

Assortative mating without assortative preference

Yu Xie^{a,b,1}, Siwei Cheng^a, and Xiang Zhou^a

^aInstitute for Social Research and Department of Sociology, University of Michigan, Ann Arbor, MI 48104; and ^bCenter for Social Research, Peking University, Beijing 100871, China

Contributed by Yu Xie, March 10, 2015 (sent for review September 8, 2014; reviewed by Scott E. Page and Zhenchao Qian)

Assortative mating—marriage of a man and a woman with similar social characteristics—is a commonly observed phenomenon. In the existing literature in both sociology and economics, this phenomenon has mainly been attributed to individuals’ conscious preferences for assortative mating. In this paper, we show that patterns of assortative mating may arise from another structural source even if individuals do not have assortative preferences or possess complementary attributes: dynamic processes of marriages in a closed system. For a given cohort of youth in a finite population, as the percentage of married persons increases, unmarried persons who newly enter marriage are systematically different from those who married earlier, giving rise to the phenomenon of assortative mating. We use microsimulation methods to illustrate this dynamic process, using first the conventional deterministic Gale–Shapley model, then a probabilistic Gale–Shapley model, and then two versions of the encounter mating model.

assortative mating | structural effect | Gale–Shapley model | encounter mating model | composition heterogeneity

We take assortative mating to mean the tendency of men and women who marry to have similar social characteristics (i.e., homogamy), although, broadly speaking, assortative mating may refer to any nonrandom mixing of spousal characteristics. Assortative mating is a commonly observed phenomenon in human societies (1–10). In the existing literature in both sociology and economics, this phenomenon has mainly been attributed to individuals’ conscious preferences for homogamy. On the one hand, sociologists have argued that people with similar attributes are likely to have similar values, interests, tastes, economic resources, and lifestyles, and individuals often value these similarities when selecting marriage partners. On the other hand, economists have explained assortative mating by stressing complementarities of marriage partners’ attributes. For example, Becker (11) showed that, under a market equilibrium, marriage partners are likely to be associated in traits that are complementary in producing household goods. In other words, people with similar attributes tend to enter marriage as long as these attributes reinforce each other in improving family welfare.

Sociologists have long been aware that assortative mating, as in friendship formation, may also result from structural forces that sort individuals into separate social contexts by similarity in attributes (2, 4, 12–17). In a modern society, it is safe to assume that experiences and social activities leading to marriage, such as dating, require that a man and a woman get to know each other and interact. This is true even when dating takes place in cyberspace. Social structure may induce or constrain assortative mating because it defines the social spaces in which such interactions take place. In the sociology literature, social structure is said to define the exposure to mating opportunities (i.e., potential persons with whom to interact). When persons with different attributes are segregated into different social contexts, assortative mating ensues even when they do not prefer to marry persons with similar attributes, because they do not have a chance to meet, or are not exposed to, persons with dissimilar attributes.

In this paper, we show that patterns of assortative mating may arise from another structural source even if individuals do not have assortative preferences or possess complementary attributes:

dynamic processes of marriages in a closed population. For a given cohort of youth in a finite population, as the percentage of married persons increases, unmarried persons who newly enter marriage are systematically different from those who married earlier, giving rise to the phenomenon of assortative mating. To put it in the causal framework of recent literature, a dynamic process could lead to changes in the population composition of the untreated pool (i.e., the pool of unmarried individuals) when treatment propensity (i.e., the propensity of getting married) is systematically associated with treatment effects (i.e., the utilities of the couple getting married) (18). We use microsimulation methods to illustrate this dynamic process, using first the conventional deterministic Gale–Shapley model, then a probabilistic model, and then two versions of the encounter mating model.

Assumptions and Models

We start with a hypothetical finite population with a sex ratio of one, including N males and N females. We assume that all marriages are heterosexual. To preserve parsimony and for convenience, we assume that individuals consider a unidimensional attribute of potential marriage partners: mate desirability. To demonstrate the influence of changing population structure alone in generating assortative mating, we assume that preference for mate desirability is invariant across all members of the opposite sex. Let X_m and X_f represent mate desirability, respectively, for males and females. In this hypothetical setup, a person’s desirability as a marriage partner, X_m or X_f , affects the utility from marrying him/her independently. That is, the contribution of the desirability of a potential partner to the overall utility for the decision maker does not depend on the attributes of the decision maker. This is what we mean by “no assortativeness” in individuals’ preference. Below, we discuss three models with which to characterize the dynamic process of marriage formation. In all three models, we treat monogamous marriage as an absorbing state and do not allow polygamy, divorce, or remarriage.

The Gale–Shapley Model. The Gale–Shapley model was originally introduced to solve the problem of stable matching (19). It proposes an iterative algorithm that consists of a number of

Significance

Assortative mating, the tendency of men and women who marry to have similar social characteristics, is a commonly observed phenomenon in human societies. This study shows that assortative mating could result from structural causes independent of human agents’ preference, because unmarried persons who newly enter marriage are systematically different from those who married earlier. Thus, assortative mating could result from selection, not by rational choice, but by the dynamics of social structures.

Author contributions: Y.X., S.C., and X.Z. designed research; S.C. and X.Z. performed research; and Y.X., S.C., and X.Z. wrote the paper.

Reviewers: S.E.P., University of Michigan; and Z.Q., Ohio State University.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. Email: yuxie@umich.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1504811112/-DCSupplemental.

rounds. In the first round, each male proposes to his preferred female. Each female who receives more than one proposal rejects all but her preferred choice among those who have proposed to her. In the original algorithm proposed by Gale and Shapley (19), she “does not accept him yet, but keeps him on a string to allow for the possibility that someone better may come along later.” In other words, they are provisionally engaged but not necessarily going to marry. Here, the Gale–Shapley algorithm requires that the males but not the females have full information about their potential partners. We slightly modify the process to assume that both sexes have full information about all potential partners. Thus, if the female’s preferred suitor happens to be her preferred choice among all available males, they instantly get married and exit the marriageable pool. In this modification, both sexes, instead of only the males, act with full information at each round; thus, we expedite the matching process without altering the final results. In each subsequent round, all unengaged males, that is, those who failed to be matched during the last round, propose to their next choices. As in the first round, each female receiving proposals rejects all but her preferred suitor among those unmarried. If he is also her preferred choice among all available males, they get married and exit the marriageable pool; otherwise, they return to the marriageable pool. This process continues until everyone is married. It can be proved that as long as each person has a fixed preference ranking of potential spouses, this algorithm is guaranteed to produce complete and stable matching. That is, everyone is married (assuming equal numbers of males and females), and there are no two people of opposite sexes who are not married to one another but would both prefer marrying each other to remaining with their current spouses.

A crucial assumption that enables the Gale–Shapley algorithm to produce stable marriages is that each person has a fixed preference ranking of potential spouses. The preference ranking, in turn, can be derived from a utility model. Here, we assume that all individuals of one sex assign utility to each potential mate of the opposite sex based on the person’s desirability. Incorporating our assumption of universal preference, we can write the i th male’s utility for marrying the j th female as a univariate function of X_j^f :

$$U_{ij}^m = \alpha_0 + \alpha_1 X_j^f, \tag{1a}$$

where X_j^f denotes the j th female’s attribute. Similarly, U_{ij}^f gives the j th female’s utility from marrying the i th male:

$$U_{ij}^f = \beta_0 + \beta_1 X_i^m, \tag{1b}$$

where X_i^m denotes the i th male’s attribute. Here, linear function is used without loss of generality; the functional form that links utility with mate desirability will not affect our major findings.

Combined with the above specification for utility, the Gale–Shapley algorithm also predicts assortative marriages. In fact, as long as X_m and X_f are truly continuous, that is, there are no ties in X_m and X_f , marriage partners will be matched to each other exactly according to their relative ranks within their respective sexes. To see this, consider the sequential process. At the beginning, all males prefer the most desirable female, and all females prefer the most desirable male, resulting in a match between the top-ranked male and the top-ranked female. Because marriage is an event leading to an absorbing state, a matched pair is immediately removed from the marriageable pool. Thus, in the second round, marriage occurs between the second-ranked male and the second-ranked female. Consequently, in the r th round of marriage, male of rank r according to X_i^m is matched to female of rank r according to X_j^f . This process is replicated until everyone in the population is married. Because

this model does not involve any randomness in utility based on mate desirability, we call this model the deterministic Gale–Shapley model (i.e., model 1).

The Probabilistic Gale–Shapley Model. In our second model, we modify the conventional Gale–Shapley model by adding a noise component to the male’s and female’s utility from marriage (Eq. 1). Thus, we rewrite the i th male’s utility from marrying the j th female at time t as

$$U_{ijt}^m = \alpha_0 + \alpha_1 X_j^f + \epsilon_{ijt} \tag{2a}$$

and the j th female’s utility from marrying the i th male at time t as

$$U_{ijt}^f = \beta_0 + \beta_1 X_i^m + \eta_{ijt}. \tag{2b}$$

It is evident that Eqs. 2a and 2b are of the same form as the random utility model widely used for studying discrete choice in econometrics (20). Following convention for the discrete-choice model, we assume that ϵ_{ijt} and η_{ijt} are independent and identically distributed as a type I extreme value distribution.

For reasons to be given later, the choice set at time t consists of all unmarried individuals of the opposite sex and the option of remaining single. For males, we denote the state of being single by $j=0$. For females, we denote the state of being single by $i=0$. For the state of being single, we normalize the systematic part of the utility function of Eq. 2 at zero, yielding

$$U_{i0t}^m = \epsilon_{i0t} \text{ for all males,} \tag{2c}$$

$$U_{0jt}^f = \eta_{0jt} \text{ for all females.} \tag{2d}$$

With the utility model thus modified by the time-varying noise component, we again apply the same Gale–Shapley matching algorithm as was discussed earlier. In each round, each male makes a marriage proposal to his most preferred female or prefers to remain single, and the female will accept the proposal if and only if this male is her most desired male and she considers marriage to him preferable to being single. That is, a marriage occurs if the preferences between a male and a female are reciprocal. However, owing to the noise component, there is no guarantee that exactly one pair will be matched in each round. Because this model adds randomness to the first model, we call it the probabilistic Gale–Shapley model (also called model 2).

The Encounter Mating Model. The Gale–Shapley model, in either its original form or our modified form, assumes that some or all individuals have full information about the attributes of all potential marriage partners in the entire population. Such an assumption is hardly true in reality, because marriage possibility is conditional on exposure to and encounter of potential partners. As an alternative to the two Gale–Shapley algorithm-based models, we learn from models in biological science (21) and propose the third model: the encounter mating model (also called model 3A).

The encounter mating model also breaks the matching process into consecutive rounds, with each round comprising two steps. In the first step, a randomly selected pair of male i and female j encounter one another. In the second step, male i evaluates the desirability of female j , and female j simultaneously evaluates the desirability of male i . A marriage occurs only when both potential partners are satisfied with each other, akin to Logan’s two-sided logit model (22, 23). To implement this idea, we keep the same utility model as in the probabilistic Gale–Shapley model (i.e., Eq. 2). One important difference between the encounter mating model and the probabilistic Gale–Shapley model, however, pertains to the size of choice set in each round. Whereas individuals compare all

available marriage partners and the state of being single in the probabilistic Gale–Shapley model, they are faced with a binary choice in the encounter mating model: marry or not marry. As before, to model this binary choice, we normalize the systematic part of the utility function for being single to zero (Eqs. 2c and 2d). As a result, the probability that the i th male is willing to marry the j th female given their encounter can be expressed simply as a binary logit model (24):

$$P_{ijt}^m = \Pr(U_{ijt}^m > U_{i0t}^m) = \frac{e^{\alpha_0 + \alpha_1 X_j^f}}{1 + e^{\alpha_0 + \alpha_1 X_j^f}}. \quad [3a]$$

Similarly, P_{ijt}^f gives the probability that the j th female is willing to marry the i th male given their encounter:

$$P_{ijt}^f = \Pr(U_{ijt}^f > U_{j0t}^f) = \frac{e^{\beta_0 + \beta_1 X_i^m}}{1 + e^{\beta_0 + \beta_1 X_i^m}}. \quad [3b]$$

Obviously, the intercept terms α_0 , β_0 affect the marriage probability in Eq. 3. In other words, they represent the utility gain from getting married relative to staying single that is unaffected by the potential partner's attributes. The higher the intercept terms α_0 , β_0 , the higher utility of getting married relative to staying single, respectively, for men and women.

By construction the two potential partners' decisions are assumed to be independent, and the joint probability that the i th male and the j th female marry given their encounter should be the product of P_{ijt}^m and P_{ijt}^f :

$$P_{ijt} = P_{ijt}^m P_{ijt}^f. \quad [4]$$

For our simulation exercise to be shown later in the paper, we further characterize the dynamic features of this model with two conditions. First, an encounter between a male and a female is assumed to be random and sequential. That is, a random pair of potential mates encounter first and then decide whether to marry before the next random encounter occurs. Second, as before, marriage is considered an absorbing state, and thus we do not allow polygamy, divorce, or remarriage. In our implementation, a marriage removes the two marrying partners from the marriageable pool for subsequent rounds. If an encountered couple do not marry, however, they return to the unmarried pool. We then start over to let another pair (i', j') encounter and simulate their marital decision. We iterate this process until all people in the hypothetical population are married.

Waiting Cost. In extending the deterministic Gale–Shapley model to the probabilistic Gale–Shapley model and the encounter mating model, we aimed to add more realism. However, the two extensions

still fall short of realism because they ignore the cost of waiting. If waiting is costless, the rational behavior under the encounter mating model would be to reject until a highly desirable (i.e., optimal or nearly optimal) match shows up. This is impossible in reality, of course, because it might take unrealistically too many iterations for such a match to occur. In actual marriage markets, young persons wishing to get married have limited time in which to consider a limited number of potential marriage partners in a very large population. Given the time constraint, waiting is costly, because each unsuccessful encounter shortens the remaining time and probably the quality of the marriageable pool. As we will illustrate later, another reason why waiting is costly is that our model predicts that the general quality of available mates declines with time. This is particularly true if one objective of marriage is childbearing, because fertile years are limited, especially for women.

Hence, it is essential to a realistic model of marriage that individuals incur a cost for waiting to marry, comparable to our earlier modification that they do not have full information about all potential marriage partners. Thus, to our baseline encounter mating model we now propose an a priori time-increasing cost of being single (C_t) for every unmarried individual. Because we wish to keep time scale-free, we parameterize the cost (C_t) as an increasing function of the proportion of individuals of a given sex who have married. In a closed system, C_t is determined by the number of persons who have exited the state of being single by time t , denoted as n_t . In other words, individuals become increasingly impatient as they observe an increasing number of peers getting married in the population. However, we assume that individuals do not have more information about the remaining population beyond n_t . In reality, the cost of waiting may not be equal between men and women owing to sex differences in the age pattern of fecundity. Thus, the specification for C_t could be sex-specific. However, for simplicity, simulation in this paper assumes the same parameterization for men and women. More specifically, we parameterize this cost of being single as

$$C_t = c \cdot n_t,$$

where c takes a positive value. This amounts to a decrease in the systematic part of utility of being single with time:

$$U_{i0t}^m = -c \cdot n_t + \epsilon_{i0t} \text{ for all males,} \quad [5a]$$

$$U_{j0t}^f = -c \cdot n_t + \eta_{j0t} \text{ for all females.} \quad [5b]$$

For the encounter mating model, the incorporation of the cost factor amounts to increasing the intercept terms α_0 , β_0 in the utility models (Eq. 2) so as to increase the overall probability of realization in the binary logit models (Eq. 3). Therefore, the binary logit models of Eqs. 3a and 3b translate into

$$P_{ijt}^m = \Pr(U_{ijt}^m > U_{i0t}^m) = \frac{e^{\alpha_0 + c \cdot n_t + \alpha_1 X_j^f}}{1 + e^{\alpha_0 + c \cdot n_t + \alpha_1 X_j^f}}, \quad [6a]$$

$$P_{ijt}^f = \Pr(U_{ijt}^f > U_{j0t}^f) = \frac{e^{\beta_0 + c \cdot n_t + \beta_1 X_i^m}}{1 + e^{\beta_0 + c \cdot n_t + \beta_1 X_i^m}}. \quad [6b]$$

This model explicitly compares the utility of marrying a potential partner to the utility of being single, which declines with time. Henceforth, we call this model the extended encounter mating model (also called model 3B).

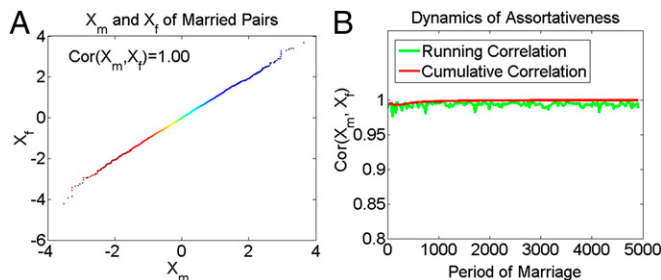


Fig. 1. (A and B) Assortative mating in the deterministic Gale–Shapley model. In A, gradations of color are used to represent marriages formed in different time periods. Specifically, bluer points correspond to earlier marriages, and redder points correspond to later marriages.

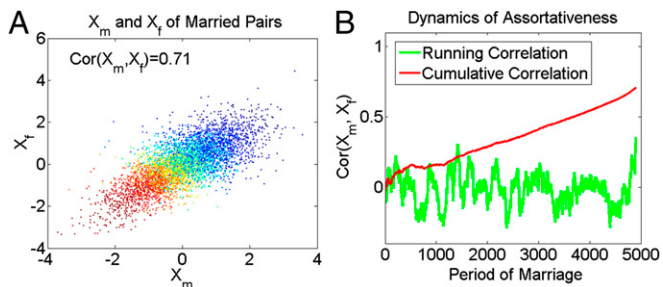


Fig. 2. (A and B) Assortative mating in the probabilistic Gale–Shapley model. In A, gradations of color are used to represent marriages formed in different time periods. Specifically, bluer points correspond to earlier marriages, and redder points correspond to later marriages.

Microlevel Simulation

We demonstrate the dynamic process of assortative mating using microlevel simulation (or an agent-based model). In our simulation, a hypothetical population is composed of 5,000 males and 5,000 females (i.e., $n = 5,000$). Individual characteristics, X_m^i and X_f^i , are assumed to follow standard normal distributions. We set $\alpha_0 = \beta_0 = -1$, $\alpha_1 = \beta_1 = 2$, and $c = 0.0005$. We simulate the dynamic matching process according to the deterministic Gale–Shapley model, the probabilistic Gale–Shapley model, and the encounter mating model, as described in the previous section. For each of these models, we track the attributes of married couples at each time point and present them in two time series, one for each sex. Time points at which no successful marriage occurs are removed, because they provide no information to our study.

Before we discuss the results, we define several quantities of interest. The first quantity is the correlation (ρ) between the husband’s and wife’s attributes, X_m and X_f , across all of the matched pairs. The correlation is a simple, crude measure of assortative mating. The second quantity is the running correlation ($\rho_{n_t}^r$) between X_{m,n_t} and X_{f,n_t} within an interval of 100 pairs matched adjacently in time, that is,

$$\rho_{n_t}^r = \text{corr}(X_{m_k}, X_{f_k} \mid k \in [n_t, n_t + 99]).$$

The third quantity is the cumulative correlation ($\rho_{n_t}^c$) between X_m and X_f for the matched pairs up to n_t , that is,

$$\rho_{n_t}^c = \text{corr}(X_{m_k}, X_{f_k} \mid k \in [1, n_t]).$$

Obviously, $\rho_{n_t}^c$, the cumulative correlation up to the last marriage, is the same as the correlation for the overall population. The fourth quantity is the cumulative correlation ($\widetilde{\rho}_{n_t}^c$) between X_m and X_f that is encountered up to n_t . Because it takes a very large number of encounters for a match to occur, for computational easiness we randomly select one of the encountered pairs for inclusion in the calculation of this quantity between two consecutive matches (n_{t-1} and n_t), that is,

$$\widetilde{\rho}_{n_t}^c = \text{corr}(X_{m_k, \text{encounter}}, X_{f_k, \text{encounter}} \mid k \in [1, n_t]).$$

The fifth quantity is the running averages (X_m^* and X_f^*) of the attributes of the newly married X_m and X_f , within an interval of 100 pairs matched adjacently in time, that is,

$$X_{m,n_t}^* = \frac{\sum_{n_t}^{n_t+99} X_m}{100},$$

$$X_{f,n_t}^* = \frac{\sum_{n_t}^{n_t+99} X_f}{100}.$$

The sixth quantity is the average time to marriage for a certain group, measured as the average number of rounds before one gets married.

In Fig. 1, we present the assortative mating results produced by the deterministic Gale–Shapley model (model 1). Fig. 1A is a scatterplot of X_m and X_f , showing that males and females in marriages are perfectly matched by their respective ranks on X_m and X_f . The correlation between X_m and X_f is 1.00. The colors of the dots represent time of marriage, with bluer points corresponding to earlier marriages and redder points corresponding to later marriages. The results clearly reveal that individuals with higher X_m and X_f scores are married earlier, and those with lower scores are married later. Fig. 1B is the running correlation and cumulative correlation by period of marriage. It suggests that the running correlation between the male’s and female’s attribute remains high and fairly stable, even when we examine it in small time windows.

In Fig. 2, we present analogously the assortative mating resulting from the probabilistic Gale–Shapley model (model 2). Fig. 2A is a scatterplot of X_m and X_f . The overall correlation between X_m and X_f is 0.71, indicating a high assortativeness in marriage. Again, the colors of the dots show that individuals with higher X_m and X_f scores are married earlier, and those with lower scores are married later. Fig. 2B suggests that the addition of a probabilistic component in model 2 shrinks the running correlation between the male’s and female’s attribute to almost zero when we examine assortativeness in small time intervals. The cumulative correlation, however, increases as more pairs are accumulated over time. This pattern suggests that assortativeness of attributes between husbands and wives mainly results from time selectivity of persons entering marriage (with persons of higher attributes entering marriage earlier), but it disappears if we control for waiting time to marriage.

As we stated earlier, we consider the encounter mating model a more realistic model for marriage. We present simulation results for this model, in both the baseline version (model 3A) and the extended version with waiting cost (model 3B). We display the results from the baseline encounter mating model (model 3A) in Fig. 3 in four panels. Fig. 3A and B show the overall correlation at 0.53, indicating a much smaller, but more plausible, degree of assortativeness than in the previous two models based on the Gale–Shapley algorithm. In Fig. 3B, the cumulative correlation among the matched pairs increases gradually as more individuals are married (red line). A large part of this increase is

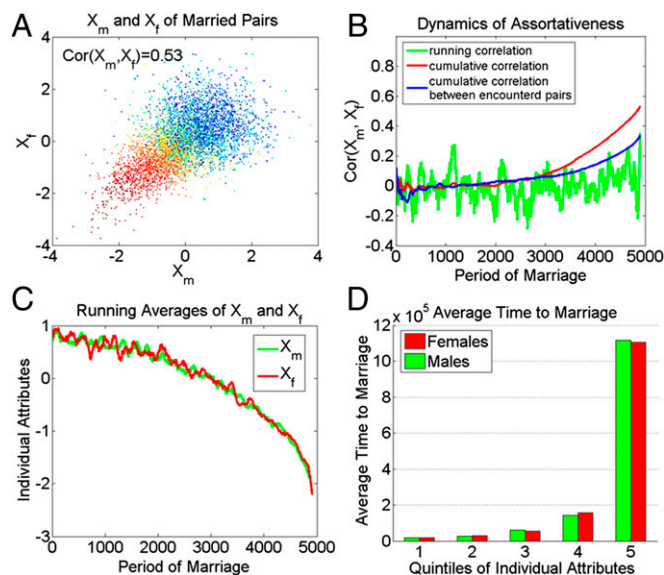


Fig. 3. (A–D) Assortative mating in the encounter mating model. In A, gradations of color are used to represent marriages formed in different time periods. Specifically, bluer points correspond to earlier marriages, and redder points correspond to later marriages.

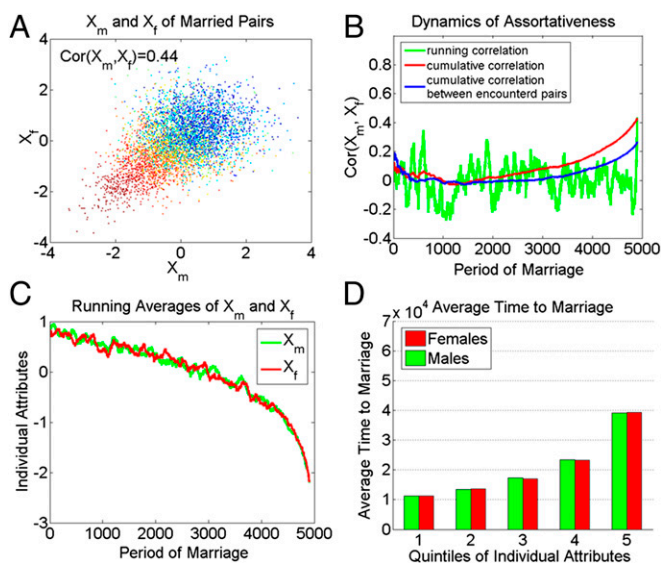


Fig. 4. (A–D) Assortative mating in the encounter mating model with increasing cost of being single. In A, gradations of color are used to represent marriages formed in different time periods. Specifically, bluer points correspond to earlier marriages, and redder points correspond to later marriages.

driven by an increase in homogeneity in the unmarried pool, which is reflected in the increasing assortativeness in X_m and X_f among the encountered pairs (blue line). Qualitatively speaking, the results are similar to those from model 2, in that the running correlation stays close to zero (except for the end), but the cumulative correlation rises with time, particularly at the end. Fig. 3C shows the running average of male's and female's attributes by time of marriage. It reveals steady declines in the scores of the attributes for husbands and wives over time. Fig. 3D shows the average time to marriage by quintiles of individual attributes. Consistent with our proposition of a sequential selection of available population, individuals ranked in lower quintiles on attributes are married much later than those ranked in higher quintiles.

Fig. 4 shows the results from the extended encounter mating model (model 3B). The pattern is qualitatively similar to that of the simpler encounter mating model, model 3A. However, a notable finding is that the inclusion of an increasing cost of being single with time further dilutes assortativeness, leading to an

overall correlation of 0.44. The reason is that a rising waiting cost causes the entry into marriage to be less selective in spouses' attribute attractiveness, which in turn reduces assortativeness owing to compositional changes. For the same reason, shown in Fig. 4D, differences in average marriage time across the five quintiles are less pronounced in the extended encounter mating model than in the baseline encounter mating model.

Our results are not sensitive to the choice of our measure using the correlation coefficient. In the sociological literature on assortative mating, frequency tables cross-classifying categorical attributes of husbands and wives are typically analyzed with log-linear models. Table S1 displays four such cumulative frequency tables with X_m and X_f discretized into quintiles at four times: 25, 50, 75, and 100% of the population having been married under the encounter mating model with waiting cost. In Table S2, we present key results of a log-linear analysis of the tables, which are consistent with findings revealed in Fig. 4.

Conclusion

In this paper, we show that patterns of assortative mating may arise from a structural source even if individuals do not have assortative preferences or possess complementary attributes: dynamic changes of those waiting for marriage in a closed system. For a given cohort of youth in a finite population, as the percentage of the married increases, unmarried persons who newly enter marriage are systematically different from those who married earlier, giving rise to the phenomenon of assortative mating. We have used microsimulation methods to illustrate this dynamic process, using first the conventional deterministic Gale–Shapley model, then a probabilistic Gale–Shapley model, and then two more realistic versions of the encounter mating model.

Our consideration of more realistic model specifications, such as the addition of a probabilistic element to the conventional Gale–Shapley model, the adoption of the encounter mating model due to the lack of full information, and the addition of the time-increasing waiting cost, all lead to a lessening of assortativeness between husband and wife. In our final model, the extended encounter mating model with waiting cost, the overall correlation in attributes between husband and wife is only 0.44, far below the perfect correlation in the case of the conventional deterministic Gale–Shapley model.

ACKNOWLEDGMENTS. We thank Robert Mare and Christine Schwartz for their comments on an earlier version of the paper. This work was supported by National Institutes of Health Grant R01-HD-074603.

- Kalmijn M (1991) Shifting boundaries: Trends in religious and educational homogeneity. *Am Sociol Rev* 56(6):786–800.
- Mare RD (1991) Five decades of educational assortative mating. *Am Sociol Rev* 56(1):15–32.
- Qian Z (1997) Breaking the racial barriers: Variations in interracial marriage between 1980 and 1990. *Demography* 34(2):263–276.
- Kalmijn M (1998) Inter-marriage and homogamy: Causes, patterns, trends. *Annu Rev Sociol* 24(4):395–421.
- Smits J, Ultee W, Lammers J (1998) Educational homogamy in 65 countries: An explanation of differences in openness using country-level explanatory variables. *Am Sociol Rev* 63(2):264–285.
- Raymo JM, Xie Y (2000) Temporal and regional variation in the strength of educational homogamy. *Am Sociol Rev* 65(5):773–781.
- Schwartz CR, Mare RD (2005) Trends in educational assortative marriage from 1940 to 2003. *Demography* 42(4):621–646.
- Qian Z, Lichter DT (2007) Social boundaries and marital assimilation: Interpreting trends in racial and ethnic intermarriage. *Am Sociol Rev* 72(1):68–94.
- Schwartz CR (2010) Earnings inequality and the changing association between spouses' earnings. *AJS* 115(5):1524–1557.
- Torche F (2010) Educational assortative mating and economic inequality: A comparative analysis of three Latin American countries. *Demography* 47(2):481–502.
- Becker GS (1973) A theory of marriage: Part I. *J Polit Econ* 81(4):813–846.
- Blau PM, Schwartz JE (1984) *Crosscutting Social Circles: Testing a Macrostructural Theory of Intergroup Relations* (Academic, New York).
- Lichter DT, LeClere FB, McLaughlin DK (1991) Local marriage markets and the marital behavior of black and white women. *Am J Sociol* 96(4):843–867.
- Lichter DT, McLaughlin DK, Kephart G, Landry DJ (1992) Race and the retreat from marriage: A shortage of marriageable men? *Am Sociol Rev* 57(6):781–799.
- Lichter DT, Anderson R, Hayward M (1995) Marriage markets and marital choice. *J Fam* 16(4):412–431.
- McPherson JM, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. *Annu Rev Sociol* 27:415–444.
- Zeng Z, Xie Y (2008) A preference-opportunity-choice framework with applications to intergroup friendship. *AJS* 114(3):615–648.
- Xie Y (2013) Population heterogeneity and causal inference. *Proc Natl Acad Sci USA* 110(16):6262–6268.
- Gale D, Shapley LS (1962) College admissions and the stability of marriage. *Am Math Mon* 69(1):9–14.
- McFadden D (1974) *Frontiers in Econometrics*, ed Zarembka P (Academic, New York), pp 105–142.
- Gimelfarb A (1988) Processes of pair formation leading to assortative mating in biological populations: Encounter-mating model. *Am Nat* 131(6):865–884.
- Logan JA (1996) Opportunity and choice in socially structured labor markets. *Am J Sociol* 102(1):114–160.
- Logan JA, Hoff PD, Newton MA (2008) Two-sided estimation of mate preferences for similarities in age, education, and religion. *J Am Stat Assoc* 103(482):559–569.
- Powers DA, Xie Y (2008) *Statistical Methods for Categorical Data Analysis* (Emerald, Bingley, UK), 2nd Ed.