Axiomatic Measures of Assortative Matching

Hanzhe Zhang*

January 29, 2024†

Abstract

An active debate concerns the direction of change in assortative matching on education in the US, because different measures yield different conclusions. To identify appropriate measures of assortative matching, I adopt an axiomatic approach: Start with the properties a measure should satisfy and identify the measures that satisfy them. I find that normalized trace (the proportion of pairs of like types) and the aggregate likelihood ratio (Eika, Mogstad and Zafar, 2019) satisfy my basic equivalence and monotonicity axioms and are uniquely characterized by axioms with cardinal interpretations. They also naturally extend to markets with singles and one-sided markets. The relation induced by the odds ratio is the unique total order on two-type markets that satisfies marginal independence, but it yields ordinal interpretation only and does not have a multi-type extension (Chiappori, Costa-Dias, and Meghir, 2024). For multi-type markets, I show the impossibility result whereby there is no total order that satisfies monotonicity and the additional requirement of robustness to categorization (i.e., the assortativity order holds regardless of the categorization of types). I apply the measures to shed light on the evolution of educational assortativity in the US and other countries.

Keywords: assortative matching, sorting, axiomatic approach, trace

JEL: C78, J01

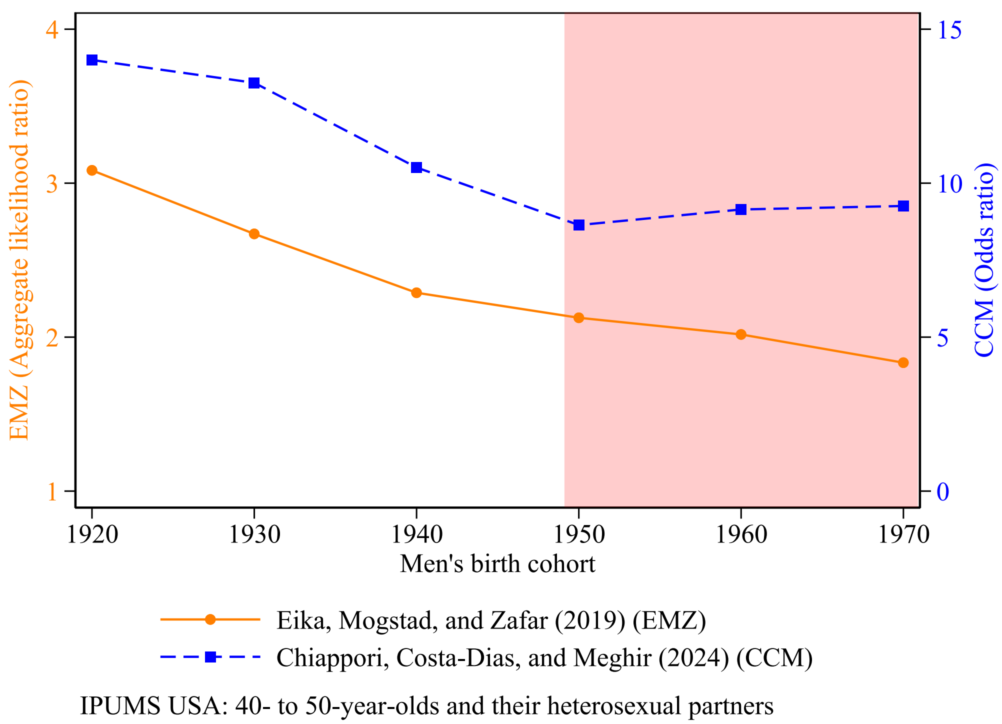
*Department of Economics, Michigan State University; hanzhe@msu.edu.

†I thank Hector Chade, Willy Chen, Pierre-André Chiappori, James Heckman, Juanna Joensen, Maciej Kotowski, Jingfeng Lu, Antonio Nicolò, Edward Schlee, Mark Whitemeyer, Basit Zafar, Mu Zhang, Jin Zhou, and seminars audiences at Michigan State, Arizona State, MEA/SOLE, Chicago, Michigan, and SEA. I acknowledge the National Science Foundation and Michigan State University Delia Koo Endowment Fund for financial support.

1 Introduction

There is an active debate about how assortative matching on education has evolved in the United States over the past few decades. Figure 1 shows that based on different measures, it can be concluded to have decreased (Eika, Mogstad, and Zafar, 2019, henceforth EMZ) or increased (Chiappori, Costa-Dias, and Meghir, 2024, henceforth CCM). All of the measures used are defensible by a list of nice properties they possess, but different measures yield different conclusions. Drawing a definitive conclusion is especially difficult, given that educational composition differs across time and location.

Figure 1: Assortative matching on college education in the US



Motivated by this specific empirical debate on assortative matching on education, I address a general theoretical question that applies to assortative matching on any observable individual characteristics: How do we rank and/or quantify the assortativity of matching markets with potentially different distributions of educational and socioeconomic characteristics?

Rather than starting with specific measures and justifying them, I take an *axiomatic approach*: To discover the appropriate measure(s) of assortative matching, I start with a set of the properties measures should satisfy. The axiomatic approach has found extensive application in economics: voting, bargaining, utility representation, and market design, to name a few (Arrow, 1950; Nash, 1950; von Neumann and Morgenstern, 1953; Sonmez, 2023). However, to the best of my knowledge, it has not been employed in the examination of assortative matching.

Technically, I start with two sets of basic axioms: (i) axioms of *equivalence* (to specify when two markets are equally assortative) and (ii) axioms of *monotonicity* (to specify when one market is more assortative than another). Equivalence axioms are *scale invariance* (invariance to scaling the market size), *side symmetry* (invariance to swapping sides), and *type symmetry* (invariance to swapping types). Two sets of notions of monotonicity I consider are (i) *diagonal* and *off-diagonal monotonicity* (assortativity increases when more like types match and decreases when more unlike types match) and (ii) *marginal monotonicity* (defining assortativity order on markets with the same marginal distributions).

These basic axioms will help eliminate some measures but they will not uniquely pin down a measure. Besides the basic axioms, I state more crucial axioms that will characterize different measures (*characterization axioms*). First, I define *decomposability* axioms: The assortativity measure for any market is a weighted average or sum of the measures for submarkets decomposed from the market. The *normalized trace*—the proportion of pairs of like types—is the unique measure (up to linear transformation) that satisfies the basic axioms plus *population decomposability*, i.e, the weight to average is population size (Theorem 1). The *aggregate likelihood ratio* (EMZ)—the excess likelihood of like types pairing up relative to the hypothetical likelihood under random matching—is the unique measure (up to multiplication) that satisfies the basic axioms plus *random decomposability*, i.e., the weight to sum is the hypothetical proportion of pairs of like types under random matching (Theorem 2). In addition, these measures have natural extensions to two-sided markets with singles, one-sided markets (e.g., homosexual matching markets), and multi-type markets (Propositions 1, 2, and 3, respectively).

In two-type markets, the *odds ratio* (CCM)—the ratio of the odds to match with the same type and with a different type—is the unique measure (up to monotonic transformation) that satisfies the basic axioms plus *marginal independence* (Edwards, 1963) (Theorem 3). Because all measures that are monotonic transformations of the odds ratio (e.g., Yule's Q, Yule's Y, and log odds ratio) provide the same ordering of markets in terms of assortativity, the result clarifies that the odds ratio carries ordinal meaning only. In addition, the odds ratio does not have a natural extension to multi-type markets and should be treated as a local measure of assortativity.

In general, in markets with more than two types (e.g., there are 14 detailed education categories in the US Census), *robustness to categorization* is a desired property: Assortativity comparison should be robust to how education levels are categorized into two or several groups. I show that there exists no total order that satisfies the minimal set of axioms of monotonicity and robustness to categorization (Theorem 4). As a result, we must resort to partial orders.

Using the axiomatized measures, I document the evolution and magnitude of assortativity on education in the US. I find that in recent decades, assortative matching on college has decreased under the aggregate likelihood ratio (EMZ) and normalized trace, and the decrease was more prominent when considering singles as part of the market. In contrast, it has increased under the odds ratio (CCM). The type-specific likelihood ratio based on college-college pairs and that based on noncollege-noncollege pairs (EMZ) provide conflicting conclusions and should be used with caution. The empirical analyses also motivate more careful consideration of what constitutes a marriage market.

The rest of the paper is organized as follows. Section 2 presents the binary-type setup, gives examples of measures that have been used in the literature, and describes basic axioms. Section 3 presents characterization results for binary-type markets and discusses extensions. Section 4 presents the results for multi-type markets. Section 5 documents assortative matching on education in the US and other countries, and Section 6 concludes.

2 Binary types

Each man and woman possesses a trait (e.g., college education or any other phenotype such as a psychological, biological, socioeconomic trait) and can be one of two types, high θ_1 or low θ_2. Consider the matching between men and women described by matrix $M = (a, b, c, d)$:

$$M = \begin{array}{ccc} m\backslash w & \theta_1 & \theta_2 \\ \theta_1 & a & b \\ \theta_2 & c & d \end{array}.$$

A cell describes the mass of pairs between a specific combination of types of men and women. To recover the mass of (matched) individuals of a specific gender and type, we sum a column or a row. For example, there is mass $a + b$ of high-type men. The marginal distribution for men is then described by $(a + b, c + d)$ and that for women is $(a + c, b + d)$. Assume a full support of types on both sides of the market: $(a + b)(a + c)(b + d)(c + d) > 0$. I refer to M as the matching, matrix, or market, and let $|M| \equiv a + b + c + d$ denote its population size.

A few matching arrangements will appear repeatedly. A *strictly positive assortative matching* has $b = c = 0$ and should be judged the most assortative by any reasonable measure. Any matching with $bc = 0$ should also be considered *positive assortative matching*

(PAM), because the maximal feasible mass of pairs of like types has been reached. Analogously, a *strictly negative assortative matching* has $a = d = 0$ and should be the least assortative by any reasonable measure; a matching with $ad = 0$ should be considered *negative assortative matching (NAM)*, because the maximal feasible mass of pairs of unlike types has been reached. In the main text, I will treat, as a consequence of the axioms below, all matchings with $bc = 0$ and $b \neq c$ (resp., $ad = 0$ and $a \neq d$) as the most (resp., least) assortative; in the appendix, I will consider the implications when they are not considered the most (resp., least) assortative.

Random matching is reached when all individuals are uniformly randomly matched. For market $M = (a, b, c, d)$, the hypothetical random matching is

$$R(M) \equiv \frac{\frac{a+b}{|M|}\frac{a+c}{|M|}|M| \;\Big|\; \frac{a+b}{|M|}\frac{b+d}{|M|}|M|}{\frac{a+c}{|M|}\frac{c+d}{|M|}|M| \;\Big|\; \frac{c+d}{|M|}\frac{b+d}{|M|}|M|}.$$

Objective. Most ideally, we would like to find an *index*, i.e., a function $I : \mathbb{R}^4_+ \to \mathbb{R} \cup \{-\infty, +\infty\}$, that measures the matching assortativity. Alternatively, we resort to finding a *total order* $\succeq$ on the set of markets, $\mathbb{R}^4_+$. At least we would like to find a *partial order* on the set of markets. Total and partial orders have ordinal meanings only. An index I has an ordinal interpretation, because real numbers are endowed with a natural order. In addition, an index may also have a cardinal interpretation.

2.1 Examples of measures

Existing measures of assortativity can be roughly divided into three categories: (i) measures that relate to the proportion of pairs of like types, (ii) measures that are compared with the hypothetical benchmark of random matching, and (iii) measures that induce the same order as the odds ratio. I will also mention additional measures and approaches in the literature. Readers may skip ahead to the axioms; these measures provide some concrete ideas before introducing the abstract axioms.

2.1.1 Proportion of pairs of like types

Normalized trace equals 1 if the matching is positive assortative, equals 0 if the matching is negative assortative, and equals the proportion of like types in the market when it is neither positive assortative nor negative assortative.

(NT) Normalized trace

$$I_{tr}(a,b,c,d) = \begin{cases} 1 & \text{if } bc = 0 \\ \frac{a+d}{a+b+c+d} \in (0,1) & \text{if } abcd \neq 0 \\ 0 & \text{if } ad = 0 \end{cases}.$$

2.1.2 Normalization based on random matching

(LR) Type-specific likelihood ratio (EMZ)[1]

$$I_{L1}(M) \equiv \frac{\text{observed } \#\theta_1\theta_1}{\text{random baseline}} = \frac{a}{|M|} \Big/ \frac{a+b}{|M|}\frac{a+c}{|M|} = \frac{a(a+b+c+d)}{(a+b)(a+c)}.$$

The likelihood ratio based on θ_1 compares the realized likelihood of high-type pairs with the benchmark of the hypothetical likelihood of high-type pairs under random matching. The likelihood ratio based on θ_2 is analogously defined:

$$I_{L2}(M) \equiv \frac{\text{observed } \#\theta_2\theta_2}{\text{random baseline}} = \frac{d}{|M|} \Big/ \frac{d+b}{|M|}\frac{d+c}{|M|} = \frac{d(a+b+c+d)}{(d+b)(d+c)}.$$

One issue with the type-dependent likelihood ratio is that it requires the choice of a type as a benchmark; in other words, it fails *type symmetry*, which I will define later on. Empirically, we will see that the likelihood ratio based on different types will yield opposite conclusions on the direction of change in assortative matching.[2]

The aggregate likelihood ratio is a weighted average of the type-specific likelihood ratios, in which the weight on each type-specific likelihood ratio is the expected mass of pairs of like types under random matching.

(ALR) Aggregate likelihood ratio (EMZ)

$$\begin{aligned} I_L(M) &\equiv \frac{(a+b)(a+c)}{(a+b)(a+c)+(d+b)(d+c)} I_{L1}(M) + \frac{(d+b)(d+c)}{(a+b)(a+c)+(d+b)(d+c)} I_{L2}(M) \\ &= \frac{\text{observed } \#\theta_1\theta_1 + \#\theta_2\theta_2}{\text{random baseline}} = \frac{a+d}{|M|} \Big/ \left(\frac{a+b}{|M|}\frac{a+c}{|M|} + \frac{d+b}{|M|}\frac{d+c}{|M|} \right). \end{aligned}$$

Simplified, this is the ratio between the observed mass of couples of all like types and the counterfactual mass of pairs if individuals matched randomly.

[1] Ciscato, Galichon, and Goussé (2020) call this the homogamy rate; the measure has been applied to study assortative matching on career ambition (Almar et al., 2023) and on childhood family income percentile (Binder et al., 2023).

[2] This is the key point of criticism of EMZ by CCM.

2.1.3 The odds ratio and its monotonic transformations

The *odds ratio* is originally used to quantify the strength of the association between two events. In matching, the odds ratio is interpreted as the ratio of the odds of matching with an educated versus uneducated woman for an educated man and the same odds for an uneducated woman, or equivalently (due to symmetry), the ratio of the odds of marrying an educated versus uneducated man for an educated woman and the same for an uneducated woman.

(OR) odds ratio (CCM); cross-ratio (Edwards, 1963)

$$I_O(M) \equiv \frac{a}{b}/\frac{c}{d} = \frac{a}{c}/\frac{b}{d} = \frac{ad}{bc}.$$

The index ranges from 0 to $+\infty$. A few monotonic transformations of the odds ratio are used in the literature. For our purpose, they all yield the same ranking of assortativity over all matchings.

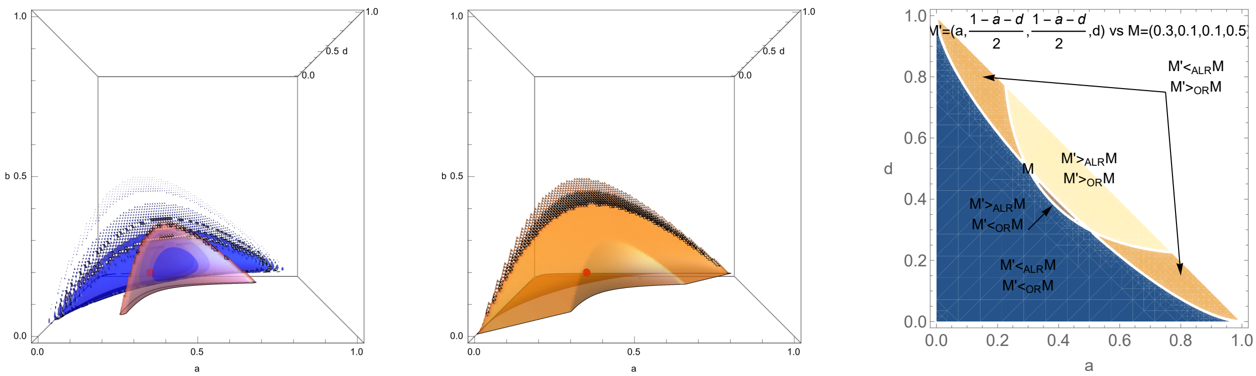
(LOR) Log odds ratio; log cross-ratio (Siow, 2015) $I_o(M) \equiv \ln I_O(M)$.

(Q) Yule's Q; Coefficient of association (Yule, 1900) $I_Q(M) \equiv \frac{I_O(M)-1}{I_O(M)+1}$. This is $+1$ when PAM, -1 when NAM, and 0 when random matching (uncorrelated).

(Y) Yule's Y; Coefficient of colligation (Yule, 1912) $I_Y(M) \equiv \frac{\sqrt{I_O(M)}-1}{\sqrt{I_O(M)}+1}$.

CCM provide a structural interpretation of the measure. In addition, Almar and Schulz (2023) use a weighted average of the odds ratio to document assortative matching on education in Denmark.

Differences between ALR and OR. The empirical debate in the motivation centers on the different conclusions drawn from the aggregate likelihood ratio and the odds ratio. When will the two measures lead to different conclusions about two markets' order of assortativity? Figure 2 provides an illustration. Fix a reference market, $M = (a, b, c, d) = (0.3, 0.1, 0.1, 0.5)$. Figure 2a shows what I call *iso-assortative surfaces* in the (a, b, d) three-dimensional space: The red iso-assortative surface collects the markets $(a, b, c = 1-a-b-d, d)$ that are equally assortative to M according to the aggregate likelihood ratio, and the blue iso-assortative surface collects the markets that are equally assortative to M according to the odds ratio. Because the iso-assortative surfaces are different, the sets of markets that are judged to be more assortative than M by the two measures will differ. The orange region in Figure 2b consists of the markets that are more assortative than M under the aggregate likelihood ratio but are less assortative than M under the odds ratio. Figure 2c illustrates *iso-assortative curves* and conflicting regions between the aggregate likelihood ratio and the odds ratio in a plane, by restricting the attention to markets in which $b = c$. In other words, Figure 2c illustrates the sliced plane of $b = (1 - a - d)/2$ in the (a, b, d) space of Figures 2a and 2b. The iso-assortative surfaces or curves can be drawn for any measure, and can be used to compare any pairs of measures.



(a) Iso-assortative surface

(b) Conflicting region

(c) Iso-assortative curves and conflicting region

Figure 2: Iso-assortative curves and conflicting conclusions

2.1.4 Additional measures

In addition, correlation, its squared expression (Spearman's rank correlation), and pure-random normalization all normalize random matching to zero.

(Corr) Correlation[3]

$$I_{Corr}(M) = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}.$$

A measure that generates the same order is its square, which is called chi-square or Spearman's rank correlation:

$$I_{\chi}(M) = [I_{Corr}(M)]^2 = \frac{(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}.$$

(PR) Pure-random normalization

$$I_{PR}(M) = \frac{ad - bc}{(\max\{b, c\} + d)(a + \max\{b, c\})}.$$

It sets random matching to have an assortativity of 0 and positive assortative matching an

[3]Using correlations, Ciscato, Galichon, and Goussé (2020) compare assortative matching of homosexual and heterosexual partners; Hou et al. (2022) study assortative matching on blood type; and Bailey and Lin (2023) document educational assortativity in the 19th century and early 20th century US.

assortativity of 1.[4]

Most empirical papers use several measures of assortative matching and several definitions of education level as robustness checks (e.g., Hou et al. (2022)). In addition, many papers take a structural approach to use observed matching patterns to identify surpluses generated by different combinations of types (Choo and Siow, 2006; Ciscato, 2023). Many papers study assortative matching because it is an important cause of inequality (Eika, Mogstad, and Zafar, 2019; Chade and Eeckhout, 2023; Pilosoph and Wee, 2023), and consequently the decomposition of causes of inequality involves assortative matching.

2.2 Basic axioms

I start with the following two sets of basic axioms a measure should satisfy. They provide ways to describe (i) when two markets are equally assortative and (ii) when one market is more assortative than another. For generality, whenever possible, I define axioms that an order—rather than an index—should satisfy. In those cases, an index is said to satisfy an axiom if the total order induced by the index satisfies it.

Definition 1. *I say that total order $\succeq$ is **induced by** index I if for all markets M and M', $I(M) > I(M') \Leftrightarrow M \succ M'$ and $I(M) = I(M') \Leftrightarrow M \sim M'$. Two indices I and I' are **order-equivalent** if the total orders $\succeq_I$ and $\succeq_{I'}$ they induce are equivalent.*

2.2.1 Equivalence axioms (EQV)

I start with axioms that define two equally assortative markets. First, when a matching market is doubled or scaled up or down by a constant without changing the relative composition, the assortativity should not be evaluated as changed.

[SInv] Scale Invariance. *For all M and $\lambda > 0$, $M \sim \lambda \cdot M$.*

Note that given SInv, we can transform the matrix into a contingency table by dividing each element by $\lambda = |M|$.

Next, the market should be considered equally assortative if we swap the labels of the two types, which is essentially swapping high-high (a) and low-low (d) pairs and swapping high-low (b) and low-high (c) pairs.

[4] Fernández and Rogerson (2001) and Wu and Zhang (2021) use this measure. Liu and Lu (2006) express the measure for markets with a discrete number of agents. Shen (2020) demonstrates numerically the peril of using only normalization compared with random matching, and advocates for pure-random normalization.

[TSym] Type Symmetry. *The market is equally assortative when types are switched:*

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array} \sim \begin{array}{c|c} d & c \\ \hline b & a \end{array}.$$

Neither should changing the sides of the markets affect the measure of assortativity. Effectively, it switches the proportion of high-low (b) and low-high (c) pairs.

[SSym] Side Symmetry. *The market is equally assortative when sides are switched:*

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array} \sim \begin{array}{c|c} a & c \\ \hline b & d \end{array}.$$

Most of the measures will satisfy these three axioms, and they are basic axioms an appropriate measure should satisfy, so I will refer to all three together as the **equivalence axioms (EQV)**.

2.2.2 Monotonicity axioms

Next, I define notions of monotonicity that specify when one market is more assortative than another. I define two notions of monotonicity, one based on comparisons of markets of different marginal distributions and the other based on comparisons of markets of the same marginals. First, when pairs of like types (i.e., the terms on the diagonal of the matrix) increase, the matching becomes more assortative.

[DMon] Diagonal Monotonicity. *For all M and $\epsilon > 0$,*

$$\begin{array}{c|c} a+\epsilon & b \\ \hline c & d \end{array} \succeq \begin{array}{c|c} a & b \\ \hline c & d \end{array} \text{ and } \begin{array}{c|c} a & b \\ \hline c & d+\epsilon \end{array} \succeq \begin{array}{c|c} a & b \\ \hline c & d \end{array},$$

where the equalities hold if and only if $bc = 0$.

DMon implies the following property that defines PAM.

⇒**[PAM] Positive Assortative Matching.** *For any positive symbols,*

$$\begin{array}{c|c} a & 0 \\ \hline 0 & d \end{array} \sim \begin{array}{c|c} a' & b' \\ \hline 0 & d' \end{array} \sim \begin{array}{c|c} a'' & 0 \\ \hline c'' & d'' \end{array} \succ \begin{array}{c|c} a''' & b''' \\ \hline c''' & d''' \end{array}.$$

Note that the qualifier "if and only if $bc = 0$" is the key for the implication. In the extension in the appendix, I will explore a modification of DMon and its implications on PAM.

I analogously define off-diagonal monotonicity: When pairs of unlike types (i.e., the terms off the diagonal of the matrix) increase, the matching becomes less assortative.

[ODMon] Off-Diagonal Monotonicity. *For all M and $\epsilon > 0$,*

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array} \succeq \begin{array}{c|c} a & b+\epsilon \\ \hline c & d \end{array} \text{ and } \begin{array}{c|c} a & b \\ \hline c & d \end{array} \succeq \begin{array}{c|c} a & b \\ \hline c+\epsilon & d \end{array},$$

where the equalities hold if and only if $ad = 0$.

Analogous to the implication of DMon on PAM, ODMon implies the following property that defines NAM.

⇒**[NAM] Negative Assortative Matching.** *For any positive symbols,*

$$\begin{array}{c|c} 0 & b \\ \hline c & 0 \end{array} \sim \begin{array}{c|c} a' & b' \\ \hline c' & 0 \end{array} \sim \begin{array}{c|c} 0 & b'' \\ \hline c'' & d'' \end{array} \prec \begin{array}{c|c} a''' & b''' \\ \hline c''' & d''' \end{array}.$$

Again, the qualifier "if and only if $ad = 0$" is the key for the implication. In the extension in the next section, I will explore a modification of ODMon and its implication on NAM.

Now, consider an alternative notion of monotonicity that compares markets with the same marginal distributions of men and women. Markets that share the same marginal distributions as (a, b, c, d) are $(a+\epsilon, b-\epsilon, c-\epsilon, d+\epsilon)$—essentially, a one-parameter family.

[MMon] Marginal Monotonicity (CCM). *Consider two markets M and M' with the same marginal distributions (i.e., $a+c = a'+c'$, $a+b = a'+b'$, $d+b = d'+b'$, and $d+c = d'+c'$). $M \succ M'$ if and only if $a > a'$ (equivalently, $b < b'$, $c < c'$, or $d > d'$).*

Claim 1. *DMon and ODMon imply MMon.*

Proof of Claim 1. Suppose $M = (a, b, c, d)$ and $M' = (a', b', c', d')$ have the same marginal distributions, and suppose $a > a'$. Market M' can be represented as

$$M' = \frac{a-(a-a') \;\big|\; b+(a-a')}{c+(a-a') \;\big|\; d-(a-a')} \prec \frac{a \;\big|\; b+(a-a')}{c+(a-a') \;\big|\; d} \prec \frac{a \;\big|\; b}{c \;\big|\; d} = M,$$

where the first $\prec$ follows from DMon and the second $\prec$ follows from ODMon. □

Conversely, however, MMon does not necessarily imply DMon or ODMon, because MMon solely specifies relations for markets with the same marginal distributions and by itself has no implications for markets with different marginal distributions.

Table 1 summarizes whether the measures satisfy the basic equivalence and monotonicity axioms. (Type-specific) likelihood ratio does not satisfy type symmetry, because relabeling the types would change the measure (a point first highlighted by CCM). The few measures that are based on random matching—aggregate likelihood ratio, aggregate homogamy rate, correlation, and pure-random normalization—do not satisfy DMon and/or ODMon. In particular, when a and/or d is relatively small, diagonal monotonicity tends to fail. Normalized trace and odds ratio are the only measures that satisfy all six basic axioms.

Table 1: Do the measures satisfy the basic equivalence and monotonicity axioms?

	Equivalence axioms			Monotonicity axioms		
	SInv	TSym	SSym	DMon	ODMon	MMon
Normalized trace	✓	✓	✓	✓	✓	✓
Type-specific likelihood ratio (EMZ)	✓	X	✓	✓	✓	✓
Aggregate likelihood ratio (EMZ)	✓	✓	✓	X	X	✓
Odds ratio (CCM)	✓	✓	✓	✓	✓	✓
Correlation	✓	✓	✓	X	X	✓
Pure-random normalization	✓	✓	✓	✓	X	✓

3 Characterization results for binary types

Table 1 shows that, except for the type-specific likelihood ratio, all of the measures listed satisfy the equivalence axioms and at least one notion of monotonicity: marginal monotonicity. Hence, the basic axioms are not sufficient to distinguish the measures and provide a definitive answer regarding what measure to use.

3.1 Characterization axioms

We will need additional axioms to distinguish and characterize the different measures. I will introduce what I call *characterization axioms*, such that a measure will be the unique one that satisfies both the basic axioms and an additional axiom, which essentially characterizes the special property of the measure.

The first set of characterization axioms will endow the measures with cardinal interpretations. Namely, the assortativity measure of a market will be a weighted sum or average of the measures of submarkets decomposed from the original market. Depending on the weights used, I will have different decomposability axioms that correspond to different measures.

3.1.1 Population decomposability and normalized trace

Consider the following axiom in which the weight to average is the population size of submarkets.

[PDec] Population Decomposability. For all markets $M \gg 0$ and $M' \gg 0$,

$$I(M + M') = \frac{|M|}{|M| + |M'|} I(M) + \frac{|M'|}{|M| + |M'|} I(M').$$

While, as explained above, MMon and the equivalence axioms cannot imply DMon and ODMon. However, MMon and PDec and the equivalence axioms (SInv and TSym, to be exact) imply DMon and ODMon.

Claim 2. *SInv, TSym, MMon, and PDec imply DMon and ODMon.*

I show in the theorem below that RPDec and PDec are characterization axioms for NT. That is, NT is the unique index, up to linear transformation, that satisfies PDec and the basic equivalence and monotonicity axioms. Because Claims 1 and 2 combine to imply the equivalence of the two notions of monotonicity—DMon+ODMon and MMon—given PDec, the characterization result holds for both notions of monotonicity.

Theorem 1. *An index satisfies EQV, MMon (or DMon and ODMon), and PDec if and only if it is a linear transformation of normalized trace.*

The index is uniquely determined if the range is specified: When the range is $[0, 1]$, the unique index is normalized trace (NT).

3.1.2 Random decomposability and aggregate likelihood ratio

The aggregate likelihood ratio is characterized by a decomposability axiom in which the weight to sum is the expected mass of randomly matched pairs of like types.

> **[RDec] Random Decomposability.** For all markets $M \gg 0$ and $M' \gg 0$,
>
> $$I(M+M') = \frac{r(M)}{r(M+M')}I(M) + \frac{r(M')}{r(M+M')}I(M'),$$
>
> where $r(M)$ indicates the expected mass of pairs of like types under random matching in market M:
>
> $$r(M) \equiv \frac{a+b}{|M|}\frac{a+c}{|M|}|M| + \frac{d+b}{|M|}\frac{d+c}{|M|}|M| = \frac{(a+b)(a+c)+(d+b)(d+c)}{a+b+c+d}.$$

Theorem 2. *An index satisfies EQV, MMon, and RDec if and only if it is proportional to the aggregate likelihood ratio.*

I make two comments. First, note that ALR does not satisfy DMon or ODMon, so I do not have a characterization result that links those axioms and ALR. Second, note that the class of measures that satisfy the axioms in Theorem 2 must be a constant multiple of ALR. The class has to be a constant multiple of ALR because the weights $r(M)/r(M+M')$ and $r(M')/r(M+M')$ do not necessarily add up to be 1. Nonetheless, this characterization result gives ALR a cardinal interpretation.

To address the second comment, consider an axiom that involves weighted averages that add up to 1 and compares markets with proportional marginal distributions.

> **[MRDec] Marginal Random Decomposability.** For all $M \gg 0$ and $M' \gg 0$ such that $M/|M|$ and $M'/|M'|$ have the same marginal distributions,
>
> $$I(M+M') = \frac{r(M)}{r(M)+r(M')}I(M) + \frac{r(M')}{r(M)+r(M')}I(M').$$

ALR satisfies this axiom, but it is not the unique measure that does so. For example, the normalized trace—and as a result, any linear combination of ALR and NT—also satisfies this axiom.

3.1.3 Marginal independence and odds ratio

I now provide the characterization result regarding marginal independence (Edwards, 1963) and the odds ratio (CCM). An order satisfies marginal independence if multiplying any row or any column of any market does not change the assortativity order.

[MInd] Marginal Independence (Edwards, 1963). *For all $M \gg 0$ and $\lambda > 0$,*

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array} \sim \begin{array}{c|c} \lambda a & \lambda b \\ \hline c & d \end{array} \sim \begin{array}{c|c} a & b \\ \hline \lambda c & \lambda d \end{array} \sim \begin{array}{c|c} \lambda a & b \\ \hline \lambda c & d \end{array} \sim \begin{array}{c|c} a & \lambda b \\ \hline c & \lambda d \end{array}.$$

Note that while decomposability axioms contain both ordinal and cardinal contents of assortative matching, in contrast, MInd is an ordinal property. Nonetheless, MInd is a strong condition. MInd implies SInv, TSym, and SSym. MInd and MMon together imply DMon and ODMon.

Claim 3. *MInd and MMon imply DMon and ODMon.*

In fact, more strongly, MMon, DMon, and ODMon are equivalent under MInd:

Claim 4. *Suppose an order or index satisfies MInd. It satisfies DMon if and only if it satisfies ODMon if and only if it satisfies MMon.*

Note that an index and its monotonic transformation are order-equivalent. CCM argue using the odds ratio primarily because it satisfies MInd. While CCM show that the odds ratio satisfies MInd and MMon, I show that the odds ratio is the unique total order that satisfies MInd and MMon (which together imply the basic axioms).

Theorem 3. *A total order satisfies MMon and MInd if and only if it is the order induced by the odds ratio. Equivalently, an index satisfies MMon and MInd if and only if it is a monotonic transformation of the odds ratio.*

3.2 Extensions

3.2.1 Singles

Consider markets with singles. Expand the market without singles by adding a row and a column to indicate singles. Let μ_{i0} denote the mass of single type-θ_i men and μ_{0j} the mass of single type-θ_j women (recall μ_{ij} indicates the mass of pairs of (θ_i, θ_j)):

$$M = \begin{array}{cccc} m \backslash w & \theta_1 & \theta_2 & \emptyset \\ \theta_1 & \mu_{11} & \mu_{12} & \mu_{10} \\ \theta_2 & \mu_{21} & \mu_{22} & \mu_{20} \\ \emptyset & \mu_{01} & \mu_{02} & \end{array}.$$

If I do not consider singles, the following three markets have the same assortativity because they are all perfectly positive assortative.

$$M_1 = \begin{array}{cccc} m \backslash w & \theta_1 & \theta_2 & \emptyset \\ \theta_1 & 100 & 0 & 50 \\ \theta_2 & 0 & 100 & 0 \\ \emptyset & 50 & 0 & \end{array}; \quad M_2 = \begin{array}{cccc} m \backslash w & \theta_1 & \theta_2 & \emptyset \\ \theta_1 & 100 & 0 & 0 \\ \theta_2 & 0 & 100 & 0 \\ \emptyset & 0 & 0 & \end{array}; \quad M_3 = \begin{array}{cccc} m \backslash w & \theta_1 & \theta_2 & \emptyset \\ \theta_1 & 150 & 0 & 0 \\ \theta_2 & 0 & 100 & 0 \\ \emptyset & 0 & 0 & \end{array}.$$

Arguably, M_2 is more assortative than M_1 because there are no singles who could have matched with each other; M_3 is more assortative than M_1 because unmatched individuals in M_1 are assortatively matched in M_3.

[SMon] Singles Monotonicity. *Consider $M = (\mu_{ij})$ and $M' = (\mu'_{ij})$. If $\mu_{i0} > \mu'_{i0}$ for an $i \neq 0$ and $\mu_{jk} = \mu'_{jk}$ for any other combination of j and k, $M \succ M'$.*

Normalized trace and aggregate likelihood ratio are naturally extended to markets with singles: (i) Normalized trace can be defined as the proportion of *individuals* in pairs of like types, and (ii) Aggregate likelihood ratio can be defined as the ratio of the realized mass of individuals in pairs of like types to the hypothetical mass of individuals in pairs of like types. With these modifications, $\widetilde{I}_{tr}(M_1) = 400/500 = 4/5$ and $\widetilde{I}_{tr}(M_2) = \widetilde{I}_{tr}(M_3) = 1$.

Proposition 1. *(i) An index on two-type markets with singles satisfies EQV, MMon (or DMon and ODMon), SMon, and PDec if and only if it is a linear transformation of normalized trace. (ii) An index on two-type markets with singles satisfies EQV, MMon, SMon, and RDec if and only if it is a constant multiple of aggregate likelihood ratio.*

3.2.2 One-sided markets

Consider that there is a one-sided market in which anyone can match with anyone; for example, a homosexual marriage market. With natural modifications of the axioms, I can show that ALR and NT continue to be appropriate measures in this market, even when there are singles and multiple types. The appendix presents the modifications of the axioms and general results with singles and multiple types. The odds ratio does not have an extension to one-sided markets.

Proposition 2. *(i) An index on one-sided markets satisfies EQV, MMon (or DMon and ODMon), and PDec if and only if it is a linear transformation of normalized trace.*

(ii) An index on one-sided markets satisfies EQV, MMon, and RDec if and only if it is a constant multiple of the aggregate likelihood ratio.

3.2.3 Summary of results

Table 2 summarizes whether the key measures of assortative matching satisfy the basic axioms of equivalence and monotonicity, the characterization axioms of the measures, and whether they can be extended to apply to the extensions of the markets. While the aggregate likelihood ratio can be extended to markets, it does not satisfy diagonal monotonicity or off-diagonal monotonicity. While the odds ratio satisfies all the basic axioms, it cannot be extended to any of the markets beyond the binary-type markets. The normalized trace satisfies all the basic axioms, and it can be naturally extended to study all the extensions, including markets with singles, one-sided (same-sex) markets, as well as markets with multiple types. Measuring assortative matching in markets with multiple types will be the subject of the next section.

Table 2: Do the measures satisfy the basic axioms and apply to extensions of markets?

	Axioms				Extensions to markets		
	EQV	MMon	DMon ODMon	Characterization	Singles	Same-sex	Multiple types
normalized trace	✓	✓	✓	PDec	✓	✓	✓
Aggregate LR	✓	✓	X	RDec	✓	✓	✓
Odds ratio	✓	✓	✓	MInd	X	X	X

4 Multiple types

Consider market $M = (\mu_{ij})_{i,j\in\{1,\cdots,n\}}$ for $n \geq 3$.

4.1 Extensions from two-type markets

We have some straightforward extensions from two-type markets. The aggregate likelihood ratio and normalized trace can be naturally extended to multi-type markets, but the odds ratio does not have a natural extension.

(NT) Normalized trace for multi-type markets

$$I_{tr}(M) = \begin{cases} 1 & \textit{if } \mu_{ij}\mu_{ji} = 0 \ \forall i \textit{ and } j \neq i \\ 0 & \textit{if } tr(M) \equiv \sum_{i=1}^{n} \mu_{ii} = 0 \\ tr(M)/|M| & \textit{otherwise.} \end{cases}$$

(ALR) Aggregate likelihood ratio for multi-type markets

$$I_L(M) \equiv \frac{tr(M)/|M|}{\sum_{i=1}^{n} \left(\frac{\sum_{j=1}^{n} \mu_{ij}}{|M|}\right)\left(\frac{\sum_{j=1}^{n} \mu_{ji}}{|M|}\right)} = \frac{|M|(\sum_{i=1}^{n} \mu_{ii})}{\sum_{i=1}^{n}(\sum_{j=1}^{n} \mu_{ij})(\sum_{j=1}^{n} \mu_{ji})}.$$

Theorems 1 and 2 can be naturally extended so that normalized trace and the aggregate likelihood ratio are the unique indices that satisfy the sets of axioms stated in those theorems. Hence, normalized trace and the aggregate likelihood ratio can still be used to measure the assortativity of multi-type markets.

Proposition 3. *(i) An index on multi-type markets satisfies EQV, MMon (or DMon and ODMon), and PDec if and only if it is a linear transformation of normalized trace. (ii) An index on multi-type markets satisfies EQV, MMon, and RDec if and only if it is a constant multiple of the aggregate likelihood ratio.*

4.2 Robustness to categorization

Consider three types θ_i, θ_j, and θ_k. When two types θ_i and θ_j merge so that the three categories are partitioned to $\{\{i,j\}\{k\}\}$, the market given this categorization becomes

$$M|_{\{\{i,j\}\{k\}\}} = \begin{pmatrix} \mu_{ii} + \mu_{ij} + \mu_{ji} + \mu_{jj} & \mu_{ik} + \mu_{jk} \\ \mu_{ki} + \mu_{kj} & \mu_{kk} \end{pmatrix}.$$

I say that M is more assortative than M' if and only if $M|_{\{\{i,j\}\{k\}\}}$ is more assortative than $M|'_{\{\{i,j\}\{k\}\}}$ for all i and all $j \neq i$. With normalized trace, for M and M' such that $|M| = |M'|$ and $\text{tr}(M) = \text{tr}(M')$, M is more assortative than M' if and only if $\mu_{12} + \mu_{21} \geq \mu'_{12} + \mu'_{21}$, $\mu_{13} + \mu_{31} \geq \mu'_{13} + \mu'_{31}$, and $\mu_{23} + \mu_{32} \geq \mu'_{23} + \mu'_{32}$.

Formally,

[RCat] Robustness to Categorization. Let $\mathcal{C}$ denote the collection of categorizations of types to be considered. Let $M|_C$ denote the market under categorization $C \in \mathcal{C}$. $M \succeq M'$ if and only if $M|_C \succeq M'|_C$ for any categorization $C \in \mathcal{C}$, and $M \succ M'$ if and only if $M|_C \succ M'|_C$ for any categorization $C \in \mathcal{C}$.

Theorem 4. *No total order satisfies MMon (or DMon and ODMon) and RCat.*

A counterexample suffices for the claim. Although the counterexample is one of three-type markets, it suffices for the claim for markets with more than three types.

Proof of Theorem 4. Consider markets

$$M = \begin{array}{c|c|c} 1/9 & 1/9 & 1/9 \\ \hline 1/9 & 1/9 & 1/9 \\ \hline 1/9 & 1/9 & 1/9 \end{array} \text{ and } M' = \begin{array}{c|c|c} 1/9-\epsilon & 1/9+\epsilon & 1/9 \\ \hline 1/9+\epsilon & 1/9 & 1/9-\epsilon \\ \hline 1/9 & 1/9-\epsilon & 1/9+\epsilon \end{array} = M + \begin{array}{c|c|c} -\epsilon & +\epsilon & 0 \\ \hline +\epsilon & 0 & -\epsilon \\ \hline 0 & -\epsilon & +\epsilon \end{array}.$$

When I group θ_1 and θ_2,

$$M|_{\{\{1,2\}\{3\}\}} = \begin{array}{c|c} 4/9 & 2/9 \\ \hline 2/9 & 1/9 \end{array} \text{ and } M'|_{\{\{1,2\}\{3\}\}} = \begin{array}{c|c} 4/9+\epsilon & 2/9-\epsilon \\ \hline 2/9-\epsilon & 1/9+\epsilon \end{array} = M|_{\{\{1,2\}\{3\}\}} + \begin{array}{c|c} +\epsilon & -\epsilon \\ \hline -\epsilon & +\epsilon \end{array}.$$

When I group θ_2 and θ_3,

$$M|_{\{\{1\}\{2,3\}\}} = \begin{array}{c|c} 1/9 & 2/9 \\ \hline 2/9 & 4/9 \end{array} \text{ and } M'|_{\{\{1\}\{2,3\}\}} = \begin{array}{c|c} 1/9-\epsilon & 2/9+\epsilon \\ \hline 2/9+\epsilon & 4/9-\epsilon \end{array} = M|_{\{\{1\}\{2,3\}\}} + \begin{array}{c|c} -\epsilon & +\epsilon \\ \hline +\epsilon & -\epsilon \end{array}.$$

By MMon,

$$M|_{\{\{1,2\}\{3\}\}} \prec M'|_{\{\{1,2\}\{3\}\}} \text{ and } M|_{\{\{1\}\{2,3\}\}} \succ M'|_{\{\{1\}\{2,3\}\}}.$$

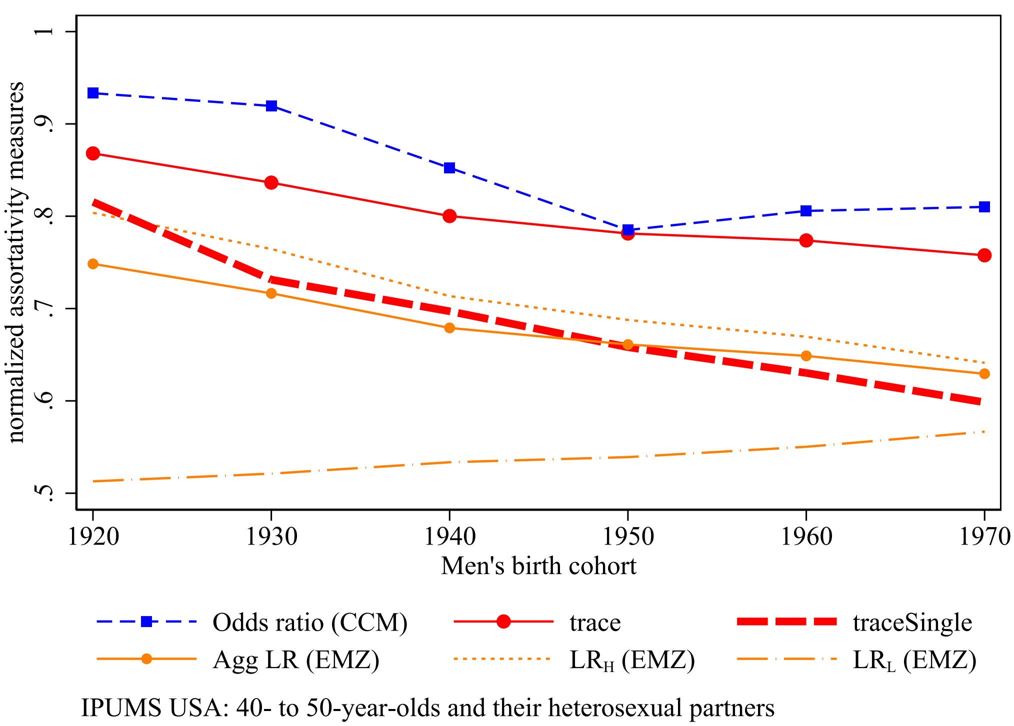
Hence, there does not exist a total order that satisfies MMon and RCat. □

The impossibility result suggests that we must resort to partial orders to satisfy RCat. Existing partial orders such as supermodular stochastic order (Meyer and Strulovici, 2012, 2015) and positive quadrant dependence order (Anderson and Smith, 2022) are natural candidates. However, they usually only compare markets with the same marginals. Hence, further investigations to compare markets with differential marginal distributions are needed.

5 US and global patterns of assortative matching

I use the axiomatized measures to study the patterns of assortative matching on education in the US and other countries across different birth cohorts. All data are from IPUMS USA (Ruggles et al., 2023) and IPUMS International (Minnesota Population Center, 2020). Figure 3 shows changes in the degree of assortative matching under different measures. I consider men of different decadal birth cohorts and their spouses forming a marriage market; more on this restriction later. I define highly educated individuals as those who finished 4 years of college. Under the aggregate likelihood ratio (EMZ) and normalized trace, assortative matching has decreased across all decades. The decrease has been even steeper under normalized trace with singles. In contrast, if I use the odds ratio, I would draw the conclusion that assortative matching has increased in recent decades. In addition, note that the likelihood ratio based on a specific education type yields conflicting conclusions: The likelihood ratio based on high education shows decreasing assortativity over time, but that based on low education shows the opposite conclusion of an increasing assortativity over time.

Figure 3: Assortative matching of heterosexual couples on college education in the US



Because both the aggregate likelihood ratio and normalized trace have cardinal interpretations, I can use them to consider the assortative matching of submarkets, i.e., states.

Figure B1 shows the results for a few states. Generally, the patterns continue to hold when I consider each state. The odds ratio shows more fluctuations across states. I can apply the same measures to provide global evidence; Figure B2 illustrates the assortative matching pattern in several other countries in the world.

Figure 4: Assortative matching of homosexual couples on college education in the US

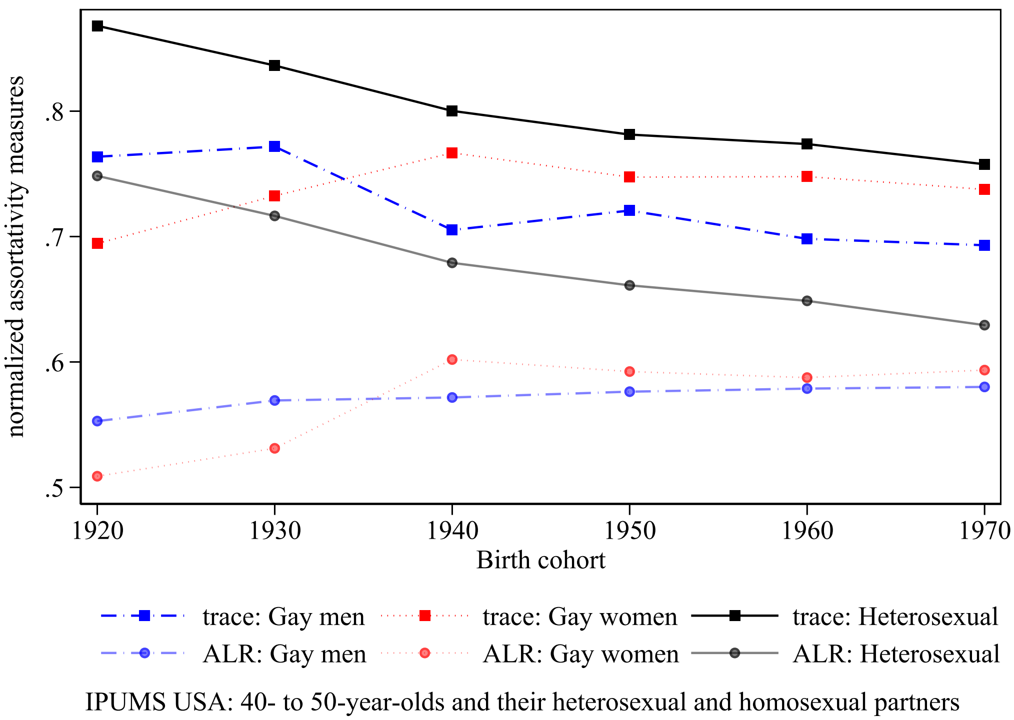


Figure 4 shows the evolution of the assortative matching of homosexual couples in the US. Assortativity has remained steady over time for both male and female same-sex couples, and is lower than that of heterosexual couples when I use either normalized trace or aggregate likelihood ratio to compare.

These empirical analyses motivate the question: What is a marriage market? I can restrict the sample in several ways, for example, to (i) 40- to 50-year-old individuals and their spouses; (ii) 40- to 50-year-old men and their wives; (iii) 40- to 50-year-old women and their husbands; (iv) all those of various birth cohorts who marry in the same year/decade; or (v) couples who cohabit and marry. The definition of a marriage market is outside the scope of this paper, but is empirically relevant. I compute the measures under these alternative definitions of a marriage market, and the conclusions are broadly consistent.

6 Conclusion

Motivated by the recent debate regarding the direction of change of assortative matching on education in the US, I study the appropriate measures of assortative matching. Rather than starting with and justifying any particular measures, I take an axiomatic approach to characterize the appropriate measures—indices, total orders, and partial orders—of matching markets that satisfy a set of justifiable properties. The axioms I consider can be classified as equivalence axioms (those that specify equally assortative markets), monotonicity axioms (those that specify orders of assortativity on markets with or without the same marginal distributions), and characterization axioms (those that characterize the measures).

In summary, I provide axioms to support existing indices, total orders, and partial orders, as well as their modifications. First, the proportion of like types (normalized trace) and its comparison with that under counterfactual random matching (aggregate likelihood ratio) are the indices that are supported by the decomposability axioms that provide cardinal interpretations: A market's assortativity measure is the weighted sum or average of those of its decomposed submarkets. They have natural extensions to markets with singles, one-sided markets, and multi-type markets. The odds ratio is the unique total order on two-type markets that satisfies equivalence axioms, marginal monotonicity, and marginal independence; however, it does not have an extension to markets with singles, one-sided markets, or multi-type markets. When I consider robustness to categorization (assortativity order does not change regardless of categorization of multiple types into fewer types), no total order (hence, no index) satisfies both it and monotonicity axioms. Therefore, we must resort to partial orders.

I propose a new perspective on the different measures on assortative matching. Additional axioms may be proposed as well to further generalize the results. For example, the diagonal monotonicity axiom essentially gives full credit to pairs of like types and no credit to pairs of unlike types, regardless of how "distant" these types are (e.g., PhD-HS pairs are treated the same as college-HS pairs in counting assortativity). Additional generalizations of the basic axioms will bring additional insights. In addition, it may be fruitful to take the axiomatic approach in studying measures of homophily in many-to-one and many-to-many matching markets and networks. I hope that this study's results will create a foundation for further research.

References

Almar, Frederik, Benjamin Friedrich, Ana Reynoso, Bastian Schulz, and Rune Vejlin. 2023. "Marital sorting and inequality: How educational categorization matters." Mimeo.

Almar, Frederik and Bastian Schulz. 2023. "Optimal weights for marital sorting measures." Mimeo.

Anderson, Axel and Lones Smith. 2022. "The comparative statics of sorting." Mimeo.

Arrow, Kenneth J. 1950. "A difficulty in the concept of social welfare." *Journal of Political Economy* 58 (4):328–346.

Bailey, Martha J. and Peter Z. Lin. 2023. "Marital matching and women's intergenerational mobility in the late 19th and early 20th century U.S." Mimeo.

Binder, Ariel J., Caroline Walker, Jonathan Eggleston, and Marta Murray-Close. 2023. "Race, class, and mobility in U.S. marriage market." Mimeo.

Chade, Hector and Jan Eeckhout. 2023. "Stochastic sorting." Mimeo.

Chiappori, Pierre-André, Monica Costa-Dias, and Costas Meghir. 2024. "Measuring assortativeness in marriage." Mimeo.

Choo, Eugene and Aloysius Siow. 2006. "Who marries whom and why." *Journal of Political Economy* 114 (1):175–201.

Ciscato, Edoardo. 2023. "Assessing racial and educational segmentation in large marriage markets." Mimeo.

Ciscato, Edoardo, Alfred Galichon, and Marion Goussé. 2020. "Like attract like? A structural comparison of homogamy across same-sex and different-sex households." *Journal of Political Economy* 128 (2):740–781.

Edwards, A. W. F. 1963. "The measure of association in a 2 x 2 table." *Journal of the Royal Statistical Society A* 126:109–114.

Eika, Lasse, Magne Mogstad, and Basit Zafar. 2019. "Educational assortative mating and household income inequality." *Journal of Political Economy* 127 (6):2795–2835.

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

Hou, Yao, Ke Tang, Jingyuan Wang, Danxia Xie, and Hanzhe Zhang. 2022. "Assortative mating on blood type: Evidence from one million Chinese pregnancies." *Proceedings of the National Academy of Sciences* 119 (51):e2209643119.

Liu, Haoming and Jingfeng Lu. 2006. "Measuring the degree of assortative mating." *Economics Letters* 92:317–322.

Meyer, Margaret and Bruno Strulovici. 2012. "The supermodular stochastic ordering." *Journal of Economic Theory* .

———. 2015. “Beyond correlation: Measuring independence through complementarities.” Mimeo.

Minnesota Population Center. 2020. “Integrated Public Use Microdata Series, International: Version 7.3 [dataset].” https://doi.org/10.18128/D020.V7.3.

Nash, John F. Jr. 1950. “The bargaining problem.” *Econometrica* 18 (2):155–162.

Pilosoph, Laura and Shu Lin Wee. 2023. “Assortative matching and household income inequality: A structural approach.” Mimeo.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler. 2023. “Integrated Public Use Microdata Series, USA: Version 14.0 [dataset].” https://doi.org/10.18128/D010.V14.0.

Shen, Jenny. 2020. *Essays in the economics of gender*. Ph.D. thesis, Princeton University.

Siow, Aloysius. 2015. “Testing Becker’s theory of positive assortative matching.” *Journal of Labor Economics* 33 (2):409–441.

Sonmez, Tayfun. 2023. “Minimalist market design.” Mimeo.

von Neumann, John and Oskar Morgenstern. 1953. *Theory of Games and Economic Behavior*. Princeton University Press.

Wu, Jiabin and Hanzhe Zhang. 2021. “Preference evolution in different matching markets.” *European Economic Review* 137:103804.

Yule, G Udny. 1900. “On the association of attributes in statistics: With illustrations from the material of the Childhood Society.” *Philosophical Transactions of the Royal Society of London Series A* 194:257–319.

———. 1912. “On the methods of measuring association between two attributes.” *Journal of the Royal Statistical Society* 75 (6):579–652.

A Omitted proofs and details

A.1 Omitted proofs for characterization results

Proof of Claim 2. Consider $M = (a, b, c, d) \gg 0$ and $M' = (a', b, c, d) \gg 0$, where $a' > a$. I want to show that $I(M') > I(M)$. By TSym,

$$I(a, b, c, d) = I(d, c, b, a).$$

By PDec,

$$\frac{1}{2}I(a, b, c, d) + \frac{1}{2}I(d, c, b, a) = I\left(\frac{1}{2}(a + d), \frac{1}{2}(b + c), \frac{1}{2}(b + c), \frac{1}{2}(a + d)\right).$$

By SInv,

$$I(a, b, c, d) = I\left(\frac{a + d}{a + b + c + d}, \frac{b + c}{a + b + c + d}, \frac{b + c}{a + b + c + d}, \frac{a + d}{a + b + c + d}\right).$$

By the same sequence of arguments by TSym, PDec, and SInv,

$$I(a', b, c, d) = I\left(\frac{a' + d}{a' + b + c + d}, \frac{b + c}{a' + b + c + d}, \frac{b + c}{a' + b + c + d}, \frac{a' + d}{a' + b + c + d}\right).$$

Note that the two matrices on the right-hand side of the two equations above have the same marginals (each row or column sums to 1). By MMon, $a' > a$ implies

$$I(M) = I(a', b, c, d) > I(M') = I(a, b, c, d).$$

Hence, DMon is proved. ODMon can be shown analogously. □

Proof of Theorem 1. I first show that any index I that satisfies SInv, TSym, SSym, DMon, ODMon, and PDec is order-equivalent to—i.e., a monotonic transformation of—NT. Consider $M = (a, b, c, d) \gg 0$ and $M' = (a', b', c', d') \gg 0$. Here, I prove a lemma.

Lemma 1. *Any index that satisfies SInv, TSym, SSym, DMon, ODMon, and PDec is an increasing function of $a + d$ and a decreasing function of $b + c$.*

Proof of Lemma 1. Consider $M = (a, x, x, d)$, and consider $M_1 = (a/2, x - \lambda/2, \lambda/2, d/2)$ and $M_2 = (a/2, \lambda/2, x - \lambda/2, d/2)$ for some $\lambda \in (0, 2x)$. Note that $M_1 + M_2 = M$ and M_1

and M_2 have the same total mass. Hence, by PDec,

$$I(M) = \frac{1}{2}I(M_1) + \frac{1}{2}(M_2).$$

By SSym, $I(M_1) = I(M_2)$. Hence, $I(M) = I(M_1) = I(M_2)$. By SInv,

$$I(M_1) = I(2M_1) = I(a, 2x - \lambda, \lambda, d).$$

Hence, for all $\lambda \in (0, 2x)$,

$$I(a, 2x - \lambda, \lambda, d) = I(a, x, x, d).$$

Hence, I have shown that $I(a, b, c, d) = I(a, b', c', d)$ whenever $b + c = b' + c'$. By the same logic, and by SSym and TSym, $I(a, b, c, d) = I(a', b, c, d')$ whenever $a + d = a' + d'$. I, by DMon, is strictly increasing in $a+d$ whenever $bc \neq 0$, and, by ODMon, is strictly decreasing in $b + c$ whenever $ad \neq 0$. □

If (i) $a + d > a' + d'$ and $b + c \leq b' + c'$ or (ii) $a + d \leq a' + d'$ and $b + c > b' + c'$, then, by Lemma 1, (i) $I(M) > I(M')$ or (ii) $I(M) < I(M')$, respectively. Suppose $a + d > a' + d'$ and $b + c > b' + c'$. Define

$$M'' = (a'', b'', c'', d'') = M' \cdot (b + c)/(b' + c').$$

By definition of M'', $b'' + c'' = b + c$, and $a'' + d'' = (a' + d') \cdot (b + c)/(b' + c')$. By SInv of I, $I(M'') = I(M')$. The comparison of $a + d$ and $a'' + d''$ pins down the ordinal assortativity relation between M and M'. That is,

$$\frac{a + d}{a'' + d''} = \frac{a + d}{b + c} \Big/ \frac{a' + d'}{b' + c'} > 1 \Leftrightarrow I(M) > I(M').$$

When $a + d < a' + d'$ and $b + c < b' + c'$, I can similarly pin down the ordinal assortativity relation between M and M'. Note that for any $M = (a, b, c, d) \gg 0$,

$$I_{tr}(M) = \frac{(a + d)}{(a + b) + (c + d)} = \frac{a + d}{b + c} \Big/ \left(\frac{a + d}{b + c} + 1\right).$$

Hence, $I(M) > I(M')$ if and only if $I_{tr}(M) > I_{tr}(M')$.

Any nonlinear transformation of NT would violate PDec or RPDec, so any index that satisfies the stated axioms must be not only a monotonic transformation but also a linear transformation of NT. □

Proof of Theorem 2. It is straightforward to check that ALR satisfies the axioms, so it remains to show the other direction. I first show that any index I that satisfies SInv, TSym, SSym, MMon, and RDec is proportional to ALR. Consider $M = (a, b, c, d)$. By TSym,

$$I(M) = I\begin{pmatrix} a & b \\ c & d \end{pmatrix} = I\begin{pmatrix} d & c \\ b & a \end{pmatrix}. \tag{1}$$

Recall

$$r(M) \equiv \frac{a+b}{|M|} + \frac{a+c}{|M|} + \frac{d+b}{|M|}\frac{d+c}{|M|}|M| = \frac{(a+b)(a+c)+(d+b)(d+c)}{a+b+c+d}.$$

By RDec,

$$I\begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix} \cdot r\begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix} = I\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot r\begin{pmatrix} a & b \\ c & d \end{pmatrix} + I\begin{pmatrix} d & c \\ b & a \end{pmatrix} \cdot r\begin{pmatrix} d & c \\ b & a \end{pmatrix},$$

which, by (1), is simplified to

$$I\begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix} \frac{|M|^2+|M|^2}{2|M|} = 2I(M)r(M) \Rightarrow I(M) = \frac{1}{2}\frac{|M|}{r(M)} I\begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix}. \tag{2}$$

Because for any $\epsilon < \min\{a+d, b+c\}$,

$$\begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix} = \begin{pmatrix} a+d-\epsilon & \epsilon \\ \epsilon & a+d-\epsilon \end{pmatrix} + \begin{pmatrix} \epsilon & b+c-\epsilon \\ b+c-\epsilon & \epsilon \end{pmatrix},$$

by RDec, for any $\epsilon < \min\{a+d, b+c\}$,

$$|M| I\begin{pmatrix} a+d & b+c \\ b+c & a+d \end{pmatrix} = (a+d) I\begin{pmatrix} a+d-\epsilon & \epsilon \\ \epsilon & a+d-\epsilon \end{pmatrix} + (b+c) I\begin{pmatrix} \epsilon & b+c-\epsilon \\ b+c-\epsilon & \epsilon \end{pmatrix}.$$

Plugging in the expression of $I(M)$ in (2), I get, for any $\epsilon < \min\{a+d, b+c\}$,

$$I(M) = \frac{1}{2r(M)}\left[(a+d)\cdot I\begin{pmatrix} a+d-\epsilon & \epsilon \\ \epsilon & a+d-\epsilon \end{pmatrix} + (b+c)\cdot I\begin{pmatrix} \epsilon & b+c-\epsilon \\ b+c-\epsilon & \epsilon \end{pmatrix}\right].$$

Take $\epsilon \to 0$ and by SInv and $I(0, 1, 1, 0) = 0$, I have

$$I(M) = \frac{1}{2}\frac{a+d}{r(M)} I\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence, any index that satisfies the above axioms is proportional to $(a+d)/r(M)$, the aggregate likelihood ratio. $\square$

Proof of Claim 3. Suppose $M=(a,b,c,d)$ and $M'=(a',b',c',d')$ have the same marginal distributions and $M \succ M'$. By MInd,

$$\begin{aligned}(a,b,c,d) &\sim (a,b,c\frac{c'}{c},d\frac{c'}{c}) \sim (a,b\frac{c}{c'}\frac{d'}{d},c',d\frac{c'}{c}\frac{c}{c'}\frac{d'}{d}) \sim (a\frac{b'}{b}\frac{c'}{c}\frac{d}{d'},b\frac{b'}{b}\frac{c'}{c}\frac{d}{d'}\frac{c}{c'}\frac{d'}{d},c',d') \\ &\sim (a'\frac{a}{a'}\frac{b'}{b}\frac{c'}{c}\frac{d}{d'},b',c',d').\end{aligned}$$

Let $\delta \equiv \frac{a}{a'}\frac{b'}{b}\frac{c'}{c}\frac{d}{d'}$. By MMon, $a>a'$, $b<b'$, $c>c'$, and $d>d'$. Hence, $\delta>1$. $(a\delta,b',c',d') \sim (a',b'/\delta,c',d') \sim (a',b',c'/\delta,d') \sim (a',b',c',d'\delta) \succ (a',b',c',d')$ for all a',b',c',d' implies DMon and ODMon. $\square$

Proof of Claim 4. Suppose order $\succeq$ satisfies MMon. Then I have for any δ,

$$(a,b,c,d) \succ (a-\delta,b+\delta,c+\delta,d+\delta).$$

By (repeatedly applying) MInd,

$$\begin{aligned}(a-\delta,b+\delta,c+\delta,d+\delta) &\sim (a,b+\delta,(c+\delta)\frac{a}{a-\delta},d-\delta) \\ &\sim (a,b,(c+\delta)\frac{a}{a-\delta},(d-\delta)\frac{b}{b+\delta}) \\ &\sim (a,b,c,d\frac{a-\delta}{a}\frac{b}{b+\delta}\frac{c}{c+\delta}\frac{d-\delta}{d}) \equiv (a,b,c,d\lambda(\delta)),\end{aligned}$$

where $\lambda(\delta)$ is strictly smaller than 1 for any $\delta>0$ and is continuously decreasing in δ. With similar transformations, I have

$$\begin{aligned}&(a-\delta,b+\delta,c+\delta,d+\delta) \\ \sim\ &(a,b,c,d\lambda(\delta)) \sim (a\lambda(\delta),b,c,d) \sim (a,b/\lambda(\delta),c,d) \sim (a,b,c/\lambda(\delta),d).\end{aligned}$$

Hence, for any $\delta>0$, DMon holds:

$$(a,b,c,d) \succ (a,b,c,d\lambda(\delta)) \sim (\lambda(\delta)a,b,c,d),$$

and ODMon holds:

$$(a,b,c,d) \succ (a,b/\lambda(\delta),c,d) \sim (a,b,c/\lambda(\delta),d).$$

Reversely,

$$(a,b,c,d) \succ (a,b,c,d\lambda(\delta)) \sim (\lambda(\delta)a,b,c,d)$$

for any δ implies

$$(a,b,c,d) \succ (a,b/\lambda(\delta),c,d) \sim (a,b,c/\lambda(\delta),d)$$

for any δ, and

$$(a,b,c,d) \succ I(a-\delta,b+\delta,c+\delta,d-\delta)$$

for any δ, where MInd is repeatedly used again. $\square$

Proof of Theorem 3. It is straightforward to check that the odds ratio and its monotonic transformation satisfy DMon, ODMon, and MInd.

It remains to show that there does not exist an index I that satisfies DMon, ODMon, and MInd, but is not order-equivalent to the odds ratio. Suppose by way of contradiction that such an index exists. Then, I must have for some $M = (a,b,c,d)$ and $M' = (a',b',c',d')$. One of the following four cases occurs: (i) $I(M) > I(M')$ and $Q(M) < Q(M')$, (ii) $I(M) < I(M')$ and $Q(M) > Q(M')$, (iii) $I(M) = I(M')$ and $Q(M) \neq Q(M')$, and (iv) $I(M) \neq I(M')$ and $Q(M) = Q(M')$.

First, suppose that case (i) $I(M) > I(M')$ and $Q(M) < Q(M')$ occurs. Because $I(M) > I(M')$, by the implication of DM, $ad \neq 0$, and by the implication of ODM, $b'c' \neq 0$, and because $Q(M) < Q(M')$, similarly, $bc \neq 0$ and $a'd' \neq 0$. For each of the following steps, I invoke a part of MInd:

$$\begin{aligned}
I(a,b,c,d) &= I(\frac{b'}{b}a, \frac{b'}{b}b, c, d) \\
&= I(\frac{b'}{b}a\frac{a'}{a}\frac{b}{b'}, b', c\frac{a'}{a}\frac{b}{b'}, d) \\
&= I(a', b', c\frac{a'}{a}\frac{b}{b'} \cdot \frac{c'}{c}\frac{a}{a'}\frac{b'}{b}, d \cdot \frac{c'}{c}\frac{a}{a'}\frac{b'}{b}\frac{d'}{d'}) \\
&= I(a', b', c', d' \cdot \frac{ad}{bc}/\frac{a'd'}{b'c'}).
\end{aligned}$$

By premise,

$$I(a', b', c', d' \cdot \frac{ad}{bc}/\frac{a'd'}{b'c'}) > I(a', b', c', d').$$

By DM, this implies

$$\frac{ad}{bc} > \frac{a'd'}{b'c'}.$$

However, this implies $Q(M) > Q(M')$, which contradicts the premise that $Q(M) < Q(M')$.

For each of the four possibilities, if no cell of matrices M and M' is zero, using the same logic as above, a contradiction can be derived.

Suppose there is a cell that is zero. Suppose $I(M) = I(M')$. If $bc = 0$, then $I(M) = I(a, 0, 0, d) = I(M')$, then $b'c' = 0$. In this case, by DM, $Q(M) = Q(M')$. If $ad = 0$ instead, then $I(M) = I(0, b, c, 0) = I(M')$ implies $a'd' = 0$. In this case, by ODM, $Q(M) = Q(M')$. Hence, whenever there is a cell with zero in one of the matrices, $I(M) = I(M')$ and $Q(M) = Q(M')$, which prevents all four cases from happening. $\square$

A.2 Extension: A stricter notion of perfect assortativity

Suppose I desire a stricter notion of perfect assortativity by specifying that a market is perfectly positive assortative only if there is no pair of different types and is perfectly negative assortative only if there is no pair of like types.[5] Namely,

⇒**[PAM'] Strict Positive Assortative Matching.** *For any positive symbols,*

$$\begin{array}{c|c} a & 0 \\ \hline 0 & d \end{array} \succ \begin{array}{c|c} a' & b' \\ \hline 0 & d' \end{array}, \begin{array}{c|c} a'' & 0 \\ \hline c'' & d'' \end{array}, \begin{array}{c|c} a''' & b''' \\ \hline c''' & d''' \end{array}.$$

⇒**[NAM'] Strict Negative Assortative Matching.** *For any positive symbols,*

$$\begin{array}{c|c} 0 & b \\ \hline c & 0 \end{array} \prec \begin{array}{c|c} a' & b' \\ \hline c' & 0 \end{array}, \begin{array}{c|c} 0 & b'' \\ \hline c'' & d'' \end{array}, \begin{array}{c|c} a''' & b''' \\ \hline c''' & d''' \end{array}.$$

The following modifications of DMon and ODMon will imply PAM' and NAM', and they will characterize a slightly modified version of the normalized trace.

[DMon'] Strict Diagonal Monotonicity. *For all $\epsilon > 0$,*

$$(a + \epsilon, b, c, d) \succeq (a, b, c, d) \text{ and } (a, b, c, d + \epsilon) \succeq (a, b, c, d),$$

where the equalities hold if and only if $b = c = 0$.

[5]For example, men and women could have chosen their education level to start, and any gender difference in education distribution and consequent mass of pairs of unlike education can be interpreted as an undesirable outcome. The level of nonassortativity can also be interpreted as the level of frictions in the market.

[ODMon'] Strict Off-Diagonal Monotonicity. *For all $\epsilon > 0$,*

$$(a, b+\epsilon, c, d) \succeq (a, b, c, d) \text{ and } (a, b, c+\epsilon, d) \succeq (a, b, c, d),$$

where the equalities hold if and only if $a = d = 0$.

Arguably, in certain situations, this "stricter" notion of perfect assortative matching is needed. For example, when there is a pre-matching investment stage in which individuals can choose their matching types, although individuals may achieve the most possible pairs of like types, the fact that there is an imbalance in the types of agents indicates that there is coordination failure in the investment stage. Hence, there may be a need to distinguish the assortativity of $(a, 0, 0, d)$ and $(a, b, 0, d)$, where $b > 0$.

[PDec'] Generalized Population Decomposability. For any markets M and M',

$$I(M + M') = \frac{|M|}{|M + M'|} I(M) + \frac{|M'|}{|M + M'|} I(M').$$

(NT') Continuous normalized trace

$$I'_{tr}(M) = \frac{\mathrm{tr}(M)}{|M|}.$$

Claim 5. *An index satisfies EQV, DMon', ODMon', and PDec' if and only if it is a linear transformation of NT'.*

B Additional empirical analyses

Figure B1 considers the assortative matching of heterosexual couples in different American states. The patterns in individual states align with the national pattern: By odds ratio, the recent cohorts have witnessed an increase in assortative matching on education, but by other measures (the aggregate likelihood ratio and normalized trace), assortative matching has declined over time.

Figure B2 considers the assortative matching of heterosexual couples in other countries in the world. I use IPUMS International's unified coding of education. Overall, there has been a decline in assortative matching worldwide.

Figure B1: Assortative matching of heterosexual couples on college education in US states

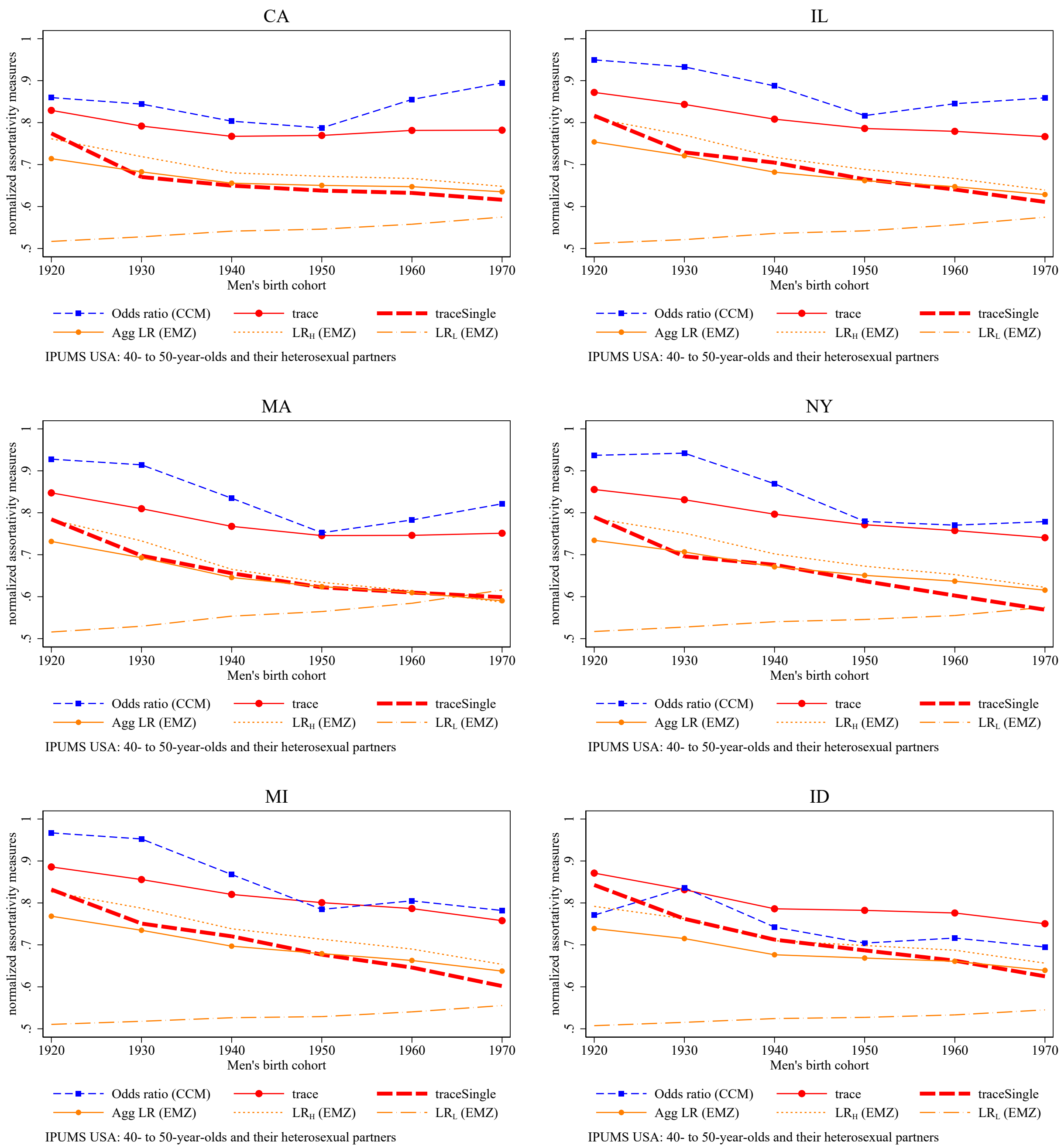


Figure B2: Assortative matching of heterosexual couples on college education in the world