Marriage market polarization in the time of college educational expansion

Iñaki Permanyer: Corresponding Author (e-mail: ipermanyer@ced.uab.es). Centre d’Estudis Demogràfics, Campus de la Universitat Autònoma de Barcelona, Edifici E-2, 08193 Bellaterra, Spain.

Albert Esteve: Centre d’Estudis Demogràfics, Campus de la Universitat Autònoma de Barcelona, Edifici E-2, 08193 Bellaterra, Spain.

Joan García: Minnesota Population Center, University of Minnesota 225 19th Ave S # 50, United States.

Abstract:
Are societies becoming more polarized between the college and the non-college educated? Is the force of homogamy generating more egalitarian unions at the cost of more polarized societies? In this paper we examine patterns of assortative mating to investigate the extent to which the force of homogamy (propensity to marry within the same educational group), the expansion of college education and the gender gap in education are contributing to the polarization in the marriage market between college and non-college educated populations worldwide. To this end, we have assembled census and survey microdata from 120 countries and 408 samples from 1960 to 2011, in which 98% of the world’s population is represented. We have developed a simple – yet effective – decomposition model that neatly assess the impact of these three forces on couples’ education distribution and the corresponding polarization levels. Results show that the force of homogamy and the gender gap in education have played a very limited role in determining the levels of marriage market polarization between college and non-college populations. Such levels, which are increasing worldwide, are being rather mechanically driven by the share of the population with college education, which continues to expand in most societies. The feared scenarios predicting the shrinkage and gradual disappearance of mixed couples in favor of compartmentalized partitions between college educated couples and below-college educated ones are not occurring.
1. Introduction

In the last few years, there has been an upsurge of interest among social scientists regarding the implications that increasing levels of educational homogamy may have in terms of mounting social distance between social strata. The tendency of individuals to look for and marry partners with similar characteristics (i.e. educational homogamy) might contribute to generate increasingly unequal, impoverished and polarized societies between those who are multiply advantaged and those who are multiply deprived. In this paper we investigate the extent to which marriage markets are polarized between college and non-college educated couples and how (i) the expansion of college education, (ii) the force of homogamy, and (iii) the gender gap in education are contributing to such levels of polarization. For that purpose, we have developed a simple – yet effective – decomposition model that neatly assess the contributions of these three forces to patterns of assortative mating and to the corresponding polarization levels.

To this end, we have assembled a database with more than 400 samples of 120 countries all over the world since the 1960s to the present day. Data mostly came from the IPUMS international census microdata samples database with a complementary use of different household surveys. Results show mounting levels of marriage market polarization between college and non-college educated populations, which are mostly and mechanically driven by the expansion of college education. By contrast, the force of homogamy and the gender gap in education have played a very limited role when determining the changes in polarization levels that have been observed worldwide. The feared scenarios predicting the shrinkage and gradual disappearance of mixed couples in favor of compartmentalized partitions between college educated couples and below-college educated ones are not occurring.

2. Background

As the world gradually embraces the tenets of an increasingly globalized and competitive knowledge-based economy, college education will become the most salient educational boundary of the 21st century, as it was literacy during the 19th and a significant part of the 20th centuries. The importance of being college educated is noticeable in many dimensions of our lives. College educated populations have systematically higher levels of employment and better paid jobs than the non-college ones (Atkinson et al 2011). The academic outcomes of children from college educated parents tend to be higher than those of children from parents with lower levels of education. And education is among the most important stratification variables of demographic behavior to the extent that Lutz and colleagues claim, education should be routinely added to all sorts of population analyses together with age and sex (see Lutz, Butz and KC 2014). These analyses include the role of education in contemporary marriage markets.
Within this context, researchers are increasingly interested in studying how assortative mating patterns (i.e. who marries whom) impact on the distribution of welfare across populations in contemporary societies. Given the significant and positive implications that educational attainment has on individuals and their families, this impact can be particularly strong when at least one of the members of the couple has college education. Educational homogamy may have relevant implications in terms of growing social distance between social strata (Schwartz 2013, Schwartz and Mare 2005, Esping-Andersen 2009). Positive assortative mating may contribute to generate increasingly unequal, impoverished and polarized societies between those who are multiply advantaged and those who are multiply deprived – a matter of concern that has echoed in popular media as well (e.g.: Paul 2006).

Such concerns are based on the fact that the force of homogamy (i.e. the tendency to marry within the same educational group) has strengthened over the last years. If not the most important, educational attainment has become one of the main structuring dimensions of contemporary marriage markets. Education influences the age at which individuals enter the marriage market and shapes their context of opportunities and expectations towards partnership formation (Kalmijn 1998, Mare 1991, Blossfeld 2009, Blossfeld and Timm 2003). The allocation of spouses/partners according to their levels of educational attainment is far from a random process. Men and women tend to look for partners with the same educational attainment, so the number of educationally homogamous couples is always higher than the expected number under random assortative mating. By educational groups, the highest levels of educational homogamy are often found among the college educated, a trend that shows no signs of weakening over time (Smits, Ultee and Lammers 2000, Schwartz and Mare 2005, Schwartz 2013).

The rise of the force of homogamy has gone hand-in-hand with the dramatic expansion of education. It is widely acknowledged that the Western model of mass education has been diffusing around the world for many decades (Meyer et al 1977). Such expansion includes rising literacy rates (Crafts 2002), increases in schooling enrolment rates and in completed years of primary, secondary and college education (Benavot and Riddle 1988, Benavot et al. 1991; Meyer, Ramirez, and Soysal 1992; Ramirez and Meyer 1980; Barro and Lee 2000; Cohen and Soto 2007; Morrisson and Murtin 2009). The gains in virtually all education indicators have benefited all major regions of the world without exception. Regarding college education, by year 1970 6.4% of the world’s 25-29 population had obtained a college degree. Three decades later, this proportion had increased to 13% and the expected figure for 2050 is 29.4% (KC et al. 2010). Interestingly, the impressive spread of college education also increases, ceteris paribus, the opportunities for non-college educated individuals to find a college-educated partner and form a ‘mixed’ (i.e. heterogamous) couple. In this way, the process of education expansion benefits not only those who receive college education but also their potentially less educated partners[[[Note#1]]].
An important feature of the expansion of college education that may have an impact of assortative mating is the gender gap in education. The education expansion that has swept the world during the last century has not been gender neutral (Dorius and Firebaugh 2010; Grant and Behrman 2010). While initially favoring males, we are currently witnessing the opposite pattern: the gender gap in education is not only closing but in many cases even reversing in favor of women (see Figure 1) (Esteve, Garcia and Permanyer 2012). Expert worldwide education projections suggest that these patterns will not only continue but also accentuate over the next decades (KC et al. 2010; Lutz and KC, 2011). In 1970, 63.6% of the college-educated were men. In year 2000, that percentage reduced to 52.6%. By 2050, the overall percent of men among the college-educated is expected to reduce even further to 44%, with many high-income countries experiencing much larger relative reductions (KC et al 2010). Therefore, the ‘excess’ of college educated women can potentially reduce the share of college educated couples while increasing the number of mixed couples [[[Note#2]]]. As a matter of fact, the reversal of the gender gap in education is responsible for the increase in the number of couples in which the woman has more education than the man (Esteve, Garcia and Permanyer 2012).

In this context, this paper examines trends and patterns in marriage market polarization between college and non-college populations and investigates the extent to which the purported polarizing and disequalizing effects in the marriage market are driven by the expansion of education, the force of homogamy or the gender gap in educational attainment. Since education expansion increases both the probability that college educated individuals partner among themselves and the probability that an individual without college education partners a college educated one, the implications of college education expansion regarding polarization are not a priori clear. Whether or not such education expansion will translate into higher or lower polarization levels will be mediated by the force of homogamy and the gender gap in education. Large differences in college education rates between women and men should contribute to increase the number of couples between college and non-college populations and reduce the number of couples among college educated populations, therefore lowering polarization levels.

Will we witness a story of divergent destinies where societies break apart in two opposing and gradually distant poles? Or will the unprecedented global changes in the access to education in favor of women tilt the balance in the opposite direction? In order to investigate these issues, we introduce a straightforward and novel methodology building on simple counterfactual and benchmarking models that compare ‘real’ (i.e.: observed) education distributions with alternative hypothetical distributions that would be observed if other structural conditions or behavioral traits had prevailed. This way, we are able to neatly decompose the contribution of the force of homogamy, the expansion of college education and the gender gap in education to the changes in shares of college educated, mixed and non-college educated couples and the educational
polarization levels associated to them. Before presenting our methodology and main results, we first present the data.

3. Data

Our analysis is based on a vast collection of census and survey microdata samples from 120 countries, spanning from 1960 to 2011. We have gathered data from 408 samples of microdata: 187 census microdata samples were obtained from the Integrated Public Use of Microdata Series international project (Minnesota Population Center, 2014); 131 from Demographic Health Surveys; 45 from the European Labor Force Surveys; 27 from the European Statistics on Income and Living Conditions; and 18 from the Generations and Gender Survey. The IPUMS and DHS datasets are available online and for free to researchers. By decade, we have 4 samples from the 1960s, 29 from the 1970s, 41 from the 1980s, 98 from the 1990s, 175 from the 2000s, and 61 from the 2010s. Our data represents over 98% of the world’s population.

The final dataset only includes samples in which the education of the spouses can be identified. The analysis is restricted to the population in heterosexual unions (marriage and cohabitation) in which women were 25-34 years old. In this way, we avoid the possible biases and distortions that arise when population overlaps across different points in time. In addition, the potential distortion effects of union dissolution, educational upgrades and remarriage are minimized (Schwartz and Mare 2012). The final dataset includes more than 14 million individual records.

Educational attainment was dichotomized into non-college and college education. Educational systems vary widely across the globe and their harmonization is always problematic (Esteve and Sobek 2003). However, there are some educational levels that are fairly standard across societies and college education is one of them. Most cross-national comparability problems regarding educational systems are related to the transition from primary to secondary levels and to the diversity of study tracks that exist at the secondary level. By focusing on the non-college vs. college divide, we avoid most of the comparability challenges. Virtually all censuses and surveys employed in our research identify college / tertiary education without ambiguities.

4. Analytical strategy

We introduce some basic definitions and notation that will be used throughout the paper. Since we are interested in assessing and decomposing polarization levels in the marriage market, from now onwards we restrict our attention to the population living in union. In this article we are basically interested in the top of the education distribution, so we only consider two educational groups: those without college education and the college educated. This generates a $2 \times 2$ contingency table with 4 possible combinations depending on the education level of the partners [[Note#3]]. The first combination corresponds to the couples in which none of the members has college education; their
share among the population in union will be denoted by ‘a’. Analogously, ‘d’ represents the share in which both partners are college educated, ‘b’ represents the share of couples in which men have no college education and women have and ‘c’ represents the opposite combination. Technically speaking, the couples counted in ‘a’ and ‘d’ are homogamous, whereas the couples counted in ‘b’ and ‘c’ are heterogamous. More specifically, the couples in ‘b’ are hypogamous, and the couples in ‘c’ are hypergamous [[[Note#4]]]. For the sake of simplicity, this distribution will be referred to as educational assortative mating, which from now onwards will be shortly denoted as \((a,b,c,d)\). Since \(a,b,c,d\) are shares, their sum adds up to 1.

4.1. Three basic determinants of assortative mating

Of primary interest for the purposes of this paper are the shares of couples where both members are college and non-college educated (i.e.: the groups that ‘threaten’ to race ahead or being left behind of the rest respectively – ‘d’ and ‘a’), the share of mixed couples (i.e.: the group that bridges the previous two – ‘b+c’

and the education polarization levels associated to the distribution \((a,b,c,d)\) (which are defined below in section 4.2). As is clear, these shares are directly influenced by several factors, among which we highlight the following ones:

(i) The expansion of college education, measured as the share of college educated population among the population in union (denoted by \(E\)). Other factors kept constant, higher values of \(E\) increase the share of college educated couples and decrease the share of non-college educated ones. Analogously, increases in \(E\) should lead to increases (respectively, decreases) in the share of mixed couples whenever \(E\) is ‘small’ (respectively, ‘big’). Formally, we have that

\[
E = \frac{b + c + 2d}{2} \quad [1]
\]

In addition, the share of college educated women among women in union is

\[
E_w = b + d \quad [2]
\]

and the share of college educated men among men in union is

\[
E_m = c + d \quad [3]
\]

Clearly, one has that

\[
E = \frac{E_w + E_m}{2} \quad [4]
\]

(ii) The force of homogamy (that is: the extent to which couples are made of individuals with similar characteristics), which will be denoted by \(H\). When
individuals mate within the same educational group, homogamy levels are, by construction, higher. *Ceteris paribus*, higher levels of homogamy tend to increase the shares of couples where both members are either college or non-college educated while decreasing the share of mixed couples. Formally, the force of homogamy can be measured with the following indicator [[[Note#5]]]

\[ H = ad - bc \quad [5] \]

(iii) The *gender gap in college education* among the population in union (denoted by \( G \)). When there are big imbalances in the education distribution of women and men, the number of homogamous couples that can be formed diminishes while the potential number of mixed couples increases. Formally, the gender gap in education is defined as:

\[ G = b - c = E_w - E_m \quad [6] \]

As shown in equations [1], [5] and [6], each education distribution \((a,b,c,d)\) has associated its corresponding college education \((E)\), force of homogamy \((H)\) and gender gap \((G)\) levels. Interestingly, the opposite is also true, that is: when these three factors are fixed, there is one *and only one* education distribution \((a,b,c,d)\) satisfying the aforementioned equations. This is clearly shown in the following equations:

\[
\begin{align*}
    a &= \varphi_a (E, H, G) = (1 - E)^2 + H - \left( \frac{G}{2} \right)^2 \\
    b &= \varphi_b (E, H, G) = E(1 - E) - H + \frac{G}{2} \left( \frac{G}{2} + 1 \right) \\
    c &= \varphi_c (E, H, G) = E(1 - E) - H + \frac{G}{2} \left( \frac{G}{2} - 1 \right) \\
    d &= \varphi_d (E, H, G) = E^2 + H - \left( \frac{G}{2} \right)^2 
\end{align*}
\]  

[7]

These fundamental identities are one of the key contributions of the paper: they show in a clear cut way how the three factors considered here (share of college education ‘\(E’\), force of homogamy ‘\(H’\) and gender gap in college education ‘\(G’\)) are related to the educational distribution of couples and they allow different sorts of counterfactual analysis (see section 4.4). The derivation of these identities is a bit involved, so the details are explained in the appendix.

4.2. Measurement of polarization

As discussed elsewhere, polarization is defined as the grouping of the population into significantly-sized clusters such that each cluster has members with similar attributes and different clusters have members with dissimilar ones (e.g.: Esteban and Ray 1994).
In our context, the groupings of the population in union are based on the couples’ education distribution \((a,b,c,d)\). Roughly speaking, an index of polarization aims to assess how far a given distribution is from a hypothetical scenario in which the population is split in two equally sized and antagonistic groups (i.e.: those with college education vs. those without college education). Drawing from the recent literature on social polarization, in this paper we will work with the following polarization index

\[
P^\alpha(a,b,c,d) = 1 - 2^{\alpha-1} \left( \frac{1}{2} - a \right)^\alpha + \frac{1}{2} \left( a + b + c \right)^\alpha \tag{8}
\]

where \(\alpha\) is a non-negative parameter measuring the importance given to the median category \([\text{Note#6}]\) (in the main empirical applications of the paper we use \(\alpha=2\), but other values have been investigated as robustness checks in section 5.4). This index is an \textit{ad hoc} adaptation of the ordinal polarization index suggested by Apouey (2007: 885) for the case in which one deals with 3 categories \([\text{Note#7}]\). \(P^\alpha\) is a standard index of polarization \([\text{Note#8}]\) that essentially measures the distance between a given education distribution \((a,b,c,d)\) and the bipolar case \((1/2,0,0,1/2)\) where the population is split in two equal-sized groups concentrated at the opposite extremes of the education distribution.

\subsection*{4.3. Benchmarking exercises}

One simple way of assessing the impact of the force of homogamy on assortative mating is to investigate the extent to which couples’ education distribution and polarization levels would have been different under alternative homogamy levels. As shown in the empirical section of the paper, the observed shares of homogamous couples is always higher than the hypothetical homogamous couples’ shares that would be observed in case couples were formed purely at random (i.e.: in the absence of homogamy and regardless of the educational attainment of their spouses). At the same time, it can be shown that prevailing homogamy levels do not maximize the number of homogamous couples that can \textit{a priori} be formed. Stated otherwise, while individuals have a tendency to partner other individuals with the same educational attainment, such tendency is not universal and the share of mixed couples is higher than the one that would be observed in a maximal homogamy scenario in which individuals absolutely prioritized partners with the same educational attainment. In this context, one might be interested in benchmarking the effect of homogamy on assortative mating and polarization levels by framing the observed values of the latter within a range of hypothetical values that would be observed if alternative homogamy levels had prevailed. For that purpose, we will derive the education distribution and the corresponding polarization levels that would be observed in the two extreme and hypothetical scenarios of ‘absence’ and ‘maximal’ educational homogamy. In the appendix we show how to derive these two hypothetical education distributions, which are denoted by \((a_0, b_0, c_0, d_0)\) and \((a_{\text{max}}, b_{\text{max}}, c_{\text{max}}, d_{\text{max}})\) respectively. The observed
shares in education distributions \((a,b,c,d)\) and the corresponding polarization levels \(P^\alpha(a,b,c,d)\) are bounded from below and from above by the aforementioned hypothetical education distributions with ‘absence’ and ‘maximal’ educational homogamy, as can be seen in the following inequalities (see the Appendix for details):

\[
\begin{align*}
\begin{cases}
a_0 \leq a \leq a_{\text{max}} \\
(b+c)_{\text{max}} \leq b + c \leq (b+c)_0 \\
d_0 \leq d \leq d_{\text{max}}
\end{cases}
\end{align*}
\]

\[P^\alpha_0 \triangleq P^\alpha(a_0,b_0,c_0,d_0) \leq P^\alpha(a,b,c,d) \leq P^\alpha(a_{\text{max}},b_{\text{max}},c_{\text{max}},d_{\text{max}}) \triangleq P^\alpha_{\text{max}}\]

As expected, the observed shares of homogamous couples (‘\(a\)’ and ‘\(d\)’) and the levels of polarization are higher than the ones that would be observed in the absence of homogamy but smaller than the ones that would be observed in case of maximal homogamy. Alternatively, the observed share of mixed couples (‘\(b+c\)’) is smaller than the one that would be observed in the absence of homogamy but higher than the one that would be observed in case of maximal homogamy. In order to benchmark these variables between the corresponding bounds, we normalize them to a \([0,1]\)-scale via the standard transformations

\[
\begin{align*}
\begin{cases}
a^* = \frac{a-a_0}{a_{\text{max}}-a_0} \\
(b+c)^* = \frac{(b+c)_0 - (b+c)}{(b+c)_0 - (b+c)_{\text{max}}} \\
d^* = \frac{d-d_0}{d_{\text{max}}-d_0} \\
P^* = \frac{P^\alpha - P^\alpha_0}{P^\alpha_{\text{max}} - P^\alpha_0}
\end{cases}
\end{align*}
\]

Since \(P^\alpha_0\) measures the polarization levels that would be observed in a hypothetical scenario where the force of homogamy were zero (i.e.: something that might be referred to as ‘baseline polarization’), \(P^\alpha - P^\alpha_0\) measures \textit{in absolute terms} the ‘amount of polarization’ that is attributable to the force of homogamy. In order to assess the magnitude of the previous quantity with respect to the maximal amount of polarization that could be attributable to assortative mating patterns, we further divide it by \(P^\alpha_{\text{max}} - P^\alpha_0\), therefore obtaining \(P^*\): a measure \textit{in relative terms} of the amount of polarization that can be attributable to the force of homogamy. Analogous reasoning applies to \(a^*\), \((b+c)^*\) and \(d^*\).

Interestingly, it turns out that the force of homogamy has exactly the same relative impact on the different shares of the education distribution (i.e.: the first three equations in [10] – \(a^*\), \((b+c)^*\) and \(d^*\) – are indeed the same, so they will simply be denoted by \(m^*\); see the Appendix). In words, \(m^*\) measures the extent to which the observed education
distribution differs from the one that would be observed in the absence of homogamy relative to the maximal possible change that could take place under extreme levels of homogamy. Using such extreme scenarios we can ascertain whether the observed education distribution shares are malleable by the tendency toward homogamy or are coarsely determined by the educational distribution of spouses.

4.4. Counterfactual modelling

So far, we have proposed a simple methodology that compares observed education distributions and the corresponding polarization levels with the ones that would be observed under alternative hypothetical levels of homogamy. While informative and interesting on its own right, this method fails to take into account the dynamics of change over time: among other things, one does not know whether changes in the force of homogamy have been more or less decisive than changes in the gender gap in education or the changes in the share of the population with college education when shaping education distributions and the corresponding polarization levels.

In order to assess the separate impact of the expansion of college education, force of homogamy and gender gap in college education on the educational distribution of couples and the corresponding polarization levels we have carried out several simulation exercises. Taking advantage of the multiple observations over time that are available for most countries included in our dataset, we ask what would have happened to the shares $a$, $b+c$, $d$ and the corresponding polarization levels $P^d$ if we held constant two of the three quantities that appear in [7] (‘$E$’, ‘$H$’ and ‘$G$’) at their value in an earlier period of time ($t_1$) and allowed the third to take a value observed later in time ($t_2$). In this way, we generate a counterfactual education distribution and the corresponding polarization level for the later period of time ($t_2$), and by comparing this with the real values from earlier periods, we can assess the impact of change on that third quantity on the variable of interest (see equations in [11] and [13]).

Let us denote the education distributions in times $t_1$ and $t_2$ as $(a_1,b_1,c_1,d_1)$ and $(a_2,b_2,c_2,d_2)$ respectively. Analogously, define the college education levels, force of homogamy and gender gap in college education observed in times $t_1$ and $t_2$ as $E_1$, $H_1$, $G_1$ and $E_2$, $H_2$, $G_2$ respectively. Define also $d^E_2 = \varphi_d(E_2,H_1,G_1)$, $d^H_2 = \varphi_d(E_1,H_2,G_1)$ and $d^G_2 = \varphi_d(E_1,H_1,G_2)$ as the counterfactual share of college educated couples that would be observed in $t_2$ if we only changed over time the dimensions of education expansion, homogamy and gender gap in education respectively while keeping the other two factors constant (see equation [7] for the definition of $\varphi_d(.....)$). Analogously, we can define the counterfactual shares of non-college educated couples $a^E_2$, $a^H_2$, $a^G_2$ and of mixed couples $(b+c)^E_2$, $(b+c)^H_2$, $(b+c)^G_2$. In this article, we are interested in the differences between the observed education distributions and the counterfactual ones, that is:
Interestingly, the observed (i.e.: ‘real’) difference between education distributions over time can be written as

\[ a_2 - a_1 = \Delta_E a + \Delta_H a + \Delta_G a \]

\[ (b + c)_2 - (b + c)_1 = \Delta_E (b + c) + \Delta_H (b + c) + \Delta_G (b + c) \]

\[ d_2 - d_1 = \Delta_E d + \Delta_H d + \Delta_G d \]

that is: the changes in education distributions over time can be neatly decomposed as the sum of the changes attributable to the three factors considered in this article: expansion of college education ‘E’, force of homogamy ‘H’ and the gender gap in education ‘G’. In this way, it is easy to quantify which of the three factors has been more decisive in driving the changes in education distributions (to illustrate: the percent contribution of, say, homogamy to the share of college educated couples can be simply calculated as 100·|ΔHd|/(|ΔEd|+|ΔHd|+|ΔGd|)). Following the same logic, we also quantify the polarization changes that would be observed if we changed one of the three factors while leaving the other two constant:

\[ \Delta_E P^a = P^a (a_2^E, b_2^E, c_2^E, d_2^E) - P^a (a_1, b_1, c_1, d_1) \]

\[ \Delta_H P^a = P^a (a_2^H, b_2^H, c_2^H, d_2^H) - P^a (a_1, b_1, c_1, d_1) \]

\[ \Delta_G P^a = P^a (a_2^G, b_2^G, c_2^G, d_2^G) - P^a (a_1, b_1, c_1, d_1) \]

In order to take advantage of the decompositions shown in equation [11], [12] and [13], we need to work with the set of coupled observations that one can construct when more than one observation over time is available for a given country. That is, if for a certain country we have \( n \) observations over time (in times \( t_1, t_2, \ldots, t_n \)), we will then consider \( n - 1 \) of these coupled observations (\( t_1 \) coupled with \( t_2, t_2 \) coupled with \( t_3 \), and so on).

5. Empirical Results

5.1. Descriptive findings

First, we present descriptive findings that are useful to overview the existing relationship between the main variables we are working with. Figure 2 shows the percentage of men and women with college education in the horizontal axis (\( E \)) and the force of homogamy (\( H \)) in the vertical axis. Each data point in the scatterplot represents one country-specific observation. In total, there are 408 observations from 120 countries between 1960 and 2011. The force of homogamy is always positive (even if, a priori, it could also take negative values) and it rises with college education. When college education is less than 5%, the force of homogamy is positive but very close to 0. From
5% to 20% of the college education, the force of homogamy increases in a linear way. Beyond 20%, there is larger variability across countries and the relationship between college education and the force of homogamy flattens. These results are consistent with the findings of Smits, Ultee and Lammers (2000), Scwhartz and Mare (2005) and Schwartz (2013) reporting increasing homogamy among the college educated.

(((FIGURE 2)))

Figure 3 shows the relationship between college education (horizontal axis) and the share of college educated couples among all couples (red triangles), the share of non-college educated couples (yellow squares) and the share of mixed couples (black dots). As can be seen, the share of non-college educated couples decreases linearly with the expansion of college education: it starts around 100% when there are no college educated individuals and it falls near 30% when college education approaches 50%. On the other hand, college education is related in a positive way to both the share of college educated couples and the share of mixed couples. Either the share of college educated couples or the share of mixed couples are around 30% when the percent of college education is at 50%. In most cases, at the same level of college education, the share of mixed couples is slightly higher that the share of college educated couples. The variability across observations increases at higher levels of college education. Despite the increasing force of homogamy that accompanies education expansion, the corresponding share of mixed couples does not show signs of decline.

(((FIGURE 3)))

Figure 4 shows the relationship between college education and the gender gap in college education among the population in union. When levels of college education are low, the gender gap in education systematically favors men. However, as college education increases, the gender gap shrinks and reverses in favor of women. The gender gap in education is favorable to women in virtually all cases where the percentage of college education exceeds 20%. These results for the population in union mirror quite faithfully the overall trends reported in Figure 1 referring to the entire population.

(((FIGURE 4)))

Finally, Figure 5 shows the relationship between college education levels (horizontal axis) and the observed levels of polarization in the marriage market (shown in the vertical axis with rounded dots). The value of 0 in the polarization index indicates absence of polarization and 1 maximum polarization. High levels of college education are associated with high levels of polarization in a well-behaved curvilinear fashion. After a certain level of college education (around 30%), the levels of polarization are not increasing as fast. While the observed levels of polarization might look surprisingly high (i.e. hovering around 0.85 out of 1 when college education approaches 50%), it should be noted that, even if the force of homogamy played no role, the corresponding
polarization levels – indicated by the lower whiskers in Figure 5 – would be high as well (reaching 0.75 when college education approaches 50%). Analogously, the upper whiskers in Figure 5 plot the polarization levels that would be observed in a hypothetical scenario of maximal educational homogamy. In the following section, we benchmark the observed levels of polarization with respect to the ones that would be observed under these alternative assortative mating scenarios.

5.2. Benchmarking

In this section, we investigate how the shares of the education distribution and the corresponding polarization levels would look like if alternative forces of homogamy had prevailed. More specifically, we start comparing the observed shares of the education distribution (a, b+c and d) with the ones that would be observed in the absence of homogamy (a₀, b₀+c₀ and d₀). For that purpose, Figure 6 plots the density functions associated to the values of 100(a−a₀)/a, 100(b+c−(b₀+c₀))/(b+c) and 100(d−d₀)/d for all the samples included in our dataset. These indices measure by how much the observed values in the shares of each type of couples (a, b+c, and d) can be attributed to the force of homogamy. As can be seen, that force has contributed rather modestly in shaping the share of non-college educated couples: on average, 6.5% of the values of those shares can be attributed to the force of homogamy (see the vertical lines in Figure 6). On the other hand, the contribution to the shares of mixed couples and college-educated couples has been substantially high. If couples were formed purely at random, the shares of mixed couples would increase, on average, 77% of their observed values while the shares of college educated couples would decrease, on average, by 78.4%.

Lastly, we compare the observed polarization levels \(P^α(a,b,c,d)\) with the ones that would be observed in the absence of homogamy \(P^α(a₀,b₀,c₀,d₀)\). Figure 6 also plots the percent contribution of the force of homogamy to existing polarization levels \((100(P^α(a,b,c,d)−P^α(a₀,b₀,c₀,d₀))/P^α(a,b,c,d))\). As can be seen, the contribution is relatively low, with an average value of 6.5%. Interestingly, the force of homogamy does not seem to play an influential role when determining observed polarization levels.

The left panel in Figure 7 jointly plots the values of college education (E) against those of \(m^*\). As can be seen, the average values of the latter hover around an average of 0.58. This means that if couples were formed purely at random, the shares of homogamous (resp. heterogamous) couples would have decreased (resp. increased) by an average of 58% when compared to its maximal scope for potential change. These results suggest that, while being a factor that greatly contributes to increase (resp. decrease) the share of homogamous (resp. mixed) couples, the force of homogamy is not at its full strength, as it could – on average – further change that share by an extra 42% (=100% − 58%). In
addition, with higher levels of college education the contribution of homogamy to the relative changes in the shares of the education distribution declines slightly.

We now repeat the same benchmarking exercise with respect to polarization. Despite the relatively low contribution of homogamy to polarization levels (see Figure 6), it turns out that such contribution is close to the maximum it could potentially achieve. As shown in the right panel of Figure 7, the average level of the \( P^* \) distribution equals 0.72 out of a maximum of 1. This suggests that, even if the force of homogamy is close to its full strength to pull up the levels of polarization, the overall effect of the former on the latter is rather modest. Another interesting pattern observed in Figure 7 is that at higher levels of college education, the relative contribution of homogamy to polarization levels tends to decrease slightly.

5.3. Change over time: Counterfactual decompositions

The results of the previous sections show pooled cross-sections using all available samples (including those countries with a single observation in time), which are helpful to give a rough idea of the relationship between couples of variables but fail to give an accurate assessment of the dynamics of change over time. Indeed, the relationships that are shown in Figures 2 to 5 are not informative on the direction of change of our parameters of interest and they could a priori be compatible with increasing or decreasing patterns over time. In addition, the results in Figures 6 and 7 are not reporting ‘true’ changes over time, but rather indicate how different our parameters of interest would look like under alternative homogamy assumptions. In order to investigate the dynamics of change appropriately, we will focus on those countries with at least two observations over time and consider the corresponding set of coupled observations (see paragraph after equation [13]). This leaves us with 289 coupled observations over time for 94 different countries covering all regions of the world.

Given the decomposition formulas [11]-[13], we can infer the percent contribution of education expansion (\( E \)), the force of homogamy (\( H \)) and the gender gap in education (\( G \)) to the changes over time in \( a, b+c, d \) and \( P^a \). Since these contributions add up to 100% (in absolute value), they can be easily represented via ternary plots [[[Note#9]]]. In each of these plots there are 289 ‘dots’, whose relative positions represent the percent contribution (in absolute terms) of \( E, H \) and \( G \) to the changes in the corresponding parameter of interest for all coupled observations over time in our dataset [[[Note#10]]]. When the dots are represented with a ‘+’ sign the corresponding parameter of interest has increased over time, while to opposite holds when it is represented with a ‘−’ sign. The share of ‘+’ signs in the four panels of Figure 8 corresponding to the changes in \( a, b+c, d \) and \( P^a \) are 28%, 72.3%, 67.8% and 71.6% respectively. That is: except for the share of non-college educated couples, our parameters of interest in this paper have
tended to increase over time (around 70% of the times in our coupled observations dataset).

A common feature of the four ternary plots shown in Figure 8 is the marginal role played by the gender gap in education in driving changes to our parameters of interest (as can be inferred from the fact that the corresponding clouds of 289 points never approach the ‘G’-vertex). Indeed, the average contributions of the gender gap to the changes over time in $a$, $b+c$, $d$ and $P^q$ are 1.3%, 2.6%, 3.1% and 1.2% respectively. In contrast, the contribution of the force of homogamy to the changes over time in the shares of the education distribution has been more notable: for the changes in $a$, $b+c$ and $d$, the contributions of ‘$H$’ have been 23.2%, 39.4% and 68.6%. However, when it comes to assess the impact of the force on homogamy on changes in polarization, the contribution falls to a mere 9.6%. Finally, the percent contribution of education expansion to the changes over time in $a$, $b+c$ and $d$ are 75.5%, 58%, 28.3% and 89.2% respectively. Therefore, besides the share of college educated couples, the education expansion has been by far the most decisive factor driving change over time in our parameters of interest. In particular, the changes in polarization levels are overwhelmingly accounted for by changes in educational attainment.

\[(((FIGURE 8)))\]

In Figure 8, one is unable to infer whether each of the forces $E$, $H$ and $G$ are separately contributing to increase or decrease our parameters of interest over time (i.e. the ‘overall’ increases / decreases of $a$, $b+c$, $d$ and $P^q$ reported in Figure 8 with ‘+’ and ‘–’ signs can be the result of separate positive or negative forces stemming from $E$, $H$ and $G$, see equations [11]-[13]). We conclude this section comparing the force of $E$ vis-à-vis $H$ when driving directional change in our parameters of interest. Given their negligible impact, we have decided not to report the changes attributable to the gender gap in education: they are available upon request. The different panels in Figure 9 jointly plot the values of ($\Delta_Ea$, $\Delta_Ha$), ($\Delta_E(b+c)$, $\Delta_H(b+c)$), ($\Delta_Ed$, $\Delta_Hd$) and ($\Delta_EP^q$, $\Delta_HP^q$). In each panel, the values of $\Delta_E$ and $\Delta_H$ are on the horizontal and vertical axes respectively and the corresponding averages are highlighted in the graph.

As can be seen in the first panel of Figure 9, the influence of education expansion has tended to decrease the share of non-college educated couples while the force of homogamy has tended to increase it. However, the relative force of the former has been much stronger in absolute terms than that of the latter (average impacts of $\bar{\Delta_E}a = -0.035$ and $\bar{\Delta_H}a = 0.007$ respectively), therefore resulting in a net decrease of that particular group in the education distribution. It turns out that the influence of education expansion has been particularly strong in increasing the share of mixed couples and the force of homogamy has tended to reduce them. As can be seen in the second panel of Figure 9, the ‘positive’ force of education expansion has been stronger in absolute terms than the ‘negative’ force of homogamy (average impacts of
\( \Delta_e(b+c) = 0.026 \) and \( \Delta_h(b+c) = -0.014 \) respectively, this is why the shares of mixed couples have tended to increase over time rather than decreasing. On the other hand, both education expansion and the force of homogamy have influenced in the increase of college educated couples. As can be seen in the third panel of Figure 9, their relative force in driving the shares of college educated couples is approximately the same (on average they have contributed to increase that share by \( \Delta_e d = 0.009 \) and \( \Delta_h d = 0.007 \) points respectively). Lastly, most changes in polarization are driven by education expansion rather than the force of homogamy. It turns out that assortative mating hardly plays a role in increasing education polarization, which is basically driven by increases in college education. This is illustrated in the fourth panel in Figure 9: the average impact of education expansion and assortative mating are \( \Delta_e P^\alpha = 0.044 \) and \( \Delta_h P^\alpha = 0.002 \) respectively.

(((FIGURE 9)))

5.4. Robustness checks

The results presented so far are based on a very concrete specification of our polarization index \( P^\alpha \) (i.e.: when \( \alpha=2 \)). One might wonder whether the results remain unaltered when choosing alternative specifications of alpha, the parameter measuring the importance given to the median category. As shown in the first three rows of Table 1, the average changes over time in polarization levels that can be attributed to \( E, H \) and \( G \) (\( \Delta_e P^\alpha, \Delta_h P^\alpha, \Delta_G P^\alpha \)) vary substantially with alpha. However, rows 4-6 show the corresponding percent contribution of each factor, which remains relatively stable with alternative specifications. Therefore, no matter what value of alpha we choose, the conclusions are essentially the same: the expansion of education is the main force driving changes over time in observed polarization levels, with the force of homogamy playing a rather secondary role.

(((TABLE 1)))

Experimentation with other social polarization indices – like the RQ polarization index proposed by Montalvo and Reynal-Querol (2005) – lead as well to very similar results which are not presented here for the sake of brevity: they are available upon request.

6. Summary and Concluding Remarks

In this article we have developed a neat methodology to disentangle and gauge the intertwined effects of three key social forces impacting on the levels of polarization in the marriage market between college and non-college populations. These forces are: the force of homogamy, the expansion of college education and the gender gap in education. We have applied this methodology to data from 120 countries and over 408
observations to investigate worldwide trends in polarization between low and high educated couples. Education was classified into two categories: non-college and college education.

We have shown that as college education expands, marriage markets become more polarized between college and non-college populations, as demonstrated by the growing absolute and relative numbers of college educated couples in the marriage markets. Together with the expansion of college education, the tendency to marry within the same educational group (i.e: the force of homogamy) is positively related to the percentage of college educated populations in the marriage market. In addition, at high levels of college education, the gender gap in education reverses in favor of women. A series of benchmarking and counterfactual exercises have allowed us to investigate the extent to which the shares of college educated couples and the corresponding polarization levels would have been different under different forces of homogamy. We found that if couples were formed purely at random, non-college and college educated couples would have decreased, on average, by 6.5% and 78.4% respectively and mixed couples would have increased by 77%. Interestingly polarization levels would have only decreased by an average 6.5% of their observed values. This shows a rather modest contribution of homogamy to polarization, which would have not been substantially different even if one hypothetically maximized the force of homogamy.

We have quantified the role of homogamy, education expansion and the gender gap in education on assortative mating and polarization levels. Our findings suggest that trends in polarization have been mainly driven by the process of education expansion rather than by the force of homogamy. The gender gap in education has played a very limited role. The contribution of the force of homogamy to the changes over time in non-college, mixed and college educated couples and polarization levels are 23.2%, 39.4%, 68.6% and 9.6% respectively, while the corresponding contributions of education expansion equal 75.5%, 58%, 28.3% and 89.2%. In sum, the polarizing trends we are observing worldwide are not overly influenced by the tendency towards homogamy, but seem to be rather mechanically driven by the share of the population with college education, which continues to expand in most societies. Inter alia, our results suggest that the gloomy scenarios of growing social divide through the gradual disappearance of mixed couples in favor of their homogamous counterparts are simply not occurring.

Aren’t the results presented in this paper “unavoidable” given the bounded nature of our variables and the parsimony of our models? After all, when no one is educated and when everyone is educated, there is no variability so polarization is zero, while in the process of education expansion, when some population groups get extra education there is increasing variability and, therefore, polarization. Hence, it is not surprising that our findings suggest the existence of an inverted U-shape in polarization as education expands (indeed, analogous results are reported by Dorius (2013), who finds an inverted U-shaped trajectory for the evolution of inequality in education explained by the fact that low-educated nations catch up with high-educated ones in many education
indicators). Far less obvious here are the following conclusions reached from our analysis: (i) Despite its purported importance and the widespread attention it has caught, assortative mating patterns seem to play a secondary role in driving the levels of polarization – a result that seems in line with the findings of Breen and Salazar (2011) in the US context; (ii) Even if the global reversal of the gender gap in higher education heralds the promise of massive social change, its effect on the polarizing trends analyzed here have been, so far, rather marginal; (iii) While the global expansion of college education has been the main driving force behind current polarizing trends, it has also greatly contributed to increase the share of mixed couples over time. Even if our findings are, no doubt, influenced by the admitted parsimony of the underlying models, the latter are good enough to: (a) neatly analyze the contribution of the key social forces that drive changes in polarization; and (b) give a broad overview of the macro-level trends that are taking place at the global level.

In conclusion, education polarization inevitably increases because increasing shares of the population have college education, but not because of homogamy patterns, which play a minor role. If trends were to continue and the college educated started becoming a majority, polarization levels would mechanically go down. Something similar to this might have already happened with basic literacy skills (Permanyer et al. 2013). At first the literate population was very scarce, then it started growing and now it is almost universal in many places. Was there any similar concern of increasing education polarization when literacy started to become widespread or it was celebrated as a measure of social progress?

References


Paul, A. 2006. “The real marriage penalty: Husbands and wives are increasingly likely to have similar incomes; Is a more divided society the result?.” New York Times November 19.


Appendix
In this appendix we show: (i) How to obtain the hypothetical education distributions \((a_0, b_0, c_0, d_0)\) and \((a_{\text{max}}, b_{\text{max}}, c_{\text{max}}, d_{\text{max}})\) that would be observed under extreme assortative mating assumptions, and (ii) How to derive the identities shown in [7].

(i) Deriving \((a_0, b_0, c_0, d_0)\) and \((a_{\text{max}}, b_{\text{max}}, c_{\text{max}}, d_{\text{max}})\).

A simple way of measuring the force of homogamy is to compare the observed education distribution \((a, b, c, d)\) with the hypothetical distribution \((a_0, b_0, c_0, d_0)\) that would be observed in case individuals did not care about their partners’ education (i.e.: couples were formed purely at random) while keeping the marginal education distribution of women and men unchanged. It is well known that under such independence assumption, one has that

\[
a_0 = (a + b)(a + c); b_0 = (a + b)(b + d); c_0 = (c + d)(a + c); d_0 = (c + d)(b + d)
\]  

[A1]

Since these are the expected frequencies that would be observed in the case that partners’ education played no role in the process of union formation, the difference between observed and expected values could be interpreted as measuring the force of homogamy. These differences will be labeled as

\[
a_p = a - a_0; b_p = b - b_0; c_p = c - c_0; d_p = d - d_0
\]  

[A2]

As shown in Permanyer et al (2013), one has that

\[
a_p = d_p = ad - bc \\
b_p = c_p = bc - ad
\]  

[A3] [A4]

so if one defines \(H=ad-bc\), then any education distribution \((a, b, c, d)\) can be rewritten as

\[
\begin{align*}
a &= a_0 + H \\
b &= b_0 - H \\
c &= c_0 - H \\
d &= d_0 + H
\end{align*}
\]  

[A5]

Equation [A5] shows a decomposition of observed cell frequencies as a sum of frequencies that would be observed in the case in which education status was irrelevant for couples’ formation plus a term \(H\) that can be interpreted as the force of homogamy. Positive values of \(H\) indicate that in the population under study, there is a tendency toward homogamy (indeed, this is the case for all observations in our sample).

Thus far, we have compared the education distribution shares with a hypothetical education distribution that results from assuming an absence of relationship between education status and couples’ formation. A conceptually related but somewhat different way of approaching the same problem is to try to answer the following question: To
what extent could the education distribution shares be different if maximal assortative mating patterns prevailed? It is straightforward to verify that when the marginal education distributions of women and men are fixed, the distribution that maximizes the force of homogamy is the one that concentrates the maximum number of couples in the main diagonal of the couples’ education distribution table, that is:

<table>
<thead>
<tr>
<th></th>
<th>Non-college Woman</th>
<th>College Woman</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-college Man</td>
<td>$a + \min{b,c}$</td>
<td>$b - \min{b,c}$</td>
<td>$a + b$</td>
</tr>
<tr>
<td>College Man</td>
<td>$c - \min{b,c}$</td>
<td>$d + \min{b,c}$</td>
<td>$c + d$</td>
</tr>
<tr>
<td>Total</td>
<td>$a + c$</td>
<td>$b + d$</td>
<td>$1$</td>
</tr>
</tbody>
</table>

Therefore, we define $(a_{\text{max}}, b_{\text{max}}, c_{\text{max}}, d_{\text{max}})$ as

$$
\begin{align*}
    a_{\text{max}} &= a + \min\{b,c\} \\
    b_{\text{max}} &= b - \min\{b,c\} \\
    c_{\text{max}} &= c - \min\{b,c\} \\
    d_{\text{max}} &= d + \min\{b,c\}
\end{align*}
$$

With these definitions, it is straightforward to check that the first three identities in [10] are indeed the same, that is:

$$
\frac{a - a_0}{a_{\text{max}} - a_0} = \frac{(b + c)_0 - (b + c)}{(b + c)_0 - (b + c)_{\text{max}}} = \frac{d - d_0}{H + \min\{b,c\}} = m^* \quad [A7]
$$

(ii) Derivation of [7].

The derivation of [7] is long and involved; it is explained in the following steps.

**Step 1.** Write $a,b,c$ and $d$ in terms of $E_m, E_w$ and $H$ (see equations [2], [3] and [5] for definitions). This involves solving the following equations system:

$$
\begin{align*}
    a + b &= 1 - E_m \\
    c + d &= E_m \\
    a + c &= 1 - E_w \\
    b + d &= E_w \\
    ad - bc &= H
\end{align*} \quad [A8]
$$

Solving [A8] we obtain
\[
\begin{align*}
 a &= (1 - E_w)(1 - E_m) + H \\
b &= E_w(1 - E_m) - H \\
c &= E_m(1 - E_w) - H \\
d &= E_w E_m + H
\end{align*}
\]  
\[ [A9] \]

Therefore, *any* education distribution \((\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})\) with gender equality (i.e. \(E_m = E_w = E\)) can be written as

\[
\begin{align*}
 \tilde{a} &= (1 - E)^2 + H \\
 \tilde{b} &= E(1 - E) - H \\
 \tilde{c} &= E(1 - E) - H \\
 \tilde{d} &= E^2 + H
\end{align*}
\]  
\[ [A10] \]

**Step 2.** Starting from the education distribution \((a, b, c, d)\) we derive another distribution \((a_g, b_g, c_g, d_g)\) with the same marginals and the same force of homogamy as the original one but with no gender gap in education. For that purpose, we need to solve the following equations system:

\[
\begin{align*}
 a_g + b_g &= a + b \\
 c_g + d_g &= c + d \\
 a_g + c_g &= a + c \\
 b_g + d_g &= b + d \\
 a_g d_g - b_g c_g &= ad - bc \\
 b_g &= c_g
\end{align*}
\]  
\[ [A11] \]

Solving [A11] we obtain:

\[
\begin{align*}
 a_g &= a + \left(\frac{G}{2}\right)^2 \\
b_g &= b + c - \left(\frac{G}{2}\right)^2 \\
c_g &= b + c - \left(\frac{G}{2}\right)^2 \\
d_g &= d + \left(\frac{G}{2}\right)^2
\end{align*}
\]  
\[ [A12] \]

**Step 3.** Since [A12] is obtained after imposing gender equality, \((a_g, b_g, c_g, d_g)\) can be seen as a particular case of \((\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})\). From [A10] and [A12] we can deduce that
Lastly, the identities in [7] obtain after basic algebraic manipulations of [A13].

Notes

1. As will be analyzed in detail below, the presence of mixed couples will be mediated by the levels of educational homogamy and gender inequality in education.

2. Clearly, whether these different welfare outcomes are finally observed or not is highly contingent upon the returns to education (Bradley 2001) and the family formation and living arrangement patterns. The potential benefits of education expansion in favor of women can be jeopardized in case the correlation between women’s education and earnings is small—a problem that might actually arise in highly segregated education systems and labor markets (Charles and Grusky 2004).

3. While finer partitions of the education distribution could be feasible with the available data, in this paper we are solely interested in the boundary between college and non-college education. The additional detail that could be gained with further refinements would not add to the paper’s main research aim.

4. Recall that in this paper, the partitioning of the education distribution into two groups is very crude. Even if the educational distribution partition upon which these notions are built is typically finer—for instance, defined over four or more educational groups—the meaning of the terms is nonetheless preserved.

5. This way of measuring the force of homogamy was already suggested in Permanyer et al (2013). It bears some resemblance with the classical odds ratio parameter that is the basis of loglinear models $\Omega=(a/c)/(b/d)=ad/bc$.

6. As $\alpha \to 0$, the relative contribution of the median category increases while for increasing values of $\alpha$, the contribution of the median category decreases.

7. The ordinal polarization index suggested by Apouey (2007:885) requires a complete ordering of the categories one is working with. This is not the case here because in our 4 category framework $(a, b, c, d)$ $b$ and $c$ cannot be ranked vis-à-vis each other. In order to remedy this problem we consider the partition in 3 groups $(a, b \cup c, d)$, which is completely ordered (all elements can be ranked vis-à-vis each other). The polarization
index presented in equation [8] is an adaptation of Apouey’s index for this 3 group partition of the population in union.

8. $P_{\alpha}$ satisfies the following classical properties expected from a polarization index: (i) $P_{\alpha}(1/2,0,0,1/2)=1$ (i.e.: polarization is maximized in the bipolar case where half of the couples are college educated and the other half are non-college educated) and $P^{\alpha}(1,0,0,0)=P^{\alpha}(0,1,0,0)=P^{\alpha}(0,0,1,0)=P^{\alpha}(0,0,0,1)=0$ (i.e.: polarization is minimized when all population is concentrated in a given cell and there is no variability).

9. A ‘ternary plot’, ‘ternary graph’, ‘triangle plot’ or ‘simplex plot’ is a barycentric plot on three variables which add up to a certain constant. It graphically depicts the ratios of the three variables as positions in an equilateral triangle.

10. Intuitively, whenever a dot is close to the vertex ‘$E$’ (for instance), it means that the change over time in the corresponding parameter has been mainly caused by the education expansion effect.
Figure 1. Education expansion and gender equality. Observed data from 1970 to 2000 (continuous lines) are taken from Lutz et al (2007). Estimated data from 2010 to 2050 (dashed lines) come from KC et al (2010). We have highlighted in bold the trajectories of a few selected countries labelled with the ISO 3166 codes.
Figure 2. Relationship between college education ($E$) and the force of homogamy ($H$). Authors’ calculations using 408 observations from 120 countries.
Figure 3. Relationship between college education \( (E) \) and the share of non-college, mixed and college educated couples \( (a, b+c \text{ and } d) \). Authors’ calculations using 408 observations from 120 countries.
Figure 4. Relationship between college education levels in the horizontal axis ($E$) and the gender gap in education ($G$). Authors’ calculations using 408 observations from 120 countries.
Figure 5. Relationship between college education levels in the horizontal axis ($E$) and levels of polarization ($P^\alpha$) under alternative assumptions in the vertical axis. Authors’ calculations using 408 observations from 120 countries.
Figure 6. Density functions of the percentual contribution of homogamy to the observed shares of non-college, mixed and college educated couples and observed polarization levels (corresponding averages marked with vertical lines at -77%, 6.5%, 6.5% and 78.4%). Authors’ calculations using 408 observations from 120 countries.
Figure 7. College education (E) vs benchmarked education distribution \( m^* \) (right panel) and college education (E) vs benchmarked polarization \( P^* \) (left panel). Authors’ calculations using 408 observations from 120 countries.
Figure 8. Percent contributions of education expansion ($E$), the force of homogamy ($H$) and the gender gap in education ($G$) to changes in the education distribution shares (top two and bottom left panels) and changes in polarization levels (bottom right panel). Authors’ calculations using 289 observations from 94 countries.
Figure 9. Share changes in non-college educated couples (top left panel), mixed couples (top right panel), college educated couples (bottom left panel) and polarization changes (bottom right panel) due to college education (horizontal axis) vs changes due to the force of homogamy (vertical axis). Authors’ calculations using 289 observations from 94 countries.
<table>
<thead>
<tr>
<th></th>
<th>0.1</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.006</td>
<td>0.024</td>
<td>0.036</td>
<td>0.044</td>
<td>0.045</td>
<td>0.044</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.002</td>
<td>0.003</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>91.1%</td>
<td>92.4%</td>
<td>95%</td>
<td>89.2%</td>
<td>85.4%</td>
<td>82%</td>
<td>79.1%</td>
</tr>
<tr>
<td></td>
<td>7.9%</td>
<td>6.6%</td>
<td>4.2%</td>
<td>9.6%</td>
<td>13.2%</td>
<td>16.4%</td>
<td>19.3%</td>
</tr>
<tr>
<td></td>
<td>1.1%</td>
<td>1%</td>
<td>0.8%</td>
<td>1.2%</td>
<td>1.4%</td>
<td>1.5%</td>
<td>1.6%</td>
</tr>
</tbody>
</table>

Table 1. Robustness checks using alternative specifications of the polarization index $P^\alpha$. Average changes in polarization over time attributable to $E$, $H$ and $G$ (rows 1 – 3) and the corresponding average percent contributions (rows 4 – 6). Authors’ calculations using 289 observations from 94 countries.