HCEO WORKING PAPER SERIES

Assortative Matching on Income

Pierre-André Chiappori
Carlo Fiorio
Alfred Galichon
Stefano Verzillo

Working Paper

2022-006

03/2022

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Assortative Matching on Income*

Pierre-André Chiappori†, Carlo Fiorio‡, Alfred Galichon§ and Stefano Verzillo¶

Abstract

We analyze marital matching on income using an extremely rich Dutch data set containing all income tax files over four years. We develop a novel methodology that directly extends previous contributions to allow for highly flexible matching patterns. Investigating all marriages that took place between 2011 and 2014, we find that marital patterns remain remarkably stable over the period. While a majority of couples match assortatively, a small but significant minority display negative assortative matching. We also show that standard approaches, which consider all married couples using current incomes, may generate misleading conclusions. Finally, we find that, in contrast with recent results, whether his income exceeds her does not seem to play any significant role.

*We are particularly grateful for comments and suggestions to J.J. Heckman and F. Bourguignon. We also thank Matteo Bodini for technical assistance. We would also like to thank seminar and conference participants at the Cowles Foundation Conference on Matching: Optimal Transport and Beyond (2019) and at the Toulouse School of Economics Conference on Econometrics of Games, Matching, Networks (2021). The information and views set out in this paper are those of the authors and do not reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein. This paper is part of the European Commission, Joint Research Centre's Research Project #8303 which uses non-public microdata from Statistics Netherlands to study Segregation in the Netherlands. Alfred Galichon acknowledges support from European Research Council Grant ERC-CoG N. 866274

†Department of Economics, Columbia University, New York, USA. email: pc2167@columbia.edu

‡Department of Economics, Management and Quantitative Methods, University of Milan, Milano, Italy. email: carlo.fiorio@unimi.it

§Economics and Mathematics Departments, New York University, and Economics Department, Sciences Po, Paris. email: alfred.galichon@nyu.edu

¶European Commission, Joint Research Centre (JRC), Ispra, Italy. email: stefano.verzillo@ec.europa.eu

1 Introduction

The issue of assortative matching on the marriage market has attracted considerable attention in recent years. From a static perspective, such matching patterns have a direct impact on inequality across households. For given levels of income inequality among men and among women, the level of inequality *between couples* may significantly increase when agents match assortatively rather than randomly. Even more important is the long-term impact of matching patterns. Educated, high-income couples tend to invest more (and more efficiently) into their children's human capital, in terms of money and time. As a result, inequality between individuals coming from different social backgrounds may rise, creating the risk of an "inequality spiral".[1] Measuring the degree of assortativeness on a given marriage market is however challenging, as demonstrated by recent discussions.[2] While the econometrics of matching models has recently experienced significant advances,[3] several questions remain open; moreover, the availability of relevant data has remained somewhat problematic.

1.1 Measuring assortativeness: theoretical issues

Many empirical works, following the initial contribution of Choo and Siow (2006b), adopt the so-called Separable Extreme Value (SEV) approach, based on a model of frictionless matching under Transferable Utility (TU). In such a framework, individuals belong to a (small) number of (large) categories. Any pair of potential mates would, if matched, generate a pair-specific surplus, which would then be shared between spouses; this surplus is the sum of a deterministic component depending only on the spouses' categories, and two random shocks describing each spouse's idiosyncratic preferences. It can be shown that, in a one-dimensional context, matching patterns generated by such a model display positive assortative matching (PAM) whenever the deterministic

[1] See for instance Aiyagari et al. (2005), Fernandez et al. (2005), Chiappori et al. (2017), Chiappori (2017).

[2] See for instance Chiappori et al. (2020).

[3] See for instance the surveys by Chiappori and Salanié (2016) and Galichon and Salanié (2017).

part of the surplus is supermodular, and negative assortative matching (NAM) under submodularity.

The SEV approach has been applied to a host of specific contexts, including matching on age (Choo and Siow, 2006a), education (Chiappori et al., 2017), ability (Chiappori et al., 2020), as well as race (Schwartz and Graf, 2009) or psychological profiles (Dupuy and Galichon, 2014). Yet, the notion that individuals belong to a small number of homogeneous categories may in some cases be unduly restrictive. It is convenient when considering matching patterns by race, religion or possibly education. Quite often, however, the traits under consideration are intrinsically continuous; this is the case, for instance, with income, age or human capital. While a continuous variable can always be discretized, such discretization implicitly requires that people within a category are perfect substitutes, whereas they may markedly differ across categories - an assumption all the more debatable when the number of categories is small and discontinuities at thresholds may be large. To address this problem, some recent contributions (for instance Dupuy and Galichon 2014, Bojilov and Galichon 2016, Chiappori et al. 2017 and others) have extended the SEV model to continuous variables. While the various solutions differ in several aspects, they share a common feature, namely that the (deterministic) surplus is a quadratic function of individual characteristics. The corresponding, "affinity" matrix can then be identified from matching patterns.

While convenient and tractable, however, a quadratic surplus comes at a cost: it strongly restricts the way assortativeness varies with income. This is because incentives to match assortatively depend on the second cross derivative of the surplus function, which is constant in a quadratic formulation. In other words, choosing a quadratic surplus amounts to assuming that preferences for homogamy are exactly identical at the top or at the bottom of the income distribution (and anywhere in-between). Yet, several theoretical arguments, as well as numerous empirical works based on discretized models, strongly suggest that preferences for homogamy need not be constant over the income distribution - a fact that has strong implications, in particular in terms of current and future inequality. In principle, matching could even positive assortative over a portion of the joint distribution of incomes, but negative assortative elsewhere.

Finally, an interesting issue is the potential asymmetry between male and female income, as several contributions have emphasized. For instance, Bertrand et al. (2015), analyzing a panel of married couples, find that the ratio of female over total income exhibits a discontinuity at 50%: there are very few couples in which her self-reported income (slightly) exceeds his. The interpretation they suggest is that prevailing social norms tend to penalize such situations. While this issue will be discussed in some detail later on, it can immediately be remarked that imposing a quadratic form for the surplus rules out such phenomena a priori, which is unacceptably restrictive.

1.2 Measuring assortativeness: which data?

As any researcher involved in the empirical analysis of marital patterns knows for a fact, good data is extremely difficult to obtain. Standard approaches typically consider a population of married couples and analyze the corresponding matching patterns using the spouse's *current* incomes. Convenient as it may be, this approach raises serious endogeneity problems. Incomes depend, among other things, on the spouses' respective labor supply decisions, which themselves are strongly influenced by marital status; as such, current earnings are as much an outcome of the marital relationship as a determinant of marriage patterns. Ideally, therefore, the analysis of matching patterns by income should use income *at* (or, even better, *just before*) marriage, rather than income observed years or even decades later. In most cases, however, such a requirement is impossible to meet, if only because of sample size constraints. Most data are cross-sectional, meaning that individual incomes before marriage is not recorded. Even for panel data (such as the PSID), pre-marital incomes are typically observed for one spouse only. And in any case, the number of marriages in the panel taking place during a given year is typically too small to allow for robust estimations. Another problem relates to the way incomes are recorded. In survey data, income is typically self-reported. This standard problem may be particularly serious when considering matching patterns, because any social norm affecting marital roles may well influence (and possibly bias) self-reporting.

Disentangling the exact gender patterns of the matching process from either labor supply responses or reporting bias is both very important and extremely difficult to do with most data. The ideal data set for studying matching on income would have the following features. It should have a panel structure; moreover, it should allow the observation of pre-marital incomes for *both* spouses. The sample should be large, so that the number of marriages in any given year enables a robust estimation of marital patterns that year. Finally, income should ideally be third–party reported to minimize reporting bias (e.g. aimed at reducing the due tax). The data base we use has all these properties, as discussed below.

1.3 This paper

A first goal of the present paper is to propose a general methodology for analyzing matching patterns based on continuous variables. We first consider a non parametric approach that extends the SEV framework to a finite but "large" number of categories, thus reducing the discontinuity issues arising from discretization. In practice, we analyze matching patterns by income, and we consider 30 income brackets for both men and women, thus generating a 30×30 matching matrix. The resulting patterns are fully summarized by a piecewise representation of the surplus function - i.e., a set of 900 coefficients, each computed using a direct extension of the seminal approach of Choo and Siow (2006b) that can readily be plotted. As we shall see, despite our very large sample size, these data are way too noisy to allow for a robust identification of matching patterns.

Next, we consider a semi-parametric model in which the surplus is estimated using a flexible functional form; in practice, we use a high degree polynomial in individual incomes. We show how this model can be identified from available data. These estimates, in turn, allow to directly compute local supermodularity as the second cross derivative of the surplus function. In particular, one can observe which income ranges exhibit significant preferences for positive or negative assortative matching by computing the sign of this secod cross derivative; moreover, an extension of existing methods allows to assess the statistical significance of these estimates. All in all, we provide a methodology that enables a flexible, yet robust identification of (positive or negative) assortativeness on large data set when the underlying trait is continuous.

We then apply this methodology to an exceptional data set, consisting of all residents in the Netherlands for which we have info on their taxable income over four years (2011 to 2014). For each year, we observe about 17 million individuals with their incomes, including those of about 140,000 individuals that married during that year. In our analysis, we concentrate on wage and salary incomes; these are third-party reported, implying that reporting biases are minimal. Finally, we observe both spouses' pre- and post-marital incomes for most newly wed couples.

We first apply the non-parametric strategy described above to all marriages that took place over the 4-year period, using income at the year of marriage as the matching trait. As expected, while the resulting surplus function is reasonably smooth (due to the large number of categories), direct estimates of supermodularity from second differences are extremely noisy. On the other hand, the semi-parametric approach leads to an estimated supermodularity function (defined as the second cross derivative of the surplus function) that is a polynomial of degree 4 in male and female income. Interestingly, while the polynomial coefficients significantly vary over time, the general patterns described by the second cross derivative remain remarkably stable. Specifically, strong preferences for homogamy appear to dominate for middle-income couples, less so for richer households. However, a small percentage of newly wed couples exhibits both large asymmetries between male and female income *and* strong preferences for negative assortative matching. Not only are these general patterns observed for each of the 4 years under consideration, but the exact location of the PAM and NAM areas are remarkably similar. Finally, we check the power of our approach by running placebo regressions on a virtual sample in which the same individuals are randomly matched into virtual couples. Our estimates do not reject the null of a zero cross derivative (equivalent to random matching) at each point.

We then proceed to a series of additional estimations. First, we address the endogeneity issue of labor income by estimating matching patterns based on incomes in a different year. Indeed, a possible concern is that incomes observed at the year of marriage could partly reflect labor supply adjustments due to the change in marital status. We therefore estimate the model for couples married in 2014 using 2011 individual incomes. Interestingly, the qualitative patterns remain essentially similar. Conversely, we analyze the impact on matching patterns of possible changes in labor income generated by marriage by estimating the model on couples wed in 2011, using individual incomes reported in 2014. While most conclusions - and in particular the presence of strong preferences for homogamy among middle-income couples - remain valid, some changes can be stressed; in particular, the "traditional couples" region, where his income largely exceeds her, exhibits substantially less tendency towards negative sorting. In the same spirit, we estimate the same model on the total population, using current incomes as matching traits. Again, areas of negative assortative matching appear to be significantly different, suggesting either cohort effects or labor supply responses to marital status.

Yet another concern is related to the use of current (as opposed to permanent) income as a matching traits. While the issue cannot be fully addressed in our context, we investigate the consequences of replacing current income with the individual's average income over the 4 years in our estimation process - so as to partially smooth out transitory income shocks. Again, results remain unchanged, suggesting that transitory shocks may not affect our conclusions.

Some models of frictionless matching under TU - particularly those where marital gains stem from the joint consumption of public goods - predict that marital gains should essentially depend on total household income. Such a framework is nested within our general formulation. We estimate the model under this restriction, and find that it is strongly counterfactual. Not only do usual test unambiguously reject the restriction, but the patterns described by this one-dimensional model markedly differ from the picture depicted by the general version.

Finally, we investigate the asymmetry between male and female incomes, and more specifically the existence of a discontinuity around equal incomes. For that purpose, we add, as a regressor, a dummy variable equal to 1 when her income exceeds his. While the coefficient is significant, its magnitude is extremely small, and the matching patterns remain unchanged. A natural interpretation is that, far from reflecting social norms affecting marital patterns, the corresponding variable simply captures some non linearities that even our highly flexible form cannot fully take into account. To test this hypothesis, we run a series of placebo estimations in which the dummy variable equals 1 when the fraction of her income to his income exceeds some threshold value different from 1 (in practice, from 80% to 120%). In each case, the coefficient is significant, typically larger than for the threshold 1 (although still negligible in absolute value), and estimated marital patterns remain unchanged. We conclude that the "50% threshold" has no specific impact on marital patterns.

The paper is organized as follows. Section 2 details the theoretical background of our empirical approach. Section 3 provides a description of the data and presents the main conclusions of the non-parametric version. The semi-parametric approach and the corresponding estimation strategy are discussed in Section 4, while Section 5 presents the results. Section 6 concludes.

2 Theoretical background

2.1 Measuring homogamy: a simple example

A first task is to define an index measuring the degree of homogamy (or positive assortativeness) of a given marital distribution. For that purpose, it is useful, following Chiappori et al. (2020), to start with the simplest possible case. Let us thus consider a population in which men and women belong to either of two categories - say, high (H) versus low

(L) income. Ignoring singles and normalizing total population size to 1, matching patterns are fully summarized by a Table of the following type:

$T =$ (1)

$_w\backslash^m$	H	L
H	a	b
L	c	d

with $a+b+c+d=1$; here, a is the proportion of couples in which both spouses have a high income, and so on. As a benchmark, it is natural to consider the structure that would obtain, with the same marginal distributions (i.e. the same total proportion of high income men and high income women), under random matching. The matrix would then be:

$T_R =$ (2)

$_w\backslash^m$	H	L
H	$(a+b)(a+c)$	$(a+b)(b+d)$
L	$(a+c)(c+d)$	$(b+d)(c+d)$

In this simple framework, defining Positive Assortative Matching (PAM) is straightforward: Table T exhibits PAM if and only if it has more couples on the diagonal (and less off diagonal) than under random matching. Formally, PAM obtains if and only if:

$$a \geq (a+b)(a+c) \text{ or equivalently } d \geq (b+d)(c+d)$$

In the opposite case, the Table exhibits Negative Assortative Matching (NAM).

Quantifying the degree of PAM is a more difficult question. A standard assortativeness index, that has been used across various fields,[4] is the following:

$$I_A = \ln\left(\frac{ad}{bc}\right) = (\ln a - \ln b) - (\ln c - \ln d)$$

Various properties of this index can be mentioned at this point. First,

[4]This measure is exactly the SEV index introduced by Choo and Siow (2006b,a). It is also known in the sociological and demographic literature as the "log linear model". See Chiappori et al. 2020 for a detailed presentation.

it is positive (resp. negative) if and only if the Table exhibits PAM (NAM). Second, it is maximum (and equal to $+\infty$) if the Tables displays Perfect Assortative Matching (in the terminology of Chiappori et al. 2020), i.e. if

$$\min(b, c) = 0$$

A third and crucial property is its structural interpretation in terms of frictionless matching model under TU, which is described next.

2.2 The SEV model

The SEV (for Separable Extreme Value) approach, initially introduced by Choo and Siow (2006b), postulates that observed patterns are derived as the stable matching of a frictionless matching model under TU. Specifically, assume that the match between a woman x with income I and a man y with income J (where $I, J \in \{H, L\}$) generates a surplus of the form:

$$s_{xy} = S_{IJ} + \alpha_x^J + \beta_y^I \tag{3}$$

where S_{IJ} is a deterministic component that only depends on individual incomes, α_x^J is a random shock that describes x's idiosyncratic preferences for a husband with income J and β_y^I is a random shock that describes y's idiosyncratic preferences for a wife with income I.[5] Finally, it is assumed that all random shocks follow independent, Type 1 extreme value distributions.

There exists a close relationship between the SEV approach and the assortativeness index just defined. Indeed, an important result (Chiappori et al., 2017) is that, if one compares two different structural matrices S and S', the distribution generated by S displays more PAM than that generated by S' (in the sense that the index I_A is larger) if and only if S is more supermodular than S', i.e. if and only if:

$$S_{HH} + S_{LL} - (S_{HL} + S_{LH}) \geq S'_{HH} + S'_{LL} - (S'_{HL} + S'_{LH})$$

[5]Note that the random shocks only affect the total surplus. In particular, an alternative but ultimately equivalent interpretation is that α_x^J represents how unobserved characteristics of Mrs x are valued by a male with income J.

One can readily check that random matching corresponds to the linear case in which the structural matrix S is neither super- nor submodular:

$$S_{HH} + S_{LL} - (S_{HL} + S_{LH}) = 0$$

In words, this situation obtains when the contribution of an individual's income to the total surplus does not depend on the spouse's income; that is, there exists numbers A_H, A_L, B_H, B_L such that

$$S_{IJ} = A_I + B_J \text{ for all } I, J \in \{H, L\}$$

It follows from the previous results that, for a large enough matching game generated by a SEV model, the index I_A is positive, implying PAM, if and only if the structural matrix $S = (S_{IJ}, I, J \in \{H, L\})$ is supermodular, that is if:

$$S_{HH} + S_{LL} - (S_{HL} + S_{LH}) \geq 0$$

Finally, this structure can readily be extended to an arbitrary number N of categories and to include singlehood. Then each woman x is characterized by her income category and a vector of $(N+1)$ random shocks $\left(\alpha_x^1, ..., \alpha_x^N, \alpha_x^\emptyset\right)$, where α_x^J denotes x's idiosyncratic preference for a husband in category J and $\alpha_x^\emptyset$ her idiosyncratic preference for singlehood (a similar property applies to each man y). The surplus generated by the match of a woman x whose income belongs to category I and a man y with income in category J takes the form:

$$s_{xy} = S_{IJ} + \alpha_x^J + \beta_y^I$$

where again S_{IJ} is the deterministic component of the surplus and α_x^J and β_y^I are random variables following a Type 1 extreme value distribution. It can then be shown (see for instance Choo and Siow 2006a) that:

$$S_{IJ} = 2 \ln \left(\frac{\mu_{IJ}}{\sqrt{\mu_{I0}}\sqrt{\mu_{0J}}} \right) = 2 \ln \mu_{IJ} - \ln \mu_{I0} - \ln \mu_{0J} \qquad (4)$$

where μ_{IJ} is the number of couples such that her income belongs to category I and his to category J, and similarly μ_{I0} (resp. μ_{0J}) is the number of single women (men) with income in I (J).

An important advantage of the SEV interpretation is that it clarifies the interpretation of observed changes. Comparing matching patterns between two populations (or two periods for the same population) raises a standard issue, namely distinguishing between the mechanical impact of changes in the marginal distributions of matching traits (income in our case), on the one hand, and possible structural changes in "preferences for homogamy" on the other hand. In a SEV context, the latter are fully described by variations in the structural matrix S, and more precisely in its supermodularity; these can readily be recovered from the data, as explained below.

Finally, the properties described above lie at the core of the empirical analysis that follows. In practice, we shall first adopt a non parametric approach by computing the structural index S_{IJ} for each pair (I, J) of income categories. Then, from a semi-parametric perspective, we estimate S as a flexible function of individual incomes. Note, finally, that for such a smooth (actually infinitely differentiable) representation, supermodularity can directly be computed as the second cross derivative of S with respect to individual incomes.

3 Data: description and non parametric analysis

3.1 The Data

Our analysis aims at overcoming the data limitations discussed in previous sections by taking advantage of population administrative data for a whole country in Northwestern Europe, i.e. the Netherlands. Data are provided by Statistics Netherlands, namely the Dutch *Centraal Bureau voor de Statistiek* (CBS). The Dutch administrative data provide a panel structure at the individual level of a large set of variables using the municipal population register (GBA) over several years for the whole population living in the country. We distinguish between couples and singles as CBS registries report the individual position in the household (categorized as single, partner in a married or unmarried couple, child or other member) and disentangle newlyweds from old couples in a year when two people change their position from single to married with the same household identifier on the exact same day. We focus on legally married couples considering two unmarried spouses as singles consistently with the literature. Using the country of origin information, we define as old couples those for which one of the two spouses is from abroad and appears for the first time in the GBA registry at the very same day the other changes status from single to married. This because we do not have enough information to rule out the possibility that marriage had happened before in another country. We also checked if newlyweds were not recorded previously as a couple in past years, as individual position in the household could change if one of the spouses moves abroad (hence outside of the GBA) and comes back later on. If this happened, the newlywed is classified as an old couple. The total number of new marriages identified this way is consistent with the official CBS marriages statistics that is computed using the marriages registry: there remain a difference, which on average is lower than 0.8%. The number of newlywed couples in the Netherlands in any given year is around 70,000. We link information on personal incomes, collected by the Dutch Tax and Customs Administration without trimming or censoring, to information on household composition and individual demographics (such as age, gender and country of origin).

Administrative tax data are ideal data for minimizing measurement errors often found in survey data. Moreover, detailed information on individual and household income allows us to focus on third-party reporting income (i.e. wages and salaries), avoiding possible biases due to self-reporting and tax evasion that might be substantial also in northern Europe.[6] We select data for four years, starting from 2011 that is the first year where full detail of individual incomes (namely its components) became available.[7]

Finally, from the full population we selected individuals within the

[6] See Kleven et al. (2011) for an analysis on Danish data.

[7] Results using data on a longer period are qualitatively similar to the ones reported in this paper. A replication package can be provided by the authors upon request.

Age	Sample	2011	2012	2013	2014
Any	All	16,960,774	17,022,404	17,075,035	17,145,159
Any	In couple	6,628,006	6,598,712	6,560,912	6,526,930
Any	Newlywed	144,124	140,120	129,528	132,368
Both 20-59	All	9,172,184	9,162,667	9,151,123	9,150,432
Both 20-59	In couple	4,042,306	3,956,468	3,866,242	3,782,366
Both 20-59	Newlywed	135,146	131,260	120,840	125,986

Notes: Individual observations. Couples are legally married couples only.

Table 1: Total number of residents by year and marital status

age range 20-59 only, as the group where most marriages occur, keeping all singles and all couples with both spouses within this age range. Table 1 shows that, in 2014 out of a total of 132,368 individuals in 66,184 newlywed couples, about 95% are in this age group. By focusing on 2014 only and splitting the population by gender, Table 2 shows that, on average, women get married at the age of 32.5 and men at 35 (panel a). Newlywed women earn an average (wage and salary) income that is 11% larger than that of the average woman in this age group (€25,070 vs €22,590) while the newlywed men earn a slightly larger (3%) wage (around €40,000) if compared with the average man (€38,870) in the same age group (panel b).

(a) Age

	Women			Men		
	Obs.	Mean	St.Dev.	Obs.	Mean	St.Dev.
All residents	4,545,795	40.18	11.52	4,604,637	40.12	11.53
In couple	1,891,183	43.42	8.86	1,891,183	45.85	8.65
Newlywed	62,993	32.45	8.30	62,993	35.18	8.72

(b) Employment income (wages and salaries), in thousand euro

	Women			Men		
	Obs.	Mean	St.Dev.	Obs.	Mean	St.Dev.
All residents	2,853,994	22.59	21.10	3,023,169	38.87	47.69
In couple	1,151,314	23.54	20.68	1,259,509	51.12	58.69
Newlywed	46,213	25.07	18.54	46,883	40.05	41.40

Notes: Individual observations. Couples are legally married couples only. Employment (self-employment) income statistics are presented only for those with positive employment income.

Table 2: Some descriptive statistics by gender for individuals aged 20-59, year 2014

3.2 Matching Matrices

To analyze the matching pattern of our population, we group the populations of couples in 30 gender-specific income groups as follows. We compute the 50^{th} and the 99^{th} percentiles of the 2011 gender-specific log-income distribution (respectively, $d_{x,50}$ and $d_{x,99}$ for females and $d_{y,50}$ and $d_{y,99}$ for males). Then, for each year considered, we group all women with income lower than $d_{x,99}$ in 25 groups, using $\{d_{x,50}, d_{x,99}\}$ and a step defined as $(d_{x,99} - d_{x,50})/25$, where all women with income below $d_{x,50}$ are assigned to group 1. The group of women with log-income larger than $d_{x,99}$ is partitioned further using the percentiles 99.2, 99.4, 99.6, 99.8 and 99.95, assigning to group 30 all women with income larger than $d_{x,99.95}$. Similarly are defined the income groups of male individuals.[8]

Table 3 shows the discretization of the gender-specific income spaces reporting mean wages and number of individuals by gender for all residents in the Netherlands in some selected income groups in 2011. Mean

[8]Results have been tested for robustness replacing $d_{x,50}$ with the 25^{th}, the 45^{th} and the 55^{th} gender-specific percentiles of log-income, as well as using different quantiles to partition the top 1% group and results are widely consistent. See the Appendix 7 for further details.

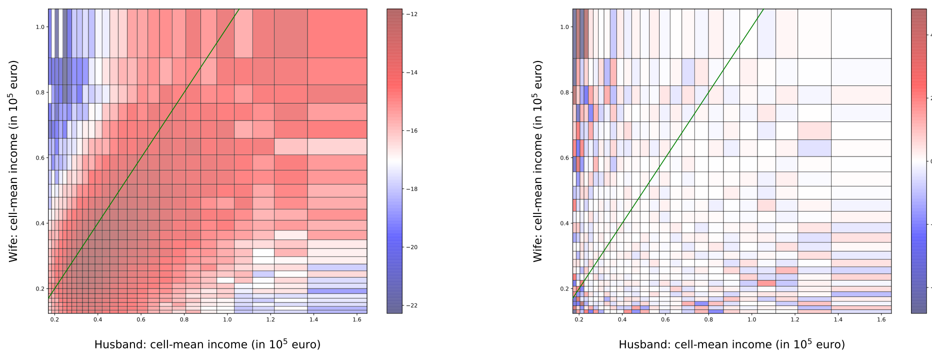
	Women					Men				
Group	Mean	St.Dev.	Frequency of couples			Mean	St.Dev.	Frequency of couples		
			New	Old	All			New	Old	All
2	12,419	286	676	24,503	33,522	17,086	408	233	4,444	32,953
...										
5	15,771	361	990	32,748	43,047	21,831	522	706	9,247	55,626
...										
10	23,464	538	2,194	42,431	66,174	32,868	777	2,791	42,274	93,098
...										
15	34,916	801	2,241	26,773	61,875	49,403	1,172	1,876	41,580	54,521
...										
20	51,925	1,185	690	9,383	22,632	74,383	1,766	570	18,176	18,285
...										
25	76,641	1,370	116	2,088	4,290	111,660	2,387	124	5,204	4,813
26	82,225	1,903	104	2,269	4,273	121,784	3,520	170	5,469	4,895
27	90,502	3,019	118	2,392	4,553	137,034	5,455	139	5,381	4,813
28	105,438	6,088	91	2,596	4,570	164,781	11,387	138	5,064	4,740
29	137,250	14,877	71	1,940	3,374	229,307	33,416	89	3,669	3,593
30	263,716	268,433	10*	670	1,082	594,455	571,235	29	953	1,207

Notes: Income groups are defined in Subsection 3.2. "New" stands for newlyweds, "Old" for old couples, "All" for all residents. All individuals considered are aged 20-59. Means and standard deviations are calculated using all residents and are measured in thousands of euro. Only a selection of groups is reported.*Exact frequency is below 10 but cannot be reported for confidentiality reasons.

Table 3: Descriptive statistics by selected income groups for the whole population, year 2011.

wages started from €12,419 for women and €17,086 for men growing to a maximum mean wage of €263,716 (€594,455) in the top bin. The log-income discretization of spouses' employment incomes obtained in 2011 is then used to generate the same income bins in all the considered years to ensure comparability. We then generate the final set of matching matrices of size 30×30 aggregating the data by each combination of the husband-wife pair of income groups. Each matching matrix is composed by 900 cells reporting couples' frequencies and average incomes for each husband-wife group combination as well as the marginals which will be used in our semi-parametric estimations.

Different binning structures (discretizing from the 25^{th}, 40^{th} or 45^{th} - instead of from the 50^{th}- percentile up to the 99^{th} percentile, and dividing the top 1% in either five or three groups) are generated for robustness estimations. Robustness results of our semi-parametric models are provided in the Appendix.



(a) Surplus (b) Second differences

Figure 1: Surplus and Second Cross Differences, all newly wed (in b), red: PAM, blue: NAM)

3.3 Non parametric analysis

For each of the 900 cells of our matching table, we compute the corresponding structural coefficient S_{IJ}. Figure 1a provides a projection of the corresponding values on the plane representing his and her income, with the convention that the corresponding value of the surplus increases from blue to red.

In particular, the surplus is larger for couples with similar incomes, particularly for low to medium income levels; it is much smaller for couples whose spouses have widely different incomes.

We are ultimately interested in matching patterns, therefore in the supermodularity of the surplus function. A first possibility is to approximate it by computing *local supermodular cores*, defined for $I, J \in \{2, 30\}$ as the second differences:

$$D_{I,J} = S_{I,J} + S_{I+1,J+1} - (S_{I,J+1} + S_{I+1,J})$$

The outcome is plotted in Figure 1b. Despite the large sample size, the graph is extremely noisy; no clear pattern emerges (except possibly a significant propensity to NAM when spouses' incomes are very different).

This finding, unfortunately, is not surprising. Despite the large population (about 70,000 marriages each year), each of the 900 cells contains on average only 77 couples, and a double difference on such small samples is doomed to give noisy results.

4 A semi-parametric model

We now present our parametric approach. As discussed above, the basic idea is to approximate the surplus function by a flexible function of spousal incomes. I_x and J_y denote respectively the income categories of a woman x and of a man y. In practice, thus, we assume that the surplus can be expanded under a basis of K^2 polynomial terms:

$$S^\lambda(I_x, J_y) = \sum_{0 \leq l+r \leq L} \lambda^{l,r} \Sigma^{l,r}(I_x, J_y), \text{ where } \Sigma^{l,r}(I_x, J_y) = (I_x)^l . (J_y)^r \tag{5}$$

and $\lambda = \left(\lambda^{0,0}, ..., \lambda^{L,0}, ..., \lambda^{0,L}\right)$ are parameters to be estimated. In that case, the supermodularity of the surplus directly translates into the sign of the second cross-derivative:

$$\frac{\partial^2 S^\lambda(I, J)}{\partial I \partial J} = \sum_{l+r \leq L} lr \lambda^{l,r} \Sigma^{l,r}(I, J),$$

where by convention $\Sigma^{l-1,r-1}(I, J) = 0$ if $\min(l, r) = 0$.

4.1 Estimation

We let $\mathcal{K} = \mathcal{I} \cup \mathcal{J}$ be the set of categories of incomes I and J of female and male individuals. For $K \in \mathcal{K}$, we define q_K as the mass of individuals of type k (either female or male) in the population. If $K \in \mathcal{I}$, then q_K is the mass of females x with income $I_x = K$; while if $K \in \mathcal{J}$, then q_K is the mass of males y with income $J_y = K$. Similarly for $K \in \mathcal{K}$, we define p_K the (endogenous) average utility payoff obtained by an individual of income category K at equilibrium.

A household, or marital arrangement $a \in \mathcal{A}$ is either a matched couple $IJ \in \mathcal{I} \times \mathcal{J}$, or a single woman of income category I whose household is denoted $I0$, or a single man J, whose household is denoted $0J$. Hence the set of marital arrangements $\mathcal{A}$ is $\mathcal{A} = (\mathcal{I} \times \mathcal{J}) \cup (\mathcal{I} \times \{0\}) \cup (\{0\} \times \mathcal{J})$. Given $a \in \mathcal{A}$, we denote w_a the number of individuals in household a: $w_{IJ} = 2$, $w_{I0} = 1$ and $w_{0J} = 1$.

We index by $\beta \in \mathcal{B}$ the vector λ, so that each β is associated with a pair of indices (l, r) with $0 \leq l + r \leq L$, and if β is associated with (l, r), we denote $S_a^\beta = \Sigma_a^{l,r}$ if $a \in \mathcal{I} \times \mathcal{J}$ is a married household, and $S_a^\beta = 0$ if a is a single household. S is the $\mathcal{A} \times \mathcal{B}$ matrix whose a, β cell is S_a^β. As a result, $S_a^\lambda = (S\lambda)_a$.

The observed number of households a is denoted $\hat{\mu}_a$. We represent $\hat{\mu}$ as the vectorization in the row-major order of the matrix $(\hat{\mu}_{IJ})$, followed by the vector $\hat{\mu}_{I0}$, followed by the vector $\hat{\mu}_{0J}$.

The number of marital arrangements $\hat{\mu}$ and the number of female and male individuals of each type are related by the relationship $q_K = \sum_{a \in \mathcal{A}} \hat{\mu}_a 1 \{K \in a\}$, which we denote

$$q = M\hat{\mu}$$

where M is the *margining-out-matrix* (MOM), which is constructed as

$$M = \begin{pmatrix} \mathbf{I}_{\mathcal{X}} \otimes \mathbf{1}_{\mathcal{Y}}^\top & \mathbf{I}_{\mathcal{X}} & \mathbf{0}_{\mathcal{X} \times \mathcal{Y}} \\ \mathbf{1}_{\mathcal{X}}^\top \otimes \mathbf{I}_{\mathcal{Y}} & \mathbf{0}_{\mathcal{Y} \times \mathcal{X}} & \mathbf{I}_{\mathcal{Y}} \end{pmatrix}.$$

It is known (Becker 1973, Choo and Siow 2006b, Galichon and Salanie 2021) that if the surplus formed from a household a is S_a^λ, the equilibrium matching is the solution to the following maximization problem

$$W(\lambda) = \max_{\mu \geq 0} \sum_{a \in \mathcal{A}} \mu_a S_a^\lambda + T \sum_{a \in \mathcal{A}} w_a \mu_a \log \mu_a$$
$$s.t. \; M\hat{\mu} = q$$

which has dual

$$W(\lambda) = \min_{p \in \mathbb{R}^{\mathcal{K}}} \sum_{K \in \mathcal{K}} p_K q_K + \sum_{a \in \mathcal{A}} w_a \exp\left(\frac{S_a^\lambda - (M^\top p)_a}{w_a}\right)$$

The primal solution μ_a and the dual solution p are related by $\mu_a = \mu_a^\theta$

where

$$\mu_a^\theta := \exp\left(\frac{S_a^\lambda - \left(M^\top p\right)_a}{w_a}\right)$$

and $\theta^\top = \left(\lambda^\top, p^\top\right)$, and we note that

$$\frac{\partial W(\lambda)}{\partial \lambda^\beta} = \sum_{I \in \mathcal{I}, J \in \mathcal{J}} \mu_{IJ}^{\lambda,p} \Sigma^\beta (I,J) = \left(S^\top \mu^\theta\right)_\beta$$

As a result, it follows (still from Galichon and Salanie 2021) that λ is estimated by $\min_\lambda \left\{W(\lambda) - \hat{\mu}^\top S \lambda\right\}$ that is, λ appears in the solution to

$$\min_{\lambda,p} \sum_{K \in \mathcal{K}} p_K q_K + \sum_{a \in \mathcal{A}} w_a \exp\left(\frac{S_a^\lambda - \left(M^\top p\right)_a}{w_a}\right) - \sum_{\substack{I \in \mathcal{I} \\ J \in \mathcal{J}}} \hat{\mu}_{IJ} \Sigma^{k,l} (I,J)$$

which rewrites in a more condensed form as

$$\min \hat{\mu}^\top \left(M^\top p - S\lambda\right) + \sum_{a \in \mathcal{A}} w_a \exp\left(\frac{S_a^\lambda - \left(M^\top p\right)_a}{w_a}\right)$$

We let $X = diag(w)^{-1} M^\top$, and we recall that $\theta^\top = \left(\lambda^\top, p^\top\right)$, and the problem rewrites as

$$\min_\theta \sum_{a \in \mathcal{A}} w_a \left\{\exp\left((X\theta)_a\right) - \hat{\mu}_a (X\theta)_a\right\}$$

which is a weighted Poisson regression, which we can estimate using standard packages on generalized linear models[9].

5 Results

We now present the main results of our semi-parametric approach.

[9] For instance, R's glmnet, or scikit-learn in Python.

5.1 Main parametric model

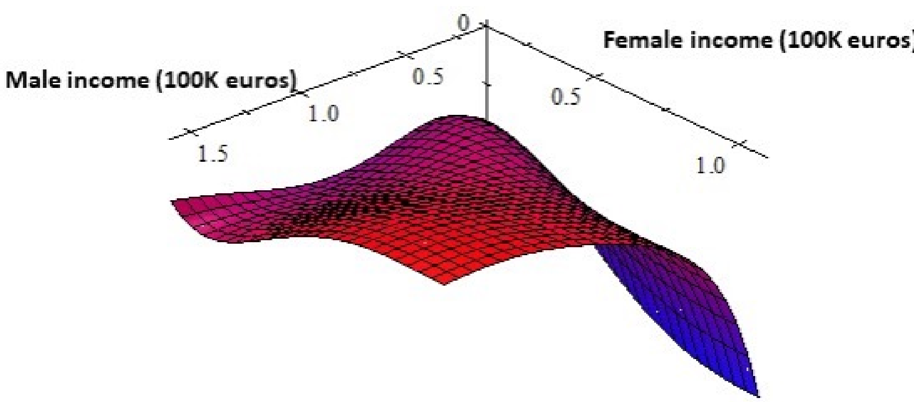
5.1.1 All newly wed

We first analyze the total sample of 256,616 marriages over the 4 years; in each case, we use incomes "at marriage" (i.e., at the year of marriage) as the matching characteristic. As discussed before, the surplus is estimated as a flexible, polynomial function of degree 6 in male and female income; as a robustness test, we shall later consider an extension to a polynomial of degree 7.

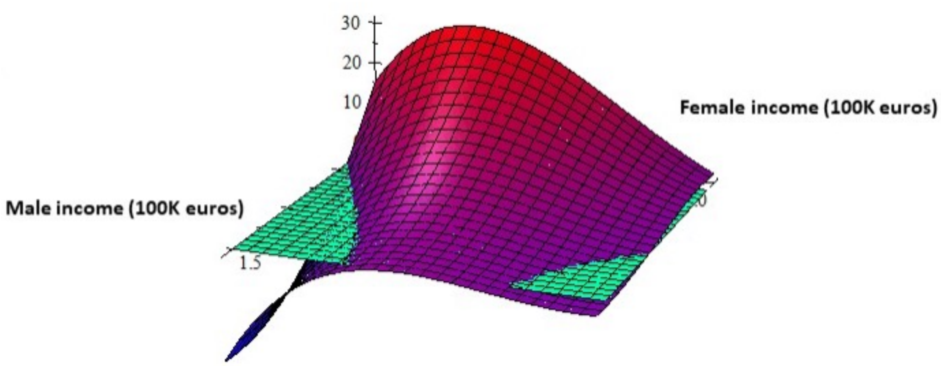
The 3D shape of the surplus function is plotted in Figure 2a for male incomes between 10,000 and 160,000 euros and female incomes between 10,000 and 100,000 euros per year. The shape is consistent with the non parametric estimates given earlier. In particular, the surplus increases with income, and is larger for couples with similar incomes.

More interesting is the graph of the second cross derivative, which indicates the super- (when positive) or sub- (when negative) modularity of the surplus function. This is Plotted 3D in Figure 2b, where the green plane corresponds to zero. Figure 3 gives a projection of the level curves of the previous graph. Here, a yellow line indicates a second cross derivative equal to zero, while the areas of significantly positive (resp. negative) second cross derivative are limited by red (blue) *solid* lines. In addition, the background color increases (from dark grey to white) with population density, thus allowing a first intuition of how populated the various regions are; and the green line indicates equal male and female incomes.

The marital patterns appear clearly from these graphs. For a large fraction of the population, there is a strong tendency to homogamy: the second cross derivative is significantly positive, and actually quite large. This is in particular the case for middle-income couples when spouses have similar incomes. As shown by Figure 3, the corresponding area, located in the south-west part of the graph, includes the vast majority of the newly wed population. Yet, a small but significant fraction of couples displays the exact opposite feature, namely a significantly negative second cross derivative, corresponding to negative assortative matching.



(a) Surplus



(b) Supermodular Core

Figure 2: All newly married

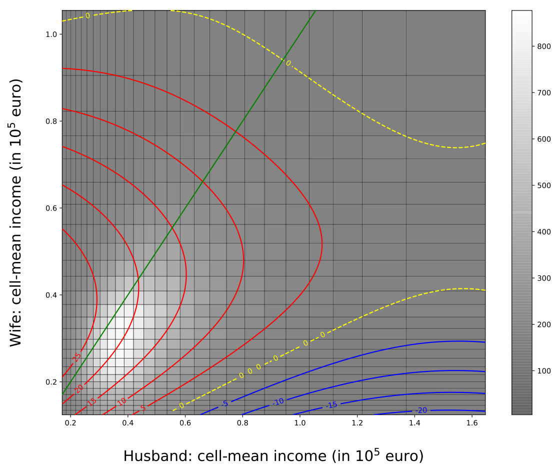


Figure 3: Supermodular core, surface levels, all newly wed

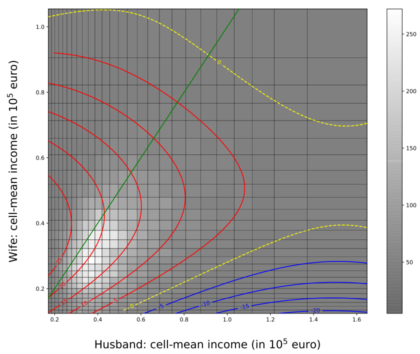
These couples can be called "traditional", in the sense that they follow the "man as breadwinner" model: his income is fairly high (above 60,000 euros), while her is much lower (less than 30,000). They represent about 5% of the population, i.e. more than 5,500 couples.

All in all, the main message is that there exists a considerable amount of heterogeneity across couples, including in the very nature of the marital gain. This supports the view that when considering matching on income, only a flexible form, estimated on a very large population, can allow to fully understand marital patterns.

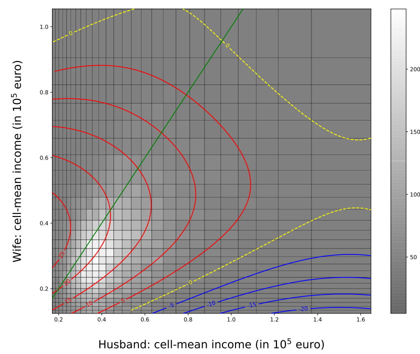
5.1.2 Year-by-year estimation

Our sample is large enough to allow for an independent estimation of each year's marriages; these are plotted in Figure 4.[10] A striking feature is the remarkable stability of the qualitative patterns depicted by these graphs. In all cases, one observes strongly positive assortativeness in the South-West part of the graph and significantly negative assortativeness for "traditional" couples at the South of the graph. This pattern stability is all the more remarkable that the coefficients of the polynomials (given

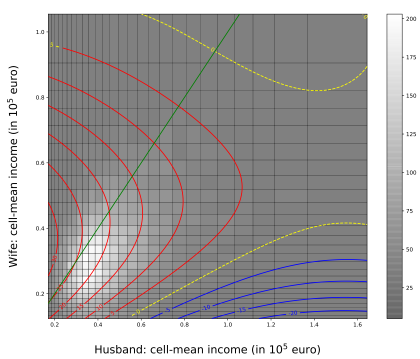
[10]For brevity, we only include the projections graphs. All other representations are available from the authors.



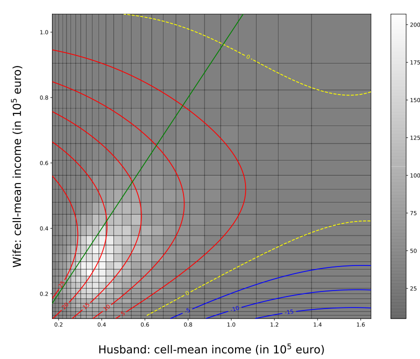
(a) Year 2011



(b) Year 2012



(c) Year 2013



(d) Year 2014

Figure 4: Supermodular Core, surface levels, newly wed 2011-2014

in Table A4 in Appendix) vary significantly between years, as can be expected in our highly flexible formulation.

5.1.3 Matching on past or future income

The size of our data set, as well as its panel structure, enables us to investigate further the structure of marital patterns and to address some potential problems. A first concern is that income at marriage may partly reflect labor supply decisions induced by the change of marital status. In that case, our estimates would suffer from an endogeneity problem, since it would amount to attributing to the matching game patterns that are actually ex-post *generated* by the match itself.

To further investigate this issue, we estimate the same model on

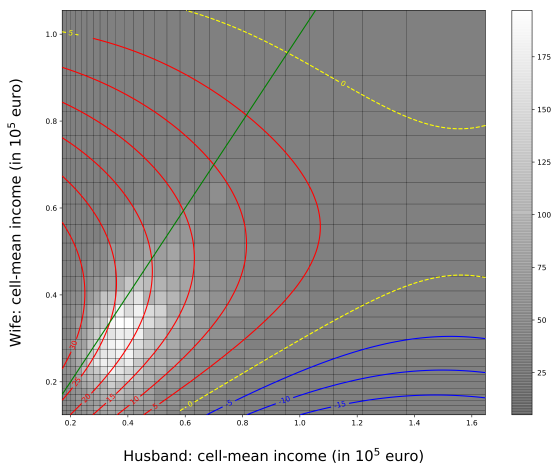


Figure 5: Supermodular core, surface levels, married in 2014, 2011 incomes

the newly wed in 2014, *using income in 2011 as the matching trait.* If endogeneity of labor supply is a serious issue, one would expect that the same marriages, using income data from three years before as a matching trait, would generate significantly different patterns. The graph is given in Figure 5. One can note that the qualitative patterns are exactly similar to those obtained using 2014 incomes. This suggests that labor supply responses to changes in marital status are unlikely to affect our analysis of matching patterns based on income at marriage.

Next, we perform the opposite exercise; that is, we consider couples that married in 2011 and use individual incomes in 2014 - i.e., three years post marriage - as a matching trait. The concern, here, is that our analysis might underestimate the importance of specialization as a component of marital gain, because specialization may only appear several months (or perhaps years) *after* marriage. If that is the case, we could expect the area of NAM to become larger when post-marital incomes are considered.

The estimates, as depicted in Figure 6, contradict this hypothesis. If anything, supermodularity appears to be even more related to equal incomes; moreover, the "traditional couples" region, at the South of the

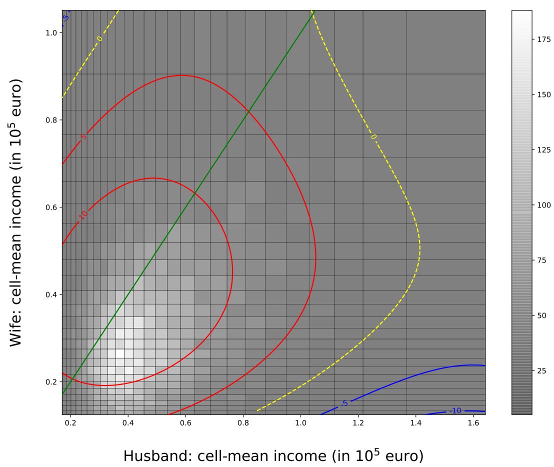


Figure 6: Supermodular core, surface levels, married 2011, 2014 incomes

graph, substantially shrinks and becomes barely significant.

A possible objection is that the time interval allowed by our data - three years between 2011 and 2014 - is too short. This will be a problem if specialization mostly takes place after the birth of children, which may happen several years after marriage. While our data do not allow us to precisely test this explanation, we can replicate our analysis on a representative sample of the whole population, therefore including (mostly) couples who have been married for several years; we use current income as the matching trait. An additional advantage of this analysis is that it replicates what most studies do - i.e., consider a sample of married couples and analyze matching patterns using currently observed characteristics, thus neglecting the fact that these characteristics (and particularly income) may in fact be endogenous to marriage. While one can readily understand the *potential* bias generated by such an approach, its exact magnitude and its qualitative consequences remain largely unknown. Our data allows us to precisely analyze this issue.

The outcome is given in Figure 7. While the qualitative features remain similar to those of the newly wed, preferences for assortativeness (whether positive or negative) appear to be much weaker. One still observes preferences for homogamy in the South-West part of the dis-

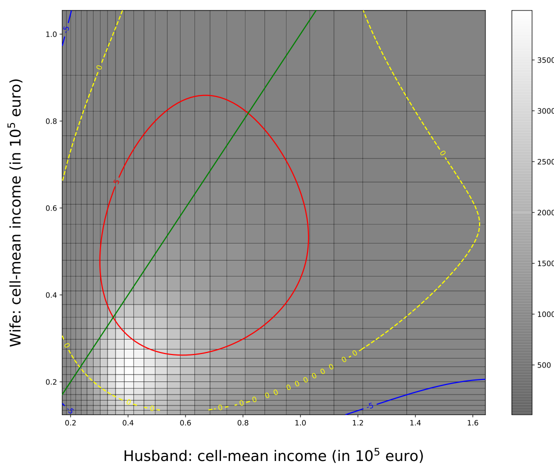


Figure 7: Supermodular core, surface levels, whole population

tribution. However, the corresponding area clearly decreases, and PAM only remains significant for a minority of the total population; whereas the NAM preferences of "traditional" couples (with high male and low female incomes) all but disappears.

These findings are tricky to interpret, if only because the whole population obviously consists of different age groups; the impact of length of marriage (and resulting changes in behavior) cannot be distinguished from possible cohorts effects. At any rate, these results confirm that an estimation based on a sample of existing couples of heterogeneous ages can give at best a biased vision of actual matching patterns.

5.2 Robustness tests and extensions

5.2.1 Placebo estimates

In order to have an idea of the power of the implemented tests, we run a series of additional tests. A first concern is that our method might systematically conclude to positive (resp. negative) assortativeness whenever spouses' income are very similar (very different), even when these patterns are in fact purely random. In principle, the structural model underlying the approach is a protection against such a risk, since accidental aspects are captured by the random terms, whereas the deterministic

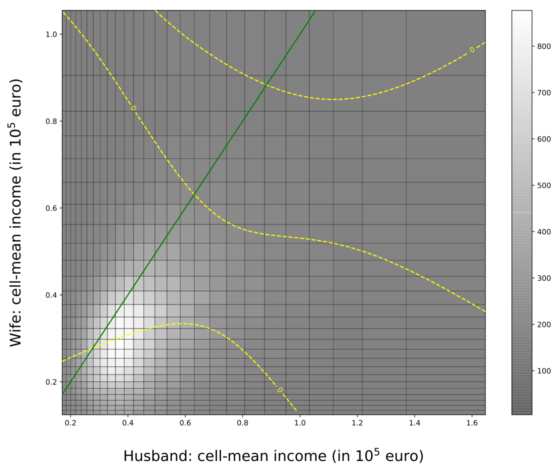


Figure 8: Placebo regression, all newly wed

component, from which the second cross derivative is computed, characterizes systematic patterns. Yet, this argument may be vulnerable to any misspecification bias.

In order to empirically assess the importance of this problem, we run a placbo regression in which we consider the same individuals as before, but match them randomly; we then perform the same analysis as before on this placebo sample. The results are quite reassuring. Figure 8 plots the outcome of our placebo test on all couples resulting from random matching of all individuals married over the four years.[11] One can see that nowhere on the whole income range does the second cross derivative of the surplus function significantly differ from zero. In other words, while random matching does produce some couples with very unequal incomes, our methods is able to identify these as outliers and to recognizes that no systematic pattern is at work. Conversely, this confirms that the results obtained on the actual population indeed reflect systematic matching patterns in the population.

[11] Similar placebo tests have been performed on the other samples, and give similar results.

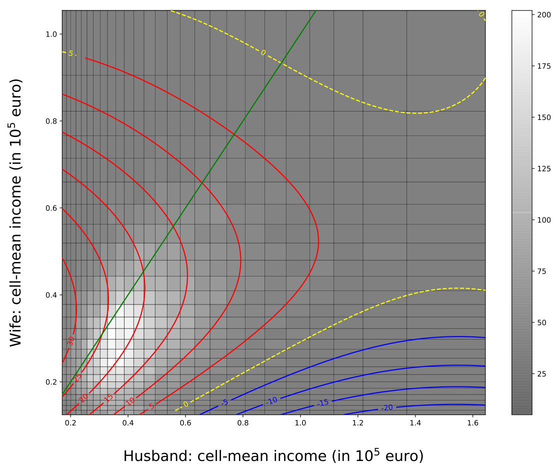


Figure 9: Supermodular core, surface levels, married 2013, average incomes

5.2.2 Current versus permanent income

Another possible weakness of our approach (and, for that matter, of all matching models based on income data) is that, in principle, one may expect people to match on *permanent* income, of which the matching trait we use, namely current income, is a noisy measure. Unfortunately, there is not much we can do to fully address this concern given the data we currently have at disposal; in particular, we observe neither education, nor any alternative measure of individual human capital. Yet, we can use the panel structure of the data to generate an average of income over the 4 years. Estimating the same model using average incomes as matching traits will, to some extent, smooth out temporary shocks. It is therefore important to check whether the qualitative conclusions remains valid.

In practice, we estimate the model on the sample of newly wed couples in 2013, replacing current incomes with average (wage and salary) income over the period 2011-2014 (discounting incomes with the CPI). The corresponding graph is provided in Figure 9. The density of income distribution is significantly affected; in particular, extreme income realizations are, as expected, less frequent. However, the supermodular core graph is basically unchanged. This suggests that the patterns emerging from the structural analysis are unlikely to be affected by temporary shocks.

5.2.3 Surplus determined by total income only

Several applied theory models use a simple version of the matching game in which the surplus generated only depends on total income (this would be the case, for instance, if individual labor supplies are fixed and the surplus is entirely generated by public consumptions within the household). In our setting, this corresponds to a set of restrictions on the polynomial form, reflecting the fact that it can be expressed as a function of the sum $(x + y)$ only.

The results are quite clear. First, the implied restrictions are very strongly rejected. Second, the resulting matching patterns are totally different from what the general model indicates. In the restricted version, supermodularity is equivalent to the convexity of the surplus function (which only depends on one variable, namely total income). Here, the surplus is estimated as a 6^{th} degree polynomial in total income for all newlywed; Figure 10 plots its second derivative with respect to total income. It is negative (indicating NAM) below a threshold of approximately $90,000$ euros per year, positive beyond that level - in sharp contrast with the findings of the general model.

This conclusion is by no means surprising. As Figure 3 clearly indicates, a given level of total household income gathers couples with very different characteristics, some of which (with very different incomes) display NAM where others clearly match assortatively. Regressing on total income amounts to ignoring this heterogeneity, which results in essentially meaningless conclusions.

5.3 Gender asymmetry and the 50% threshold

Lastly, our data allow to reconsider an intriguing finding of the previous literature - namely the existence of a discontinuity, in the income distribution of married couples, when the wife's income reaches 50% of total household income (Bertrand et al., 2015).

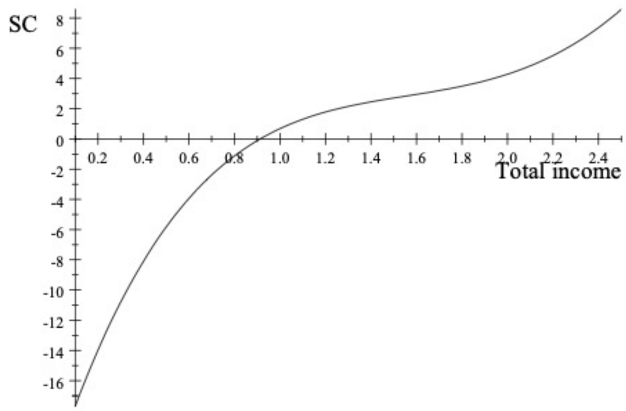
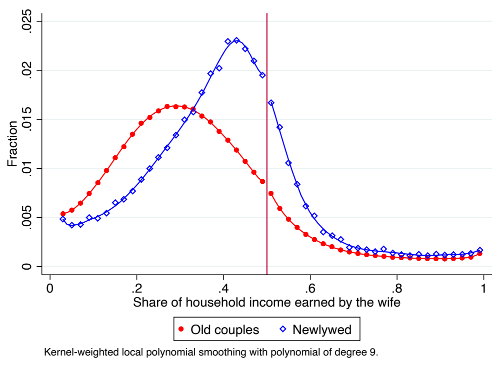


Figure 10: Supermodular core, total income

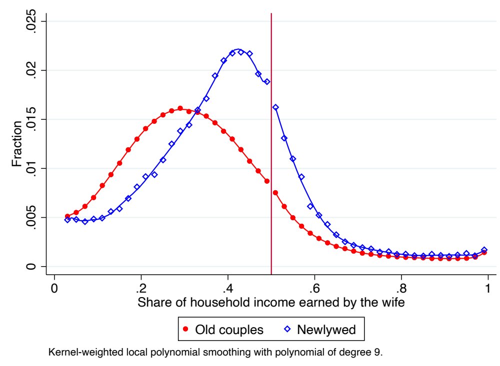
This phenomena, however, has several possible explanations. One, supported by the authors, relies on social norms affecting marital patterns: the husband was historically supposed to be the breadwinner providing for the entire family, and most men are still reluctant to marry a spouse whose income is larger than their own. Alternatively, the finding may result from specific, post-marriage labor supply decisions. For instance, women, when married, may choose to reduce their working time, so that their earnings become inferior to their husband's. Such a mechanism, although interesting per se, should certainly not be considered as a direct feature of the prevailing matching patterns. Finally, the discontinuity might simply reflect biases in self-reporting of earnings. To the extent that the "husband as the breadwinner" social norm plays a significant role, it may as well generate incentives for a wife to underreport her income (and/or a husband to overreport his) whenever the former exceeds the latter.

In our data, all incomes are third-party reported, which basically eliminates the third explanation; and we use data at or before marriage, thus avoiding the second. In other words, our data allow us to concentrate on the possible existence of the first explanation, namely the existence of a social norm affecting decisions on the matching market.

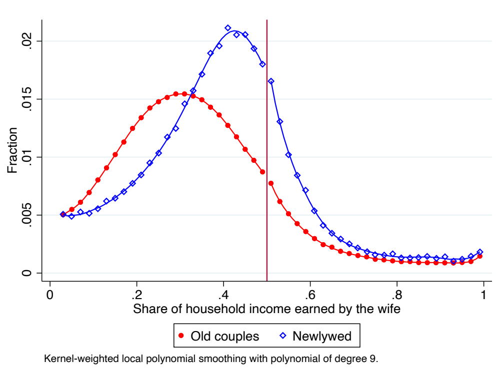
As a preliminary investigation, one can have a brief look at the in-



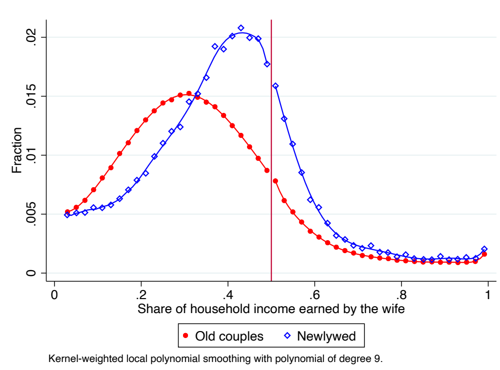
(a) Year 2011



(b) Year 2012



(c) Year 2013



(d) Year 2014

Figure 11: Income Ratio, distribution per year

come distribution of all newly wed couples (using income at marriage), and check whether a discontinuity is visible for the equal income line. This is depicted in Figure 11, where the equal income threshold is in red. While a majority of couples sit below the line, where his income exceeds her, a significant fraction belongs to the opposite part of the graph; more importantly, the 50% threshold does not seem to correspond to a clear discontinuity in population density.

A clear advantage of our approach is that it allows us to formally test this conclusion. In practice, thus, we estimate on all newly wed couples the same model as before with a twist - namely, we introduce a dummy variable equal to 1 if his income exceeds hers, 0 otherwise.

We find that the coefficient of the dummy is positive and significant, but the magnitude of the effect is very small. Our estimate is 0.162

(with a standard deviation of 0.013), while the total surplus variations over the sample exceeds 10; that is, the threshold effect represents less that 2% of observed variations. Moreover, the estimated polynomial coefficients in the new model are almost exactly identical to the initial ones. In other words, the importance of the so-called "50% threshold" phenomenon appears to be at best negligible.

One can however go one step further. A possible explanation for the significance of this coefficient is that, far from reflecting any social norm, it simply captures some non linearities in the surplus function that, given our large sample size, our polynomial representation is not flexible enough to fully take into account.

To explore this possibility, we run a series of placebo tests in which the dummy "1 if his income exceeds hers" is replaced with "1 if his income exceeds λ% of hers", with λ takes the values .8, .9, 1, 1.1 and 1.2 (corresponding to ratios of her to total household income respectively equal to 56%, 53%, 50%, 48% and 45%). For each regression, the coefficient of the dummy variable significantly differs from 0; the values are plotted on Figure 12. If anything, surplus is slightly larger when his income represents at least 80% of hers, but decreases when it exceeds 110%, although the exact interpretation of this result is not clear. At any rate, the 50% ratio does not seem to play a particular role.

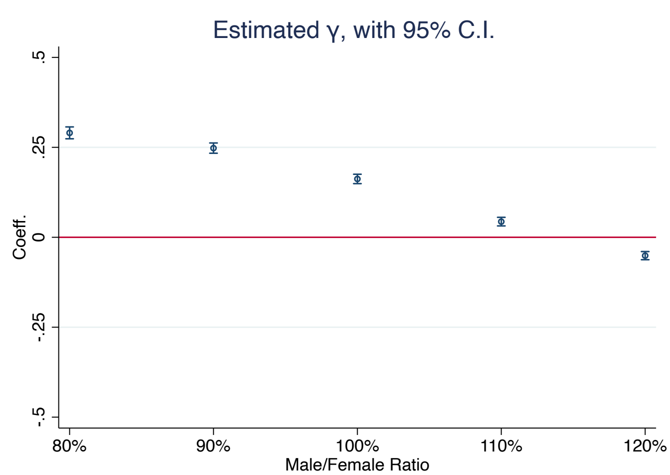


Figure 12: Coefficient of various income ratios

Again, in order to assess the power of these results, we run the same placebo regression as before, using randomly matched individuals. Not surprisingly, the coefficients of the various dummies never significantly differ from zero, as indicated in Figure 13.

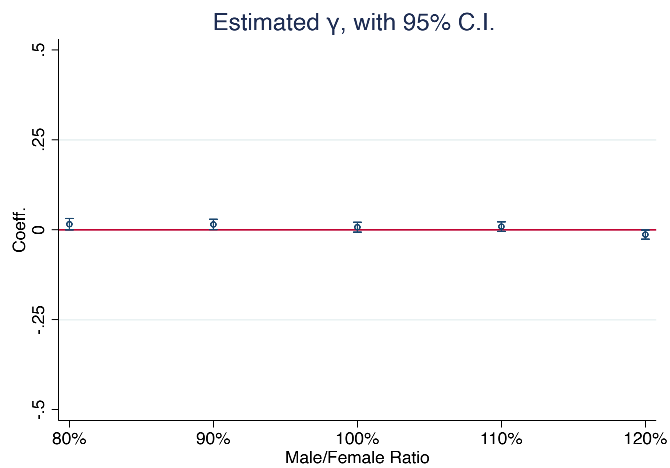


Figure 13: Coefficient of various income ratios - Placebo regression

Interestingly, our conclusion is in line with recent contributions in the literature. As noted by Hederos and Stenberg (2019), several data sets exhibit an important spike in the distribution of the wife's share of household income at the point where spouses earn *exactly* the same - a fact that is suggestive of declaration biases. Such biases are all the more likely when individual incomes are not paid by a third party but generated within the household, as is the case with self-employed persons, family businesses or co-working of spouses. Using linked employer-employee data from Finland, Zinovyeva and Tverdostup (2021) document the existence of a discontinuity of the same magnitude as in the U.S. and show that it can be fully explained by the convergence of *declared* earnings by spouses who start working together; the convergence, in turn, can be explained by various factors (including tax optimization). Finally, Slotwinski and Roth (2020) exploit a unique Swiss data set combining survey and administrative information for the same individual and their partner to demonstrate that individuals misreport incomes in surveys to comply with the male breadwinner norm, and that this misreporting essentially explains the observed discontinuity in reported earnings. Since incomes, in our data set, are third-party reported, one can hardly expect to find a significant effect of the 50% threshold - and, indeed, we don't.

6 Conclusions

Several conclusions can be drawn from our analysis. First, an analysis of matching patterns based on income requires both highly specific data and adequate empirical approaches. Income is a continuous variable; while discretization is certainly feasible, the number of categories must be quite large. Yet, one should concentrate on marriage taking place over a short period of time, since there is no reason to a priori expect different cohorts to exhibit identical patterns. These two requirements are compatible only for very large sample sizes. Moreover, in order to avoid endogeneity problems generated by labor supply response, income should be observed at, or ideally, before marriage for both spouses. Last but not least, incomes should be third-party rather than self-reported, since reporting biases are probably a serious issue. Our data set, which exhibits all these characteristics, arguably constitutes an ideal starting point.

Second, marital patterns are unlikely to be uniform over the entire population. From a theoretical perspective, the exact nature of the economic gains generated by marriage have been extensively discussed. Becker (1973) early contribution emphasizes benefits generated by specialization within the couple - one spouse specializing in domestic work while the other spends most of their time on the market. As recognized by Becker himself, specialization typically leads to *negative* assortative matching. In contrast, recent works have pointed out other sources of marital gains, such as risk sharing and investments into children's human capital - and more generally the existence of public goods within the household; these generate strong preferences for positive assortative matching.[12]

In practice, both types of gains are likely to coexist in most cases.

[12] See Chiappori (2017) for a detailed discussion.

Many couples exhibit at least some kind of specialization (for instance, women typically spend more time than men on domestic chores); yet, most of them have children and invest into their human capital, and basically all benefit from public consumptions such as shared housing. The crucial issue is to determine which effect dominates in any particular couple; and there is no reason to believe that the answer should be the same for all incomes.

A different (but equivalent) way to consider the issue relies on the idea that, in frictionless matching models, assortativeness of the stable matching is related to supermodularity of the surplus function, i.e., under a differentiability assumption, to the sign of its second cross derivative with respect to the male and female traits (here, incomes). Models based on specialization typically generate a negative sign for all incomes, while those involving public consumption or children investment predict a positive value.

In real life, however, there is no reason to expect that the sign be constant over the whole domain of individual incomes. One should therefore aim at estimating this sign *locally*, i.e. at (or in a small neighborhood of) any possible pair of incomes. This task is however quite difficult, because this local property - a negative sign for a second cross derivative - drives behavioral patterns that are by no means local. While spousal incomes are indeed positively correlated, individuals do not exclusively marry a spouse with a similar income; on the contrary, one can observe, on a large enough data set, that the support of the distribution of the husband's income conditional on the wife often covers the entire range of male incomes (and conversely). From that perspective, estimating (a flexible approximation to) the global surplus, then computing the assortativeness patterns implied by these estimates, appears as an interesting possible solution. This is the approach we adopt in our paper.[13]

This strategy, however, comes with specific requirements. In par-

[13] This situation is reminiscent of the estimation of the distribution of risk aversion over the population: while risk aversion certainly is a local measure, it is often estimated from behavioral responses to non-local lotteries. Several authors have actualy argued that risk aversion could *not* be estimated from choices regarding local lotteries - see for instance Rabin (2000).

ticular, imposing a quadratic form for the surplus, convenient as that may be in applied theory models, typically provides poor empirical results, because the restrictions it imposes - uniform tendencies towards homogamy over the whole sample - are way too restrictive and actually counterfactual. We therefore argue that one should adopt a formulation that is flexible enough to allow for significant variations in preferences for homogamy - that is, formally, in the sign of the second cross derivative of the surplus function. A key purpose of our paper is precisely to provide a consistent approach for formulating and implementing a flexible strategy well adapted to this type of data. We show that a semi-parametric model, based on a flexible, high-degree polynomial representation of the surplus function, can be estimated.

Regarding the results, we first find that the existing population exhibits a large degree of heterogeneity. In particular, while most couples match assortatively, a small but significant minority exhibits strong preferences for negative assortative matching. Typically, these are "traditional" couples, where the husband is the main breadwinner and the wife's income plays a marginal role at best. These patterns are remarkably stable over time, at least for the four years under consideration; while the precise estimates of polynomial coefficients differ a lot between years, the resulting areas of positive and negative assortativeness remain essentially constant. Moreover, our analysis provides a precise estimate of the number of "traditional" couples in the population. The proportion we find, around 5%, is surprisingly small. It would be quite interesting to repeat the analysis on older data, and check whether (and how) these patterns vary across generations.

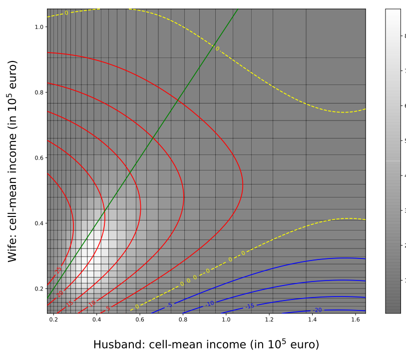
Interestingly, a similar analysis applied to the whole married population - thus including couples from several different cohorts - gives somewhat different conclusions. This suggests that for many standard approaches, which consider already married couples and current incomes, cohort effects and/or labor supply response to marital status may significantly bias the conclusions. Finally, while permanent incomes cannot be robustly estimated from our data, our robustness tests tend to suggest that transitory income shocks have no impact on marital patterns, as expected.

All in all, we view this paper as an additional step towards a better understanding of matching patterns in developed economies. Obviously, much work remains to be done. In particular, could existing data be supplemented by additional information related to individual human capital (such as education or working experience), a more robust investigation would become possible. Moreover, observing actual behavior can provide additional information about the surplus itself, thus leading to more robust estimations. One possibility, in the spirit of Chiappori et al. (2018), is to analyze post-marital behavior, which allows to recover the household's (total) utility. Less structurally, one can observe specific outcomes, such as divorce, that are arguably (negatively) correlated with the match-specific surplus. We believe that the methodology developed in this work could fruitfully be applied to a more ambitious endeavour of this kind.

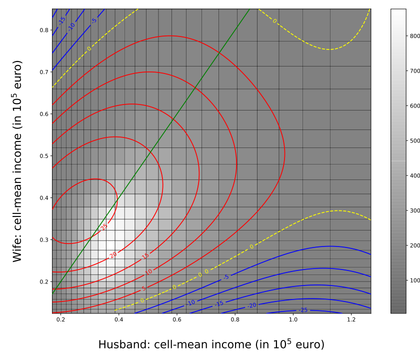
7 Appendix: robustness checks

7.1 Changing the bins' structure

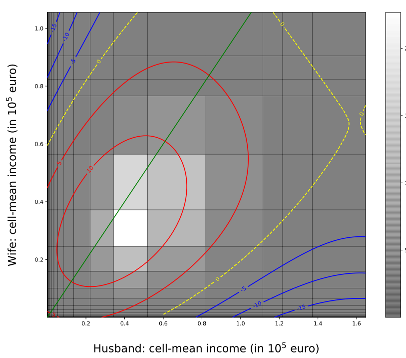
Here we present a selection of robustness checks of results changing the discretization in bins of the income domain, as described in Section 3.2, focusing on all newly weds. In panel (a) we report the same picture presented in Figure 3, just for the sake of a clear comparison with other bin's structures. In panel (b) we keep the same bin structure for the 50^{th}-99^{th} income range but divide the top 1% in three groups instead of five, using the 99.25, 99.5 and 99.75 quantiles. In panels (c) and (d) we keep the same partitioning of the top 1% as in panel (a), but change the 25-binning structure for incomes below the 99^{th} percentile. In particular, in panel (c) we use the range $[q^z_{25}, q^z_{99}]$ and a step defined as $(q^z_{99} - q^z_{25})/25$, whereas in panel (d) we use the range $[q^z_{40}, q^z_{99}]$ and a step defined as $(q^z_{40} - q^z_{25})/25$.



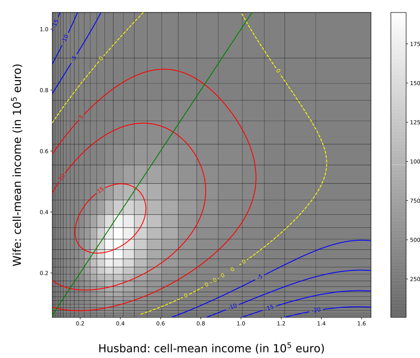
(a) As in Figure 3



(b) Changing the discretization of the top 1%



(c) Changing the discretization of the bottom 99%, using the q^z_{25}, q^z_{99} quantile range.

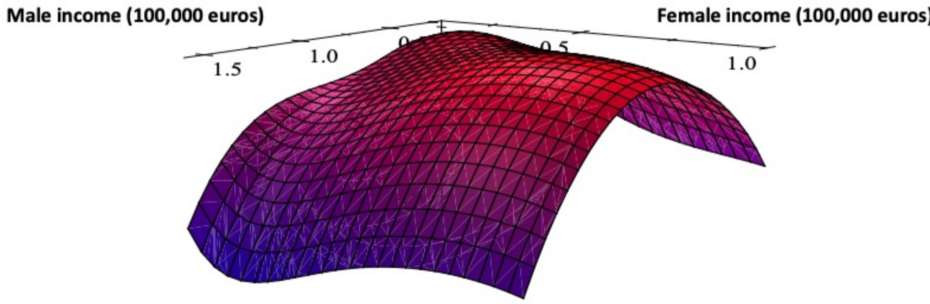


(d) Changing the discretization of the bottom 99%, using the q^z_{40}, q^z_{99} quantile range.

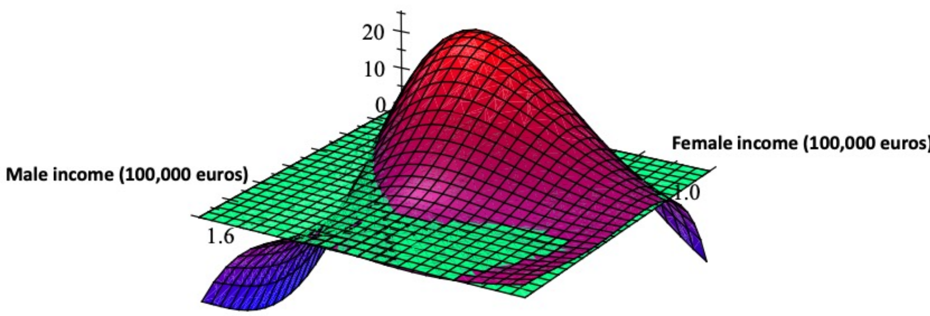
Figure 14: Supermodular core, surface levels, all newly wed, for different bins' structures

7.2 Changing the degree of the polynomial

While we did not do formal tests regarding the optimal polynomial degree, we performed a simple verification by implementing the same regression with a polynomial of degree 7 (instead of 6). The resulting estimate of the supermodular core and its projection for all newly wed couples are shown in Figures 15 and 16, respectively. The main patterns are largely similar; in particular, we still observe strong preferences for homogamy for most couples but a significant minority of "traditional" couples displaying negative assortative matching. The only notable dif-



(a) Surplus



(b) Supermodular Core

Figure 15: All newly married

ference is a second zone of NAM in the North-West part of the graph, i.e. for couples where her income is very large (above €70,000) while he is making less than €50,000. An important point is that, as can be seen from the density background, such couples are very rare - less than 250 couples, i.e. around 0.3% of newly wed couples - and can safely be ignored.

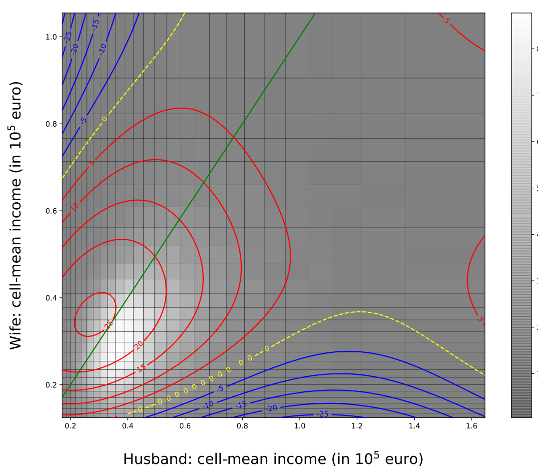


Figure 16: Supermodular core, all newly wed, polynomial of degree 7

	2011		2012		2013		2014		2011-2014	
Var.	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*
$\lambda^{0,0}$	-8,248	0,272	-7,943	0,273	-7,045	0,278	-6,510	0,265	-7,429	0,135
$\lambda^{1,0}$	28,615	1,359	25,677	1,350	23,105	1,412	18,171	1,320	23,823	0,676
$\lambda^{0,1}$	17,768	2,431	19,557	2,509	12,923	2,571	12,783	2,455	15,719	1,242
$\lambda^{1,1}$	3,786	4,483	1,221	4,654	14,360	4,758	15,674	4,385	9,146	2,260
$\lambda^{2,0}$	-75,756	3,557	-67,517	3,618	-68,272	3,815	-52,191	3,567	-65,741	1,810
$\lambda^{0,2}$	-80,609	10,490	-90,262	10,862	-68,264	11,230	-67,292	10,645	-76,614	5,389
$\lambda^{1,2}$	104,726	10,709	103,280	12,143	94,237	11,918	80,220	11,112	95,435	5,681
$\lambda^{2,1}$	-30,955	6,302	-23,999	5,904	-40,904	6,627	-36,586	5,607	-33,672	3,009
$\lambda^{3,0}$	88,546	4,543	77,263	4,740	84,550	5,006	61,070	4,677	77,594	2,360
$\lambda^{0,3}$	86,912	22,230	113,607	23,153	67,083	24,016	69,246	22,567	84,125	11,463
$\lambda^{1,3}$	-143,314	13,939	-144,136	15,794	-135,807	15,284	-115,339	14,059	-134,956	7,305
$\lambda^{2,2}$	14,591	2,651	16,529	2,626	19,086	2,875	15,261	2,256	16,678	1,228
$\lambda^{3,1}$	4,368	4,983	-1,856	4,776	8,627	5,415	8,473	4,376	5,256	2,400
$\lambda^{4,0}$	-47,478	2,842	-40,324	3,010	-46,942	3,198	-32,065	2,959	-41,540	1,494
$\lambda^{0,4}$	-31,816	23,528	-63,290	24,784	-13,857	25,650	-19,780	23,844	-31,850	12,184
$\lambda^{1,4}$	68,112	7,832	68,333	8,855	65,438	8,495	53,724	7,671	64,304	4,052
$\lambda^{2,3}$	3,791	0,900	4,477	1,086	2,091	0,983	3,661	1,065	3,279	0,478
$\lambda^{3,2}$	-6,022	0,834	-6,900	0,773	-6,210	0,952	-5,749	0,607	-6,229	0,368
$\lambda^{4,1}$	1,671	1,585	3,721	1,556	0,597	1,773	0,197	1,382	1,434	0,771
$\lambda^{5,0}$	11,433	0,792	9,508	0,849	11,613	0,908	7,586	0,828	9,995	0,420
$\lambda^{0,5}$	0,382	11,556	16,608	12,324	-7,995	12,686	-3,717	11,628	1,024	5,999
$\lambda^{1,5}$	-11,126	1,481	-11,119	1,681	-10,650	1,606	-8,430	1,414	-10,434	0,761
$\lambda^{2,4}$	-1,135	0,177	-1,426	0,230	-1,199	0,213	-1,378	0,231	-1,227	0,101
$\lambda^{3,3}$	0,180	0,079	0,268	0,084	0,453	0,092	0,378	0,068	0,313	0,038
$\lambda^{4,2}$	0,522	0,090	0,590	0,086	0,447	0,107	0,422	0,066	0,499	0,040
$\lambda^{5,1}$	-0,281	0,154	-0,488	0,156	-0,183	0,177	-0,124	0,135	-0,258	0,076
$\lambda^{6,0}$	-0,944	0,072	-0,783	0,079	-0,984	0,084	-0,625	0,076	-0,831	0,039
$\lambda^{0,6}$	1,222	2,016	-1,666	2,179	2,665	2,225	1,741	1,999	1,063	1,047
N. Obs	2,329,769		2,287,373		2,226,968		2,226,631		9,070,741	
N. cells	899		899		899		899		899	

Notes: Surplus S^{λ}_{xy} (see equation 5) is estimated via a flexible polynomial function of degree 6 in male and female incomes using yearly data or pooling all considered years (2011-2014) together.

Table 4: Coefficients of the parametric surplus function, S^{λ}_{xy}

	2011		2012		2013		2014		2011-2014	
Var.	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*	*Coeff.*	*Std. Err.*
$\lambda^{0,0}$	-10,759	0,540	-9,713	0,537	-9,926	0,554	-9,101	0,534	-9,810	0,269
$\lambda^{1,0}$	53,396	3,317	46,879	3,270	44,576	3,397	42,810	3,263	46,375	1,646
$\lambda^{0,1}$	19,052	5,773	15,940	5,938	26,831	6,166	17,246	5,933	19,527	2,963
$\lambda^{1,1}$	-52,484	10,965	-33,442	11,122	-49,062	11,570	-37,819	10,987	-41,988	5,484
$\lambda^{2,0}$	-154,635	11,304	-141,287	11,448	-133,266	11,875	-133,255	11,480	-138,799	5,739
$\lambda^{0,2}$	-40,894	32,192	-40,670	33,219	-105,603	34,624	-53,314	33,027	-59,589	16,580
$\lambda^{1,2}$	220,531	30,096	184,060	34,875	222,779	35,047	240,448	34,703	216,883	16,494
$\lambda^{2,1}$	47,528	21,822	25,759	20,412	57,720	22,381	14,848	19,962	33,372	10,442
$\lambda^{3,0}$	222,228	19,967	208,494	20,723	188,977	21,416	204,767	20,815	203,257	10,326
$\lambda^{0,3}$	-92,970	93,113	-76,435	96,199	119,798	100,362	-48,352	94,880	-25,815	47,926
$\lambda^{1,3}$	-340,506	57,717	-316,086	68,531	-405,292	68,024	-440,919	68,089	-374,413	32,107
$\lambda^{2,2}$	18,079	12,803	42,547	12,508	43,614	12,933	31,464	10,344	31,855	5,753
$\lambda^{3,1}$	-103,400	25,944	-84,801	25,111	-131,673	27,633	-70,011	24,020	-92,777	12,678
$\lambda^{4,0}$	-163,787	19,073	-158,373	20,131	-131,593	20,788	-161,979	20,199	-151,756	9,984
$\lambda^{0,4}$	308,929	146,223	288,845	151,881	-21,790	158,161	259,436	148,516	210,046	75,341
$\lambda^{1,4}$	191,461	57,510	191,414	68,928	291,109	67,446	334,609	67,597	252,368	31,973
$\lambda^{2,3}$	53,099	7,417	38,428	8,330	29,013	7,909	32,437	7,423	37,944	3,750
$\lambda^{3,2}$	-34,387	7,274	-40,289	6,409	-35,333	7,263	-28,593	4,877	-33,030	2,947
$\lambda^{4,1}$	74,869	15,092	63,976	15,051	92,626	16,571	53,504	14,014	68,081	7,475
$\lambda^{5,0}$	63,584	9,620	64,055	10,287	46,801	10,636	68,307	10,278	59,874	5,079
$\lambda^{0,5}$	-299,470	123,353	-297,567	129,262	-37,556	134,039	-284,724	125,155	-230,604	63,711
$\lambda^{1,5}$	-45,421	27,010	-50,808	32,559	-99,597	31,486	-123,645	31,350	-80,873	14,955
$\lambda^{2,4}$	-26,823	2,980	-23,071	3,570	-18,175	3,256	-18,614	3,459	-21,126	1,583
$\lambda^{3,3}$	2,651	1,279	4,282	1,389	3,640	1,401	3,979	1,120	3,478	0,621
$\lambda^{4,2}$	6,324	2,121	7,384	1,890	6,548	2,204	3,887	1,374	5,574	0,859
$\lambda^{5,1}$	-20,491	3,982	-17,649	4,066	-25,446	4,479	-14,517	3,697	-18,619	1,990
$\lambda^{6,0}$	-12,174	2,306	-12,872	2,499	-8,047	2,587	-14,247	2,480	-11,695	1,227
$\lambda^{0,6}$	124,361	51,320	130,147	54,363	22,887	56,033	128,356	51,935	101,878	26,574
$\lambda^{1,6}$	3,722	4,578	5,006	5,555	13,343	5,320	17,806	5,212	10,283	2,523
$\lambda^{2,5}$	3,337	0,547	3,197	0,673	2,465	0,628	2,206	0,655	2,649	0,297
$\lambda^{3,4}$	1,023	0,194	0,638	0,217	0,502	0,234	0,679	0,178	0,750	0,100
$\lambda^{4,3}$	-0,757	0,127	-0,755	0,136	-0,592	0,146	-0,729	0,108	-0,711	0,062
$\lambda^{5,2}$	-0,306	0,204	-0,405	0,191	-0,386	0,222	-0,079	0,137	-0,247	0,086
$\lambda^{6,1}$	1,787	0,351	1,563	0,368	2,264	0,405	1,261	0,329	1,635	0,178
$\lambda^{7,0}$	0,870	0,192	0,966	0,213	0,523	0,220	1,095	0,209	0,855	0,104
$\lambda^{0,7}$	-18,738	8,013	-20,494	8,594	-3,804	8,789	-20,398	8,041	-15,963	4,156
N. Obs	2,329,769		2,287,373		2,226,968		2,226,631		9,070,741	
N. cells	899		899		899		899		899	

Notes: Surplus S^{λ}_{xy} (see equation 5) is estimated via a flexible polynomial function of degree 7 in male and female incomes using yearly data or pooling all considered years (2011-2014) together.

Table 5: Coefficients of the parametric surplus function, S^{λ}_{xy}

References

Aiyagari, S. R., J. Greenwood, and N. Guner (2005, May). Looking Back: Marriage, Divorce and Out-of-Wedlock Births. RCER Working Papers 516, University of Rochester - Center for Economic Research (RCER).

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

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics 130*(2), 571–614.

Bojilov, R. and A. Galichon (2016). Matching in closed-form: equilibrium, identification, and comparative statics. *Economic Theory 61*(4), 587–609.

Chiappori, P.-A. (2017). *Matching with transfers.* Princeton University Press.

Chiappori, P.-A., M. Costa Dias, and C. Meghir (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy 126*(S1), S26–S72.

Chiappori, P.-A., M. Costa Dias, and C. Meghir (2020, April). Changes in assortative matching: Theory and evidence for the us. Working Paper 26932, National Bureau of Economic Research.

Chiappori, P.-A., R. J. McCann, and B. Pass (2017). Multi-to one-dimensional optimal transport. *Communications on Pure and Applied Mathematics 70*(12), 2405–2444.

Chiappori, P.-A. and B. Salanié (2016). The econometrics of matching models. *Journal of Economic Literature 54*(3), 832–61.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review 107*(8), 2109–67.

Choo, E. and A. Siow (2006a). Estimating a marriage matching model with spillover effects. *Demography 43*(3), 463–490.

Choo, E. and A. Siow (2006b). Who marries whom and why. *Journal of Political Economy 114*(1), 175–201.

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

Fernandez, R., N. Guner, and J. Knowles (2005). Love and money: A theoretical and empirical analysis of household sorting and inequality. *The Quarterly Journal of Economics 120*(1), 273–344.

Galichon, A. and B. Salanié (2017). The econometrics and some properties of separable matching models. *American Economic Review 107*(5), 251–55.

Galichon, A. and B. Salanie (2021). Cupid's invisible hand: Social surplus and identification in matching models. *Review of Economic Studies*.

Hederos, K. and A. Stenberg (2019, June). Gender Identity and Relative Income within Households-Evidence from Sweden. Working Paper Series 3/2019, Stockholm University, Swedish Institute for Social Research.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011, 05). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica 79*(3), 651–692.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Schwartz, C. R. and N. L. Graf (2009). Assortative matching among same-sex and different-sex couples in the united states, 1990–2000. *Demographic Research 21*, 843.

Slotwinski, M. and A. Roth (2020). Gender norms and income misreporting within households. *ZEW-Centre for European Economic Research Discussion Paper* (20-001).

Zinovyeva, N. and M. Tverdostup (2021, October). Gender identity, coworking spouses, and relative income within households. *American Economic Journal: Applied Economics 13*(4), 258–84.