Economica

Economica (2022) **89**, 29–43
doi:10.1111/ecca.12386

Marriage Market Counterfactuals Using Matching Models

By ARNAUD DUPUY* and SIMON WEBER†

**CREA, University of Luxembourg, IZA, †University of York*

Final version received 25 June 2021.

We use a simple structural matching model with unobserved heterogeneity to produce counterfactual marriage patterns, and thus quantify the contribution of changes in marital patterns in rising income inequality. We propose an algorithm that allows us to fix the degree of assortative mating without changing the level of marital gains and hence isolate the intensive and extensive margins (i.e. isolate changes in assortative mating from changes in marriage rates). We apply this approach to US data from 1962 to 2017, and show that marital patterns can explain about a quarter of the rise in income inequality, the intensive margin contributing 7%, the extensive margin the remaining 93%. Our algorithm also allows us to show that the extensive margin is itself driven for three-fifths by a change in the total number of singles and for two-fifths by a change in the distribution of types among singles (in particular low-educated women).

INTRODUCTION

Most developed countries have witnessed a dramatic rise in income inequality over the last six decades (Piketty and Saez 2003). As a result, the study of the sources of income inequality has been increasingly popular among economists, with plausible explanations coming from unexpected fields such as family economics. A recent stream of research (e.g. Burtless 1999; Eika *et al*. 2019; Schwartz 2010) studies the extent to which the marriage market, and specifically the degree of education assortative mating, contributes to rising between-household income inequality over time. In this literature, as in our paper, a marriage market is defined for each year at the national level so that market and year are interchangeable. The main exercise then consists in creating, for each year, a sample of counterfactual marriage patterns that would have occurred if the degree of assortative mating had been held constant and equal to that observed in a reference year. The resulting differences in income inequality over time can no longer be attributed to differences in the degree of assortative mating, hence comparing the observed differences in income inequality to the counterfactual ones quantifies the contribution of changes in the degree of assortative mating over time.

This decomposition exercise relies crucially on the ability to *measure* the degree of assortative mating in any given year. The proportion of perfectly assortatively matched couples is not a good enough measure because it depends on the marginal distribution of education of men and women, which has changed over time. For this reason, the bulk of the literature[1] measures the degree of assortative mating as the ratio of the observed proportion of perfectly assortatively matched couples to the proportion that would have occurred under random matching.[2] However, as argued by Chiappori *et al*. (2020b), although this is the correct way of defining assortative mating, it is not a satisfying way of measuring it. It produces biased quantifications of the true degree of assortative mating as it confounds the distribution of men and women with marital preferences. This is easily seen using the following simple two-type two-market example. Consider two large marriage markets where men and women are in equal number and are equally distributed over two types: low or high education. Let p (respectively, q) be the share of high type men and women in the first (respectively, second) market. Suppose that marital preferences are the same in the two markets, and are such that like attracts like. Then even though marital preferences are constant across markets, the degree of assortative mating as computed in, for example, Eika *et al.* (2019), differs, with, for instance, for couples whose spouses are both of low type, a value of $1/(1-p)$ in the first market and a value of $1/(1-q)$ in the second.

For this reason, a recent stream of research uses structural models of the marriage market to measure the degree of assortative mating—for example, Chiappori *et al.* (2017, 2020b) and Ciscato and Weber (2020). There are two main advantages of using a structural model. First, as argued in Chiappori *et al.* (2020a,b), a structural model allows one to define a measure of the degree of assortative mating from the properties of the marital surplus and in particular, its supermodularity. Suppose that the marital surplus is denoted ϕ_{xy}, where x is the type of the man and y is the type of the woman (say, their education level). Chiappori *et al.* (2020b) propose the index $\phi_{xy}+\phi_{x'y'}-\phi_{x'y}-\phi_{xy'}$ as a local measure of the degree of assortative mating between men of types x and x', and women of types y and y'. Second, with this definition in mind, the structural model can be used to construct the counterfactual marriage patterns that would have occurred if the indices $\phi_{xy}+\phi_{x'y'}-\phi_{x'y}-\phi_{xy'}$ for all x, x', y, y' would have been equal to those of a reference year. The differences between the observed marriage patterns and the counterfactual ones are due merely to the differences in the degree of assortative mating between the year considered and the reference one.

One possibility to fix the degree of assortative mating equal to that of a reference year is to set the whole surplus itself equal to that of the reference year (see Chiappori *et al.* 2020a), and thus derive a first series of counterfactual marriage patterns. However, a disadvantage of this approach is that by changing the whole marital surplus, one changes not only the degree of assortative mating (i.e. the intensive margin) but also the extent to which people gain from marriage (i.e. the extensive margin). In other words, fixing the marital surplus constant over time is sufficient to fix the degree of assortative mating but not necessary. In fact, all marital surpluses of the form $\phi_{xy}+a_x+b_y$ share the same degree of assortative mating as ϕ_{xy} but have higher or lower gains from marriage depending on whether a_x+b_y is positive or negative. Hence the counterfactual marriage patterns derived using marital surplus $\phi_{xy}+a_x+b_y$ and ϕ_{xy} differ from one another even though they have been generated with the same degree of assortative mating, merely because the incentives to marry, and hence the marriage rates by types of men and women, differ.

Our main methodological contribution is to propose an algorithm to compute counterfactual marriage patterns obtained when the observed marital surplus is replaced by $\phi_{xy}+a_x+b_y$, where ϕ_{xy} is the marital surplus in the reference year, hence fixing the degree of assortative mating equal to that of the reference year, while solving for the parameters a_x and b_y so that the marriage rates by types of men and women are equal to the observed ones in the data. The differences in marriage patterns between the observed ones and those produced by this second series of counterfactuals are then due only to the differences in the degree of assortative mating (the intensive margin), whereas the difference between the two series of counterfactual marriage patterns reveals the extensive margin. Computing these two types of counterfactuals hence allows us to decompose the change in marriage patterns into an intensive and an extensive margin. Moreover, this algorithm can also be used to decompose the extensive margin into a size effect and a composition effect. Indeed, one can construct a third series of counterfactual marriage patterns obtained so that the observed marital surplus is still replaced by $\phi_{xy}+a_x+b_y$, where ϕ_{xy} is the marital surplus in the reference year, hence fixing the degree of assortative mating equal to that of the reference year, while solving for the parameters a_x and b_y so that the total mass of single men (respectively, women) is equal to that obtained in the first counterfactual series, but the distribution of singles by types is as observed in the data. As a result, the difference in marriage patterns obtained from the first and third counterfactuals can be due only to differences in relative marital gains across types (composition effect), whereas the difference in marriage patterns between the second and third counterfactuals can be due only to the absolute level of gains from marriage (size effect).

In this paper, we follow Chiappori *et al.* (2017, 2020a,b) and use the seminal structural matching market model due to Choo and Siow (2006) to derive our series of counterfactuals. In this model, men and women who differ along discrete observable characteristics x and y, respectively, called *types*, and unobservable idiosyncratic tastes, meet on a frictionless marriage market and aim at forming heterosexual pairs. The appeal of this model lies in its tractability:

1. The marital surplus ϕ_{xy} for a couple of type (x, y) is identified from the mass of couples of type (x, y), the mass of single men of type x and the mass of single women of type y.
2. Equilibrium is fully characterized by a set of non-linear equations (Chen *et al.* 2021; Galichon *et al.* 2019) that can be used to construct the equilibrium marriage patterns associated with any given counterfactual marital surplus ϕ_{xy} for all (x, y), while fully accounting for the distribution of types of men and women.

Following Chiappori *et al.* (2020a), we make use of point (ii) above to solve for counterfactual marriage patterns.

In our empirical contribution, we apply our approach and study marriage patterns and income inequality in the USA since the 1960s. We draw inspiration from recent important papers such as Eika *et al.* (2019) and Greenwood *et al.* (2014). In particular, we follow Eika *et al.* (2019) and use data for the USA from the Current Population Survey between 1962 and 2017. We assume that each year in the data, which is representative of the US population in that year, constitutes a separate marriage market. We compute counterfactuals for each year t in the data using 1962 as the reference year. We document the rise in between-household income inequality as measured by the Gini coefficient, where households include both singles and couples. We show that the associated Gini coefficient rose from 0.385 in 1962 to 0.495 in 2017. Had the marital surplus remained constant since 1962, our model would predict that the Gini coefficient would have increased from 0.385 to 0.465. Hence changes in marriage patterns (through the intensive and extensive margins) contribute to about a quarter of the observed increase in income inequality over the period considered. Using our three series of counterfactual marriage patterns, we show that this contribution is mainly due to the extensive margin (93%), the intensive margin explaining only 7%, the extensive margin itself being driven for three-fifths by a change in the total number of singles (size effect) and for two-fifths by a change in the distribution of types among the population of singles (composition effect).

Relation to the literature

Our model of the marriage market is a frictionless matching model with perfectly transferable utility. Early contributions to the transferable utility matching literature include Koopmans and Beckmann (1957), Shapley and Shubik (1971) and Becker (1973).

Recently, several papers extended this framework by adding unobserved heterogeneity in tastes—for example, Choo and Siow (2006), Dupuy and Galichon (2014), and Galichon and Salanié (2020). Our setting relies heavily on the Choo and Siow model, in particular its reformulation as a matching function equilibrium (Chen *et al.* 2021; Galichon *et al.* 2019).

We use the model to study educational assortative mating and income inequality between households. For a theoretical discussion of the connection between educational assortative mating and income inequality, see, for example, Fernández *et al.* (2005). There is a substantial literature in economics, sociology and demography documenting trends in educational assortative mating—for example, Gihleb and Lang (2016), Kalmijn (1991), Liu and Lu (2006), Mare (1991), and Schwartz and Mare (2005). The effect on income inequality is studied in Burtless (1999), Greenwood *et al.* (2014) and more recently Eika *et al.* (2019). These papers rely on a purely statistical approach. Ciscato and Weber (2020)—who make use of the continuous and multidimensional version of the Choo and Siow model developed by Dupuy and Galichon (2014)—and Pilossoph and Wee (2020) both rely on structural methods. However, in the former, singles are excluded from the analysis. Recently, Chiappori *et al.* (2020b) proposed an approach similar to ours.

Organization of the paper

The paper is organized as follows. Section II introduces briefly the Choo and Siow (2006) matching model, and presents the classes of counterfactuals used throughout the paper. Section III provides an application of our methodology. It presents the data and descriptive statistics, and discusses the decomposition of (trends in) income inequality. Section IV concludes. Additional computational details and results can be found in an Online Appendix.

I. Matching and Counterfactuals

The first subsection introduces succinctly the Choo and Siow (2006) model (see Online Appendix C for a more detailed account), from which the methodology to derive counterfactual matchings, presented in the second subsection, is inspired.

The Choo and Siow model

A marriage market à la Choo and Siow consists of men indexed by $i \in \mathfrak{I}$ and women indexed by $j \in \mathfrak{J}$ who meet on the market and may form heterosexual couples, holding the option to remain single. Men and women can be grouped into *types*, each type containing individuals with similar observable (to the analyst) characteristics. The set of types for men is denoted $\mathfrak{X}$, and the set of types for women is denoted $\mathfrak{Y}$. Man i is then said to be of observable type x_i and woman j of observable type y_j. The set of types of both men and women is extended to include the option of remaining single, that is, $\mathfrak{X}_0 = \mathfrak{X} \cup \{0\}$ and $\mathfrak{Y}_0 = \mathfrak{Y} \cup \{0\}$.

A matching μ is a vector $(\{\mu_{xy}\}_{xy \in \mathfrak{X} \times \mathfrak{Y}}, \{\mu_{x0}\}_{x \in \mathfrak{X}}, \{\mu_{0y}\}_{y \in \mathfrak{Y}})$, where μ_{xy} is the mass of marriages between type x men and type y women, μ_{x0} is the mass of single men of type x, and μ_{0y} is the mass of single women of type y. The total mass of men of type x (women of type y) is denoted n_x (m_y). A matching must satisfy the scarcity constraints

(1) $$n_x = \sum_y \mu_{xy} + \mu_{x0}, \quad m_y = \sum_x \mu_{xy} + \mu_{0y}.$$

Utility is perfectly transferable, and a couple formed of a man i of type x and a woman j of type y derives a joint utility

$$\phi_{ij} = \phi_{x_i y_j} + \varepsilon_{i y_j} + \eta_{x_i j},$$

where ε and η are interpreted as random tastes of man i for women of type y, and of woman j for men of type x, and assumed to follow (centred) Gumbel Type I distributions with unit scaling factors. The utilities of being single for a man i of type x and a woman j of type y are defined similarly as $\phi_{x0} + \varepsilon_{i0}$ and $\phi_{0y} + \eta_{0j}$, respectively.[3]

Under the above assumptions, the equilibrium matching μ can be expressed as an aggregate matching function of the form

(2) $$\mu_{xy} \equiv M_{xy}(\mu_{x0}, \mu_{0y}) = \sqrt{\mu_{x0}\mu_{0y}} \exp\left(\frac{\Phi_{xy}}{2}\right),$$

where $\Phi_{xy} = \phi_{xy} - \phi_{x0} - \phi_{0y}$ is the marital surplus of type (x, y) couples.

This equation can be inverted to yield

(3) $$\Phi_{xy} = \log\left(\frac{\mu_{xy}^2}{\mu_{x0}\mu_{0y}}\right).$$

Equation (3) links the marital surplus on the left-hand side to matching data on the right-hand side. The marital surplus of couples of type (x, y) is therefore non-parametrically point identified using data on a single cross-section of matches μ_{xy} and singles μ_{x0} and μ_{0y}.

Moreover, note that using equation (2) and system (1), one can compute the equilibrium matching for any given marital surplus. This feature is exploited in the next subsection to produce counterfactual marriage patterns.

Producing counterfactuals

Let $t = 0, \ldots, T$ index years (or 'marriage markets') in which we observe marriage patterns, and assume that for each year, data consist of observed marriage patterns

$$\hat{\mu}^t = \left(\{\hat{\mu}^t_{xy}\}_{xy \in \mathfrak{X} \times \mathfrak{Y}}, \{\hat{\mu}^t_{x0}\}_{x \in \mathfrak{X}}, \{\hat{\mu}^t_{0y}\}_{y \in \mathfrak{Y}}\right),$$

so that the population supplies $(\hat{n}^t_x, \hat{m}^t_y)$ are observed as well.

Note that for each year t, making use of the Choo and Siow (2006) framework, one can identify non-parametrically the surplus $\hat{\Phi}^t = (\{\hat{\Phi}^t_{xy}\}_{xy \in \mathfrak{X} \times \mathfrak{Y}})$ from the data $\hat{\mu}^t$ using equation (3).

We call year t=0 the reference year. Our aim is to construct three counterfactual equilibrium matchings (marriage patterns) for each year $t \geq 1$.

1. *Counterfactual 1* (Observed−(Intensive+Extensive)). With this counterfactual, we aim to fix the whole surplus constant over time and equal to that of the reference year. The resulting counterfactual marriage patterns may differ from the observed ones along an intensive margin (because the degree of assortative mating may be different between the observed market and the reference one) and along an extensive margin (because the gains from marriage may also be different). Hence the difference between the observed marriage patterns $\hat{\mu}^t$ and the counterfactual ones quantifies the sum of the intensive and intensive margins.
2. *Counterfactual 2* (Observed−Intensive). With this counterfactual, we aim to disentangle the intensive margin from the extensive margin. To do so, one replaces the marital surplus in year t by that of the reference year but solves for the (reservation) utilities of singles, adding $a_x + b_y$ to the surplus so that the equilibrium masses of singles coincide with those observed in year t, that is, $(\hat{\mu}^t_{x0}, \hat{\mu}^t_{0y})$. The difference between the observed marriage patterns $\hat{\mu}^t$ and those produced by this counterfactual is due merely to the difference in the degree of assortative mating in year t and the reference year, and hence quantifies the intensive margin.
3. *Counterfactual 3* (Observed−(Intensive+Size)). With this counterfactual, we aim to decompose the extensive margin into a size and a composition effect. When comparing the observed marriage patterns to the ones obtained under Counterfactual 1, there are changes in not only the total mass of singles (the size effect), but also the distribution of types among the singles (the composition effect, or 'who are the singles?' effect). To disentangle these two components, one produces a third counterfactual obtained by replacing the marital surplus in year t by that of the reference year but solving for the (reservation) utilities of singles, adding $a_x + b_y$, so that (i) the total mass of singles is equal to that in Counterfactual 1, and (ii) the distribution of types within the single population is the same as in the observed market. The difference between the marriage patterns under Counterfactual 1 and this third counterfactual is due merely to the difference in the distribution of types among singles, and hence quantifies the composition effect in the extensive margin. The difference between the marriage patterns under Counterfactual 2 and this third counterfactual is due merely to the difference in the total mass of singles, and hence quantifies the size effect in the extensive margin.

Below, we explain in greater details how each of these counterfactuals is designed, along with the algorithms that they require.

Counterfactual 1 (Observed−(Intensive+Extensive)). For any market $t \geq 1$, this counterfactual refers to the counterfactual marriage market in which we have the following.

1. We set the marital surplus equal to that of the reference year (i.e. we set the surplus equal to $\hat{\Phi}^0_{xy}$).
2. We set the population supplies equal to the observed ones (i.e. equal to $\hat{n}^t_x$ and $\hat{m}^t_y$).

This counterfactual (unique) equilibrium matching is computed using Algorithm 1 in Online Appendix A.

Counterfactual 1, together with the observed matching, allows us to isolate the sum of the intensive and extensive margins, since setting the surplus equal to that of the reference year induces individuals in year t to have marital patterns similar to those in the reference year.

Intuitively, Algorithm 1 in Online Appendix A makes use of equation (2) which, together with the constraints (1), yields

$$n_x = \sum_y \sqrt{\mu_{x0}\mu_{0y}} \exp\left(\frac{\Phi_{xy}}{2}\right) + \mu_{x0}, \quad m_y = \sum_x \sqrt{\mu_{x0}\mu_{0y}} \exp\left(\frac{\Phi_{xy}}{2}\right) + \mu_{0y}. \tag{4}$$

This system of equations fully characterizes equilibrium. It allows us to solve for the equilibrium mass of singles of each type and each gender, and hence the equilibrium marriage patterns given the distribution of men and women and the surplus function (see Algorithm 1 in Online Appendix A).

Counterfactual 2 (Observed−Intensive). For any market $t \geq 1$, this counterfactual refers to the counterfactual marriage market in which we have the following.

1. We set the marital surplus equal to that of the reference year, up to two additive terms (a_x, b_y) that are chosen so that the equilibrium masses of singles coincide with those observed in the market, that is, $(\hat{\mu}^t_{x0}, \hat{\mu}^t_{0y})$.
2. We set the population supplies equal to the observed ones (i.e. equal to $\hat{n}^t_x$ and $\hat{m}^t_y$).

The resulting counterfactual (unique) equilibrium matching is computed using Algorithm 2 in Online Appendix A.

Counterfactuals 1 and 2 allow us to disentangle the intensive margin from the extensive margin. The equilibrium matching under Counterfactual 2 preserves the degree of assortative mating for those who marry, but keeps the number of singles and the distribution of types in the single population equal to their observed levels in year t. As a result, the observed matching and the equilibrium matching under Counterfactual 2 differ only because of the intensive margin, whereas the equilibrium matching under Counterfactual 2 and the equilibrium matching under Counterfactual 1 differ only because of the extensive margin.

As to computing the terms (a_x, b_y), the intuition is as follows. Plugging the marital surplus $\hat{\Phi}^t_{xy} + a_x + b_y$ into equations (4), after simple algebra, one obtains

$$\sum_y \exp\left(\frac{\kappa_{xy} + a_x + b_y}{2}\right) = \dot{n}^t_x, \quad \sum_x \exp\left(\frac{\kappa_{xy} + a_x + b_y}{2}\right) = \dot{m}^t_y, \tag{5}$$

where $\kappa_{xy} = \hat{\Phi}^t_{xy} + \log \hat{\mu}^t_{x0}\hat{\mu}^t_{0y}$, and $\dot{n}^t_x = \hat{n}^t_x - \hat{\mu}^t_{x0}$ and $\dot{m}^t_y = \hat{m}^t_y - \hat{\mu}^t_{0y}$ are respectively the masses of married men of type x and women of type y imposed by the fixed distribution of singles and the distribution of types in counterfactual market t. Since the masses of single men and women of each type are known (as given by the target), the matching model degenerates to a matching model without singles. Indeed, the system of equations (5) is in fact the discrete version of the system that can be found in Dupuy and Galichon (2014). One can then solve this system for $\{a_x\}_{x \in \mathfrak{X}}$ and $\{b_y\}_{y \in \mathfrak{Y}}$, up to a normalization $a_{x_0} = 0$. The vectors $\{a_x\}_{x \in \mathfrak{X}}$ and $\{b_y\}_{y \in \mathfrak{Y}}$ can be thought of as the utilities of being single in market t.

Counterfactual 3 (Observed−(Intensive+Size)). For any market $t \geq 1$, this counterfactual refers to the counterfactual marriage market in which we have the following.

1. We set the marital surplus equal to that of the reference year, up to two additive terms (a_x, b_y) that are chosen so that the equilibrium masses of singles coincide with a specific target. This target is chosen such that (a) the total mass of singles is equal to that in Counterfactual 1, and (b) the distribution of types within the single population is the same as in the observed market.[4]
2. We set the population supplies equal to the observed ones (i.e. equal to $\hat{n}_x^t$ and $\hat{m}_y^t$).

The resulting counterfactual (unique) equilibrium matching is computed using Algorithm 2 in Online Appendix A.

Counterfactual 3 therefore shares the same surplus and same distribution of types in the population of singles as Counterfactual 2, but has different targeted masses of singles. As a result, the equilibrium matching under Counterfactual 2 and the equilibrium matching under Counterfactual 3 differ merely because of their different masses of singles (i.e. the size effect), whereas the equilibrium matching under Counterfactual 3 and the equilibrium matching under Counterfactual 1 differ merely because of the distribution of types in the population of singles (i.e. the composition effect).

II. Application

In this section, we use our framework to study the US marriage market from 1962 to 2017, and quantify the contribution of different margins of changes in marital patterns to rising income inequality.

Data

We analyse the US marriage market between 1962 and 2017 using the Current Population Survey (CPS) March supplements data. We assume that each year constitutes a separate marriage market. For each year, we select heterosexual couples in which the husband is aged between 26 and 60 years old, and the wife is aged between 24 and 58 years old.[5] Thus there is some overlap across marriage markets.[6] However, in this application, our focus is principally on the long-term comparison between 1962 and 2017, in which case there is no overlap.

In the baseline sample, our definition of couples includes legally married and cohabiting individuals. Note, however, that prior to 1995, the data do not allow us to identify those individuals who are in a cohabiting relationship, and they are therefore counted as singles. This creates a slight discrepancy in 1995. We argue that this is not an issue, since our results depend mainly on our marital surplus estimates from 1962, when cohabiting couples were rare.[7]

In addition to couples, we select single, never-married[8] men aged 26–60, and single, never-married women aged 24–58. We exclude all couples and singles with missing information on age, income (defined as labour income from wages, salary and self-employment) and education. In our application, we define four types corresponding to four education levels: below high school degree, high school degree, some college, and college graduates (and above).[9]

Descriptive statistics on couples and singles are reported in Table 1 for selected years.[10] The table depicts the dramatic decrease in the share of men and women with less than a high school degree, and the increase in the share of men and women with some college education or a college degree. These are major structural changes that are fully taken into account in our model. The table indicates a growing income gap between singles and couples. Since the proportion of singles has steadily increased since 1962, we expect the size effect of the extensive margin to play an important role in rising income inequality. Table 1 also suggests that low-education types were under-represented among singles in 1962 but are now over-represented, while the opposite is true for the highly educated. Thus we also expect the composition effect of the extensive margin to play some role in our results.

The time period under scrutiny has also seen dramatic changes in fertility, women's labour market participation and divorce laws, to name only a few (Stevenson and Wolfers 2007). Each of these changes likely had an impact on the US marriage market, and this will be reflected in our estimates of the marital surplus. Our model does not allow us to identify these impacts separately, but this is not relevant for the purpose of our application.

TABLE 1
DESCRIPTIVE STATISTICS

	1962		1990		2017	
	Women	Men	Women	Men	Women	Men
A. Couples						
Age	39.53	42.56	38.93	41.40	41.10	43.16
Below high school	42.27	49.21	13.68	15.39	7.22	8.45
High school degree	40.56	28.65	43.78	36.64	23.95	28.85
Some college	10.02	9.37	20.77	19.87	27.25	25.84
College degree	7.15	12.77	21.77	28.10	41.58	36.86
Income (in $1000)	6.90	46.06	22.12	59.13	34.12	68.04
Household income (in $1000)	52.96	52.96	81.25	81.25	102.16	102.16
N	11,593	11,593	23,270	23,270	26,651	26,651
B. Singles						
Age	38.32	38.75	32.36	34.29	35.38	37.46
Below high school	35.19	47.43	16.00	17.35	8.45	10.67
High school degree	38.11	25.59	32.57	35.41	28.56	36.04
Some college	9.81	10.76	20.73	19.18	29.53	25.13
College degree	16.89	16.22	30.70	28.05	33.46	28.15
Income (in $1000)	22.29	28.42	29.11	36.94	27.96	37.04
Household income (in $1000)	22.29	28.42	29.11	36.94	27.96	37.04
N	1131	1227	5426	5532	8192	7413
C. Difference (Couples−Singles)						
Age	1.20	3.81	6.34	6.83	5.72	5.69
Below high school	7.08	1.78	3.22	2.73	−1.23	−2.22
High school degree	2.45	3.06	13.81	7.07	−4.61	−7.20
Some college	0.21	−1.39	−1.81	−3.33	−2.28	0.71
College degree	−9.74	−3.44	−15.23	−6.47	8.12	8.70
Income (in $1000)	−15.39	17.64	−12.60	23.50	6.16	30.99
Household income (in $1000)	30.67	24.54	46.75	38.55	74.20	65.11

Notes: The table provides descriptive statistics for some selected years (1962, 1990 and 2017). Our sample include single (never-married) and married men aged 26–60, and single (never-married) and married women aged 24–58. Income figures are expressed in 2014 dollars.

Figure 1 indicates that marriage rates have changed dramatically over the period 1962–2017. While marriage rates were roughly similar across types of men at the beginning of the period and have generally declined, the decline is more pronounced for some types. In recent years, men with college education are more likely to be married than men with a high school degree or less. The picture is similar for women, except that highly educated women married significantly less at the beginning of the period. These changes are fully accounted for in our decomposition exercise that accounts explicitly for singles.

Counterfactual analysis

We now investigate to what extent changes in marital patterns contributed to changes in income inequality in marriage market t compared to the reference year of 1962. To do so, for each market t and each of the three counterfactual matchings considered in the second subsection of Section II, we generate the associated sample of households following the methodology described in Subsection A.3 of Online Appendix A. These counterfactual samples are then used to compute (counterfactual) measures of inequality (e.g. the Gini coefficient).[11] For any market t, we compute the measure of inequality $\hat{g}^t$ obtained from using the observed sample (i.e. matching $\hat{\mu}^t$), as well as from using the samples under Counterfactuals 1, 2 and 3, respectively (we denote these counterfactual measures g_1^t, g_2^t and g_3^t, respectively).

By comparing these inequality measures, we are able to identify several channels through which changing marital patterns between year t and 1962 contribute to changing income inequality. Recall that the difference between $\hat{g}^t$ and g_1^t captures the total contribution of marital patterns (the sum of the intensive and extensive margin effects) to

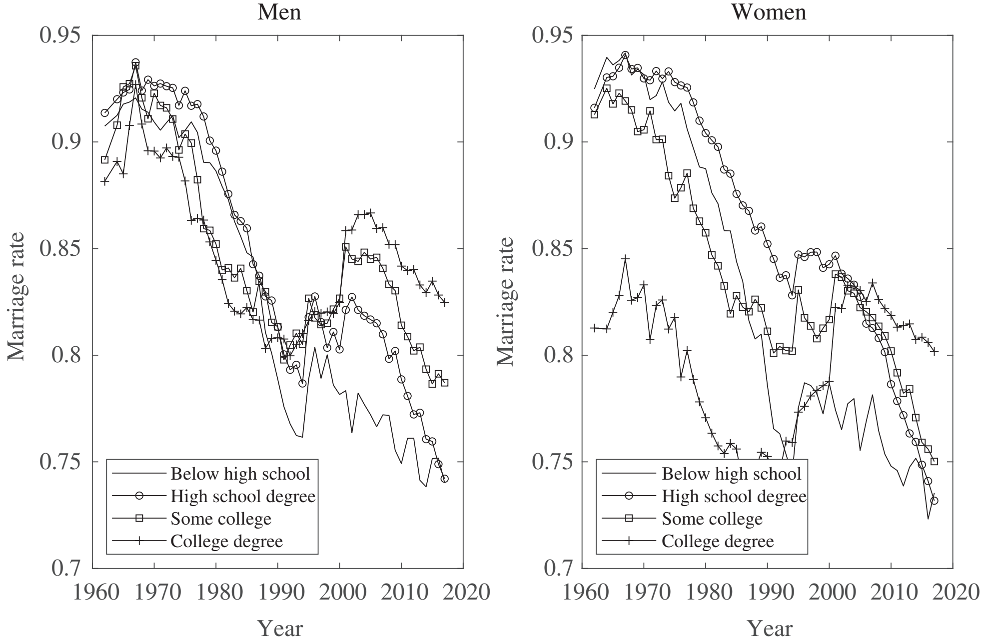

FIGURE 1. Marriage rates by types of men and women, 1962–2017.

changing income inequality between market t and the reference year 1962. The difference between $\hat{g}^t$ and g_2^t captures the effect of the intensive margin, whereas the difference between g_2^t and g_1^t captures the extensive margin. Finally, the difference between g_2^t and g_3^t captures the size effect at the extensive margin, and the difference between g_3^t and g_1^t captures the composition effect.

The main results are displayed in Figure 2. It shows that over the period 1962–2017, the actual Gini coefficient increased by 0.11 points, from about 0.385 to 0.495. Had the surplus remained constant since 1962 (Counterfactual 1), our model would predict that the Gini coefficient would have increased from 0.385 to 0.465. Under this counterfactual, the evolution of the Gini coefficient is driven only by factors external to the marriage market, except for changes in the distribution of types, which are assumed exogenous to our model. We conclude that changes in marital patterns (through the intensive and extensive margins) contribute to an increase of 0.03 points in the Gini coefficient, which represents about a quarter of the observed increase in income inequality over the period considered.

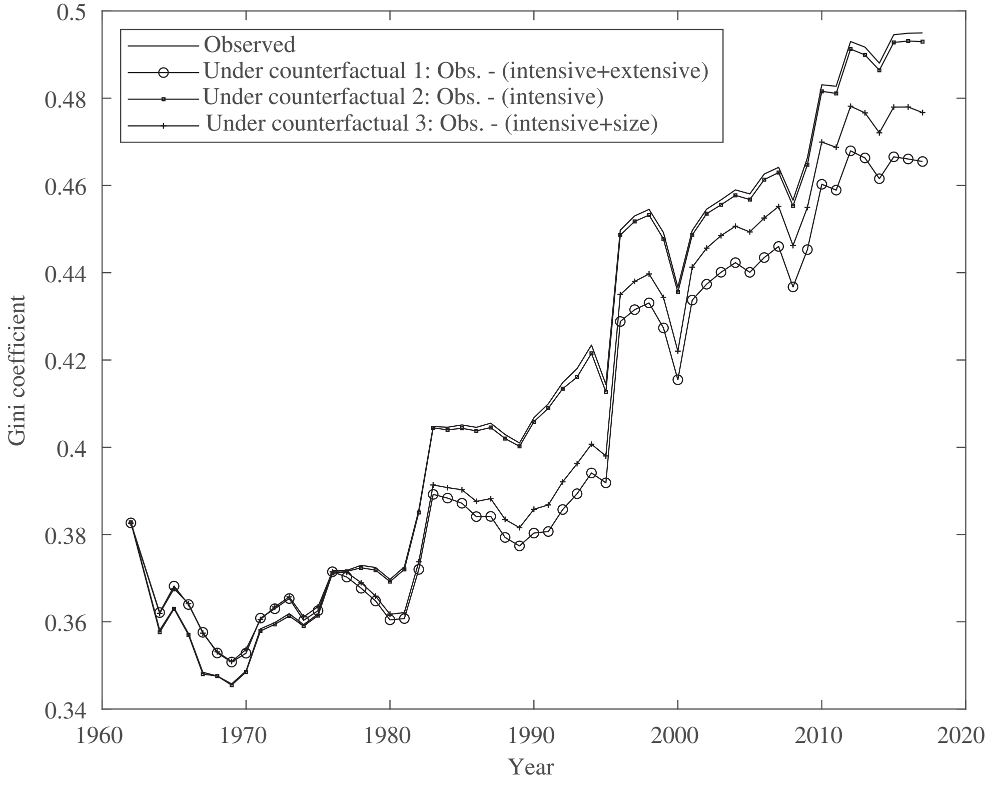

FIGURE 2. Trend in income inequality, 1962–2017.
Notes: In this figure, we plot measures of between-household income inequality (Gini coefficients) between 1962 and 2017 obtained using the observed sample of households (labelled 'Observed'), using the sample under Counterfactual 1 (which identifies the sum of the intensive and extensive margin, and is labelled 'Observed−(Intensive+Extensive)'), using the sample under Counterfactual 2 (which identifies the intensive margin, and is labelled 'Observed−Intensive'), and using the sample under Counterfactual 3 (which identifies the size effect at the extensive margin (plus the intensive margin), and is labelled 'Observed−(Intensive+Size)').

Figure 3 plots the evolution of the overall contribution of marital patterns in wage inequality as well as its decomposition into the intensive and extensive margins. The figure clearly shows that out of the total contribution of marital patterns of 0.03 points, the intensive margin accounts for 0.002 points, and the extensive margin accounts for 0.028 points. The extensive margin itself can be further decomposed (using Counterfactual 3) into a size effect that accounts for 0.016 points and a composition effect that accounts for 0.012 points.[12] The contribution of marital patterns to the rise in income inequality is therefore due mainly to the extensive margin, at 93%, the intensive margin explaining only 7%.[13] The extensive margin itself is driven for three-fifths by a change in the total number of singles (size effect), and for two-fifths by a change in the distribution of types among the population of singles (composition effect).

Our analysis shows three important results. First, the intensive margin has had a modest impact on the increase in between-household inequality. This is very much in line with Eika *et al.* (2019) and Chiappori *et al.* (2020a). This is because overall, assortative mating in education has not increased by much. Other factors, such as changes in the returns to education, have likely played a much bigger role in rising household income inequality. Second, our results clearly indicate the overwhelming importance of the extensive margin. Indeed, in 2017, our sample of households contains 37% of singles, but this proportion would fall to 27.5% had the surplus remained constant since 1962 as in Counterfactual 1 (see Figure A2 in Online Appendix D for yearly estimates). Third, the extensive margin is driven by both the size and composition effects. In 2017, about 48% of single men and 38% of single women have a high-school degree or less, but had the

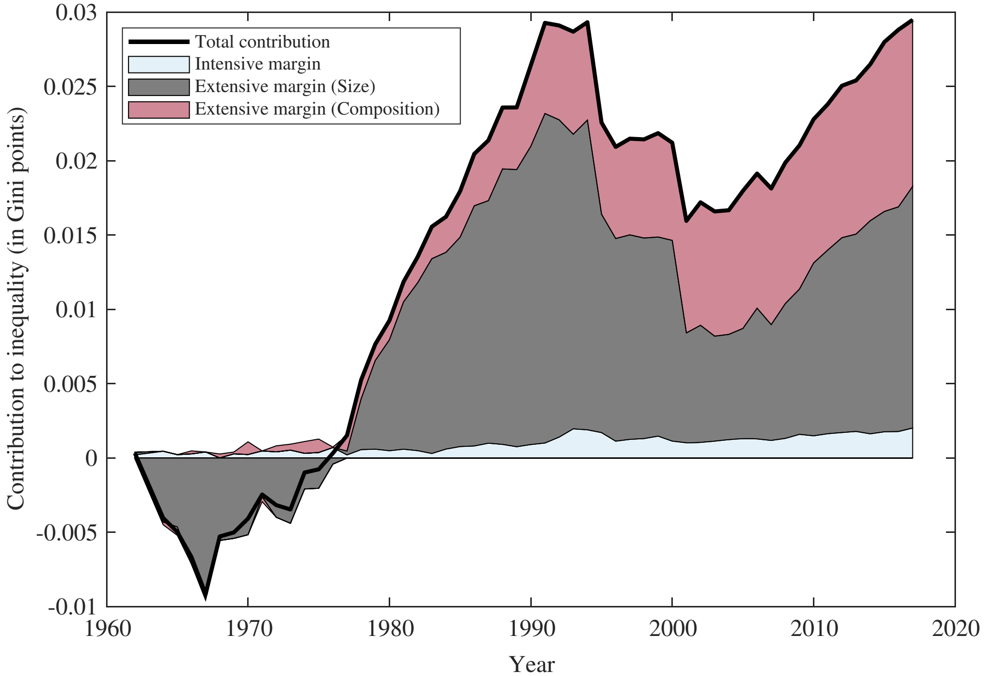

FIGURE 3. Trend in income inequality, 1962–2017.
Notes: In this figure, we plot the contribution of the intensive margin and the extensive margin (the size and composition effects) to income inequality, as measured by the Gini coefficient, between 1962 and 2017.

surplus not changed since 1962, these proportions would fall to 38% and 8%, respectively (see Figure A3 in Online Appendix D for yearly estimates). The decrease in the number of singles and the change in the distribution of types among singles under the counterfactual matching are explained by the dramatic drop in the marital surplus since 1962. As a result, more individuals of lower education are pushed into marriage in the counterfactual matching (which holds the marital surplus constant to its 1962 level) compared to the observed matching in the data.

III. CONCLUSION

In this paper, we use a structural matching model (à la Choo and Siow 2006) to construct marriage market counterfactuals in order to quantify the contribution of changes in marital patterns in rising income inequality. Our main methodological contribution is to propose an algorithm to compute counterfactual marriage patterns obtained when the degree of assortative mating is fixed to some reference level but the marital gains are so that the marriage rates by types of men and women are equal to the observed ones in the data. We argue that our approach is advantageous because singles are fully integrated in the model, and counterfactuals have a clear interpretation and can be used to decompose differences in marital patterns along an extensive margin and an intensive margin.

Using our counterfactuals, we study the effect of marital patterns on the evolution of between-household income inequality in the USA between 1962 and 2017. We find that had the marital surplus remained constant over time, between-household income inequality would have been lower in 2017 by almost 0.03 Gini points. Marital patterns can therefore explain about a quarter of the observed increase in income inequality over the period considered. Interestingly, our decomposition exercise shows that most of this effect comes from the extensive margin (93%), and that the intensive margin (assortative mating in education) contributed only very little (7%). Moreover, our results also show that the extensive margin itself is driven for three-fifths by a change in the total number of singles and for two-fifths by a change in the distribution of types among the population of singles. We thus conclude that while, as one might have expected, the sharp increase in singlehood has had a significant contribution in the increase in income inequality, the changes in the educational composition of singles has had an almost as important role. Uncovering the relative magnitude of these effects is the main contribution of our methodology.

ACKNOWLEDGMENTS

The authors thank the Editor and three anonymous referees for their comments, as well as Jim Heckman and Marco Francesconi for insightful discussions, and seminar participants at the University of Luxembourg for comments.

Dupuy acknowledges funding from the FNR under the CORE grant no. C14/SC/8337045. Weber is grateful for the hospitality of the University of Chicago, BFI, CEHD and KU Leuven, which hosted him while most of this paper was written.

NOTES

1. There are other ways to proceed, of course. There is a large literature in economics, demography and sociology aiming at measuring assortative mating on the marriage market (in education) using a battery of models and statistical techniques. See, for example, Schwartz and Mare (2005) in sociology or Liu and Lu (2006) in economics.

2. The joint distribution of couples under random matching is obtained as the product of the marginal distributions of married men and women by education.
3. In a single market, the systematic utility of being single can be normalized to 0 without loss of generality, interpreting ϕ_{xy} as the marital surplus. However, this normalization is not justified when comparing across markets unless the utility of being single does not vary across markets.
4. Formally, we denote this target as $(\tilde{\mu}^t_{x0}, \tilde{\mu}^t_{0y})$, which is such that $\sum_x \tilde{\mu}^t_{x0} = \sum_x \mu^t_{x0}(\bar{\Phi})$ and $\tilde{\mu}^t_{x0}/\left(\sum_x \tilde{\mu}^t_{x0}\right) = \hat{\mu}^t_{x0}/\left(\sum_x \hat{\mu}^t_{x0}\right)$. The same constraints are imposed on the women's side of the market. Subsection B.1 of Online Appendix B provides details on how this can be done.
5. This age selection differs from that in Eika *et al.* (2019), which selects all couples in which at least one partner is aged between 26 and 60. Our approach is more consistent with the rest of the matching literature; see, for example, Chiappori *et al.* (2017).
6. Ideally, one would like to use data on newlyweds only to estimate the marital surplus. However, to the best of our knowledge, there are no datasets that (i) consistently provide information on the date of marriage, (ii) have sufficiently large sample size, and (iii) go back several decades.
7. Fitch *et al.* (2005) estimate that at most 2% of all couples households included unmarried couples in 1960. We checked our results using an alternative sample in which those who cohabit are always counted as singles (both before and after 1995). Our results are left unchanged, and are available on request.
8. Since divorce and remarriage are not the focus of this paper, we exclude divorced and separated (as well as widowed) individuals, as in Chiappori *et al.* (2017). We thus implicitly assume that divorced and separated individuals match on a different market. If this assumption were violated, then our estimates of the marital surplus would be upward biased. However, note that this should not affect the estimates of the degree of assortativeness.
9. Following the bulk of the literature assessing the role of matching in rising inequality, types of couples are defined using spouses' education only. There are three main reasons for that. First, education is the single most important factor on which sorting in the marriage market occurs, apart from age, and educational attainment has increased dramatically over time, especially for women. Second, while the CPS has the advantage of providing data over a long period of time, it does not (systematically) provide information on spouses' relevant variables such as height, weight, subjective health or personality traits on which sorting occurs. Third, creating richer types for both men and women increases the likelihood of observing types of couples with low frequencies, the so-called 'thin cell' problem; using our definition of types, the cell with the lowest frequency in our data has about 50 couples in each single year.
10. From Table 1, it is apparent that our sample differs only slightly from that of Eika *et al.* (2019) because of the differences in the age restriction that we imposed.
11. Note that in this application, we focus only on between-household income inequality. To study income inequality at the individual level, one must first estimate consumption economies of scale and sharing rules (how resources are divided among members of a household). This is beyond the scope of this application. See Lise and Seitz (2011).
12. The intensive and extensive margin effects, as well as the total effect of marital patterns on inequality, are displayed with confidence intervals in Figure A1 of Online Appendix D.
13. These results are in line with those of Eika *et al.* (2019) when singles are included; see panel e of Figure A12 in their paper.

REFERENCES

BECKER, G. S. (1973). A theory of marriage: Part I. *Journal of Political Economy*, **81**(4), 813–46.

BURTLESS, G. (1999). Effects of growing wage disparities and changing family composition on the U.S. income distribution. *European Economic Review*, **43**(4–6), 853–65.

CHEN, L., CHOO, E. S. Y., GALICHON, A. and WEBER, S. (2021). Matching function equilibria with partial assignment: existence, uniqueness and estimation. Working Paper.

CHIAPPORI, P.-A., COSTA-DIAS, M., CROSSMAN, S. and MEGHIR, C. (2020a). Changes in assortative matching and inequality in income: evidence for the UK. *Fiscal Studies*, **41**(1), 39–63.

CHIAPPORI, P.-A., COSTA-DIAS, M. and MEGHIR, C. (2020b). Changes in assortative matching: theory and evidence for the US. NBER Technical Report.

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

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

CISCATO, E. and WEBER, S. (2020). The role of evolving marital preferences in growing income inequality. *Journal of Population Economics*, **33**(1), 307–47.

DUPUY, A. and GALICHON, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, **122**(6), 1271–319.

EIKA, L., MOGSTAD, M. and ZAFAR, B. (2019). Educational assortative mating and household income inequality. *Journal of Political Economy*, **127**(6), 2795–835.

Fernández, R., Guner, N. and Knowles, J. (2005). Love and money: a theoretical and empirical analysis of household sorting and inequality. *Quarterly Journal of Economics*, **120**(1), 273–344.
Fitch, C., Goeken, R. and Ruggles, S. (2005). The rise of cohabitation in the United States: new historical estimates. Minnesota Population Center Working Paper no. 3.
Galichon, A., Kominers, S. D. and Weber, S. (2019). Costly concessions: an empirical framework for matching with imperfectly transferable utility. *Journal of Political Economy*, **127**(6), 2875–925.
Galichon, A. and Salanié, B. (2020). Cupid's invisible hand: social surplus and identification in matching models. *Review of Economic Studies*, forthcoming; available online at https://ssrn.com/abstract = 1804623 (accessed 14 July 2021).
Gihleb, R. and Lang, K. (2016). Educational homogamy and assortative mating have not increased. NBER Working Paper no. 22927; available online at https://www.nber.org/papers/w22927 (accessed 26 June 2021).
Greenwood, J., Guner, N., Kocharkov, G. and Santos, C. (2014). Marry your like: assortative mating and income inequality. *American Economic Review*, **104**(5), 348–53.
Kalmijn, M. (1991). Status homogamy in the United States. *American Journal of Sociology*, **97**(2), 496–523.
Koopmans, T. and Beckmann, M. (1957). Assignment problems and the location of economic activities. *Econometrica*, **25**(1), 53–76.
Lise, J. and Seitz, S. (2011). Consumption inequality and intra-household allocations. *Review of Economic Studies*, **78**(1), 328–55.
Liu, H. and Lu, J. (2006). Measuring the degree of assortative mating. *Economics Letters*, **92**(3), 317–22.
Mare, R. D. (1991). Five decades of educational assortative mating. *American Sociological Review*, **56**(1), 15–32.
Piketty, T. and Saez, E. (2003). Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, **118**(1), 1–39.
Pilossoph, L. and Wee, S. L. (2020). Assortative matching and household income inequality: a structural approach. Working Paper.
Schwartz, C. R. (2010). Earnings inequality and the changing association between spouses' earnings. *American Journal of Sociology*, **115**(5), 1524–57.
Schwartz, C. R. and Mare, R. D. (2005). Trends in educational assortative marriage from 1940 to 2003. *Demography*, **42**(4), 621–46.
Shapley, L. S. and Shubik, M. (1971). The assignment game I: The core. *International Journal of Game Theory*, **1**(1), 111–30.
Stevenson, B. and Wolfers, J. (2007). Marriage and divorce: changes and their driving forces. *Journal of Economic Perspectives*, **21**(2), 27–52.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

A Algorithms
B Further remarks
C The matching model: a detailed account
D Additional results