PAA2011

(Work in progress)

The reversal of the gender gap in education and its impact on union formation: the end of hypergamy[1]

Albert Esteve (aesteve@ced.uab.es)
Joan Garcia (jgarcia@ced.uab.es)
Iñaki Permanyer (inaki.permanyer@uab.es)

Centre d'Estudis Demogràfics (www.ced.uab.es)

The educational expansion that most countries in the world have witnessed in recent decades has been accompanied by a reduction of the gender gap in educational attainment. The consequences of this unprecedented phenomenon have yet to be explored in many of the dimensions of social life. In this paper we examine the impact that the advances in women's education has had on gender symmetry in union formation and, more specifically, on female educational hypergamy (women's tendency to marry men with a higher educational attainment than themselves). We use newly integrated IPUMS census micro-data from 103 samples and taken in 38 countries. Results from multilevel linear regression models show that female educational hypergamy is lower in societies with a lower gender gap in education and that where the gender gap reverses female hypogamy becomes the norm. Thus, if current trends in education are to continue, the pervasiveness of hypergamy will tend to disappear.

---

[1] This research project is funded by the European Research Council Starting Grant *WORLDFAM ERC-2009-StG-240978.*

Introduction

In recent decades the world has witnessed an impressive expansion in education. This expansion has been accompanied by a significant decrease of the gender gap in education (Hausmann et al. 2009; UNESCO, 2007; Grant & Behrman 2010; Dorius & Firebaugh 2010). In some high and middle income countries, younger cohorts of women are attaining higher levels of education than men, and if current trends are to continue, this will happen in other countries very soon. The consequences of this unprecedented phenomenon have yet to be explored in the many dimensions of social life. Among others, it may have large implications for union formation and might even change the meaning of education and the mechanisms by which education influences union formation: Will the generalization of higher education among women change the relationship between marriage and education? How will this affect gender symmetry in union formation? Will the traditional tendency of women to marry men with a higher education than themselves persist given the reversal of the gender gap in education?

We explore the impact the advances in women's education has had on gender symmetry in union formation and, more specifically, on female educational hypergamy (unions in which women have a lower level of educational attainment than their partners). We use newly harmonized census micro-data samples from the IPUMS international database. Our analysis is based on 103 census samples from 38 countries covering all major regions of the world from the 1970s to the 2000s. We use multilevel linear regression to model female hypergamy as a function of the gender gap in education. The model includes two additional variables that may mediate the relationship: the marriage gradient and the social preference for hypergamy. The first takes into account the fact

that the propensity to marry may vary by educational attainment. The second provides a net measure of hypergamy, assuming no constraints in the marriage market owing tothe number of eligible partners by educational attainment exist.

Gender symmetry in union formation

The pervasive differences between the roles that women and men have in society as a whole have always been reflected in the different degrees of gender asymmetry in union formation (both formal and informal). This has been present both in arranged and free-love marriage societies (Coontz 2005; Goody 1983). While in arranged marriages symmetry is the rule in certain dimensions (as for example ethnic group, race, caste, religion), the pre-conditions to marriage for women and men are asymmetrical in many instances, i.e. in the ancient custom of dowry. In the transition from arranged to romantic (free-choice) unions, gender asymmetry in union formation evolved and achieved individual characteristics that gradually took precedence over ascribed ones in the marriage market. In this context, education became a major structuring dimension in modern marriage markets (Blossfeld & Timm 2003; Kalmijn 1998).

Education affects union formation in many different ways. Beyond its impact on the timing and quantum of marriage, we must consider its influence on partner choice, which leads to specific patterns of assortative mating. Following Kalmijn (1998), partner choice is driven by three major groups of forces: individual preferences, third party influences, and the structural opportunities of the marriage market. As the pressure of third parties lessens, individual preferences and the context of opportunities grow to become the major forces driving assortative mating. Both are responsible for the

widespread strength of educational homogamy over time. On the one hand, if there is an extension of the years of schooling, the power of educational systems as marriage markets increases. On the other hand, individuals tend to mate in an assortative manner (McPherson et al. 2001, Mare, 1991; Schwartz and Mare 2005; Esteve and López 2010; Smits and Park 2010; Smits et al. 1998).

Despite educational homogamy being the rule, interestingly, among heterogamous couples hypergamy prevails (women marry men that are more educated than themselves). For many years and from a high-income country perspective, this was consistent with the Beckerian breadwinner model according to which men specialized in production and women in reproduction (see Becker, 1981). Hence, highly educated women are less attractive marriage partners but also benefit less from marriage as they have less need of the husband's financial support. This leads to the economic independence hypothesis which assumes that high-educated women are less likely to marry (Blossfeld 2009; Kaufman & Goldscheider 2007).

The unprecedented increases in education and female labor participation are challenging the consistency of the breadwinner model in modern societies (Oppenheimer 1994). According to Oppenheimer, the economic value of a union lies in the pooling of resources, rather than in specialization. However, given women's tendency to marry men older than themselves, women with high incomes will have to face the opportunity cost of remaining single because later matches may not be as suitable as the ones refused at an earlier time (Oppenheimer 1988). On the contrary, men with high income would stand a high chance of marrying.

Despite differences in views regarding the economic value of marriage between Becker and Oppenheimer, both predict that the chance of marriage will increase with socioeconomic status for men but decrease for women. This pattern is consistent with the existence of female educational hypergamy. Both theories were originated in the United States and soon spread to many other high-income countries. Theoretical developments of these theories from low and middle income countries are rather scarce (Mensch 2005). This may give a good reason to explain why the gender gap in access to and duration within the educational systems did not play a central role in any of these formulations. By zooming out of high-income countries and examining change within countries over time, we adopt a broader perspective to assess to what extent the gender gap on education accounts for most of the observed variations in educational hypergamy.

Hypotheses

The overarching hypothesis of this research is that female educational hypergamy is mostly influenced by the structural constraints of the marriage market. These, in turn, are a consequence of women and men's access to the education system and their respective performance therein. The reason being that the enrolment, progress and graduation of women and men in the education system impose important restrictions on the marriage market that will greatly influence the observed levels of hypergamy. Other dimensions may mediate the relationship although we expect they have less impact. These are the educational gradient in union formation and the social preference for hypergamy. More specifically, we expect to observe the following relationships between variables.

H_1: Advances in women's education will erode the pervasiveness of the female educational hypergamy pattern. Gender roles at the macro society level are mirrored at the micro individual level in the process of union formation. If female education is highly valued in a given society, women's education will be appreciated in the marriage market. Therefore we do not hypothesize to observe hypergamy in educationally advanced societies with no gender gap in education.

H_2: If highly educated women are more prone to marry, the levels of educational hypergamy will decline. Differences in the tendency to marry by educational attainment affect the pool of eligible partners. If women's tendency to marry decreases by educational attainment, the gender-gap in education becomes accentuated as for what the marriage market is concerned. Thus, societies in which the marriage gradient is less steep will present lower levels of hypergamy.

H_3: Net hypergamy will be positively related to crude hypergamy. If we define crude hypergamy as the observed percentage of couples that are educationally hypergamous and net hypergamy as measuring the odds of a union of being hypergamous net of the constraints of the marriage market (thus indicating a *social preference* for hypergamy), we hypothesize that net hypergamy will be positively related to crude hypergamy. Since we assume that gender roles at the macro society level are mirrored at the individual level, we expect the net preference for hypergamy to play a minor role in explaining the total variation of the observed hypergamy across countries. Most of the norms and values that make women and men more or less attractive in the marriage market are

already embedded in the mechanisms that make them enroll and persist through the educational system.

Data and methods

IPUMSi samples

We use the newly harmonized IPUMS micro-level data from the Minnesota Population Center (Minnesota Population Center 2010). The dataset used in this paper contains 103 samples (they are the basic units of our analysis) from 38 countries (some countries have several observations) taken from the 1970s to the 2000s Census Rounds (see the Appendix 2 for more details on the countries included in the dataset). All possible IPUMS samples have been included in the dataset except for those cases in which information was not organized at the household level or when the variable "Educational Attainment" (on which more will be said later) was not available.

Within each sample we have selected a random subsample of at the most 200000 households (following the approach taken in Ruggles 2008). These subsamples are weighted but not expanded[2], so that the weighted members of the subsample do not add up to the whole population. For each sample we have selected individuals between 25 and 34 years old (on this choice more will be said below). Between these ages, virtually all individuals have reached their highest educational level.

[2] In STATA we have used the command "*Analytical Weight*".

In order to compute the dependent and independent variables in our model we need the following basic variables: age, sex, educational attainment, marital status (MARST) and spouse location (SPLOC). The educational attainment variable has been created and harmonized by IPUMS according to the following four categories: "Less than Primary", "Primary completed", "Secondary completed" and "University level". While being somewhat crude, this partitioning allows for a great comparability across different countries that typically have a wide variety of educational systems (Esteve & Sobek 2003). MARST simply identifies the marital status of the different household members. SPLOC identifies the position occupied by the spouse/partner within the list of household members. This variable does not distinguish between formal marriages and consensual unions, so in this paper no distinction will be made between them either.

Variables

The dependent variable we use is hypergamy in education. A couple is called hypergamic (hypogamic) if the educational attainment level of the female (male) is lower than the educational attainment level of her (his) spouse. Hence, hypergamy in education for a given sample will be defined as $H=ln(A/B)$, where A (resp. B) is the number of hypergamic (resp. hypogamic) couples in the sample and *ln* denotes the natural log. This way, $H=0$ whenever the number of hypergamic couples equals the number of hypogamic couples and $H<0$ (resp. $0<H$) when hypogamic (resp. hypergamic) couples outnumber hypergamic (resp. hypogamic) ones. By defining H as

the log of *A/B* we guarantee that our theoretical distribution is symmetrically distributed around 0[3].

When defining hypergamy levels, the choice of the spouses/partners' range of ages that will be taken into account is not exempt from certain difficulties. If the range of ages is too short it might not include the whole set of relevant couples. If it is too long, marital dissolution or mortality processes may modify somehow the set of couples that should be taken into account. In order to avoid specification problems we have performed some sensitivity analysis picking different age ranges (25-34, 25-40): the results we have obtained are basically the same, so they will not be reported here (they are available upon request). The results shown in this paper are based on the range of ages between 25 and 34.

In order to control the levels of hypergamy in education we will introduce three independent variables. First, the *gender gap in education*: for each sample it is defined as

$$G := \ln\left(\frac{\left(\frac{W_3}{(W_3 + M_3)}\right)}{\left(\frac{W_0}{(W_0 + M_0)}\right)}\right) \qquad (1)$$

[3] While *A/B* and *B/A* are *not* symmetrically distributed around *1*, ln(*A/B*) and ln(*B/A*) *are* symmetrically distributed around *0*.

where W_3 (M_3) is the number of women (men) in the sample with the highest educational attainment (*i.e*: university level) and W_0 (M_0) is the number of women (men) with the lowest educational attainment (*i.e*: less than primary completed). G is thus the log of the ratio between the probability of finding a woman among the most highly educated and the probability of finding a woman among the least educated. Whenever $G=0$, women are equally represented among the most and the least educated. When $G<0$ (resp. $G>0$) women are less (resp. more) represented among the highly educated. For this reason $G<0$ represents the "classical" status quo situation in which the educational attainment level of men is, on average, higher than that of women and $G>0$ reflects those incipient cases in which a reversal has occurred. Again, we are working with logs to make sure that our distribution of gender gaps is symmetrically distributed around 0. It must be pointed out that whenever any of the W_0, W_3, M_0, M_3 represents less than 5% of the members of the sample we have merged those individuals with the ones of the contiguous category (*i.e*: those of W_0 (and M_0) with those of W_1 (and M_1) and those of W_3 (and M_3) with those of W_2 (and M_2)). It is worth mentioning that the proportion of women in the intermediate education categories ($W_1/(W_1+M_1)$, $W_2/(W_2+M_2)$) play no role in the definition of the gender gap. Sensitivity analysis reveals that the inclusion of these proportions in alternative definitions of the gender gap produces quantitatively and qualitatively analogous results.

Secondly we introduce the marriage gradient variable. It is defined as follows

$$M = \ln\left(\frac{\left(\frac{UW_3}{SW_3}\right)}{\left(\frac{UW_0}{SW_0}\right)}\right) \qquad (2)$$

where UW_3 (resp. UW_0) is the number of women with highest (resp. lowest) educational level that have ever been in union and SW_3 (resp. SW_0) is the number of women with highest (resp. lowest) educational level that are single. The marriage gradient compares the propensity to be in union among the highly educated women with the propensity to be in union among the least educated women. $M=0$ indicates that these propensities are the same, while $M<0$ ($M>0$) indicates that the propensity to marry is higher among the least (most) educated. As before, the log makes our distribution symmetrical around 0 and whenever one category is not very representative (less than 5% of the sample) it is fused with the contiguous category. In order to construct this index we have selected individuals between 25 and 34 years old. However, in order to discard any bias that might be caused by the fact that the highest educational levels are achieved at higher ages and therefore one might suspect the existence of "late arrivers", we have performed various sensitivity analyses. The marriage gradient index has also been computed using couples between 30 and 34 in the nominator and couples between 25 and 29 in the denominator. The results we have obtained are basically the same so they will not be shown here: they are available upon request for the interested reader.

Thirdly, we introduce a net indicator of hypergamy that aims to control the fact that in some cases, the marginal distributions in education for women and men are not the same, thus conditioning the observed hypergamy/hypogamy patterns. In this way, an index of net hypergamy can be conceived as an indicator of the social preference for

hypergamy that is not determined by the education distribution of the marriage market. We rely on log-linear net estimates of hypergamy that are constructed as follows. First, we depart from the gender symmetry (S) model that assumes that the probability of a marriage between two individuals of unequal schooling with the same pairing does not depend on whether it is the husband or wife who has more schooling. Second, we test the existence of a uniform tendency among men or women to marry up or down, we estimate a female hypergamy parameter, following Mare's example (1991). The net measure of hypergamy will be denoted as *A (see Appendix 1 for model specification)*.

Hypergamy model

Based on the previous theoretical discussion and on the variables presented so far, in this subsection we present the different multilevel models used to explain hypergamy that will be explored in this paper. Multilevel models are particularly appropriate when data have a hierarchical or clustered structure and the behavior of the units at the lower level of analysis can be influenced by the higher level unit to which they belong (Goldstein 1995). In this context, individual samples constitute the first (lower) level of analysis. The units of the second level of analysis are the countries in which the different samples have been taken. In our dataset, each country (level 2 unit denoted by *j*) has at most four samples (level 1 unit denoted by *i*) taken in different periods of time. Ignoring the hierarchical structure of the dataset will overlook the importance of group effects and may render invalid some of the traditional statistical analysis techniques used for studying data relationships. Introducing multilevel models, it will be possible to determine precisely what part of the variation of observed hypergamy is due to a variation *across* countries and what part is due to a sample variation *within* countries.

In the first multilevel model we simply explore the variability of observed hypergamy (measured by H_{ij}) when its values are allowed to vary across the different samples of our dataset. In a second model we incorporate an explanatory variable (the gender gap in education, measured by G_{ij}) to investigate its effects on observed hypergamy. According to $\mathbf{H_1}$, we expect the regression coefficient associated to G_{ij} to be negative. In a third model we incorporate two new explanatory variables: the marriage gradient (measured by M_{ij}) and our measure of net hypergamy (measured by A_{ij}). According to $\mathbf{H}_2$ and $\mathbf{H}_3$ we expect the regression coefficients associated to M_{ij} and A_{ij} to be negative and positive respectively. The equation of model 3 is:

$$\begin{aligned} H_{ij} &= \beta_{0j} + \beta_1 G_{ij} + \beta_2 M_{ij} + \beta_3 A_{ij} + e_{ij} \\ \beta_{0j} &= \beta_0 + \mu_{0j} \\ \mu_{0j} &\approx N(0, \sigma^2_{u0}) \\ e_{ij} &\approx N(0, \sigma^2_e) \end{aligned} \qquad (3)$$

where H_{ij}, G_{ij}, M_{ij} and A_{ij} are the values of hypergamy, gender-gap in education, marriage gradient and social preference for hypergamy respectively. The parameter β_{0j} is the intercept, which consists of β_0, the overall sample mean of the level of hypergamy, and μ_{0j} representing one country's specific differences from the overall mean. These differences are assumed to be Normally distributed with mean 0 and variance σ^2_{u0}. The term e_{ij} is the residual or error term. Residuals are assumed to be Normally distributed with mean 0 and variance σ^2_e. In multilevels, the variance between groups is partitioned into components corresponding to each level of the hierarchy, which in our case correspond to samples and countries.

Findings

In order to explore the relationship between the observed distribution of hypergamy (*H*) and the values of the gender gap in education (*G*) we show their values in a scatter plot (see Figure 1). As we see in Figure 1, there is a negative relationship between the gender gap and hypergamy, which means that higher levels of women's education are associated with lower values of hypergamy. Crude hypergamy is strongly tied to the gender gap in educational attainment. When the reversal of gender gap has occurred ($G > 0$), hypogamous couples outnumber the number of hypergamous ones ($H < 0$).

A quick glance at Figure 1 might lead one to think that the negative relationship between *G* and *H* might be explained by other contextual macro variables like, say, the GDP, the stage of the demographic transition, and so on. In order to rule out that possibility, we have examined the evolution in time for those countries with more than one sample. In Figure 1 we highlight the cases of India, Kenya, Brazil and the US. For all of them we observe a consistent pattern that is also observed for those countries with more than one sample in time: in the successive census rounds women's education increases and crude hypergamy decreases (see Appendix 2).

[[[Insert Figure 1 around here]]]

It is worth noting that, as the gender gap closes or even is being reversed, the variance in hypergamy reduces as well. This basically shows that there is some heteroskedasticity that needs to be explained. We expect that the variables included in the different explanatory models will account for most of this variance differences. Using a multilevel ordinary least square regression model, we examine the relationship

between the crude measure of hypergamy (dependent variable) and the different explanatory variables presented in the previous section. The results of this multilevel regression are presented in Table 2.

[[[Insert Table 2 around here]]]

Our first model is simply a benchmark model that describes the distribution of observed hypergamy (*H*) across all samples included in the dataset without introducing any explanatory variable. The results (see second column in Table 2) show that the average value of *H* equals 0.391 (thus indicating that the number of hypergamic couples outnumber hypogamic ones at an aggregate level including all samples at the same time) while the variance across countries (σ_{u0}^2) equals 0.417 and the variance across samples (σ_e^2) equals 0.163. This should *not* lead one to conclude that hypergamy in education is nowadays the rule in union formation (0.391). It is simply an average that mixes the values of *H* for different periods of time from the 1970s to the 2000s).

In our second model (see third column in Table 2) we only introduce the gender-gap (*G)* as an explanatory variable. As expected, the regression coefficient associated to *G* is negative (-1.77) (as hypothesized by $\mathbf{H_1}$) and highly significant. Interestingly, the cross-country variance and the cross-sample variance drastically reduce to σ_{u0}^2=0.089 and σ_e^2=0.036 respectively. This means that the gender gap in education accounts for almost 80% of the cross-country and within country variance in observed hypergamy[4]. Sensitivity analysis shows that *G* is by far the variable that accounts for most of the

[4] The first benchmark model with no explanatory variables had a cross-country variance of 0.417. After introducing *G* as an explanatory factor, this value equals 0.089, a reduction of about 80% of its original value. The same can be said about the cross-sample variance.

cross-country and within country variance in observed hypergamy[5]. If we refine our model by further incorporating the marriage gradient and the social preference for hypergamy as explanatory variables, the results are quantitatively and qualitatively similar (see fourth column in Table 2). When incorporating several variables into the model we have not encountered multicollinearity problems[6]. The sign of the regression coefficient for the marriage gradient (M) is negative, while the corresponding one for the social preference for hypergamy (A) is positive, as expected by $\mathbf{H_2}$ and $\mathbf{H_3}$. Moreover, the coefficients are statistically significant. Given the fact that multilevel statistical models do not allow to construct goodness of fit (R^2) coefficients, in Figure 2 we plot the observed values of crude hypergamy against the hypergamy values predicted by our third model. The scatterplot shown in Figure 2 illustrates the accuracy of model 3 to predict the values of hypergamy.

[[[Insert Figure 2 around here]]]

Discussion

In this paper we have explored the relationship between the gender gap in education and the tendency of women to marry men with more education than themselves. Using internationally comparable data we show that: i) societies with a low gender gap in

[5] When the marriage gradient (*M*) is chosen as the sole explanatory variable of the model we obtain σ^2_{u0}=0.199 and σ^2_e=0.179, and when the social preference for hypergamy (*A*) is the sole explanatory variable, we obtain σ^2_{u0}=0.283 and σ^2_e=0.107. Hence, the variation explained by these models is much lower than the one that is obtained when *G* is the sole explanatory factor (see previous footnote).

[6] The correlation coefficients between the independent variables are not very large: Corr(G,A)=-0.568, Corr(G,M)=0.495 and Corr(A,M)=-0.322. Moreover, the Variance Inflation Factors (VIFs) for these variables are extremely low: VIF(A)=1.48, VIF(M)=1.328, VIF(G)=1.756.

education tend to have lower levels of education hypergamy (this holds true at a worldwide level and for the different cross-sections of our dataset) and ii) there is a tendency, observed in all countries with several observations in time, towards a *joint* increase of women's educational levels together with a decrease in educational hypergamy. Moreover, we have shown that the reversal of the gender gap in education is associated with a reversal of the pattern of assortative mating leading to more hypogamous than hypergamous unions. We have verified the robustness of this relationship using different multilevel regression models taking into account two other dimensions that are associated with hypergamy: the marriage gradient and the social preference for hypergamy. The marriage gradient shows that in societies where highly educated women are less likely to marry (a pattern observed in most of the countries), we find higher levels of hypergamy; the logic being that the fact that higher educated women marry less reinforces the gender gap in education in these groups.

As for the social preference for hypergamy, although the effect of net hypergamy on crude hypergamy is positive and statistically significant, the contribution of this variable to the model is substantially modest, especially when the gender gap in education is considered. There is no single country in our analysis at any observed period of time in which crude and net levels of hypergamy yield different results. This result suggests that there are no contradictions either between the net and crude measures of hypergamy and therefore or between the gender gap and the net measure of (or social preference for) hypergamy. A careful look at the values provided in Appendix 2 illustrates this point. For those countries where more than one observation exists and the gender gap in education has reversed, the social preference for hypergamy has also reversed, suggesting a preference for hypogamy.

According to our results, if current trends in education are to continue the end of hypergamy is near. This unprecedented scenario demonstrates the important implications that women's education may have for the erosion of traditional patterns in assortative mating. In broader perspective, the consequences that these changes might have on the gender roles distribution within and outside the couple will need to be addressed. Will these transformations lead to more egalitarian couples and alter the economic foundations of marriage? Will gender asymmetry in assortative mating become apparent on other dimensions? In addition, future research will have to address the meaning of education in a changing context.

References

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

Blossfeld, H.-P., & Timm, A. (2003). *Who MarriesWhom? : Educational Systems as Marriage Markets in Modern Societies* (Vol. 12). Dordrecht: Kluwer Academic.

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

Buchmann, C., & Hannum, E. (2001) Education and Stratification in Developing Countries: A Review of Theories and Research. *Annual Review of Sociology, 27*, 77-102.

Coontz, S. (2005). *Marriage, a History: From Obedience to Intimacy, or How Love Conquered Marriage*. New York: Viking.

Dorius, S.F., & Firebaugh, G. (2010) Trend in Global Gender Inequality. *Social Forces*, 88(5): 1941-1968.

Esteve, A., & Sobek, M. (2003) Challenges and Methods of International Census Harmonization. *Historical Methods, 36*(2), 66-79.

Esteve, A., López, L., (2010) "Union Formation Implications of Race and Gender Gaps in Educational Attainment: The Case of Latin America", *Population Research and*

Policy Review, 29 (5), pp. 609-637.Goldstein, H. (1995). *Multilevel Statistical Models*. London: John Wiley & Sons.

Goody, J. (1983) *The Development of the Family and Marriage in Europe*. Cambridge: Cambridge University Press.

Grant, M. and Behrman, J. (2010) Gender Gaps in Educational Attainment in Less Developed Countries. *Population and Development Review, 36, 71-89.*

Hausmann, R., Tyson, L., Zahidi, S. (2009) The Global Gender Gap Report, World Economic Forum. Cologny/Geneva, Switzerland.

Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology, 24*, 395-421.

Kaufman, G., Goldscheider, F. (2007) Do men need a spouse more than a women?: percpetion of the importance for marriage form men and women. *The Sociological Quarterly*, 48:29-46.

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

Mason, K. O. (2001). Gender and Family Systems in the Fertility Transition. *Population and Development Review, 27*, 160-176.

McPherson, M., Smith-Lovin, L., Cook, J.M., (2001) Birds of a feather: Homophily in Social Networks. *Annual Review of Sociology*, 27: 415-444.

Mensch, B. S. (2005) "Transition to marriage" in Lloyd, C.B. (ed.) *Growing Up global. The Changing Transitions to Adulthood in Developing Countries*. Washington, D.C.: The National Academies Press, pp. 416-505.

Minnesota Population Center. (2010). Integrated Public Use Microdata Series - International: Version 6.0. Minneapolis: University of Minnesota.

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

Oppenheimer, V.K. (1994). Women's Rising Employment and the Future of the Family in Industrial Societies. *Population and Development Review*, 20,2: 293-342.

Ridgeway, C. L., & Smith-Lovin, L. (1999). The Gender System and Interaction. *Annual Review of Sociology, 25*(1), 191-216.

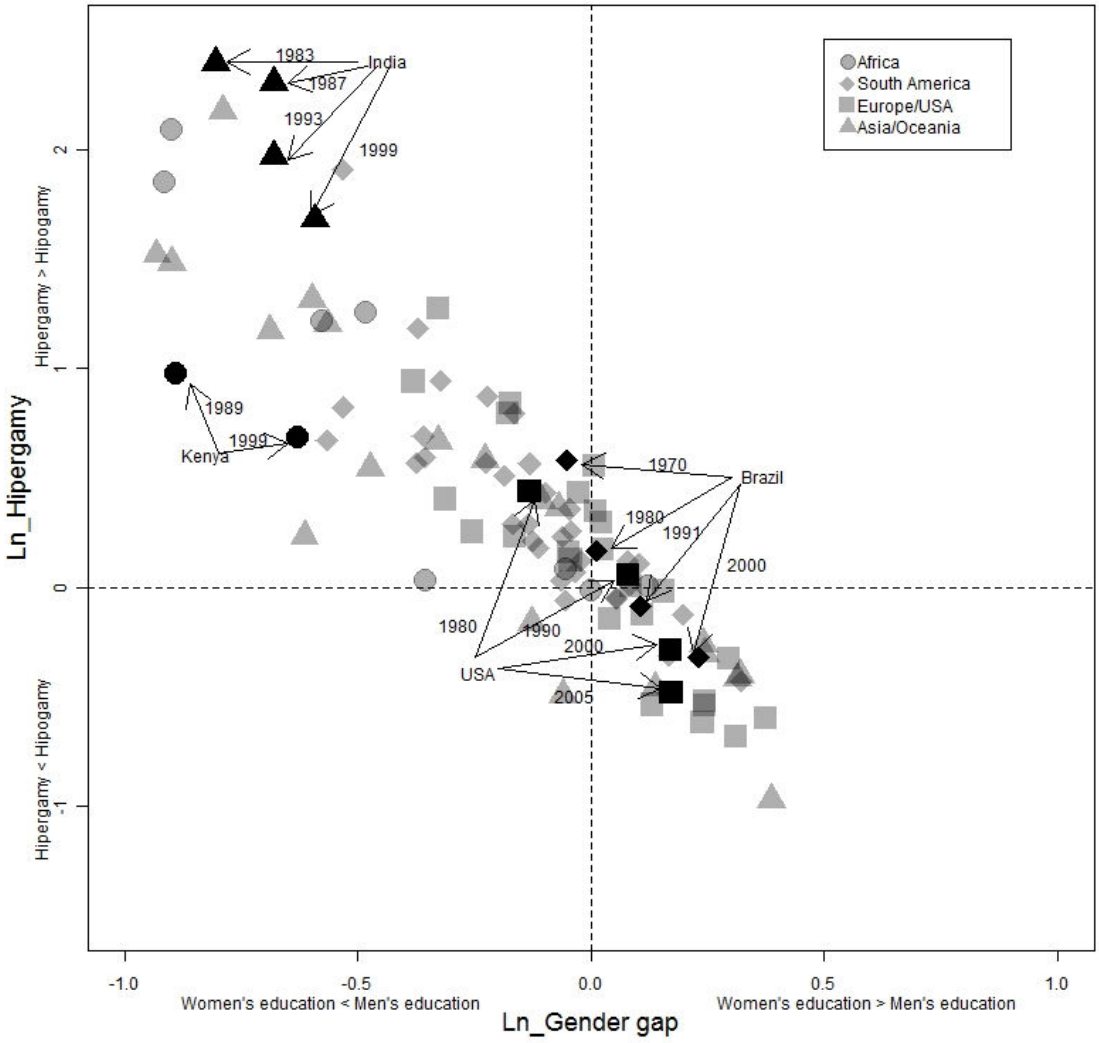
Ruggles, S., & Heggeness, M. (2008). Intergenerational Coresidence in Developing Countries. *Population and Development Review, 34*(2), 253-281.

Schwartz, C. R., & Mare, R. D. (2005). Trends in Educational Assortative Marriage from 1940 to 2003. *Demography, 42*(4), 621-646.Smits, J., Park, H. (2009). Five Decades of Educational Assortative Mating in 10 East Asian Societies. *Social Forces*, 88(1), 227-255.

Smits, J., Ultee, W., & Lammers, J. (1998). Educational Homogamy in 65 Countries: An Explanation of Differences in Openness Using Country-Level Explanatory Variables. *American Sociological Review, 63*(2), 264-285.

United Nations Educational Scientific and Cultural Organization (UNESCO). (2007). *Global Education Digest*. Montreal, Quebec: UNESCO Institute for Statistics.

Figure 1. The relationship between the gender gap in education and hypergamy (selection of countries). See Appendix 2 for country specific data.



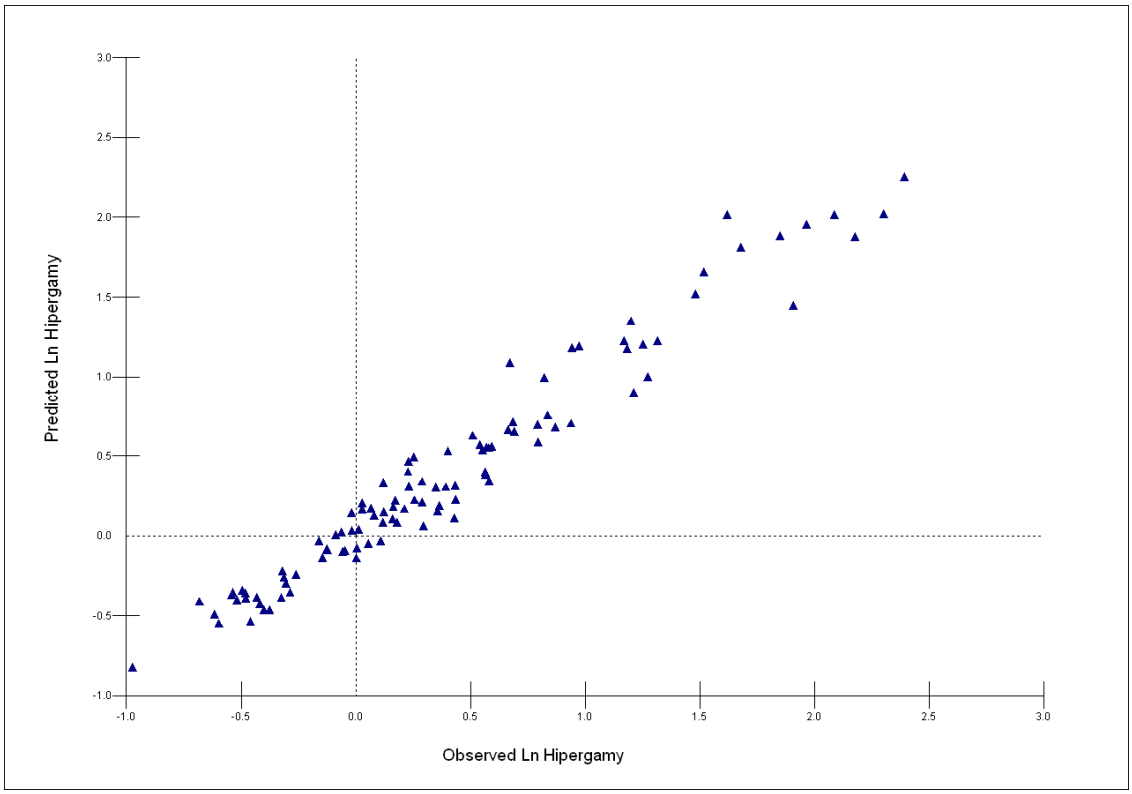
Source: Authors' calculations using IPUMS international data.

Table 1. Multilevel regression results.

	Model 1	Model 2	Model 3
Fixed effects			
Constant	0,391***	0,07	-0,086
Gender gap		-1,77***	-1,547***
Marriage gradient			-0,131**
Social preference for hypergamy			0,237***
Random intercept			
Cross-country variance σ^2_{u0}	0,417***	0,089***	0,072***
Cross-sample variance σ^2_e	0,163***	0,036***	0,034***
-2 Log Likelihood	180.414	22.826	11.941

Source: Authors' calculations using IPUMS data.

Figure 2. Observed hypergamy and expected values of hypergamy predicted by model 3.



Source: Authors' calculations using IPUMS International data.

Appendix 1. Log-linear model specification to estimate social preference for hypergamy (A)

Recall we are considering four education categories {1,2,3,4} corresponding to "Less than Primary", "Primary completed", "Secondary completed" and "University level". The education level of women (men) will be denoted by E_w(E_m). The net measure of hypergamy is obtained from the following log-linear model:

$$\log(f_{ij}) = \beta + \beta_i^H d_i^H + \beta_j^W d_j^W + \sum_k \beta_k^S d_k^S + \sum_l \beta_l^A d_l^A$$

where f_{ij} is the expected number of marriages between husbands in education category 'i' and wives in education category 'j'. Moreover,

$$d_i^H = \begin{cases} 1 \; if \; husband \; is \; in \; education \; category \; 'i' \\ 0 \; otherwise \end{cases}$$

$$d_j^W = \begin{cases} 1 \; if \; wife \; is \; in \; education \; category \; 'j' \\ 0 \; otherwise \end{cases}$$

$$d_1^S = if \{E_W = 1 \; and \; E_M = 2\} \; or \; \{E_W = 2 \; and \; E_M = 1\}$$

$$d_2^S = if \{E_W = 1 \; and \; E_M = 3\} \; or \; \{E_W = 3 \; and \; E_M = 1\}$$

$$d_3^S = if \{E_W = 1 \; and \; E_M = 4\} \; or \; \{E_W = 4 \; and \; E_M = 1\}$$

$$d_4^S = if \{E_W = 2 \; and \; E_M = 3\} \; or \; \{E_W = 3 \; and \; E_M = 2\}$$

$$d_5^S = if \{E_W = 2 \; and \; E_M = 4\} \; or \; \{E_W = 4 \; and \; E_M = 2\}$$

$$d_6^S = if \{E_W = 3 \; and \; E_M = 4\} \; or \; \{E_W = 4 \; and \; E_M = 3\}$$

$$d_i^A = \begin{cases} 1 \; if \; E_W < E_M \\ 0 \; otherwise \end{cases}$$

In this model, $\beta_i^H d_i^H + \beta_j^W d_j^W$ accounts for the marginal distributions. On the other hand, $\sum_k \beta_k^S d_k^S$ accounts for the symmetric part of the model that that assumes that the probability of a marriage between two individuals of unequal schooling with the same pairing does not depend on whether it is the husband or wife who has more schooling. Finally, $\sum_i \beta_i^A d_i^A$ accounts for the asymmetric part of the model that tests the existence of a uniform tendency among women to marry up. Recall that in this model, the d_i^A do not depend on *l*, so all the $\beta_i^A = \beta^A$ for some β^A. Our net measure of hypergamy (A) is simply defined as β^A.

Appendix 2. Variables included in the model and sample size by country and sample

	Hipergamy	Gender gap in education	Marriage gradient	Social preference for hypergamy	Sample size. Pop 25-34
EUROPE & NORTHAMERICA					
Armenia					
2001	0.72	1.34	0.49	-0.27	42,468
France					
1968	1.49	0.73	0.51	0.43	73,761
1975	1.34	1.02	0.36	0.45	78,817
1982	1.17	0.95	0.43	0.29	86,219
1990	0.87	1.04	0.57	-0.01	75,333
1999	0.58	1.14	0.72	-0.19	69,750
Greece					
1971	2.56	0.68	0.29	0.59	86,736
1981	1.26	0.85	0.24	0.22	83,708
1991	1.13	0.96	0.17	0.15	81,960
2001	0.58	1.28	0.23	0.00	86,025
Hungary					
1970	1.19	1.03	0.57	-0.26	70,399
1980	3.57	0.72	0.51	0.65	77,761
1990	2.21	0.84	0.79	0.60	63,683
2001	0.98	1.17	0.42	0.08	65,663
Italy					
2001	0.54	1.27	0.25	0.18	79,000
Portugal					
1981	1.54	0.97	0.38	0.08	65,674
1991	0.88	1.11	0.25	-0.06	65,659
2001	0.51	1.37	0.25	-0.13	60,732
Romania					
1977	1.42	1.01	0.20	-0.48	85,190
1992	2.31	0.84	0.28	0.40	75,197
2002	1.74	1.01	0.22	0.52	97,492
Slovenia					
2002	0.55	1.45	0.32	0.29	25,731
Spain					
1991	1.00	1.09	0.29	0.23	100,868
2001	0.60	1.28	0.26	-0.18	95,908
United States					
1980	1.55	0.88	0.44	-0.11	78,710
1990	1.06	1.08	0.64	0.31	78,930
2000	0.75	1.18	0.72	-0.33	64,074
2005	0.62	1.19	0.76	-0.30	62,305
AFRICA					
Ghana					
2000	3.50	0.62	0.22	1.02	135,380
Guinea					
1983	6.36	0.40	0.07	0.62	66,138
1996	8.06	0.41	0.09	1.36	100,226
Kenya					
1989	2.65	0.41	0.18	0.91	124,158
1999	1.98	0.53	0.32	0.91	125,911

	(H) Hipergamy	(G) Gender gap in education	(M) Marriage gradient	(A) Social preference for hypergamy	Sample size, Pop 25-34
Rwanda					
2002	1.03	0.70	0.33	-0.10	103,808
South Africa					
1996	1.08	0.95	0.92	0.07	117,047
2001	0.98	1.00	0.90	0.04	121,901
2007	1.00	1.13	0.93	0.12	93,762
Uganda					
1991	5.05	0.31	0.09	0.61	121,813
2002	3.36	0.56	0.13	0.04	126,847
SOUTH AMERICA					
Argentina					
1970	1.77	0.69	0.30	0.29	65,537
1980	1.34	0.85	0.33	0.26	116,167
1991	0.94	0.95	0.28	0.10	98,332
2001	0.62	1.18	0.35	-0.09	97,000
Bolivia					
1976	6.74	0.59	0.48	0.56	61,878
1992	3.26	0.69	0.44	0.42	89,200
2001	2.57	0.73	0.20	0.33	94,445
Brazil					
1970	1.79	0.95	0.34	0.65	126,621
1980	1.18	1.01	0.34	0.41	124,720
1991	0.92	1.11	0.32	0.21	136,789
2000	0.73	1.26	0.31	0.07	120,926
Chile					
1970	1.76	0.88	0.67	0.40	114,465
1982	1.29	0.96	0.60	0.19	120,695
1992	1.01	1.09	0.70	0.31	126,282
2002	1.07	0.97	0.38	-0.23	98,625
Colombia					
1973	1.99	0.70	0.53	0.60	137,306
1985	1.33	0.88	0.26	0.38	143,538
1993	0.95	1.06	0.39	-0.01	142,268
2005	0.73	1.18	0.28	-0.16	116,383
Costa Rica					
1973	1.03	0.94	0.71	0.93	22,350
1984	1.24	0.88	0.49	0.33	37,850
2000	0.95	1.06	0.40	0.27	59,543
Ecuador					
1974	2.39	0.80	0.33	0.54	79,736
1982	2.21	0.85	0.35	0.58	109,533
1990	1.67	0.83	0.26	0.44	119,188
2001	1.13	0.98	0.30	0.34	102,093
Mexico					
1970	1.96	0.57	0.31	1.52	58,811
1990	1.81	0.70	0.19	0.37	143,042
1995	1.76	0.80	0.18	0.48	51,647
2000	1.20	0.89	0.21	0.03	139,109

	(H) Hipergamy	(G) Gender gap in education	(M) Marriage gradient	(A) Social preference for hypergamy	Sample size, Pop 25-34
Panama					
1970	1.43	0.95	0.26	-0.29	19,465
1980	1.26	0.94	0.21	0.55	28,037
1990	1.12	1.08	0.19	0.06	36,738
2000	0.88	1.22	0.17	0.03	46,466
Venezuela					
1971	2.27	0.59	0.47	0.29	97,744
1981	1.54	0.91	0.43	0.23	136,994
1990	1.11	1.11	0.38	-0.01	124,264
2001	0.65	1.38	0.42	-0.03	112,412
ASIA & OCEANIA					
Cambodia					
1998	4.39	0.41	0.35	0.68	149,520
China					
1982	8.82	0.45	0.05	0.18	136,889
1990	3.32	0.57	0.20	0.13	124,799
India					
1983	10.95	0.45	0.08	0.51	88,000
1987	9.99	0.51	0.10	0.50	97,759
1993	7.13	0.51	0.11	0.28	86,697
1999	5.36	0.55	0.09	0.15	91,250
Iraq					
1997	3.73	0.55	0.30	0.15	215,406
Israel					
1972	1.72	0.62	0.39	-0.12	39,114
1983	0.85	0.88	0.60	-0.15	61,573
1995	0.63	1.15	0.86	-0.35	75,499
Jordan					
2004	0.61	0.94	1.06	-0.01	85,699
Kyrgyz Republic					
1999	0.67	1.38	0.72	-0.26	72,631
Malaysia					
1970	4.56	0.39	0.13	0.29	21,587
1980	3.22	0.50	0.22	0.35	26,927
1991	1.94	0.72	0.26	0.42	58,230
2000	1.44	0.93	0.31	0.19	68,689
Mongolia					
1989	0.69	1.06	1.73	-0.35	29,362
2000	0.38	1.47	0.71	-0.19	42,180
Palestine					
1997	1.26	0.54	0.59	0.19	35,427
Philippines					
1990	0.77	1.27	0.20	0.19	157,209
1995	0.74	1.27	0.24	0.06	155,415
2000	0.66	1.37	0.26	0.06	148,974
Vietnam					
1989	1.79	0.80	0.70	-0.04	155,993
1999	1.48	0.90	0.79	-0.22	145,156

Source: IPUMS international