Educational Assortative Mating and Household Income Inequality

Lasse Eika

Statistics Norway

Magne Mogstad

University of Chicago, Statistics Norway, and National Bureau of Economic Research

Basit Zafar

Arizona State University and National Bureau of Economic Research

We use data from Denmark, Germany, Norway, the United Kingdom, and the United States to document the degree of educational assortative mating, how it evolves over time, and the extent to which it differs between countries. This descriptive analysis motivates and guides a decomposition analysis in which we quantify the contribution of various factors to the distribution of household income. We find that assortative mating accounts for a nonnegligible part of the cross-sectional inequality in household income in each country. However, changes in assortative mating over time barely move the time trends in household income inequality.

I. Introduction

It is often argued that individuals are increasingly sorting into internally homogeneous marriages, and that this assortative mating has led to a rise in household income inequality. This widespread view stems from two

We thank James J. Heckman, three referees, and seminar participants at several universities and conferences for comments and suggestions. We are grateful to Rasmus Landersø

Electronically published September 26, 2019
[*Journal of Political Economy*, 2019, vol. 127, no. 6]

empirical facts. The proportion of couples who share the same level of education (i.e., educational homogamy) has been growing over the past few decades (see, e.g., Pencavel 1998; Schwartz and Mare 2005). Accompanying this increase in educational homogamy, there has been a rise in household income inequality (see, e.g., Western, Bloome, and Percheski 2008). In the United States, for example, the share of couples in which both spouses have a college degree increased by 22 percentage points between 1962 and 2013, while the Gini coefficient in household income among married couples increased from 33.9 to 43.2 over this period.

Measuring educational assortative mating and its impact on household income inequality has proven difficult for several reasons.[1] One challenge is to determine whether the increase in educational homogamy arises from secular changes in educational attainment of men and women, or because of shifts in educational assortative mating. For example, the closing of the gender gap in higher education may increase the probability that a college graduate is married to someone with a college degree, even if there were no changes in the assortativeness of marriage (see Liu and Lu 2006). Another challenge is that the economic returns to education have increased considerably over the past few decades (see, e.g., Autor, Katz, and Kearney 2008). As a result, educational assortative mating may become increasingly important for the distribution of household income, even if there were no changes in the mating pattern.

This paper tries to address these challenges, making two main contributions. The first is to examine the degree of educational assortative mating, how it evolves over time, and the extent to which it differs between countries. The coverage and detailed nature of the data we are using allows us to bring new evidence on the assortativeness of marriage over time and between countries. This evidence motivates and guides our second contribution: a quantification of the explanatory power of various factors to household income inequality. In particular, we employ the semiparametric decomposition method proposed by DiNardo, Fortin, and Lemieux (1996) to quantify the relative importance of changes in educational composition, returns to education, and educational assortative mating for the rise in household income inequality.

While our study is centered on the United States, we also provide evidence from Denmark, Germany, the United Kingdom, and Norway. The

and Michael Graber for helping us with the analysis of the Danish data and German data. The project received financial support from the Norwegian Research Council. Data are provided as supplementary material online.

[1] The term “assortative mating” is used in different ways by different authors. Commonly, positive (negative) educational assortative mating is defined as men and women with the same level of education marrying more (less) frequently than what would be expected under a marriage pattern that is random in terms of education. In most of our analysis, we follow this definition of educational assortative mating. As a robustness check, we also use alternative measures of assortative mating.

data available in each of these countries allow us to study educational assortative mating and household income inequality over several decades. The US data are available for the longest time period, going back to 1940. By comparing the results across countries, we shed light on whether our findings are specific to the United States or common across several western economies that differ considerably in the incentives to choose spouses by their education levels.[2]

Our paper offers three sets of results. First, we present new evidence on the long-run evolution of educational assortative mating in the United States. We show that ever since 1940 (our first year of data), Americans with the same level of education marry much more frequently than what would be expected under a marriage pattern that is random in terms of education. This positive assortative mating occurs at all levels of education throughout the entire sample period. However, the time trends are heterogeneous and vary depending on where in the educational distribution one looks. On one hand, assortative mating among the highly educated has been steadily declining over time. In the early 1960s, for example, Americans with a college degree were five times as likely to be married to a spouse with a college degree, compared to the counterfactual situation in which spouses were randomly matched with respect to education; in 1980 and 2013, they were only three and two times as likely, respectively. On the other hand, assortative mating has gradually increased among the lowly educated. In the early 1960s, Americans without a high school degree were 1.6 times as likely to be married to one another as compared to the probability with random mating; in 1980 and 2013, they were 2.6 and 7.2 times as likely. Aggregating the measures of assortative mating across education levels, we obtain a measure of the overall educational assortative mating. This measure suggests that on average, the degree of assortative mating increased gradually from 1940 to the 1980s, after which it has changed relatively little. This conclusion is robust to employing alternative measures of assortative mating as well as to accounting for sorting by age and changes in the probability of marriage by education level.

Our second set of results comes from the decomposition method used to quantify the contribution of various factors to the distribution of household income in the United States. We find that educational assortative mating accounts for a nonnegligible part of the cross-sectional inequality in household income. For example, our results suggest that the Gini coefficient in 2013 is almost 5 percent higher compared to the counterfactual

[2] For example, Landersø and Heckman (2017) show that the relationships linking education and income differ greatly between Denmark and the United States. In particular, the economic returns to education are relatively low in Denmark, in part due to the compressed wage distribution (possibly due to unions and minimum wage standards) but also because of the progressivity of the tax-transfer system.

situation in which spouses were randomly matched. However, the changes in assortative mating over time barely move the time trends in household income inequality. This is because the inequality contribution from the increase in assortative mating among the lowly educated is offset by the equalizing effect from the decline in assortative mating among the highly educated. By comparison, increases in the returns to education generate a considerable rise in household income inequality, but these price effects are partly mitigated by increases in college attendance and completion rates among women. For example, our estimates suggest that the Gini coefficient in 2013 would have been around 25 percent lower if returns to education had remained at the levels we observe in the early 1960s. By way of comparison, the Gini coefficient would have been 12 percent higher had the educational composition in 2013 been like that in the 1960s.

Our third set of results shows that the main findings about educational assortative mating and household income inequality are not specific to the United States, but generalize to the other developed countries for which we have data. In each country, we find evidence of positive assortative mating at all levels of education. In addition, the time trends are qualitatively similar across countries: Among college graduates, assortative mating has been declining over time, whereas the lowly educated are increasingly sorting into internally homogeneous marriages. On top of this, the conclusions about household income inequality are broadly consistent across countries. Educational assortative mating is important for the cross-sectional inequality in household income, whereas changes in assortative mating over time matter little for the time trends in household income inequality. By comparison, changes in the returns to education over time are a key factor behind the evolution of household income inequality.

In interpreting these three sets of results, it is important to keep in mind the descriptive nature of our analysis. While our study offers new facts about educational assortative mating over time and across countries, it is silent on the causes of sorting in marriage. For example, sorting along educational attainment might reflect a preference for a spouse with a certain education level or that people could be more likely to meet someone with a similar level of education in school, college, or at work. Our decomposition method is also best understood as a descriptive approach or accounting exercise, abstracting from several potentially important partial equilibrium considerations (e.g., self-selection into education by comparative advantage) and general equilibrium conditions (e.g., simultaneous determination of education choices, returns to education, and marriage decisions).[3]

[3] See Fortin, Lemieux, and Firpo (2011) for a review of decomposition methods in economics and a discussion of these considerations.

Our study is primarily related to the empirical literature on educational assortative mating (see, e.g., Mare 1991; Pencavel 1998; Fernandez and Rogerson 2001; Breen and Salazar 2011; Greenwood et al. 2014, 2016). Our paper expands and clarifies this prior research in several important ways. First, our analysis of the United States considers the trends over a longer time period, showing that the degree of assortative mating increased gradually from 1940 to the 1980s, after which it has changed relatively little. Second, we study educational assortative mating for a broad range of countries, revealing a common pattern in educational assortative mating over time. Third, we study educational assortative mating in the United States by race, finding qualitatively similar results for blacks and whites. Fourth, we provide a detailed analysis of educational assortative mating among the college educated. This analysis shows a stronger decline in assortative mating among American couples with graduate degrees as compared to those with only undergraduate degrees. Using the rich Norwegian data, we also demonstrate that assortative mating is even stronger by postsecondary field of study (college major) than by education level, suggesting that the choice of field of study could be an important but neglected pathway through which individuals sort into homogeneous marriages. And lastly, our analysis highlights the importance of using measures of educational assortative mating that adjust for changes over time in the marginal distributions of education of men and women. As discussed in detail later, this issue is important to understand why Greenwood et al. (2014, 2016) conclude that educational assortative mating increased after 1980, whereas our study suggests it changed little or declined modestly over the past three decades.

Our paper also contributes to the literature trying to explain or account for the rise in economic inequality observed in many developed countries since the early 1980s. Most of the evidence is on the factors behind the increase in earnings inequality among males. A smaller body of work has examined the trends in household income inequality. Some studies decompose the inequality in household income by income sources and subgroups (see, e.g., Karoly and Burtless 1995; Cancian and Reed 1998; Aslaksen, Wennemo, and Aaberge 2005; Western et al. 2008; Breen and Salazar 2011). Other studies use shift-share approaches to examine the change in income inequality accounted for by changes in male and female labor earnings distributions and changing household characteristics (see, e.g., Burtless 1999; Daly and Valletta 2006; Greenwood et al. 2014; Larrimore 2014). Our study complements this body of work by quantifying the relative importance of changes in educational composition, returns to education, and educational assortative mating to household income inequality over time for a broad set of developed countries. From this analysis, two new conclusions emerge: Educational assortative mating accounts for a modest but nonnegligible part of the cross-sectional inequality in household income, and changes in assortative mating over time barely move the time trends in household income inequality.[4]

The remainder of this paper proceeds as follows. Section II describes our data and reports descriptive statistics. Section III presents our findings on educational assortative mating. Section IV describes the decomposition method before using it to explore factors behind the evolution of household income inequality. The final section offers some concluding remarks.

II. Data and Descriptive Statistics

Below we describe the data, select the estimation samples, and present descriptive statistics. Details about data sources, sample restrictions, and each of the variables are given in appendix A.

A. *Data for the United States*

Our analysis of the United States relies on two data sources. For the period 1962–2013, we use the public-use March Current Population Survey (CPS). In every year, the survey covers a nationally representative sample of households. The variables captured in the survey include individual demographic information (such as gender, date of birth, and marital status) and socioeconomic data (including educational attainment and income). The data contain unique family identifiers that allow us to match spouses.

Using the CPS data, we describe educational assortative mating and examine its impact on household income inequality over the period 1962–2013. To study assortative mating over an even longer period of time, we take advantage of the public-use samples from the US decennial censuses. The primary advantage of the census data is that they allow us to describe educational assortative mating in 1940. However, spouses cannot be linked in the 1950 census, preventing us from analyzing that year. Additionally, the 1940 census does not offer a comprehensive measure of income. For these reasons, the analysis of household income inequality only uses the CPS data.

[4] These conclusions differ from those drawn by Greenwood et al. (2014). Using US data, they conclude that moving from the observed pattern of assortative mating to a random pattern has little discernible impact on income inequality in 1960, whereas repeating this experiment for 2005 has a marked impact on the income distribution. Additionally, they argue that forcing men and women to sort into marriages in 2005 as they did in 1960 would generate a large drop in income inequality in 2005. Both these conclusions are wrong. In a corrigendum titled "Corrigendum to Marry Your Like: Assortative Mating and Income Inequality," Greenwood et al. point out these errors. This corrigendum is dated June 2015, a year after our NBER working paper was posted. While the corrigendum does not explain what caused the errors, it reports results that are consistent with our conclusions.

B. Data for Other Countries

Our analysis also uses data from Denmark, Germany, the United Kingdom, and Norway. In each country, we have individual demographic information (gender, date of birth, and marital status), socioeconomic data (including educational attainment and income), and family identifiers to match spouses. For Denmark and Norway, we access administrative data containing records for every individual and household. The Norwegian data allow us to go back to 1967, whereas the Danish data we have access to only begin in 1980. For Germany and the United Kingdom, we have to rely on survey data. The German Socio-Economic Panel (SOEP) is an annual longitudinal survey that covers the period from 1984 to 2013. The UK Labour Force Survey (LFS) collects data biannually from 1979 to 1983, and then annually from 1984 until 2013. However, the LFS does not ask questions about income before 1993.

C. Sample Selection and Variable Definitions

Our main analysis uses information on the educational attainment and income of married couples. In every year, the sample is restricted to married couples in which at least one of the spouses is between 26 and 60 years old.[5] In our main analysis, these individuals are assigned to one of four mutually exclusive groups according to the highest level of education completed: high school dropouts (< 12 years of schooling), high school graduates (12 years of schooling), individuals with some college (13–15 years of schooling), and college graduates (> 15 years of schooling). Our measure of individual income consists of wages and income from self-employment.[6] In each year, we exclude individuals with missing information on income, and set negative income to zero.[7] We measure household income by pooling the individual income of the spouses.

In table 1, we document key characteristics of the samples of husbands and wives in the United States. As expected, female labor force participation has grown over time. As a result, the incomes of females have increased, both in absolute levels and as shares of household income. At the same time, we can see a convergence in educational attainment of men

[5] In a robustness check, we examine how the inclusion of singles or cohabitants affects the conclusions about assortative mating. We also check the sensitivity of the results to restricting the sample to younger couples in which at least one of the spouses is between 26 and 30 years old in a given year.

[6] In the United Kingdom, however, the survey does not ask about income from self-employment (but records whether people are self-employed). We therefore exclude the self-employed (and their spouses) from the UK sample when analyzing household income inequality (Sec. IV).

[7] Less than 1 percent of the observations in the CPS have negative income in a given year. Dropping these observations does not affect our results.

TABLE 1
US Data: Summary Statistics

	1962		1980		2013	
	Women	Men	Women	Men	Women	Men
Sample means:						
Age	40.5	43.8	40.4	43.2	43.5	45.6
College degree	.071	.129	.154	.236	.391	.370
Income (2014 USD)	7,200	46,102	15,358	62,285	29,979	62,028
Labor force participation	.409	.955	.625	.933	.714	.892
Number of observations	13,754		31,549		31,570	

Source.—CPS (1962–2013).

Note.—This table reports average characteristics of our estimation sample of married couples aged 26–60 in the United States. Labor force participation is defined as having positive labor income from wages or self-employment.

and women. The characteristics (and trends over time) for the full sample that also includes singles, shown in table D1 (tables D1–D8 are available online), are qualitatively similar. Table D2 reports the same characteristics for the samples of husbands and wives in Denmark, Germany, the United Kingdom, and Norway. These countries have also experienced a closing in the gender gap in higher education and an increase in the incomes and labor force participation of women.

D. *Descriptive Statistics*

Before turning to the examination of educational assortative mating, we describe a few important features of our data.

We begin by displaying the education distributions of husbands and wives over time. Figure 1 presents the US time trends. The proportion of husbands with a college degree starts out at around 13 percent in 1962 and increases to about 37 percent in 2013. By comparison, only 7 percent of the wives had a college degree in 1962. Over time, however, the educational attainment of women caught up with that of men, and the wives in 2013 are actually more likely to have a college degree. The increase in college education is accompanied by a substantial decline in the proportion of the population without a high school diploma. As evident from figure D1 (figs. D1–D17 are available online), broadly similar patterns are observed in Denmark, Germany, the United Kingdom, and Norway.

Figure 2 shows how the closing of the gender gap in higher education is accompanied by an increase in homogamy among the college educated. In the United States, the share of couples in which both spouses have a college degree increased by 22 percentage points between 1962 and 2013. This figure reveals a similar pattern for the other countries we consider, where the proportion of couples in which both spouses are college educated has increased substantially over time.

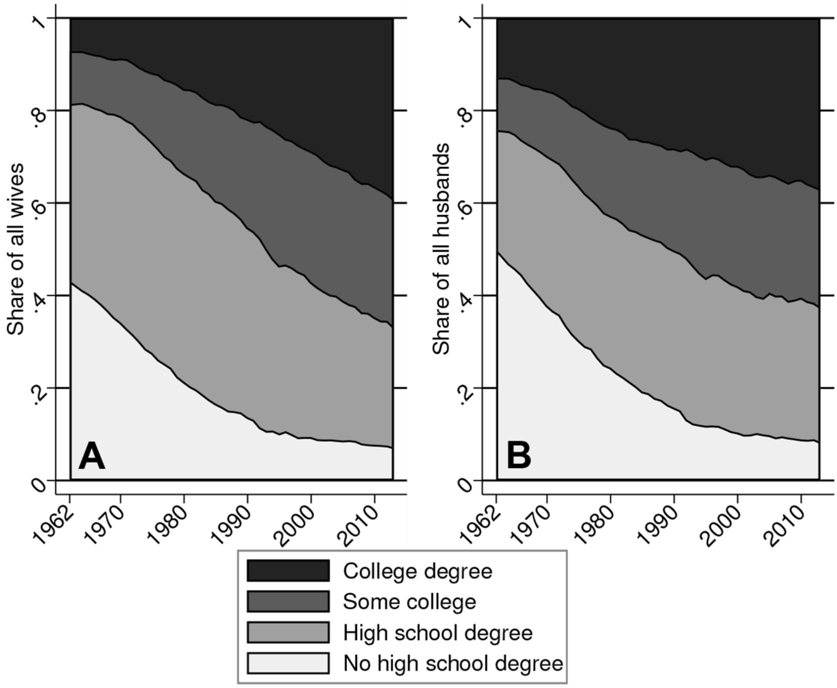

FIG. 1.—US time trends in educational attainment of wives (*A*) and husbands (*B*). Source: CPS (1962–2013), married couples aged 26–60.

Alongside the changes in the education composition of husbands, wives, and couples, a large body of work documents significant changes in the labor market returns to education (see, e.g., Autor et al. 2008; Acemoglu and Autor 2011). Table D3 confirms this pattern in our data, reporting income differentials of husbands and wives from ordinary least squares (OLS) regressions of annual income on education levels (conditional on potential experience). We see sizable income premiums for high school and college degrees in each country. Looking at changes over time in table D3, it is evident that the positive association between income and education has been increasing in the United States, especially since the 1980s.

III. Educational Assortative Mating

A. Baseline Estimates for the United States

Assortative mating is typically defined as a mating pattern in which individuals with similar traits mate with one another more frequently than would be expected under a random mating pattern. This definition suggests that educational assortative mating can be quantified by comparing

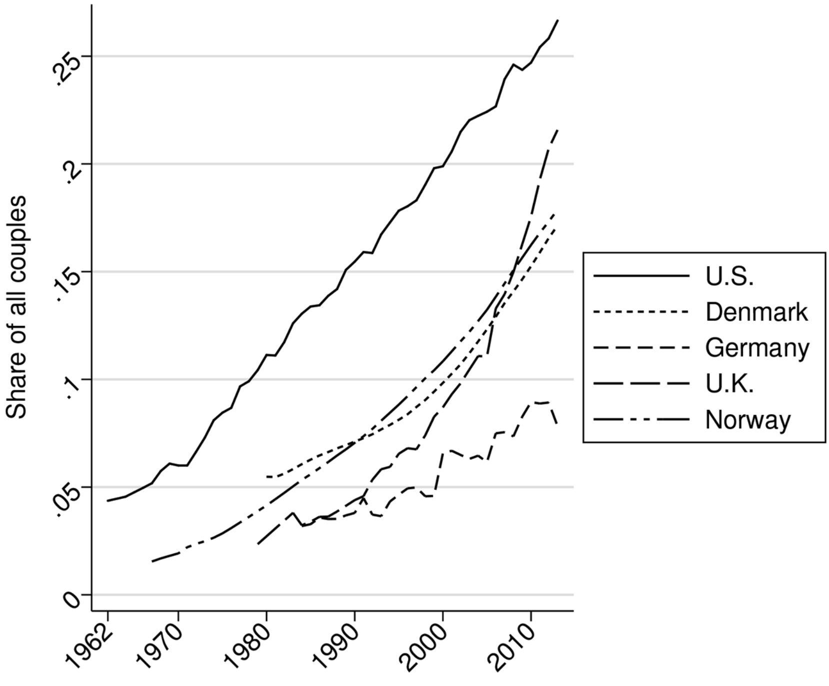

FIG. 2.—Proportion of couples in which both spouses have a college degree: the United States, Denmark, Germany, the United Kingdom, and Norway. Sources: CPS (1962–2013); Danish registry data (1980–2013); SOEP (Germany, 1984–2013); LFS (United Kingdom, 1979–2013); Norwegian registry data (1967–2013). Sample: married couples aged 26–60.

the contingency table for the wife's and husband's educational levels to a contingency table generated by random matching for husbands and wives. Based on these contingency tables, it is possible to measure marital sorting between education levels e_f and e_m as the observed probability that a husband with education level e_m is married to a wife with education level e_f, relative to the probability under random matching with respect to education:

$$s(e_f, e_m) = \frac{P(E_f = e_f, E_m = e_m)}{P(E_f = e_f)P(E_m = e_m)}, \tag{1}$$

where E_f (E_m) denotes the education level of the wife (husband). Thus, the magnitude of $s(e_f, e_m)$ can be interpreted as the likelihood of a particular match as compared to what the probability of the match would be with random matching. Positive (negative) assortative mating means that men and women with education levels e_m and e_f marry more (less) frequently than what would be expected under a marriage pattern that is random in terms of education: that is, the marital sorting parameter $s(e_f, e_m)$ is larger (smaller) than 1 when e_f is equal to e_m. The joint education distribution of the spouses is fully described by the marital sorting parameters and the marginal education distributions of wives and husbands.

In each year, we use CPS data to estimate the sorting parameters $s(e_f, e_m)$ for every combination of education of the husbands and wives. Table D4 reports the full set of estimates of $s(e_f, e_m)$ for the United States in the years 1962, 1980, and 2013. Figure 3 complements this table by displaying the time trend in the sorting parameters on the diagonal (where husbands and wives have the same education level, $e_f = e_m$). This figure shows there is evidence of positive assortative mating at all levels of education during the entire period 1962–2013. The time trends, however, are heterogeneous and vary depending on where in the educational distribution one looks. We can see that assortative mating has declined among the highly educated. In 1962, for example, Americans with a college degree were nearly five times as likely to be married to a spouse with a college degree, compared to the counterfactual situation in which spouses were randomly matched with respect to education; in 1980 and 2013, they were only three times and twice as likely, respectively. Conversely, assortative mating has increased among the lowly educated. In 1962, Americans without a high school degree were 1.6 times as likely to be married to one another as compared to the probability with random mating; in 1980 and 2013, they were 2.6 and 7.2 times as likely, respectively.

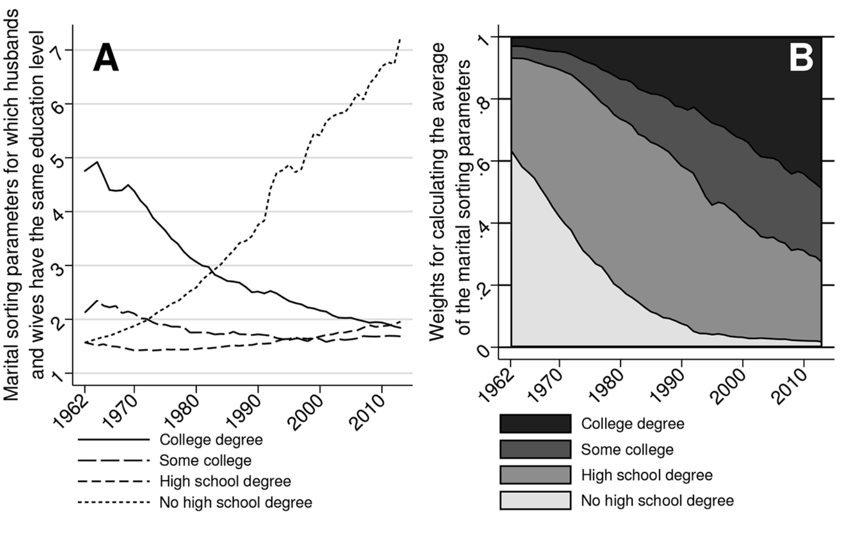

FIG. 3.—US trends in assortative mating by educational level. *A*, Time trends in the marital sorting parameters $s(e_f, e_m)$ for which the husbands and wives have the same education level. *B*, Weights used to calculate the weighted average of the marital sorting parameters along the diagonal (see fig. 4). Source: CPS (1962–2013), married couples aged 26–60.

To obtain a measure of the overall educational assortative mating, we compute the weighted average of the marital sorting parameters along the diagonal. Figure 4 shows how this measure of aggregate assortative mating changes over time. The results suggest that, on average, the degree of educational assortative mating increased steadily from 1962 to the mid-1980s, after which it has changed relatively little. In both 1980 and 2013, Americans with the same level of education were about 1.9 times more likely to be married to one another as compared to the probability with random mating. By way of comparison, Americans in 1962 were only 1.7 times as likely to be married to someone with the same level of education as compared to the probability with random mating.

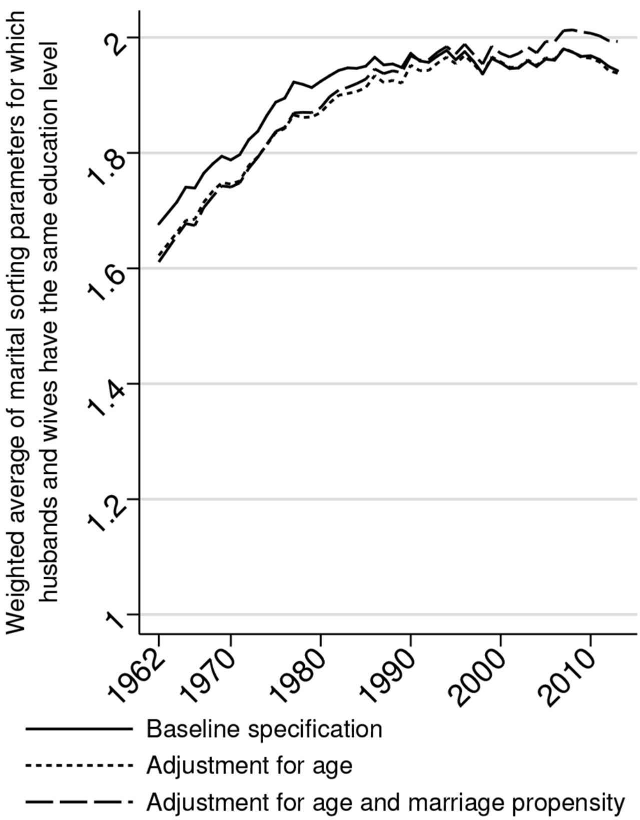

FIG. 4.—US trends in aggregate educational assortative mating. This figure displays the time trends in the weighted average of the marital sorting parameters $s(e_f, e_m)$ along the diagonal (where husbands and wives have the same education level). Source: CPS (1962–2013), married couples aged 26–60.

B. Robustness Checks

This section shows that our results are robust to employing alternative measures of assortative mating as well as to accounting for sorting by age and changes in the probability of marriage by education level.

1. Alternative Measures

For continuous multivariate distributions, the univariate marginals and the dependence structure can be completely separated and the latter may be represented by a copula. When the marginals are continuous, any question about the dependence structure can be answered with knowledge of the copula alone. For discrete multivariate distributions, there exists no unique copula that summarizes all information about dependence, and many common measures of dependence cannot be represented by a copula. Thus, it is natural to ask how sensitive the results are to using measures of educational assortative mating other than $s(e_f, e_m)$.

One possibility is to use Altham's index, which is based solely on the odds ratios.[8] Applied to educational assortative mating, this index summarizes the distance between the row-column associations in the contingency table for wife's and husband's educational level and the row-column associations in a contingency table generated by random matching of husbands and wives. Using our notation and the four categories of education, Altham's index can be written as

$$\left[\sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{k=1}^{4}\sum_{l=1}^{4}\left|\log\left(\frac{P(E_f = i, E_m = j)P(E_f = k, E_m = l)}{P(E_f = i, E_m = l)P(E_f = k, E_m = j)}\right)\right|^2\right]^{1/2}, \quad (2)$$

where the index takes a smaller value the closer the row-column association is to what we would observe under independence, and it is equal to zero if matches are random. It is difficult, however, to give the Altham index a cardinal interpretation. It is also worth noting that the Altham index assigns the same weight to all cells in the contingency table. In comparison, the measure of aggregate assortative mating based on $s(e_f, e_m)$ focuses attention on the diagonal, where husbands and wives have the same education level. Despite these differences, the two measures offer a similar picture of the trends in assortative mating over time. Figure 5*B* reports estimates of the Altham index for the United States during the period 1962–2013. For comparison, figure 5*A* repeats our baseline measure of aggregate sorting over the same time period. Both measures suggest that on average, the degree of educational assortative mating increased steadily

[8] Altham's index was proposed by Altham (1970) to measure association or dependence in contingency tables, and it is frequently used in studies of mobility with unordered variables such as occupations (see, e.g., Long and Ferrie 2013).

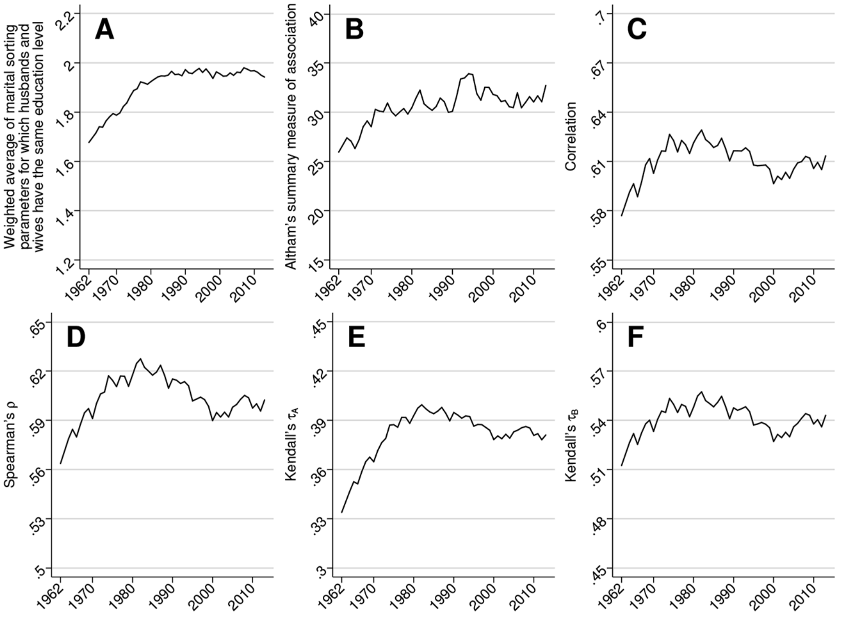

FIG. 5.—Educational assortative mating in the United States: alternative measures. *A*, Our baseline measure of aggregate educational assortative mating; *B*, estimates of Altham's index; *C*, estimates of Pearson's correlation coefficient; *D*, estimates of Spearman's correlation coefficient; *E*, estimates of Kendall's τ without adjustment for ties; *F*, estimates of Kendall's τ with adjustment for ties. Source: CPS (1962–2013), married couples aged 26–60.

from 1962 to the mid-1980s, while it has changed little if any over the past two decades.

Both Altham's index and the marital sorting parameters $s(e_f, e_m)$ do not impose any ordering of the different education groups.[9] There are, however, several commonly used measures of dependence for ordered contingency tables, including Spearman's rank-order correlation coefficient, Pearson's product-moment correlation coefficient, and Kendall's τ (see, e.g., Nelsen 2006; Agresti 2010). The Spearman correlation between the education levels of husbands and wives is simply the Pearson correlation between the rank values of those two variables. Kendall's τ is an alternative measure of rank correlation, given by the difference between the number of concordant and discordant pairs (of couples) relative to the total

[9] Arguably, there is no uniform ordering of education levels across individuals. For example, Carneiro, Heckman, and Vytlacil (2011) estimate that returns to college differ across individuals in magnitude and even sign, and that people select into education based in part on their idiosyncratic returns.

number of pairs (of couples).[10] As a result, the Kendall correlation ranges from −1 to 1, and it will be closer to 1 the more similar the ranks of the spouses are in the marginal distribution of education of husbands and wives.

Figures 5*C* and 5*D* show that both the Spearman and the Pearson correlation coefficients suggest that educational assortative mating increased from 1962 to the 1980s. In the 1990s, the coefficients fell somewhat, but they remained considerably larger in 2013 as compared to 1962. Similarly, the estimates of Kendall's τ presented in figures 5*E* and 5*F* suggest that educational assortative mating increased until the 1980s, after which it declined modestly. This conclusion holds true whether we consider τ_A, which ignores ties (i.e., pairs of couples for which husbands or wives have the same level of education), or make adjustment for ties, as in τ_B.[11]

2. Age and Marriage Adjustments

The baseline results in figure 4 capture a variety of marriage patterns, including first-time marriages among young individuals as well as divorces and remarriages later in the life cycle. Given that we only have repeated cross-sectional data, we cannot investigate the dynamics of marriage and divorce. Instead, we perform a robustness check comparing estimates of educational assortative mating in the United States for the full sample (age range 26–60 in a given year) to the subsample in which at least one of the spouses is between 26 and 30 years old in a given year. Figure D2 presents the results. We find evidence of positive assortative mating at all levels of education in both samples. Moreover, the time trends are qualitatively the same and quantitatively quite similar across the two samples. Among college graduates, assortative mating has been declining over time, whereas the lowly educated are increasingly sorting into internally homogeneous marriages.

Another way to take the ages of spouses into account in constructing the measure of assortative mating is to adjust the sorting parameter in equation (1). Let the probability that we observe a household in which the wife has education e_f and age a_f and the husband has education e_m and age a_m be denoted $P(E_f = e_f, A_f = a_f, E_m = e_m, A_m = a_m)$, where E_f (E_m) and A_f (A_m) denote education and age group for wives (husbands), respectively. As before, we measure marital sorting by the ratio between the probability of observing such a household and the probability of observing a similar household if matching is random with respect to education, except that it is now conditional on the spouses' ages. The martial sorting parameter is then

[10] In our setting, a pair of couples is said to be concordant if both the wife and husband in one couple have higher education than the wife and husband in the other couple. The pair of couples is discordant, on the other hand, if one couple has a wife with lower education and a husband with higher education as compared to the other couple.

[11] See Agresti (2010) for a discussion of how to adjust for ties in computing Kendall's τ.

$$s(e_f, a_f, e_m, a_m) = \frac{P(E_f = e_f, A_f = a_f, E_m = e_m, A_m = a_m)}{P(E_f = e_f|A_f = a_f)P(E_m = e_m|A_m = a_m)P(A_f = a_f, A_m = a_m)}, \quad (3)$$

where the denominator is the probability of observing this couple if matching is random with respect to education. The expression in the denominator is derived from the fact that under random matching with respect to education, the wife's education e_f is not directly dependent on the husband's education e_m and age a_m, and the husband's education e_m is not directly dependent on the wife's education e_f and age a_f.

To construct age-adjusted marital sorting parameters, we divide the sample into 144 groups. Specifically, for each gender we use three age groups in addition to the four educational levels. The first group consists of individuals younger than 37 years, the second group includes individuals between the ages of 37 and 48, and the last group consists of individuals older than 48 years. We then estimate a full set of marital sorting parameters for all possible combinations of age and education. In order to develop an age-adjusted measure of marital sorting for a given combination of husband's and wife's education, we aggregate the age and education specific measures across ages. Details on the age adjustment and aggregation are presented in appendix B. Figure 4 shows the weighted average of the age-adjusted marital sorting parameters along the diagonal (i.e., for groups where the husbands and wives have similar education). Adjusting for age does not materially change the trends in marital sorting.

As a final robustness check, we modify the sorting parameter to account for changes in the probability of marrying (conditional on age and education). For this purpose, we now also include single men and women aged 26–60 in the sample, add separate education and age categories corresponding to single women (no husband) and single men (no wife), and reestimate the marital sorting parameters. The denominator in equation (3) now corresponds to the marriage patterns if both the probability of marrying and who marries whom are independent of education. Figure 4 shows the weighted average of the marital sorting parameter (adjusted by age and marriage propensity) along the diagonal (i.e., for groups where the husbands and wives have similar education). Adjusting for age and marriage propensity does not materially change the trends in marital sorting.

3. Incorporate Cohabiting Couples

The increasing tendency for couples to cohabit (instead of marrying) could potentially affect the trends in educational assortative mating. To examine this, we perform two checks. We first exploit that cohabitation is recorded in the CPS for the years 2007–2013. This allows us to add cohabiting couples (aged 26–60) to the sample of married couples for these years. During this period, 6.5 percent of all men and women aged 26–60 are recorded as cohabitants. Figure D3 shows that classifying these cohabiting couples as married does not materially change the estimates of the marital sorting parameters. The second check we perform uses the Norwegian register data, where we have information about cohabitation for the years 1988–2013. Figure D4 shows that the trends in the sorting parameters over this period barely move when cohabiting couples are classified as married.

C. *Additional Analysis*

1. Comparison with Existing Research

All the measures in figure 5 adjust for changes over time in the marginal distributions of education of husbands and wives. In comparison, some studies construct measures of educational assortative mating from regressions of the wife's years of education ($\mathrm{Y\,edu}_f$) on the husband's years of education ($\mathrm{Y\,edu}_m$):

$$\mathrm{Y\,edu}_f = \alpha + \gamma \mathrm{Y\,edu}_m + u. \tag{4}$$

After estimating this regression model separately for different years, these studies interpret the (change over time in the) coefficient on husband's education as a measure of the (change over time in) educational assortative mating. For example, Greenwood et al. (2014, 2016) use estimates of γ to study educational assortative mating in the United States over the period 1960–2005.[12]

For comparison, we use our data for the period 1962–2013 to estimate (4) separately for each year. Figure 6 presents the yearly estimates of γ. In line with Greenwood et al. (2014, 2016), we find that this coefficient increases steadily over time, also after 1980. However, the evolution of this coefficient over time does not accurately measure the changes in educational assortative mating, as it confounds changes in the assortativeness of marriage with shifts in the marginal distributions of education.[13]

[12] Greenwood et al. (2014) report three different measures of educational assortative mating for the period 1960–2005: the regression coefficient for husband's education (γ), Kendall's τ (τ_A), and a weighted average of the marital sorting parameters ($s(e_f, e_m)$) along the diagonal. All these measures suggest that educational assortative mating increased from 1960 to 1980. In contrast, only γ—which is confounded by changes in the marginal distributions of education—suggests assortative mating continued to increase after 1980. Indeed, their estimate of τ_A actually declines somewhat from 1980 to 2005, whereas the estimated $s(e_f, e_m)$ barely moves from 1980 to 2000, after which it increases a little.

[13] In an attempt "to control for the secular rise in the educational levels for the married population," Greenwood et al. (2014, 348) allow the intercept to change from year to year

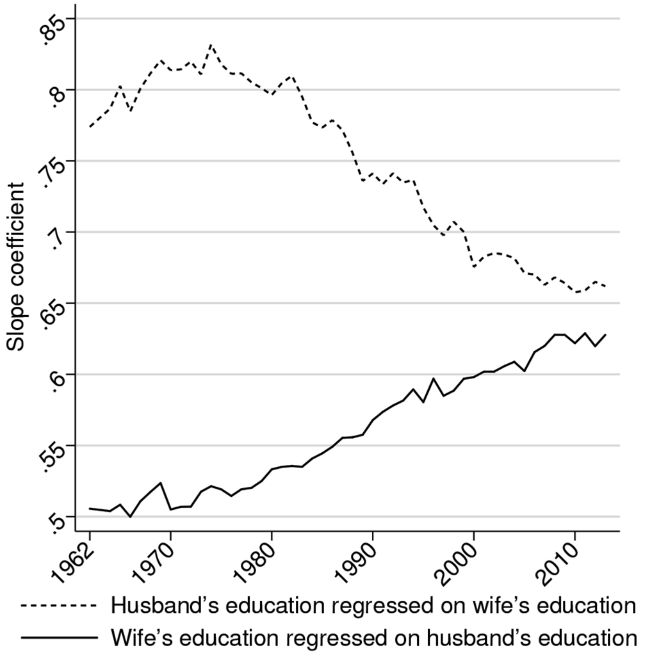

FIG. 6.—US estimates of the association between the educational attainment of spouses. This figure displays the slope coefficients from separate regressions for each year of (a) the wife's years of education on an intercept and the husband's years of education, and (b) the husband's years of education on an intercept and the wife's years of education. Source: CPS (1962–2013), married couples aged 26–60.

To illustrate this point, we switch the regressand and the regressor in (4), running regressions of husband's years of education on an intercept and wife's years of education. Figure 6 shows the results. While the coefficient on wife's education declines over the time, the coefficient on husband's education increases over the same period. The reason is that the variance in years of education among husbands falls substantially over this time period (compared to the variance in years of education among wives). Comparing the results in figure 6 to those in figure 5 shows how changes in the marginal distributions of education, if ignored, may lead to unwarranted conclusions about the evolution of educational assortative mating.

(as we do by estimating eq. [1] separately for different years). While this accounts for shifts in the means, it does not account for changes in the variances of education—which is what confounds the interpretation of γ as measuring educational assortative mating.

2. Longer Time Trend in Educational Assortative Mating

While the CPS data begin in 1962, the decennial censuses allow us to go back to 1940. In figure 7, we compare estimates of the overall educational assortative mating based on the censuses and the CPS. Both sets of estimates suggest that on average, the degree of educational assortative mating increased steadily from the 1960s to the mid-1980s, after which it changed relatively little. What the 1940 census reveals, however, is that educational assortative mating increased considerably prior to the 1960s. In 1940, we estimate that Americans with the same level of education were about 1.3 times more likely to be married to one another as compared to the probability with random mating. By way of comparison, Americans in the 1960s were 1.7 times as likely to be married to someone with the same level of education as compared to the probability with random mating.

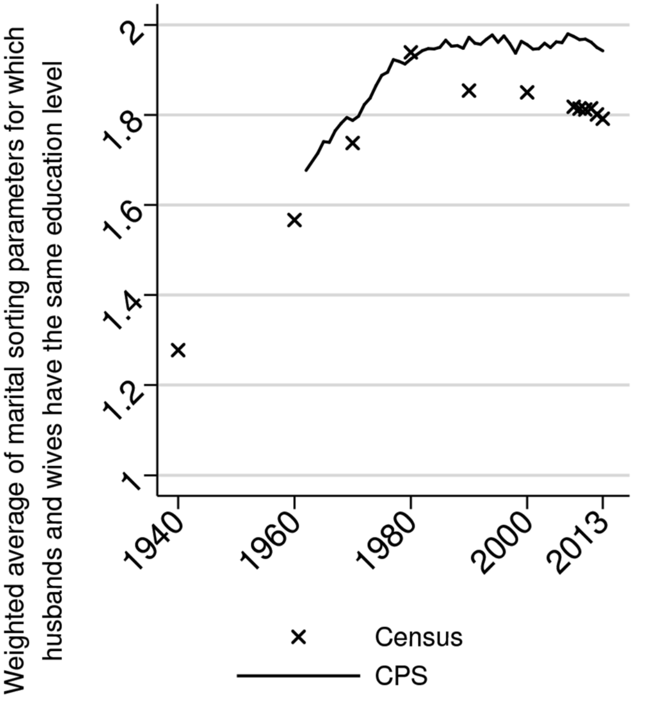

FIG. 7.—Trends in aggregate educational assortative mating using the CPS and the census. This figure displays the time trends in the weighted average of the marital sorting parameters $s(e_f, e_m)$ along the diagonal (where husbands and wives have the same education level). The figure compares results when using the CPS (1962–2013) and the census (1940–2013). Sample: married couples aged 26–60.

3. Educational Assortative Mating by Race and Cohabitation Status

Figure 8 presents estimates of educational assortative mating by race. In this figure, the estimation sample is restricted to same-race couples. Separately for white and black couples, we estimate a full set of the sorting parameters $s(e_f, e_m)$ for each year in the period 1962–2013. To obtain a measure of the overall educational assortative mating by race, we compute the weighted average of the race-specific marital sorting parameters along the diagonal. The patterns are qualitatively similar for blacks and whites: Assortative mating increased gradually from 1962 to the 1980s, after which it has changed relatively little. Among blacks, however, educational assortative mating was relatively low in 1962, but this group experienced a relatively large increase in the assortativeness of marriages until the mid-1980s. In figure D5, we explore the differences by race further,

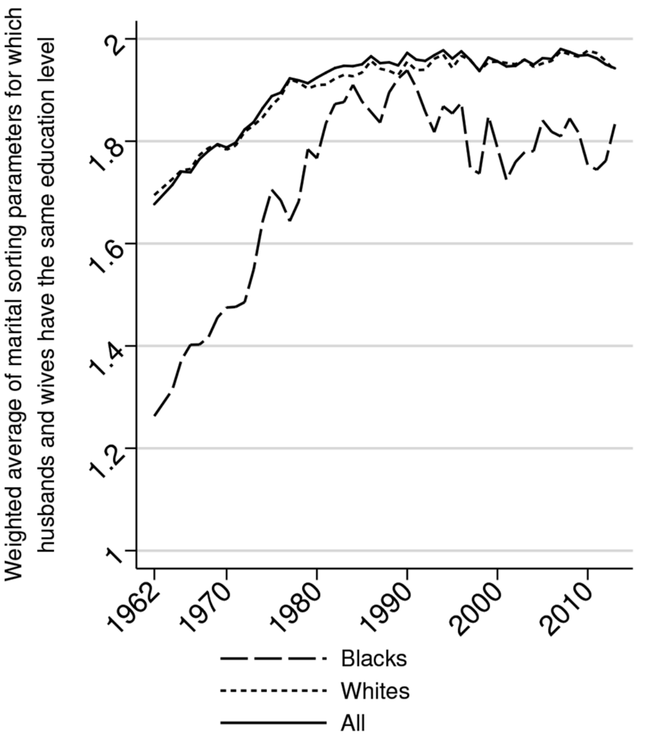

FIG. 8.—US trends in aggregate educational assortative mating, by race. This figure displays the time trends in the weighted average of the marital sorting parameters $s(e_f, e_m)$ along the diagonal (where husbands and wives have the same education level), calculated for all couples and couples in which both are either black or white. Source: CPS (1962–2013), married couples aged 26–60.

presenting race-specific sorting parameters by educational level. The time trends are qualitatively similar by the level of education: For both groups, assortative mating has been declining over time among the college educated, whereas the lowly educated are increasingly sorting into internally homogeneous marriages.

In figure D3, we report estimates of educational assortative mating for married couples and cohabitants. We only perform this analysis for the years 2007–2013, since this is the period in which we observe cohabitation in our data. Overall, educational assortative mating is relatively similar for cohabitants as compared to married couples. A notable exception is the sorting among the lowly educated, which is less pronounced for cohabitants as compared to married couples. In figure D6, we disaggregate educational assortative mating by race and cohabitational status. The prevalence of cohabitation is much higher among blacks (12.9 percent of all black couples) as compared to whites (8.1 percent of all whites couples). However, the differences in educational assortative mating between married couples and cohabitants do not vary significantly by race. Thus, the relatively high cohabitation rate among blacks contributes to lower assortative mating among lowly educated blacks as compared to lowly educated whites.

4. Educational Assortative Mating across Countries

A natural question is whether our findings about educational assortative mating are specific to the United States or whether they generalize to other developed countries. To investigate this question, we examine assortative mating in Denmark, Germany, the United Kingdom, and Norway. The data available in each of these countries allow us to study educational assortative mating over several decades.

For each country and every year, we estimate the sorting parameters $s(e_f, e_m)$ for all possible combinations of education of the husbands and wives. Figure 9 displays the time trends in the sorting parameters on the diagonal (where husbands and wives have the same education level). The estimates correspond to those in figure 3, except we are now using data for countries other than the United States. In each country, we find evidence of positive assortative mating at all levels of education. In addition, the time trends are qualitatively similar across countries: Among college graduates, assortative mating has been declining over time, whereas the lowly educated are increasingly sorting into internally homogeneous marriages. Comparing across countries, the increase in assortative mating among the lowly educated is strongest in the United States. By comparison, assortative mating among college graduates has declined more in the United Kingdom and Norway as compared to the United States.

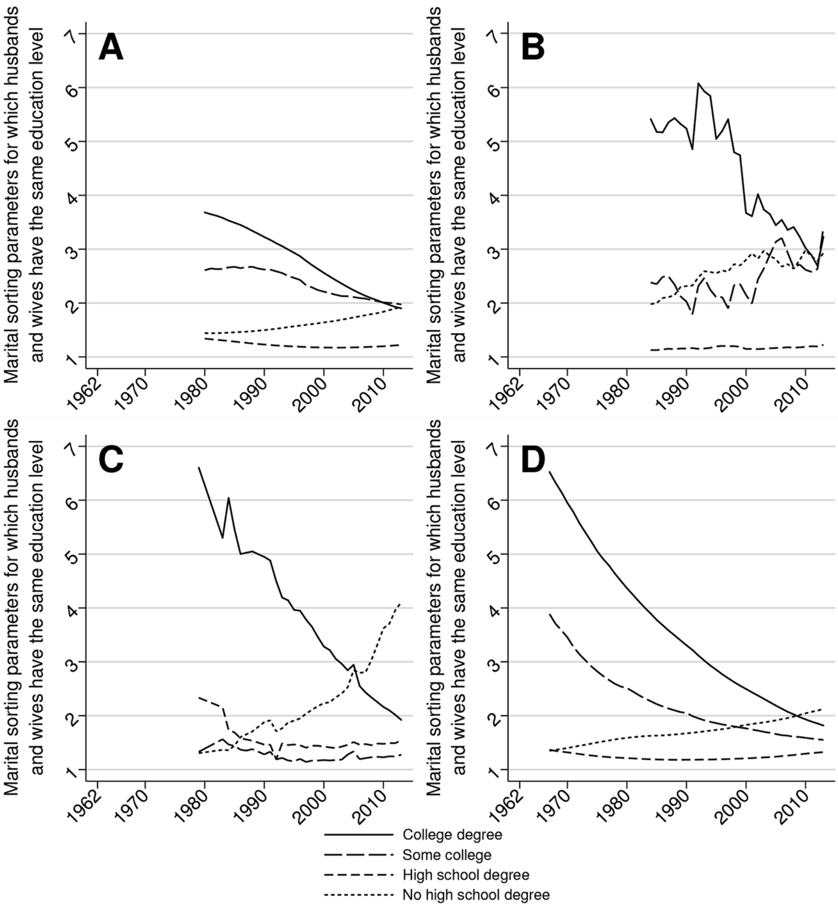

FIG. 9.—Trends in assortative mating by educational level in Denmark (*A*), Germany (*B*), the United Kingdom (*C*), and Norway (*D*). This figure displays the time trends in the marital sorting parameters $s(e_f, e_m)$ for which the husbands and wives have the same education level. Source: Danish registry data (1980–2013); SOEP (Germany, 1984–2013); LFS (United Kingdom, 1979–2013); Norwegian registry data (1967–2013). Sample: married couples aged 26–60.

5. Educational Assortative Mating among the College Educated

So far, we have focused exclusively on assortative mating by education levels. However, the type of education that people acquire is potentially as important as their level of education. For example, the earnings differences observed across college majors rival the earnings gap between individuals with and without a college degree (see, e.g., Altonji, Blom, and Meghir 2012; Kirkeboen, Leuven, and Mogstad 2016). While neither the

CPS data nor the censuses offer information about college major, they do make a distinction between undergraduate and graduate degrees. This allows us to split the college graduates category into two subcategories according to whether college graduates have undergraduate or graduate degrees. Using this new classification with five education groups of the husbands and wives, we estimate a full set of sorting parameters $s(e_f, e_m)$ for all combinations of e_f and e_m.

Figure D7 displays the time trends in the sorting parameters for couples with undergraduate degrees and couples with graduate degrees. We can see that assortative mating declined for both types of the college educated. In 1962, for example, Americans with an undergraduate college degree were 4.2 times as likely to be married to a spouse with an undergraduate college degree, compared to the counterfactual situation in which spouses were randomly matched with respect to education; in 1980 and 2013, they were only 3.0 and 1.9 times as likely, respectively. The decline in assortative mating has been even more pronounced among couples with graduate degrees. In 1962, Americans with a graduate degree were 8.4 times more likely to be married to one another as compared to the probability with random mating; in 1980 and 2013, they were about 5.2 and 3.1 times more likely, respectively.

While the US data do not allow us to look more closely at educational assortative mating among the college educated, the rich Norwegian data record postsecondary field of study. This allows us to split the college category into nine mutually exclusive subcategories by postsecondary field of study. In line with previous evidence, table D5 shows that medicine, law, engineering, science, and business command high income premiums, whereas individuals with humanities, nursing, and education degrees tend to have relatively low income. It is also notable that these income differentials have become more pronounced over time. Figure 10 displays the sorting parameters for couples with the same field of study. This figure reveals that assortative mating is much stronger by field of study than by education level. The assortativeness is strongest for law and medicine, the fields with the highest economic returns. In 1967, for example, a graduate in law was 73 times as likely to be married to a college graduate with a law degree, compared to the counterfactual situation in which spouses were randomly matched. By comparison, college graduates as a whole were only 6.5 times as likely to be married to one another as compared to the probability with random mating. The assortative mating by field of study declines over time but remains sizable. In 2013, graduates in law were still 26 times as likely to be married to one another, relative to the probability under random matching. Taken together, the findings from Norway suggest that the choice of postsecondary field of study could be an important but neglected pathway through which individuals sort into internally homogeneous marriages.

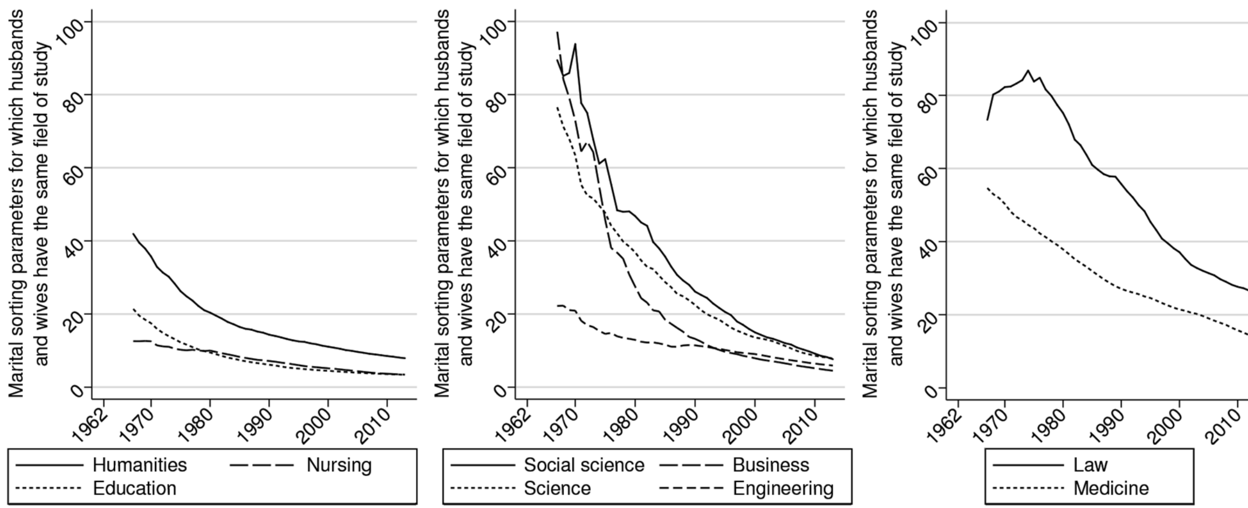

FIG. 10.—Trends in assortative mating by postsecondary field of study in Norway. This figure displays the time trends in marital sorting parameters $s(e_f, e_m)$ for which husbands and wives have the same college major. Source: Norwegian registry data (1967–2013), married couples aged 26–60.

IV. Determinants of Household Income Inequality

A. Decomposition Method

To quantify the contribution to household income inequality of changes in returns to education, educational composition, and educational assortative mating, we adopt the decomposition method proposed by DiNardo et al. (1996). This approach produces income distributions under counterfactual scenarios in which the distribution of one factor is fixed at a base year, while the other factors vary over time.

The joint distribution of household income and couples' education in year t is $F_{Y,X}(y, x|t)$, where y denotes household income and x denotes the couples' educational attainments E_f, E_m. The distribution of income in year t is given by

$$F_Y(y|t) = \int F_{Y|X}(y|x, t)\, dF_X(x|t), \tag{5}$$

where $F_{Y|X}(y|x, t)$ is the conditional distribution of income for couples with characteristics x in year t (i.e., the returns to education) and $F_X(x|t)$ is the joint distribution of spouses' education in year t.

To define the counterfactual scenarios, let t_y denote the year in which the economic returns are measured, t_x denotes the year in which the couples' educational attainments are measured, and t_s denotes the year the marital sorting parameters are measured. Depending on when we measure these three factors, we obtain different counterfactual scenarios. In general, the income distribution under a counterfactual scenario is given by

$$\tilde{F}_Y\left(y|t_y, t_x, t_s\right) = \int F_{Y|X}\left(y|x, t_y\right) \Psi_x\left(x|t_y, t_x, t_s\right) dF_X\left(x|t_y\right), \tag{6}$$

where Ψ_x is a reweighting function defined as

$$\Psi_x\left(x|t_y, t_x, t_s\right) = \frac{d\tilde{F}_X(x|t_x, t_s)}{dF_X\left(x|t_y\right)}, \tag{7}$$

where $d\tilde{F}_X(x|t_x, t_s)$ denotes the joint distribution of spouses' education that would have occurred if the couples' educational attainments were measured in t_x and the marital sorting parameters were measured in t_s.

In the empirical analysis, we hold the distribution of one variable fixed at base year t_0, while we let the distributions of the other factors vary over time. This informs us about how household income inequality is affected by changes in that variable over time. In the case of marital sorting, we construct the counterfactual income distribution if couples matched according to the sorting parameter in base year t_0. For example, $\tilde{F}_Y(y|t_y = t, t_x = t, t_s = t_0)$ represents the income distribution in a scenario in which the returns to education and the educational composition are measured in year t, whereas the marital sorting parameters are measured in year t_0. By comparing this counterfactual income distribution to the actual income distribution in year t, we are able to assess how household income inequality is affected by changes in educational assortative mating between year t_0 and t.

To obtain the counterfactual income distribution, we estimate the reweighting function as

$$\tilde{\Psi}_x\left(x|t_y, t_x, t_s\right) = \frac{\tilde{P}(x|t_x, t_s)}{P\left(x|t_y\right)}, \tag{8}$$

where $P(x|t_y)$ denotes the proportion of couples with educational attainments $x = \{e_f, e_m\}$ in year t_y, and the counterfactual $\tilde{P}(x|t_x, t_s)$ is the proportion of couples who would have had characteristics x if the marginal distributions of education of husbands and wives were as in year t_x and couples matched according to the marital sorting parameter of year t_s. The stochastic matching procedure used to estimate the counterfactual proportion is described in appendix C.

We construct income distributions under alternative counterfactual scenarios, including keeping the marital sorting parameter used to match couples, the education distribution of men and women, or the economic returns to education, fixed at base year t_0.

B. Assortative Mating and Household Income Inequality in the United States

Figure 11 graphs household income inequality over time in the United States. This figure measures inequality according to the much used Gini coefficient. The solid line shows the growth in household income inequality, while the dashed line gives the time trends in household income inequality for a counterfactual scenario in which all the marital sorting parameters $s(e_f, e_m)$ are set equal to 1; this means that men and women with the same level of education marry as frequently as what would be expected under a marriage pattern that is random in terms of education.

As expected, assortative mating leads to an increase in household income inequality. For example, educational assortative matching increased the Gini coefficient in 2013 by 5 percent, from .412 to .432. How large is such an increase in inequality? Abstracting from behavioral responses, a 5 percent increase in the Gini coefficient corresponds to introducing an equal-sized lump sum tax of 5 percent of the mean household income and redistributing the derived tax as proportional transfers in which each household receives 5 percent of its income (Aaberge 1997). Interpreted

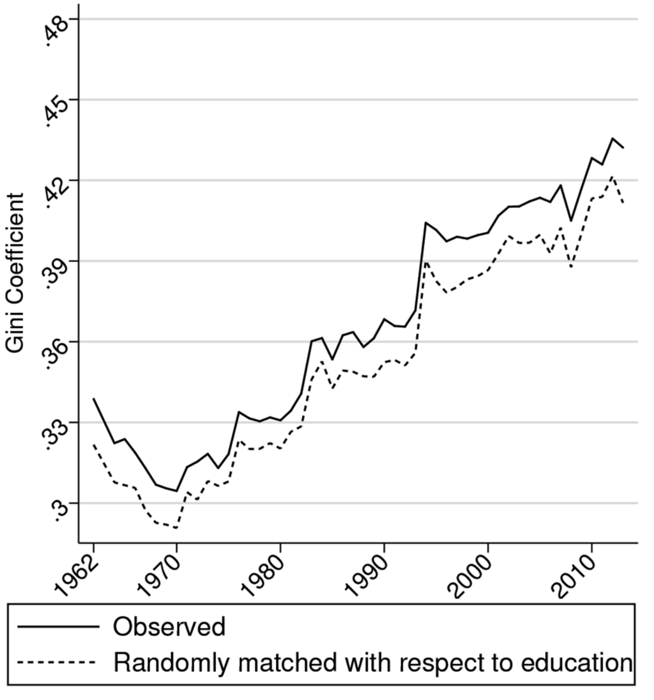

FIG. 11.—US household income inequality and educational assortative mating. The solid lines show the Gini coefficient in the actual distribution of household income. The dashed line shows the Gini coefficient in a counterfactual scenario in which husbands and wives are matched randomly with respect to education. Source: CPS (1962–2013), married couples aged 26–60.

in this way, a key insight from figure 11 is that educational assortative mating has a nonnegligible impact on the distribution of household income in the United States.

Table D6 complements figure 11 by showing how assortative mating affects different parts of the distribution of household income. The 90/10 ratio measures the income at the 90th percentile of the household income distribution relative to that of the 10th percentile, while the 90/50 and 50/10 ratios illustrate whether an increase in the 90/10 ratio is due to the rich getting richer or the poor getting poorer. We compare the percentile ratios in the actual distribution of household income to those that would have occurred if husbands and wives are randomly matched with respect to education. The results suggest that assortative mating matters most for inequality in the lower part of distribution, particularly at the beginning of the period we investigate.

C. Evolution in Household Income Inequality in the United States

We now examine the importance of various factors for the time trend in household income inequality in the United States, including changes in educational assortative mating (fig. 12), returns to education (fig. 13), and educational composition (fig. 14). Each figure compares the actual evolution of household income inequality to the counterfactual levels of inequality, where we hold the distribution of one factor fixed at its level in 1962 while we let the distributions of the other factors vary over time. Figure D8 shows that the conclusions from the decomposition analysis are relatively robust to whether we use 1962, 1984, or 2013 as the base year.

Figure 12 suggests that changes in assortative mating over time matter little for the time trends in household income inequality. This finding refutes the widespread view that changes in assortative mating have led to

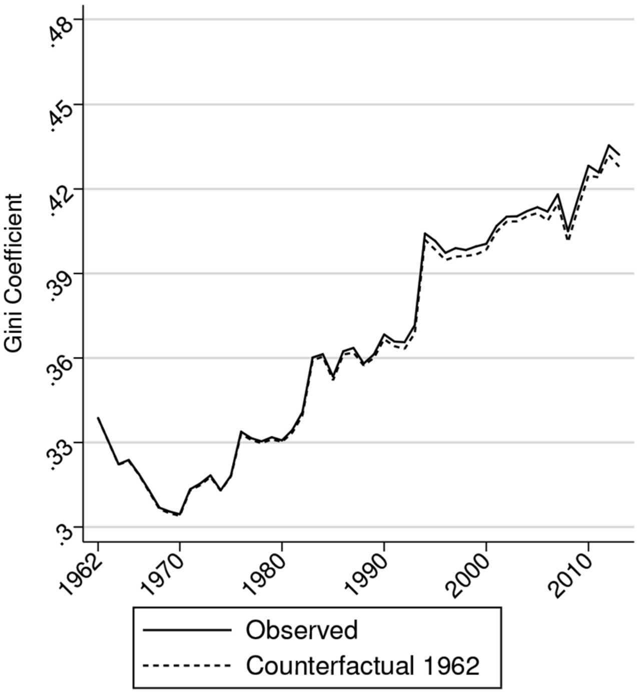

FIG. 12.—US household income inequality and changes in educational assortative mating. The solid line shows the Gini coefficient in the actual distribution of household income. The dotted line shows the Gini coefficient in a counterfactual scenario in which spouses are matched with respect to the 1962 marital sorting parameters, while we let the distributions of the other factors vary over time. Source: CPS (1962–2013), married couples aged 26–60.

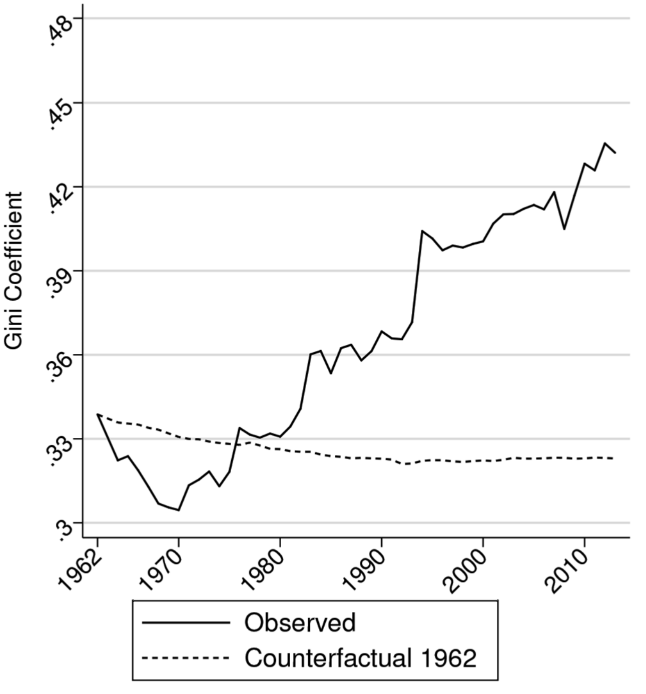

FIG. 13.—US household income inequality and changes in returns to education. The solid line shows the Gini coefficient in the actual distribution of household income. The dotted line shows the Gini coefficient in a counterfactual scenario in which the returns to education are kept fixed at their levels in 1962, while we let the distributions of the other factors vary over time. Source: CPS (1962–2013), married couples aged 26–60.

a rise in household income inequality. The Gini coefficients in the actual and the counterfactual distributions of household income barely differ.

Figure 13 shows that increases in the returns to education seem to be a key driver behind the rise in household income inequality. Indeed, the decomposition analysis suggests that the Gini coefficient in household income would have been steadily declining if the returns to education had remained at their levels in 1962. In 2013, for example, the Gini coefficient is predicted to be 25 percent lower in the absence of changes to the returns to education. This reduction in the Gini coefficient corresponds to introducing a 25 percent proportional tax on income and then redistributing the derived tax revenue as equal-sized amounts to the households (Aaberge 1997). This finding suggests that changes in the returns to education are not only important in explaining the growth in income inequality among males (see, e.g., Autor et al. 2008; Acemoglu and Autor 2011), but also a key factor behind the rise in household income inequality over the past few decades. Table D7 demonstrates that the lower

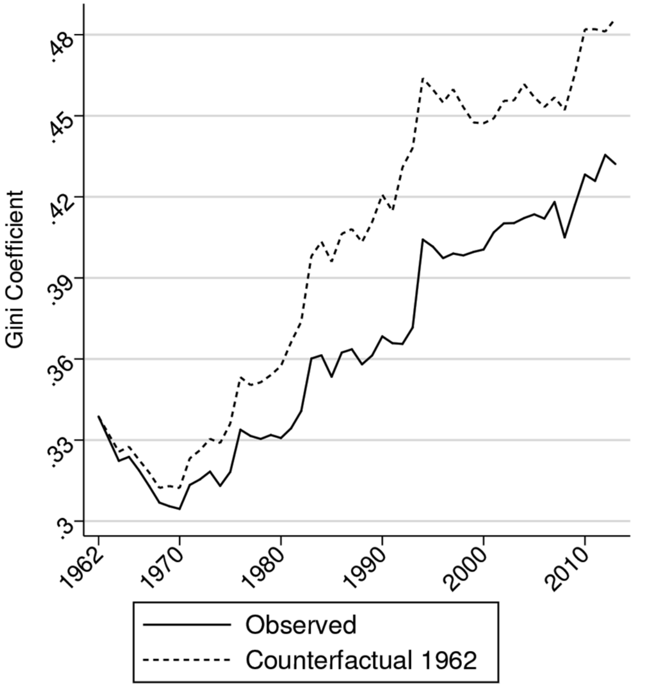

FIG. 14.—US household income inequality and changes in educational composition. The solid line shows the Gini coefficient in the actual distribution of household income. The dotted line shows the Gini coefficient in a counterfactual scenario in which the education distributions of husbands and wives are kept fixed at their levels in 1962, while we let the distributions of the other factors vary over time. Source: CPS (1962–2013), married couples aged 26–60.

part of the household income distribution has been most influenced by changes in education returns. For example, if the returns to education remained at their levels in 1962, we estimate that the 50/10 ratio would have been 47 percent lower in 2013 (a reduction from 4.17 to 2.23), whereas the 90/50 ratio would have been 15 percent lower (a reduction from 2.27 to 1.92).

Figure 14 suggests that changes in the educational composition offset some of the increase in household income inequality. The decomposition results suggests that over the period 1962–2013, Americans would have experienced a 55 percent larger increase in inequality if the education distributions of husbands and wives were as in 1962. In 2013, for instance, we find that the Gini coefficient would have been 12 percent higher in the absence of the changes in educational composition. These compositional effects are distinct from the standard price effects that are often invoked to explain changes in inequality (see, e.g., Juhn, Murphy, and Pierce

1993; Lemieux 2006). Holding returns to education constant, changes in composition can mechanically raise or lower income inequality by changing the population shares of different education groups. Table D8 shows that it is primarily the lower end of the household income distribution that has benefited from the changes in the educational composition.

D. Robustness Checks

We performed several specification checks to examine the robustness of the decomposition results.

1. Adjusting for Age

We perform two checks to examine the sensitivity of the results to accounting for age in the measurement of marital sorting. The first is to restrict the sample to couples in which at least one of the spouses is between 26 and 30 years old in a given year. Figure D2 shows that the estimates of educational assortative mating are quite similar for this subgroup as compared to the full sample. Figure D9 shows that also the decomposition results are quite robust to restricting the sample to young couples in a given year.

As a second robustness check, we use the full sample but assign individuals to groups by the combination of their gender, education, and age. For each gender, we use the three age categories in addition to the four educational levels, dividing the sample into 144 mutually exclusive groups. By comparison, the baseline specification—in which we abstracted from age—gives 16 mutually exclusive groups. Except for the additional groups, we use the same decomposition method as outlined above. Figure D10 shows that the decomposition results barely move when we account for age in the measurement of marital sorting.

2. Adjusting for Changes in Marriage Propensity

We also check whether our results are robust to accounting for changes over time in the likelihood of getting married according to the age and education of males and females. For this purpose, we also include single men and women aged 26–60 in the sample. This allows us to characterize each individual by their age, educational level, gender, and marital status. The inclusion of singles adds another gender-specific sorting parameter for each age and educational group, which represents not being married. Individuals in groups in which this parameter is larger than 1 are more likely to be single than an average individual.

In a counterfactual joint distribution of the household's characteristics, we now let the marginal distribution of education and age as well as the gender-specific probability of being married vary over time. However, couples are matched according to marital sorting parameters of year t_s, which take into account both the relative probability of being married at all (based on gender, age, and education) and who an individual is likely to be married to.

To directly compare the robustness check to the main results, we exclude singles from the actual and counterfactual income distributions. The results from this analysis are presented in figure D11. It is reassuring to find that accounting for changes in the probability of being married by education level does not affect our conclusion: Changes in assortative mating over time continue to barely move the time trends in household income inequality.

We further report results for the entire sample of married and single households.[14] The results from this analysis are presented in figure D12. Panel E displays trends in household income inequality when keeping the probability of being married at a given age fixed as in a base year. This counterfactual allows us to investigate how the rise in singlehood and delay in timing of marriage may have affected household income inequality. The results suggest that the Gini coefficient in 2013 would have been about 7 percent lower if people married at the same rate and at the same age as in 1962.

3. Robustness to Categorization of Education

Figure D13 examines the impact of making a distinction between undergraduate and graduate degrees. We split the college graduates category into two subcategories according to whether college graduates have undergraduate or graduate degrees. Except for the additional groups, we use the same decomposition method as outlined above. We find that the conclusions about household income inequality are not materially affected by whether or not we make a distinction between undergraduate and graduate degrees.

Figure D14 uses the rich Norwegian data set to assess how accounting for heterogeneity by postsecondary field of study affects the evidence on the determinants of household income inequality. This figure shows that the conclusions about the evolution of household income inequality in Norway hold: Changes in educational composition and returns to education remain the key factors, while educational assortative mating continues to play a minor role. This finding is reassuring given that one cannot

[14] To adjust for differences in household size, household income is equivalized using the EU scale. The EU scale divides total household income by the sum of 1 for the first adult, 0.5 for each other adult, and 0.3 for each child under the age of 14.

link spouses in the US data with information on postsecondary fields of study.

E. Evidence from Countries other than the United States

Section III.C showed a common pattern in educational assortative mating across the countries for which we have data. This section shows that the same holds true for the conclusions about household income inequality. For each country, we use the same decomposition method as outlined in Section IV.A. To easily compare the results across countries, we choose common base years, 1984 and 2013. The reason we select these base years is that they are the first and last time periods for which we have data for nearly all countries.[15]

Figure 15 graphs the time trends in household income inequality as measured by the Gini coefficient. The solid lines show the growth in household income inequality, while the dashed lines give the time trends in household income inequality for a counterfactual scenario in which all the marital sorting parameters $s(e_f, e_m)$ are set equal to 1; this means that men and women with the same level of education marry as frequently as what would be expected under a marriage pattern that is random in terms of education. As expected, assortative mating leads to an increase in household income inequality in all countries.

Next, we examine the importance of various factors for the time trends in household income inequality, including changes in educational assortative mating (fig. D15), returns to education (fig. D16), and educational composition (fig. D17). Each figure compares the actual evolution of household income inequality to the counterfactual levels of inequality in a country, where we hold the distribution of one factor fixed at its level in 1984 or in 2013, while we let the distributions of the other factors vary over time. Consistently across countries, we find that changes in assortative mating over time matter little for the time trends in household income inequality. By comparison, changes in the returns to education over time are a key factor behind the evolution of household income inequality.

V. Summary and Discussion of Results

This paper documented the degree of educational assortative mating, how it evolves over time, and the extent to which it differs between countries. Our analysis focused on the United States but also used data from Denmark, Germany, the United Kingdom, and Norway. We found evidence of positive assortative mating at all levels of education in each country.

[15] The only exception is the United Kingdom, for which we lack income data prior to 1993.

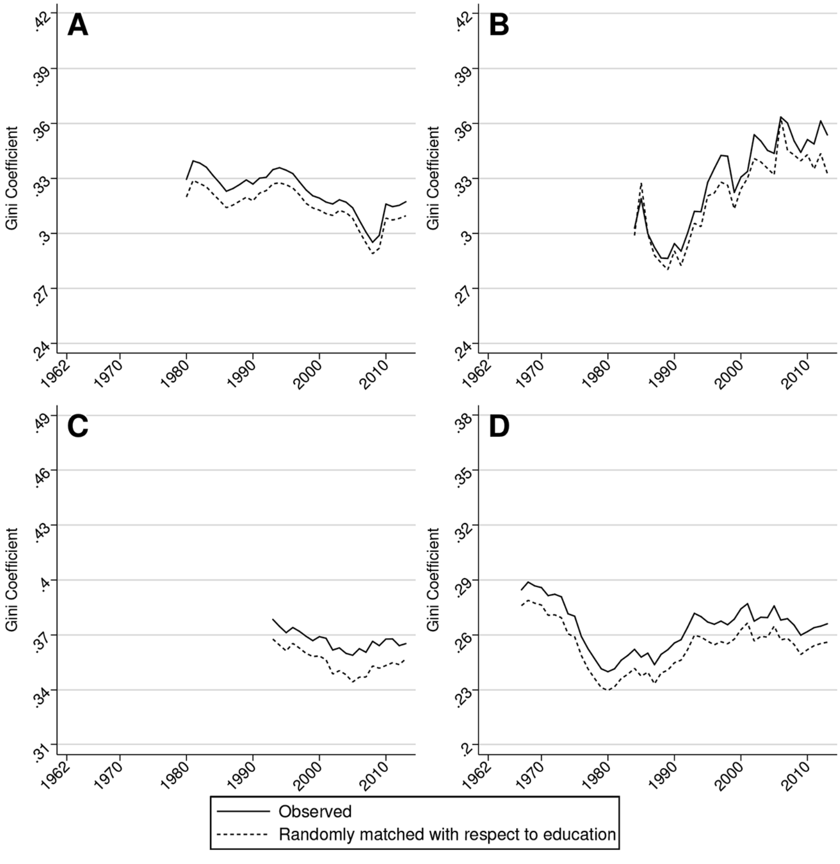

FIG. 15.—Household income inequality and educational assortative mating in Denmark (*A*), Germany (*B*), the United Kingdom (*C*), and Norway (*D*). The solid lines show the Gini coefficient in the actual distribution of household income. The dashed lines show the Gini coefficient in a counterfactual scenario in which husbands and wives are matched randomly with respect to education. Source: Danish registry data (1980–2013); SOEP (Germany, 1984–2013); LFS (United Kingdom, 1993–2013); Norwegian registry data (1967–2013). Sample: married couples aged 26–60.

However, the time trends vary by the level of education: Among college graduates, assortative mating has been declining over time, whereas the lowly educated are increasingly sorting into internally homogeneous marriages.

These findings motivated and guided a decomposition analysis, where we quantified the contribution of various factors to the distribution of household income. We found that educational assortative mating accounts for a nonnegligible part of the cross-sectional inequality in household income in each country. However, changes in assortative mating over time barely move the time trends in household income inequality.

This is because the inequality contribution from the increase in assortative mating among the lowly educated is offset by the equalizing effect from the decline in assortative mating among the highly educated. By comparison, increases over time in the returns to education generate a considerable rise in household income inequality, but these price effects are partly mitigated by increases in college attendance and completion rates among women.

In interpreting these results, it is important to keep in mind the descriptive nature of our analysis. While our study carefully describes educational assortative mating over time and across countries, it is silent on the causes of sorting in marriage. The observed sorting patterns can arise because of several distinct reasons. First, sorting can arise because of search frictions, independently of preferences. For example, sorting along educational attainment might not reflect a preference for a spouse with a certain education level, but rather that people could be more likely to meet someone with a similar level of education in school, college, or at work. Alternatively, sorting can arise in the absence of any search frictions as an equilibrium outcome of preferences and the market mechanism.[16]

Quantifying the importance of the many possible causes of assortative mating is difficult given our data, and it is, therefore, beyond the scope of the paper. Chiappori and Salanié (2016) review the literature on assortative mating, concluding that given only data about matching patterns, it is hardly possible to distinguish between models with frictions and models with unobserved heterogeneity. Given this identification challenge, empirical models of assortative mating typically consider a frictionless environment (see, e.g., Choo and Siow 2006). For example, recent work by Chiappori, Salanié, and Weiss (2017) interpret changes in educational assortative mating over time through the lens of a frictionless matching framework with transferable utility. The paper attributes changes in matching patterns to changes in the additional surplus generated by assortativeness, which in turn are driven by increasing returns to human capital.

In reality, the observed sorting patterns could reflect both search frictions and spouse preferences. Hitsch, Hortaçsu, and Ariely (2010) try to sidestep this identification challenge by examining sorting in the setting of online dating. The authors argue that this setting provides a unique opportunity to study the causes of assortative mating, in part due to search frictions being minimal but also because researchers observe both the choice sets faced by the users and the decisions they make from these

[16] For example, people may prefer a spouse with the same education, in which case sorting is due to "horizontal" spouse preferences. Alternatively, preferences might be purely "vertical," in the sense that everyone ranks all potential spouses in the same way. In the equilibrium of a frictionless market, the ranks of the matched men and women will then be perfectly correlated. If the ranks are monotonically related to the spouse's education level, there will also be sorting by education.

choice sets. Hitsch et al. (2010) begin by estimating the extent to which the actual matches achieved at an online dating site can be explained by preferences alone. Their estimates suggest that preferences are a key cause of sorting. Next, the authors use these estimates to form predictions for off-line matches, i.e., marriages. The predicted sorting patterns are qualitatively, and for some attributes also quantitatively, quite similar to the actual sorting patterns in marriages. The authors argue that the degree to which they underpredict sorting along some attributes—education in particular—is consistent with search frictions. Another possible explanation of the prediction errors is that the preferences expressed in the setting of online dating may not be generalizable to the marriage context of the whole population.

Our decomposition method is also best understood as a descriptive approach, where observed outcomes for one group are used to construct counterfactual scenarios for another group. In constructing these scenarios, we follow the literature on decomposition methods in abstracting from potentially important partial equilibrium considerations (e.g., self-selection into education by comparative advantage) and general equilibrium conditions (e.g., simultaneous determination of education distributions and returns). As a result, we are reluctant to give the decomposition a strict causal interpretation or explore mechanisms that may lead to the patterns we find. We rather think of our analysis as an accounting exercise of the contribution of different factors to inequality. Interpreted in this way, an insight from our analysis is that changes in assortative mating may account for relatively little of the rise in household income inequality. An important question for future research is how the conclusions about educational assortative mating and household income inequality might change if one tries to model the joint determination of premarital schooling, female labor force participation, and marriage patterns of men and women while accounting for search frictions. Greenwood et al. (2016) and Chiappori et al. (2017) make important progress on this question, assuming a frictionless marriage market. However, unless search frictions or other limitations on choice sets are carefully considered, interpretation of preferences based on actual matching patterns could be misleading.

Appendix A

Data

A. Denmark

Our results for Denmark are based on registry data from 1980 to 2013. The mapping of years of education to educational categories is similar to the one used for the United States, with two exceptions. Because of the system for vocational education at the upper secondary level, individuals with 12 to 13.5 years of schooling are classified as having a "high school degree." Individuals with 15 years of education are classified as having a "college degree" as this corresponds to a bachelor's degree in Denmark. Household income is the sum of wages and income from self-employment for the husband and the wife.

B. Germany

We use the German Socio-Economic Panel (SOEP) for each year from 1984 to 2013. The SOEP is a longitudinal panel in which the same households are followed year after year. The sample is weighted to be representative of the West German population until reunification in 1990, and for reunified Germany afterward. Household income is the sum of wages and income from self-employment for the husband and the wife.

The mapping of years of education to educational categories is similar to the one used for the United States. We separate between less than high school, high school, and more than high school. The latter is classified as a college degree for individuals that have completed at least 15 years of education. Note that the educational system in Germany is characterized by a relatively high fraction of individuals with vocational degrees at the upper secondary level and a relatively low fraction of individuals with tertiary education.

C. United Kingdom

The Labour Force Survey (LFS) was carried out in 1979, 1981, and each year starting in 1983. The last year we have data from is 2013. The sample is representative of the UK population and is restricted to individuals aged 26–60 years and their spouses. We drop individuals if the information about education or spousal identifier is missing. The resulting sample, the *education sample*, is used to calculate marginal and joint distributions of education. In addition, since the survey started collecting information about income in 1993, we are constrained to using data from that year onward in the decomposition analysis. The household's income is the sum of the husband's and the wife's gross weekly pay in the main job and gross weekly pay in the second job. Income from self-employment is not reported in the LFS. Households with missing information on income are excluded when analyzing household income inequality (Sec. IV).

Classification of education in the UK data is based on various qualifications that are not directly translated into years of education. Individuals with no qualifications are categorized as having "no high school degree." "High school degree" includes "GCSE A–C or equivalent" and "CSE below grade 1 or equivalent." "Some college" includes "A level or equivalent." "College degree" includes "bachelor's degree or equivalent" and "above bachelor's degree." Individuals with unknown or uncategorized qualifications are assigned a level depending on the age when education was completed.

D. Norway

Our analysis employs several registry databases from 1967 to 2013 that are maintained by Statistics Norway. The data contain unique family identifiers that allow us to link spouses. To enhance comparability with the other countries, we construct a measure of individual income that consists of wages and income from self-employment. In each year we exclude individuals with missing information on income, and set negative incomes to zero. Household income is measured by pooling the individual income of the spouses.

Educational attainment is reported by the educational establishment directly to Statistics Norway, thereby minimizing any measurement error due to misreporting. We have information not only about years of schooling and highest completed degree, but also field of study or academic major in postsecondary education. To enhance comparability between the education systems in the United States and Norway, we make two adjustments to the definition of education levels based on years of schooling. In Norway, certain types of high school degrees require only 10 or 11 years of schooling; we count individuals with these degrees as high school graduates. Several bachelor's degrees in Norway take only 3 years of postsecondary study; we record all individuals with a 3 year or more postsecondary credential as college graduates.

Appendix B

Adjusted Measures of Educational Assortative Mating

We begin by showing how we refine the measures of marital sorting to take spouses' age into account. The martial sorting parameter in equation (3) is the ratio of the probability of observing a couple in which the wife has education e_f and age a_f and the husband has education e_m and age a_m, relative to the probability of observing this couple if matching is random with respect to education. Let the probability that a household has characteristics e_f, a_f, e_m, a_m be denoted $P(e_f, a_f, e_m, a_m)$, which can be rewritten as $P(e_f, e_m|a_f, a_m)P(a_f, a_m)$. If (1) the wife's and husband's educations e_f and e_m are conditionally independent given ages a_f and a_m, and (2) e_f (e_m) is independent of a_m (a_f), then the denominator in equation (3) is

$$\tilde{P}^{\text{random}}\left(e_f, a_f, e_m, a_m\right) = P\left(e_f|a_f\right)P(e_m|a_m)P\left(a_f, a_m\right). \tag{B1}$$

In our analysis, we construct an education-specific measure of sorting as the weighted average of the education- and age-specific measures. For educations e_f, e_m, the age group a_f, a_m is weighted by

$$w\left(a_f, a_m|e_f, e_m\right) = P\left(a_f|e_f\right)P(a_m|e_m)\frac{P\left(a_f, a_m\right)}{P\left(a_f\right)P\left(a_m\right)}. \tag{B2}$$

The first two terms on the right-hand side give us the expected share of couples with ages a_f, a_m among those with educations e_f, e_m, if individuals matched randomly with respect to age. This share is multiplied with the "assortative mating parameter" of the age group a_f, a_m. This product is approximately the share of couples with ages a_f, a_m, among those with educations e_f, e_m, if sorting on age is as in the overall sample. The sorting parameter for education levels e_f and e_m is then

$$\tilde{s}\left(e_f, e_m\right) = \sum_{a_f}\sum_{a_m}\frac{w\left(a_f, a_m|e_f, e_m\right)}{\sum_{a_f}\sum_{a_m}w\left(a_f, a_m|e_f, e_m\right)}s\left(e_f, a_f, e_m, a_m\right). \tag{B3}$$

Positive (negative) assortative mating would mean that men and women with the same level of education match more (less) frequently than what would be expected under a matching pattern that is random in terms of education: that is, the marital sorting parameter $\tilde{s}(e_f, e_m)$ is larger (smaller) than one when e_f is equal to e_m. Our main analysis reports an overall measure of sorting as the average parameter for all groups in which the husbands and wives have similar education. That is simply given by

$$\hat{s} = \sum_e \frac{P(E_f = e)P(E_m = e)}{\sum_e P(E_f = e)P(E_m = e)} \tilde{s}(e, e). \tag{B4}$$

In the same way, we refine the measures of marital sorting to take changes in the probability of marriage into account. We now also include single men and women aged 26–60 in the sample, add separate education and age categories corresponding to single women (no husband) and single men (no wife), and re-estimate the marital sorting parameters. The denominator in equation (3) now corresponds to the marriage pattern if both who marries whom and who is married at all are independent of education. That is, the parameter takes into account both the probability of a match conditional on being married, and the probability of being married at all.

Appendix C

Stochastic Matching Procedure

This appendix describes the stochastic matching procedure we use to estimate $\tilde{P}(x|t_e, t_a, t_s)$, which is the proportion of couples who would have educational attainments e_f, e_m and ages a_f, a_m if the distributions of education were measured in year t_e, the joint distribution of age were measured in year t_a, and couples were formed according to the marital sorting parameters of year t_s. This is a generalized version of the stochastic matching procedure used for our benchmark results.

Consider a finite number of households, N, which initially are matched according to the distribution when assortative mating is random with respect to education. Given the fact that matching is independent of education, the number of households with characteristics e_f, e_m, a_f, a_m is

$$N^0(e_f, e_m, a_f, a_m) = N\left[P(e_f|a_f, t_e)P(e_m|a_m, t_e)P(a_f, a_m|t_a)\right]. \tag{C1}$$

Recall that $s(e_f, a_f, e_m, a_m|t_s)$ is defined as the actual probability of the match relative to the probability under random matching with respect to education. The stochastic matching procedure rematches the women and men in the distribution corresponding to equation (C1) using the marital sorting parameter $s(e_f, a_f, e_m, a_m|t_s)$. The procedure takes two steps:

STEP 1. We draw a man and a woman from the distribution corresponding to equation (C1).

STEP 2. With probability proportional to $s(e_f, a_f, e_m, a_m|t_s)$, the pair is matched and forms a household. With the inverse probability, the pair remains unmatched.

We repeat these steps until all men and women have achieved a match. Whenever a household is formed we adjust the probabilities used in step 1 as the composition of remaining men and women changes.

When producing the counterfactual in which the probability of being married is kept fixed as in a base year, we keep the fractions of households headed by single men, single women, and married couples fixed, while the marginal distributions of women's and men's education and age are as in year *t*. Women and men are assigned to households using the same approach as the one above.

References

Aaberge, R. 1997. "Interpretation of Changes in Rank-Dependent Measures of Inequality." *Econ. Letters* 55 (2): 215–19.

Acemoglu, D., and D. Autor. 2011. "Skills, Tasks, and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, vol. 4, edited by O. Ashenfelter, R. Layard, and D. Card, 1043–1171. Amsterdam: Elsevier.

Agresti, A. 2010. *Analysis of Ordinal Categorical Data*. 2nd ed. Hoboken, NJ: Wiley.

Altham, P. M. E. 1970. "The Measurement of Association of Rows and Columns for an $r \times s$ Contingency Table." *J. Royal Statis. Soc., Series B* 32 (1): 63–73.

Altonji, J. G., E. Blom, and C. Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Ann. Rev. Econ.* 4 (1): 185–223.

Aslaksen, I., T. Wennemo, and R. Aaberge. 2005. "Birds of a Feather Flock Together: The Impact of Choice of Spouse on Family Labor Income Inequality." *Labour* 19 (3): 491–515.

Autor, D., L. Katz, and M. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Rev. Econ. and Statis.* 90 (2): 300–323.

Breen, R., and L. Salazar. 2011. "Educational Assortative Mating and Earnings Inequality in the United States." *American J. Sociology* 117 (3): 808–43.

Burtless, G. 1999. "Effects of Growing Wage Disparities and Changing Family Composition on the US Income Distribution." *European Econ. Rev.* 43 (4): 853–65.

Cancian, M., and D. Reed. 1998. "Assessing the Effects of Wives' Earnings on Family Income Inequality." *Rev. Econ. and Statis.* 80 (1): 73–79.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil. 2011. "Estimating Marginal Returns to Education." *A.E.R.* 101 (6): 2754–81.

Chiappori, P.-A., and B. Salanié. 2016. "The Econometrics of Matching Models." *J. Econ. Literature* 54 (3): 832–61.

Chiappori, P.-A., B. Salanié, and Y. Weiss. 2017. "Partner Choice, Investment in Children, and the Marital College Premium." *A.E.R.* 107 (8): 2109–67.

Choo, E., and A. Siow. 2006. "Who Marries Whom and Why." *J.P.E.* 114 (1): 175–201.

Daly, M. C., and R. G. Valletta. 2006. "Inequality and Poverty in United States: The Effects of Rising Dispersion of Men's Earnings and Changing Family Behaviour." *Economica* 73 (289): 75–98.

DiNardo, J., N. M. Fortin, and T. Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001–44.

Fernandez, R., and R. Rogerson. 2001. "Sorting and Long-Run Inequality." *Q.J.E.* 116 (4): 1305–41.

Fortin, N., T. Lemieux, and S. Firpo. 2011. "Decomposition Methods in Economics." In *Handbook of Labor Economics*, vol. 4, 1–102. New York: Elsevier.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos. 2014. "Marry Your Like: Assortative Mating and Income Inequality." *A.E.R.* 104 (5): 348–53.

———. 2015. "Corrigendum to Marry Your Like: Assortative Mating and Income Inequality." http://www.jeremygreenwood.net/papers/ggksPandPcorrigendum.pdf.

———. 2016. "Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation." *American Econ. J.: Macroeconomics* 8 (1): 1–41.

Hitsch, G. J., A. Hortaçsu, and D. Ariely. 2010. "Matching and Sorting in Online Dating." *A.E.R.* 100 (1): 130–63.

Juhn, C., K. M. Murphy, and B. Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *J.P.E.* 101 (3): 410–42.

Karoly, L. A., and G. Burtless. 1995. "Demographic Change, Rising Earnings Inequality, and the Distribution of Personal Well-Being, 1959–1989." *Demography* 32 (3): 379–405.

Kirkeboen, L., E. Leuven, and M. Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *Q.J.E.* 131 (3): 1057–1111.

Landersø, R., and J. J. Heckman. 2017. "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US." *Scandinavian J. Econ.* 119 (1): 178–230.

Larrimore, J. 2014. "Accounting for United States Household Income Inequality Trends: The Changing Importance of Household Structure and Male and Female Labor Earnings Inequality." *Rev. Income and Wealth* 60 (4): 683–701.

Lemieux, T. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *A.E.R.* 96 (3): 461–98.

Liu, H., and J. Lu. 2006. "Measuring the Degree of Assortative Mating." *Econ. Letters* 92 (3): 317–22.

Long, J., and J. Ferrie. 2013. "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *A.E.R.* 103 (4): 1109–37.

Mare, R. D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Rev.* 56 (1): 15–32.

Nelsen, R. B. 2006. *An Introduction to Copulas*. 2nd ed. New York: Springer.

Pencavel, J. 1998. "Assortative Mating by Schooling and the Work Behavior of Wives and Husbands." *A.E.R.* 88 (2): 326–29.

Schwartz, C. R., and R. D. Mare. 2005. "Trends in Educational Assortative Marriage from 1940 to 2003." *Demography* 42 (4): 621–46.

Western, B., D. Bloome, and C. Percheski. 2008. "Inequality among American Families with Children, 1975 to 2005." *American Sociological Rev.* 73 (6): 903–20.