University of Limerick
Department of Sociology Working Paper Series

Working Paper WP2019-01
July 2019

Brendan Halpin

Department of Sociology
University of Limerick

Partner Choice and Parameter Estimates: Modelling the Effect of Preferences

Partner choice and parameter estimates: modelling the effect of preferences

Brendan Halpin

9th July 2019

Contents

1 Introduction

Assortative mating, the process through which people interactively choose each other as spouses, is a complex, consequential and informative phenomenon.[1] It has important consequences for the structure of society, for inequality and integration, and it tells a lot about social structures, patterns of interactions and preferences. However, it is a complex process, and while tempting it is problematic to read preferences directly from the empirical outcomes. The mapping of preferences to outcomes in a process subject to constraints is not straightforward. Outcomes are constrained in complex ways, not least the gendered distribution of the characteristics of interest, the competition among peers for alters with the desired characteristics, and the fact that alters are agents with preferences too. Additionally, the fact that the process is dynamic, where the context changes continually, not only because of external historical reasons, but also endogenously, as the distribution of single people changes as others marry.

Research on assortative mating has a long history of looking at the data, usually tables of spouse pairs classified by the characteristic of interest, and making inferences about the data generating processes in terms of homogamy preference, competition for desirable characteristics, female hypergamy preference, or relative desirability of different characteristics leading to status exchange. Earlier work tends to read directly from patterns in the tables, or simple summaries. But loglinear modelling of these tables, predominant in the literature since the 1980s, is markedly superior, by controlling for the marginal distributions and correlated characteristics. For instance, raw data may show lots of women marrying up, but loglinear analysis may correctly detect that this is explained entirely by a gendered distribution where men are higher on average.

The research question that motivates this research is whether loglinear models of tables of achieved marriages correctly capture evidence of underlying preferences, such as homophily, competition, female hypergamy, or relative ranking of attributes in the context of status exchange. This is partly a question about loglinear models, but more fundamentally it concerns the extent to which the theoretical mechanisms leave detectable evidence in the resulting data, the tables of spouse pairs usually used in analysis.

The paper focuses on the specific issue of female hypergamy in assortative mating by education. This is of particular contemporary interest in the context of a reversing gender gap in education. While some commentators are convinced that hypergamy preferences are entrenched in most societies, much research based on loglinear models shows that most observed hypergamy is driven by the inequality in the marginal distributions, in which case it should disappear with time as the marginal distributions of education equalise (or even reverse). We ask whether, under a number of different data generating processes, hypergamy preference is accurately detected by loglinear models of the square table of marriages, and thus whether we can trust the conclusions based on empirical research about the evidence for hypergamy preference.

This is addressed by simulating a number of ideal-typical mechanisms, overlaying varying levels of hypergamy preference over a number of distinct processes driving assortative mating. The simulations show that realistic patterns of assortative mating can be generated by quite different processes. Asymmetric preferences superimposed on these simulations yield data from which loglinear models reliably detect evidence of hypergamy, in a cleanly monotonic fashion, but with two important caveats. First, under certain circumstance loglinear analysis of the table of resulting marriages will report evidence of

---

[1] Id: prefpar.org,v 1.4 2019/07/09 12:11:03 brendan Exp $

non-zero net hypergamy even in the absence of a preference, and second, the relationship between the strength of the preference and the size of the parameter estimate differs according to the type of data generating process. These features are largely due to the temporal nature of the process: through repeated operation through time, with dynamically changing pools of remaining single agents, symmetric processes may cumulate into tables of achieved marriages that have asymmetric properties. If longitudinal data is available (e.g., cohorts with date of first marriage and spouse information) it will permit better models to be fitted. Restricting the data to recent marriages will also help; this is commonly advocated, but potentially costly in data terms. However, in the absence of such data, the bias due to modelling cumulated marriages is relatively small in magnitude.

2 Assortative mating and preferences

Research on assortative mating has a long history in sociology, demography, economics and biological anthropology. Sociological interest in assortative mating can be motivated variously, by the extent to which societies are open or closed, by the structure of inequality, or by the structuring of the life course. Currently there is significant activity on issues such as the interplay between homogamy or heterogamy and factors such as

- the income distribution (e.g., Salverda and Brals, 2012; Ravazzini, Kuhn and Suter, 2019)
- inter- and intra-generational inequality (Schwartz, 2013)
- migration, ethnicity and status exchange (Choi & Mare, 2012; Schwartz, Zeng & Xie, 2016; Telles & Esteve, 2019; Azzolini & 22:20:25, 2013),
- effects on outcomes such as transitions between cohabitation and divorce and their dissolution (Mäenpää & Jalovaara, 2013, 2014), or paid work in later life (Visser & Fasang, 2018)

2.1 Loglinear models and international comparisons

Assortative mating became reinvigorated as a topic in the early 1990s, when the use of loglinear models of square tables of marriages became established (Mare, 1991; Kalmijn, 1991b, 1991a, 1994, 1998; Qian, 1998; Smits, Ultee & Lammers, 1998; Raymo & Xie, 2000; Smits, Ultee & Lammers, 2000; Halpin & Chan, 2003; Schwartz, 2013). Loglinear models effectively estimate the underlying pattern of assortative mating, controlling for the strong and often uninformative effect of the marginal distributions. They also allow the exploitation of readily accessible data, namely tables of spouse characteristics (ideally of recent marriages), to draw conclusions about the processes and structures generating them.

Assortative mating remains a fertile base for development of loglinear models of square tables: Bucca and Urbina (2019) present a Lasso regularisation approach to selecting models, which they show to be more sensitive (particularly with small or sparse data sets) to the data generating process than conventional goodness of fit statistics like BIC. Similarly, Schwartz et al. (2016) demonstrates very clearly that loglinear models are necessary to detect evidence of status exchange by controlling not only for the marginal distributions but also the association between the various status dimensions. Bouchet-Valat (2014) presents a single index, the Mean Absolute Odds Ratio, a transformation of loglinear coefficients, which can be interpreted as a measure of the strength of homogamy.

Large-scale international comparisons are also common, testing theories about transitions to "modernity", changes in the economic position of females, or differences between cultures (e.g., Confucian vs Protestant vs Catholic), and using country characteristics to argue about the processes driving assortative mating. Much, but not all, of this work has been in terms of parameters of loglinear models, and thus separates effects of the marginal distributions from the processes that operate within their constraints (Smits et al., 1998; Raymo & Xie, 2000; Smits et al., 2000; Smits & Park, 2009; Katrňák, Fučík & Luijkx, 2012; Monaghan, 2014). A certain amount of this comparison is still conducted in terms of raw levels of homogamy or heterogamy (e.g., Esteve, McCaa & López, 2013). A completely different approach was taken by Blossfeld and Timm (2003), where duration models on longitudinal data, estimating the hazard of marrying up, etc. (see also Chan & Halpin, 2003). While it is very attractive to incorporate the dynamic dimension, as the duration models do, the loss of control for the marginal distributions is a problem: the hazard of marrying up does not depend just on the individual's characteristics and the time-dependent hazard function, but also the number of single alters with a higher education, and the distribution of peers. Loglinear models lose the individual dynamic focus, but control for margins.

Schwartz et al. (2016) make an interesting contribution to a debate on status exchange (Rosenfeld, 2005; Kalmijn, 2010; Gullickson & Fu, 2010; Rosenfeld, 2010) demonstrating convincingly why loglinear models are necessary to detect the presence of status exchange, by controlling not only for the marginal distributions, but also for correlation between the two status factors, and that simpler indices of association are misleading. This is a valuable paper in that it tackles in detail how to relate complex specifications of loglinear models to theoretical arguments, how to estimate very specific quantities (in particular, the extent to which there are excess cases in status exchange cells, over and above what would be observed without status exchange preferences). The current paper represents a similar exercise, if at a more fundamental level: the research question concerns whether square tables of spouse pairs actually contain enough information to detect the theoretical quantity by means of loglinear models.

Economics has made a contribution to this area, both at theoretical and empirical levels. Naturally, economists tend to regard the partnership process as a market-like, competitive process. Monaghan (2014) and Nielsen and Svarer (2009) are informed by economic theory, and look for economic rationality (regarding income expectations) as explainers of educational assortative mating but with little success. Monaghan finds only modest evidence of a link between returns to education and assortative mating within countries. Nielsen and Svarer (2009) similarly find that theoretical expectations based on economic models of rational agents maximising expected future income have little power, and rather, that the tendency for individuals to marry those they meet in the educational environment (the same or a nearby institution) accounts for half the pattern of sorting.

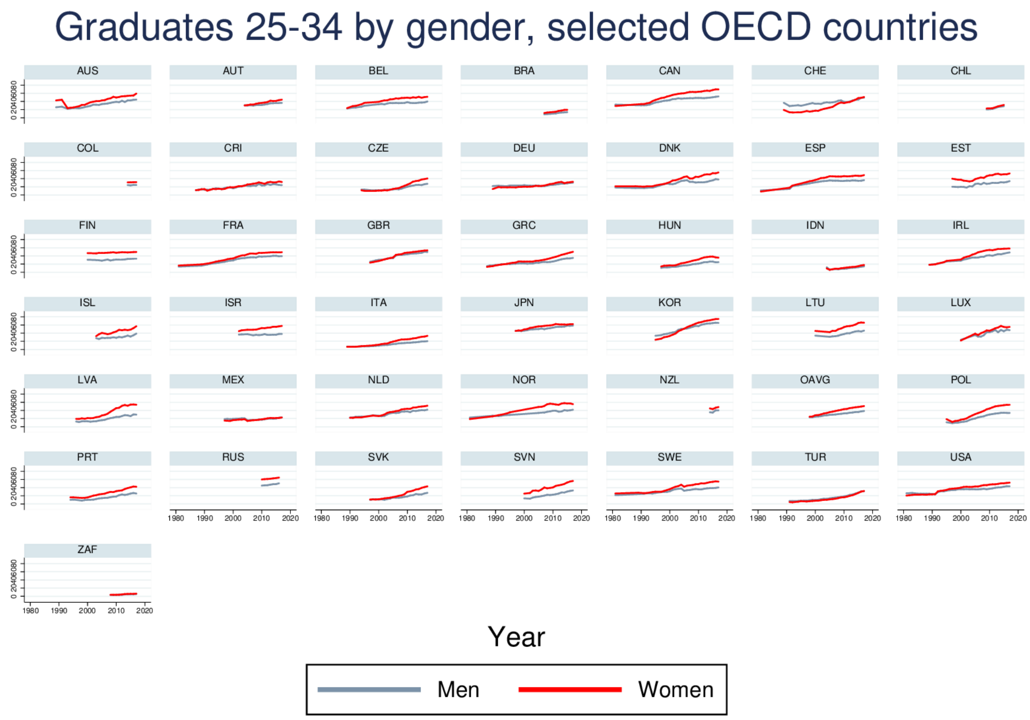
The distinctive contribution of economics is more at a theoretical level. Following Becker's extension of the boundaries of economics to include family formation matters (Becker, 1981, esp ch 4), Burdett and Coles (1997) adapt job-search theory to create an abstract model of assortative mating from first principles, and show that the familiar sorting we observe empirically can emerge from rational agents looking for the most desirable partner, using heuristics to decide when to settle for their current choice or to continue searching in a context of uncertainty.

In other fields including what might broadly be referred to as biological anthropology, much of the current concerns have to do with the interplay between social processes and genetic distributions. For instance, Hugh-Jones, Verweij, St. Pourcain and Abdellaoui, 2016 show that assortative mating by education leads to correlations within couples of polygenetic scores that predict educational attainment.

2.2 The reversing gender gap in education and female hypergamy norms

Over the past several decades many countries have experienced a significant shift in the impact of gender on educational attainment. The traditional gender gap at the expense of females has reduced, and in many cases reversed with females achieving higher levels of education on average than males (Vincent-Lancrin, 2008; OECD, 2008). Figure 1 uses data from a selection of OECD countries on the proportion of graduates aged 25–34, and shows that a female lead in participation emerged widely between about

1990 and 2010, driven by a combination of factors including the growth in the third level system, in female labour participation and in the sort of white-collar occupations where higher qualifications are typically required.



Source: https://data.oecd.org/eduatt/population-with-tertiary-education.htm

Figure 1: Gender and participation in higher education, OECD countries

While in earlier cohorts in most countries, men tended to be better educated than women, with the result that women marrying “up” was relatively common, current cohorts are faced with a situation where females are better educated on average, and pairs exhibiting female hypergamy are correspondingly rarer. The outcome of what we might term “excess female hypergamy” in the past tended to be taken as normative, and affected ideas of what constituted a suitable partner. Thus, it was widely considered desirable that the husband in a couple should be better educated. Hence the shift in the gender distribution of education has lead to fears that desirable matches will be harder to come by, with well-educated women and poorly-educated men potentially having a harder time finding partners. However, the question arises whether this preference is what we might call a “driving norm”, i.e., something that is resistant to change and drives structural outcomes, or a “reflective norm”, something that reflects what is currently observed and portrays it as desirable, but which will change as the experienced reality changes. A second question is to what extent there has been, empirically, excess female hypergamy: given the traditional male advantage in educational outcomes, there will be more female-hypergamous couples even if matching is random with respect to education, but do we observe even more than this?

This question has usually been tested via modelling tables of current (or first) marriages. There is widespread agreement that the bulk of female hypergamy is driven by the margins. For instance, Kalmijn (1998, p 413) says “once such differences are taken into account through loglinear analyses, researchers generally find little evidence of asymmetry”, and Schwartz and Mare (2005, fn 11) find that parameters accounting for female hypergamy do not improve the fit of the model. As I show below, application of similar models to current European Social Survey data on a large number of European countries is consistent with this point of view, showing at most some evidence of weak but declining relative female hypergamy.

However, popular discourse and some researchers interpret the empirical excess of female hypergamous couples as evidence of a strong and persistent norm for women to "marry up". A relatively early paper in the literature is even titled "Do American women marry up?"; even without the benefit of log-linear models, it concludes "not much" (Rubin, 1968). Rather more recently, Hakim baldly states:

> "Most women prefer economic dependence on men, if it is on offer, and increasing educational attainment among women has had remarkably little effect on women's preference for marrying 'up' if they can, to a better-educated and higher-earning spouse" (Hakim, 2007, p 130)

and:

> "Women's aspiration to marry up, if they can, to a man who is better-educated and higher-earning, persists in most European countries." (Hakim, 2011, p 30)

Other writers are more measured in their assessment of the persistence of the aspiration to marry up in the face of radical change in educational distributions. Van Bavel (2012) considers a number of possible outcomes, including both the persistence of the aspiration in the face of declining opportunities (leading perhaps to increased status exchange, or increased spinsterhood), and adaptation to the new margins (e.g., more male hypergamy). Esteve, García-Román and Permanyer (2012) directly address the effect of the gender gap reversal on hypergamy in a mostly empirical paper, using IPUMS census extract data on 56 countries. They show a very strong correlation between an index of female hypergamy and female disadvantage and use this to infer that preferences have changed. However, this is an unsurprising finding, given the structural constraints of tables of marriages (one that is perhaps descriptively interesting), and doesn't necessarily say anything about preferences, but rather the changing constraints within which they operate. See also Esteve et al. (2016), which uses even more extensive data on 120 countries.

2.3 Research using simulation

André Grow, Jan van Bavel and colleagues have been working for the past several years on the reversal of the gender gap in education and its implications for female hypergamy, using agent-based simulation models (van Bavel, 2012; Grow & van Bavel, 2015; van Bavel, Schwartz & Esteve, 2018). Their simulations are sophisticated, taking into account homophily (in terms of age and education level), general desirability of the potential partner (represented as earnings prospects), the possibility of divorce and repartnering, mortality, the structuring role of the educational system, etc. Unlike the simulations described below they allow male–female differences in the parameterisation. Their primary goal is to approximate the observed empirical distributions and their changes across cohort, in a way that is consistent with plausible theoretically informed mechanisms, with a view to determining whether changing patterns are evidence of changing preferences, or are consistent with the same preferences operating with different distributions. Their broad conclusion is that it is the latter.

This is very interesting work because it simulates preferences to generate outcomes, and compares simulated outcomes with empirical reality. By virtue of using simulation, it is one of the few streams of AM research with an explicitly stated generative model. Much existing research clearly has sophisticated theoretical ideas about how the processes operate, and some research explicates part of the processes in detail, but most effort is put into creating and justifying the statistical model of the outcome rather than the generative process. Burdett and Coles (1997) would be an exception: as is often the case with economists, they work from a relatively simple but explicit mathematical model, and derive a reduced-form statistical model that estimates parameters of the theoretical model. By virtue of being an explicit statement of the data-generating process, simulation can also be seen as embodying a theoretical model, rather than a statistical one.

The Grow–van Bavel simulations differ from those used in this paper in a number of important respects, primarily in being empirically focused and having a commitment to being at least approximately realistic, while the simulations in this paper address the analytic question of the extent to which the parameters of the data generation process can be detected by modelling the outcome data.

3 Four forms of assortative mating process

We commonly observe patterns of assortative mating such that spouse pairs are disproportionately similar. These patterns can theorised as arising from a number of different processes. As a rough generalisation, different disciplines will see different sorts of processes as being at least plausible: much biologically-informed work tends to expect the explanation to involve choosing mates with high reproductive fitness, while economists often assume the desired characteristic is earning potential. Sociology tends to prefer arguments based on culturally-mediated tastes and norms (such as religious or ethnic endogamy), or structural factors such as segregation by social class.

We can differentiate between processes that involve preferences, and those that involve structure. That is, patterns can be driven by individuals choosing between potential partners, or by individuals being exposed to different distributions of potential partners. We can further divide the preference processes into two: preferences for high values of certain attributes of the alter, leading to competition, and preference for the attributes of the ego and alter to be close, or homophily.[2] While homophilous processes will directly lead to like marrying like, competitive processes do it indirectly: while everyone desires the most desirable alter, only the most desirable egos are reciprocated. That is, while homophilous preferences map directly onto homogamous patterns, homogamy *emerges* from "competitive" preferences, without being implicit in them.

Patterned assortative mating will also emerge due to structural reasons, such as geographic or social segregation, or homophilous patterns of voluntary interaction. If people marry people in their own neighbourhood, and educational attainment is unevenly distributed geographically, educational homogamy will be higher than under independence as a side effect. Mare (1991) points out that the educational system itself can be a source of such structure, in that in spending time in education you interact with people with similar educational opportunities. Not only do schools serve as meeting places, but the time-structure of education has a strong effect: the higher your level of education, the later you leave the system. Thus by the time you start to search for a partner, many of those who left the system earlier have already made matches.

4 Excess educational homogamy and the ESS

Before embarking on the simulations, let us consider an empirical case. In this section, we model the outcomes of assortative mating by education across a number of cohorts and a wide range of European countries, using the European Social Survey. The European Social Survey is a valuable source of data on many aspects of life in European countries, including partnership. It contains data on many different countries (36 in at least one wave), with data collected in a fashion that is strictly comparable across

---

[2] Sometimes the preference may not centre on a zero difference, but on a positive value. Age and height might be examples of this, such that husbands are often older and taller, but not too much. As with educational assortative mating, it is a difficult problem to go from spouse distributions by height or age to claims about preferences, but separate work on simulation of partnering by height suggests that it is quite likely that individual couples have a propensity to replicate the average height difference in the population.

time and country, over eight waves between 2002 and 2016. While it is not longitudinal, the repeated collection means that we can build up quite substantial decadal cohorts, sampled at different time points. Data on education is coded according to ISCED (Schneider, 2010), reduced to four categories (primary or less, incomplete secondary, complete secondary, third level). While for many reasons it is considered ideal to model only recent first marriages (to exclude the effects of divorce and mortality, and recognising that processes affecting re-marriage may be quite different), research has shown that analysing tables of current marriages does not result in significant bias (Schwartz & Mare, 2012).

4.1 Excess female hypergamy

We can define an index of "raw" hypergamy as the ratio of hypergamous to hypogamous marriages, i.e., the number of marriages where the husband is better educated divided by the number where he is less educated.[3] The left panel of Figure 2 shows the evolution of this measure of observed excess hypergamy across the cohorts and ESS countries. Since it is a ratio, the value 1 implies no excess. As can be seen there is a very large variation in the observed level of hypergamy, but on average the ratio is substantially above one. However, there is a strong secular decline in the ratio, such that for a not insignificant number of country–cohort observations, there is hypogamy rather than hypergamy.

Of course, both the levels of and temporal trends in hypergamy will be very strongly affected by the marginal distributions. In most of these cohorts, men are better educated, so even if education had no bearing on partner choice there will be an excess of hypergamous couples, and this excess will decline across cohort. The right panel of Figure 2 shows the excess-hypergamy index calculated on the table of independence (i.e., as if couples were formed at random relative to education). This shows the same temporal pattern of temporal decline (or even reversal), but on a smaller scale: some but not all of the excess hypergamy is due to the marginal inequality. In other words, there is more hypergamy that expected under independence, but it is declining as the educational distributions change.

4.2 Log-linear models

The excess female hypergamy index is a crude measure. We are more interested in assessing the extent to which there is greater hypergamy than one would expect given the marginal distributions. We get an impression of this by comparing the two panels of Figure 2, but loglinear models provide the means to address this question more formally, by allowing us to predict data given the margins plus a parsimonious model of the association between husbands' and wives' education that separates out the symmetric and the asymmetric effects.

Given a two-way table of spouse pairs, an imparsimonious saturated model can be described as follows:

$$\log(F_{ij}) = \mu_o + \mu_i^R + \mu_j^C + \mu_{ij}^{RC}$$

or, in multiplicative terms:

$$F_{ij} = \tau_o \tau_i^R \tau_j^C \tau_{ij}^{RC}$$

That is, the frequency in each cell is affected by a grand mean effect, a row effect, a column effect, and a cell-specific association effect. Since this model has at least as many parameters as there are cells in the table, it is saturated, reproduces the data exactly, and tells us little. More useful models are arrived at by putting constraints on the $\mu_{\mathrm{ij}}^{\mathrm{RC}}$ term, using fewer degrees of freedom and potentially yielding information about the form of the association. Three are in common use:

---

[3] Esteve et al. (2012) use the log of this quantity as the dependent variable in their models; note in the graphs below the y-axis has a log scale.

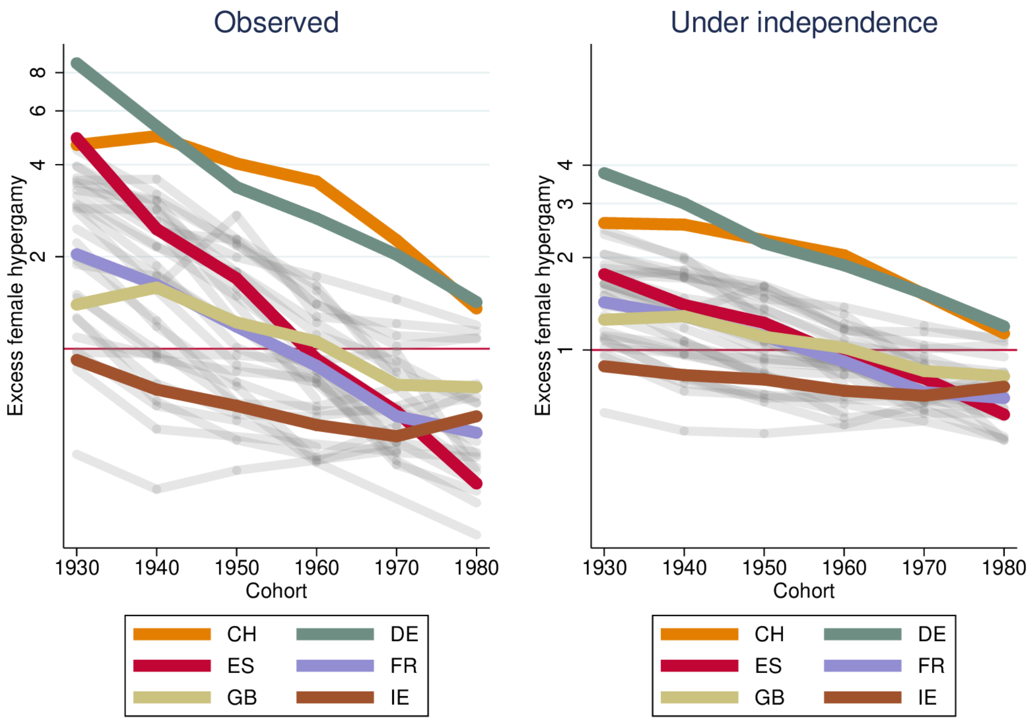


Figure 2: Observed and expected excess hypergamy by country and cohort, ESS waves 1 to 8, weighted data. Y-axis is log-scale.

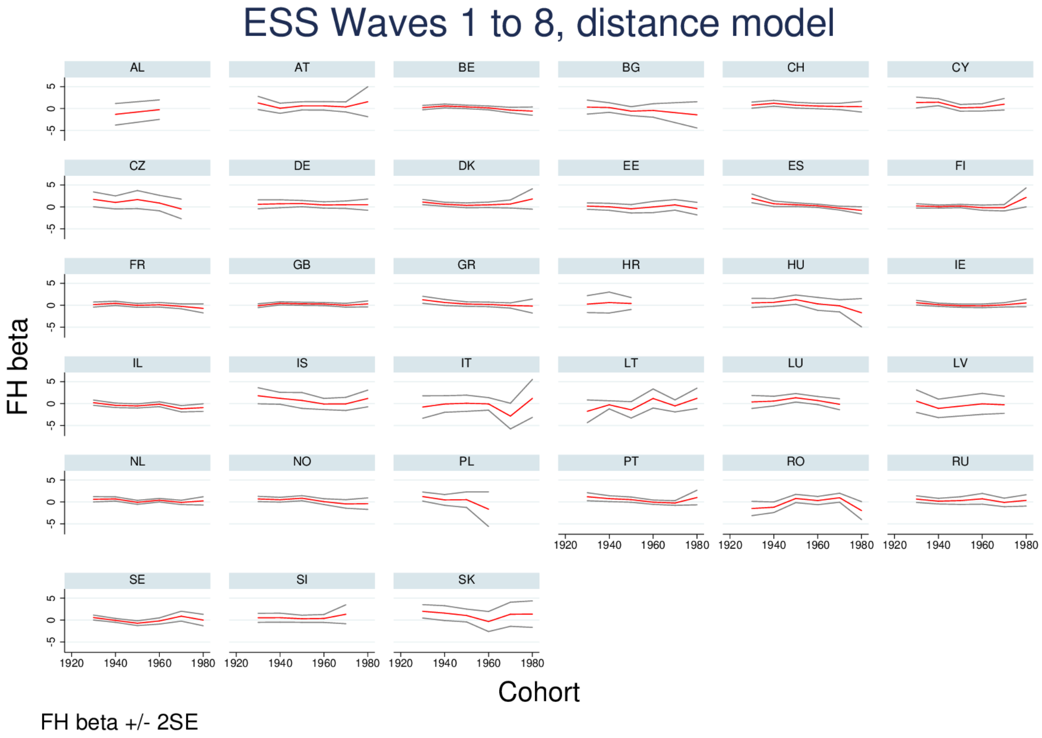


Figure 3: Female hypergamy loglinear estimates by country and cohort, ESS waves 1 to 8

- the distance model
- the crossings model, and
- the quasi-symmetry model.

The distance model assumes the educational variable is ordinal (in fact, with equal intervals), and pairings are more common or rarer depending on the distance between the two categories:

$$\log(n_{ij}) = \beta_0 + \beta_i^R + \beta_j^C + \delta_{ij}^{|R-C|}$$

The crossings model is also ordinal, but in a sequential sense. The values in the variable are thought of not as being located on an unobserved scale, but as having a series of barriers between them. Thus a 1/4 match requires stepping over the 1/2, 2/3 and 3/4 barriers, where as 3/4 only requires passing the 3/4 barrier. The effect of each barrier is estimated.

The quasi-symmetry model doesn't require ordinality, but assumes that, conditional on the row and column margins, the tendency for i/j and j/i matches are equal. Full symmetry requires the two marginal distributions to be the same.

$$\log(F_{ij}) = \mu_o + \mu_i^R + \mu_j^C + \delta_{ij}^{RC}, \quad \delta_{ij} = \delta_{ji}$$

See Table 1 for a visualisation of the association parameters.

Each of these models replaces the unconstrained association terms, $\mu_{\mathrm{ij}}^{\mathrm{RC}}$, with a smaller number of parameters. However, conditional on the margins, all are symmetric in effect. Female hypergamy is an asymmetric effect, and we fit it in these cases by adding a single parameter which applies to all cells one one side of the diagonal (for female hypergamy, for cells where the husband's education is higher). The interpretation of a hypergamy parameter in this context is the extra tendency for couples to be female-hypergamous after controlling for the margins and a parsimonious symmetric association. It is conventional in the literature to interpret this term as a measure of net female hypergamy.

4.3 Modelling the data

If we apply these models to the ESS data, country by country, we find that the crossings model fits poorly, but the quasi-symmetry and distance models perform better (BIC favours the distance model for 27 out of 36 countries, quasi-symmetry for 8, and independence for 1; see appendix). The female hypergamy parameter is significant at 5% for 15 of the 36 countries, and at 1% for 7. If we add an interaction with cohort (as a linear trend) this is significant at 5% for 4 countries for QS, and 7 for the distance model (and shows hypergamy is declining for all but one of these). Fitting separate hypergamy parameters by country and cohort (see Figure 4) shows no very strong patterns, with some evidence of decline.

The empirical picture is broadly consistent with the observations of Kalmijn (1998), Schwartz and Mare (2005) mentioned above, that taking account of hypergamy is often statistically unnecessary, once symmetric association is taken account of. That is not to say there is no evidence of hypergamy, but that it is not a clear pattern.

5 Simulations

This leads us to the central research question of the paper: do loglinear models faithfully capture information about the data generation process, in particular the operation of preferences? The specific question we focus on is whether the asymmetric association term in a loglinear model really captures preferences

Table 1: Symmetric and asymmetric association parameterisation

Distance model: 3 additional parameters on a 4x4 table

	a	b	c	d
a	0	1	2	3
b	1	0	1	2
c	2	1	0	1
d	3	2	1	0

Crossings model: 3 additional parameters on a 4x4 table

	a	b	c	d
a	0	1	1+2	1+2+3
b	1	0	2	2+3
c	1+2	2	0	3
d	1+2+3	2+3	3	0

Quasi-symmetry model: 6 additional parameters

	a	b	c	d
a	0	1	2	3
b	1	0	4	5
c	2	4	0	6
d	3	5	6	0

Asymmetric effect: 1 additional parameter

	a	b	c	d
a	0	1	1	1
b	0	0	1	1
c	0	0	0	1
d	0	0	0	0

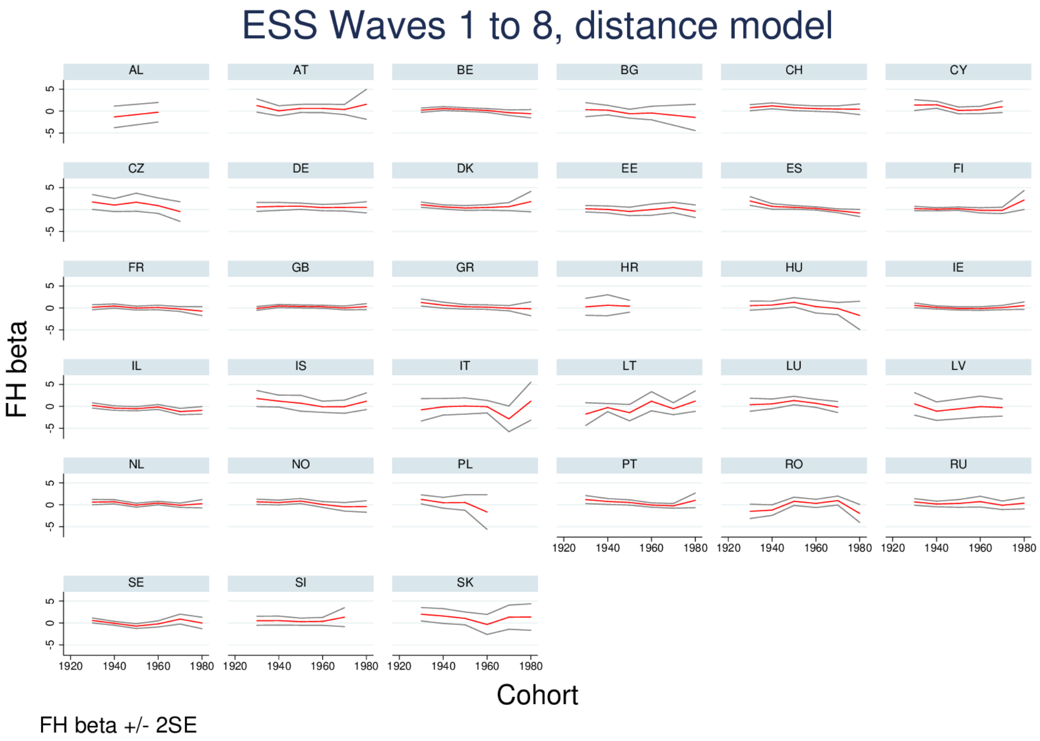


Figure 4: Female hypergamy estimates, distance model, ESS data

for hypergamy (or hypogamy). Is the presence of a significant hypergamy term in the model evidence of its presence in the data generating process? Is absence of significance good evidence for absence of an effect? We address this problem through simple agent-based simulation, generating spouse pairs through a variety of mechanisms, with hypergamy preferences varying from zero to strong, and applying loglinear models to the resulting data.

We consider four basic simulations, each capturing a different sort of theoretical mechanism, as considered above:

- Homophily
- Competition
- Social structure
- Time structure

The core structure of the simulations creates separate populations of men and women, with educational status as four ordinal categories, distributed such that men have substantially higher education (see Table 2). It then iterates through a series of random pairings, where if both candidates are “happy” with the pairing they are moved to the “married” pool (which is an absorbing state). The two preference-based simulations (homophily and competition) determine happiness based on the pairs’ educational characteristics, while the two structural simulations have pairs marry at random, but constrain the pairings. All of the simulations have a modifiable parameter which allows a preference on the part of women for husbands with higher levels of education than theirs (hypergamy preference). Multiple runs are made, with this parameter at different settings, including a hypergamy preference of zero.

All four simulations will generate more or less strong patterns of assortative mating, but they are symmetrical in design (except for the hypergamy parameter). To the extent that excess female hypergamy is observed, it is likely a consequence of the the uneven marginal distributions and, when used, the hypergamy-preference parameter. However, as will be seen, hypergamy levels that deviate from expectations under independence are observed for the homophily and time-structure simulations.

Note that unlike the agent-based simulations produced by the van Bavel/Grow group (van Bavel et al., 2018; van Bavel, 2012; Grow & van Bavel, 2015), which are intended to at least broadly capture the actual historical/empirical processes taking place in the countries studied, these simulations have the analytical goal of explicating the relationship between preferences, mechanisms and the resulting patterns of assortative mating.

Table 2: Levels of education in the simulations

	Base distribution	1	2	3	4
	Men	6.25	18.75	31.25	43.75
	Women	25.00	25.00	25.00	25.00
	Social structure distributions	1	2	3	4
Group 1	Men	9.77	24.19	32.26	33.87
	Women	33.33	27.78	22.22	16.67
Group 2	Men	6.25	18.75	31.25	43.75
	Women	25.00	25.00	25.00	25.00
Group 3	Men	3.03	13.64	30.30	53.03
	Women	14.29	21.43	28.57	35.71

5.1 Homophily

The homophily simulation implements the notion that people wish to form partnerships with people like themselves. Homophily may be qualitative, where people prefer partners from the same category (religion, ethnicity), or quantitative, where difference on some dimension (height, age, years of education) is disfavoured. Qualitative homophily may imply no preference between groups other than the in-group, while quantitative homophily implies that while perfect matches are preferred overall, near matches are preferred to far matches. For a characteristic like education, quantitative homophily is more intuitively attractive, so under this scenario the probability of proposing is inversely related to the difference in education. Necessarily, this is symmetric, as the distance is the same for each partner of a candidate pair. Clearly, homophily will tend to overpopulate the diagonal, and will make hypergamy as unpopular as hypogamy. We can see this structure as consonant with the "distance" loglinear model.

In the simulation, couple formation depends on their similarity on the education dimension, their values on a dimension of attractiveness that is independent of education, and an ephemeral pairwise random factor: Define σ_{ij} as the satisfaction of individual i with potential match j: this is a weighted sum of the difference in educational levels ϵ (where $\epsilon \in (1, 2, 3, 4)$), scaled from -0.5 (maximal mismatch) to 0.5 (exact match), and individual j's non-educational characteristics (random uniform on the interval (-0.5, 0.5), independent of education). If σ_i plus μ times a match-specific random uniform value on the interval (-0.5, 0.5), ζ, is greater than 0, ν_{ij} is true. If both ν_{ij} and ν_{ji} are true, the match takes place.

$$\sigma_{ij} = \omega(\frac{1}{2} - \frac{|\epsilon_i - \epsilon_j|}{3}) + (1 - \omega)\eta_j \tag{1}$$

$$\nu_{ij} = \sigma_{ij} + \mu\zeta_{ijt} > 0 \tag{2}$$

$$\nu_{ij} \wedge \nu_{ji} \Rightarrow \text{married}_{ij} \tag{3}$$

Note that in this simulation, the education levels of both ego and alter matter in for ego's decision, as well as alter's non-educational characteristics. The simulation has two analyst-set parameters, ω which controls the weight given to education versus other characteristics, and μ which controls the effect of the match-specific random component. The match-specific random component is important in introducing an element of non-determinacy.

5.2 Competition

The competition simulation implements the notion that higher spouse education is desirable, and thus the probability of either individual proposing is positively related to the education of the other. The simulation assumes the individual has regard only to the other's education, without reference to his/her own (except for the hypergamy-preference effect). Burdett and Coles (1997) make the quite reasonable assumption that even where choices are made in terms of spouse-desirability, individuals will have reference to their own characteristics and refrain from proposing to someone "out of their league"; while this is realistic, it is not necessary for the emergence of strong assortative patterns, and indeed is probably a weak effect in that people who fail to propose for this reason would very likely have been refused anyway. Thus in the basic simulation no reference is made to own characteristics, and this enters only in the hypergamy parameter.

In practice, couple formation depends on their desirability on the education dimension, their values on a dimension of attractiveness that is uncorrelated with education, and an ephemeral pairwise random factor.

Define σ_{ij} as the satisfaction of individual i with potential match j: this is a weighted sum of j's educational levels ϵ_j, scaled to -0.5 – 0.5, and individual j's non-educational characteristics (random uniform on the interval (-0.5, 0.5), independent of education). If σ_i plus a match-specific random uniform value on the interval (-0.5, 0.5), ζ, is greater than 0, ν_{ij} is true. If both ν_{ij} and ν_{ji} are true, the match takes place.

$$\sigma^C_{ij} = \omega(\frac{\epsilon_j - 1}{3} - \frac{1}{2}) + (1 - \omega)\eta_j \tag{4}$$

We determine the outcome by putting σ^C_{ij} into equations 2 and 3, as with the homophily simulation.

5.3 Social structure

Strong patterns of assortative mating may emerge even where the characteristic plays no role in the pairing decision, when the population is segregated into more-or-less endogamous groups where the characteristic of interest has different distributions. Thus if people tend to marry within their own social class, or their own region, and the distribution of education differs by class or region, there will be association between spouses' education levels even if education has no bearing on pairing decisions. This is simulated by creating three sub-populations with low, medium and high levels of education (with the same male–female asymmetry in each, see Table 2) and allowing marriage at random within each population. We achieve this using either of the previous algorithms with ω set to 0, so that only the non-education random factor affects decision making.

5.4 Time structure

The time-structure simulation is based on Mare's insight that the educational system itself structures pairing processes (Mare, 1991). The simulation focuses on the effect of the temporal structure, the notion that in many life courses partnership formation happens more-or-less soon after completion of education (it ignores Mare's second observation that educational institutions serve as meeting places). That is, typically people will seek to form a partnership after leaving education (perhaps with a delay), and typically people will not marry while still in education. Consequently, those who leave education early have a pool of potential spouses who have also left early, and those who leave later find that many early-leavers are already partnered. In the simulation, pairing is random with respect to education, except that before a threshold time (which rises with level of education) the probability of proposing is near zero, and substantially above zero after.

Individual i's satisfaction with a match depends on timing: if simulation time t is greater than or equal to the "release time" for i's educational group, ρ_{ϵ_i} it depends on the non-educational characteristics of the matched individual, η_j. If not, it is reduced by 1.0. The release times, ρ, are strictly increasing with level of education. The value σ^T can then be substituted into equations 2 and 3 as with the homophily and competition simulations. Note that if the value μ in equation 2 is sufficiently high, it will permit marriages to take place before the relevant release time, but at a lower rate than after.

$$\sigma_{ij}^T = \begin{cases} \eta_j - 1 & \text{if } t < \rho_{\epsilon_i} \\ \eta_j & \text{if } t >= \rho_{\epsilon_i} \end{cases} \tag{5}$$

5.5 Hypergamy

Female hypergamy preference is added to the foregoing simulations in the form of a parameter γ that raises the female satisfaction with the match if the male's education is higher than hers. In effect we replace equation 2 with the following:

$$\nu_{ij} = \sigma_{ij} + \mu\zeta_{ijt} + \gamma\delta_{ij} > 0 \tag{6}$$

where δ_{ij} takes the following values:

$$\delta_{ij} = \begin{cases} 0 \text{ if } i \text{ is male} \\ 0 \text{ if } \epsilon_j \leq \epsilon_i \\ 1 \text{ if } i \text{ is female and } \epsilon_j > \epsilon_i \end{cases} \tag{7}$$

6 Results

The simulations are written in Stata (predominantly using Mata, Stata's matrix language), and code will be made available. For the purposes of this paper, very large populations are used, 5 million men and 5 million women, and the simulations are run until 4.5 million couples are formed. The large population means that run-to-run variability is less, and the loglinear models return more stable results; it also means that small effects are detectable, as long as they are consistent. Seven to twelve iterations are needed for most of the simulations, but the time-dependent simulation is a little slower due to its staging of entry to the partnership process.

6.1 Patterns of assortative mating

Our first concern is whether the several simulations result in realistic patterns of assortative mating. The top row of Figure 5 shows, for each simulation, the departure from independence as Pearson residuals ($\frac{O-E}{\sqrt{E}}$). An "independence" simulation is included for reference: this pairs people without any reference to education (except for the hypergamy preference, when applied). While, as expected, the independence simulation shows no significant residuals, all four of the main simulations show the familiar patterns

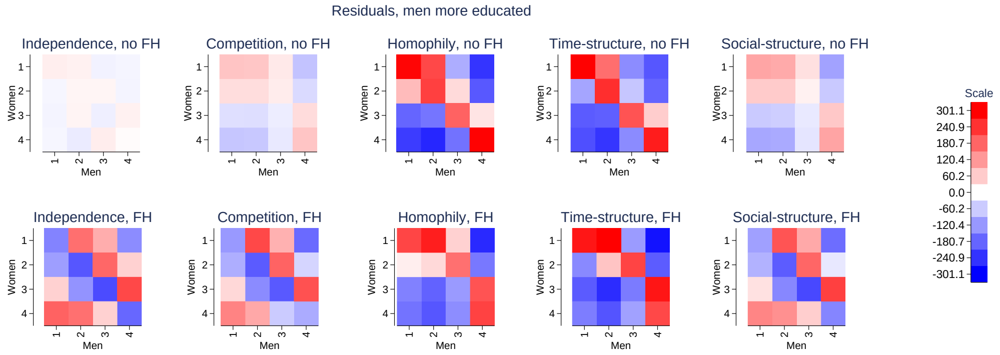


Figure 5: Simulations without and with a strong hypergamy preference, Pearson residuals

of assortative mating (heavier diagonal, sparse in top right and bottom left), but homophily and time structure are more marked than competition and social structure. For homophily versus competition, this is largely because under homophily both individuals' education is taken into account in each individual's assessment of the match, and thus ego and alter are more likely to have the same assessment of the match. For competition, only alter's characteristics are taken into account, so any given match is less likely to involve agreement (i.e., everyone is happy with an attractive alter, but the alter reciprocates only where the ego is also attractive). Similarly, the temporal segregation implied by the time-structure is stronger than that by the social segregation. A further point to note is that the pattern of residuals is very similar across time-structure and homophily, and across competition and social structure.

If we add a strong female hypergamy preference, the picture changes. The second row of Figure 5 shows the corresponding picture with a hypergamy preference parameter of 1 (a strong effect). The main effect of the hypergamy preference is consistent across the four simulations (and independence, where the only association is driven by the hypergamy preference). The main effect is to shift the weight of the tables up and right, so the highest residuals tend to be in cells where the man is one level above the woman. One interesting feature evident in the independence, competition and social structure panels is that there is an excess of level 4 women marrying men in levels 1 and 2. While superficially counter-intuitive, this is a direct consequence of the hypergamy preference: level 4 women cannot marry up, and thus more educated men will be preferentially taken by less educated women, leaving the less educated men for level 4 women to marry (if more slowly). The stronger homophilous patterns created by the homophily and time-structure simulations mask this effect.

6.2 Model detection of the effect of hypergamy preference

Eyeballing the residuals in this manner is helpful, but it is difficult to clearly assess the effect of the hypergamy preference. We can get a sharper picture of this by loglinear modelling. We create $4 \times 4 \times 2$ tables for each simulation, where one panel has no hypergamy preference and the other does, and fit a model which implies a common pattern of association (i.e., one that does not account for the hypergamy preference). Figure 6 shows the residuals for the hypergamy-effect panel, indexing the difference due to the hypergamy-preference parameter. This provides a more consistent picture, with all simulations showing the pattern of highly educated women marrying down. Moreover, the strength of the effect is similar across simulation, with the exception that the time-structure simulation has a very strong one-level-up effect. This is largely due to the time structuring. That is, when a woman with lower education emerges onto the partnership market, there are lots of males at the same level as she is, but she will have to wait for the opportunity to marry up. And then, the opportunity to marry one level up (which,

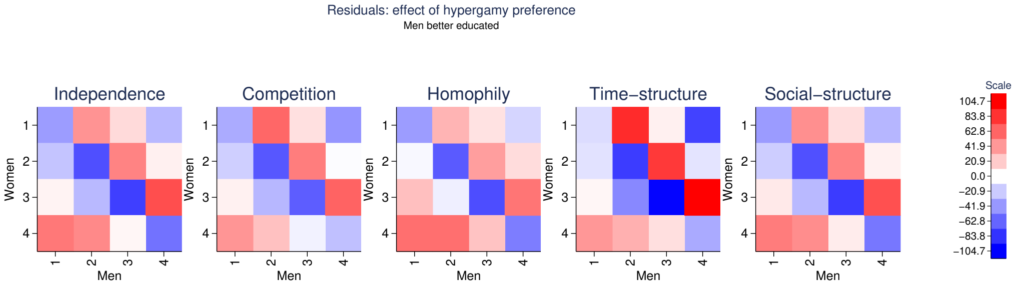


Figure 6: Loglinear residuals, indexing the difference due to hypergamy preference

by design, constitutes hypergamy just as much as marrying more levels up) arrives sooner than the opportunity to marry more than one level up.

6.3 Estimating the hypergamy parameter

What happens when we apply the loglinear association models described above to the simulation data? We consider the quasi-symmetry and distance models only, as in general the crossings model fits less well. For each simulation, and for each of five levels of the hypergamy preference parameter (from 0 to 1 in steps of 0.25) we estimate the distance and quasi-symmetry models with the additional asymmetric parameter. The magnitude of the loglinear hypergamy estimate is shown in Figure 7 (see also appendix). The results from the two models are quite similar and have three main features:

- the loglinear estimate responds strongly and monotonically to the preference parameter, from near zero at zero to relatively high levels at a preference factor of 1.0
- independence, competition and social structure are similar, with time-structure showing higher and homophily lower estimates of the hypergamy effect
- though estimates are close to zero for zero preference, time-structure shows a significant positive hypergamy effect and homophily a smaller but significant negative effect, in the absence of any hypergamy preference.

Thus we see that the practice of estimating hypergamy preference from the loglinear parameter is justified, but only to a first approximation. In particular, the fact that we observe significant effects for two simulations even in the absence of hypergamous preference depends on the very large sample size of 4.5 million couples. Normal survey samples do not have the power to detect these effects. However, it is nonetheless the case that we observe hypergamy (as estimated by loglinear models) in its absence as a preference, and furthermore, the strength of the parameter varies according to the data generating process. Thus our interpretation of the hypergamy parameter has to be qualified: factors other than a hypergamy preference can affect its size.

6.4 Time dynamics and a simpler homophily simulation

How could the data generating process produce hypergamy (as measured by loglinear models) over and above that predicted by a symmetric association model? The answer lies in the time dynamics. The loglinear model estimates the patterns in the table in a flat cross-sectional manner, but the processes generating the table operate through time, and while these processes might be completely symmetric as

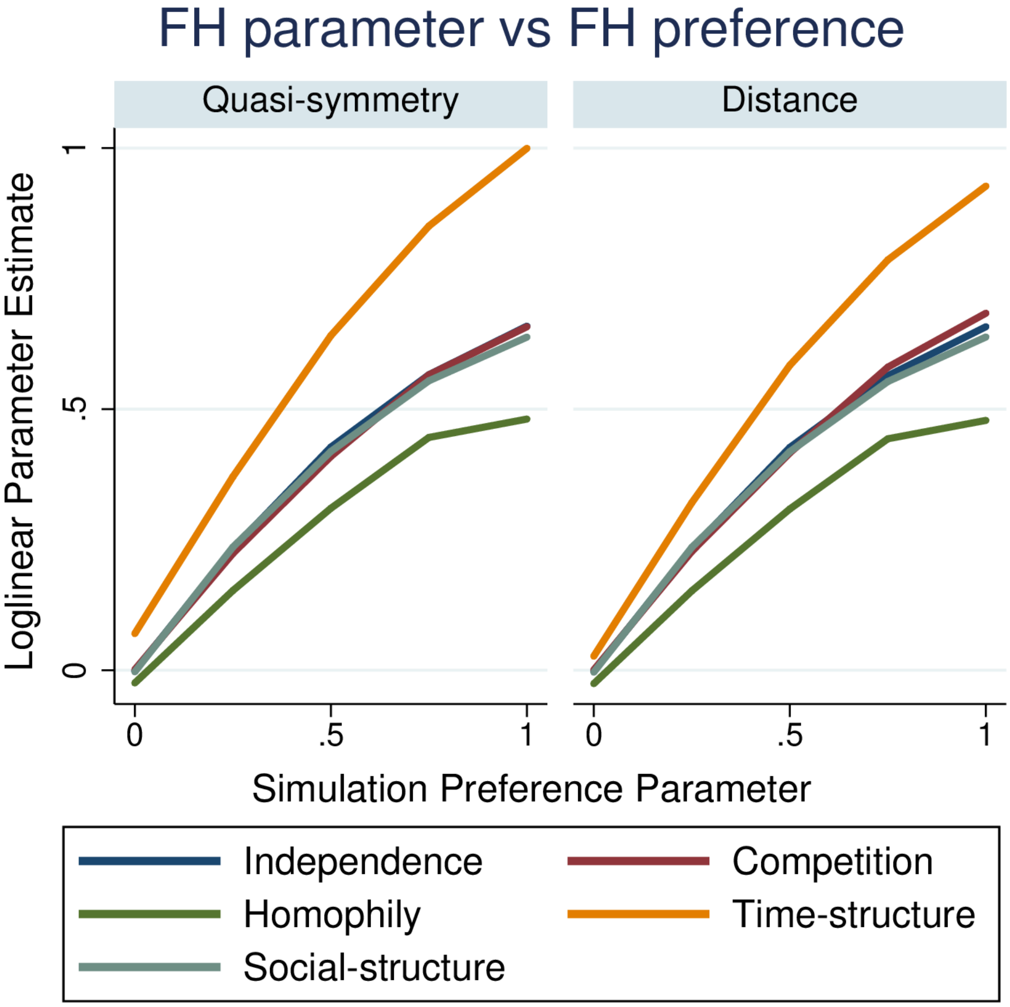


Figure 7: The response of the loglinear hypergamy parameter to hypergamy preference

they happen, the manner in which they dynamically alter the distribution of single people can sum over time to an outcome that appears to have positive hypergamy (or negative, depending on the process). We can examine this in the case of homophily with a much simpler simulation than the foregoing. We start with the same uneven distributions of single males and females (160,000 of each), and iterate 20 times, pairing off numbers at each iteration proportional to the product of the proportions of single males and females in each category, scaled inversely by the distance between the categories (all homogamous matches marry, 75% of those one step apart, 56.25% for those two steps apart, and 42.2% of those three steps apart). This is equivalent to people meeting at random, with a probability of marriage declining with distance. We can use this to generate a three-way table of marriages by spouse characteristics by iteration, and a two-way table collapsing across iterations.

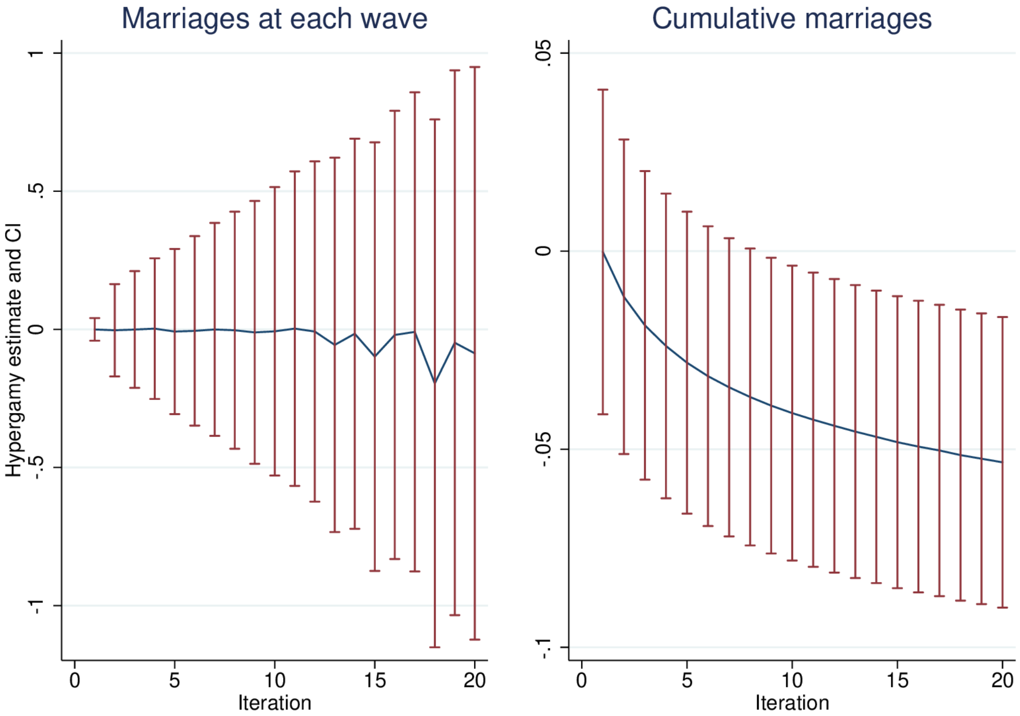


Figure 8: Female hypergamy parameter, simple simulation, by iteration. The left panel shows the parameter estimate ($\pm 1.96 \times \mathrm{SE}$) for new pairings at each wave, using a distance model. The right panel shows estimates for cumulative marriages at each wave.

If we fit a distance model with a hypergamy parameter to the resulting data collapsed across iterations (i.e., a 4×4 marginal table), we get a significant negative female hypergamy parameter, but if we model any one wave on its own (i.e., new pairings that wave) we cannot reject the null. Figure 8 shows estimates with confidence intervals, with in the left panel the effect for new marriages each wave, and in the right panel the cumulative marriages by each wave. The effects for new marriages are consistently null, while the cumulative marriages steadily build to a small but significant negative effect. While the vast majority of pairings happen early, the later pairings (based on distributions of singles that have changed, with proportionally more highly-educated men and more less-educated women) have the effect of creating a cumulative pattern with apparent evidence of female hypogamy.

If we fit a model that interacts the margins with wave, the hypergamy effect disappears again. In other words, while the data collapses over the wave margin into a pattern that suggests hypogamy, when the data is modelled in a way that takes time into account, the correct inferences can be made.

7 Discussion and conclusion

The mapping of preferences to outcomes in the context of constraints (such as the gendered distribution of education) and of competition with peers, for alters who also have preferences and agency, is complex. The foregoing simulations have demonstrated that a number of conceptually distinct mechanisms can generate data with realistic patterns of assortative mating, and have thrown light in particular on the consequence of asymmetric preferences for the outcomes. These are ideal-typical simulations, however, and do not seek to generate adequate accounts of the empirical process. In the real world it is likely that many of the mechanisms described are in operation in parallel, with social segregation by neighbourhood, network, institutional affiliation, varying patterns of life course ordering across groups and so on being very important, both homophilic and competitive preferences operating, and a more complex set of attributes being taken into account. In particular there is likely to be more complex gender asymmetry than a simple hypergamy preference. Another important difference is that while the simulations operate batch-style on a single cohort, in the real world partnership formation is a rolling continuous-production process.

However, empirical adequacy is not the primary goal. Rather, the purpose is analytical: to map from preferences, in particular an asymmetric female hypergamy preference, to outcomes, and see how well preferences can be detected by loglinear modelling of the resulting square table of marriages. In so far, loglinear models have some success: broadly speaking the hypergamy preference is detected monotonically in analysis of each simulation's output. However, the relationship between the strength of the hypergamy preference and the size of the corresponding parameter estimate varies across the simulation mechanisms. More concerningly, for some mechanisms, the loglinear estimates are non-zero when the hypergamy preference is zero. The difference in magnitude of the response may be because within the mechanics of the different simulations there is more or less scope for a hypergamy preference to operate. We have shown that for the homophily simulation it arises directly from the time dynamics: the simulation mechanisms operate through time in a fashion that changes the structure of opportunity (the distributions of the remaining single agents) such that a zero hypergamy preference operating at each time point cumulates into a pattern of hypogamy. While a similar analysis has not been presented for the time-structure simulation, the same is true: the staged release of individuals onto the marriage market means the structure of opportunity has an even more marked temporal pattern, such that symmetric preferences at each time point cumulate into an asymmetric hypergamous pattern.

The fact that time-dynamics yield unexpected results is most clearly evident in the relationship between the strength of the hypergamy preference and the magnitude of its parameter estimate, particularly in the context of the simple homophily simulation (see Figure 8). That simplified simulation shows that taking time into account solves the problem: the operation of preferences at each time point can be captured correctly by loglinear association models, but a spurious hypogamy effect emerges when we collapse along the time dimension. We can also see the complex nature of the association, and how it varies with the strength of the preference parameter, by looking at the goodness of fit of the loglinear association models of the main simulated data (see Table 3). (*NOTE: currently this calculates BIC using the number of table cells for N. Using the sample size gives different results, favouring parsimonious models more. For loglinear models, there are arguments in favour of using the number of cells rather than the number of individuals, but in this case it is likely far too conservative*.) Looking at the four mechanisms, plus the independence simulation (where education is not considered, except via hypergamy), we use BIC to compare the following models (with a hypergamy parameter):

- Independence (row and column effects only)
- Distance

- Quasi-symmetry
- the saturated model.

Table 3: Association model preferred by BIC, across simulation mechanisms and strength of hypergamy preference

FH pref	Simulation				
	Independence	Competition	Homophily	Time-structure	Social structure
0	Independence	Distance	Saturated	Saturated	Distance
.25	Distance	Saturated	Saturated	Saturated	Distance
.5	Distance	Saturated	QS	QS	Distance
.75	QS	Saturated	QS	Saturated	QS
1	QS	Saturated	Saturated	Saturated	Saturated

The independence simulation is informative: while the independence model is preferred when there is no hypergamy preference, as the preference rises, loglinear models see either a distance or quasi-symmetric pattern, though our naive expectation would be that the association should be entirely described by the hypergamy parameter. The competition and social structure simulations show sensible results at zero preference, but as the preference rises (immediately for competition, later for social structure) the pattern of association becomes too complex for distance or QS to describe it adequately, and BIC prefers the saturated model. Interestingly, homophily and time-structure show over-complex association from the start, with a role for QS in the middle only of the preference range. Compare this to the models of the real-world ESS data, where the distance and (to a lesser degree) QS models were preferred in nearly all cases.

In short, the introduction of female hypergamy preferences into these dynamic simulation brings about more complex association than might be naively expected, and for some mechanisms the association is more complex even without hypergamy preferences. This is due to the dynamic nature of the simulations, where the operation of even relatively simple preferences/mechanisms cumulates over time into patterns that are more complex than expected.

The simulations as presented are simple, but capable of extension. The formulation of hypergamy preference as a simple block asymmetry is potentially limiting, and it might be more realistic to model it as a more complex gender asymmetry in preferences. Another path for further work could include combining mechanisms, and calibrating against data, converging on the sort of work done by the Grow/van Bavel group (e.g., Grow & van Bavel, 2015). Extending the simulations to address other theoretical issues would also be attractive, status exchange being an obvious candidate. The simulations currently take account of education and an unspecified attribute uncorrelated with education; it would be straightforward to extend them to incorporate other dimensions, and to explicitly model the impact of relative preferences on behaviour, and on outcomes, relating to work in the tradition of Schwartz et al. (2016).

The simulations have a number of implications for the practice of assortative mating research. First (and this is not a particularly new observation but has perhaps new force), different mechanisms can generate quite similar patterns of assortative mating, so we need to use information additional to that in square tables of marriages to differentiate between the mechanisms. Second, while loglinear models are clearly necessary and far superior to simpler summaries of the structure of tables, we need to be cautious about making inferences from parameters to preferences. Processes happen through time, while loglinear models are one-shot. If at all possible, we should use data with a time dimension, and incorporate time in the modelling. In the absence of properly longitudinal data, this can be read as a renewed injunction to focus on recent marriages where at all possible.

References

Azzolini, D. & 22:20:25, R. G. 2. 2. (2013). *And they all lived happily ever after. evidence of status exchange among mixed-nativity couples in Italy*. Conference presentation, ECSR Tilburg.

Becker, G. S. (1981). *A treatise on the family*. Cambridge, MA.: Harvard University Press.

Blossfeld, H.-P. & Timm, A. (Eds.). (2003). *Who marries whom: Educational systems as marriage markets in modern societies*. Kluwer.

Bouchet-Valat, M. (2014). Les évolutions de l'homogamie de diplôme, de classe et d'origine sociales en france (1969-2011) : Ouverture d'ensemble, repli des élites. *Revue française de sociologie*, *55*(3), 459–505. doi:10.2307/24382304

Bucca, M. & Urbina, D. R. (2019). Lasso regularization for selection of log-linear models: An application to educational assortative mating. *Sociological Methods and Reseach*, *(online first 2019)*.

Burdett, K. & Coles, M. G. (1997). Marriage and class. *Quarterly Journal of Economics*, *112*(1), 141–168.

Chan, T. W. & Halpin, B. (2003). Who marries whom in Great Britain? In H.-P. Blossfeld & A. Timm (Eds.), *Who marries whom: Educational systems as marriage markets in modern societies*. Kluwer.

Choi, K. H. & Mare, R. D. (2012). International migration and educational assortative mating in Mexico and the United States. *Demography*, *49*(2), 449–476.

Esteve, A., García-Román, J. & Permanyer, I. (2012). The gender-gap reversal in education and its effect on union formation: The end of hypergamy? *Population and Development Review*, *38*(3), 535–546.

Esteve, A., McCaa, R. & López, L. Á. (2013). The educational homogamy gap between married and cohabiting couples in Latin America. *Population Research and Policy Review*, *32*(81), 81–102.

Esteve, A., Schwartz, C. R., van Bavel, J., Permanyer, I., Klesment, M. & García-Román, J. (2016). The end of hypergamy: Global trends and implications. *Population and Development Review*, *42*(4), 615–625. doi:10.1111/padr.12012. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/padr.12012

Grow, A. & van Bavel, J. (2015). Assortative mating and the reversal of gender inequality in education in Europe an agent-based model. *PLOS One*. doi:https://doi.org/10.1371/journal.pone.0127806

Gullickson, A. & Fu, V. K. (2010). Comment: An endorsement of exchange theory in mate selection. *American Journal of Sociology*, *115*(4), 1243–1251. Retrieved from http://www.jstor.org/stable/10.1086/649049

Hakim, C. (2007). Dancing with the Devil? essentialism and other feminist heresies. *The British Journal of Sociology*, *58*(1), 123–132. doi:10.1111/j.1468-4446.2007.00142.x

Hakim, C. (2011). *Feminist myths and magic medicine: The flawed thinking behind calls for further equality legislation*. Centre for Policy Studies. London.

Halpin, B. & Chan, T. W. (2003). Educational homogamy in Ireland and Britain: Trends and patterns. *British Journal of Sociology*, *54*(4), 473–495.

Hugh-Jones, D., Verweij, K. J., St. Pourcain, B. & Abdellaoui, A. (2016). Assortative mating on educational attainment leads to genetic spousal resemblance for polygenic scores. *Intelligence*, *59*, 103–108.

Kalmijn, M. (1991a). Shifting boundaries: Trends in religious and educational homogamy. *American Sociological Review*, *56*, 786–800.

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

Kalmijn, M. (1994). Assortative mating by cultural and economic occupational status. *American Journal of Sociology*, *100*(2), 422–52.

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology*, *24*, 395–421.

Kalmijn, M. (2010). Educational inequality, homogamy, and status exchange in black-white intermarriage: A comment on Rosenfeld. *American Journal of Sociology*, *115*(4), 1252–1263.

Katrňák, T., Fučík, P. & Luijkx, R. (2012). The relationship between educational homogamy and educational mobility in 29 European countries. *International Sociology*, *27*(4), 551–573.

Mäenpää, E. & Jalovaara, M. (2013). The effects of homogamy in socio-economic background and education on the transition from cohabitation to marriage. *Acta Sociologica*, *56*(3), 247–263.

Mäenpää, E. & Jalovaara, M. (2014). Homogamy in socio-economic background and education, and the dissolution of cohabiting unions. *Demographic Research*, *30*(65), 1769–1792. doi:10.4054/DemRes.2014.30.65. eprint: http://www.demographic-research.org/volumes/vol30/65/30-65.pdf

Mare, R. D. (1991). Five decades of educational assortative mating. *American Sociological Review*, *56*, 15–32.

Monaghan, D. (2014). *The impact of economic inequality and gender parity on educational assortative mating: Evidence from the Luxembourg Income Study* (LIS Working Paper Series No. 607). Luxembourg Income Study (LIS). Luxembourg. Retrieved from http://hdl.handle.net/10419/95400

Nielsen, H. S. & Svarer, M. (2009). Educational homogamy: How much is opportunities? *Journal of Human Resources*, *44*(4), 1066–1086. eprint: http://jhr.uwpress.org/content/44/4/1066.full.pdf+html. Retrieved from http://jhr.uwpress.org/content/44/4/1066.abstract

OECD. (2008). *Higher education to 2030 (vol 1): Demography*. Paris: OECD.

Qian, Z. (1998). Changes in assortative mating: The impact of age and education, 1970–1990. *Demography*, *35*(3), 279–292.

Ravazzini, L., Kuhn, U. & Suter, C. (2019). Les opposés ne s'attirent pas – le rôle de la formation et du revenu dans la mise en couple en Suisse. *Social Change in Switzerland*, (17).

Raymo, J. M. & Xie, Y. (2000). Temporal and regional variation in the strength of educational homogamy. *American Sociological Review*, *65*, 773–780.

Rosenfeld, M. J. (2005). A critique of exchange theory in mate selection. *American Journal of Sociology*, *110*, 1284–1325.

Rosenfeld, M. J. (2010). Still weak support for status-caste exchange: A reply to critics. *American Journal of Sociology*, *115*(4), 1264–1276.

Rubin, Z. (1968). Do American women marry up? *American Sociological Review*, *33*(5), 750–760. Retrieved from https://www.jstor.org/stable/2092885

Salverda, W. & Brals, D. (2012). *Income inequality and educational homogamy in household formation in Europe*. University of Amsterdam.

Schneider, S. (2010). Nominal comparability is not enough: (in-)equivalence of construct validity of cross-national measures of educational attainment in the European Social Survey. *Research in Social Stratification and Mobility*, *38*(3), 343–357. doi:10.1016/J.RSSM.2010.03.001

Schwartz, C. R. (2013). Trends and variation in assortative mating: Causes and consequences. *Annual Review of Sociology*, *39*(1), 451–470. doi:10.1146/annurev-soc-071312-145544. eprint: http://dx.doi.org/10.1146/annurev-soc-071312-145544

Schwartz, C. R. & Mare, R. D. (2005). Trends in educational assortative marriage from 1940 to 2003. *Demography*, *42*(4), 621–646. doi:10.1353/dem.2005.0036

Schwartz, C. R. & Mare, R. D. (2012). The proximate determinants of educational homogamy: The effects of first marriage, marital dissolution, remarriage, and educational upgrading. *Demography*, *49*(2), 629–650.

Schwartz, C. R., Zeng, Z. & Xie, Y. (2016). Marrying up by marrying down: Status exchange between social origin and education in the United States. *Sociological Science*, *3*(1003-1027).

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.

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*(5), 781–788.

Telles, E. & Esteve, A. (2019). Racial intermarriage in the Americas. *Sociological Science*, *6*. doi:10.15195/v6.a12

van Bavel, J. (2012). The reversal of gender inequality in education, union formation and fertility in Europe. *Vienna Yearbook of Population Research*, *10*, 127–154. Retrieved from http://epub.oeaw.ac.at/0xc1aa500d_0x002d68a4.pdf

van Bavel, J., Schwartz, C. R. & Esteve, A. (2018). The reversal of the gender gap in education and its consequences for family life. *Annual Review of Sociology*, *44*, 341–360.

Vincent-Lancrin, S. (2008). The reversal of gender inequalities in higher education: An on-going trend. In *Higher education to 2030 (vol 1): Demography*. Paris: OECD.

Visser, M. & Fasang, A. E. (2018). Educational assortative mating and couples' linked late-life employment trajectories. *Advances in Life Course Research*, *37*, 79–90.

A Appendix

A.1 Simple homophily/hypogamy simulation

From simple homophily

- Cumulative with distance and fh
- Pooled with distance and fh
- Fit stats, beta & se for graph

A.1.1 Hypergamy effect in cumulative marriages after 20 waves

```
Iteration 0:   log likelihood = -1433856.4
Iteration 1:   log likelihood = -1433296.7
Iteration 2:   log likelihood = -1433295.7
Iteration 3:   log likelihood = -1433295.7

Poisson regression                              Number of obs     =        320
                                                LR chi2(10)       =  288481.28
                                                Prob > chi2       =     0.0000
Log likelihood = -1433295.7                     Pseudo R2         =     0.0914

------------------------------------------------------------------------------
       delta |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         men |
          2  |   1.045278   .0057249   182.59   0.000     1.034057    1.056498
          3  |   1.584068   .0066197   239.29   0.000     1.571094    1.597043
```

```
          4  |   2.015038   .0087688   229.80   0.000     1.997852    2.032225
             |
       women |
          2  |  -.1575084    .004482   -35.14   0.000     -.166293   -.1487238
          3  |  -.2762028   .0062689   -44.06   0.000    -.2884895   -.2639161
          4  |  -.2929823   .0087363   -33.54   0.000    -.3101051   -.2758594
             |
        dist |
          1  |  -.2626679   .0054513   -48.18   0.000    -.2733521   -.2519836
          2  |  -.5467012   .0051368  -106.43   0.000    -.5567692   -.5366333
          3  |  -.7974926   .0057183  -139.46   0.000    -.8087003   -.7862849
             |
        1.fh |   -.052431   .0083606    -6.27   0.000    -.0688176   -.0360445
       _cons |   6.885158   .0055418  1242.41   0.000     6.874297     6.89602
------------------------------------------------------------------------------
```

A.1.2 Hypergamy effect in wave-specific marriages pooled over 20 waves

```
Iteration 0:   log likelihood = -1768128.4
Iteration 1:   log likelihood = -495991.54  (backed up)
Iteration 2:   log likelihood = -352273.96  (backed up)
Iteration 3:   log likelihood = -193197.73
Iteration 4:   log likelihood = -83065.516
Iteration 5:   log likelihood = -21841.527
Iteration 6:   log likelihood = -16650.813
Iteration 7:   log likelihood = -16441.933
Iteration 8:   log likelihood = -16440.415
Iteration 9:   log likelihood = -16440.414

Poisson regression                                Number of obs     =        320
                                                  LR chi2(29)       = 3122191.81
                                                  Prob > chi2       =     0.0000
Log likelihood = -16440.414                       Pseudo R2         =     0.9896

------------------------------------------------------------------------------
       delta |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         men |
          2  |   1.045278   .0057249   182.59   0.000     1.034057    1.056498
          3  |   1.584068   .0066197   239.29   0.000     1.571094    1.597043
          4  |   2.015038   .0087688   229.80   0.000     1.997852    2.032225
             |
       women |
          2  |  -.1575084    .004482   -35.14   0.000     -.166293   -.1487238
          3  |  -.2762028   .0062689   -44.06   0.000    -.2884895   -.2639161
          4  |  -.2929823   .0087363   -33.54   0.000    -.3101051   -.2758594
```

```
             |
        wave |
          2  |  -2.643416   .0050972  -518.60   0.000    -2.653406   -2.633425
          3  |  -3.066574   .0062257  -492.57   0.000    -3.078776   -3.054372
          4  |  -3.400671    .007311  -465.15   0.000    -3.415001   -3.386342
          5  |  -3.678796   .0083687  -439.59   0.000    -3.695198   -3.662393
          6  |  -3.918313   .0094087  -416.46   0.000    -3.936754   -3.899873
          7  |  -4.128714   .0104331  -395.73   0.000    -4.149163   -4.108266
          8  |  -4.316729   .0114459  -377.14   0.000    -4.339163   -4.294296
          9  |  -4.486997   .0124501  -360.40   0.000    -4.511399   -4.462595
         10  |  -4.642042   .0134429  -345.32   0.000     -4.66839   -4.615695
         11  |  -4.785364   .0144324  -331.57   0.000    -4.813651   -4.757077
         12  |  -4.918865   .0154207  -318.98   0.000    -4.949089   -4.888641
         13  |   -5.04233   .0163957  -307.54   0.000    -5.074465   -5.010195
         14  |  -5.158612   .0173711  -296.97   0.000    -5.192659   -5.124565
         15  |   -5.26855   .0183472  -287.16   0.000     -5.30451    -5.23259
         16  |  -5.371087   .0193076  -278.19   0.000    -5.408929   -5.333244
         17  |  -5.468849   .0202704  -269.79   0.000    -5.508578   -5.429119
         18  |  -5.560932   .0212216  -262.04   0.000    -5.602525   -5.519338
         19  |  -5.650025   .0221847  -254.68   0.000    -5.693507   -5.606544
         20  |  -5.733876   .0231313  -247.88   0.000    -5.779212   -5.688539
             |
        dist |
          1  |  -.2626679   .0054513   -48.18   0.000    -.2733521   -.2519836
          2  |  -.5467012   .0051368  -106.43   0.000    -.5567692   -.5366333
          3  |  -.7974926   .0057183  -139.46   0.000    -.8087003   -.7862849
             |
        1.fh |   -.052431   .0083606    -6.27   0.000    -.0688176   -.0360445
       _cons |   9.619298   .0055775  1724.67   0.000     9.608367     9.63023
------------------------------------------------------------------------------
```

A.1.3 Hypergamy effect in wave-specific marriages pooled over 20 waves, allowing margins to change by wave

```
Iteration 0:   log likelihood = -6647341.9
Iteration 1:   log likelihood =   -4693624  (backed up)
Iteration 2:   log likelihood = -4255288.3  (backed up)
Iteration 3:   log likelihood = -3878117.7  (backed up)
Iteration 4:   log likelihood = -2367627.1  (backed up)
Iteration 5:   log likelihood = -1872727.4  (backed up)
Iteration 6:   log likelihood = -761507.35
Iteration 7:   log likelihood = -247554.88
Iteration 8:   log likelihood = -8975.8025
Iteration 9:   log likelihood = -1531.4589
Iteration 10:  log likelihood = -1218.5439
Iteration 11:  log likelihood = -1214.7338
```

```
Iteration 12:  log likelihood = -1214.7256
Iteration 13:  log likelihood = -1214.7256

Poisson regression                                Number of obs     =        320
                                                  LR chi2(143)      = 3152643.19
                                                  Prob > chi2       =     0.0000
Log likelihood = -1214.7256                       Pseudo R2         =     0.9992

------------------------------------------------------------------------------
       delta |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         men |
          2  |   1.098607   .0065035   168.93   0.000      1.08586    1.111353
          3  |   1.609508   .0072405   222.29   0.000     1.595317    1.623699
          4  |   1.946142   .0092729   209.87   0.000     1.927968    1.964317
             |
       women |
          2  |  -.0002545   .0049653    -0.05   0.959    -.0099862    .0094773
          3  |   -.000379   .0066298    -0.06   0.954    -.0133731    .0126152
          4  |  -.0004467   .0090516    -0.05   0.961    -.0181874    .0172941
             |
 [ ...output suppressed... ]
             |
        dist |
          1  |  -.2876985   .0054818   -52.48   0.000    -.2984426   -.2769545
          2  |  -.5755819   .0051842  -111.03   0.000    -.5857428    -.565421
          3  |  -.8637119   .0058504  -147.63   0.000    -.8751784   -.8522454
             |
        1.fh |  -.0001772   .0084316    -0.02   0.983    -.0167029    .0163484
       _cons |   9.433832   .0064976  1451.89   0.000     9.421097    9.446568
------------------------------------------------------------------------------
```

A.1.4 Hypergamy effect in wave-specific marriages and cumulative marriages, by wave

	Delta			Cumulative		
Wave	beta	se	p	beta	se	p
1	-0.00003	0.00935	0.99718	-0.00003	0.00935	0.99718
2	-0.00027	0.03809	0.99427	-0.01119	0.00906	0.21660
3	0.00044	0.04815	0.99268	-0.01840	0.00889	0.03842
4	0.00003	0.05802	0.99957	-0.02368	0.00877	0.00696
5	0.00006	0.06786	0.99933	-0.02780	0.00869	0.00139
6	-0.00362	0.07781	0.96285	-0.03120	0.00863	0.00030
7	-0.00297	0.08770	0.97301	-0.03406	0.00859	0.00007
8	-0.00516	0.09773	0.95788	-0.03652	0.00855	0.00002
9	-0.00362	0.10783	0.97322	-0.03867	0.00852	0.00001
10	0.00098	0.11783	0.99334	-0.04054	0.00849	0.00000
11	-0.00328	0.12816	0.97957	-0.04223	0.00847	0.00000

```
12   0.00121   0.13871   0.99302  -0.04375   0.00845   0.00000
13   0.00026   0.14919   0.99862  -0.04513   0.00843   0.00000
14  -0.00312   0.16008   0.98443  -0.04641   0.00842   0.00000
15  -0.01684   0.17105   0.92160  -0.04761   0.00841   0.00000
16   0.00644   0.18105   0.97162  -0.04869   0.00840   0.00000
17  -0.01463   0.19288   0.93955  -0.04972   0.00839   0.00000
18  -0.01095   0.20373   0.95715  -0.05068   0.00838   0.00000
19  -0.02023   0.21452   0.92488  -0.05159   0.00837   0.00000
20  -0.01219   0.22683   0.95715  -0.05243   0.00836   0.00000
```

A.2 Association models and simulations

Association models such as distance or quasi-symmetry attempt to capture the association between the male and female marginals with a restricted set of parameters (i.e., more parsimoniously than the saturated model which predicts perfectly). How well do these models capture the association in the simulated data when there is a hypergamy parameter, across the various simulations? We consider fit for:

- independence (row and column parameters only)
- distance
- crossings
- quasi-symmetry
- the saturated model

with a parameter for asymmetry (aliased in the saturated model), for simulations with hypergamy preferences between 0 and 1 in steps of 0.25. We consider two tests: BIC, and the LR test versus the saturated model.

```
     +-----------------------------------------------------+
     |               sim    fh   igof   dgof   xgof   qgof |
     |-----------------------------------------------------|
  1. |      Independence     0    . 5   0.65    . 5   0.46 |
  2. |      Independence   .25    .     0.80    . 9   0.38 |
  3. |      Independence    .5    .     0.33    .     0.81 |
  4. |      Independence   .75    .      .      .     0.39 |
  5. |      Independence     1    .      .      .     0.46 |
     |-----------------------------------------------------|
  6. |       Competition     0    .     0.66    .     0.75 |
  7. |       Competition   .25    .      .      .      .   |
  8. |       Competition    .5    .      .      .      .   |
  9. |       Competition   .75    .      .      .      .   |
 10. |       Competition     1    .      .      .      .   |
     |-----------------------------------------------------|
 11. |         Homophily     0    .      .      .      .   |
 12. |         Homophily   .25    .      .      .      . 1 |
 13. |         Homophily    .5    .      .      .     0.32 |
 14. |         Homophily   .75    .      .      .     0.34 |
```

```
15. |        Homophily     1    .      .      .      .   |
    |----------------------------------------------------|
16. |   Time-structure     0    .      .      .      .   |
17. |   Time-structure   .25    .      .      .      .   |
18. |   Time-structure    .5    .      .      .      . 9 |
19. |   Time-structure   .75    .      .      .      . 4 |
20. |   Time-structure     1    .      .      .      . 2 |
    |----------------------------------------------------|
21. | Social-structure     0    .     0.15    .     0.79 |
22. | Social-structure   .25    .     0.25    .      . 4 |
23. | Social-structure    .5    .     0.90    .     0.83 |
24. | Social-structure   .75    .      .      .     0.36 |
25. | Social-structure     1    .      .      .      . 2 |
    +----------------------------------------------------+
```

Looking first at goodness of fit relative to the saturated model, we see that the independence+FH model fits very poorly everywhere except in the independence simulation for FH=0 (and there it fits moderately poorly). The crossings model is similar, with bad fit everwhere except for low FH in the independence simulation. Distance and QS perform better, being preferable to the saturated model in a number in 7/25 models for distance and 12 of 25 for QS.

```
    +----------------------------------------------------------------------------------------------
    |              sim    fh       ibic       dbic       xbic       qbic       sbic         bicmod
    |----------------------------------------------------------------------------------------------
 1. |     Independence     0    260.106   260.4023   268.1608   266.9487   270.9531   Independence
 2. |     Independence   .25   277.9634    259.237   266.3331   267.1175   270.7275       Distance
 3. |     Independence    .5   476.1939   262.3817   352.9701   265.3734   270.4884       Distance
 4. |     Independence   .75   890.2222    276.042   480.7443    266.583   270.2645             QS
 5. |     Independence     1   1408.964   274.2286   725.4441   266.0664   270.0765             QS
    |----------------------------------------------------------------------------------------------
 6. |      Competition     0   1767.738   259.5566   1013.452   265.1974   270.1637       Distance
 7. |      Competition   .25   2795.203   400.4484   1136.241   302.7711   269.9359      Saturated
 8. |      Competition    .5   3320.112   616.1035   1263.077   337.1529   269.6623      Saturated
 9. |      Competition   .75   5193.396   1360.008   1431.358   500.2571   269.4427      Saturated
10. |      Competition     1    11120.6   3054.274   2802.189   742.4145    269.328      Saturated
    |----------------------------------------------------------------------------------------------
11. |        Homophily     0   310922.6   321.0991   104601.6   275.0697   269.9989      Saturated
12. |        Homophily   .25   293157.1   298.1049   96126.66   274.0708   269.9484      Saturated
13. |        Homophily    .5     276065   294.0702   89035.85   266.5513   269.8317             QS
14. |        Homophily   .75   256420.7   319.8618    82086.5   266.2836   269.6795             QS
15. |        Homophily     1   220071.2   318.7644   69555.01      282.3   269.5828      Saturated
    |----------------------------------------------------------------------------------------------
16. |   Time-structure     0   240381.1   8649.161    75816.8   353.3171   270.1278      Saturated
17. |   Time-structure   .25   287027.5   9419.938   89744.59   295.1207   270.1098      Saturated
18. |   Time-structure    .5   326200.4   10661.17   100480.1   269.3514   270.0306             QS
19. |   Time-structure   .75   362678.7   12592.48     114144    270.671   269.9221      Saturated
20. |   Time-structure     1   389190.7   14467.85     124622   272.4056   269.8129      Saturated
```

```
     |---------------------------------------------------------------------------------------------
 21. | Social-structure     0   8807.955   265.2584    4539.38    265.924   270.9947       Distance
 22. | Social-structure   .25   7935.145   263.5725   3887.832   271.6302   270.8062       Distance
 23. | Social-structure    .5   6439.438   258.3769   3201.119   265.4231   270.5913       Distance
 24. | Social-structure   .75   5292.988   279.5011   2810.488   266.9176   270.3918             QS
 25. | Social-structure     1   4254.736   277.6159   2349.266   272.3946   270.2299      Saturated
     +---------------------------------------------------------------------------------------------
```

BIC tells a broadly similar story. Independence is favoured only where it makes sense (independence simulation, FH=0). Distance and QS each are judged best in 6 of the other 24 models, but the saturated model is favoured in 12. In other words, the restricted patterns of association are inadequate about half the time, and more so for higher levels of the FH preference.

For the independence simuation, this is an interesting finding. If there is no FH preference, this is properly modelled by the independence model, confirmed by BIC. However, if the only association mechanism is the hypergamy preference, one would be inclined to expect that independence+FH should fit well here. Instead distance and QS are favoured. In other words, the introduction of a simple asymmetry seems to result in a more complex pattern of association. The same feature is visible in most of the other simulations: the saturated model is often preferred, perhaps more so with higher levels of hypergamy preference. Notably, the homophily and time-structure simulations favour the saturated model, even at FH=0.

Two remarks: first, adding a simple asymmetry seems to result in a more complex pattern of association. Second, homophily and time-structure seem to have complex association even at FH=0, due to the dynamics.

A.3 Modelling ESS: fit

How well do different association models fit the ESS data?

```
     +------------------------------------+
     | country   igof   dgof   xgof   qgof |
     |------------------------------------|
  1. |       1   0.00   0.15   0.69   1.00 |
  2. |       2   0.00   0.00   0.00   0.22 |
  3. |       3   0.00   0.00   0.00   0.70 |
  4. |       4   0.00   0.38   0.00   0.88 |
  5. |       5   0.00   0.00   0.00   0.89 |
     |------------------------------------|
  6. |       6   0.00   0.00   0.00   0.68 |
  7. |       7   0.00   0.05   0.00      . |
  8. |       8   0.00   0.35   0.00   0.58 |
  9. |       9   0.00   0.00   0.00   0.15 |
 10. |      10   0.00   0.49   0.00   0.35 |
     |------------------------------------|
 11. |      11   0.00   0.00   0.00   0.05 |
 12. |      12   0.00   0.00   0.00   0.02 |
 13. |      13   0.00   0.00   0.00   0.13 |
 14. |      14   0.00   0.00   0.00   0.01 |
```

```
15. |      15   0.00   0.03   0.00   0.83 |
    |-------------------------------------|
16. |      16   0.00   0.76   0.00   0.30 |
17. |      17   0.00   0.00   0.00   0.72 |
18. |      18   0.00   0.00   0.00   0.02 |
19. |      19   0.00   0.00   0.00   0.33 |
20. |      20   0.00   0.00   0.00   0.33 |
    |-------------------------------------|
21. |      21   0.00   0.61   0.00   0.85 |
22. |      22   0.00   0.66   0.00   0.61 |
23. |      23   0.00   0.00   0.00   0.85 |
24. |      24   0.03   0.97   0.05   0.42 |
25. |      25   0.00   0.00   0.00   0.41 |
    |-------------------------------------|
26. |      26   0.00   0.00   0.00   0.74 |
27. |      27   0.00   0.00   0.00   0.09 |
28. |      28   0.00   0.00   0.00   0.81 |
29. |      29   0.00   0.03   0.00   0.59 |
30. |      30   0.00   0.66   0.00   0.64 |
    |-------------------------------------|
31. |      31   0.00   0.01   0.00   0.93 |
32. |      32   0.00   0.00   0.00   0.94 |
33. |      33   0.00   0.48   0.00   0.98 |
34. |      34   0.00   0.25   0.00   0.93 |
35. |      35   0.00   0.02   0.00   0.01 |
    |-------------------------------------|
36. |      36   0.00   0.54   0.67      . |
    +-------------------------------------+
```

```
    +------------------------------------------------------------------------------------------+
    | country       ibic       dbic       xbic       qbic       sbic     minbic         bicmod |
    |------------------------------------------------------------------------------------------|
 1. |       1   562.8256   511.4749   526.5027    547.345   551.6077   511.4749       Distance |
 2. |       2   1280.064    746.838   898.7849   751.2217   763.2118    746.838       Distance |
 3. |       3   2314.339   947.1699    1449.76   925.5885   971.2773   925.5885             QS |
 4. |       4   1646.285   679.3859   1223.437   729.4262   745.7615   679.3859       Distance |
 5. |       5   2186.063   880.1889   1265.897   874.6835   919.0762   874.6835             QS |
    |------------------------------------------------------------------------------------------|
 6. |       6   1259.433   749.7919   881.0875   776.6011   810.6279   749.7919       Distance |
 7. |       7   928.7445   640.5676   724.1208   677.1447   677.1447   640.5676       Distance |
 8. |       8   1929.398   784.5774   1085.264   844.2591   873.7538   784.5774       Distance |
 9. |       9   1531.736   855.2949   1138.263     849.37    876.663     849.37             QS |
10. |      10   981.1846   691.7349   914.9291   736.6401   749.7192   691.7349       Distance |
    |------------------------------------------------------------------------------------------|
11. |      11   2537.548   948.1536   1464.013    930.139   963.6594    930.139             QS |
12. |      12   1645.423   876.8029   1200.683   881.0716   901.1255   876.8029       Distance |
```

```
 13. |       13   1943.432   969.8442    1242.51   926.8728   964.2173   926.8728             QS |
 14. |       14   1832.756   924.3208   1147.812   974.2914   1003.018   924.3208       Distance |
 15. |       15   2281.903   832.0522    1346.86   875.0264   918.5381   832.0522       Distance |
     |-------------------------------------------------------------------------------------------|
 16. |       16   771.3423   560.7589   605.6583   607.5621   610.7843   560.7589       Distance |
 17. |       17   1726.921   723.3979   1112.925   757.7154   769.6356   723.3979       Distance |
 18. |       18   2160.087    914.295   1545.973   960.7098   991.8074    914.295       Distance |
 19. |       19   1969.702   912.0849   1504.929    909.736   947.3417    909.736             QS |
 20. |       20   629.8578   578.1218   595.1238   599.6061   605.9771   578.1218       Distance |
     |-------------------------------------------------------------------------------------------|
 21. |       21   727.1707   574.9028    699.286   622.6119   626.9083   574.9028       Distance |
 22. |       22   746.8091   600.2107   737.3238   641.5223   649.1426   600.2107       Distance |
 23. |       23   974.2497   632.9633   730.4144   646.1884   669.6323   632.9633       Distance |
 24. |       24   498.4127   517.8903   534.7498   560.9384    564.551   498.4127   Independence |
 25. |       25   2602.309   885.9625   1507.976   924.0046   966.3445   885.9625       Distance |
     |-------------------------------------------------------------------------------------------|
 26. |       26   1438.913   789.7984   942.4892   781.7081    801.149   781.7081             QS |
 27. |       27   1246.255   679.8405   921.4111   687.8567   691.6807   679.8405       Distance |
 28. |       28   2346.949   906.6887    1621.46   886.9941   930.2557   886.9941             QS |
 29. |       29   684.5633    669.968   717.4544   716.5854   738.7454    669.968       Distance |
 30. |       30   923.7282    677.389   811.8058   722.7408   737.7679    677.389       Distance |
     |-------------------------------------------------------------------------------------------|
 31. |       31   1179.344   822.0948   999.8318   860.2596   905.3884   822.0948       Distance |
 32. |       32   1691.828   735.5416   1104.374   747.6829   760.4804   735.5416       Distance |
 33. |       33   1039.892   612.6021   750.1014   650.8402   659.4267   612.6021       Distance |
 34. |       34   1104.231   644.2053   697.2482   684.6597   705.4229   644.2053       Distance |
 35. |       35    913.905   688.8611   860.2595   723.7098   728.7034   688.8611       Distance |
     |-------------------------------------------------------------------------------------------|
 36. |       36   579.1015   517.2425   522.4376    557.395    557.395   517.2425       Distance |
     +-------------------------------------------------------------------------------------------+
```