FISCAL STUDIES, vol. 41, no. 1, pp. 39–63 (2020) 0143-5671

Changes in Assortative Matching and Inequality in Income: Evidence for the UK*

PIERRE-ANDRÉ CHIAPPORI,† MONICA COSTA-DIAS,‡
SAM CROSSMAN§ and COSTAS MEGHIR◊

†*Columbia University; NBER*
(pc2167@columbia.edu)
‡*Institute for Fiscal Studies; University of Porto; IZA; CEPR*
(monica_d@ifs.org.uk)
§*Institute for Fiscal Studies*
(sam_c@ifs.org.uk)
◊*Yale University; NBER; IFS; IZA; CEPR*
(c.meghir@yale.edu)

Abstract

The extent to which like-with-like marry is important for inequality as well as for the outcomes of children who result from the union. In this paper, we present evidence on changes in assortative mating and its implications for household inequality in the UK. Our approach contrasts with others in the literature in that it is consistent with an underlying model of the marriage market. We argue that a key advantage of this approach is that it creates a direct connection between changes in assortativeness in marriage and changes in the value of marriage

* Submitted December 2019.

The authors are grateful to Bernard Salanié, the Editor Jim Ziliak, two anonymous referees and participants in seminars and conferences at Stony Brook University, University of Pennsylvania, University of Oxford, University of Essex and University of Bristol for insightful comments and discussions. Financial support from the Economic and Social Research Council (ESRC Centre for the Microeconomic Analysis of Public Policy at IFS – grant number ES/M010147/1 – and grant numbers ES/K00624X/1 and ES/N015304/1) is gratefully acknowledged. Costas Meghir also thanks the Cowles Foundation and the ISPS at Yale for financial support. All errors are the authors' own.

Keywords: marital sorting, income inequality.

JEL classification numbers: D31, J12.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

for the various possible matches by education group. Our empirical results do not show a clear direction of change in assortativeness in the UK between the birth cohorts of 1945–54 and 1965–74. We find that changes in assortativeness pushed income inequality up slightly, but that the strong changes in education attainment across the two cohorts contributed to scale down inequality.

I. Introduction

Over the past 40 years, inequality has been rising in the US, the UK and other developed countries. Understanding the causes of such changes can help uncover the underlying social trends and even shape policy responses. One of the important factors driving inequality is sorting in the marriage market. Sorting refers to the extent to which like marries like in certain dimensions, such as education. Sorting itself is likely driven by the potential social and economic benefits of marrying someone similar to or, alternatively, very different from oneself.

Becker (1981) has noted that if partners specialise in different activities such as housework and labour market work, there will be a tendency toward negative sorting, where one partner is better educated and works in the paid labour market while the other has a lower skill level as required by housework. On the other hand, if two better-educated spouses can together produce children of higher skills and ability, there will be strong incentives to sort positively, where like marries like. Within this insight lies both the increased social tendency to marry within education groups and the consequence of such a tendency – namely, increased inequality in intergenerational outcomes: at a time when the returns to education and human capital have increased, making the returns to investing in children ever higher, the push towards a social equilibrium with more and more positive sorting becomes ever more powerful and consequential. Other factors that can have large impacts on sorting include household technology, relative labour market opportunities for men and women, and divorce laws.[1] Thus just understanding the direction of sorting is important in itself, as argued by Eika, Mogstad and Zafar (2019).

Detecting sorting, and particularly positive assortative matching, is not particularly complicated; however, quantifying the degree of (and measuring changes in) sorting is challenging. Indeed, the literature includes different measures that do not necessarily provide the same answers. As discussed in Chiappori, Costa-Dias and Meghir (2020), changes in sorting relate to underlying changes in the relative benefit of marrying the same type of person (say in educational attainment and/or ability) compared with marrying different types – for example, how much more beneficial is it now for college graduates to marry each other relative to marrying across education groups? When all we

[1]Reynoso, 2019.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

observe are the frequencies with which individuals marry within and across groups, we need to find a way to convert changes in these frequencies to changes in values. The problem is relatively easy when the proportions of people obtaining each level of education do not vary over the period we are making the comparison. When these proportions change, however, things become particularly hard, because one must disentangle the mechanical effects of these changes from possible structural variations in the underlying values and benefits. A model-free approach to answering this question is generally not available.

In this paper, we analyse changes in sorting in the marriage market in the UK. For this purpose, we use an index derived by Chiappori, Costa-Dias and Meghir (2020), which in turn is based on the model of Choo and Siow (2006), to study changes in sorting in the UK and the roles of marital matching and education attainment in driving changes in earnings inequality.

II. Assortative matching

We will consider sorting in the marriage market by education, asking whether it is the case that like tends to marry like and whether such tendency has changed over time. While ascertaining whether a set of matches is assortative in education or not is simple and very intuitive, establishing whether one set of matches is *more* assortative than another generates unexpected difficulties when the proportions of men and women attaining various education levels are different across the groups or time periods being compared. These difficulties are far from innocuous and it is fair to say that there is (still) some ambiguity over the exact meaning of a claim such as 'assortative matching has increased over a given period'. Chiappori, Costa-Dias and Meghir (2020) discuss the issues that arise with the measurement of *changes* in assortative matching. The key point is that ambiguity arises unless we provide a direct link between changes in assortative matching and changes in the value of marriage. In this paper, we develop a structural method for assessing changes in assortativeness, and we show how to implement it empirically based on the framework of Choo and Siow (2006).[2]

To define assortativeness, it is convenient to start with the simple example of an economy in which men and women can only belong to two education categories: college graduate (C) and less than college, which we label as high school (HS). Conditional on marriage, the matching patterns for this population can be fully summarised by Table 1.

[2]In Chiappori, Costa-Dias and Meghir (2020), we discuss various alternatives to measuring changes in assortative matching, including that by Fernández and Rogerson (2001), Greenwood, Guner and Knowles (2003), Greenwood et al. (2014) and Shen (2019).

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 1

Matching patterns in a marriage market with two types, conditional on marriage

Women: / *Men*	C	HS
C	r	$m - r$
HS	$n - r$	$1 - n - m + r$

TABLE 2

Random matching in a marriage market with two types, conditional on marriage

Women: / *Men*	C	HS
C	mn	$m(1 - n)$
HS	$n(1 - m)$	$(1 - n)(1 - m)$

In Table 1, m and n are the proportions of college graduate men and women respectively and r is the proportion of couples where spouses are both college graduates. The example may seem excessively simple, if only because it only considers two education classes. However, it turns out that the measures of sorting will always have to be local rather than considering all education categories together. So, for example, we will be able to answer the question on sorting and changes thereof in the group of college graduates and high-school graduates, or high-school graduates and high-school dropouts. The overall picture can then be pieced together from these local comparisons. Indeed, it is easy to provide examples where assortative matching is positive (or increasing) at one end of the distribution while negative (or decreasing) at the other end. This implies that most measures will have to be *local*, in the sense that they consider a (series of) 2×2 case(s) similar to Table 1. If all possible such cases imply positive sorting, then we can say that matching is globally assortative. Only very simple models will imply perfect sorting and they typically neglect a number of important features of reality such as preferences for alternative matches beyond the economic benefits.

Defining positive assortative matching

Positive assortativeness is defined in relation to random matching. For the same population of college and non-college educated as above, random matching holds if the proportion of couples where both spouses are college educated equals mn. In such a case, the sorting patterns are as summarised in Table 2.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

We then say that the matching patterns in Table 1 exhibit positive assortativeness when the *proportion of couples with equal education* (the 'diagonal' of the table) *is larger than what would obtain under random matching* (the 'diagonal' of Table 2). Hence, we have positive assortative matching if and only if

$$(1) \qquad r \geq mn,$$

which also implies that the frequency of couples where both spouses have less than college education (HS) exceeds what one would expect from random matching, i.e.

$$(2) \qquad 1 - n - m + r \geq (1 - n)(1 - m).$$

In other words, positive assortativeness arises where extra forces generate more matches between equally educated people than would happen for random reasons.

Much more challenging is the *comparison* of assortative matching across two different tables, each representing an alternative society or different periods. The typical approach in the literature is to use an index to quantify and compare assortativeness. A few alternative such indices have been proposed, but it remains unclear how to best assert that one matching table 'displays more assortative matching' than the other. Chiappori, Costa-Dias and Meghir (2020) provide a comprehensive discussion of the issues. Here we use one of the criteria derived in that paper – namely, the separable extreme value (SEV) index that we discuss below.

III. Measuring assortative matching

We propose a measure of assortativeness that relies on a structural interpretation of sorting by relating the observed sorting patterns to their underlying determinants – namely, the surplus of the match, which defines and quantifies the extent to which two people gain by matching together relative to remaining single or matching with others. In this section, we start by deriving our structural measure of assortativeness and then develop its empirical counterpart within the Choo and Siow (2006) framework.

1. A structural index of assortativeness

We consider a population of men and women, indexed by i and j respectively, who meet in the marriage market. In our framework, people match on a trait that produces value in marriage, such as education, ability or skills. We denote by X_i and Y_j the relevant trait of man i and woman j, respectively.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 3

Surplus matrix in a marriage market with two types

Women: / *Men*	C	HS
C	$Z^{C,C}$	$Z^{C,HS}$
HS	$Z^{HS,C}$	$Z^{HS,HS}$

In a transferable utility framework, the economic value produced by the match (i, j) can be defined as a function $Z(X_i, Y_j)$. Z is measured in relation to the outside option of not getting married, and hence quantifies the economic surplus from marriage. It is precisely the structure of Z that determines sorting in marriage. Specifically, if the value of increasing a male characteristic (such as his education) is itself increasing when the corresponding female characteristic increases (her education), it generates an incentive to sort positively on education: more-educated individuals gain more than their less-educated peers from matching with more-educated spouses. In this case, we call the surplus Z 'supermodular'.

The concept of supermodularity is the basis for our index of assortativeness. Without loss of generality, we discuss it in the context of the college and high-school marriage market example presented above, where X and Y stand for the education of the spouses and assume values C and HS. The corresponding structural surplus matrix is represented in Table 3, where the superscripts I, J (for $I, J = C, HS$) denote the education of the husband and wife, respectively.

We now say that the surplus function Z is supermodular if this structural surplus matrix has a positive supermodular core, SM, which is simply defined as

$$SM = Z^{HS,HS} + Z^{C,C} - Z^{C,HS} - Z^{HS,C}. \tag{3}$$

The condition $SM > 0$ can be rearranged to yield

$$Z^{C,C} - Z^{C,HS} > Z^{HS,C} - Z^{HS,HS}, \tag{4}$$

which precisely expresses the fact that the incremental value of marrying a more-educated woman increases with the education of the man. Many factors may lead to such complementarity in the traits of spouses. For example, a father with a college degree may be more valuable in bringing up children when his partner is equally educated, compared with being less educated, if there are complementarities in the education of parents for raising children.

SM is the structural index of assortativeness we propose to use. It plays a crucial role in the analysis of assortativeness. Indeed, a standard theoretical

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

result of the matching literature is that assortative matching is related to supermodularity of the surplus function.[3] Specifically, a strictly positive core, $SM > 0$, reflects a tendency for positive assortativeness, and the more positive the core the stronger is such tendency. However, SM is impractical since the values of marrying different types are not observed. We will now show how to use the framework of Choo and Siow (2006) to derive an empirical counterpart of our index.

2. The separable extreme value index of assortativeness

The seminal contribution of Choo and Siow (2006) explicitly relates the economic forces for sorting under transferable utility – namely, the marital surplus – to the observed sorting patterns. In this framework, people match on both observed and unobserved characteristics in a frictionless marriage market. The resulting matches, together with the allocation of the value produced by each match, are the outcome of a stable equilibrium in which no matched individual would be better off by remaining single, and no two individuals would be better off by matching to each other rather than staying with their current match (or remaining single).

Our separable extreme value (SEV) model draws from Choo and Siow (2006). We assume that the marriage market is populated by a large number of men and women who are classified into a small number N of observed classes, such as education groups. Let I and J denote the class of man i and woman j respectively, with $I, J = 1, \ldots, N$.

A central assumption in this model is that the surplus generated by matching man i in class I with woman j in class J, s_{ij}, is the sum of two terms:

$$s_{ij} = Z^{IJ} + \gamma_{ij}, \tag{5}$$

where Z is the systematic influence of the observed traits on matching patterns and γ_{ij} is a random term reflecting unobserved heterogeneity. γ_{ij} represents the preference that the individuals have for each other beyond the value of the match that is driven by observed characteristics. Moreover, the random term γ_{ij} is assumed to be separable in the following sense:

$$\gamma_{ij} = \varepsilon_i^J + \omega_j^I, \tag{6}$$

where $(\varepsilon_i^J, \omega_j^I)$ are idiosyncratic random variables that only depend on the class of the spouse. Intuitively, ε_i^J can be interpreted as the specific preference of man i for marrying a woman in class J or, conversely, as the common preferences of women in class J for man i.

[3]Chiappori, 2017.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

What matters to the individual deciding whether and who to marry is the *share* of the surplus they will obtain as a result of marriage.[4] In equilibrium, these shares are endogenously determined to clear the marriage market. In a transferable utility framework, they add up to the total surplus s. We denote by u_i and v_j the values of marriage for man i and woman j. If they marry together in a stable match, it must be that $s_{ij} = u_i + v_j$.

The key insight of the SEV approach is that only the sharing rule over the economic part of the surplus matters for assortativeness in marriage.[5] Specifically, it can be shown that there exists a collection of values U^{IJ} and V^{IJ} that add up to the economic surplus Z^{IJ} for all possible matching combinations, and such that the individual utilities of man i and woman j married to each other in the stable equilibrium are, respectively,

$$u_i = U^{IJ} + \varepsilon_i^J \tag{7}$$

and

$$v_i = V^{IJ} + \omega_j^I. \tag{8}$$

In the above, U^{IJ} is the part of the deterministic surplus obtained by a man of type I marrying a woman of type J, and V^{IJ} is the part of the surplus that goes to her. For completeness, the spouse's class is set to zero among singles, so the economic values of remaining single are U^{I0} and V^{0J} for men and women, respectively. Consistently with our definition of marital surplus in relation to the outside option of not marrying, these are set to zero. Below, we explain that one such normalisation is needed for identification.

Given the structure of the individual problem, we can now establish that for a man of type I to marry a woman of type J, it must be that

$$U^{IJ} + \varepsilon_i^J \geq U^{IS} + \varepsilon_i^S \quad \text{for all } S = 0, \ldots, N. \tag{9}$$

In the SEV model, the random components of the value of marriage (ε, ω) are assumed to be extreme value distributed, which implies that the probability of a man of type I marrying a woman of type J, conditional on his type, is

$$P^{IJ} = \frac{\exp(U^{IJ})}{\sum_{S=0}^{N} \exp(U^{IS})}. \tag{10}$$

[4]One may wonder where attraction and love fit in this framework. They clearly play a role in determining which specific people will come together, but we assume this occurs after matching by class has been determined.

[5]See Choo and Siow (2006), Chiappori, Salanié and Weiss (2017) and Chiappori, Costa-Dias and Meghir (2018).

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

Equivalently, we can define the conditional probabilities of women of type J marrying men of type I as Q^{IJ}. Thus, for example, the proportion of total marriages where both are college graduates (r in the notation of Table 1) is given by

$$(11) \qquad r = m \times P^{C,C} = n \times Q^{C,C}.$$

Using these probabilities, we can recover the utilities of both men and women (U^{IJ} and V^{IJ}) and hence the surplus. The probabilities (P^{IJ}, Q^{IJ}) are observed from the matching patterns in the data. For example, in our simple example with two types of men and women, take the case of college-educated men. We observe three probabilities including the one for being single, ($P^{C,0}$, $P^{C,HS}$, $P^{C,C}$); these add up to 1. There are three related values, including the one for being single ($U^{C,0}$, $U^{C,HS}$, $U^{C,C}$). These make a system of three equations similar to equation 10 and three unknowns, which are the three values. But since only two of the equations are independent as the three conditional probabilities sum to 1, we can only identify the values of marrying different spouses for each partner *relative* to the value of being single, which we normalise to zero without loss of generality. Thus, just as in McFadden (1981), the choice probabilities reveal the values of the marital options relative to the baseline of remaining single, which is what is needed to establish changes in sorting. Adding these up provides us with the surplus, whose structure determines the extent of sorting.

Once the surplus of the match is identified, we can use our structural index in equation 3 to quantify assortativeness. In the context of our example, the link between the observed frequencies and the supermodular core of the surplus matrix is

$$(12) \qquad \begin{aligned} SM &= Z^{HS,HS} + Z^{C,C} - Z^{HS,C} - Z^{C,HS} \\ &= 2\ln\left(\frac{r\,(1+r-m-n)}{(n-r)\,(m-r)}\right), \end{aligned}$$

where r, m and n are defined in Table 1. We therefore define the SEV index of assortativeness, Φ_{SEV}, below. The point now is that whenever Φ_{SEV} increases, sorting also increases.

Definition.[6] The SEV assortativeness index Φ_{SEV} corresponding to matching Table 1 is defined by

$$(13) \qquad \Phi_{SEV}\,(m, n, r) = \ln\left(\frac{r\,(1+r-m-n)}{(n-r)\,(m-r)}\right).$$

[6]Chiappori, Costa-Dias and Meghir, 2020.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

It is worth highlighting that our structural approach deals explicitly with changes in the proportion of singles across groups or cohorts as it relates the unconditional (on marriage) probabilities of marrying different types of spouses to the underlying value of these marriages *relative to the value of remaining single*. Hence, changes in marital surplus will reflect both changes in the probability of marriage and changes in marital sorting conditional on marriage. The extreme value assumption underlying our index, however, allows us to quantify assortativeness using only marital patterns conditional on marriage. This is a reflection of the 'independence of irrelevant alternatives' property of the multinomial logit model, and would not hold under alternative distributional assumptions on the random terms (ε, ω). In our context, it implies that changes in the proportion of singles that do not show in the sorting patterns *conditional on marriage* must mean that the systematic value of marrying a spouse of one type relative to another type remains unaltered.

Beyond the direct link with the underlying causes of sorting (namely, the structure of the gains from marriage), this index satisfies two intuitive properties: (a) whenever the proportions obtaining a college degree (m, n) remain the same but the proportion r of college-to-college marriages increases, the index grows; and (b) the index can never be larger for any matching pattern than when everyone is married to a person of their own type, in which case matching is perfectly assortative. In the case of SEV, the value of the index under perfectly assortative matching is $+\infty$. It turns out that these properties are key for any index of assortativeness to satisfy if it is to avoid inconsistencies, such as declaring a matching table with non-zero off-diagonal elements as more assortative than one with no off-diagonal matches.[7]

IV. The data

We use data from the UK Labour Force Survey[8] (LFS) to bring evidence on the changes in assortativeness in marriage by education and of their role in explaining changes in earnings inequality over time. The LFS is the largest representative household survey in the UK, collecting quarterly data on a rotating panel of households since 1992. Each household is interviewed for up to five quarters before being dropped and replaced by a new household. In each quarter, the survey records detailed demographic and socio-economic information for all adult household members, including date of birth, gender, education and employment status. Crucially, the LFS contains family identifiers and reports the relationship between all members of the household, which allows us to link couples. Employment earnings information has also been collected since 1993, but only on quarterly waves 1 and 5. Our sample period therefore runs from 1993 till the first quarter of 2019.

[7]Chiappori, Costa-Dias and Meghir, 2020.
[8]Office for National Statistics, 2019.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 4

Sample sizes: individuals aged 40–49 by cohort and gender

	Men	*Women*	*All*
Entire sample	142,359	154,906	297,265
Birth cohort 1945–54	45,337	47,538	92,875
Birth cohort 1965–74	39,627	44,307	83,934

In this paper, we use the unrestricted version of the LFS, which does not support the longitudinal linkage of individuals or families. To avoid having repeated observations on some but not all families, which we cannot link because a time-consistent family identifier is missing from our data, we keep only one observation per family. We select the observation for the wave 5 interview, one of the two that measure earnings.[9] Furthermore, we drop all observations of individuals who are not the head of the household or his/her spouse. We also drop same-sex couples, which constitute a small subset of our sample (less than 1 per cent of the families).

Ideally, we would like to observe the long-lasting marital circumstances and permanent income of each individual in our sample as these are closer reflections of their economic well-being. In practice, however, we see a single snapshot of the life of each respondent. To capture the time changes in the persistent family circumstances, our focus is on adults in the middle of their working life, between the ages of 40 and 49. The following considerations drive this choice. First, marital rates have a strong positive age gradient during early adult years before they flatten out as individual marital status settles. This means that while young adults may be observed as singles, many will move on to marry in the future. Using information on their marital status would produce downward-biased estimates of marital rates. Second, earned income grows and is more volatile at the start of working life, as individuals move up their careers while also facing more precarious working conditions. This implies that the earnings of younger adults are a noisier measure of their permanent income than those observed further down the line. And third, observations in later stages of life are more likely to reflect the declines in earnings that accompany retirement choices or the onset of health problems. In contrast, most marital choices are complete by age 40, and earnings plateau over the following decade of life.

Overall, our sample has nearly 300,000 individuals, with more women than men. For studying changes in assortativeness by education and their role in changing inequality in earnings, we will focus on two cohorts born about 20 years apart, in 1945–54 and 1965–74. Sample sizes for each of these groups are shown in Table 4.

[9]We prefer wave 5 to wave 1 because it is less afflicted by non-responses in the earnings variable. Our results are not affected by this choice.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

V. Definition of the main variables and descriptive statistics

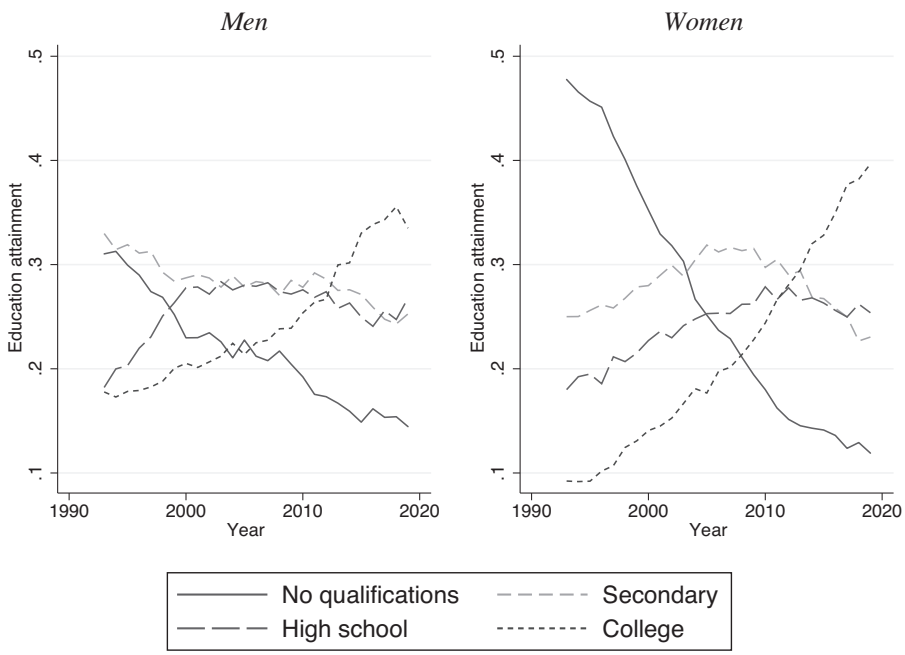
1. Education

We look at sorting patterns by education using a four-category measure. Individuals are assigned to one of the following groups, according to the highest completed qualification they acquired: 'no qualifications', for those leaving school prior to completing secondary-school qualifications (known as GCSEs in the UK); 'secondary' education, for those leaving school by the age of 16 with GCSE qualifications; 'high school' qualifications or equivalent, generally completed at age 18 (A levels in the UK); and 'college' degree (three years or more).

Figure 1 shows how education attainment has changed over time for men and women in the 40–49 age range. It displays some quite dramatic shifts from the bottom education group to the top, particularly for women. It also demonstrates that the distribution of education is much more balanced across genders today than it used to be, as women have caught up with men over the years. For the two cohorts we are comparing, the proportion of

FIGURE 1

Distribution of education, 40- to 49-year-olds by gender



© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

college-educated women almost tripled, from 11 per cent in the earlier cohort to 30 per cent in the later one, while the proportion leaving school without qualifications shrank from 43 per cent to 15 per cent of the cohort size. For men over the same period, college graduation rates increased from 19 per cent to 30 per cent, while the proportion without formal qualifications dropped from 28 per cent to 17 per cent.

These changes imply that direct comparisons of sorting patterns across cohorts do not reveal the direction of change in assortativeness in marriage. Indeed, in a purely mechanical sense, we expect to see more homogamous couples when the distributions of education for men and women become more similar, even if assortativeness in education does not increase.

2. Marriage

One practical issue is how to define marriage. The LFS does not collect historical demographic information, so we study marital status at the time of the interview. Given the increasing prevalence of cohabitation, we study both legal marriages and all unions. In the former case, cohabiting spouses are considered as separate single individuals. In both cases, divorced and widowed individuals are counted as singles. Figure 2 shows the proportions of married individuals by year, education and gender. There is a steep decline in marital rates over time, particularly for legal marriages and for those with lower levels of formal education. Hence, it is not only the case that individuals are getting more educated over time, but also that marriage is becoming relatively less common among those with lower levels of education. We will be able to account for changes in the composition of married and single families, as well as for changes in sorting patterns conditional on marriage, when assessing the role of marriage in driving changes in inequality.

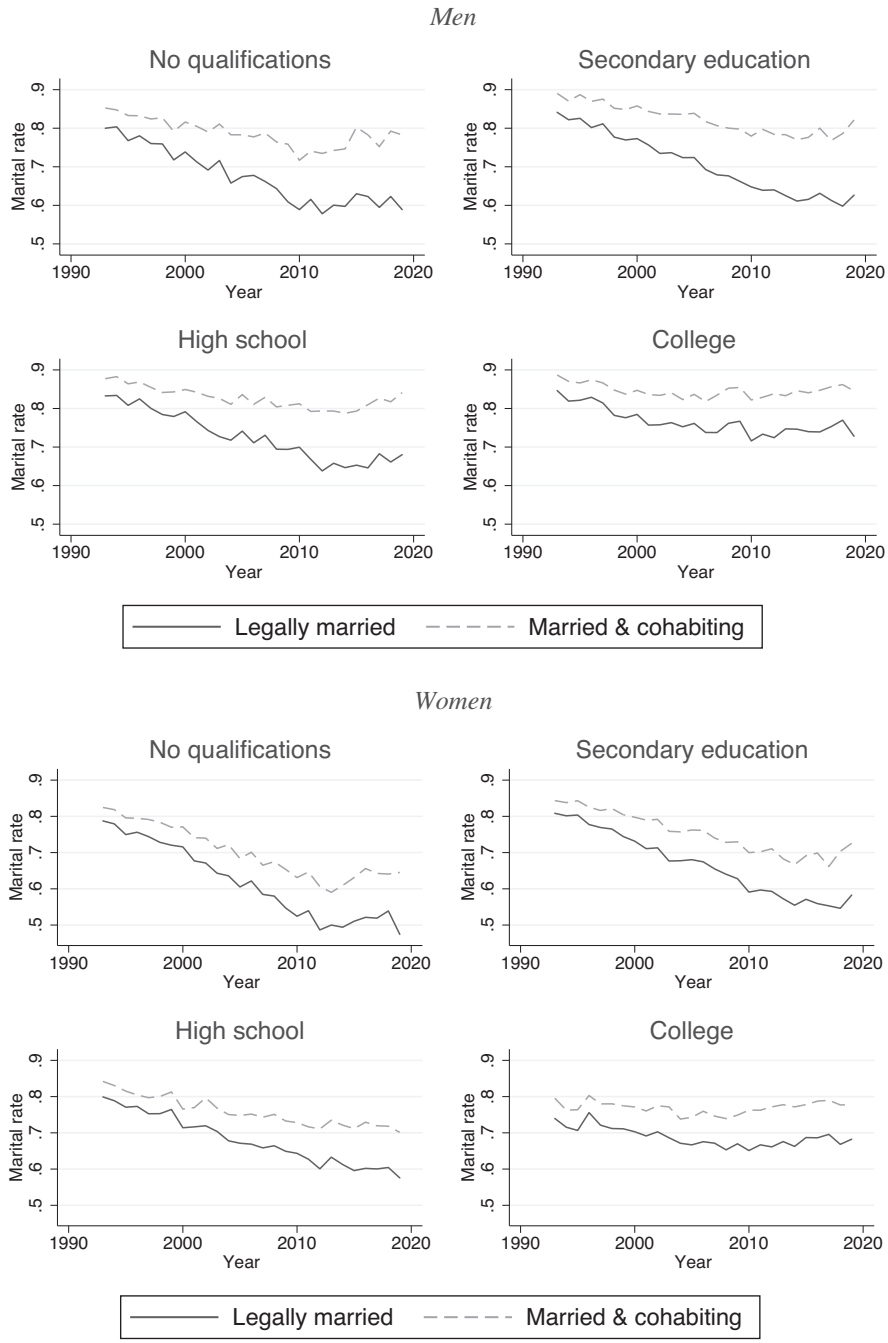
It is also clear from the lines in Figure 2 that men are more likely to be found in marriage than women. Census data for the same period reveal that women and men of this age group are nearly equally populous. In our sample, however, the number of women exceeds the number of men by over 8 per cent (see Table 4). The entirety of this difference is loaded on the population of singles, since both spouses are observed in couples. We attribute this difference to some uneven success in sampling single men and women on the ground. To make up for it, we reweight our sample to ensure that both genders are equally represented in each cohort when studying sorting patterns.

When counting families instead of individuals, we consider all couples with at least one spouse from the relevant birth cohort and observed aged 40–49. In all cases, single individuals are only included in the sample if they belong to

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

FIGURE 2

Marital rates, 40- to 49-year-olds by gender and education



© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

the relevant age range and birth cohort. Based on this definition, the proportion of cohabiting couples increased from 7 per cent of all couples for birth cohort 1945–54 to 17 per cent for birth cohort 1965–74.

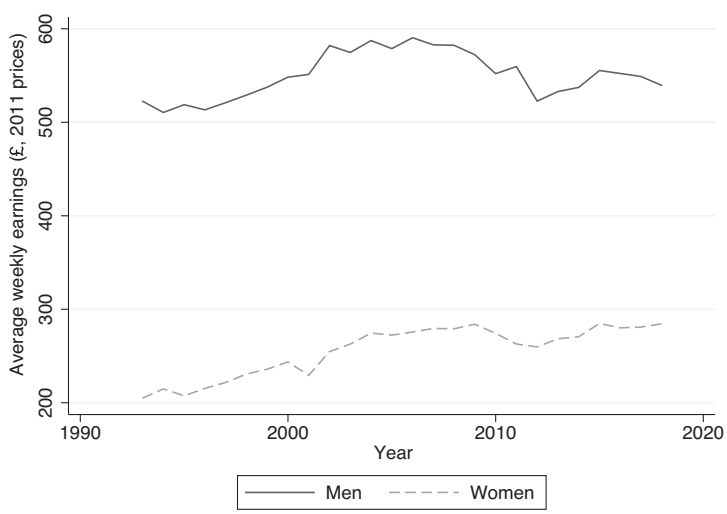
3. Earned income

The LFS collects gross weekly earnings for all employees, but it does not include income information for the self-employed. We keep the self-employed in our sample for studying sorting patterns, but do not consider families with self-employed individuals when studying income inequality. Our measure of family income is the pooled weekly earnings of both spouses divided by 2 for couples, or the weekly earnings of the individual for singles. Earnings of non-working individuals are set to zero. We deflate income using the Consumer Prices Index (CPI) and trim the top 0.1 per cent of its distribution.

Figure 3 shows that the average weekly earnings of men and women increased modestly over this period. The gains are more noticeable for women, as men have lost earlier gains in the years since the recession. Figure 4 splits the groups by education. It shows no gains in real earnings for any of the education groups. On the contrary, it shows a generalised decline in earnings that is more pronounced for educated men. All aggregate gains in earnings can therefore be attributed to the gradual increase in education attainment of this age group.

FIGURE 3

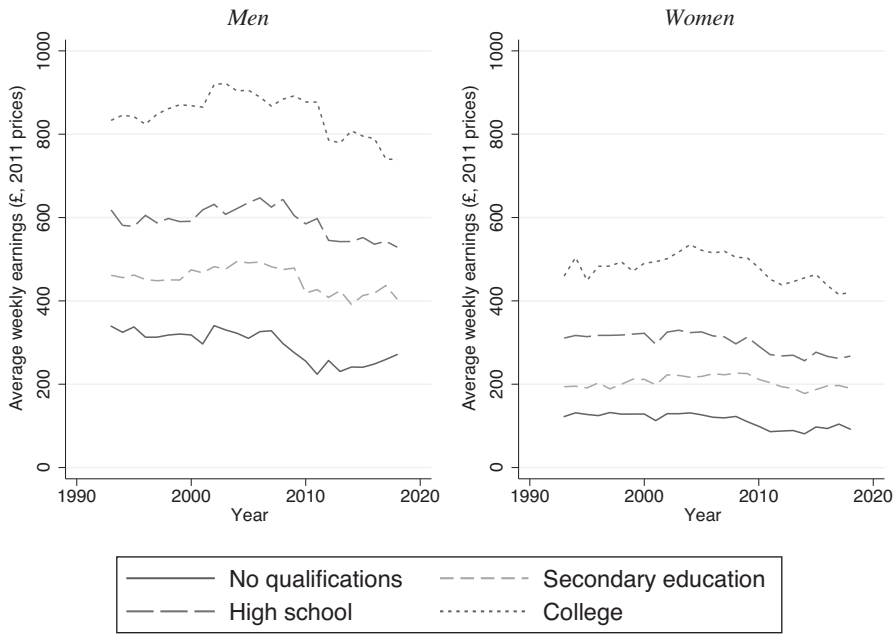
Average real weekly earnings, 40- to 49-year-olds by gender



© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

FIGURE 4

Average real weekly earnings, 40- to 49-year-olds by gender and education



VI. Education and marital sorting

Table 5 shows the sorting patterns conditional on union for the two cohorts we are considering. Here we use the broader definition of couples that includes those in legal unions and those cohabiting. Figures for the case where cohabiting spouses are treated as single individuals can be found in the online appendix and show similar patterns.

The table shows signs of positive assortative matching for both cohorts, with a disproportionate share of unions happening along the main diagonal. For instance, under random matching, we would expect about 11.9 per cent of the unions in the earlier cohort to be between men and women with no formal qualifications (the product of the proportion of partnered women with no qualifications by the proportion of partnered men without qualifications), while the data reveal that the size of this cell in Panel A is in fact almost 18.5 per cent of the population of couples. The table also reveals a sharp increase in the proportion of unions between equally educated individuals in the top half of the education distribution, from the earlier to the later cohort, with a compensating drop in the proportion of homogamous low-educated couples. This, of course, is partly the result of the shift in the distribution of

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 5

Sorting matrices conditional on union (married and cohabiting couples, %)

Man's education	*Woman's education*			
	No qual.	Secondary	High school	College
		Panel A. Birth cohort 1945–54		
No qualifications	18.46	6.64	3.07	0.72
Secondary	14.07	10.51	5.44	1.28
High school	6.59	6.97	6.23	2.11
College	2.23	3.31	5.81	6.57
		Panel B. Birth cohort 1965–74		
No qualifications	6.30	5.20	3.70	2.04
Secondary	3.85	10.65	8.09	4.42
High school	2.44	7.61	9.06	7.31
College	1.41	3.43	6.22	18.25

education across cohorts, particularly among women, towards higher levels of education.

What we do now is assess whether these changes in sorting patterns across cohorts can be reliably attributed to changes in assortative behaviour. We proceed by applying the SEV approach developed in Section III.2 to bring new evidence on the nature of changes in assortativeness in the UK.

To keep our results concise while bringing evidence to bear on the most relevant sorting patterns, we focus on sorting along the main diagonal. Given the patterns of positive assortative matching revealed in Table 5, the main diagonal is the key part of the sorting matrix where most of the mass of unions is concentrated. In practice, this means that we will study the full subset of 2×2 submatrices that keep homogamous matches in the main diagonal. To be precise, for the following sorting matrix over four levels of education (1 to 4):

$m \backslash w$	1	2	3	4
1	s_{11}	s_{12}	s_{13}	s_{14}
2	s_{21}	s_{22}	s_{23}	s_{24}
3	s_{31}	s_{32}	s_{33}	s_{34}
4	s_{41}	s_{42}	s_{43}	s_{44}

we would study changes in assortativeness for each of the following submatrices:

$$S_{ij} = \begin{array}{c|cc} m \backslash w & i & j \\ \hline i & s_{ii} & s_{ij} \\ j & s_{ji} & s_{jj} \end{array} \quad \text{for } (i, j) = (1, 2), (2, 3), (3, 4), (1, 3), (2, 4), (1, 4).$$

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 6

Difference in the SEV index between the 1965–74 and 1945–54 cohorts

	Submatrix					
	S_{12} No qual. Secondary *(1)*	S_{23} Secondary High school *(2)*	S_{34} High school College *(3)*	S_{13} No qual. High school *(4)*	S_{24} Secondary College *(5)*	S_{14} No qual. College *(6)*
	Panel A. Married and cohabiting couples					
Difference	0.484	–0.106	0.123	0.120	–0.288	–0.708
Std error	0.048	0.047	0.057	0.061	0.066	0.085
p-value	0.000	0.067	0.059	0.051	0.000	0.000
	Panel B. Legally married couples					
Difference	0.579	–0.076	0.150	0.171	–0.269	–0.717
Std error	0.054	0.051	0.059	0.065	0.069	0.094
p-value	0.000	0.135	0.019	0.025	0.000	0.000

Note: Labels in the top row identify the education levels in the respective submatrix. Figures in the first row in each panel show estimates of the difference in the SEV criterion between the later and earlier cohorts. The next two rows show the bootstrapped standard errors based on 1,000 repetitions and the *p*-values for two-sided significance testing adjusted for multiple hypotheses.

The study of these submatrices brings evidence on changes in the proportion of homogamous couples relative to each of the alternative matches.

Table 6 reports the differences across the two most distant cohorts in the supermodular core of all submatrices that keep homogamous unions in the main diagonal – there are six of these submatrices for the four education groups that we are considering. The table also reports the bootstrapped standard error of the difference, based on 1,000 repetitions, and the *p*-value for the two-sided significance test, corrected for multiple testing using the stepdown method.[10] Panel A in the table looks at all unions, while Panel B studies legal marriages only.

The table shows a mixed picture, with very similar patterns for the two types of unions considered in the top and bottom panels. The first three columns detail estimates for the submatrices where the alternative to marrying one's equal is to marry someone close (a one-level difference) in terms of education. These are the submatrices for which the off-diagonal options have an important weight, as seen in Table 5. For these submatrices, our estimates show an increase in assortativeness except among the two middle-educated groups, for whom the change is not statistically significant at conventional levels. For the other comparisons, the evidence is mixed. It shows more assortativeness towards the bottom of the distribution of education (column 4), less towards the top (column

[10]Romano and Wolf, 2005; Romano, Shaikh and Wolf, 2008; Romano and Wolf, 2016.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

5) and for the most unequal matches (column 6). All these three cases show a high degree of assortativeness in both cohorts, with infrequent off-diagonal unions. Columns 5 and 6 show, however, that unions between college graduates and those with no or basic qualifications have become relatively more common. A closer inspection of the sorting patterns for cohabiting and legal unions in Table 5 shows that it is the high-educated women with three or more years of college education who are increasingly marrying down and driving the drop in assortativeness at these margins. Table A1 in the online appendix reveals a similar pattern for legal marriages only.

Although Table 6 does not provide unequivocal evidence of increased marital sorting in education, it does show marital sorting has increased in the submatrices that represent the largest shares of the market.[11] This evidence is based on the SEV criterion, which uses the way the gains from sorting have changed over time, as well as assumptions on how preferences are distributed. In further analysis, Chiappori, Costa-Dias and Meghir (2020) show that unambiguous statements about changes in sorting – i.e. those that do not depend on auxiliary assumptions such as the distribution of preferences – are uncommon. There, we exploit different criteria to bring further evidence on the changes in assortativeness in marriage.

We now turn to the implications of the changes in sorting that we have identified for inequality across households. This is important for many reasons and particularly so for the intergenerational transmission of inequality: if resources are allocated across households in more unequal ways, investments in children may become more unequal, implying larger inequality in human capital in the next generation.

VII. Changes in education attainment, marital sorting and income inequality

In this section, we quantify the contribution of changes in education attainment and assortativeness in marriage to inequality in family earned income. To do so, we need to calculate the counterfactual marriage rates and sorting patterns under different assumptions about assortativeness and education attainment. We explain how this can be done before discussing the results.

1. Counterfactual sorting distributions in the SEV framework

One key feature of the SEV approach is that it explicitly relates the sorting patterns that can be observed in the data to the underlying forces that determine that exact outcome in the marriage market, given the distribution of

[11] Blundell et al. (2018) and Chiappori, Costa-Dias and Meghir (2020) find stronger evidence of an increase in assortativeness in the US than in Britain.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

classes – education in our case – on each side of the market. Under the SEV assumptions, the deterministic component of the marital surplus and how it is shared between spouses can be determined from the equilibrium sorting patterns. These quantities reveal the value of marriage and of marrying different classes of spouses. In turn, they can be used together with the distribution of classes to determine the equilibrium marital sorting patterns in a market.

We say that two marriage markets are equally assortative if their equilibrium sorting patterns are generated by the same structural matrix of marital surplus. This does not mean that the sorting matrices look identical across the two markets, and indeed they will not be if the distributions of classes are different. But it implies that, for any two cohorts with different distributions of classes, we can use the SEV framework to determine the counterfactual sorting matrix that emulates the assortativeness of cohort 1 under the marginal distributions of cohort 2, or vice versa. This is what we will do to separate the roles that changes in assortativeness in marriage and education attainment play in inequality in family earnings.

To illustrate how it can be done, suppose we want to calculate the marriage rates and marital sorting in a market for which we know the structural surplus matrix Z and the marginal distributions (m, M) of N classes of men and women. Under the SEV assumptions, the deterministic part of the marital surplus of a union between a man of type I and a woman of type J (Z^{IJ}) is simply[12]

$$(14) \qquad Z^{IJ} = 2\ln\left(\frac{\mu_{IJ}}{\sqrt{\mu_{I0}\mu_{0J}}}\right),$$

where $(\mu_{IJ}, \mu_{I0}, \mu_{0J})$ are, respectively, the mass of men in (I, J) matches (which equals the mass of women in (I, J) matches when men and women are equally populous, as in our case), the mass of men who are single and of type I, and the mass of women who are single and of type J. We can rearrange this expression as follows:

$$(15) \qquad \mu_{IJ} = \exp\left(\frac{Z^{IJ}}{2}\right)\sqrt{\mu_{I0}}\sqrt{\mu_{0J}}.$$

We notice that the marginal distributions (m, M) of men and women are

$$(16) \qquad m_I = \mu_{I0} + \sum_{J=1}^{N} \mu_{IJ}$$

[12]See Choo and Siow (2006).

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

and

$$M_J = \mu_{0J} + \sum_{I=1}^{N} \mu_{IJ}. \tag{17}$$

Replacing μ_{IJ} in these expressions yields the following system of $2N$ equations:

$$\begin{cases} \mu_{I0} + \sum_{J=1}^{N} \exp\left(\frac{Z^{IJ}}{2}\right) \sqrt{\mu_{I0}}\sqrt{\mu_{0J}} - m_I = 0 & \text{for } I = 1, \ldots, N \\ \mu_{0J} + \sum_{I=1}^{N} \exp\left(\frac{Z^{IJ}}{2}\right) \sqrt{\mu_{I0}}\sqrt{\mu_{0J}} - M_J = 0 & \text{for } J = 1, \ldots, N \end{cases}. \tag{18}$$

The solution to the system of quadratic equations is the set of singlehood rates μ_{I0} and μ_{0J} for $I, J = 1, \ldots, N$. Once these are determined, we can replace them in equation 15 to determine the new marital sorting patterns. It is worth mentioning that, since the SEV approach links sorting patterns to the value of marriage relative to staying single, it does allow us to calculate both the counterfactual sorting patterns conditional on marriage *and* the counterfactual marriage rates.

2. Inequality in family earnings

Table 7 shows the Gini coefficients for family earned income, within each of the groups defined by cohort, marital status and education. In this table, couples can be legally married or cohabiting; similar results for legally married couples only can be found in Table A2 of the online appendix. Inequality is higher for singles and in couples of lower-educated individuals. This points to the role of marriage and education in protecting earnings. It is also clear that within-group inequality increased across the cohorts for almost all groups, and the change is larger where inequality was already high in the earlier cohort.

Interestingly, we also see an increase in inequality for the entire population, but only a modest one. This suggests that a shift in the composition of families or a decline in between-group inequality, or both, have counteracted the rise in within-group inequality. The empirical patterns we have identified so far reveal forces working in opposite directions. On the one hand, individuals in the later cohort are more likely to be single, which further pushes inequality up. The comparison of marital sorting patterns across the two cohorts did not reveal a clear direction of change, indicating an increase in assortativeness at some margins and a drop at other margins, and hence the effects of these changes on inequality are ambiguous a priori. On the other hand, the later cohort is

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 7

Gini coefficients for family earned income, by marital status (all unions), education and cohort

	Woman's education				*Single men*
	No qual.	Secondary	High school	College	
	Panel A. Birth cohort 1945–54				
Man's education					
No qualifications	0.481	0.339	0.330	0.326	0.648
Secondary	0.367	0.307	0.284	0.271	0.540
High school	0.324	0.278	0.268	0.284	0.445
College	0.333	0.278	0.258	0.236	0.383
Single women	0.686	0.518	0.430	0.344	
	All families: 0.437				
	Panel B. Birth cohort 1965–74				
Man's education					
No qualifications	0.611	0.421	0.397	0.321	0.734
Secondary	0.450	0.337	0.291	0.270	0.550
High school	0.381	0.296	0.278	0.266	0.443
College	0.465	0.293	0.279	0.281	0.418
Single women	0.777	0.584	0.512	0.411	
	All families: 0.446				

Note: Couples are those legally married and those cohabiting. Family income is divided by 2 for couples.

more educated than the earlier one, and education clearly reduces inequality in earnings. Finally, our description of the earnings of men and women by education suggests that the college premium may not have increased over the period, and indeed may have dropped for men. This too can contribute to counterbalance the rise in inequality within groups.

We now investigate how marital sorting and education attainment contributed to change inequality in family earnings across the two cohorts.

3. Education, marital sorting and income inequality

We use the procedure described in Section VII.1 to calculate two counterfactual sorting matrices, for the later birth cohort of 1965–74.[13] The first is based on the marginal distribution of education of the later cohort but superimposes the sorting patterns of the earlier one, born in 1945–54. It will reveal how marital sorting would look under the marginal distribution of the later cohort had assortativeness remained unchanged. The second does the reverse by

[13]For structural approaches measuring the effects of changes in sorting, see Fernández (2002) and Fernández, Guner and Knowles (2005) amongst others.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

TABLE 8

Counterfactual simulations

	Cohort 1945–54	*Cohort 1965–74*		
		Original	Counterfactual 1	Counterfactual 2
All unions	0.437	0.446	0.425	0.528
Legal unions	0.445	0.460	0.425	0.535

Note: Counterfactual 1 uses the distribution of education of the 1965–74 cohort and the assortativeness of the 1945–54 cohort. Counterfactual 2 uses the distribution of education of the 1945–54 cohort and the assortativeness of the 1965–74 cohort.

keeping the sorting patterns of the later cohort while imposing the marginal distribution of the earlier one. It will reveal how marital sorting would look under the value of marriage of the later cohort with the distribution of education as for the earlier cohort. We then reweight our sample for the later cohort using the counterfactual distribution of education and marital patterns and calculate the corresponding Gini coefficients.

The results of this exercise are shown in Table 8. What is clear from the figures in the table is that changes in the probability of marriage and in sorting patterns contributed to push inequality upwards, but only modestly. Indeed, counterfactual 1 shows that inequality would have been lower had the surplus from marrying different types of spouses remained unaltered at the levels of the 1945–54 cohort, but not by very much: the Gini coefficient would have been 2 percentage points lower when considering all unions, and 3.5 percentage points lower when considering legal marriages only. This result is mainly driven by the relative increase in the value of remaining single, which contributes to keep more individuals single than would have happened otherwise. Since inequality is higher among singles, particularly those with lower levels of education, such change contributes to push up inequality in family income.

In contrast, the change in education attainment contributed strongly to prevent inequality from increasing across the two cohorts, as shown in counterfactual 2. Inequality is lower for the more educated, and marital rates dropped by less for this group as well, which explains why the higher levels of education of the later cohort acted to counteract the generalised increase in within-group inequality reported in Table 7. These results are closely aligned with what Eika, Mogstad and Zafar (2019) find using a different measure of assortativeness in marriage.

VIII. Conclusion and discussion

Existing evidence on the direction of change in assortativeness in marriage over the recent past is inconclusive. Most studies use US data, and some find evidence of an increase in assortativeness by education and earnings

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

capacity[14] while others find that assortativeness changed differentially across the distribution of education and declined at the top.[15] The comparison across cohorts is complicated by large changes in the distribution of education that we document above. We have argued that changes in the composition of the marriage market limit our ability to assess changes in assortativeness without relying on specific assumptions about the structure of the matching model. In this paper, we develop the separable extreme value (SEV) approach to bring new evidence on changes in assortativeness in the UK.

More generally, the question of the structure of the marriage market is central to our understanding of inequality, child development and the intergenerational transmission of skills and wealth. Beyond love and companionship, marriage provides public goods, risk sharing and the environment in which children grow up. Thus, the evolving structure of the family can have momentous impacts on how society evolves and how we deal with and understand social inequalities. Studying assortativeness and how it changes is just one step in understanding how families are structured and how economic and social forces shape the family.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

- Appendix

References

Becker, G. S. (1981), *A Treatise on the Family*, Cambridge, MA: Harvard University Press.

Blundell, R., Joyce, R., Norris Keiller, A. and Ziliak, J. (2018), 'Income inequality and the labour market in Britain and the US', *Journal of Public Economics*, vol. 162, pp. 48–62.

Chiappori, P. A. (2017), *Matching with Transfers: The Economics of Love and Marriage*, Princeton, NJ: Princeton University Press.

—, Costa-Dias, M. and Meghir, C. (2018), 'The marriage market, labor supply, and education choice', *Journal of Political Economy*, vol. 126, pp. S26–72.

—, — and — (2020), 'Changes in assortative matching: theory and measurement in the US and the UK', University of Columbia, mimeo.

—, Salanié, B. and Weiss, Y. (2017), 'Partner choice, investment in children, and the marital college premium', *American Economic Review*, vol. 107, pp. 175–201.

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

Eika, L., Mogstad, M. and Zafar, B. (2019), 'Educational assortative mating and household income inequality', *Journal of Political Economy*, vol. 127, pp. 2795–835.

[14]Greenwood et al., 2014; Chiappori, Salanié and Weiss, 2017.

[15]Eika, Mogstad and Zafar, 2019.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies

Fernández, R. (2002), 'Education, segregation and marital sorting: theory and an application to the UK', *European Economic Review*, vol. 46, pp. 993–1022.
—, Guner, N. and Knowles, J. (2005), 'Love and money: a theoretical and empirical analysis of household sorting and inequality', *Quarterly Journal of Economics*, vol. 120, pp. 273–344.
— and Rogerson, R. (2001), 'Sorting and long-run inequality', *Quarterly Journal of Economics*, vol. 116, pp. 1305–41.
Greenwood, J., Guner, N. and Knowles, J. (2003), 'More on marriage, fertility, and the distribution of income', *International Economic Review*, vol. 44, pp. 827–62.
—, —, Kocharkov, G. and Santos, C. (2014), 'Marry your like: assortative mating and income inequality', *American Economic Review: Papers and Proceedings*, vol. 104, pp. 348–53.
McFadden, D. (1981), 'Econometric models of probabilistic choice', in C. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge, MA: MIT Press.
Office for National Statistics (2019), *Quarterly Labour Force Survey, 1992–2019*, ONS, Social Survey Division, Northern Ireland Statistics and Research Agency, Central Survey Unit.
Reynoso, A. (2019), 'The impact of divorce laws on the equilibrium in the marriage market', University of Michigan, mimeo.
Romano, J., Shaikh, A. and Wolf, M. (2008), 'Formalized data snooping based on generalized error rates', *Econometric Theory*, vol. 24, pp. 404–47.
— and Wolf, M. (2005), 'Stepwise multiple testing as formalized data snooping', *Econometrica*, vol. 73, pp. 1237–82.
— and — (2016), 'Efficient computation of adjusted p-values for resampling-based stepdown multiple testing', *Statistics and Probability Letters*, vol. 113, pp. 38–40.
Shen, J. (2019), '(Non-)Marital assortative mating and the closing of the gender gap in education', Princeton University, mimeo.

© 2020 The Authors. *Fiscal Studies* published by John Wiley & Sons Ltd. on behalf of Institute for Fiscal Studies