Educational Assortative Mating and Income Inequality among black and white families in the US, 1976-2017*

Mauricio Bucca

Cornell University

December 20, 2018

Abstract

I re-assess a puzzling finding in the social stratification literature: the null effect of increasing education assortative mating on the takeoff in income inequality. The article uses more sophisticated methods than past work to unravel potentially offsetting forces driving the null-effect of assortative mating. In particular, the article allows for separate analyses for White and Black populations, recognizing that different trends and patterns of assortative mating may have different consequences for income inequality. In addition, it decomposes the overall impact of assortative mating into two components: one due to changes in the marginal distribution of education; another one due to changes in pure assortative behavior, both of which might have evolved differently over time, and by race. Finally, it examines the impact of differential selection into marriage – the married population being increasingly whiter and more educated – on the relationship between assortative mating and inequality. The results provide an exhaustive confirmation of the minor or null effect of the educational assortative mating on income inequality, ruling out some possible explanations for this finding. Results suggest that such null effect is not due to offsetting trends for Blacks and Whites; to countervailing effects of educational expansion and changing assortative behavior; to insufficiently strong changes in assortative mating and selection into marriage; or to the use of methods that are not able to detect complex patterns.

1 Introduction

Social scientists have long been aware of the potential influence of assortative mating –the non-random allocation of individuals into romantic relationships– on income inequality and its intergenerational reproduction (Schwartz, 2013; Rosenfeld, 2008a; Blau and Duncan, 1967). Particular

*THIS IS A PRELIMINARY DRAFT. PLEASE DO NOT CITE OR DISTRIBUTE WITHOUT PERMISSION OF THE AUTHOR.

attention has been devoted to the study of educational assortative mating, given the importance of education for structuring marriage markets and for explaining myriad lifelong outcomes, for individuals as well as their offsprings. However, understanding the role of educational assortative mating as a potential cause of the rise in income inequality remains an elusive scientific endeavor. In particular, researchers have yet to square three findings: the increasing similarity in education among spouses (Schwartz and Mare, 2005b; Cancian and Reed, 1999), the continued growth in income inequality across families, and the null or neglectable effect of the former on the latter (Western et al., 2008; Breen and Salazar, 2011). In other words, why is that the increase of educational similarity among spouses has not translated into an increase in economic dissimilarity across families?

Counter to these trends, sociologists have found that black women face structural pressures toward educational heterogamy as a result of the relative shortage of similarly educated black men. Consequently, American black women – especially the highly educated – are more likely than their white counterparts to "marry down" education-wise and, likely, income-wise. Such tendency towards educational and economic hypogamy might works against the well-studied dynamic of income concentration among educated couples, limiting the capacity of black women to concentrate advantages within the family. This expectation remains a theoretical possibility since, to the best of my knowledge, no study has investigated the link between educational assortative mating and income inequality within the black population. The present article fills this gap by investigating this issue for both white and black Americans.

An additional path through which educational sorting might affect income inequality among married couples is by determining who enters marriage in the first place. In fact, much sociological research reveals profound and growing disparities in marriage chances by race and educational attainment. On a methodological level, this means that existing evidence on marriages is, by necessity, based on a fraction of the US population that is, on average, whiter, more educated, and increasingly so. On a substantive level, selection into marriage based on education shapes the opportunity structure in the marriage market by affecting the availability of potential partners of different educational levels. For these reasons, it is possible that part of what we know about assortative mating – and its relation to rising inequality – is driven by the biased and changing nature of marriage samples. This study examines this possibility by explicitly incorporating selection into marriage as an endogenous factor to the process of assortative mating.

This article goes beyond past work in three regards: first, it expands the study of educational assortative mating and its consequences for inequalities to allow for separate patterns for white and black populations. Second, by combining traditional log-linear models with micro-simulations, this work can decompose the overall impact – or the lack thereof – of educational assortative mating on income inequality into a component due to changes in the marginal distribution of education and a component due to changes in purely assortative behavior. This distinction is absent in existing research. Third, the article examines the impact of differential selection into marriage on the link between assortative mating and inequality.

2 Trends in educational assortative mating

By various accounts, educational assortative mating -and educational homogamy, in particular- has increased in the US during the last decades (Mare, 2008; Schwartz and Mare, 2005b; Greenwood et al., 2014, but Gihleb and Lang, 2016), a trend that also maintains for a number of industrialized countries (Schmidt and Winter, 2008; Breen and Salazar, 2010; Esteve and Cortina, 2006, but Birkelund and Heldal, 2003; Katrák et al., 2006; Breen and Andersen, 2012b).

This literature has hypothesized a number of factors that explain the variation in educational assortative matting trends. These can be broadly divided into changes in individual preferences and changes in the structure of opportunities. Regarding changes in individual preferences, modernization scholars argued that as societies develop individuals' ascribed characteristics (e.g., race, religion, social origin) lose importance as a source of success. Acquired traits, such as education and experience, become more consequential for individual achievement and partner selection. This together with the debilitation of third parties as matchmakers led scholars to predict an increase of educational assortative mating in modern societies (Treiman, 1970; Lash, 1987; Kalmijn, 1991). Conversely, homogamy on ascribed traits, such as race and religion, was expected to decrease (Rosenfeld, 2008b). A parallel line of research maintains that increasing symmetry between men and women may also affect educational assortative mating – less clear is in what direction it does so. On the one hand, as women increase their socioeconomic standing relative to men's they can freely choose a partner on the basis of personal affinity and lifestyles, without being concerned with their economic prospects. In other words, they can afford to choose love (Fernandez et al., 2005). On the other, women's new status might urge men to start competing for highly-educated spouses. If both parties have increasingly similar preferences for high-status partners, then educational assortative mating would increase (Oppenheimer, 1988; Sweeney, 2002a).

Other theories stress the changes in the structure of search for a partner (Xie et al., 2015). On the one hand, chances of intermarriage might depend on the degree of social distance between groups (Blau, 1977). Because the extent of inequality in a society has a correlate in the degree of spatial and institutional segregation (Reardon and Bischoff, 2011), scholars have argued and documented that societies with higher levels of inequality exhibit larger odds of educational homogamy, presumably because in such societies people with similar levels of schooling are more likely to collide in neighborhoods and institutions such as schools and the workplace (Torche, 2010). On the other hand, educational assortative mating may also arise from structural sources. At the most basic demographic level, Mare (1991) argues that educational expansion may increase assortative by lengthening the time and the likelihood of contact between equally educated men and women. Individuals might increasingly meet their partners in their last educational transition (Blossfeld, 2009) and educational institutions may play a "third-party" role on marriage decisions (Schwartz and Mare, 2005b). Similarly, the expansion of female labor participation may have transformed the workspace in an important place of meeting for people with similar education (Blossfeld, 2009).

Available research for the US consistently documents such increase in the odds of educational homogamy due to educational expansion, the narrowing -and lately reverse- of the gender education gap and the advancement of women in the labor market (Mare and Schwartz, 2006; Schwartz,

2010). Moreover, it has been shown that the increase in assortative mating has not been homogeneous across the education distribution but mostly driven by the strengthening of homogamy at both ends of the education ladder (Schwartz and Mare, 2005b). As for non-homogamous couples, there is a general tendency for women to marry a husband with higher educational attainment (Blackwell and Lichter, 2004), although this tendency is expected to decline given changes in the educational attainment of men and women.

Counter to these trends, sociologists have found that black couples are less likely to be homogamous and, in particular, highly educated black women face structural pressures toward educational heterogamy as a result of the relative shortage of similarly educated black men (Spanier and PC, 1980; Maralani, 2013). Consequently, black women are more likely than their white counterparts to "marry down" education-wise and, likely, income-wise. Such tendency towards educational and economic hypogamy might works against the well-studied dynamic of income concentration among highly educated couples, as it may limit the capacity of educated black couples to concentrate advantages within the family.

3 Differential retreat from marriage

By definition, the study of educational assortative marriage requires that individuals indeed enter marriage. However, due to the secular decline of marriage, increasing levels of marital dissolution and the emergence of alternative forms of cohabiting unions (Cherlin, 2004; Teachman et al., 2007) available evidence on the resemblance of spouses is representative of an ever smaller share of American families (Sweeney, 2002b). Moreover, because this retreat from marriage is markedly structured along racial and socioeconomic lines, our knowledge about assortative mating is likely to be biased towards specific groups of the population.

The declining prevalence of marriage can be traced to a myriad of factors, both cultural and structural (Cherlin, 2004). Especially relevant are women's advancements in education and the labor market, which have altered traditional patterns of gender specialization and potential gains from marriage. Although different theories yield different predictions regarding the consequences of these transformations (e.g., (Oppenheimer, 1988)'s career-entry theory vs. Becker (1981)'s specialization and trading model of marriage), empirical evidence generally favors theories that predict a positive effect of women's economic prospect on marriage, supporting the idea of gains-to-marriage and the resulting greater socioeconomic resemblance between partners (Oppenheimer, 1997; Sweeney, 2002b; Schoen and Cheng, 2006; Teachman et al., 2007, but Xie et al., 2003). Yet, for some groups, these transformations "meant that marriage was no longer worth the costs of limited independence and potential mismatch" (Lundberg and Pollak, 2013).

Although these changes affect society at large, the retreat from marriage is markedly heterogeneous by race and socioeconomic status. Evidence amounts to at least five decades of divergence in the marriage patterns of black and white Americans. Blacks have long exhibited lower marriage rates than whites, and this gap is increasing over time due to the rising proportion of never-married black

women but stable prevalence of their white counterparts (Bennett et al., 1989). Consequently, the retreat from marriage has taken on different meaning for the two populations, where whites have mostly retreated from early marriage and blacks have retreated from marriage altogether (Teachman et al., 2007).

One of the most influential explanations for the black-white marriage gap is the hypothesis of a shortage of "marriageable" black men. According to Wilson (1990) the decline of marriage rates and the rise of female-headed families could be explained by the shrinking pool of economically attractive men for black women to marry. Researchers have extended the notion of marriageability – originally formulated in terms of employment rate – to incorporate several dimensions of black male's labor markets and have found mixed support to Wilson's claims (Bennett et al., 1989; Lichter et al., 1991, 1992; Wood, 1995; Raley, 1996; Lichter et al., 2002). This shortage is further exacerbated by the much faster educational advancement of black women compared to black men (McDaniel et al., 2011). As a result, black women face a more difficult marriage market compared to their white counterparts, a finding that holds until today (Cohen and Pepin, 2018).

The retreat from marriage has also unfolded deferentially by socioeconomic status as the decline in marriage rates has occurred more rapidly for those with less education, fewer resources, and less economic opportunities (Lichter et al., 2002; Ellwood and Jencks, 2004; Schoen and Cheng, 2006; Harknett and Kuperberg, 2011). Also in this regard, we observe divergent trajectories with respect to marriage, especially for women: women with more education and resources increasingly cohabit before marriage but marry before conceiving children. In contrast, poor and less educated women are much less likely to marry and thus more likely to rear children in cohabitating relationships (McLanahan, 2004; Lundberg and Pollak, 2013). This socioeconomic divergence in marriage patterns is related to class differences in parenting styles and aggravates inequalities for the next generation (Lundberg and Pollak, 2013). As Schoen and Cheng (2006) put it, the retreat from marriage is being led by those with the least resources.

Furthermore, race and socioeconomic background interact at producing inequalities in marriage rates. For example, in the mid-80's Bennett et al. (1989) found that increased education negatively affected the probability of ever marrying among whites, but it was positively associated with marriage among blacks. Similarly, Schoen and Cheng (2006) find that although marriage rates have changed for most race-education groups, the retreat from marriage is the strongest among blacks with the least education. These trends in marriage patterns have significant consequences for the study of educational resemblance among spouses. Differential retreat from marriage means that, over time, patterns of educational assortative mating among married couples are not only informative of a smaller share of the population but also of a systematically distinct one: whiter, more educated and more affluent. In addition, the fact that educated individuals enter marriage at a higher rate than their less educated counterparts shapes the opportunity structure in the marriage market in ways that favor educational homogamy at the top. Thus, higher educational resemblance among potentially high-earners would mechanically result from the marginal distributions of partners' educational, although it is also plausible that individuals who enter marriage have different preferences for partners' education compared to their unmarried counterparts.

3.1 Educational assortative mating and income inequality

In a context of increased educational similarity among spouses, it is intuitive to expect a consequent increase in income inequality across families. Because of the strong link between an individual's education and her earnings, researchers have hypothesized that the rise of educational similarity would translate into increased income similarity among partners, thus reducing income dispersion within households and augmenting it between them. Such augmented resemblance in partners' earnings could reflect an increased preference for a partner with similar potential earnings – of which education is a proxy – but can also arise as a by-product of other drivers of educational assortative mating such as changes in the structure of opportunities or sorting on the basis of the lifestyles that come along with education. Proponents of this perspective even saw these demographic phenomena as a potential explanation for the take-off in income inequality in the US and other countries (Esping-Andersen, 2007).

However, the effect of increased educational assortative mating on income inequality remains largely unclear. Using decomposition methods, sociologists have found that assortative mating has a null or only marginal impact on income distribution in the United States (Western et al., 2008; Breen and Salazar, 2011), a result that holds in other national contexts (Grotti and Scherer, 2016; Breen and Salazar, 2010). In contrast, through simulation methods, (Greenwood et al., 2014) found that, compared to random mating, education-based assortative mating had a positive effect on income inequality in 2015 but none in 1960. Other researchers report that homogamy has contributed to increasing income inequality in other countries (Schmidt and Winter, 2008; Breen and Andersen, 2012a).

Three main possible explanations have been offered for the non-effect of assortative mating on cross-sectional inequality. First, that the increase in educational assortative mating has not been substantial enough to produce the expected effect on inequality (Breen and Salazar, 2011).

A second possibility is that researchers have used improper methods. In fact, research linking assortative mating to inequality typically relies on decomposition methods (Western et al., 2008; Breen and Salazar, 2011; Breen and Andersen, 2012a), but these suffer from at least two problems: they are often based on linearized, overly-simple measures of educational assortative mating (e.g., correlation between partner's education), which may mask local patterns or offsetting trends (Schwartz, 2013). This limitation might be especially consequential given evidence of increasing homogamy among the most and least educated (Schwartz and Mare, 2005b). In cases when decomposition methods can accommodate discrete educational groups, these methods are unable to differentiate between assortative mating due to the marginal distributions of education and assortative mating due to assortative behavior (e.g. (Breen and Salazar, 2011)). Such limitation represents an important shortcoming since these two sources are informative of distinct social processes.

Finally, another possible explanation for the null effect of increased assortative mating on inequality is the comparatively weaker linkage between wives' educational attainment and their labor market outcomes (Schwartz, 2013). This means that partners that, given their educational attainment have similar earning potential, in practice do not earn so similarly. For most of the past

century, the low proportion of women in the labor force (especially upon marriage) and the high sensitivity of wives' labor supply to their husbands' wages was responsible for this phenomenon (Goldin, 2006). However, as norms regarding traditional gender roles change, and women attain similar (and more) education than men, wives' labor supply should continue on the path of more strongly depend on their own potential wages rather than their husband's. In turn, increased educational assortative mating is expected to have a larger impact on economic inequality among families.

4 The present study

In the present study, I re-assess the impact of educational assortative mating and income inequality among married couples in the US. Unlike previous research, I examine this relationship separately for black and white populations, as known differences in patterns of assortative mating lead one to expect different consequences for inequality. In addition, I analyze the impact of differential selection into marriage by race and education as an endogenous factor to the process of assortative mating.

Methodologically, the article combines log-linear models and micro-simulation to overcome some of the limitations of decomposition methods used in previous research. In particular, the use of micro-simulation allows me to take advantage of the rich characterization of patterns and trends in assortative mating that log-linear models offer. This strategy permits separating partners' educational resemblance and its consequences for inequality into a part due to the marginal distribution of partners' education and a component due to purely assortative behavior. Furthermore, it allows me to treat the marginal distribution of spouses' education as a product of both differential selection into marriage and race-specific patterns of educational expansion.

In substantive terms, this analytic strategy allows me to disentangle between the concomitant processes, outlined above, by which assortative mating influences income inequality. By theories that emphasize gains-to-marriage, one would expect education to have a positive effect on the chances of marriage, thus facilitating educational homogamy among the most educated. This, combined with the educational advancement of men and – especially – women lead one to expect a decline in educational homogamy at the lower end of the educational spectrum but an increase at the top end. These effects would all arise from changes in the marginal distribution of partners' education. Complementary, theories that argue for an increased preference for similarly educated (or maximally educated) partners lead to the same expectations, but in this case, changes would be reflected in purely assortative behavior, above and beyond the structure of opportunity. Regarding the consequences for inequality, one might expect that an increase in the prevalence of educational homogamy would favor income concentration within households, thus augmenting income inequality across families. Given known patterns of marriage and assortative mating, these expectations are likely to apply especially to the white population.

In contrast, by theories of black marriage market shortage, one would expect both lower mar-

riage rates for black women and, for those who enter marriage, a higher likelihood of educational hypogamy ("marrying down"). According to this theory, such shortage arises from the marginal distribution of education for black men and women and not from a higher "propensity" towards these behaviors. In the absence of such shortage, it is to be expected, the prevalence of educational homogamy within the black population would be higher, favoring a concentration of similar earners within households and boosting inequality across families.

5 Analytic Strategy

I address the impact of educational assortative mating on income inequality for black and white Americans by means of counterfactual simulations. These simulations create artificial couples based on assumptions regarding selection into marriage and into a partner, and aim to answer questions of the following type: how much income inequality across families would we observe if those who get married today had the same socioeconomic characteristics of married people in the late 70's, at the beginning of the inequality take-off? How much inequality would we observe if educational resemblance among partners had not increased over the last four decades? What would happen to family income inequality if blacks were as likely as whites to marry similarly educated partners? The counterfactual assumptions in these simulations come from logistic regression models for selection into marriage and log-linear models that characterize trends and patterns of assortative mating. The analysis unfolds in three consecutive stages, which I detail next.

5.1 Modeling of selection into marriage

This stage of the analysis investigates changes over time in the stratifying effect of education and race with respect to marriage. More specifically, I use logistic regression to model the probability of being married as a function of period, educational attainment, race, and gender. In addition, I include a cubic polynomial to adjust for age. Models are estimated separately for men and women, blacks and whites. Moreover, both year and educational attainment are modeled as discrete variables and are allowed to interact. Estimates from these models allow me to compute, for an individual i, of gender g and race r, her predicted probability of being married (m) – i.e., entering the sample for the study of assortative mating – as a function of her education l, survey year k and age. Formally:

$$\hat{m}_{igr} = logit^{-1}(\hat{\alpha}_{gr} + \sum_{l}^{L} \hat{\beta}^{E}_{grl} * 1_l + \sum_{k}^{K} \hat{\beta}^{Y}_{grk} * 1_k + \sum_{lk}^{LK} \hat{\beta}^{EY}_{grlk} * 1_{lk} + A'_{igr}\hat{\delta}_{gr}) \tag{1}$$

where 1's are indicator variables, $\hat{\beta}^{E}_{l}$, $\hat{\beta}^{Y}_{k}$ and $\hat{\beta}^{EY}_{lk}$ are regression coefficients for categories of education, year and their interaction (indexes l and k do not include reference categories), A is

a vector of age terms and $\hat{\delta}$ the corresponding coefficient estimates. Estimates of the probability of being married are later used to assign individuals to a marriage status based on counterfactual assumptions.

5.2 Modeling of educational assortative mating

In the second stage of the analysis, I characterize the pattern of assortative mating and its evolution over time using log-linear models for contingency tables. These models provide estimates of the association of couples' education while controlling for changes in the marginal distribution of educational attainment. More specifically, I analyze the 225 cells three-way table yielded by cross-classifying husband's education (H: 5 categories), wife's education (W: 5 categories) and year of the survey (Y: 9 categories). Analyses are conducted separately for blacks and whites (R: 2 categories). Equation 2 describes a saturated model for the contingency table resulting from the cross-tabulation of these three variables for each race.

$$\log F_{ijkr} = \lambda_{0r} + \lambda^{H}_{ir} + \lambda^{W}_{jr} + \lambda^{Y}_{kr} + \lambda^{HY}_{ijr} + \lambda^{WY}_{ikr} + \lambda^{HW}_{jkr} + \lambda^{HWY}_{ijkr} \tag{2}$$

Here F_{ijk} corresponds to the count in the ijkth cell of the table, where i and j index different levels of educational attainment of husband and wife, respectively, k indexes survey years and r indicates the race of belonging. The marginal distributions of these variables are given by λ^{H}_{ir}, λ^{W}_{jr} and λ^{Y}_{kr}, while changes over time in the marginal distribution of partners' education are denoted by λ^{HY}_{ikr} and λ^{WY}_{jkr}. Finally, λ^{HW}_{ijr} and λ^{HWY}_{ijkr} captures association in partner's educational attainment both at the baseline and subsequent years, respectively.

In search for a most parsimonious yet accurate characterization of educational assortative mating, I test several model specifications, each corresponding to a different hypothesis regarding the pattern and evolution of the mating structure. More specifically, I run different specifications of *homogamy* and *crossing models* (Goodman, 1972). These include homogamy models both with constrained and unconstrained main diagonal, models with symmetric and asymmetric movements across the minors diagonals, crossing models and combinations of these (details in Appendix A.1).

In order to select a model specification that accurately describes patterns of assortative mating without over-fitting the data[1], I rely on traditional goodness of fit statistics – e.g., Deviance, Dissimilarity Index, Akaike and Bayesian Information Criteria–, as well as k-fold cross-validation (details in Appendix A.2). Because parameter estimates will serve as inputs for simulations, predictive accuracy is a critical consideration in the model selection process.

Based on the predicted counts yielded by the preferred model in this analysis, I compute, for each individual, the conditional probability of marrying a partner of each educational level given the

[1]That is, fitting a model that uses too many parameters relative to the number of observations.

person's own education, race and period. Formally,

$$
\begin{aligned}
p^{W}_{i|jkr} \mid \hat{\theta} &= P(H = i \mid W = j, Y = k, R = r, \hat{\theta}) = \frac{\hat{F}_{ijkr}}{\sum_i \hat{F}_{ijkr}} \\
p^{H}_{j|ikr} \mid \hat{\theta} &= P(W = j \mid H = i, Y = k, R = r, \hat{\theta}) = \frac{\hat{F}_{ijkr}}{\sum_j \hat{F}_{ijkr}}
\end{aligned} \tag{3}
$$

Thus, for each wife and husband $\tilde{p}^{W}_{jkr} : \{p^{W}_{1jkr}, \ldots, p^{W}_{5jkr}\}$ and $\tilde{p}^{H}_{ikr} : \{p^{H}_{i1kr}, \ldots, p^{H}_{i5kr}\}$ denotes the respective probability that she/he will marry a partner of every level of educational attainment given her/his own characteristics. Importantly, these estimates are conditional on parameter estimates from the model chosen to describe assortative mating ($\hat{\theta}$), which will be manipulated in counterfactual simulation.

5.3 Simulations using counterfactuals

The third part of the analysis investigates the impact of selection into marriage and education assortative mating on trends in income inequality and differences across races.

For this purpose, building on results from the previous stages of the analysis, I implement a series of counterfactual scenarios in which individuals are stochastically assigned into marriage or singlehood. Next, for those assigned into marriage, I form single-race artificial couples on the basis of individuals' own education, the education of their potential partners and survey year[2]. Once artificial couples have been formed, I pool their income at the couple level and measure income inequality using the Theil entropy index. I assess the influence of selection into marriage and educational assortative mating by comparing observed inequality to inequality under different scenarios.

Different counterfactual scenarios are implemented to investigate the partial contribution of specific aspects of the marriage and mating structure to income inequality across black and white families (e.g., racial gaps in marriage rates, educational expansion, changes is assortative behavior). Operationally, this is achieved by replacing parameters estimates from the marriage and assortative mating models ($\hat{\theta}$ and $\hat{\beta}$, respectively) by a modified version that impose counterfactual assumptions on specific facets of this process (e.g. "if homogamy mating today were as in 1975"), while maintaining others as observed. Throughout the article I refer to these as $\tilde{\beta}_s$ and $\tilde{\theta}_t$, where the tilde indicates that these are modified parameter vectors and s and t index particular counterfactual scenarios regarding marriage and assortative mating, respectively.

More specifically, these simulations rewire the marriage network by manipulating who is connected and to whom according to the following set of rules:

[2] I also incorporate an age range restriction so that artificial couples resemble observed age differences among partners of their same race, educational level and survey year.

i. I determine whether a person is married or not with a random draw from a Bernoulli distribution that uses the counterfactual probability of marriage as a parameter. Formally,

$$\tilde{M}_{igr} \sim Bernoulli(\tilde{m}_{igr} \mid \tilde{\beta}), \text{ where } \tilde{M}_{igr} \in \{\text{Married}, \text{Single}\}$$

As mentioned above, the probability of marriage depends on gender (g) and race (r), educational achievement (l), survey year (k) and age.

ii. If a person is assigned to marriage, I determine the educational level that is required for a potential partner of his/her with a random draw from a Multinomial distribution. The parameter of this distribution is the vector of counterfactual probabilities of marrying a same-race partner of each educational level ($\tilde{p}_{jgr}$). These probabilities depend on the person's own educational attainment, gender and survey year. More specifically, the probabilities described in Equation 3 are adjusted using weights that account for the fact that counter-factual selection into marriage changes the marginal distribution of spouses' education. When analyses are conducted on the sample of actual marriages, these weights take value 1. Formally,

$$\tilde{SH}^{W}_{jkr} \sim Multinomial(\tilde{p}^{W}_{jgr} \mid \tilde{\theta}, \tilde{\beta}), \text{ where } \tilde{SH}^{W}_{jgr} \in \{10-11, \ldots, \geq 16\}$$

$$\tilde{SW}^{H}_{ikr} \sim Multinomial(\tilde{p}^{H}_{igr} \mid \tilde{\theta}, \tilde{\beta}), \text{ where } \tilde{SW}^{H}_{igr} \in \{10-11, \ldots, \geq 16\}$$

Here $\tilde{SH}^{W}_{jkr}$ corresponds to the educational level the wife j will look for in a husband, given her education, race and survey year. $\tilde{SW}^{H}_{ikr}$ is the analogous for husbands.

iii. I assign each wife w and husband h to an artificial same race partner such that he/she is observed in the same survey year, has the required educational achievement ($\tilde{SH}^{W}_{igr}$ or $\tilde{SW}^{H}_{igr}$), and a realistic age difference. Here $\tilde{c}_w : \{w, \tilde{h}\}$ denotes an artificial couple where a husband is probabilistically assigned to a wife, and $\tilde{c}_h : \{h, \tilde{w}\}$ denotes an artificial couple where a wife is assigned to a husband.

iv. Next, the whole sample of artificial marriages – $\tilde{C}$ – is obtained by combining a 50% random sample of couples were the wife is the seed (denote set of indexes by W') and a 50% random sample of couples were the husband is the seed (denote set of indexes by H'). This procedure permits to implement the partner search as initiated by both men and women while preserving the marginal distributions of variables for wives and husbands, without artificially duplicating the sample size. Formally,

$$\tilde{C} : (\bigcup_{w \in W'} \tilde{c}_w) \bigcup (\bigcup_{h \in H'} \tilde{c}_h)$$

v. Finally, I compute family income inequality under counterfactual scenario st using the Theil Entropy Index. Formally,

$$T^{st} = \frac{1}{N} \sum_{c=1}^{N} \frac{y_c^{st}}{\mu^{st}} \ln \frac{y_c^{st}}{\mu^{st}}$$

The comparison of actual inequality T to T^{st} informs us about the impact on income inequality of the specific aspects of selection into marriage and assortative mating.

This approach allows me to take advantage of the rich tradition of log-linear modeling in sociology and incorporate it to the study of the effects of assortative mating on income inequality, thus overcoming the problem of oversimplification implicit in decomposition methods. Following the intuition behind standard permutations tests (Dufour et al., 2015), I implement a sufficiently large number of iterations for each scenario to summarize uncertainty around these counterfactual inequality estimates.

Table 1: Counter-factual scenarios

Mating (t) / Marriage (s)	0. Observed	1. As observed	2. At 1975	3. B-W switch	4. Random
0. Observed	(0,0)	x	x	x	x
1. As observed	(1,0)	(1,1)	(1,2)	(1,3)	(1,4)
2. Margins at 1975	(2,0)				
3. Assortative behavior at 1975	(3,0)				
4. B-W switch in margins	(4,0)				
5. B-W switch in assortative behavior	(5,0)				
6. Random	(6,0)				

In particular, as detailed in Table 1, my analysis implements these counterfactual scenarios in two stages: in the first set of simulations I manipulate different aspects of assortative mating but take the sample of married individuals as given. In other words, these scenarios combine factual patterns of selection into marriage with counterfactual assumptions regarding educational assortative mating. In these scenarios changing patterns of assortative mating can affect income inequality by altering the availability of potential partners of different educational levels (i.e., the marginal distributions) and/or individual propensities to marry different types of partners (i.e., assortative behavior). Column 2 in Table 1 details and indexes these scenarios.

The second set of simulations implement counterfactual assumptions regarding selection into marriage while keeping assortative mating "as observed"[3]. In these scenarios, selection into marriage might affect income inequality by shaping the opportunity structure for couples to form. In other words, if selection into marriage occurred in a different way than what observed, the distribution of education for married men and women would also change, thus altering the changes in mating based on education. It is also plausible that changing patterns of selection into marriage might translate into different patterns of assortative behavior, as non-married individuals might have different preferences for partners' education compared to their married counterparts. However, since assortative behavior is also observed for married individuals, this aspect of assortative mating cannot be counterfactually manipulated. Row 2 in Table 1 details and indexes these scenarios.

In further detail:

[3]Note that counterfactuals regarding selection into marriage cannot be combined with actual patterns of assortative mating because, by necessity, these are only available for observed couples.

- (1,0) *Mating as observed*: Artificial couples are formed probabilistically according to the observed pattern of assortative mating. This scenario aims to account for the fact that simulations are based on manipulations of a *model* for assortative mating. Consequently, the model's imperfect fit would induce a degree of randomness in the process of partner selection, potentially leading to a reduction in inequality. For this reason, the results from this scenario provide a benchmark to interpret the output of other simulations, in which specific features of the assortative mating model are counterfactually changed.
- (2,0) *Margins at 1975*: For each race and gender, the marginal distribution of education is kept at levels observed in 1975, the beginning of the studied period. This scenario aims to assess the effect of educational expansion – as an ingredient of the mating structure - on income inequality.
- (3,0) *Assortative behavior at 1975*: For each race and gender parameters that describe pure assortative behavior are kept at levels observed in 1975. This scenario aims to assess the effect of pure assortative mating on income inequality.
- (4,0) *Black-white switch in margins*: This scenario simulates income inequality for each race assuming that, in any given year, whites have the marginal distribution of blacks and vice versa. The primary aim of this scenario is to test the effect of the marriage market shortage for educated black women on family income inequality.
- (5,0) *Black-white switch in assortative mating*: This scenario simulates income inequality for each race assuming that, in any given year, whites exhibit the assortative behavior of blacks and vice versa.
- (6,0) *Random mating*: In this scenario, artificial couples are formed assuming that educational mating is not assortative and agents in the marriage market are uniquely constraint by the marginal distribution of education. The difference between observed income inequality and the one observed under this scenario provides an upper bound for the effect of education-based assortative behavior on income inequality. This scenario is equivalent to the simulation implemented by Greenwood et al. (2014).
- (1,1) *Marriage and Mating as observed*: Individuals are assigned into marriage according to the observed marriage patterns and, next, artificial couples are formed probabilistically according to the observed pattern of assortative mating. The comparison between this scenario and counterfactual (1,0) – actual marriage and mating as observed – is informative of how much randomness in marriage assignment is induced by the model's imperfect fit, and how this affects income inequality. Therefore, this scenario provides a benchmark for results from other simulations regarding selection into marriage, because these manipulate specific features of the marriage model.
- (1,2) *Marriage at 1975 and Mating as observed*: For each race and gender marriage rates are fixed at levels observed in 1975 and patterns of assortative mating are kept as observed. This scenario aims to assess the effect of the generalized retreat from marriage on income inequality.

- (1,3) *Black-white switch in marriage and Mating as observed*: This scenario simulates income inequality for each race assuming that, in any given year, whites have the marriage rates of blacks and vice versa, while patterns of assortative mating are kept as observed. The primary aim of this scenario is to test the effect of the lower and more unequal marriage chances of blacks on income inequality.
- (1,4) *Random marriage*: In this scenario individuals' chances of marriage are constrained by the marginal distribution of marriage, but any other source of selection is removed, including education. The difference between observed income inequality and the one observed under this scenario provides an upper bound for the effect of education-based selection into marriage on income inequality.

6 Data and Measures

I use the public use Annual Social and Economic Supplement of the Current Population Survey (ASEC-CPS hereafter) to study the link between marriage, educational assortative mating and income inequality for the two major racial groups in the US over the period 1976-2017.

For this purpose, I restrict the sample to black or white individuals (90% and 10% of the sample, respectively) who are aged 20 to 64 and are the householder and spouse of householder (N=3,128,309). I use this sample for the study of selection into marriage. Furthermore, for the study of patterns of educational assortative mating I use the sub-sample of prevailing marriages[4] in which both husband and wife are 20 to 64 years old (N = 1,131,635), the woman is no older than her partner by more than 25 years, the man is no older than his partner by more than 30 years and none of the partners is under the age of 18. For both substantive and practical reasons, I exclude mixed-race couples, therefore limiting the study to exclusively black or exclusively white couples. Given the extremely low prevalence of mixed race couples in my samples (0.34% of marriages), this choice does not threaten the generalizability of the results to the American demographic context, broadly defined.

The primary measure of the study are individuals' educational achievement and earnings. To facilitate comparison with previous studies I follow Schwartz and Mare (2005b) and Breen and Salazar (2011) and measure educational attainment as completed years of schooling using the following five categories scheme: less than grade 10 ("less than 10"), grades 10-11 ("10-11"), grade 12 ("12"), one to three years of college ("13-15") and four or more years of college ("16 or more"). The measure of individual income consists of wages/salary and income from self-employment (in 2017 US$). In turn, family income is obtained by adding the income of spouses. I treat zero income as a meaningful value, as it is informative of the economic standing of individuals and families. Finally, to ensure an adequate sample for all analyzed periods, I collapse survey years into half-decades.

[4]Unmarried partner can only be identified starting 1995.

7 Findings

7.1 Descriptive Statistics

Descriptive statistics reveal clear temporal changes in marriage and assortative mating, as well as black-white differences in these two dimensions. As shown in Figure 1, singleness has monotonically increased across genders and races, but its prevalence is significantly larger within the black population, especially among black women. Black women are more than twice as likely than white women to be single, and the ratio of black to white men is only slightly smaller. Nowadays, singleness is the modal state among black adults, while marriage is the modal status for white adults.

Among those who are married, homogamy – that is, having an equally educated partner – is the most common event, regardless of race. Importantly, the share of educationally homogamous couples is increasing over time, especially among whites. Trends in heterogamy – having a partner with different educational attainment–, on the other hand, differ across black and white populations. Among whites, until the 2000's it was more common for the husband to have more education than the wife, a situation that recently reversed as a consequence of the faster educational advancement of women. In contrast, for blacks, the prevalence of marriages where the wife has more education than the husband is higher than the opposite case over the entire time period.

Finally, desegregation of homogamous marriages reveals countervailing trends, with a sharp increase in the share of highly educated homogamous couples and a decline of homogamy at low educational levels. These trends hold for both black and white marriages but are more marked for whites'. Taken together, these trends in assortative mating likely reflect both black-white differences and temporal changes in the marginal distributions of education and assortative behavior. Log-linear analyses in section 7.3 will disentangle these two processes.

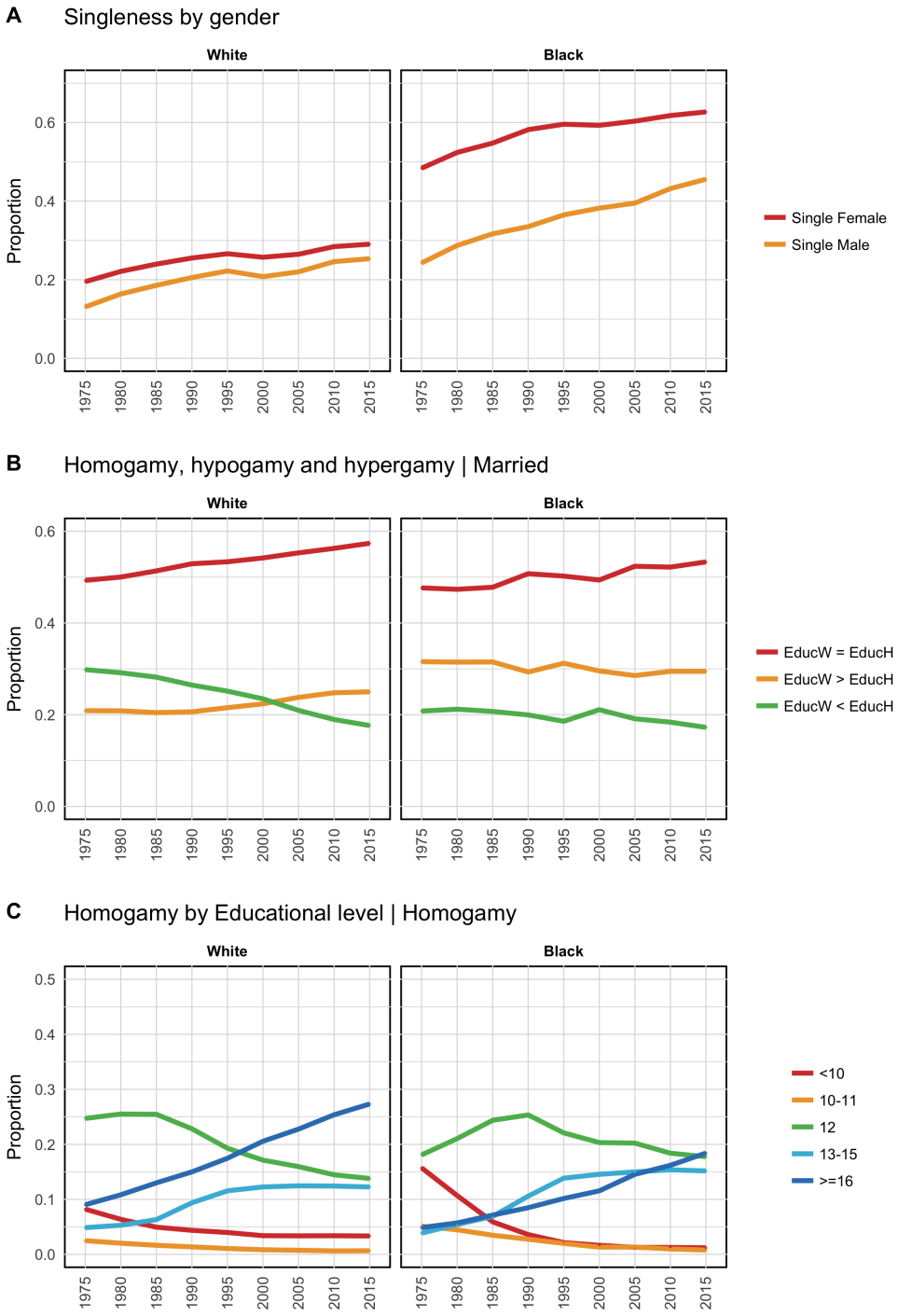

Figure 1: Descriptive statistics

The stylized facts reported here provide preliminary support for three motivating statements: first, that marriage is structured by race. The next section will further explore stratification by educa-

tional attainment, indicating that selection into marriage is an endogenous factor to the process of assortative mating. Second, that educational homogamy is rising among the most educated, especially among whites. Third, that the patterns of educational heterogamy differ by race, where black women are generally more likely to have more education than their husbands.

7.2 Trends and patterns of selection into marriage

In this section I present findings regarding trends and patterns of selection into marriage by race and educational attainment. Results from logistic regression models confirm that education and race stratify marriage chances and, for most of the population, these are declining over time. Whites are significantly more likely to be married than blacks, but the gap is much larger among women. In addition, people with more education are generally more likely to be married than those with less education and these differences are widening over time. This trend is mostly driven by the rapid decline of marriage among individuals with 10 to 11 years of schooling, and the relative stability of marriage rates among college graduates (16 years or more). I also find that education and race interact at producing different chances of marriage: the stratifying effect of education is much weaker for whites than it is for blacks. Moreover, the importance of education for marriage changes is increasing at a faster rate for the black population, especially black men. By contrast, among white college graduate females, the likelihood of marriage shows a slight increase since the mid-90's. Finally, except for black women, these results indicate that college graduates were once the least likely to be married but are now the most likely to do so.

Figure 2 displays these results.

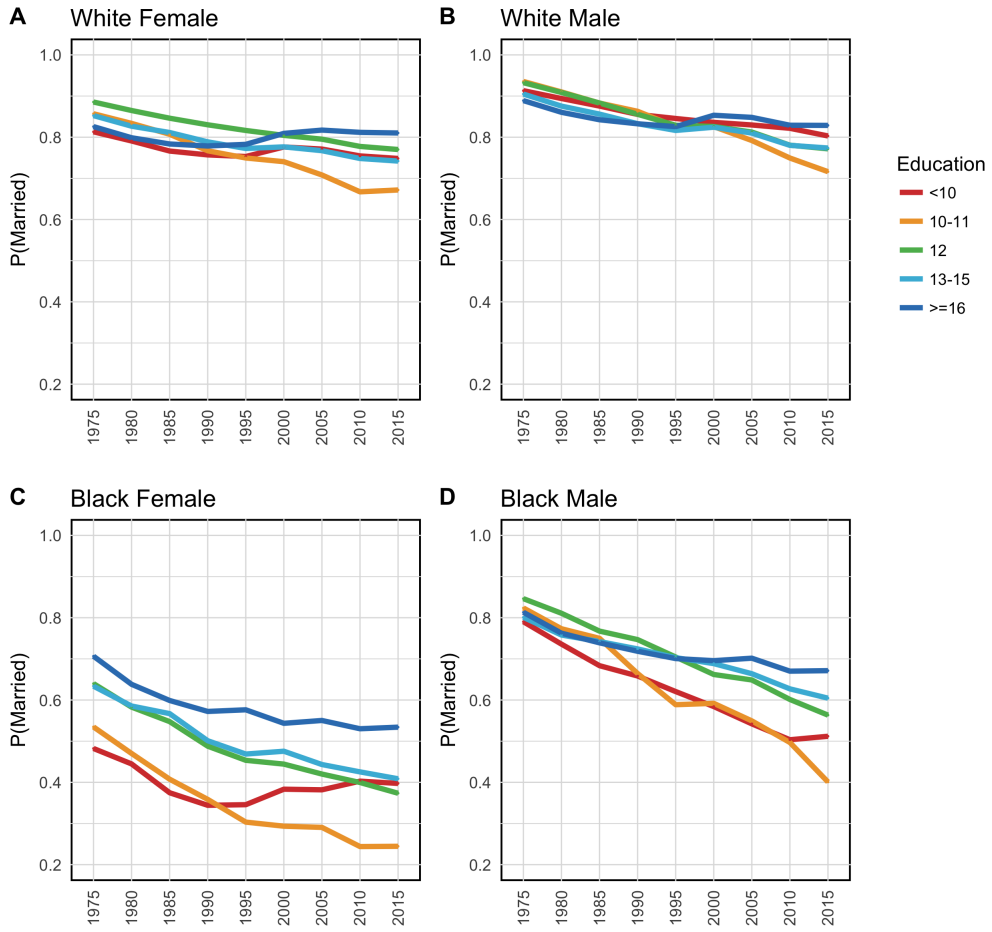

Figure 2: Predicted probability of being married at age 42 (mean) by race, gender and educational attainment

These findings have several implications for the study of educational assortative mating and family income inequality: first, the much higher rate of marriage among whites implies that studies of assortative mating in the US disproportionately represent the patterns of this group. In fact, in the general sample the white-black ratio is 8.9, while in the married sample this number is 14.3. This selection bias is aggravated by the faster decline of marriage among blacks compared to whites. Second, the fact that marriage is more strongly stratified along educational lines for blacks implies that the mating patterns we observe for black couples disproportionately represent the patterns of highly educated black Americans. Third, widening inequality in marriage by educational attainment and race indicates that this selection bias has become more severe in recent periods: compared to past decades, marriage is now systematically more likely to occur among whites, and among people with more education. Again, this issue is the most evident for blacks, but it also holds for whites.

Moreover, because of the connection between education and earnings, these trends in selection

into marriage mean that existing evidence on the relation between assortative mating and family income inequality is likely biased towards processes that take place at the top of the education and income distributions.

Finally, these findings also suggest that patterns of educational assortative mating would be different if marriage chances were structured in different ways. The reason is that education-based selection into marriage shapes the opportunity structure in the marriage market (i.e., the availability of potential partners of different educational levels), thus mechanically affecting individuals' chances of mating. In particular, given that highly educated individuals are increasingly more likely to enter marriage compared to their less educated counterparts – especially among blacks– , one might expect less educational homogamy among the highly educated and lower levels of income inequality across marriages in the absence of such selection patterns.

7.3 Trends and patterns of educational assortative mating

This section explores trends and patterns of educational assortative mating within black and white adult populations. For this purpose, I implement several log-linear model specifications, each corresponding to a particular hypothesis regarding the patterns and evolution of the mating structure. Diagnostics yielded by traditional goodness-of-fit statistics and a cross-validation procedure suggest that the pattern of educational assortative mating is best represented by Model 8 in Tables 2 and 3.

This model allows different extents of educational homogamy at different educational levels as well as changes over time in the strength of homogamy. It also incorporates asymmetric movements along the minor diagonals, which indicate that partners are likely to differ by one level of education and that education hypogamy and hypergamy vary in extent (e.g., the likelihood of observing a couple where the husband has one more level of education than his wife is different than that of finding a couple where the wife has one more level of education than her husband). This result holds in both black and white populations and is only partially consistent with existing research.

In the case of white couples, judging by BIC alone, one would choose Model 11, a specification that incorporates crossing barriers to intermarriage and a constrained diagonal for homogamy. Such choice would be consistent with previous research that finds crossing models to outperform different variants of homogamy models at describing assortative mating trends in the US (Schwartz and Mare, 2005b). However, statistics that impose less severe or none penalties to model complexity do not yield the same diagnostic. In fact, by all other measures, the best performing model is Model 8. As for assortative mating among black spouses, BIC prefers a simple model (Model 3), where educational assortative mating is entirely explained by the marginal distribution of partners' education, changes over time in these distributions and cross-sectional interaction between husband's and wife's education. Notably, this model implies that there are no relevant trends in assortative behavior. Such preference is partially explained by the fact that the sample of black marriages is rather small, a situation in which BIC tends to prefer parsimonious but sometimes ill-fitted models (Clogg, 1982; Weeden and Grusky, 2005). The choice of a specification for black

marriage patterns is further complicated by the fact that each fit statistic shows a preference for a different model.

In addition, I evaluate all models through k-fold cross-validation. By assessing the predictive performance of models out-of-sample, cross-validation prevents the problem of over-fitting, thus serving the same purpose of traditional goodness of fit statistics (details in section A.2 in Appendix). The last column of tables 2 and 3 reports the cross-validation error of each model candidate, which indicates that the most predictive model is Model 8, both for black and white marriages.

Table 2: Log-Linear Models of the Association Between White Husbands and Wives Educational Attainment: US, 1975-2015

Model	df	G2	100*D	AIC	BIC	CV-error
(1) Indep = [W][H][Y]	208	690359.09	30.36	689943.09	690581.04	1387234.85
(2) Indep-Year = [WY][HY]	144	551718.31	28.02	551430.31	552828.05	1371927.84
(3) CI = [W][H][Y][HW][WY][HY]	128	3050.93	1.66	2794.93	4382.61	1313857.06
(4) CI+[CDiag][CDiagY]	120	2131.08	1.40	1891.08	3573.74	1313757.85
(5) CI+[Diag][DiagY]	88	738.54	0.64	562.54	2625.09	1313622.01
(6) CI+[CDiag][CDiagY][Sym][SymY]	112	2104.29	1.39	1880.29	3657.92	1313758.97
(7) CI+[CDiag][CDiagY][Asym][AsymY]	104	1836.54	1.33	1628.54	3501.14	1313732.33
(8) CI+[Diag][DiagY][Sym][SymY]	80	632.75	0.58	472.75	2630.28	1313611.11
(9) CI+[Diag][DiagY][Asym][AsymY]	72	407.90	0.37	263.90	2516.40	1313592.48
(10) CI+[Cross][CrossY]	96	851.91	0.72	659.91	2627.49	1313642.59
(11) CI+[Cross][CrossY][CDiag][CDiagY]	88	569.84	0.58	393.84	2456.39	1313605.56
(12) CI+[Cross][CrossY][Diag][DiagY]	72	469.03	0.45	325.03	2577.53	1313599.41
(13) Saturated = [WHY]	0	0	0	0	3107.26	1313561.94

W: Wife's Education, H: Husband's' Education, Y: Year, CDiag: Main Diagonal constrained
Diag: Main Diagonal unconstrained, Sym: Symmetric movements minor diagonal
Asym: Asymmetric movements minor diagonal, Cross: Crossing parameters.

Table 3: Log-Linear Models of the Association Between Black Husbands and Wives Educational Attainment: US, 1975-2015

Model	df	G2	100*D	AIC	BIC	CV-error
(1) Indep = [W][H][Y]	208	49870.43	30.87	49454.43	50049.78	1387235.72
(2) Indep-Year = [WY][HY]	144	30683.34	25.26	30395.34	31580.11	1371929.54
(3) CI = [W][H][Y][HW][WY][HY]	128	324.72	2.00	68.72	1400.85	1313855.05
(4) CI+[CDiag][CDiagY]	120	273.87	1.62	33.87	1439.67	1313762.39
(5) CI+[Diag][DiagY]	88	184.53	0.91	8.53	1709.04	1313625.95
(6) CI+[CDiag][CDiagY][Sym][SymY]	112	259.24	1.62	35.24	1514.72	1313757.00
(7) CI+[CDiag][CDiagY][Asym][AsymY]	104	242.31	1.54	34.31	1587.46	1313733.00
(8) CI+[Diag][DiagY][Sym][SymY]	80	154.54	0.82	-5.46	1768.72	1313614.72
(9) CI+[Diag][DiagY][Asym][AsymY]	72	137.23	0.68	-6.77	1841.09	1313594.32
(10) CI+[Cross][CrossY]	96	204.85	1.26	12.85	1639.68	1313641.60
(11) CI+[Cross][CrossY][CDiag][CDiagY]	88	154.70	0.99	-21.30	1679.21	1313609.22
(12) CI+[Cross][CrossY][Diag][DiagY]	72	132.50	0.77	-11.50	1836.36	1313600.84
(13) Saturated = [WHY]	0	0	0	0	2510.95	1313566.38

W: Wife's Education, H: Husband's' Education, Y: Year, CDiag: Main Diagonal constrained
Diag: Main Diagonal unconstrained, Sym: Symmetric movements minor diagonal
Asym: Asymmetric movements minor diagonal, Cross: Crossing parameters.

Consequently, I choose Model 8 to characterize trends and patterns of educational assortative mating of black and white couples. This specification is the recommended choice for whites by almost all fit statistics, and it is a plausible candidate for blacks. Moreover, it is the model with better predictive performance according to cross-validation. This consideration is of critical importance, since parameter estimates will serve as inputs for counterfactual simulations.

Figures 3 and 4 depict the trends implied by the preferred model, both in terms of educational assortative mating due to changes in the marginal distributions of partners' education (Panel A) and

assortative mating due to assortative behavior (Panel B)[5]. As observed, assortative mating arises from both sources but trends over time are mostly dominated by changes in the marginal distributions of education. Because a high school degree (grade 12) is the modal educational attainment of men, in the absence of assortative behavior, having a high school graduate husband would be the most likely outcome for women of all educational levels, followed by having a husband with "some college" (grade 13-15). However, due to the continued educational advancement of women, all types of couples involving a wife with little education are expected to decline over time, while the share of highly educated couples is expected to increase. These patterns are comparable for black and white populations, but there are some important differences to note: first, the expected prevalence of homogamy among the highly educated is larger for white compared to black couples. Second, the share of highly educated homogamous couples is expected to increase faster among black couples than whites. Similarly, the fraction of poorly educated homogamous couples is expected to decline at a faster rate in the black population.

Regarding assortative behavior, I find evidence of educational homogamy, above and beyond what is expected due to the marginal distributions. Homogamy occurs at almost all educational levels, especially at both ends of the spectrum: people with 10 or fewer years of schooling, as well as people with 16 or more years of schooling, are much more likely to marry someone of the same educational level[6]. Homogamy due to assortative behavior is similar across races but somewhat higher in the white population.

In addition, marrying a partner one level above or one level below the own education is also more likely than what it would be expected in the absence of assortative behavior. Again, this tendency is stronger at the two ends of the educational spectrum. Although the model allows for asymmetric heterogamy, these results display symmetric movements along the minor diagonals. For example, a black man with a college education (16 years or more) has roughly the same odds of marrying a woman with some college (13-15) than does a black woman with a college education to marry a man with some college (13-15). In the case of white couples, however, college-educated men exhibit slightly higher odds of marrying a woman with some college than college-educated women are of marrying a partner with some college. Regarding temporal variation, patterns of assortative behavior remain, for the most part, stable over time for both black and white couples. There is, however, evidence of increasing homogamy at the two lowest educational groups and a slight decline in homogamy for highly educated black couples.

Overall, these results highlight comparability in the trends and patterns of assortative behavior across black and white marriages as well as over time. These findings make it unlikely that this aspect of assortative mating would have a significant influence on black-white differences and

[5]Given the difference in sample size, the model intercept is removed from both figures to facilitate comparison

[6]It is important to notice that the results presented in these figures are not strictly interpretable in terms of odds of assortative behavior. The reason is that the reference category in the chosen log-linear model corresponds to couples where both partners have "less than elementary school" in 1990 but, counter-intuitively, the indicator variable for homogamy takes on value zero. This problem arises from the model specification and commonly occurs in applications of log-linear models (e.g., (Schwartz and Mare, 2005b, p.636)). As a result, the intercept does not have clear substantive meaning, thus complicating the interpretation of the remaining parameters. Nevertheless, this issue does not affect findings regarding *trends* in assortative mating nor those regarding the *relative* size of parameters.

temporal trends in income inequality. On the contrary, patterns of assortative mating due to the marginal distribution of education might affect income inequality among black and white marriages.

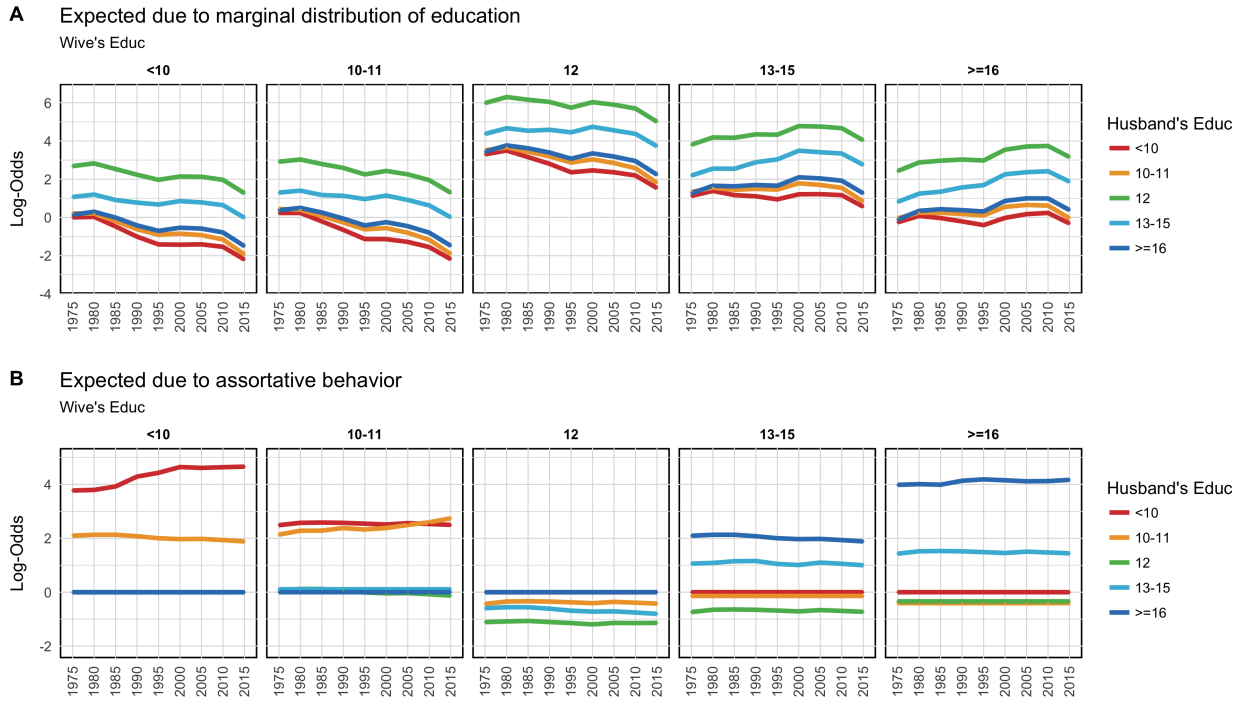

Figure 3: Patterns of educational assortative mating among white couples

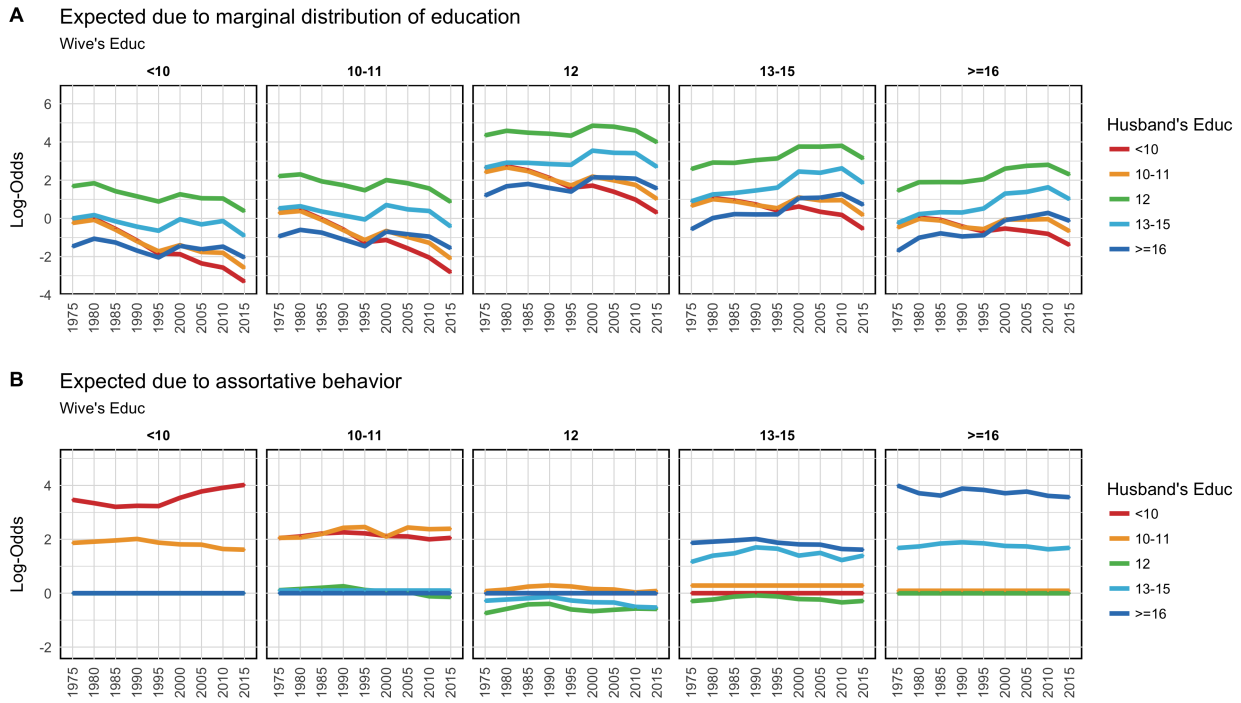

Figure 4: Patterns of educational assortative mating among black couples

7.4 The impact of educational assortative mating on income inequality

Regarding the impact of educational assortative mating on income inequality, counterfactual simulations yield the following main findings:

- The comparison between observed inequality (0,0) and inequality in the scenario in which educational assortative mating is kept "as observed" (1,0) suggests that randomness in partner selection induced by imperfect model fit has a negative impact on income inequality. Because the model fit is poorer in the black population (see Tables 2 and 3), the gap is larger in this case.

- Against what is expected by some scholarship, the increase in partners' educational similarity driven by educational expansion has not reinforced income inequality. On the contrary, if the marginal distribution of education remained as it was in 1975, income inequality would be higher or similar to what it is today (2,0). This finding holds for both whites and blacks but is acuter for the latter. This result is likely related to the sharp decline in the prevalence of black marriages in which both partners have very low educational attainment.

- Temporal changes in assortative behavior exert no influence on income inequality (3,0). This result is consistent with the finding of no temporal change in patterns of educational assortative behavior.

- These results do not support the intuition that the marriage market shortage for educated black women might have the effect of maintaining inequality lower that it would be in the absence of such shortage: if the marginal distribution of education for blacks were that of whites, inequality would have remained "as observed", not higher. The same would be the case if the marginal distribution of education for whites were that of blacks (4,0).
- Black-white differences in assortative behavior are found to exert some – small – influence on income inequality (5,0). According to these results, the white pattern of assortative behavior is more inequality enhancing than the blacks pattern. This result might reflect some of the minor black-white differences in assortative behaviors, such as the somewhat higher propensity to educational homogamy among white couples.
- Finally, if there were no assortative behavior, above and beyond the constraints imposed by the marginal distribution of education (6,0), income inequality would be reduced by about 13%-14% with respect to observed levels. This reduction is more or less constant over time for blacks but increases over time for whites. The result for the white population is consistent with what Greenwood et al. (2014) found using similar methods. In relation to the benchmark scenario, this reduction is, however, smaller: about 7% in the latest period for whites and about 4% for blacks.

Figure 5 displays these results.

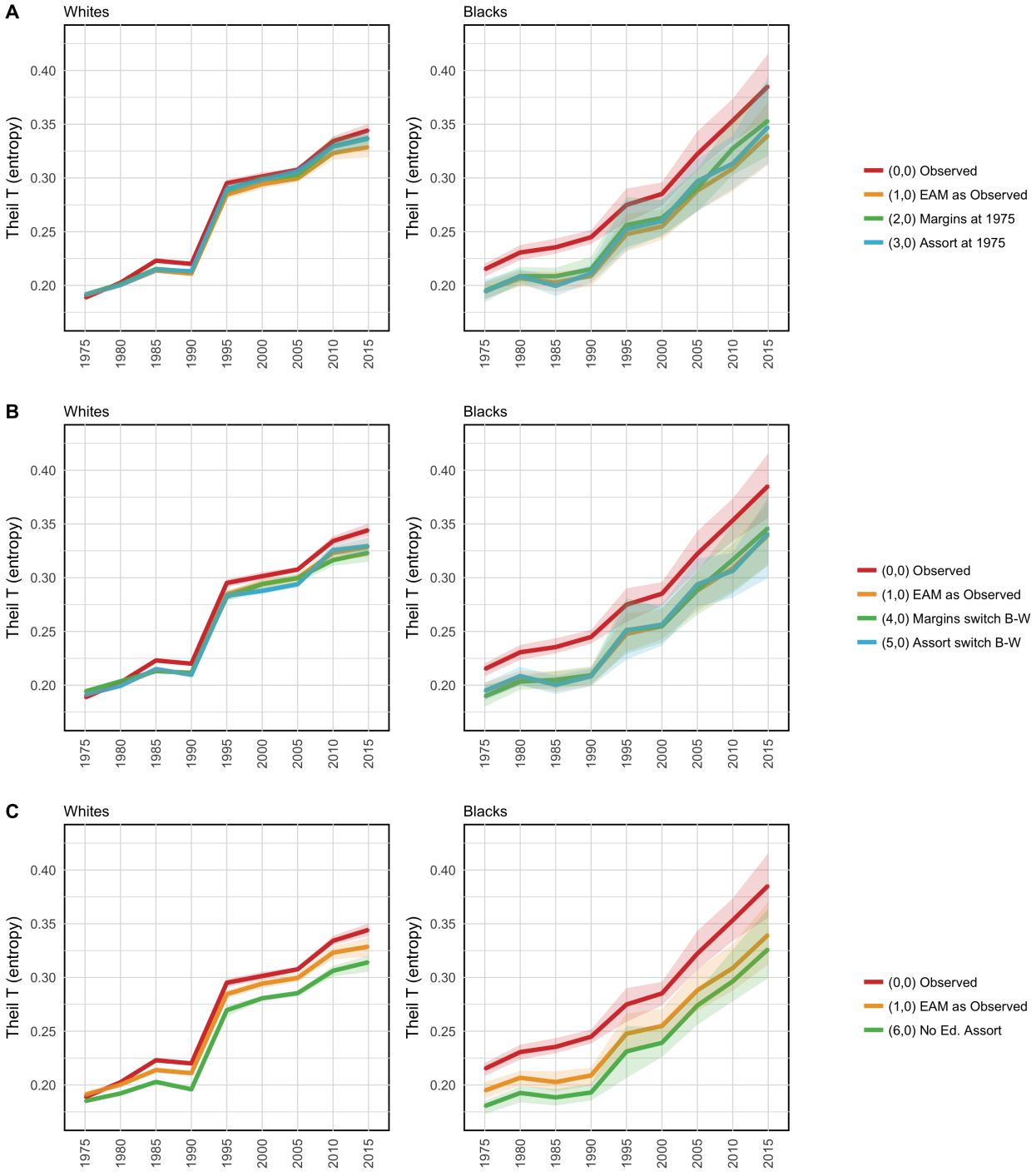

Figure 5: Income inequality under counterfactual scenarios

7.5 The impact of selection into marriage on income inequality

Regarding the impact of selection into marriage on income inequality, counterfactual simulations yield the following main findings:

- For whites, selection into marriage seems to have no or a very minor impact on income inequality. This is likely the combined result of two things: the comparatively mild retreat from marriage among all white educational groups and, as shown in the previous section, the limited impact of educational sorting on income inequality.
- In the case of the black population, the large difference between scenario (1,0) – observed marriages but mating "as observed – and scenario (1,1) – marriage and mating "as observed" – likely indicates poor predictive capacity of the marriage model. If that is the case, this scenario approaches a situation in which marriage changes are largely random. These results suggest that, under this assumption, inequality is expected to be significantly larger than if selection into marriages is taken as it is.

Figure 6 displays these results.

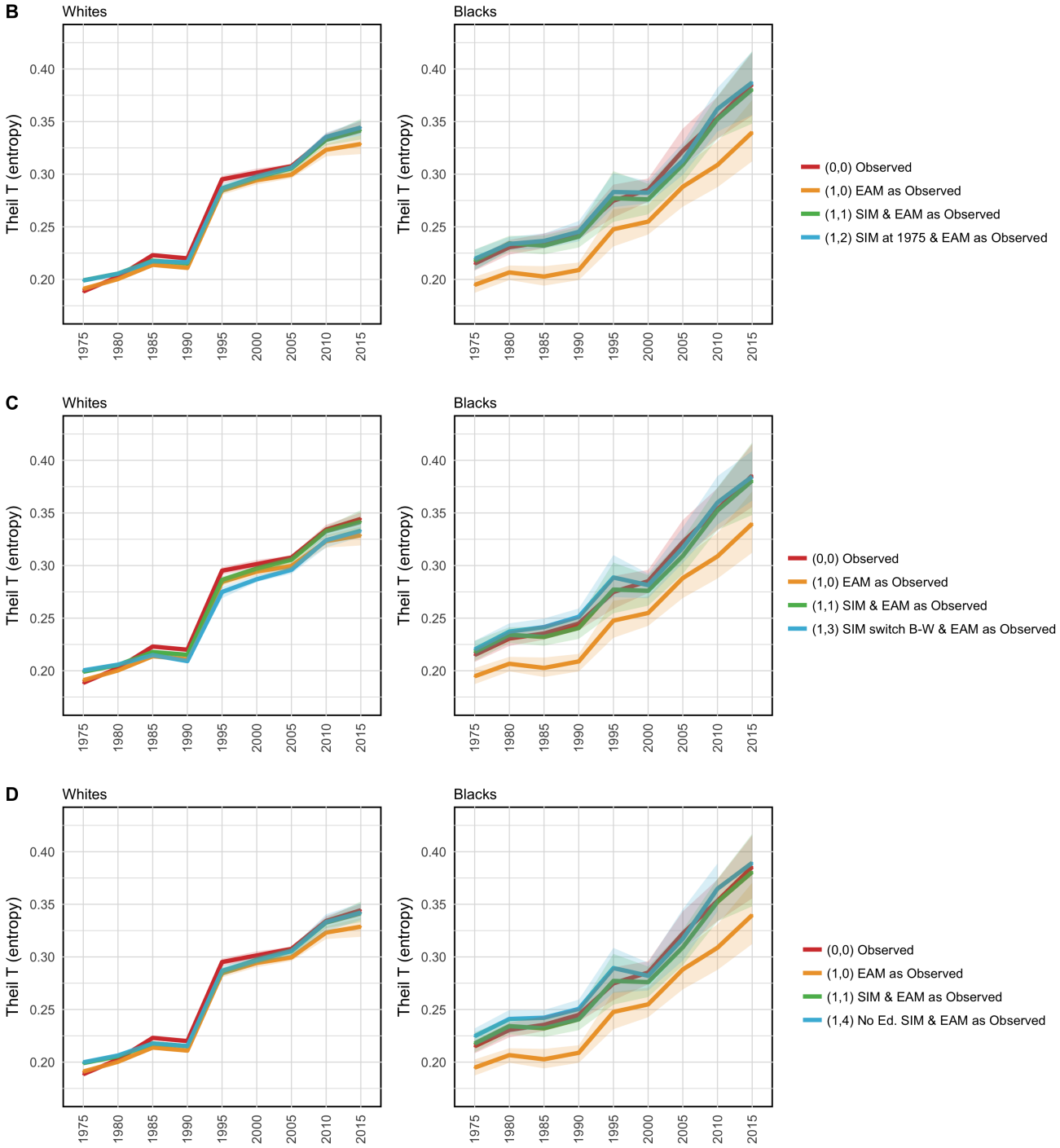

Figure 6: Income inequality under counterfactual scenarios

8 Discussion

The article compares patterns and trends of assortative mating of white and black couples in the US and assesses how these affect the distribution of income within and between the two populations. The article addresses three pressing issues in the literature on social stratification: first, it revisits the null findings or limited effects of increasing educational assortative mating on family income inequality. Second, it studies the connection between assortative mating and inequality among African Americans as a possible form of "perverse openness". In particular, it asks whether structural conditions which lead to educational hypogamy ("marrying down") among highly educated black women have the unintended effect of reducing income inequality among black families. Third, it assesses the impact of selection into marriage on inequality, as this is an endogenous factor to the process of assortative mating. To test these hypotheses, I combine traditional log-linear models for the study of assortative mating with micro-simulation. While the former offers a rich characterization of potentially complex patterns of assortative mating, the latter enables me to take advantage of such complexity to counterfactually study how different facets of the phenomenon produce the observed distribution of family income among racial groups.

Regarding the first issue, findings confirm previous research reporting minor effects of increasing educational assortative mating on family income inequality. In particular, these results suggest that educational expansion has had a small negative impact on income inequality, as scenarios that simulate no expansion display slightly higher levels of income dispersion. Similarly, if mating behavior were not assortative income disparities among families would be somewhat lower than observed. In contrast, changes in assortative behavior appear unrelated to inequality trends. This result likely reflects the fact that, contrary to previous studies, assortative behavior was found to be remarkably stable over time.

Despite these modest effects, the different facets of educational assortative mating demonstrate to be unable to explain the rising levels of income inequality. Since this finding holds using sophisticated analytic tools, it is unlikely that null effects reported by previous research were due to the use of overly simplistic methods. As for the possibility that changes in assortative mating have not been large enough to affect inequality, this study finds no changes in assortative behavior, but it does find a substantial increase in homogamy due to educational expansion. Taken together, these results make the more plausible the third possible explanation for the null effect of increasing assortative mating. That is that, despite the advancement of women in education and the labor market, the linkage between wives' educational attainment and their labor market outcomes is still not strong enough to exert a sizable effect on the distribution of resources within and between families (a future version of the paper will explore this point further).

Regarding the black-white comparison, this study finds that, although income inequality is higher among black families compared to whites, the general results outlined above are highly comparable across these two populations. Such similarity is not entirely surprising given that both groups display similar patterns of assortative behavior. And even though there are important black-white differences in the extent of assortative mating due to the marginal distributions of education, these do not explain differences in inequality across the two populations.

Finally, we find evidence of substantial and accruing education-based selection into marriage, especially among blacks. Because this biases the poll of marriageable individuals upwards in term of educational attainment, in the absence of such selection pattern one might expect less homogamy among the highly educated and less inequality across families. Quite the contrary, the results suggest that, for black families, more random allocation of individuals into marriage would lead to more, not less inequality. As for whites, the findings indicate that selection into marriage has no or a minor impact on income inequality. This result is consistent with the comparatively mild retreat from marriage among all white educational groups and the limited impact of educational sorting on income inequality.

References

Becker, G. (1981). A Treatise on the Family.

Bennett, N. G., Bloom, D. E., and Craig, P. H. (1989). The Divergence of Black and White Marriage Patterns. *American Journal of Sociology*, 95(3):692–722.

Birkelund, G. E. and Heldal, J. (2003). Who marries whom? Educational homogamy in Norway. *Demographic Research*, 8(1):1–30.

Blackwell, D. L. and Lichter, D. T. (2004). Homogamy among dating, cohabiting, and married couples. *Sociological Quarterly*, 45(4):719–737.

Blau, P. M. (1977). A Macrosociological Theory of Social Structure. *American Journal of Sociology*, 83(1):26–54.

Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure*. Wiley, New York,.

Blossfeld, H.-P. (2009). Educational Assortative Marriage in Comparative Perspective. *Annual Review of Sociology*, 35(1):513–530.

Breen, R. and Andersen, S. (2012a). Educational Assortative Mating and Income Inequality in Denmark. *Demography*, 49(3):867–887.

Breen, R. and Andersen, S. H. (2012b). Educational Assortative Mating and Income Inequality in Denmark. *Demography*, 49(3):867–887.

Breen, R. and Salazar, L. (2010). Has increased womens educational attainment led to greater earnings inequality in the United Kingdom a multivariate decomposition analysis. *European Sociological Review*, 26(2):143–157.

Breen, R. and Salazar, L. (2011). Educational Assortative Mating and Earnings Inequality in the United States. *American Journal of Sociology*, 117(3):808–843.

Cancian, M. and Reed, D. (1999). The impact of wives' earnings on income inequality: issues and estimates. *Demography*, 36(2):173–184.

Cherlin, A. J. (2004). The deinstitutionalization of American marriage. *Journal of Marriage and Family*, 66(4):848–861.

Clogg, C. C. (1982). Some Models for the Analysis of Association in Multiway Cross-Classifications Having Ordered Categories. *Journal of the American Statistical Association*, 77(380):803.

Cohen, P. N. and Pepin, J. R. (2018). Unequal marriage markets: Sex ratios and first marriage among Black and White women. *SocArXiv*.

Dufour, J.-M., Flachaire, E., and Khalaf, L. (2015). Permutation tests for comparing inequality measures *.

Ellwood, D. T. and Jencks, C. (2004). The Spread of Single-Parent Families in the United States since 1960. *SSRN Electronic Journal*, 2004(February).

Esping-Andersen, G. (2007). Sociological Explanations of Changing Income Distributions. *American Behavioral Scientist*, 50(5):639–658.

Esteve, A. and Cortina, C. (2006). Changes in educational assortative among in contemporary Spain. *Demographic Research*, 14:405–428.

Fernandez, R., Guner, N., and Knowles, J. (2005). Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality. *The Quarterly Journal of Economics*, 120(1):273–344.

Gihleb, R. and Lang, K. (2016). Educational Homogamy and Assortative Mating Have Not Increased. *NBER Working Paper*.

Goldin, C. (2006). The Quiet Revolution That Transformed Women's Employment, Education, and Family. *American Economic Review*, 96(2):1–21.

Goodman, L. A. (1972). Some multiplicative models for the analysis of cross-classified data. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). Marry your like: Assortative mating and income inequality. In *American Economic Review*, volume 104, pages 348–353.

Grotti, R. and Scherer, S. (2016). Does gender equality increase economic inequality? Evidence from five countries. *Research in Social Stratification and Mobility*, 45:13–26.

Harknett, K. and Kuperberg, A. (2011). Education, labor markets and the retreat from marriage. *Social Forces*, 90(1):41–63.

Kalmijn, M. (1991). Status Homogamy in the United States. *American Journal of Sociology*, 97(2):496–523.

Katrák, T., Kreidl, M., and Fónadová, L. (2006). Trends in educational assortative mating in central Europe: The Czech Republic, Slovakia, Poland, and Hungary, 1988-2000. *European Sociological Review*, 22(3):309–322.

Lash, S. (1987). *Love as Passion, the Codification of Intimacy - Luhmann,N*, volume 21. Harvard University Press, Cambridge, Massachusetts.

Lichter, D. T., LeClere, F. B., and McLaughlin, D. K. (1991). Local Marriage Markets and the Marital Behavior of Black and White Women. *American Journal of Sociology*, 96(4):843.

Lichter, D. T., Mclaughlin, D. K., Kephart, G., and Landry, D. J. (1992). Race and the Retreat From Marriage: A Shortage of Marriageable Men? *American Sociological Review*, 57(6):781–799.

Lichter, D. T., McLaughlin, D. K., and Ribar, D. C. (2002). Economic restructuring and the retreat from marriage. *Social Science Research*, 31(2):230–256.

Lundberg, S. and Pollak, R. (2013). Cohabitation and the Uneven Retreat from Marriage in the U.S., 1950-2010. *NBER Working Papers*.

Maralani, V. (2013). The Demography of Social Mobility: Black-White Differences in the Process of Educational Reproduction. *American Journal of Sociology*, 118(6):1509–1558.

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

Mare, R. D. (2008). Educational Assortative Mating in Two Generations.

Mare, R. D. and Schwartz, C. R. (2006). Educational Assortative Mating and the Family Background of the Next Generation. *Sociological Theory and Methods*, 21(2):253–278.

McDaniel, A., DiPrete, T. A., Buchmann, C., and Shwed, U. (2011). The Black Gender Gap in Educational Attainment: Historical Trends and Racial Comparisons. *Demography*, 48(3):889–914.

McLanahan, S. (2004). Diverging Destinies: How Children Are Faring Under the Second Demographic Transition. *Demography*, 41(4):607–627.

Oppenheimer, V. K. (1988). A Theory of Marriage Timing. *American Journal of Sociology*, 94(3):563–591.

Oppenheimer, V. K. (1997). Women's Employment and the Gain to Marriage: The Specialization and Trading Model. *Annual Review of Sociology*, 23(1):431–453.

Powers, D. A. and Xie, Y. (2000). *Statistical methods for categorical data analysis*. Academic Press, San Diego, California.

Raley, R. K. (1996). A Shortage of Marriageable Men? A Note on the Role of Cohabitation in Black-White Differences in Marriage Rates. *American Sociological Review*, 61(6):973–983.

Reardon, S. F. and Bischoff, K. (2011). Income Inequality and Income Segregation. *American Journal of Sociology*, 116(4):1092–1153.

Rosenfeld, M. J. (2008a). Racial, Educational and Religious Endogamy in the United States: A Comparative Historical Perspective. *Social Forces*, 87(1):1–31.

Rosenfeld, M. J. (2008b). Racial, Educational and Religious Endogamy in the United States: A Comparative Historical Perspective. *Social Forces*, 87(1):1–31.

Schmidt, C. M. and Winter, B. S. (2008). The Dynamics of Assortative Mating in Germany 1 1 Introduction. Technical Report 0346, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen.

Schoen, R. and Cheng, Y. H. A. (2006). Partner choice and the differential retreat from marriage. *Journal of Marriage and Family*, 68(1):1–10.

Schwartz, C. (2010). Earnings Inequality and the Changing Association between Spouses'Earnings. *American Journal of Sociology*, 115(5):1524–1557.

Schwartz, C. and Mare, R. D. (2005a). Trends in Educational Assortative Marriage from 1940 to 2003. *Demography*, 42(4):621–646.

Schwartz, C. R. (2013). Trends and Variation in Assortative Mating: Causes and Consequences. *Annual Review of Sociology*, 39(1):451–470.

Schwartz, C. R. and Mare, R. D. (2005b). Trends in Educational Assortative Marriage from 1940 to 2003. *Demography*, 42(4):621–646.

Spanier, G. and PC, G. (1980). Mate Selection Differentials between Whites and Blacks in the United States. *Social Forces*, 58(3):707.

Sweeney, M. M. (2002a). Two Decades of Family Change: The Shifting Economic Foundations of Marriage. *American Sociological Review*, 67(1):132–147.

Sweeney, M. M. (2002b). Two Decades of Family Change: The Shifting Economic Foundations of Marriage. *American Sociological Review*, 67(1):132–147.

Teachman, J. D., Tedrow, L. M., Crowder, K. D., Marriage, J., and Nov, N. (2007). The Changing Demography of America ' s Families. *Journal of Marriage and the Family*, 62(4):1234–1246.

Torche, F. (2010). Educational assortative mating and economic inequality: a comparative analysis of three Latin American countries. *Demography*, 47(2):481–502.

Treiman, D. J. (1970). Industrialization and Social Stratification. *Sociological Inquiry*, 40(2):207–234.

Weeden, K. A. and Grusky, D. B. (2005). The Case for a New Class Map. *American Journal of Sociology*, 111(1):141–212.

Western, B., Bloome, D., and Percheski, C. (2008). Inequality among American Families with Children, 1975 to 2005. *American Sociological Review*, 73(6):903–920.

Wilson, W. J. (1990). *The truly disadvantaged: the inner city, the underclass, and public policy*. University of Chicago Press.

Wood, R. G. (1995). Marriage Rates and Marriageable Men: A Test of the Wilson Hypothesis. *Journal of Human Resources*, 30(1):163–193.

Xie, Y., Cheng, S., and Zhou, X. (2015). Assortative mating without assortative preference. *Proceedings of the National Academy of Sciences*, 112(19):5974–5978.

Xie, Y., Raymo, J. M., Goyette, K., and Thornton, A. (2003). Economic potential and entry into marriage and cohabitation. *Demography*, 40(2):351–367.

A Appendix

A.1 Homogamy and crossing models

Homogamy models test whether individuals are more likely to marry partners with their *same* level of educational attainment. In its simplest version the "constrained homogamy" model assumes assortative mating patterns are captured by the marginal distribution of husbands' and wives' educational attainment, plus one terms that indicates educational homogamy across the main diagonal of the contingency table for each year (Powers and Xie, 2000). Formally,

$$\log F_{ijk} = \lambda_0 + \lambda_i^H + \lambda_j^W + \lambda_k^Y + \lambda_{ij}^{HW} + \lambda_{ik}^{HY} + \lambda_{jk}^{WY} + \gamma_{i=j,k}^{DY}, \tag{4}$$

where D=1 if husband's education is equal to his wife's education and zero otherwise.

Related versions of this model allow for more complex pattern of educational homogamy. For example, a common variant permits the strength of homogamy to differ across levels of educational attainment. Other models incorporate parameter(s) that capture (a)symmetric patterns of intermarriage along the minor diagonals of the contingency table, where partners differ by only one level of educational attainment.

By contrast, *crossing models* represent the association between spouses' as a series of barriers to inter-marriage, where different categories of education would entail varying degrees of difficulty for crossing (Powers and Xie, 2000; Schwartz and Mare, 2005a). Formally,

$$\log F_{ijk} = \lambda_0 + \lambda_i^H + \lambda_j^W + \lambda_k^Y + \lambda_{ij}^{HW} + \lambda_{ik}^{HY} + \lambda_{jk}^{WY} + \gamma_{ijk}^{HWY} \tag{5}$$

where

$$\gamma_{ijk}^{HWY} = \begin{cases} \sum_{q=j}^{i-1} \gamma_{qk} & \text{if } i > j \\ \sum_{q=i}^{j-1} \gamma_{qk} & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Here the γ_{qk} parameter represents the variation in the difficulty of crossing educational barrier q in year k relative to a baseline year. The crossing parameters capture the log odds of marriage between individuals in adjacent schooling categories relative to the log odds of homogamy, net of the marginal distributions of spouses' education (Schwartz and Mare, 2005a). Thus, log-odds of intermarriage for partners that cross more than one barrier are calculated by adding the parameters of each barrier crossed. As with homogamy models, crossing models permit relaxing some of its assumptions.

A.2 Cross-validation and K-fold cross-validation

Generally speaking, cross-validation is a technique for model validation. It provides a measure of the average error that results from using a model to predict the response on a new observation (also referred to as "test error"). For this purpose cross-validation techniques adopt a "validation set approach", which consists of the following basic steps:

i. Given a set of n observations, randomly split such set into two subsets, a "training set" and a "testing set".

ii. Fit the statistical model on the training set.

iii. Use the model fitted in the previous step to predict the outcome variable for the observations in the testing set.

iv. Compare the values of predicted and observed responses in the testing set using some error metric (typically MSE for quantitative response variables). This measure is called cross-validation error and provides an estimate of the "test error".

A popular type of cross-validation is the K-fold approach, which involves randomly dividing the set of n observations into k groups (folds) of roughly equal size. Then, the i fold is treated as a testing test, and the model is fitted on the remaining $k-1$ folds. This procedure is repeated k times using, each time, a different fold as testing set. At each repetition the error metric L_i is computed on the observations in the i fold. This process produces k estimates of the test error, $\{L_1, L_2, \ldots, L_k\}$. Finally, the k-fold cross-validation error is computed by averaging these values:

$$CV = \frac{1}{k}\sum_{i=1}^{k} L_i \tag{7}$$

In our application we use Poisson Deviance as a measure of test error, as it constitutes a proper loss function for Poisson distributed outcomes. The Poisson Deviance is defined as:

$$D_{\mathcal{M}} = 2 * \sum_{i=1}^{n} \{Y_i log(Y_i/\hat{Y}_i^{\mathcal{M}}) - (Y_i - \hat{Y}_i^{\mathcal{M}})\} \tag{8}$$

where Y_i is the observed outcome and $\hat{Y}_i^{\mathcal{M}}$ is the prediction under model $\mathcal{M}$.