \right| + \sum_{z=x,y} \sum_{p=1}^{3} \sum_{k=1}^{K} \lambda_{k,z}^{p} \left(z^{(k)} \right)^{p} + \lambda_0,$$

where $\lambda = \{ (\lambda_{kl})_{kl}, \left(\lambda_{k,x}^{p} \right)_{k,p}, \left(\lambda_{k,y}^{p} \right)_{k,p}, \lambda_0 \}$ is the vector of parameters to be estimated. The last term is a constant, the second indicates the main contribution of the attributes of both the husband and the wife and takes the form of a polynomial of order 3, while the first term indicates the contribution of the interaction between husband's and wife's attributes.[6] All attributes are standardized to facilitate comparison between coefficients without loss of generality.

Following Galichon and Salanie (2017), we use a matching moment estimator which consists in finding parameter λ such that the observed co-moments are equal to the

---

[5]Note that questions about well-being are asked in each wave between 2005 and 2016. Our results are broadly consistent with whichever wave we use to construct measures of well-being and horizon of separation.

[6]Increasing the order of the polynomial improves the fit of the model but since some coefficients λ_k^3 and λ_l^3 are significantly different from 0 we set it to 3. The absolute difference specification is shown to provide a better fit than a bilinear specification. See online appendix B.

predicted ones, that is, for instance for the term $\left|x^{(k)} - y^{(l)}\right|$,

$$\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \hat{\mu}_{xy} \left|x^{(k)} - y^{(l)}\right| - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \mu_{xy}^{\lambda} \left|x^{(k)} - y^{(l)}\right| = 0,$$

where the observed matching $\hat{\mu}$ is computed from the data and the equilibrium matching given parameter λ, μ^{λ}, is computed using Galichon, Kominers and Weber (2015)'s IPFP Algorithm.

We then use the estimated parameters λ and the resulting equilibrium matching μ^{λ} in equations 2.3 and 2.2 to compute, for each couple h observed in wave 2005, the marital surplus $\Phi_{x_h y_h}^{\lambda}$ as well as a measure of the sharing of the surplus between spouses, defined as $\left|U_{x_h y_h}^{\lambda} - V_{x_h y_h}^{\lambda}\right|$. The larger this term, the more unequally the surplus is shared between spouses.

3.3. **Results.** Our estimates of preference parameters show standard features. The interactions between spouses' characteristics $(\lambda_{kl})_{kl}$ are negative for all $k = l$, suggesting that like attracts like on age, education, subjective health, attitude towards risk, height, BMI and the big 5 personality traits. The magnitude of these interactions is larger for age and education. Lastly, for each k and $l \neq k$, $\lambda_{kk} > \lambda_{kl}$.[7]

We now evaluate the extent to which the marital surplus $\Phi_{x_h y_h}^{\lambda}$ and how it is shared $\left|U_{x_h y_h}^{\lambda} - V_{x_h y_h}^{\lambda}\right|$ correlate with measures of well-being and separation. Table 1 presents these correlations for a selected number of variables. Results show that a higher marital surplus is significantly associated with a lower probability of becoming separated by 2016, a higher probability of having children and a more equal subjective life satisfaction between the two spouses. However, we also find that the marital surplus is not correlated with own life satisfaction or household's total life satisfaction although it does correlate significantly and positively with women's satisfaction with household income. Regarding our measure of how the surplus is shared, we find that

---

[7]See table 4 in the online appendix.

unequal sharing is not significantly associated with separation. However, it is negatively and significantly associated with households' total subjective life satisfaction and positively and significantly associated with the spouses' gap in subjective life satisfaction. Moreover, the more unequally the estimated marital surplus is shared among spouses, the lower the individual spouse's subjective life satisfaction (with life and with household's income), and for men, the higher their degree of pessimism towards the future.

Finally, interestingly enough, our results show that a higher marital surplus is also associated with a higher probability that spouses agree on the fact that the husband has the last word in the household's financial decisions and a higher probability that each spouse attends church. The more unequally the estimated marital surplus is shared among spouses, the lower spouse's church attendance. We interpret these results as an indication that couples with higher marital surplus tend to be more traditional in their gender roles and that couples with higher and more equally shared marital surplus have a stronger attachment to the "institution of marriage" which, in Germany, is mainly church-based. Further research could investigate the mechanisms behind these associations.

4. Conclusion

The aim of this paper was to provide a first assessment of the empirical content of marital preferences, estimated using Choo and Siow (2006)'s methodology. Our results show that the estimated marital surplus and its (unequal) sharing among spouses is significantly associated with many variables of subjective well-being and with marial dissolution, with the intuitively expected signs. We take our results as providing further empirical validation of the Choo and Siow methodology and its extensions recently proposed in the literature. Finally, this note suggests that more insights could be gained by extending the present static framework to the dynamic case, although this is left for future research.

TABLE 1. Correlations

	Marital surplus $(\Phi^{\lambda}_{x_h y_h})$	Sharing of surplus $(|U^{\lambda}_{x_h y_h} - V^{\lambda}_{x_h y_h}|)$
A. Couples.		
Separation	-0.096***	-0.011
$|U^{\lambda}_{x_h y_h} - V^{\lambda}_{x_h y_h}|$	0.010	1.000
Children in HH	0.099***	-0.015
Gross HH income (log)	0.010	-0.019
Last word: man	0.039**	-0.018
Last word: wom	-0.007	0.019
Last word: both	0.004	-0.005
Last word: disagree	-0.029	0.007
Total life satisfaction	0.005	-0.067***
|Life satisfaction diff.|	-0.064***	0.041**
B. Married men.		
Sat. with life (men)	-0.007	-0.078***
Sat. with HH inc. (men)	-0.006	-0.058***
Pessimistic (men)	0.016	0.064***
Church attendance (men)	0.076***	-0.042**
C. Married women.		
Sat. with life (wom)	0.025	-0.035*
Sat. with HH inc. (wom)	0.039**	-0.037*
Pessimistic (wom)	0.010	0.028
Church attendance (wom)	0.046**	-0.033*

Note: *** means sig at 1%, ** sig at 5% and * sig at 10%. This table presents correlations between our measure of marital surplus $\Phi^{\lambda}_{x_h y_h}$ and selected variables. Variables at the household level include: a dummy indicating separation by 2016; the estimated utility gap (or sharing of marital surplus) $|U^{\lambda}_{x_h y_h} - V^{\lambda}_{x_h y_h}|$; a dummy indicating whether there are children in the household; the log of gross household income; a set of four dummies indicating whether the final word regarding financial decisions goes to the man, the woman, both or if they disagree on who has the last word; the sum of the partners' subjective life satisfaction and the absolute value of the difference in subjective life satisfaction. At the individual level (men and women), we use the following variables: subjective satisfaction with life; subjective satisfaction with household income; a subjective measure of pessimistic attitude towards the future; and a dummy indicating whether the individual goes to church.

10   KARINA DOORLEY[‡], ARNAUD DUPUY[§], SIMON WEBER[†]

References

[1] Pierre-Andre Chiappori, Bernard Salanie, and Yoram Weiss. "Partner Choice, Investment in Children, and the Marital College Premium". In: *American Economic Review* 107.8 (Aug. 2017), pp. 2109–67.

[2] Eugene Choo and Aloysius Siow. "Who Marries Whom and Why". In: *Journal of Political Economy* 114.1 (2006), pp. 175–201.

[3] Edoardo Ciscato, Alfred Galichon, and Marion Goussé. "Like Attract Like? A Structural Comparison of Homogamy Across Same-Sex and Different-Sex Households". In: *Working Paper* (2015), pp. –.

[4] Edoardo Ciscato and Simon Weber. "The Role of Evolving Marital Preferences in Growing Income Inequality". In: *Working Paper* (2017).

[5] Arnaud Dupuy. "Migration in china: to work or to wed?" In: *Working Paper* (2018).

[6] Arnaud Dupuy and Alfred Galichon. "Personality Traits and the Marriage Market". In: *Journal of Political Economy* 122.6 (2014), pp. 1271–1319.

[7] Neil E. Gretsky, Joseph M. Ostroy, and William R. Zame. "Perfect Competition in the Continuous Assignment Model". In: *Journal of Economic Theory* 88.1 (1999), pp. 60 –118. ISSN: 0022-0531.

[8] Neil E. Gretsky, Joseph M. Ostroy, and William R. Zame. "The nonatomic assignment model". In: *Economic Theory* 2.1 (1992), pp. 103–127. ISSN: 1432-0479.

[9] T.C. Koopmans and M.J. Beckmann. "Assignment Problems and the Location of Economic Activities". In: *Econometrica* 25.1 (1957), pp. 53–76.

[10] Lloyd S Shapley and Martin Shubik. "The Assignment Game I: The Core". In: *International Journal of Game Theory* 1.1 (1971), pp. 111–130.

Online Appendix to "The empirical content of marital surplus in matching models "

by Karina Doorley, Arnaud Dupuy and Simon Weber

Appendix A provides a more complete description of the data used to estimate marital surplus as well as descriptive statistics. Appendix B discusses the various parametric specifications of the surplus that were tested and appendix C provides parameter estimates for our preferred specification.

Appendix A. Data

We use micro-data from the German Socio-Economic Panel Study (SOEP) which is a long-ranging panel data (1984–2016) providing detailed information on a rich set of individual and household characteristics. The SOEP includes a separate marital spell module which provides details on the start and end dates of each marriage spell. The base year t_0 is set at 2005 and corresponds to the first year that personality traits are recorded in the SOEP. We keep all individuals who are aged no less than 25 and no more than 49 in 2005. These individuals are then matched to their marriage spell history.

Self-assessed height in centimeters and body weight in kilograms are recorded in 2004 and BMI is constructed from this information. Self-assessed health status is recorded every year on a scale of 1 ("very good") to 5 ("bad") and the value of this variable in 2005 (or, if this is missing, the closest previous assessment) is retained for the analysis. Education level is available in years. The individual's general willingness to take risks is recorded in 2004 on a scale of 0 ("not at all willing to take risks") to 10 ("very willing to take risks"). Life satisfaction ("How satisfied are you with your life, all things considered") is recorded every year as a categorical variable that takes the values 0-10 (where 0 is "totally unsatisfied" and 10 is "totally satisfied").

We also have information about personality traits. In 2005 (and 2009), the SOEP included a condensed assessment of established psychological personality inventories to measure the "Big five" personality factors: openness, conscientiousness, extraversion, agreeableness and neuroticism. Respondents were asked to rate their agreement with fifteen different statements about themselves, three for each personality trait, on 7-point Likert scales (1: does not apply to me at all, 7: applies to me perfectly). Following Caliendo, Künn, and Weißenberger (2016), we obtain a respondent's score for a particular personality characteristic by averaging the scores of each statement relating to that characteristic, after adjusting for reverse coding where necessary.

Donnellan and Lucas (2008) demonstrated that each of the scales in the SOEP correlates highly with the corresponding scale in the full Big Five Inventory while Boyce, Wood, and Ferguson (2016) found that the retest reliability in the German SOEP between 2005 and 2009 was satisfactory, indicating little systematic change to the scores over time.

Descriptive statistics for our sample for the year 2005 are displayed in table 2.[8] Statistics are displayed separately for married and single individuals. Panel A, which summarizes the data for women, indicates that there are 2,601 married women and 731 single women in the sample. The average age of women in the sample is 37 and the average years of education is 12.5. There is little difference between the age or education of married and single women. Among married women, 70% have children in the household while the corresponding figure for single women is 41%. Gross household income is higher for married women than for single women. Height is similar across marital status at around 167 cm while average BMI is 24.1 for married women and 23.4 for single women. Self assessed health is similar for married and single women at 2.1 for married women and 2.2 for single women on a scale of 1-5. For married women, life satisfaction is higher (7.1), satisfaction with household income is higher (6.3) and willingness to take risks is lower (4.2) than for single women (6.6, 5.1, and 4.7 respectively). When asked their attitude towards the future, married women and single women are equally optimistic/pessimistic at around 2.1 on a scale of 1 (optimistic) to 4 (pessimistic). Married women are more likely (0.5) to attend church than single women (0.3). Lastly, average values for the Big Five are very similar for married and single women. On the 7-point Likert scale, average openness is around 4.6, conscientiousness is 6, extraversion is 5, agreeableness is 5.6 and neuroticism is 4.2.

---

[8]Figures for height, BMI and willingness to take risks relate to the year 2004 (the closest available year).

Panel B depicts the same statistics for married men and single men. The number of married men in the sample is 2,601, identical to the number of married women, There are also 642 single men in the sample. Married men are older, on average, than single men (39 and 37 respectively). Years of education are similar for each group and similar to the average of 12.5 for women. 70% of married men have children in their household compared to just 3% of single men. Gross household income is larger for married men than single men. Average male height is 179 cm and there is little difference between married and single men. Married men have a higher BMI (26.2) than single men (24.9) although their self-assessed health is very similar (and similar to that of women) at around 2 on a scale of 1-5. Like women, married men have higher life satisfaction (7) and satisfaction with household income (6.2) than single men (6.6 and 5.5). While mens' willingness to take risks is higher than that of women, the pattern is similar with married men being less willing to take risks (5.2) than single men (5.4). When asked their attitude towards the future, married men and single men are equally optimistic/pessimistic at around 2.1 on a scale of 1 (optimistic) to 4 (pessimistic). Married men are more likely (0.5) to attend church than single men (0.3). Average values for the Big Five are very similar for married and single men and are comparable to the results for women. On the 7-point Likert scale, average openness is around 4.4, conscientiousness is 5.9, extraversion is 4.8, agreeableness is 5.3 and neuroticism is 3.6.

Panel C shows couple level variables. The sum of life satisfaction of both spouses is 14 while the absolute difference between the life satisfaciton of spouses is 1.2. When asked who has the last word on financial decisions in the household, most couples are in agreement that both spouses share this equally (82%). A small minority state that the woman (3%) or the man (5%) has the last word while 10% of couples disagree on who has the last word. Finally, 15% of couples separate at some point during the sample period.

TABLE 2. Descriptive Statistics

	Married			Single			Total		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
A. Women									
Age	36.95	6.2	2601	35.44	7.3	731	36.62	6.5	3332
Height	166.76	6.3	2601	167.45	6.5	731	166.91	6.3	3332
BMI	24.09	4.1	2601	23.44	4.3	731	23.95	4.2	3332
Health	2.08	0.8	2601	2.19	0.9	731	2.11	0.8	3332
Educ.	12.46	2.6	2601	12.53	2.5	731	12.47	2.6	3332
Risk-taking	4.15	2.1	2601	4.71	2.2	731	4.28	2.2	3332
Openness	4.59	1.2	2601	4.81	1.2	731	4.64	1.2	3332
Conscientious	6.03	0.8	2601	5.97	0.8	731	6.01	0.8	3332
Extraversion	5.02	1.1	2601	5.11	1.1	731	5.04	1.1	3332
Agreeable	5.56	0.9	2601	5.55	0.9	731	5.56	0.9	3332
Neurotic	4.17	1.2	2601	4.11	1.3	731	4.16	1.2	3332
Gross HH income (log)	10.85	0.8	2579	9.53	1.5	699	10.57	1.1	3278
Children in HH	0.70	0.5	2601	0.41	0.5	731	0.64	0.5	3332
Sat. with life (wom)	7.11	1.7	2595	6.60	2.0	730	7.00	1.8	3325
Sat. with HH inc. (wom)	6.31	2.2	2583	5.13	2.5	724	6.06	2.4	3307
Pessimistic (wom)	2.04	0.7	2591	2.09	0.8	729	2.05	0.8	3320
Church attendance (wom)	0.52	0.5	2589	0.33	0.5	729	0.48	0.5	3318
B. Men									
Age	39.19	6.0	2601	36.63	7.0	642	38.68	6.3	3243
Height	179.52	7.0	2601	179.40	7.0	642	179.49	7.0	3243
BMI	26.22	3.5	2601	24.92	3.6	642	25.96	3.6	3243
Health	2.03	0.8	2601	2.06	0.8	642	2.04	0.8	3243
Educ.	12.58	2.8	2601	12.73	2.6	642	12.61	2.7	3243
Risk-taking	5.23	2.1	2601	5.44	2.1	642	5.27	2.1	3243
Openness	4.38	1.1	2601	4.60	1.1	642	4.42	1.1	3243
Conscientious	5.95	0.9	2601	5.85	0.9	642	5.93	0.9	3243
Extraversion	4.77	1.1	2601	4.85	1.1	642	4.79	1.1	3243
Agreeable	5.26	1.0	2601	5.35	0.9	642	5.28	1.0	3243
Neurotic	3.64	1.2	2601	3.66	1.2	642	3.64	1.2	3243
Gross HH income (log)	10.85	0.8	2579	9.97	1.5	614	10.68	1.0	3193
Children in HH	0.70	0.5	2601	0.03	0.2	642	0.57	0.5	3243
Sat. with life (men)	7.00	1.8	2595	6.56	1.9	642	6.91	1.8	3237
Sat. with HH inc. (men)	6.19	2.2	2585	5.45	2.4	634	6.04	2.3	3219
Pessimistic (men)	2.08	0.8	2591	2.12	0.8	641	2.09	0.8	3232
Church attendance (men)	0.46	0.5	2593	0.29	0.5	640	0.42	0.5	3233
C. Couple variables									
Separation	0.15	0.4	2601	.	.	0	0.15	0.4	2601
Last word: man	0.05	0.2	2572	.	.	0	0.05	0.2	2572
Last word: wom	0.03	0.2	2572	.	.	0	0.03	0.2	2572
Last word: both	0.82	0.4	2572	.	.	0	0.82	0.4	2572
Last word: disagree	0.10	0.3	2572	.	.	0	0.10	0.3	2572
Total life satisfaction	14.07	3.0	2601	.	.	0	14.07	3.0	2601
|Life satisfaction diff.|	1.17	1.3	2589	.	.	0	1.17	1.3	2589

Descriptive statistics for married and single women and men in GSOEP wave 2005 (2004 values of height, weight, BMI and risk-taking)

TABLE 3. Specifications and selection criteria (log-likelihood values)

Spec #	Spec. ϕ_{kl}	Spec. ϕ_k^p, ϕ_l^p	$\phi_{kl}(\cdot, \cdot)$	$\phi_k^p(\cdot), \phi_l^p(\cdot)$	$\log(L)$
1	1	1	$x \times y$	x, y	-5.4855
2	1	2	$x \times y$	x, x^2, y, y^2	-5.4031
3	1	3	$x \times y$	x, x^2, x^3, y, y^2, y^3	-5.4009
4	2	1	$|x - y|$	x, y	-5.3952
5	2	2	$|x - y|$	x, x^2, y, y^2	-5.3877
6	2	3	$|x - y|$	x, x^2, x^3, y, y^2, y^3	-5.3841

The value of the log-likelihood is normalized to -1 for specification 1.

APPENDIX B. PARAMETRIC SPECIFICATIONS

Recall that the surplus is parameterized as in equation 3.1. This requires us to specify the underlying basis functions. We propose to choose among a total of 6 specifications the one that best fits the data. These 6 specifications are obtained by considering two specifications for basis functions capturing the interaction between spouses' characteristics, ϕ_{kl}, and three specifications for the basis functions capturing the main effects of husbands' characteristics, ϕ_k^p and wives' characteristics ϕ_l^p.

The first specification for the basis functions ϕ_{kl} assumes that the interaction between men's and women's characteristics is the product of these characteristics, namely

$$\phi_{kl}(x^{(k)}, y^{(l)}) = x^{(k)} y^{(l)} \qquad \forall k, l \in K \times L$$

in which case, characteristics are complement whenever $\lambda_{kl} > 0$ or substitutes whenever $\lambda_{kl} < 0$.

The second specification assumes that the interaction between men's and women's characteristics is the absolute value of the difference of these characteristics, that is

$$\phi_{kl}(x^{(k)}, y^{(l)}) = |x^{(k)} - y^{(l)}| \qquad \forall k, l \in K \times L$$

in which case λ_{kl} indicates the importance of "similarities" between the k-th attribute of husbands and the l-th attribute of wives. If λ_{kl} is positive (negative), spouses with (dis-)similar attributes attract (repel) each other.

The basis functions ϕ_k^p and ϕ_l^p, for $p \in \{1, ..., P\}$, are assumed to be linear combinations of men's and women's characteristics. We propose to use power transformations of these characteristics,

$$\phi_k^p(x^{(k)}) = (x^{(k)})^p \qquad \text{and} \qquad \phi_l^p(y^{(l)}) = (y^{(l)})^p$$

for all $k \in \{1, ..., K\}$ and $l \in \{1, ..., L\}$, and consider three specifications, $P = 1$, $P = 2$ and $P = 3$.

Table 3 provides a summary of these specifications. We are agnostic about which model is best to bring to the data. We make use of the marriage patterns predicted by each specification to compute a likelihood which is reported in the last column of Table 3. Obviously, increasing P improves the fit of the model but since some coefficients λ_k^3 and λ_l^3 are significantly different from 0 we set $P = 3$. The most drastic improvement in the fit of the model however is obtained by choosing the absolute value of the difference in characteristics rather than the product to describe the interaction between men's and women's traits. Therefore, our preferred specification for the remainder of the paper will be specification 6.

Appendix C. Estimates

The estimates of preference parameters capturing the interaction between spouses' characteristics $(\lambda_{kl})_{kl}$ are reported in table 4. Three familiar patterns appear at first glance. First, all the diagonal entries of the matrix are negative, suggesting that like attracts like. Second, consistent with previous studies, the coefficients of age (-4.23)

TABLE 4. Estimate of the affinity matrix

	Age	Height	BMI	Health	Educ.	Risk	Open.	Consc.	Extrav.	Agree.	Neuro.
Age	-4.23	0.06	0.05	0.13	0.04	0.08	0.04	0.10	-0.02	-0.06	0.12
Height	-0.17	-0.68	0.08	-0.06	-0.27	-0.04	-0.06	0.05	-0.19	0.03	-0.01
BMI	-0.03	-0.17	-0.62	-0.16	0.02	-0.15	0.23	-0.14	-0.03	-0.01	-0.05
Health	0.11	0.01	-0.06	-0.56	-0.16	-0.28	0.05	-0.03	0.07	-0.04	-0.04
Educ.	-0.23	-0.15	0.38	0.07	-1.71	-0.16	-0.04	0.01	0.21	0.18	0.12
Risk	-0.09	-0.04	0.24	0.16	-0.08	-0.80	0.03	0.09	0.13	-0.02	-0.21
Open.	-0.15	0.07	0.10	-0.03	-0.14	-0.09	-0.79	-0.03	0.22	-0.00	-0.14
Consc.	-0.09	-0.01	-0.00	0.13	0.32	0.05	-0.10	-0.63	-0.06	-0.13	0.18
Extrav.	0.02	-0.06	0.09	-0.02	-0.00	0.09	0.16	-0.17	-0.28	-0.11	0.03
Agree.	0.12	0.02	-0.05	-0.02	-0.07	0.00	-0.10	-0.10	-0.09	-0.48	-0.14
Neuro.	0.12	0.08	-0.03	-0.05	-0.02	-0.18	-0.01	-0.12	0.14	-0.07	-0.32

Estimates for the interaction between X and Y, for specification 2 of $\phi_{kl}(\cdot, \cdot)$ and specification 3 for $(\phi_k^p(\cdot), \phi_l^p(\cdot))$

and education (-1.71) stand out: with the coefficient of age being 2.5 times larger than that of education, which itself is twice as large as the third largest coefficient $(-0.8$ for the interaction of risk attitude). Third, for all eleven characteristics, the coefficient on the diagonal dominates off-diagonal entries (in magnitude).[9]

---

[9]All variables are normalized, therefore the entries in the matrix are comparable with each other.

Reference

[1] Christopher J. Boyce, Alex M. Wood, and Eamonn Ferguson. "For better or for worse: The moderating effects of personality on the marriage-life satisfaction link". In: *Personality and Individual Differences* 97 (2016), pp. 61 –66. ISSN: 0191-8869.

[2] Marco Caliendo, Steffen Künn, and Martin Weißenberger. "Personality traits and the evaluation of start-up subsidies". In: *European Economic Review* 86.C (2016), pp. 87–108.

[3] Brent Donnellan and Richard Lucas. "Age Differences in the Big Five Across the Life Span: Evidence from Two National Samples". In: *Psychology and aging* 23(3) (2008), pp. 558–566.