Centre d'Estudis Demogràfics

MARRIAGE MARKET POLARIZATION IN THE TIME OF COLLEGE EDUCATIONAL EXPANSION

Iñaki PERMANYER
Albert ESTEVE
Joan GARCÍA

454

PAPERS DE DEMOGRAFIA

2015

UNIVERSITAT AUTÒNOMA DE BARCELONA
08193 Bellaterra (Barcelona)

Centre d'Estudis Demogràfics

MARRIAGE MARKET POLARIZATION IN THE TIME OF COLLEGE EDUCATIONAL EXPANSION

Iñaki PERMANYER
Albert ESTEVE
Joan GARCÍA-ROMÁN

454

Centre d'Estudis Demogràfics

2015

Abstract.- *Marriage market polarization in the time of college educational expansion*

Are societies becoming more polarized between college-educated and non-college-educated people? 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 (the propensity to marry within the same educational group), the expansion of college education and the gender gap in education are contributing to polarization in the marriage market between college- and non-college-educated populations worldwide. To this end, we assembled census and survey microdata from 120 countries and 408 samples from 1960 to 2011 that represent 98% of the world's population. We developed a simple yet effective decomposition model that neatly assesses the impact of these three forces on couples' education distribution and the corresponding polarization levels. The 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-educated and non-college-educated populations. These levels, which are increasing worldwide, are rather mechanically driven by the share of the population with a 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.

Keywords.- Marriage market; education; homogamy: university studies; mixed couples.

Resum.- Les societats estan polaritzant-se entre les persones amb formació universitària i les que no tenen estudis universitaris? La força de l'homogàmia genera unions més igualitàries a costa de societats més polaritzades? En aquest treball examinem els patrons d'aparellament assortiu per investigar fins a quin punt la força de l'homogàmia (la propensió a casar-se dins del mateix grup educatiu), l'expansió de l'educació universitària i la bretxa de gènere en l'educació contribueixen a la polarització del mercat matrimonial entre poblacions universitàries i no universitàries. Per a això, vam reunir microdades de censos i enquestes de 120 països i 408 mostres de 1960 a 2011 que representen el 98% de la població mundial. Hem desenvolupat un model de descomposició senzill però eficaç que avalua adequadament l'impacte d'aquestes tres forces en la distribució educativa de les parelles i els nivells de polarització corresponents. Els resultats mostren que la força de l'homogàmia i la bretxa de gènere en l'educació han tingut un paper molt limitat a l'hora de determinar els nivells de polarització del mercat matrimonial entre poblacions educades en universitat i no universitàries. Aquests nivells, que creixen a nivell mundial, es veuen impulsats mecànicament per la proporció de la població amb una formació universitària, que continua expandint-se en la majoria de les societats. No es produeixen els temuts escenaris que preveuen la contracció i la desaparició gradual de parelles mixtes en favor de particions compartimentades entre parelles educades a la universitat i les que no tinguin formació universitària.

Paraules clau.- Mercat matrimonial; Educació; Homogàmia; Estudis universitaris; Parelles mixtes.

Resumen.- ¿Están las sociedades cada vez más polarizadas entre personas con educación universitaria y personas sin educación universitaria? ¿La fuerza de la homogamia está generando uniones más igualitarias a costa de sociedades más polarizadas? En este artículo, examinamos los patrones de apareamiento selectivo para investigar hasta qué punto la fuerza de la homogamia (la propensión a casarse dentro del mismo grupo educativo), la expansión de la educación universitaria y la brecha de género en la educación están contribuyendo a la polarización en el mercado matrimonio entre las poblaciones con educación universitaria y no universitaria en todo el mundo. Con este fin, reunimos microdatos de censos y encuestas de 120 países y 408 muestras de 1960 a 2011 que representan el 98% de la población mundial. Desarrollamos un modelo de descomposición simple pero efectivo que evalúa cuidadosamente el impacto de estas tres fuerzas en la distribución de la educación de las parejas y los niveles de polarización correspondientes. Los resultados muestran que la fuerza de la homogamia y la brecha de género en la educación han desempeñado un papel muy limitado en la determinación de los niveles de polarización del mercado matrimonial entre las poblaciones con educación universitaria y sin educación universitaria. Estos niveles, que están aumentando en todo el mundo, están impulsados más bien mecánicamente por la proporción de la población con educación universitaria, que continúa expandiéndose en la mayoría de las sociedades. Los temidos escenarios que predicen la contracción y la desaparición gradual de las parejas mixtas a favor de particiones compartimentadas entre parejas con educación universitaria y parejas con educación inferior a la universidad no se están produciendo.

Palabras clave.- Mercado matrimonial; Educación; Homogamia: Estudios universitarios; Parejas mixtas.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Papers de Demografia, 454 (2015), 1-33 pp.

MARRIAGE MARKET POLARIZATION IN THE TIME OF COLLEGE EDUCATIONAL EXPANSION[1]

Iñaki PERMANYER UGARTEMENDIA
ipermanyer@ced.uab.es

Albert ESTEVE PALÓS
aesteve@ced.uab.es

Joan GARCÍA ROMÁN
jgarcia@ced.uab.es

1.- Introduction

In recent years, there has been an upsurge of interest among social scientists regarding the implications of increasing levels of educational homogamy 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 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 contribute to these levels of polarization. For this purpose, we have developed a simple yet effective decomposition model that neatly assesses the contributions of these three forces to patterns of assortative mating and to the corresponding polarization levels.

We have assembled a database with more than 400 samples from 120 countries worldwide from the 1960s to the present day. The data mostly came from the IPUMS international census microdata samples database, with complementary use of various household surveys. The results show mounting levels of marriage market polarization between college- and non-college-educated populations that is mostly and mechanically driven by the expansion

[1] This paper was presented at the Population Association of America (PAA). *2015 Annual Meeting*. San Diego, CA (USA), April 30-May 2.

of college education. By contrast, the force of homogamy and the gender gap in education have played very limited roles in determining the changes in polarization levels observed worldwide. The negative 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 literacy was during the 19th century and a significant part of the 20th century. 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 do non-college-educated ones (Atkinson et al. 2011). The academic outcomes of children of college-educated parents tend to be higher than those of children of parents with lower levels of education. Furthermore, education is among the most important stratification variables of demographic behavior, to the extent that Lutz and colleagues claim that education should be routinely added to all types 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) affect the distribution of welfare across populations in contemporary societies. Given the significant and positive implications of educational attainment on individuals and their families, this effect may be particularly strong when at least one of the members of the couple has a 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 increasingly unequal, impoverished and polarized societies between those who are multiply advantaged and those who are multiply deprived. This is 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. Educational attainment has become one of the main structuring dimensions of contemporary marriage markets, if not the most important dimension. Education influences the age at which individuals enter the marriage market and shapes their context of opportunities and expectations toward 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, Scwhartz and Mare 2005, Schwartz 2013).

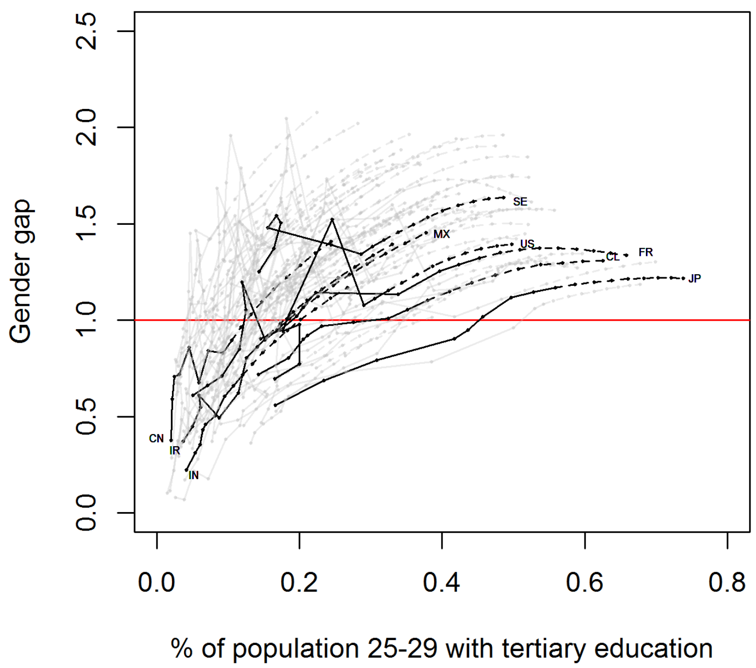
The rise of the force of homogamy has gone hand-in-hand with the dramatic *expansion of education*. It is widely acknowledged that the Western model of mass education has been diffusing around the world for many decades (Meyer et al 1977). This expansion includes rising literacy rates (Crafts 2002) and increases in schooling enrollment 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. Regarding college education, by 1970, 6.4% of the world's population aged 25-29 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 a college education but also their potentially less educated partners [2].

An important feature of the expansion of college education that may contribute to 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). Despite initially favoring males, we are currently witnessing the opposite pattern: the gender gap in education is not only closing but, in many cases, even

[2] 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.

reversing in favor of women (see Figure 1) (Esteve, Garcia and Permanyer 2012). Experts' worldwide education projections suggest that these patterns will not only continue but will also increase over the next decades (KC et al. 2010; Lutz and KC, 2011). In 1970, 63.6% of college-educated people were men. In 2000, that percentage decreased to 52.6%. By 2050, the overall percentage of men among the college-educated population is expected to decrease 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 may reduce the share of college-educated couples while increasing the number of mixed couples[3]. In 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).

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.

[3] Clearly, whether these different welfare outcomes are finally observed 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 if 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).

This paper examines trends and patterns in marriage market polarization between college- and non-college-educated 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. Because education expansion increases *both* the probability that college-educated individuals will partner among themselves *and* the probability that an individual without college education will partner with a college-educated one, the implications of college education expansion with regard to polarization are not *a priori* clear. Whether this 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 increasing the number of couples between college- and non-college-educated populations and reduce the number of couples among college-educated populations, thus lowering polarization levels.

Will we witness a situation of divergent destinies in which societies break into two opposing and gradually distant poles? Or will the unprecedented global changes in access to education in favor of women tilt the balance in the opposite direction? To investigate these issues, we introduce a straightforward and novel methodology that builds 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. In 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 with 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 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 represent more than 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 the women are 25-34 years old. In this way, we avoid the possible biases and distortions that arise when the 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 ambiguity.

4.- Analytical strategy

We introduce some basic definitions and notation that will be used throughout the paper. Because we are interested in assessing and decomposing polarization levels in the marriage market, we restrict our attention to the population living in union. In this article, we are interested in the top of the education distribution, so we consider only two educational groups: those without college education and the college educated. This approach generates a 2×2 contingency table with 4 possible combinations depending on the education level of the partners[4]. The first combination corresponds to couples in which neither of the members has a 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 a college education; 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[5]. For simplicity, this distribution will be referred to as *educational assortative mating*, which will be briefly denoted as (*a, b, c,d*). Because *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 in which both members are college and non-college educated (i.e., the groups that 'threaten' to race ahead of or to be left behind 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 with the distribution (*a, b, c,d*) (which are defined below in section 4.2). These shares are directly influenced by several factors, among which we highlight the following:

(*i*) The *expansion of college education*, measured as the share of the 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, 'large). Formally, we have:

$$E = \frac{b+c+2d}{2} \qquad [1].$$

[4] Although 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.

[5] 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.

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

$$E_w = b + d \qquad [2],$$

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

$$E_m = c + d \qquad [3].$$

Clearly, we have

$$E = \frac{E_w + E_m}{2} \qquad [4].$$

(*ii*) The *force of homogamy* (that is, the extent to which couples consist 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 in which 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[6]:

$$H = ad - bc \qquad [5].$$

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

$$G = b - c = E_w - E_m \qquad [6].$$

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

[6] This way of measuring the force of homogamy was suggested in Permanyer et al (2013). It bears some resemblance to the classical odds ratio parameter that is the basis of loglinear models Ω=(a/c)/(b/d)=ad/bc.

$$\left.\begin{aligned} a = \varphi_a(E,H,G) &= (1-E)^2 + H - \left(\frac{G}{2}\right)^2 \\ b = \varphi_b(E,H,G) &= E(1-E) - H + \frac{G}{2}\left(\frac{G}{2}+1\right) \\ c = \varphi_c(E,H,G) &= E(1-E) - H + \frac{G}{2}\left(\frac{G}{2}-1\right) \\ d = \varphi_d(E,H,G) &= E^2 + H - \left(\frac{G}{2}\right)^2 \end{aligned}\right\} \quad [7].$$

These fundamental identities are one of the key contributions of this paper. They clearly show 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 types of counterfactual analysis (see section 4.4). The derivation of these identities is complicated, 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 divided into 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:

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

where α is a non-negative number that can be interpreted as a polarization sensitivity parameter[7] (in the main empirical applications of the paper, we use the intermediate value

[7] When α 0, the relative contribution of the median category to polarization levels increases, whereas for increasing values of α, the contribution of the median category decreases.

of α=2, but other values have been investigated as robustness checks in section 5.4). This index is an *ad hoc* adaptation of the ordinal polarization index suggested by Apouey (2007: 885) for the case in which one addresses 3 categories[8]. P^{α} is a standard index of polarization[9] that 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.

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 are always higher than the hypothetical shares of homogamous couples that would be observed if 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 *a priori* be formed. Stated otherwise, although individuals have a *tendency* to partner with other individuals with the same educational attainment, this tendency is not universal, and the share of mixed couples is higher than the share that would be observed in a maximal homogamy scenario in which individuals absolutely prioritized partners with the same educational attainment. In this context, one could benchmark 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 this 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 illustrate how to derive these two hypothetical education distributions, which

[8] The ordinal polarization index suggested by Apouey (2007:885) requires a complete ordering of the categories. 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. To remedy this problem, we consider the partition in 3 groups (a, b 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.

[9] $P\alpha$ 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 in which 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 the entire population is concentrated in a given cell and there is no variability).

are denoted by (a_0, b_0, c_0, d_0) and (a_{max}, b_{max}, c_{max}, d_{max}), respectively. The observed shares of 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 seen in the following inequalities (see the appendix for details):

$$\left.\begin{array}{c} a_0 \le a \le a_{\max} \\ (b+c)_{\max} \le b+c \le (b+c)_0 \\ d_0 \le d \le d_{\max} \\ P_0^\alpha := P^\alpha(a_0,b_0,c_0,d_0) \le P^\alpha(a,b,c,d) \le P^\alpha(a_{\max},b_{\max},c_{\max},d_{\max}) := P_{\max}^\alpha \end{array}\right\} \quad [9]\cdot$$

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

$$\left.\begin{array}{rl} a^* &= \dfrac{a-a_0}{a_{\max}-a_0} \\ (b+c)^* &= \dfrac{(b+c)_0-(b+c)}{(b+c)_0-(b+c)_{\max}} \\ d^* &= \dfrac{d-d_0}{d_{\max}-d_0} \\ P^* &= \dfrac{P^\alpha-P_0^\alpha}{P_{\max}^\alpha-P_0^\alpha} \end{array}\right\} \quad [10]\cdot$$

Because P_0^α measures the polarization levels that would be observed in a hypothetical scenario in which the force of homogamy is zero (i.e., what might be referred to as 'baseline polarization'), $P^\alpha - P_0^\alpha$ measures *in absolute terms* the 'amount of polarization' that is attributable to the force of homogamy. 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 $P_{\max}^\alpha - P_0^\alpha$, thus obtaining P^*: a measure *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 those that would be observed under alternative hypothetical levels of homogamy. Although this method is informative and interesting in its own right, it fails to take into account the dynamics of *change* over time. Among other things, it is unclear whether changes in the force of homogamy have been more or less decisive than changes in the gender gap in education or changes in the share of the population with a college education when shaping education distributions and the corresponding polarization levels.

To assess the separate impact of the expansion of college education, the force of homogamy and the gender gap in college education on the educational distribution of couples and the corresponding polarization levels, we conducted 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^α 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). 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 at 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 at times t_1 and t_2 as E_1, H_1, G_1 and E_2, H_2, G_2, respectively. Define also $d_2^E = \varphi_d(E_2, H_1, G_1)$, $d_2^H = \varphi_d(E_1, H_2, G_1)$ and $d_2^G = \varphi_d(E_1, H_1, G_2)$ 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 the 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_2^E, a_2^H, a_2^G and of mixed couples $(b+c)_2^E, (b+c)_2^H, (b+c)_2^G$. In this article, we are interested in the differences between the observed education distributions and the counterfactual ones:

$$\left.\begin{array}{c} \Delta_E a = a_2^E - a_1, \Delta_E(b+c) = (b+c)_2^E - (b+c)_1, \Delta_E d = d_2^E - d_1 \\ \Delta_H a = a_2^H - a_1, \Delta_H(b+c) = (b+c)_2^H - (b+c)_1, \Delta_H d = d_2^H - d_1 \\ \Delta_G a = a_2^G - a_1, \Delta_G(b+c) = (b+c)_2^G - (b+c)_1, \Delta_G d = d_2^G - d_1 \end{array}\right\} \quad [11].$$

Interestingly, the observed (i.e., 'real') difference between education distributions over time can be written as:

$$\left.\begin{array}{c} a_2 - a_1 = \Delta_E a + \Delta_H a + \Delta_G a \\ (b+c)_2 - (b+c)_1 = \Delta_E(b+c) + \Delta_H(b+c) + \Delta_G(b+c) \\ d_2 - d_1 = \Delta_E d + \Delta_H d + \Delta_G d \end{array}\right\} \quad [12].$$

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, for example, homogamy to the share of college-educated couples can be simply calculated as $100 \cdot |\Delta_H d| / (|\Delta_E d| + |\Delta_H d| + |\Delta_G d|)$). 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:

$$\left.\begin{aligned}\Delta_E P^{\alpha} &= P^{\alpha}(a_2^E, b_2^E, c_2^E, d_2^E) - P^{\alpha}(a_1, b_1, c_1, d_1)\\ \Delta_H P^{\alpha} &= P^{\alpha}(a_2^H, b_2^H, c_2^H, d_2^H) - P^{\alpha}(a_1, b_1, c_1, d_1)\\ \Delta_G P^{\alpha} &= P^{\alpha}(a_2^G, b_2^G, c_2^G, d_2^G) - P^{\alpha}(a_1, b_1, c_1, d_1)\end{aligned}\right\} \qquad [13].$$

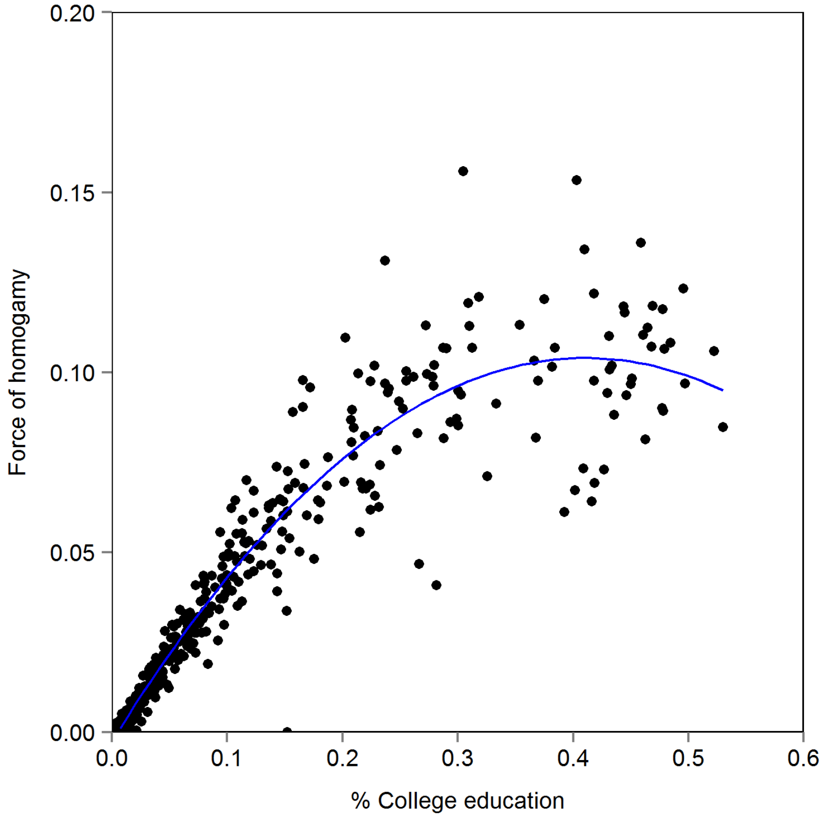
To take advantage of the decompositions shown in equations [11], [12] and [13], we need to work with the *set of coupled observations* that we can construct when more than one observation over time is available for a given country. That is, if we have n observations over time (in times $t_1, t_2, \ldots, t_n$) for a certain country, 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 outline the existing relationship between our main variables. 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. When college education is from 5% to 20%, the force of homogamy increases in a linear way. Beyond 20%, there is greater 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) that report increasing homogamy among the college educated.

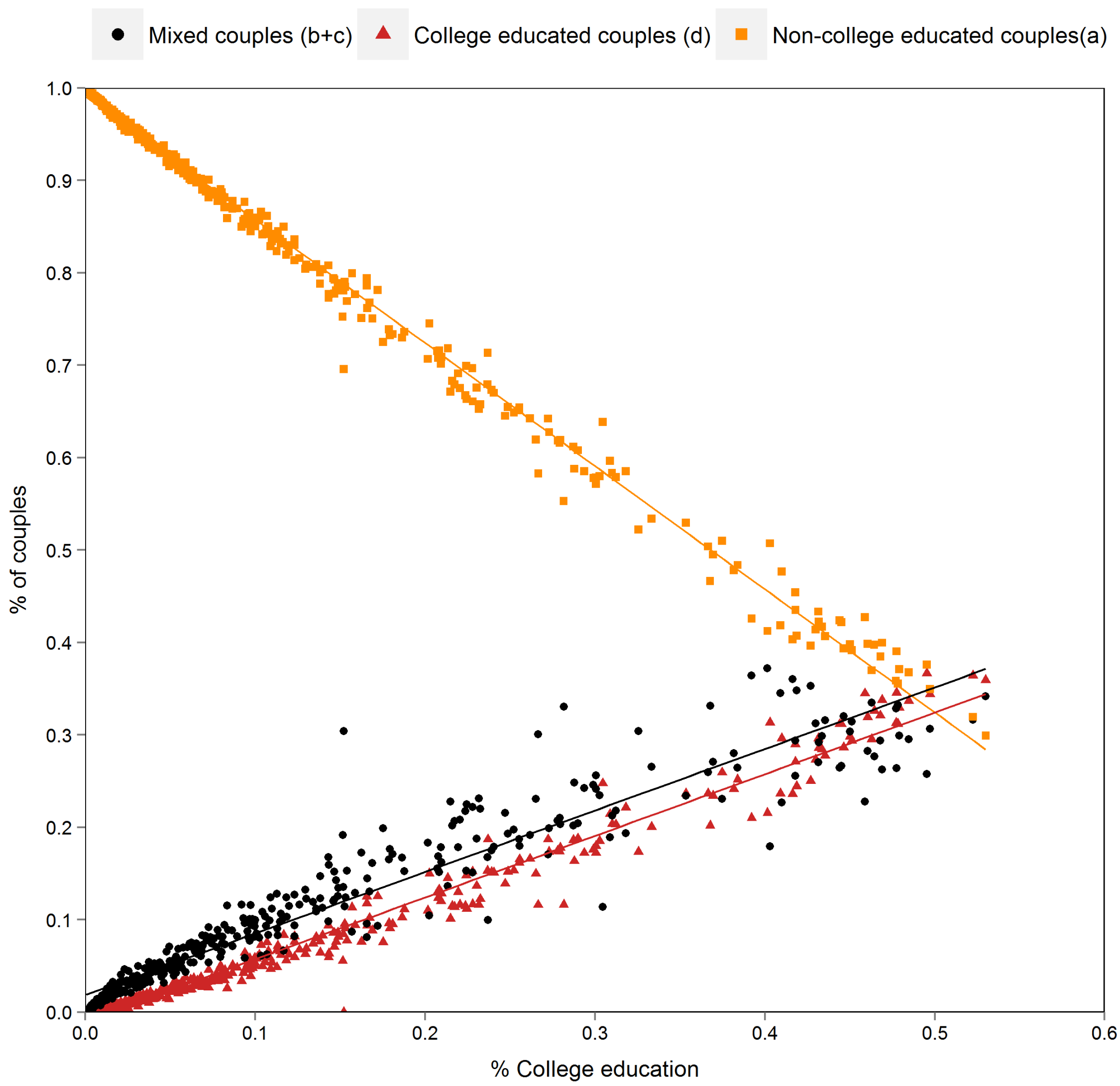
Figure 2.- Relationship between college education (*E*) and the force of homogamy (*H*)



Authors' calculations using 408 observations from 120 countries.

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 shown, the share of non-college-educated couples decreases linearly with the expansion of college education; it begins at approximately 100% when there are no college-educated individuals, and it falls near 30% when college education approaches 50%. In contrast, 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 is approximately 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 than 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.

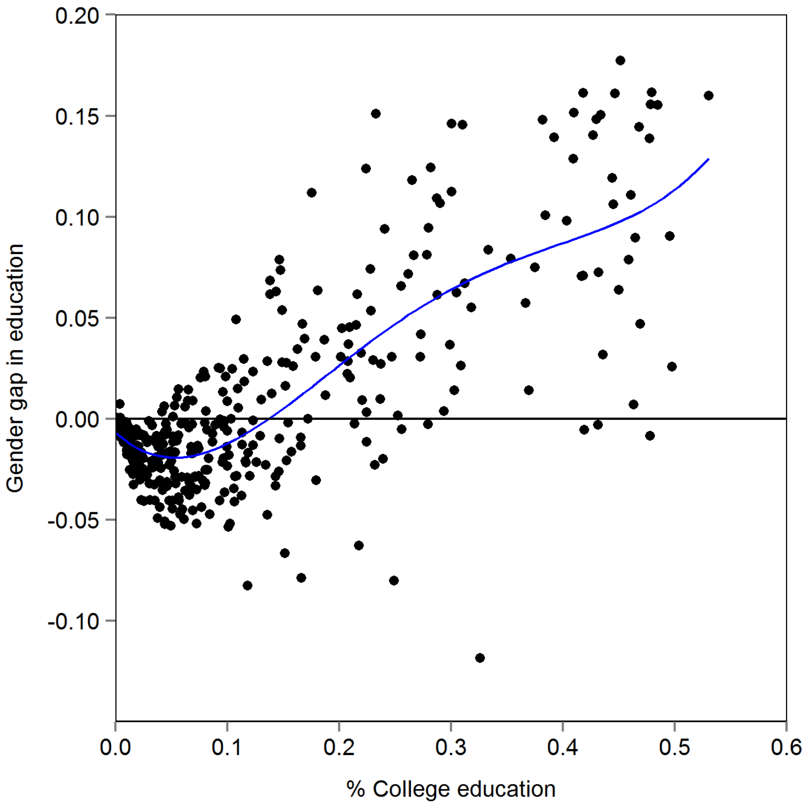
Figure 3.- Relationship between college education (*E*) and the share of non-college-educated, mixed and college-educated couples (*a*, *b*+*c* and *d*)



Authors' calculations using 408 observations from 120 countries.

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 the overall trends reported in Figure 1 referring to the entire population.

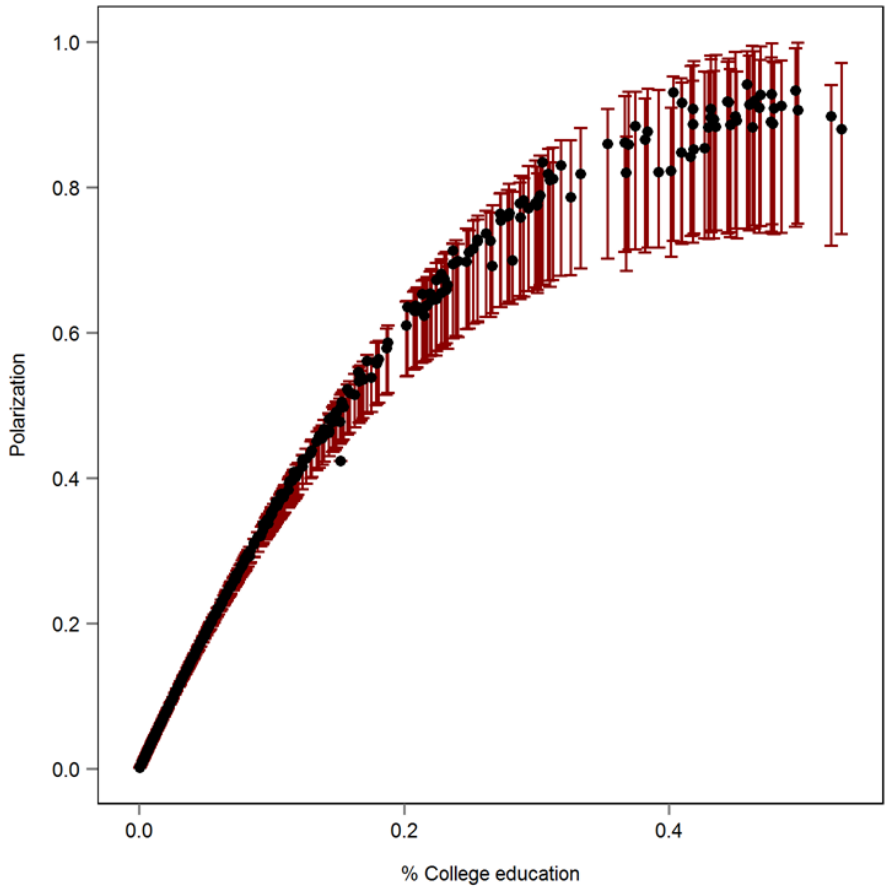
Figure 4.- Relationship between college education levels in the horizontal axis (*E*) and the gender gap in education (*G*)



Authors' calculations using 408 observations from 120 countries.

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 an 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 (approximately 30%), the levels of polarization do not increase as rapidly. Although the observed levels of polarization may look surprisingly high (i.e., hovering at approximately 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.

Figure 5.- Relationship between college education levels in the horizontal axis (*E*) and levels of polarization (P^α) under alternative assumptions in the vertical axis

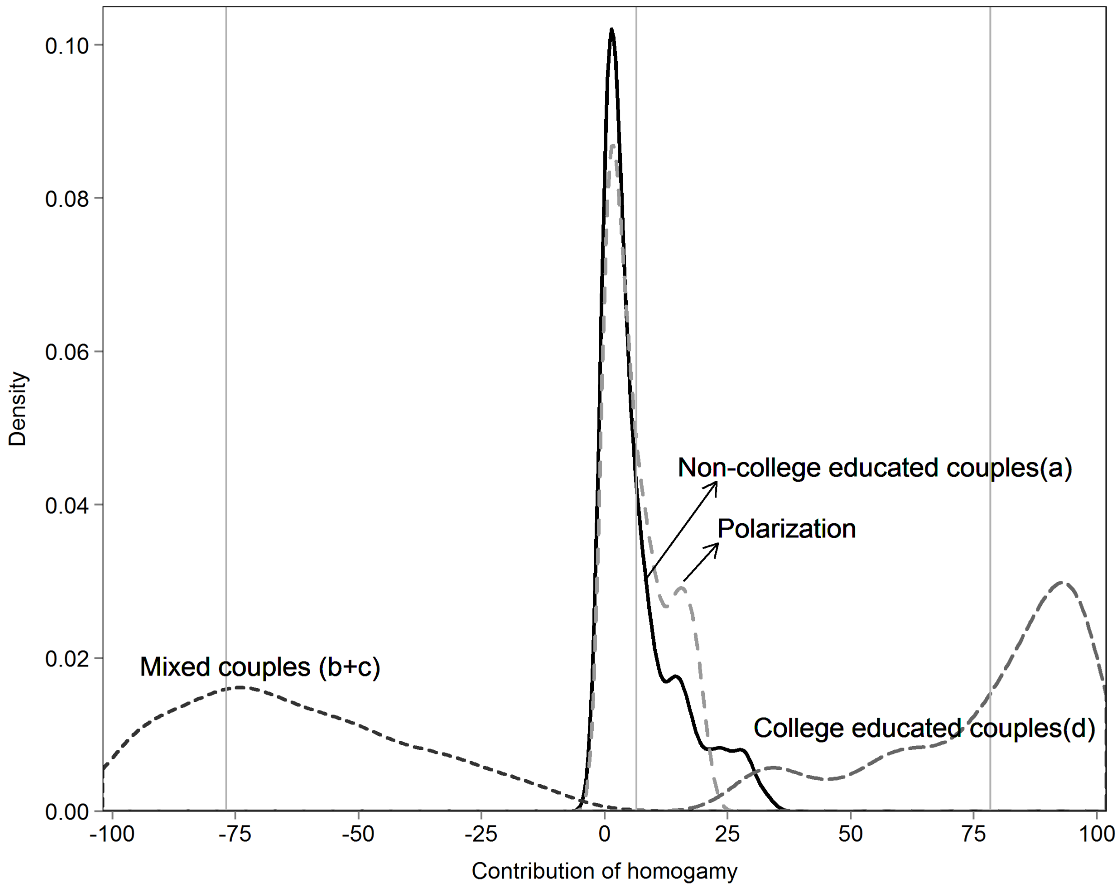


Authors' calculations using 408 observations from 120 countries.

5.2.- Benchmarking: assortative mating under extreme homogamy scenarios

In this section, we investigate how the shares of the education distribution and the corresponding polarization levels would look if alternative forces of homogamy had prevailed. More specifically, we begin by comparing the observed shares of the education distribution (a, $b+c$ and d) with those that would be observed in the absence of homogamy (a_0, b_0+c_0 and d_0). For this purpose, Figure 6 plots the density functions associated with the values of $100(a–a_0)/a$, $100(b+c–(b_0+c_0))/(b+c)$ and $100(d–d_0)/d$ for all the samples included in our dataset. These indices measure the extent to which the observed values in the shares of each type of couple (*a*, *b*+*c*, and *d*) can be attributed to the force of homogamy. As shown, 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). In contrast, 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, by 77% of their observed values, whereas the shares of college-educated couples would *decrease*, on average, by 78.4%.

Figure 6.- Density functions of the percentage contribution of homogamy to the observed shares of non-college-educated, mixed and college-educated couples and observed polarization levels (corresponding averages marked with vertical lines at -77%, 6.5%, 6.5% and 78.4%)

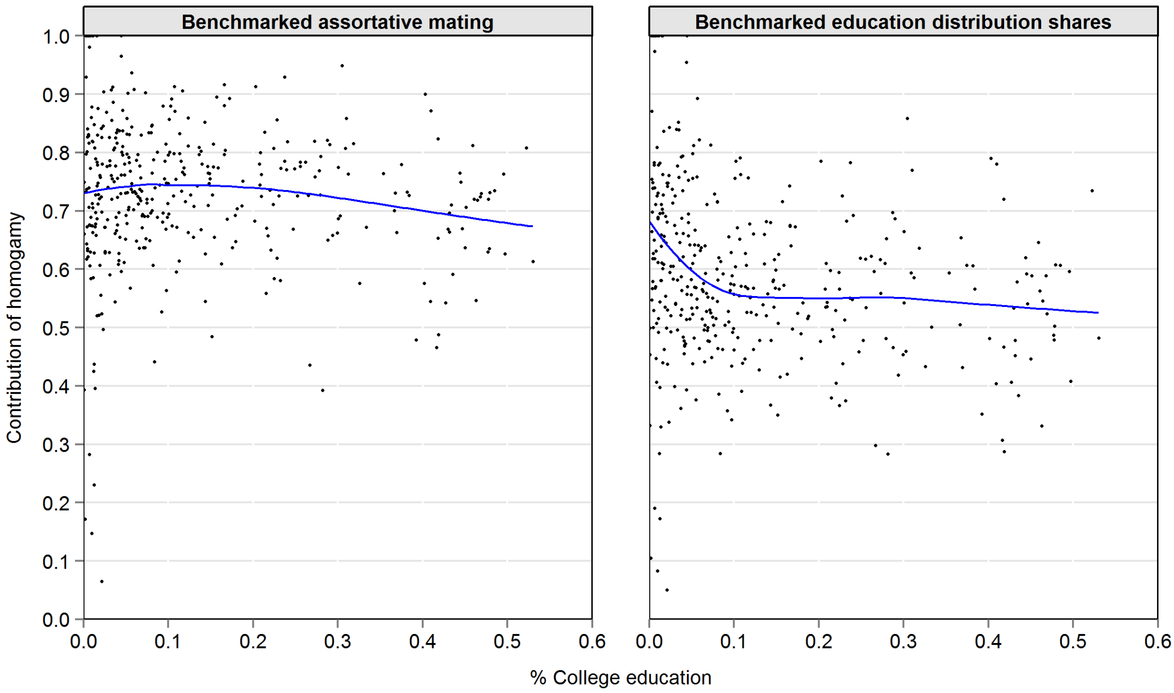


Authors' calculations using 408 observations from 120 countries.

Lastly, we compare the observed polarization levels $P^{\alpha}(a, b, c, d)$ with the ones that would be observed in the absence of homogamy $P^{\alpha}(a_0, b_0, c_0, d_0)$. Figure 6 also plots the percent contribution of the force of homogamy to existing polarization levels $(100(P^{\alpha}(a, b, c, d) - P^{\alpha}(a_s, b_s, c_s, d_s)) / P^{\alpha}(a, b, c, d))$. As shown, 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 shown, 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 despite being a factor that greatly contributes to increasing (resp. decrease) the share of homogamous (resp. mixed) couples, the force of homogamy is not at its full strength because 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.

Figure 7.- College education (*E)* vs benchmarked assortative mating m^* (right panel) and college education (*E)* vs. benchmarked polarization P^* (left panel)



Authors' calculations using 408 observations from 120 countries.

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 this 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 finding 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.- Counterfactual decompositions of change over time in assortative mating and polarization

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 provide a rough idea of the relationship between couples of variables but fail to provide an accurate assessment of the dynamics of *change over time*. Indeed, the relationships that are shown in Figures 2 to 5 are not informative with regard to 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 do not report 'true' changes over time but rather indicate how different our parameters of interest would look under alternative homogamy assumptions. 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^{α}. Because these contributions add up to 100% (in absolute value), they can easily be represented via ternary plots[10]. In each of these plots, there are 289 'dots' whose relative positions represent the percent contribution (in

[10] A 'ternary plot', 'ternary graph', 'triangle plot' or 'simplex plot' is a barycentric plot on three variables that add up to a certain constant. It graphically depicts the ratios of the three variables as positions in an equilateral triangle.

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[11]. When the dots are represented with a '+' sign, the corresponding parameter of interest has increased over time, whereas the opposite holds when the dots are 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^{α} 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 (approximately 70% of the times in our coupled observations dataset).

Figure 8.- Percentage 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.

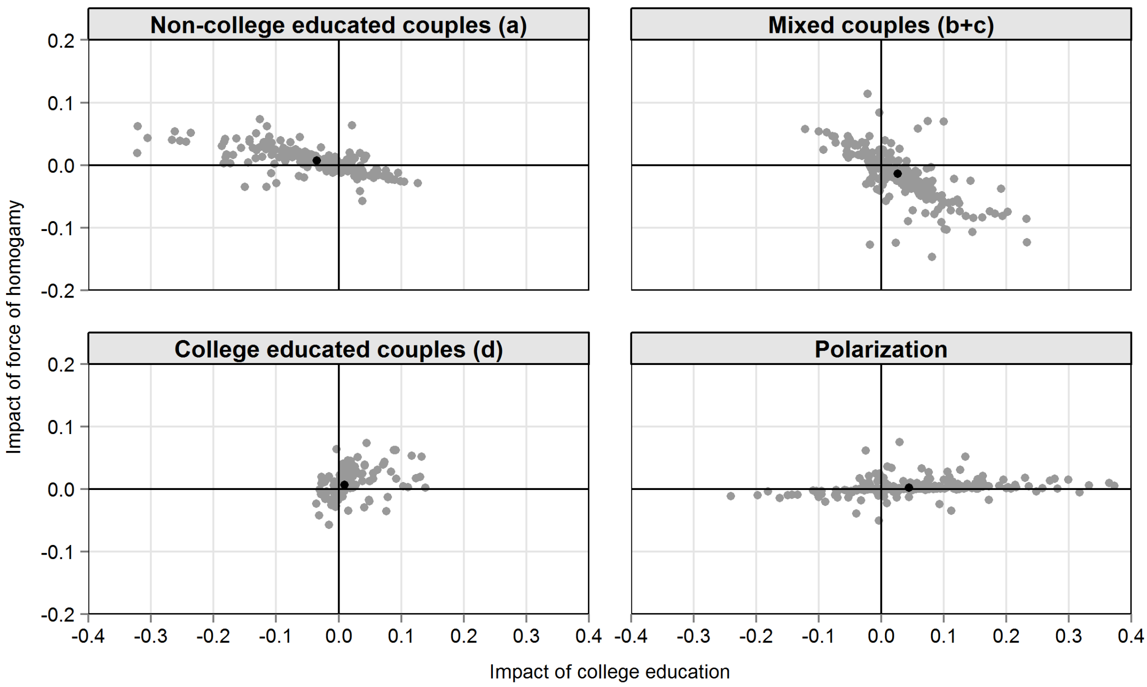
[11] Intuitively, whenever a dot is close to the vertex 'E' (for instance), it means that the change over time in the corresponding parameter has been mainly caused by the education expansion effect.

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^{α} are 1.3%, 2.6%, 3.1% and 1.2%, respectively. In contrast, the contribution of the force of homogamy to changes over time in the shares of the education distribution are more notable: for the changes in *a*, *b*+*c* and *d*, the contributions of '*H*' are 23.2%, 39.4% and 68.6%. However, with regard to assessing the impact of the force of homogamy on changes in polarization, the contribution falls to a mere 9.6%. Finally, the percent contributions of education expansion to the changes in *a*, *b*+*c*, *d* and P^{α} are 75.5%, 58%, 28.3% and 89.2%, respectively. Therefore, in addition to the share of college-educated couples, 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.

In Figure 8, we are unable to infer whether each of the forces *E*, *H* and *G* contribute separately to increase or decrease our parameters of interest over time (i.e., the 'overall' increases / decreases of *a*, *b*+*c*, *d* and P^{α} 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 by 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 panels in Figure 9 jointly plot the values of ($\Delta_E a$, $\Delta_H a$), ($\Delta_E(b+c)$, $\Delta_H(b+c)$), ($\Delta_E d$, $\Delta_H d$) and ($\Delta_E P^{\alpha}$, $\Delta_H P^{\alpha}$). In each panel, the values of Δ_E and Δ_H are on the horizontal and vertical axes, respectively, and the corresponding averages are highlighted in the graph.

As shown in the first panel of Figure 9, the influence of education expansion has tended to decrease the share of non-college-educated couples, whereas 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 $\overline{\Delta_E a} = -0.035$ and $\overline{\Delta_H a} = 0.007$ respectively), resulting in a net decrease of that particular group in the education distribution. 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 this share. As shown 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 $\overline{\Delta_E(b+c)} = 0.026$ and $\overline{\Delta_H(b+c)} = -0.014$ respectively). This is why the shares of mixed couples have tended to increase over time rather than decreasing. In contrast, both education expansion and the force of homogamy have influenced in the increase in college-educated couples. As shown 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 increasing that share by $\overline{\Delta_E d} = 0.009$ and $\overline{\Delta_H d} = 0.007$ points, respectively). Lastly, most changes in polarization are driven by education expansion rather than the force of homogamy. Assortative mating plays a very small 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 impacts of education expansion and assortative mating are $\overline{\Delta_E P^\alpha} = 0.044$ and $\overline{\Delta_H P^\alpha} = 0.002$, respectively.

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.

5.4.- Robustness checks

The results presented so far are based on a concrete specification of our polarization index P^{α} (i.e., when α=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* ($\overline{\Delta_E P^{\alpha}}, \overline{\Delta_H P^{\alpha}}, \overline{\Delta_G P^{\alpha}}$) 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 that drives changes over time in observed polarization levels, with the force of homogamy playing a secondary role.

Table 1.- Robustness checks using alternative specifications of the polarization index P^{α}. Average changes in polarization over time attributable to *E*, *H* and *G* (rows 1 – 3) and the corresponding average percent contributions (rows 4 – 6)

Alpha Values	0.1	0.5	1	2	3	4	5
$\overline{\Delta_E P^{\alpha}}$	0.006	0.024	0.036	0.044	0.045	0.044	0.042
$\overline{\Delta_H P^{\alpha}}$	0.000	0.000	0.001	0.002	0.003	0.005	0.005
$\overline{\Delta_G P^{\alpha}}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000
% contribution of *E*	91.1%	92.4%	95%	89.2%	85.4%	82%	79.1%
% contribution of *H*	7.9%	6.6%	4.2%	9.6%	13.2%	16.4%	19.3%
% contribution of *G*	1.1%	1%	0.8%	1.2%	1.4%	1.5%	1.6%

Authors' calculations using 289 observations from 94 countries.

Experimentation with other social polarization indices, such as the RQ polarization index proposed by Montalvo and Reynal-Querol (2005), lead to very similar results that 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 methodology to disentangle and gauge the intertwined effects of three key social forces that affect the levels of polarization in the marriage market between college-educated and non-college-educated populations. These forces are homogamy, the expansion of college education and the gender gap in education. We have applied this methodology to data from 120 countries and more than 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-educated and non-college-educated 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, 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 decreased by an average of only 6.5% of their observed values. This finding shows a rather modest contribution of homogamy to polarization, which would have not been substantially different even if we 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 contributions of the force of homogamy to the changes over time in non-college-educated, mixed and college-educated couples and polarization levels are 23.2%, 39.4%, 68.6% and 9.6%, respectively, whereas the corresponding contributions of education expansion are 75.5%, 58%, 28.3% and 89.2%, respectively. In sum, the polarizing trends we observe worldwide are not overly influenced by the tendency toward homogamy but seem to be rather mechanically driven by the share of the population with a college education, which continues to expand in most societies. Inter alia, our results suggest that the gloomy scenarios of a growing social divide through the gradual disappearance of mixed couples in favor of their homogamous counterparts are simply not occurring.

Are the results presented in this paper "unavoidable" given the bounded nature of our variables and the parsimony of our models? When no one is educated and when everyone is educated, there is no variability, so polarization is zero. In the process of education expansion, when some population groups receive 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 are the following conclusions obtained from our analysis: (i) despite its purported importance and the widespread attention it has attracted, 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 rather marginal so far; (iii) although the global expansion of college education has been the main driving force behind current polarizing trends, it has also greatly contributed to increasing the share of mixed couples over time. Even if our findings are undoubtedly influenced by the admitted parsimony of the underlying models, the latter are sufficient to (a) neatly analyze the contribution of the key social forces that drive changes in polarization and (b) provide 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 a college education, but not because of homogamy patterns, which play a minor role. If trends were to continue and the college educated were to become a majority, polarization levels would mechanically go down. A similar effect may have already occurred with basic literacy skills (Permanyer et al. 2013). The literate population was initially very scarce, and then it began to grow; now it is almost universal in many places. Was there similar concern about increasing education polarization when literacy began to become widespread, or it was celebrated as a measure of social progress?

References:

ATKINSON, A.; PIKETTY, T.; SAEZ, E. (2011) "Top incomes in the long run of history". *Journal of Economic Literature,* 49 (1): 3-71.

BARRO, R.; LEE, J.-W. (2000) "International data on educational attainment: Updates and implications". *Center for International Development Working Paper,* 42. Cambridge, MA: Center for International Development.

BENAVOT, A.; RIDDLE, P. (1988) "The Expansion of Primary Education, 1870–1940: Trends and Issues". *Sociology of Education,* 61 (3): 191-210.

BENAVOT, A.; CHA, Y.-K.; KAMENS, D.; MEYER, J.; WONG, S.-Y. (1991) "Knowledge for the Masses: World Models and National Curricula, 1920-1986". *American Sociological Review,* 56 (1): 85-100.

BLOSSFELD, H.-P. (2009) "Educational assortative marriage in comparative perspective". *Annual Review of Sociology,* 35: 513-530.

BLOSSFELD, H.-P.; TIMM, A. (Eds.) (2003) *Who marries whom? Educational systems as marriage markets in modern societies*. Dordrecht: Kluwer Academic.

BRADLEY, K. (2001) "The Incorporation of Women into Higher Education: Paradoxical Outcomes?". *Sociology of Education,* 73 (1): 1-18.

BREEN, R.; SALAZAR, L. (2011) "Educational assortative mating and earnings inequality in the United States". *American Journal of Sociology,* 117 (3): 808-843.

CHARLES, M.; GRUSKY, D. (2004) *Occupational Ghettos: The Worldwide Segregation of Women and Men.* Stanford University Press.

COHEN, D.; SOTO, M. (2007) "Growth and human capital: good data, good results". *Journal of Economic Growth,* 12: 51-76.

CRAFTS, N. (2002) "The Human Development Index, 1870-1999: Some revised estimates". *European Review of Economic History,* 6 (3): 395-405.

DORIUS, S.; FIREBAUGH, G. (2010) "Trends in global gender inequality". *Social Forces,* 88 (5): 1941-1968.

DORIUS, S. (2013) "The rise and fall of worldwide education inequality from 1870 to 2010: Measurement and trends". *Sociology of Education,* 86 (2): 158-173.

ESTEBAN, J.; RAY, D. (1994) "On the Measurement of Polarization". *Econometrica,* 62: 819-851.

ESTEVE, A.; GARCIA, J.; PERMANYER, I. (2012) "The gender-gap reversal in education and its effect in union formation: The end of hypergamy?". *Population and Development Review,* 38 (3): 535-546.

ESPING-ANDERSEN, G. (2009) *The incomplete revolution: Adapting welfare states to women's new roles. Polity.*

ESTEVE, A.; SOBEK, M. (2003) "Challenges and methods of international census harmonization". *Historical Methods: A Journal of Quantitative and Interdisciplinary History,* 36 (2): 66-79.

GRANT, M.; BEHRMAN, J. (2010) "Gender gaps in educational attainment in less developed countries". *Population and Development Review,* 36 (1): 71-89.

KALMIJN, M. (1998) “Intermarriage and homogamy: Causes, patterns, trends.” *Annual review of sociology,* 395-421.

KC, S. et al. (2010) “Projection of population by level of educational attainment, age, and sex for 120 countries for 2005-2050”. *Demographic Research,* 22: 383-472.

LUTZ, W.; KC, S. (2011) “Global human capital: Integrating education and population”. *Science,* 333: 587-592.

LUTZ, W.; BUTZ, W.; KC, S. (Eds.) (2014) *World Population and Human Capital in the Twenty-First Century.* Oxford: Oxford University Press.

MARE, R. (1991) “Five decades of educational assortative mating”. *American Sociological Review,* 56: 15-32.

MEYER, J.; RAMIREZ, F.; RUBINSON, R.; BOLI-BENNETT, J. (1977) “The world education revolution, 1950-1970”. *Sociology of Education,* 4: 242-258.

MEYER, J.; RAMIREZ, F.; SOYSAL, Y. (1992) “World Expansion of Mass Education, 1870-1980”. *Sociology of Education,* 65 (2): 128-149.

MINNESOTA POPULATION CENTER (2014) *Integrated Public Use Microdata Series, International: Version 6.3 [Machine-readable database].* Minneapolis: University of Minnesota.

MONTALVO, J.; REYNAL-QUEROL, M. (2005) “Ethnic polarization, potential conflict and civil war”. *American Economic Review*, 95 (3): 796-816.

MORRISON, C.; MURTIN, F. (2009). “The century of education”. *Journal of Human Capital,* 3 (1): 1-42.

PERMANYER, I.; GARCÍA, J.; ESTEVE, A. (2013) “The impact of educational homogamy on isolated illiteracy levels”. *Demography,* 50 (6): 2209-2225.

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.

RAMIREZ, F.; Meyer, J. (1980) “The Social Construction of the Modern World System”. *Annual Review of Sociology,* 6: 369-99.

SCHWARTZ, C.; MARE, R. (2012) “The Proximate Determinants of Educational Homogamy: The Effects of First Marriage, Marital Dissolution, Remarriage, and Educational Upgrading”. *Demography*, 49: 629-650. PMID: 22450676.

SCHWARTZ, C. (2013) “Trends and variations in assortative mating: causes and consequences”. *Annual Review of Sociology,* 39: 451-470.

SCHWARTZ, C.; MARE, R. (2005) “Trends in educational assortative mating from 1940 to 2003”. *Demography,* 49: 629-650.

SMITS, J.; ULTEE, W.; LAMMERS, J. (2000) “More or less educational homogamy? A test of different versions of modernization theory using cross-temporal evidence for 60 countries”. *American Sociological Review,* 65: 781-788.

Appendix:

In this appendix, we show (i) how to obtain the hypothetical education distributions (a_0,b_0,c_0,d_0) and ($a_{max}, b_{max}, c_{max}, d_{max}$) 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* ($a_{max}, b_{max}, c_{max}, d_{max}$).

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 if individuals did not care about their partners' education (i.e., if couples were formed purely at random) while keeping the marginal education distribution of women and men unchanged. It is well known that under such an independence assumption, one has

$$a_0 = (a+b)(a+c); b_0 = (a+b)(b+d); c_0 = (c+d)(a+c); d_0 = (c+d)(b+d) \qquad [A1].$$

Because these are the expected frequencies that would be observed if 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

$$a_p = a - a_0; b_p = b - b_0; c_p = c - c_0; d_p = d - d_0 \qquad [A2].$$

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

$$\begin{aligned} a_p &= d_p = ad - bc \qquad [A3] \\ b_p &= c_p = bc - ad \qquad [A4] \end{aligned}.$$

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

$$\left.\begin{aligned} a &= a_0 + H \\ b &= b_0 - H \\ c &= c_0 - H \\ d &= d_0 + H \end{aligned}\right\} \qquad [A5].$$

Equation [A5] shows a decomposition of observed cell frequencies as a sum of frequencies that would be observed if education status were 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 attempt to answer the following question: to what extent would 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:

	Non-college Woman	College Woman	Total
Non-college Man	*a+Min{b, c}*	*b-Min{b, c}*	*a+b*
College Man	*c-Min{b, c}*	*d+Min{b, c}*	*c+d*
Total	*a+c*	*b+d*	*1*

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

$$\left.\begin{aligned} a_{\max} &= a + \min\{b,c\} \\ b_{\max} &= b - \min\{b,c\} \\ c_{\max} &= c - \min\{b,c\} \\ d_{\max} &= d + \min\{b,c\} \end{aligned}\right\} \qquad [A6].$$

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

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

(*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:

$$\left.\begin{aligned} a+b &= 1-E_m \\ c+d &= E_m \\ a+c &= 1-E_w \\ b+d &= E_w \\ ad-bc &= H \end{aligned}\right\} \qquad [A8].$$

Solving [A8], we obtain

$$\left.\begin{aligned} a &= (1-E_w)(1-E_m)+H \\ b &= E_w(1-E_m)-H \\ c &= E_m(1-E_w)-H \\ d &= E_wE_m+H \end{aligned}\right\} \qquad [A9].$$

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

$$\left.\begin{aligned} \tilde{a} &= (1-E)^2+H \\ \tilde{b} &= E(1-E)-H \\ \tilde{c} &= E(1-E)-H \\ \tilde{d} &= E^2+H \end{aligned}\right\} \qquad [A10].$$

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:

$$\left.\begin{aligned} a_g+b_g &= a+b \\ c_g+d_g &= c+d \\ a_g+c_g &= a+c \\ b_g+d_g &= b+d \\ a_gd_g-b_gc_g &= ad-bc \\ b_g &= c_g \end{aligned}\right\} \qquad [A11].$$

Solving [A11], we obtain:

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

Step 3. Because [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

$$\left.\begin{aligned} (1-E)^2 + H &= a + \left(\frac{G}{2}\right)^2 \\ E(1-E) - H &= \frac{b+c}{2} - \left(\frac{G}{2}\right)^2 \\ E^2 + H &= d + \left(\frac{G}{2}\right)^2 \end{aligned}\right\} \qquad [A13].$$

Lastly, the identities in [7] obtain after basic algebraic manipulations of [A13].