IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16368

Optimal Weights for Marital Sorting Measures

Frederik Almar
Bastian Schulz

AUGUST 2023

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16368

Optimal Weights for Marital Sorting Measures

Frederik Almar
Aarhus University

Bastian Schulz
Aarhus University, Dale T. Mortensen Centre, IZA and CESifo

AUGUST 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Optimal Weights for Marital Sorting Measures*

Changing distributions of male and female types affect the measurement of educationbased marriage market sorting. We develop a weighting strategy that minimizes the distortion of sorting measures due to changing type distributions. The optimal weights reflect that female type distributions have changed relatively more in recent decades. Based on our weighted measure, we document increasing sorting in Denmark between 1980 and 2018. Alternative measures suggest flat or decreasing trends.

JEL Classification: C43, D10, J11, J12

Keywords: positive assortative mating, marriage market sorting, homophily, educational attainment, sorting measures, aggregation

Corresponding author:
Bastian Schulz
Aarhus University
Department of Economics and Business Economics
Fuglesangs Allé 4
DK-8210 Aarhus V
Denkmark

E-mail: bastian.schulz@econ.au.dk

* We thank Ana Reynoso and Rune Vejlin for comments and inspiring discussions. The Danish register data used in this study has been accessed through ECONAU, Department of Economics and Business Economics, Aarhus University. Bastian Schulz thanks the Aarhus University Research Foundation (Grant no. AUFF-F-2018-7-6), and the Dale T. Mortensen Centre, Aarhus University, for financial support.

1 Introduction

An increasing tendency of individuals to marry their like in terms of educational attainment, a phenomenon known as positive assortative mating (PAM, sorting), potentially increases inequality between households (e.g., Kremer, 1997; Fernández and Rogerson, 2001; Breen and Salazar, 2011; Chiappori et al., 2020a). However, the literature disagrees on whether PAM has increased over time (e.g., Greenwood et al., 2016; Eika et al., 2019; Almar et al., 2023). Shifting distributions of education-based types constitute a measurement challenge. While recent papers acknowledge this (e.g., Liu and Lu, 2006; Eika et al., 2019; Chiappori et al., 2020b, 2021), how changing type distributions distort commonly-used sorting measures is an open question.

To answer this question, we provide an in-depth analysis of a prominent sorting measure. We formally analyze how changing type distributions affect the measure and propose a strategy to minimize the distortion. Using weights, we take into account that the distribution of educational attainment has evolved in different ways for men and women. In Danish data, educational outcomes of females increased relatively more in recent decades (consistent with, e.g., Goldin, 2006). The optimal weights according to our decision rule eliminate the dominating effect of female-type-distribution changes on the sorting measure. We find that PAM has increased while alternative measures suggest flat or decreasing trends.

We conclude that it is important to take gender-specific trends in the underlying type distributions into account. The optimal choice of weights is context-dependent and matters for conclusions about sorting trends.

2 Data and Trends

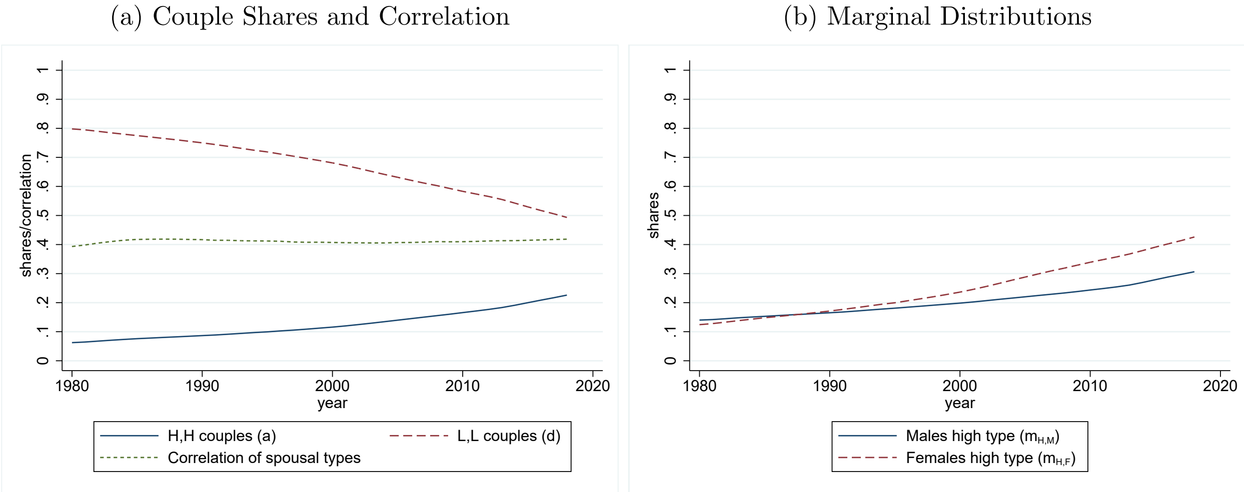
We use Danish data to illustrate how evolving type distributions affect the measurement of PAM. The population register contains demographic variables and person IDs for all residents and their (married or cohabiting) partners (Statistics Denmark BEF, 1980–2018).[1] We study the period 1980–2018 and observe on average 1,800,866 individuals in the age range 19–60 per year who are either married to or cohabiting with an individual of the opposite sex. The combined stock of couples is stable over time.[2]

We use the education register (Statistics Denmark UDDA, 1980–2018) to distinguish between highly-educated individuals (bachelor's degrees and above, ISCED 6–8) and individuals

[1]Cohabitation is identified based on a number of criteria: opposite-sex, joint children, shared address, less than 15 years age difference, no family relationship.

[2]Figure A.1 depicts the evolution of the stocks of couple types and their age composition.

Figure 1: Assortative Matches and Marginal Distributions



Note: Panel (a) shows how the shares of couples in which both spouses have either high or low educational attainment, (H, H) or (L, L), have evolved over time, along with the cross-sectional correlation of spousal types. Panel (b) shows the evolution of the fraction of highly-educated males and females. Section 2 explains how the sample and the education-based types are constructed. The symbols a, d, $m_{(H,M)}$, and $m_{(H,F)}$ are introduced in Section 3 and link the data series to the formal analysis of sorting measures.

with lower educational attainment (compulsory schooling, high school, vocational training, short-cycle tertiary programs, ISCED 1–5). Thus, the education-based type T is either H (high) or L (low). Gender is indexed M (male) and F (female).

Figure 1a shows that the share of (H, H) couples (blue, solid line) increased between 1980 and 2018. Thus, it has become more common to observe couples in which both partners are highly educated. However, the share of (L, L) couples (red, dashed line) has decreased. Moreover, Panel (b) shows the shift toward higher education. In 2018, more than 40% (30%) of women (men) are highly educated, compared to around 15% in 1980. This shift in the marginal distribution of types affects the share of (H, H) couples as it became more likely to meet highly-educated individuals. Thus, couple-type shares alone cannot provide evidence for PAM.

The cross-sectional correlation of couple types in Panel (a) (green, dotted line) is essentially flat. Note that the correlation coefficient conflates changes of couple shares and changes of marginal distributions. We show in appendix B that the correlation responds to such changes in a highly nonlinear way, which makes the trend of the correlation coefficient uninformative about PAM.[3]

In summary, we need a formal framework to measure PAM and disentangle it from changes of the marginal type distributions.

[3]See also Eika et al. (2019) and Chiappori et al. (2021).

Table 1: Contingency tables

(a)

M\F	H	L	Marginal
H	a	b	$a+b$
L	c	d	$c+d$
Marginal	$a+c$	$b+d$	1

(b)

M\F	H	L	Marginal
H	a	$m_{(H,M)}-a$	$m_{(H,M)}$
L	$m_{(H,F)}-a$	d	$1-m_{(H,M)}$
Marginal	$m_{(H,F)}$	$1-m_{(H,F)}$	1

3 Measurement and Optimal Weights

3.1 The Setup

The contingency table 1a summarizes the marriage market allocation. $a > 0$ ($d > 0$) denotes the share of couples in which both spouses have high (low) education. $a + d$ is the share of sorted couples. $b > 0$ and $c > 0$ denote the shares of couples with different levels of education. Intuitively, the higher $a + d$ relative to $b + c$, the more pronounced is PAM.

To investigate how changing marginals affect sorting measures, we substitute the share of high-type men $m_{(H,M)}$ for $a + b$ and the share of high-type women $m_{(H,F)}$ for $a + c$ in table 1b.

3.2 The Weighted Sum of Likelihood Indices

Based on table 1, the weighted sum of likelihood indices is defined as follows:

$$I_{\mathcal{S}} = \frac{a}{m_{(H,M)}m_{(H,F)}} \times w_H + \frac{d}{(1-m_{(H,M)})(1-m_{(H,F)})} \times w_L. \tag{1}$$

The measure fulfills the formal criteria for sorting measures outlined by Chiappori et al. (2020b, 2021). PAM is captured by the ratio of the actual shares of sorted couples and the expected shares based on the "supply" of different types.

w_H and w_L are weights used to aggregate along the diagonal. Chiappori et al. (2020b) suggest that these weights can be thought of as a convex combination of the shares of males and females with the same level of education, which depend on the respective marginal distribution. Let $I_{\mathcal{S}}^{convex}$ denote the measure with these weights applied, where $\lambda \in [0, 1]$ is the coefficient on the male marginal distribution:

$$\begin{aligned} I_{\mathcal{S}}^{convex} &= \frac{a}{m_{(H,M)}m_{(H,F)}} \times (\lambda m_{(H,M)} + (1-\lambda)m_{(H,F)}) \\ &+ \frac{d}{(1-m_{(H,M)})(1-m_{(H,F)})} \times (\lambda(1-m_{(H,M)}) + (1-\lambda)(1-m_{(H,F)})). \end{aligned} \tag{2}$$

To investigate the impact of changing shares of sorted couples and marginal distributions, we totally differentiate (2):

$$\Delta I_S^{convex} = \underbrace{\left(\frac{\lambda m_{(H,M)} + (1-\lambda) m_{(H,F)}}{m_{(H,M)} m_{(H,F)}}\right)}_{>0} \Delta a + \underbrace{\left(\frac{\lambda(1-m_{(H,M)}) + (1-\lambda)(1-m_{(H,F)})}{(1-m_{(H,M)})(1-m_{(H,F)})}\right)}_{>0} \Delta d$$

$$+ \underbrace{(1-\lambda)\left(\frac{d}{(1-m_{(H,M)})^2} - \frac{a}{m_{(H,M)}^2}\right)}_{\gtreqless 0} \Delta m_{(H,M)} + \underbrace{\lambda\left(\frac{d}{(1-m_{(H,F)})^2} - \frac{a}{m_{(H,F)}^2}\right)}_{\gtreqless 0} \Delta m_{(H,F)}. \tag{3}$$

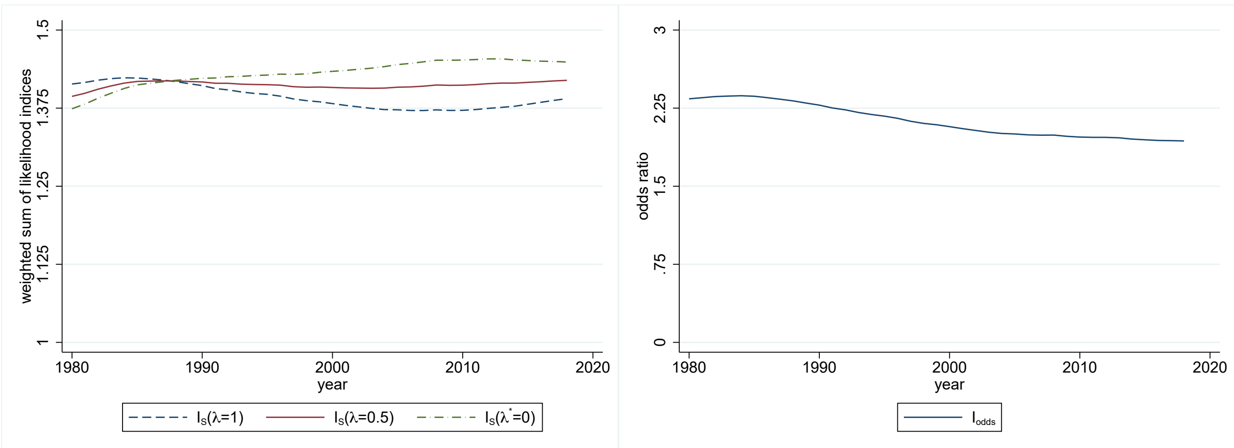
I_S^{convex} is increasing in the shares of sorted couples (a, d) because the coefficients in the first line are positive. However, the impact of changing marginal distributions depends on both the configuration of the contingency table and λ. Thus, the choice of λ allows us to take into account the importance of gender differences in the effect of changing marginals on measured sorting.

We plot the likelihood index I_S^{convex} for different values of λ in figure 2a. It indicates PAM in all cases because $I_S^{convex} > 1$. However, different values of λ lead to different trends. With weight on changes in the male type distribution ($\lambda = 1$), sorting is decreasing. With weight on changes in the female type distribution ($\lambda = 0$), sorting is increasing. For $\lambda = 0.5$, the trend is flat. Thus, the choice of λ is crucial for conclusions about the trend of PAM.

Figure 2: Sorting Trends Depend on Measures and Weights

(a) Likelihood Indices

(b) Odds Ratio



Note: Panel (a) depicts the evolution of the weighted sum of likelihood indices as defined in equation (1) for three different values of λ. Panel (b) depicts the evolution of the odds ratio as defined in equation (5). Section 2 explains how the sample and the education-based types are constructed.

The choice of λ

We propose to choose λ to minimize the impact of marginal-distribution changes on the sorting measure. This can be achieved by setting $\lambda \in [0, 1]$ such that the absolute value of the sum of the $\Delta m_{(H,M)}$ and $\Delta m_{(H,F)}$ terms in equation (3) is minimized:

$$\min_{\lambda} \left| (1-\lambda) \underbrace{\left(\frac{d}{(1-m_{(H,M)})^2} - \frac{a}{m^2_{(H,M)}} \right) \Delta m_{(H,M)}}_{=\gamma_1} + \lambda \underbrace{\left(\frac{d}{(1-m_{(H,F)})^2} - \frac{a}{m^2_{(H,F)}} \right) \Delta m_{(H,F)}}_{=\gamma_2} \right| .$$

This objective function is a convex combination of the two endpoints γ_1 and γ_2. If $sign(\gamma_1) \neq sign(\gamma_2)$, then zero lies between the two endpoints and the optimal λ^* solves $(1-\lambda)\gamma_1 + \lambda\gamma_2 = 0$. If, on the other hand, $sign(\gamma_1) = sign(\gamma_2)$, then the optimal λ^* is the endpoint with the smallest absolute value, either $|\gamma_1|$ or $|\gamma_2|$. In summary:

$$\lambda^* = \begin{cases} 0 & \text{if} \quad sign(\gamma_1) \quad = \quad sign(\gamma_2), \quad |\gamma_2| > |\gamma_1| \\ \frac{\gamma_1}{\gamma_1 - \gamma_2} & \text{if} \quad sign(\gamma_1) \quad \neq \quad sign(\gamma_2) \\ 1 & \text{if} \quad sign(\gamma_1) \quad = \quad sign(\gamma_2), \quad |\gamma_1| > |\gamma_2| \, . \end{cases} \tag{4}$$

From figure 1, we know that $\Delta m_{(H,F)} > \Delta m_{(H,M)} > 0$ and that in the base year 1980 $m_{(H,M)} \approx m_{(H,F)}$. Thus, $sign(\gamma_1) = sign(\gamma_2)$ for all years. λ^* must be either zero or one. In the data, $|\gamma_2| > |\gamma_1|$ because of the bigger change in the female marginal type distribution. Thus, $\lambda^* = 0$ is optimal for all years.

$\lambda^* = 0$ implies increasing sorting, see figure 2a. This is due to the fact that the positive contribution of more sorted high-type couples (term one in equation (3) is positive) outweighs the negative contributions from fewer sorted low-type couples (term two in equation (3) is negative) and changing marginal type distributions (term three is negative and term four drops out with $\lambda^* = 0$ in equation (3)).[4]

An advantage of the weighted sum of likelihood indices is that it can be defined for any number of types. In appendix C, we generalize the decision rule (4) for more than two types.

[4]Totally eliminating the effect of changing marginal distributions on the sorting measure—the sum of terms three and four in equation (3)—would require λ^* to be outside the unit interval. The measure would no longer fulfill the monotonicity property stated in Chiappori et al. (2021) because the measure would decrease in the share of sorted couples, see terms one and two in equation (3).

3.3 Alternative Measures

The odds ratio

An alternative measure that also fulfills the formal criteria for sorting measures outlined by Chiappori et al. (2021) is the (log) odds ratio. Based on table 1, it is defined as follows:

$$I_{odds} = \ln\left(\frac{ad}{bc}\right) = \ln\left(\frac{ad}{(m_{(H,M)} - a)(m_{(H,F)} - a)}\right). \tag{5}$$

It can be written in terms of the marginal distributions using $m_{(H,M)}$ and $m_{(H,F)}$. As before, we totally differentiate I_{odds}.

$$\begin{aligned} \Delta I_{odds} &= \underbrace{\left(\frac{m_{(H,M)}m_{(H,F)} - a^2}{a(m_{(H,M)} - a)(m_{(H,F)} - a)}\right)}_{>0}\Delta a + \underbrace{\left(\frac{1}{d}\right)}_{>0}\Delta d \\ &- \underbrace{\left(\frac{1}{m_{(H,M)} - a}\right)}_{>0}\Delta m_{(H,M)} - \underbrace{\left(\frac{1}{m_{(H,F)} - a}\right)}_{>0}\Delta m_{(H,F)}. \end{aligned} \tag{6}$$

Increasing shares of sorted couples a and d imply higher sorting while increasing shares of high-type individuals $m_{H,M}$ and $m_{H,F}$ imply lower sorting. Thus, I_{odds} can decrease over time if the increase in the shares of high-type men or women is sufficiently large.

We plot I_{odds} in figure 2b. $I_{odds} > 0$ indicates PAM. However, the odds ratio is decreasing over time. The increasing share of (H, H) couples ($\Delta a > 0$) is dominated by a decreasing share of (L, L) couples ($\Delta d < 0$) and increasing shares of highly-educated males and females ($\Delta m_{(H,M)} > 0$, $\Delta m_{(H,F)} > 0$), recall figure 1. Note that the coefficients of the male and female high-type shares are symmetric in (6). Therefore, the measure does not allow for gender-specific effects of changing marginals on measured PAM. Another limitation is that the odds ratio is defined for two types only.

Alternative Weights

Greenwood et al. (2016), Eika et al. (2019), and Almar et al. (2023) use versions of measure (1) with alternative weights. In appendix D, we show that those weights are not necessarily a convex combination of the male and female marginals. The effect of more sorted couples is thus not guaranteed to be positive. In our data, conclusions based on the optimal λ and the alternative weights used in the literature are similar, i.e., sorting is increasing. However, this is coincidental and not guaranteed to hold in other settings.

4 Conclusion

We show how to use gender-specific weights, which compensate changes of the underlying type distributions, to improve the measurement of education-based marriage market sorting. Because the female type distribution has changed more than the male one in recent decades, attaching the weight to the female side minimizes the distortion of the sorting measure.

We find increasing PAM while alternative weights and measures suggest flat or decreasing trends. Thus, both the sorting measure and the weighting scheme are important, and researchers should use weights like ours that are disciplined by the data.

References

Almar, Frederik, Benjamin Friedrich, Ana Reynoso, Bastian Schulz, and Rune Vejlin (2023) "Marital Sorting and Inequality: How Educational Categorization Matters," Discussion Paper 15912, IZA Institute of Labor Economics.

Breen, Richard and Leire Salazar (2011) "Educational Assortative Mating and Earnings Inequality in the United States," *American Journal of Sociology*, Vol. 117, pp. 808–843.

Chiappori, Pierre-Andre, Monica Costa-Dias, Sam Crossman, and Costas Meghir (2020a) "Changes in Assortative Matching and Inequality in Income: Evidence for the UK," *Fiscal Studies*, Vol. 41, pp. 39–63.

Chiappori, Pierre-Andre, Monica Costa Dias, and Costas Meghir (2020b) "Changes in Assortative Matching: Theory and Evidence for the US," Working Paper 26932, National Bureau of Economic Research.

——— (2021) "The measuring of assortativeness in marriage: a comment."

Eika, Lasse, Magne Mogstad, and Basit Zafar (2019) "Educational Assortative Mating and Household Income Inequality," *Journal of Political Economy*, Vol. 127, pp. 2795–2835.

Fernández, Raquel and Richard Rogerson (2001) "Sorting and Long-Run Inequality," *The Quarterly Journal of Economics*, Vol. 116, pp. 1305–1341.

Goldin, Claudia (2006) "The Quiet Revolution That Transformed Women's Employment, Education, and Family," *American Economic Review*, Vol. 96, pp. 1–21.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos (2016) "Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation," *American Economic Journal: Macroeconomics*, Vol. 8, pp. 1–41.

Kremer, Michael (1997) "How Much does Sorting Increase Inequality?," *The Quarterly Journal of Economics*, Vol. 112, pp. 115–139.

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

Statistics Denmark BEF (1980–2018) "BEF - Befolkningen, Statistics Denmark: Population and Education, Social statistics.," URL: https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/the-population.

Statistics Denmark UDDA (1980–2018) "UDDA - Uddannelser (BUE), Statistics Denmark: Population and Education, Social statistics.," URL: https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/highest-education-attained.

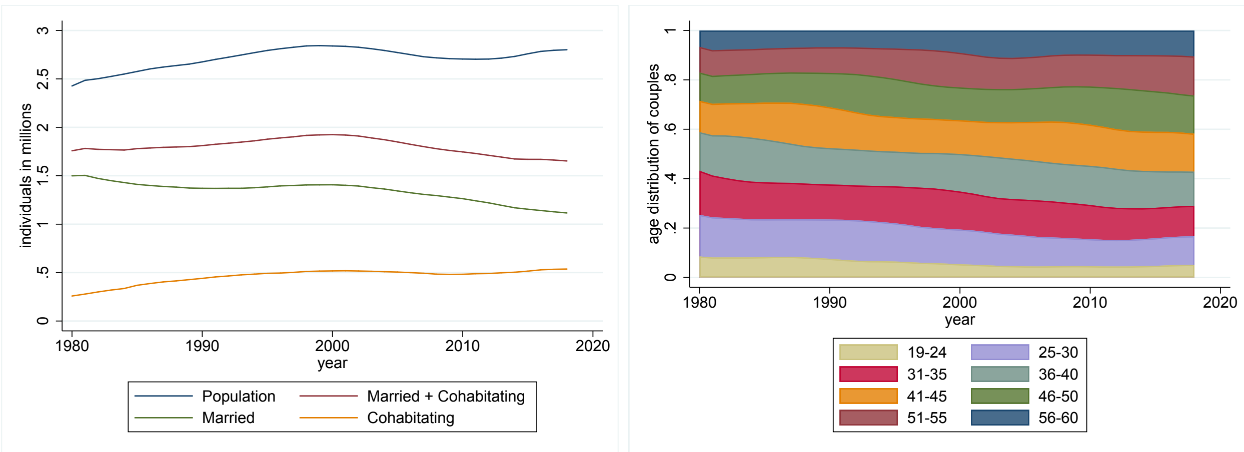
Online Appendix

(not for publication)

A Additional Results

Figure A.1: Marriage, Cohabitation, Age Composition

(a) Stocks of Couples (b) Age Composition of Couples



Note: Panel (a) reports the development in numbers of individuals by marital status. Panel (b) plots the age distribution of individuals who are either legally married or cohabiting. Section 2 explains how the sample and the education-based types are constructed.

B The correlation coefficient

Following Chiappori et al. (2021), the correlation coefficient in the 2×2 case as described in Table 1b can be written in the following way:

$$I_{corr} = \frac{ad - (m_{H,M} - a)(m_{H,F} - a)}{\sqrt{m_{H,M}(1 - m_{H,M})m_{H,F}(1 - m_{H,F})}}. \tag{A.1}$$

Applying the same approach as in Subsections 3.2 and 3.3, we totally differentiate (A.1), define $\Theta = \sqrt{m_{H,M}(1 - m_{H,M})m_{H,F}(1 - m_{H,F})}$, and obtain the following:

$$\Delta I_{corr} = \underbrace{\left(\frac{d-2a+m_{H,M}+m_{H,F}}{\Theta}\right)}_{>0}\Delta a + \underbrace{\left(\frac{a}{\Theta}\right)}_{>0}\Delta d$$
$$+ \underbrace{\left(\frac{-(m_{(H,F)}-a)\Theta - \frac{\sqrt{m_{(H,F)}}(1-2m_{(H,M)})\sqrt{1-m_{(H,F)}}}{2\sqrt{m_{(H,M)}}\sqrt{1-m_{(H,M)}}}(ad-(m_{(H,M)}-a)(m_{(H,F)}-a))}{\Theta}\right)}_{\gtreqless 0}\Delta m_{(H,M)}$$
$$+ \underbrace{\left(\frac{-(m_{(H,M)}-a)\Theta - \frac{\sqrt{m_{(H,M)}}(1-2m_{(H,F)})\sqrt{1-m_{(H,M)}}}{2\sqrt{m_{(H,F)}}\sqrt{1-m_{(H,F)}}}(ad-(m_{(H,M)}-a)(m_{(H,F)}-a))}{\Theta}\right)}_{\gtreqless 0}\Delta m_{(H,F)}. \tag{A.2}$$

As can be seen from Equation (A.2), the impacts of the changing marginal distributions $m_{(H,M)}$ and $m_{(H,F)}$ are ambiguous and highly nonlinear.

C Generalization

An advantage of the I_S^{convex} sorting measure is its generalizability to more than two types. Consider a marriage market with N types of males and N types of females. Table A.1 shows the generalized contingency table for this case.

Table A.1: Generalized contingency table

M\F	1	2	$\cdots$	$n-1$	n	Marginal
1	a_{11}	a_{12}	$\cdots$	$a_{1,n-1}$	$m_{1M}-\sum_{i=1}^{n-1} a_{1i}$	m_{1M}
2	a_{21}	a_{22}	$\cdots$	$a_{2,n-1}$	$m_{2M}-\sum_{i=1}^{n-1} a_{2i}$	m_{2M}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
$n-1$	$a_{n-1,1}$	$a_{n-1,2}$	$\cdots$	$\ddots$	$\vdots$	$m_{n-1,M}$
n	$m_{1F}-\sum_{i=1}^{n-1} a_{i1}$	$m_{2F}-\sum_{i=1}^{n-1} a_{i2}$	$\cdots$	$\cdots$	a_{nn}	$1-\sum_{i=1}^{n-1} m_{iM}$
Marginal	m_{1F}	m_{2F}	$\cdots$	$m_{n-1,F}$	$1-\sum_{i=1}^{n-1} m_{iF}$	1

The generalized version of I_S^{convex} can be written as follows:

$$I_S^{convex} = \sum_{i=1}^{n-1} \frac{a_{ii}}{m_{iM}m_{iF}}(\lambda m_{iM} + (1-\lambda)m_{iF})$$
$$+ \frac{a_{nn}}{(1-\sum_{i=1}^{n-1} m_{iM})(1-\sum_{i=1}^{n-1} m_{iF})}\left(\lambda\left(1-\sum_{i=1}^{n-1} m_{iM}\right) + (1-\lambda)\left(1-\sum_{i=1}^{n-1} m_{iF}\right)\right) \tag{A.3}$$

Like the 2×2 case in (2), we take the weighted sum over all the diagonal cells of the contingency table divided by the product of the respective male and female marginal distributions.

Next, we totally differentiate equation (A.3):

$$
\begin{aligned}
\Delta I_S^{convex} = &\sum_{k=1}^{n-1} \underbrace{\left(\frac{\lambda m_{kM} + (1-\lambda) m_{kF}}{m_{kM} m_{kF}}\right)}_{>0} \Delta a_{kk} \\
&+ \underbrace{\left(\frac{\lambda \left(1-\sum_{i=1}^{n-1} m_{iM}\right) + (1-\lambda)\left(1-\sum_{i=1}^{n-1} m_{iF}\right)}{\left(1-\sum_{i=1}^{n-1} m_{iM}\right)\left(1-\sum_{i=1}^{n-1} m_{iF}\right)}\right)}_{>0} \Delta a_{nn} \\
&+ (1-\lambda) \underbrace{\sum_{k=1}^{n-1} \left(\frac{a_{nn}}{\left(1-\sum_{i=1}^{n-1} m_{iM}\right)^2} - \frac{a_{kk}}{m_{kM}^2}\right) \Delta m_{kM}}_{=\xi_1 \gtreqless 0} \\
&+ \lambda \underbrace{\sum_{k=1}^{n-1} \left(\frac{a_{nn}}{\left(1-\sum_{i=1}^{n-1} m_{iF}\right)^2} - \frac{a_{kk}}{m_{kF}^2}\right) \Delta m_{kF}}_{=\xi_2 \gtreqless 0}.
\end{aligned}
\tag{A.4}
$$

We apply the same logic as in the 2×2 case (Equation (4)) to derive a decision rule for λ in the general case:

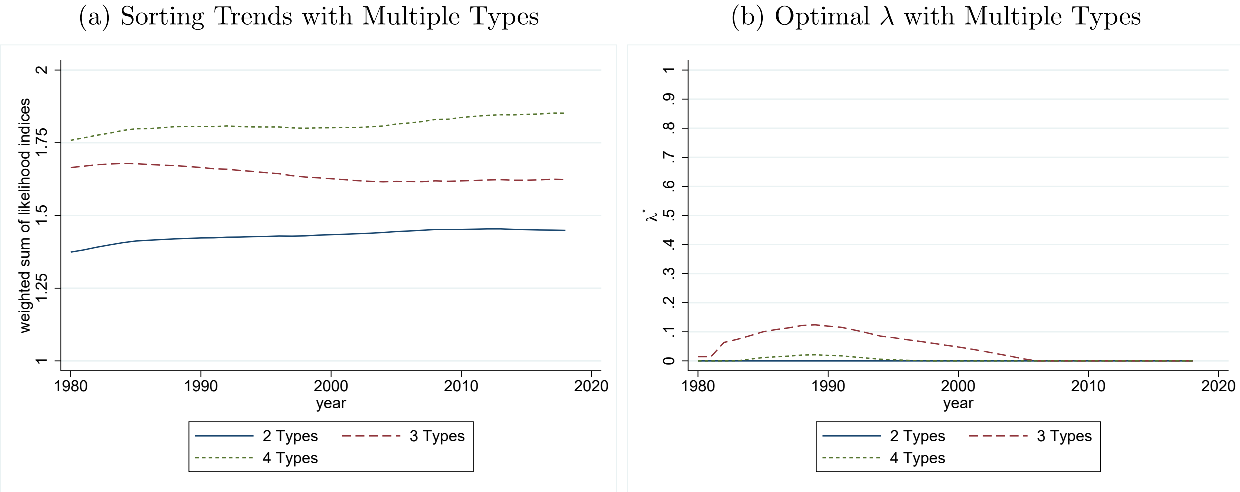
$$
\lambda^* = \begin{cases} 0 & \text{if } sign(\xi_1) = sign(\xi_2), \quad |\xi_1| < |\xi_2| \\ \frac{\xi_1}{\xi_1 - \xi_2} & \text{if } sign(\xi_1) \neq sign(\xi_2) \\ 1 & \text{if } sign(\xi_1) = sign(\xi_2), \quad |\xi_1| > |\xi_2| \end{cases}
\tag{A.5}
$$

Here, the objective function is a convex combination of the two endpoints ξ_1 and ξ_2 defined in Equation (A.4) which are equivalent to γ_1 and γ_2 in the 2×2 case.

In addition to λ^* and $I_S^{\lambda^*}$ for $N = 2$ ($H =$ Tertiary, $L =$ Non-tertiary), we compute the optimal weights and the sorting measure for $N = 3$ (Tertiary, Secondary, Primary), and $N = 4$ (Master/PhD, Bachelor, Secondary, Primary). We show the development of I_S^{convex} for all cases in figure A.2, Panel (a). The values for λ^* are shown in Panel (b).

Sorting is positive irrespective of the number of types because the indices in figure A.2a are always greater than one. For the two-type case (blue solid line), we get the exact same trend as in figure 2a with $\lambda = 0$ (green dash-dotted line) because the $\lambda^* = 0$ in the two-type case (blue solid line in figure A.2b). Interestingly, in the three-type case, we see a slightly decreasing extent of sorting (although almost flat) and some intermediate values for λ^* (red dashed lines in figure A.2). Hence, the increasing extent of sorting is not driven by sorting trends within the

Figure A.2: Marriage Market Sorting and Optimal Weights with Multiple Types



Note: Panel (a) shows the development of the sorting measure $I_S^{\lambda^*}$ for $N = 2$, $N = 3$, and $N = 4$. Panel (b) shows the values of λ^* according to the derived decision rule (A.5). Section 2 explains how the sample and the education-based types are constructed.

lower levels of education, i.e., primary and secondary education. Instead, increasing sorting is driven by more sorting among highly-educated individuals. This is evident from the increasing trend for the four-type case (green short-dashed line) in figure A.2a. In this case, λ^* is zero or close to zero throughout (see figure A.2b), which is consistent with the more pronounced changes in the female type distribution.

One could speculate that the higher level of sorting with more types is due to a mechanical effect of adding more types. This is not the case. The overall level of sorting increases because a higher extent of sorting among high types is uncovered with more granular types. To understand why there is no mechanical effect, consider the following example. Using the notation from the generalized contingency table A.1, in which the highest type has index 1, we compare the sorting measures for $N = 2$ and $N = 3$ cases under $\lambda^* = 0$

$$I_S^{convex}(N = 2) = \frac{a_{11}}{m_{1M}} + \frac{a_{22}}{1 - m_{1M}}, \tag{A.6}$$

and

$$I_S^{convex}(N = 3) = \frac{a_{11}}{m_{1M}} + \frac{a_{22}}{m_{2M}} + \frac{a_{33}}{1 - m_{1M} - m_{2M}}. \tag{A.7}$$

Let the top category be split up into two separate categories when we go from $N = 2$ to $N = 3$. The terms related to the bottom category remain unchanged, so the second terms in Equation (A.6) and the third term in Equation (A.7) can be ignored when comparing sorting levels. For $N = 2$, assume, as in Table A.2, that two-thirds of men are in the top category and that all top-category men are married to top-category women: Thus, ignoring the bottom

Table A.2: The 2×2 case - Example

Male\Female	1	2	Marginal
1	$^2/_3$	0	$^2/_3$

category, we get $I_S^{convex}(N=2) = \frac{a_{11}}{m_{1M}} = 1$. Now, what happens to I_S^{convex} when we split up the top category into two separate categories depends on the extent of sorting within the previous top category. Consider two cases. First, there could be no sorting within the previous top category, so the two-thirds of men would be uniformly distributed across matches in categories 1 and 2, see Panel (a) of Table A.3. However, there could also be perfect sorting within the

Table A.3: The 3×3 case - Example

(a) No sorting within

Male\Female	1	2	3	Marginal
1	$^1/_6$	$^1/_6$	0	$^1/_3$
2	$^1/_6$	$^1/_6$	0	$^1/_3$

(b) Perfect sorting within

Male\Female	1	2	3	Marginal
1	$^1/_3$	0	0	$^1/_3$
2	0	$^1/_3$	0	$^1/_3$

previous top category, see Panel (b) of Table A.3. In this case, the third of men in the new category 1 would be matched with category-1 women and the third of men in the new category 2 would be matched with category-2 women.

If no sorting is revealed within the top category, the sorting measures $I_S^{convex}(N=2)$ and $I_S^{convex}(N=3)$ are identical: $I_S^{convex}(N=3) = \frac{a_{11}}{m_{1M}} + \frac{a_{22}}{m_{2M}} = \frac{1}{2} + \frac{1}{2} = 1 = I_S^{convex}(N=2)$. This proves that there is no mechanical effect of adding another type that increases the sorting measure. In contrast, if perfect sorting is revealed within the top category, $I_S^{convex}(N=3)$ is twice as high as $I_S^{convex}(N=2)$: $I_S^{convex}(N=3) = \frac{a_{11}}{m_{1M}} + \frac{a_{22}}{m_{2M}} = 1+1 = 2$. The examples considered here are two extreme cases but they show that introducing another type can either leave the sorting measure unchanged or increase it, depending on the configuration of the contingency table and the sorting patterns within categories. In our setting, going from $N=2$ to $N=4$ indeed uncovers positive sorting within the tertiary-education subgroup, i.e., among graduates with Master/PhD degrees. This explains the ranking of sorting measures in Figure A.2a.

D Alternative Weights

In this appendix, we first show the equivalence of the weights used in Greenwood et al. (2016), Eika et al. (2019), and Almar et al. (2023). Second, we show how these alternative weights compare to the ones derived in this paper.

Greenwood et al. (2016) divide the sum of the diagonal elements (trace) of the matrix formed by the contingency table by the trace of the counterfactual matrix under random matching. They do not use an explicit weighting scheme. In our notation, their sorting measure is

$$I_S^{trace} = \frac{a+d}{m_{(H,M)}m_{(H,F)} + (1-m_{(H,M)})(1-m_{(H,F)})}. \tag{A.8}$$

We first show that I_S^{trace} is mathematically equivalent to the weighted sum of likelihood indices used in Eika et al. (2019) and Almar et al. (2023). In our notation and for the 2×2 case, their weights are

$$w_H = \frac{m_{(H,M)}m_{(H,F)}}{m_{(H,M)}m_{(H,F)} + m_{(L,M)}m_{(L,F)}} \tag{A.9}$$

and

$$w_L = \frac{m_{(L,M)}m_{(L,F)}}{m_{(H,M)}m_{(H,F)} + m_{(L,M)}m_{(L,F)}} = \frac{(1-m_{(H,M)})(1-m_{(H,F)})}{m_{(H,M)}m_{(H,F)} + m_{(L,M)}m_{(L,F)}}. \tag{A.10}$$

We plug these weights into the definition of the weighted sum of likelihood indices according to Equation (1). The products of the marginal distributions cancel out, and we are left with the sorting measure

$$\frac{a+d}{m_{(H,M)}m_{(H,F)} + m_{(L,M)}m_{(L,F)}} = \frac{a+d}{m_{(H,M)}m_{(H,F)} + (1-m_{(H,M)})(1-m_{(H,F)})} = I_S^{trace}, \tag{A.11}$$

which is exactly the Greenwood et al. (2016) measure. Although their weighting is not explicit, the random matching counterfactual in the denominator takes the marginal distributions and their changes over time into account.

Next, we rewrite this sorting measure as a weighted sum of likelihood indices with the weights of the same form as in equation (2):

$$\begin{aligned}
I_S^{trace} &= \frac{a}{m_{(H,M)}m_{(H,F)} + (1-m_{(H,M)})(1-m_{(H,F)})} \times \frac{m_{(H,M)}m_{(H,F)}}{m_{(H,M)}m_{(H,F)}} \\
&+ \frac{d}{m_{(H,M)}m_{(H,F)} + (1-m_{(H,M)})(1-m_{(H,F)})} \times \frac{(1-m_{(H,M)})(1-m_{(H,F)})}{(1-m_{(H,M)})(1-m_{(H,F)})} \\
&= \frac{a}{m_{(H,M)}m_{(H,F)}} \times \frac{m_{(H,M)}m_{(H,F)}}{m_{(H,M)}m_{(H,F)} + (1-m_{(H,M)})(1-m_{(H,F)})} \\
&+ \frac{d}{(1-m_{(H,M)})(1-m_{(H,F)})} \times \frac{(1-m_{(H,M)})(1-m_{(H,F)})}{m_{(H,M)}m_{(H,F)} + (1-m_{(H,M)})(1-m_{(H,F)})}.
\end{aligned}$$

This result can be generalized to the $n \times n$ case as described in Appendix C.

We now turn to comparing the weights implied by I_S^{trace} to the ones used in I_S^{convex}. The implied weights of $I_{\mathcal{S}}^{trace}$ are not necessarily a convex combination of the male and female marginals. To see this, we define λ^{trace} as the λ that equalizes $I_{\mathcal{S}}^{trace}$ and $I_{\mathcal{S}}^{convex}$:

$$\lambda^{trace} = \left(\frac{m_{(H,M)} m_{(H,F)}}{m_{(H,M)} m_{(H,F)} + (1 - m_{(H,M)})(1 - m_{(H,F)})} - m_{(H,F)} \right) \frac{1}{m_{(H,M)} - m_{(H,F)}}. \quad \text{(A.12)}$$

Evidently, λ^{trace} can lie outside the unit interval. In the data, $\lambda^{trace} < 0$ holds until 1987, which is the last year in which males had a higher share of high types than females ($m_{(H,M)} > m_{(H,F)}$). From 1988, $\lambda^{trace} > 1$ holds. For these values of λ, the effect of more sorted couples on the measure is not necessarily positive, see equation (3).

In our case, conclusions based on $I_{\mathcal{S}}^{convex}$ with λ^* and $I_{\mathcal{S}}^{trace}$ turn out to be similar, i.e., sorting is increasing, but this is coincidental and not guaranteed to hold in other settings. To scrutinize this finding, we compare the weights used in $I_{\mathcal{S}}^{\lambda^*}$ and $I_{\mathcal{S}}^{trace}$. $w_H^{\lambda^*} > w_H^{trace}$ holds, which implies that $I_{\mathcal{S}}^{\lambda^*} > I_{\mathcal{S}}^{trace}$. The reason is that the likelihood index for (H, H) couples is larger than for (L, L) couples, i.e., $\frac{a}{m_{(H,M)} m_{(H,F)}} > \frac{d}{(1-m_{(H,M)})(1-m_{(H,F)})}$. However, we see a stronger increase in $I_{\mathcal{S}}^{trace}$ compared to $I_{\mathcal{S}}^{\lambda^*}$ because the increase in w_H is larger for $I_{\mathcal{S}}^{trace}$, i.e., $0 < \Delta w_H^{\lambda^*} < \Delta w_H^{trace}$. Thus, both $I_{\mathcal{S}}^{trace}$ and $I_{\mathcal{S}}^{\lambda^*}$ increase over time.