The Strength of Absent Ties: Social Integration via Online Dating

Josué Ortega and Philipp Hergovich*

October 2, 2017

Abstract

We used to marry people to which we were somehow connected to: friends of friends, schoolmates, neighbours. Since we were more connected to people similar to us, we were likely to marry someone from our own race.

However, online dating has changed this pattern: people who meet online tend to be complete strangers. Given that one-third of modern marriages start online, we investigate theoretically, using random graphs and matching theory, the effects of those previously absent ties in the diversity of modern societies.

We find that when a society benefits from previously absent ties, social integration occurs rapidly, even if the number of partners met online is small. Our findings are consistent with the sharp increase in interracial marriages in the U.S. in the last two decades.

JEL Codes: J12 (marriage), D85 (networks), C78 (matching).

KEYWORDS: social integration, interracial marriage, online dating, matching, social networks, random graphs.

*Ortega: University of Essex, Wivenhoe Park, CO4 3SQ UK. Hergovich: University of Vienna, Oskar Morgenstern Platz 1, A-1090 Austria. We are particularly indebted to Dilip Ravindran for helpful feedback on earlier drafts. We are also grateful to Y. Bramoullé, A. Chaintreau, P.A. Chiappori, A. Clausen, B. Driesen, B. Golub, P. Harless, K. Mazur, H. Moulin, O. Tercieux, and audiences at Columbia University and the CTN Workshop 2017 for their helpful comments. J. Ortega acknowledges the hospitality of Columbia University. Errors are our own. Corresponding author: josue.ortega1@gmail.com.

1 Introduction

In the most cited article on social networks,[1] Granovetter (1973) argued that the most important connections we have may not be our close friends but our acquaintances: people that are not very close to us, either physically or emotionally, help us to relate to groups that we otherwise we would not be linked to. For example, it is from acquaintances that we are more likely to hear about job offers. Those weak ties serve as bridges between our group of close friends and other clustered groups, hence allowing us to connect to the global community in several ways.[2]

Interestingly, the process of how we meet our romantic partners in at least the last hundred years closely resembles this phenomenon. We would probably not marry our best friends, but we are likely to end up marrying a friend of a friend or someone we coincided with in the past. Rosenfeld and Thomas (2012) show how Americans met their partners in the last decades, listed by importance: through mutual friends, in bars, at work, in educational institutions, at church, through their families, or because they became neighbors. This is nothing but the weak ties phenomenon in action.[3]

But in the last two decades, the way how we meet our romantic partners has changed dramatically. Rosenfeld and Thomas (2012) argue that *" the Internet increasingly allows Americans to meet and form relationships with perfect strangers, that is, people with whom they had no previous social tie"*. To this end, they document that in the last decade online dating has become the second most popular way to meet a spouse for Americans (see Figure 1).[4]

Online dating has changed the way people meet their partners not only in America but in many places around the world. As an example, Figure 2 shows one of the author's Facebook friends graph. The yellow triangles reveal previous relationships that started in offline venues. It can easily be

[1] "What are the most-cited publications in the social sciences according to Google?", *LSE Blog*, 12/05/2016.

[2] Although most people find a job via weak ties, it is also the case that weak ties are more numerous. However, the individual value from an additional strong tie is larger than the one from an additional weak tie (Kramarz and Skans, 2014; Gee et al., 2017).

[3] Backstrom and Kleinberg (2014) reinforce the previous point: given the social network of a Facebook user who is in a romantic relationship, the node which has the highest chances to be his romantic partner is, perhaps surprisingly, not the one who has most friends in common with him.

[4] We thank M. Rosenfeld and R. Thomas for allowing us to use their figure.

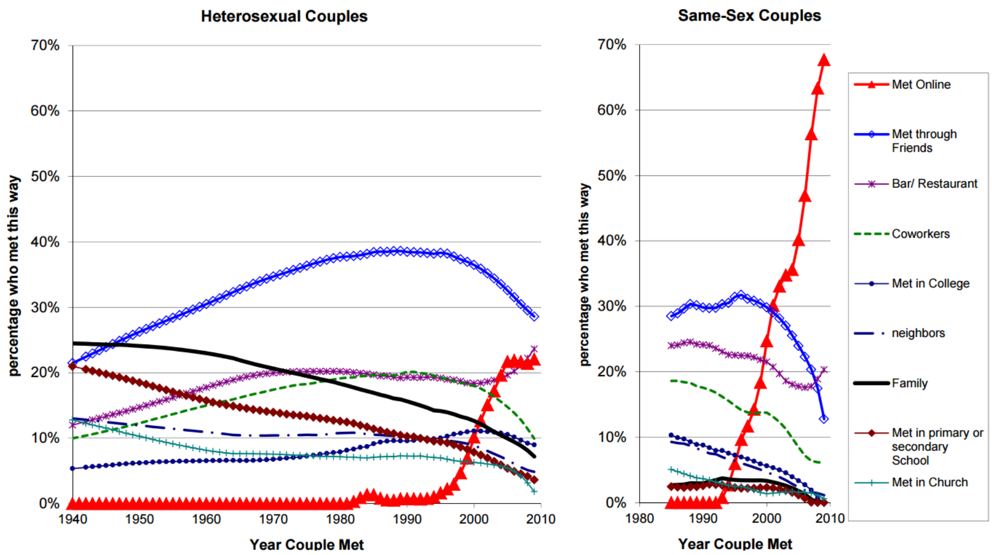


Figure 1: How we met our partners in the last decades.

seen that those ex-partners had several mutual friends with the author; in the corresponding graph, their edge had a high embeddedness in graph-theoretic jargon. In contrast, nodes appearing as red stars represent partners he met through online dating. Those have no contacts in common with him, and thus it is likely that, if it were not for online dating, those persons would have never interacted with him.

Because one-third of modern marriages start online (Cacioppo et al., 2013), and up to 70% of homosexual relationships, the way we match online with potential partners shapes the demography of our communities, in particular its racial diversity. Meeting people outside our social network online can intuitively increase the number of interracial marriages in our societies, which is remarkably low. Only 6.3% and 9% of the total number of marriages are interracial in the U.S. and the U.K., respectively.[5] The low rates of interracial marriage are expected, given that still 50 years ago these were considered illegal in the U.S., until the Supreme Court ruled out anti-miscegenation laws in the famous *Loving vs. Virginia* case.[6]

[5] "Interracial marriage: Who is marrying out", *Pew Research Center*, 12/6/2015; and "What does the 2011 census tell us about inter-ethnic relationships?", *UK Office for National Statistics*, 3/7/2014.

[6] Interracial marriage in the U.S. has increased considerably from 1970, but it is still rare (Arrow, 1998; Kalmijn, 1998; Fryer, 2007; Furtado, 2015). Interracial marriage occurs far less frequently than interfaith marriages (Qian, 1997).

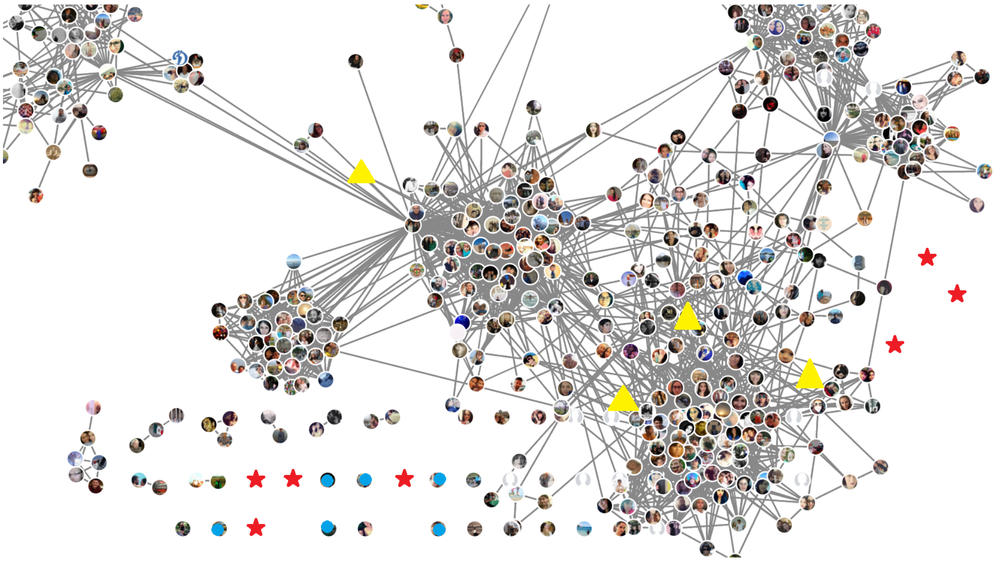


Figure 2: How one of us met his partners in the last decade.

This paper aims at improving our understanding of the impact of online dating on racial diversity in modern societies. In particular, we intend to find how many more interracial marriages, if any, will occur after online dating becomes available in a society. In addition, we are also interested in whether marriages created online are any different from those that existed before.

Understanding the evolution of interracial marriage is an important problem, for intermarriage is widely considered a measure of social distance in our societies (Wong, 2003; Furtado, 2015), just like residential or school segregation. Moreover, the number of interracial marriages in a society has important economic implications. The combined income of an White-Asian modern couple is 14.4% higher than than the combined income of an Asian-Asian marriage, and 18.3% higher than a White-White couple (Wang, 2012). Even when controlling for factors that may influence the intermarrying decision, Gius (2013) finds that all interracial couples not involving Afroamericans have higher combined incomes than a White-White couple. He obtains his findings by analyzing 636,257 observations from the American Community Survey (ACS) in 2010. Fu (2007) finds that White people in Hawaii are 65% more likely to live in poverty if they had married outside their own race. He arrives at this conclusion by also analyzing data from the ACS, but from 2000.[7]

[7]There is a large literature that analyzes the effect of marrying an immigrant. This literature is relevant because often immigrants are from different races than natives. This

Interracial marriage also affects the offspring of couples who engage in it. Duncan and Trejo (2011) find that children of an interracial marriage between a Mexican Latino and an interracial partner enjoy significant human capital advantages over children born from endogamous Mexican marriages in the U.S.[8] Those human capital advantages include a reduction of 50% in the high school dropout rate for male children. However, Pearce-Morris and King (2012) find no differences in the global well-being or behavior problems between children raised in interracial or intraracial households.

1.1 Overview of Results

This article builds a theoretical framework to explain how many more interracial marriages occur after the establishment of online dating. Our model combines non-transferable utility[9] matching *à la* Gale and Shapley (1962) with random graphs, first studied by Erdős and Rényi (1959) and Gilbert (1959). Our theoretical framework is easy to grasp and has an intuitive graphical visualization.

We consider a Gale-Shapley marriage problem, in which agents may belong to different races or communities. All agents from all races are randomly located on the same unit square. Agents want to marry the person who is closest to them, but they can only marry people who they know, i.e. to whom they are connected. As in real life, agents are highly connected with agents of their own race, but only poorly so with people from other races. Again inspired by empirical evidence, we assume that the marriages that occur in our society are those predicted by game-theoretic stability.[10]

literature has consistently found that an immigrant who married a native often has a higher probability of finding employment (Meng and Gregory, 2005; Furtado and Theodoropoulos, 2010; Goel and Lang, 2017). Interestingly, marrying a native increases the probability of employment, but not the perceived salary (Kantarevic, 2004).

[8] Although Hispanic is not a race, Hispanics do not associate with other races. In the 2010 U.S. census, over 19 million of Latinos selected to be of "some other race". See "For many Latinos, racial identity is more culture than color", *New York Times*, 13/1/2012.

[9] Most of the literature studying marriage with matching models uses transferable utility, following the seminal work of Becker (1973, 1974, 1981). A review of that literature appears in Browning et al. (2014). Although our model departs substantially from this literature, we point out similarities with particular papers from this literature in Section 2.

[10] See Banerjee et al. (2013) for the case of marriage, and Hitsch et al. (2010) for the case of romantic relationships that start online.

We introduce online dating in our societies by creating previously absent ties, obtained by a small increase in the probability that any two agents of different races are connected. We compare how many more interracial marriages are formed in the expanded society. We also keep an eye on the characteristics of those newly formed marriages. In particular, we focus on the average distance between partners before and after the introduction of online dating, which we use as a proxy for the strength of marriages in a society.

The graphical interpretation of our model is similar to the one used by the mathematics literature in matching of Poisson point processes (Holroyd et al., 2009; Holroyd, 2011; Amir et al., 2016), from which we borrow useful technical results (see the proof of Proposition 1, which establishes uniqueness of the stable marriage). Our model also roughly resembles the graphical model of residential segregation of Schelling (1969, 1971, 1972). However, unlike the famous Schelling model, our model predicts nearly complete racial integration upon the emergence of online dating, even if the number of partners that individuals meet from newly formed ties is small. Our model also predicts that marriages created in a society with online dating tend to be stronger.

We obtain these predictions by considering randomly generated societies, and analyzing the evolution of their diversity and strength over 10,000 simulations. We use simulations because we prove in Proposition 2 that our welfare measures are not edge-monotonic, meaning they may decrease when societies become more interracially connected. It is perhaps counter-intuitive that adding interracial connections may decrease the number of interracial marriages, but the logic behind this result is simple. An interracial connection may create an interracial marriage at the cost of destroying two existing ones. Similarly, we find that adding an interracial connection may even decrease the total number of marriages (interracial or not) in a society.

We contrast our theoretical results with empirical U.S. data, and find that, as predicted by our model, the number of interracial marriages substantially increases after the popularization of online dating. We discuss how the observed sharp increase cannot be purely due to changes in the composition of the U.S. population. We also present evidence that marriages created online have a lower divorce rate, as predicted by our model.

Our results contribute to clarify the relationship between social networks and interracial marriage. In a related paper, Furtado and Theodoropoulos

(2010) find that immigrants who intermarry have a higher chance of finding employment than those who marry within their own ethnic group. Interestingly, most of this effect is due to the valuable social networks that immigrants gain by marrying a local (and not because an easier chance to get a visa). In their model, intermarriage creates social networks. In ours, social networks generate intermarriage, by creating previously absent ties within races via online dating. The increase in the number of interracial marriages in our model does not require a changes in agents' preferences.

1.2 Structure of the Article

We present our model in Section 2. Section 3 introduces the welfare measures underlying the further analysis. Sections 4 and 5 analyze how these measure change when societies become more connected using theoretical analysis and simulations, respectively.

Section 6 contrasts our model predictions with observed demographic trends from the U.S. Section 7 concludes and provides an outlook on other potential applications of our theoretical framework.

2 Marriages in a Network

2.1 Agents

There are r races or communities, each with n agents. Each agent i is identified by a pair of coordinates $(x_i, y_i) \in [0,1]^2$, that can be understood as measures of agents' social and political opinions,[11] to which we refer as *personality traits*. Both coordinates are drawn uniformly and independently for all agents.[12]

[11] For a real-life representation using a 2-dimensional plane see www.politicalcompass.org. A similar interpretation appears in Chiappori et al. (2012) and in Chiappori et al. (2016), in which the traits include age, education, race, religion, weight or height.

[12] Another way to understand how agents' personality traits are drawn is to consider a Poisson point process (PPP) defined on the unit square with intensity $\lambda = n$. In a PPP the number of agents is not fixed but drawn from a Poisson distribution, although there are n in expectation. In our case, the number of agents is fixed throughout.

Each agent is either male or female. Female agents are plotted as stars and males as dots. Each race is balanced in its ratio between men and women. Each race is assigned a particular color.

2.2 Edges

Agents are connected to others of their own race with probability p: these edges are represented as solid lines and occur independently of each other. Agents are connected to others of different race with probability q: these interracial edges appear as dotted lines and are also independent. We present an illustrative example in Figure 3.

Our model is a generalization of the random graph model (Erdős and Rényi, 1959; Gilbert, 1959; for a textbook reference, see Bollobás, 2001). In our model, there are r random graphs with parameter p and n nodes. Nodes are connected across graphs with probability q. The intuition in our model is that two agents are connected if they know each other. In expectation, each agent is connected to $n(r-1)q + (n-1)p$ persons.

A *society* S is a realization from a generalized random graph model, defined by a four-tuple (n, r, p, q). A society S has a corresponding graph $S = (M \cup W; E)$, where M and W are the set of men and women, respectively, and E is the set of edges. We use the notation $E(i, j) = 1$ if there is an edge between agents i and j, and 0 otherwise. We denote such edge by either (ij) or (ji).

2.3 Agents' Preferences

All agents are heterosexual and prefer marrying anyone over remaining alone.[13] We denote by P_i the set of potential partners for i. The preferences of agent i are given by a function $\delta_i : P_i \to \mathbb{R}_+$ that has a distance interpretation.[14] An agent i prefers agent j over agent k if $\delta_i(i, j) \leq \delta_i(i, k)$. The intuition is that agents like potential partners that are close to them in terms of personality

[13] Heterosexuality is assumed for convenience, because it is well-known that in one-sided matching there may be no stable pairings.

[14] Although δ can be generalized to include functions that violate the symmetry ($\delta(x, y) \neq \delta(y, x)$) and identity ($\delta(x, x) = 0$) characteristic properties of mathematical distances.

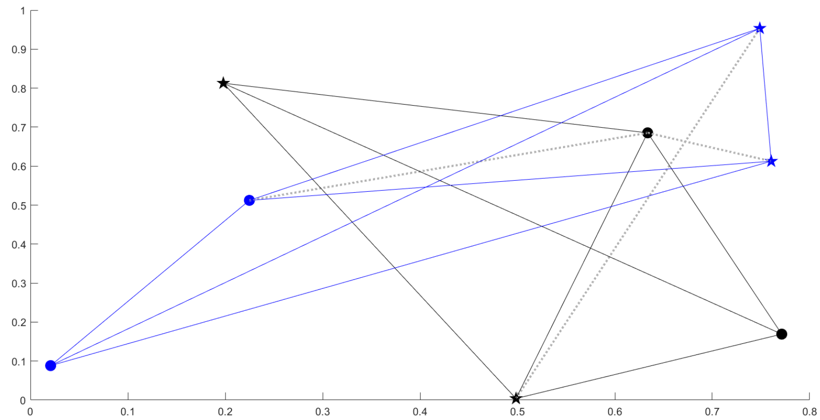


Figure 3: 4 agents, 2 races, linked with $p = 1$ and $q = 0.2$.

traits.

The function δ_i could be arbitrary, or could be the same for agents of the same race. It could also be weighted to account for strong intraracial preferences that are often observed in reality (Wong, 2003; Fisman et al., 2008; Hitsch et al., 2010; Rudder, 2014; Potarca and Mills, 2015; McGrath et al., 2016).[15] Inter or intraracial preferences can easily be incorporated into the model, as in equation (3) below, but for ease of exposition and mathematical convenience, we only consider in the main text two intuitive and simple functions that do not incorporate homophily. We incorporate homophily in Appendix B.

The first one is the Euclidean distance for all agents, so that for any agent i and every potential partner $j \neq i$,

$$\delta^E(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{1}$$

and $\delta^E(i, i) = \sqrt{2}\ \forall i \in M \cup W$.[16] Euclidean preferences are intuitive and

[15] It is not clear whether the declared intraracial preferences show an intrinsic intraracial predilection or capture external biases, which, when removed, leave the partner indifferent to match across races. Evidence supporting the latter hypothesis includes: Fryer (2007) documents that White and Black U.S. veterans have had higher intermarry rates after serving with mixed communities. Fisman et al. (2008) finds that people do not find partners of their own race more attractive. Rudder (2009) shows that online daters have a roughly equal user compatibility. Lewis (2013) finds that users are more willing to engage on interracial dating if they interacted earlier with a dater from another race.

[16] We deviate from the usual definition of Euclidean distance to allow for a more intuitive

have been widely used in social science (Bogomolnaia and Laslier, 2007). The indifference curves associated with Euclidean preferences can be described by concentric circles around each point.

The second preferences we consider are such that every agent prefers a partner close to them in personality trait x, but they all agree on the optimum value in personality trait y. The intuition is that the y-coordinate indicates an attribute like wealth, that is usually considered desirable by all partners. We call these preferences *assortative*.[17] Formally, for any agent i and every potential partner $j \neq i$,

$$\delta^A(i,j) = |x_i - x_j| + (1 - y_j) \tag{2}$$

and $\delta^E(i,i) = 2 \ \forall i \in M \cup W$. The indifference curves of assortative preferences are depicted in Figure 4.

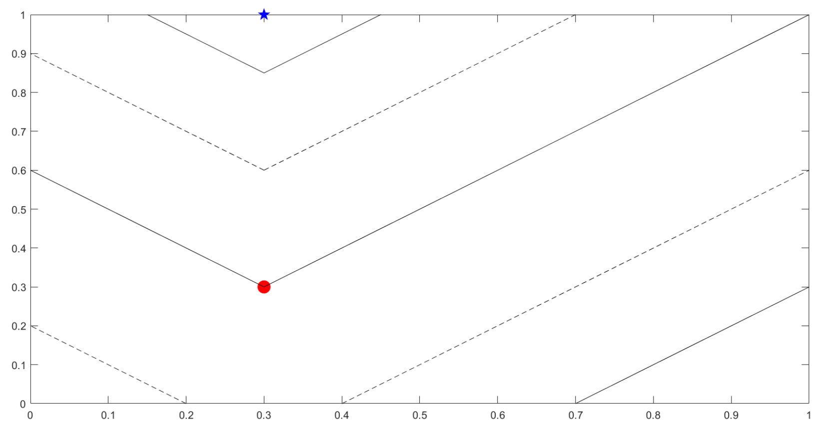


Figure 4: Indifference curves for assortative preferences (the blue star is the best partner for the red dot).

Both Euclidean and assortative preferences can be generalized by weighting them by specific constants β_{ij}, such that

$$\delta_i' = \beta_{ij}\, \delta(i,j) \tag{3}$$

definitions of marriages in the next Subsection.

[17]If we keep the x-axis fixed, so that agents only care about the y-axis, we get full assortative mating as a particular case.

The factor β_{ij} captures intraracial specific preferences whenever it is constant for all pairs i, j who belong to the same race, but different for all pairs i, k not belonging to the same race. Similarly, it can capture specific reluctance to match with agents from specific races whenever above 1. We use this function to incorporate homophily in preferences in Appendix B.

A society in which all agents have either all Euclidean or all assortative preferences will be called Euclidean or assortative, respectively. We focus on these two cases. In both cases agents' preferences are strict because we assume personality traits are drawn from a continuous distribution.

2.4 Marriages

Agents can only marry potential partners they know: i.e. if there exists a path of length at most k between them in the society graph.[18] We consider two types of marriages:

1. Direct marriages: $k = 1$. Agents can only marry if they know each other.
2. Long marriages: $k = 2$. Agents can only marry if they know each other or if they have a mutual friend in common.

To formalize the previous marriage notion, let $\rho_k(i, j) = 1$ if there is a path of at most length k between i and j, with the convention $\rho_1(i, i) = 1$. A marriage $\mu : M \cup W \to M \cup W$ of length k is a function that satisfies

$$\forall m \in M \qquad \mu(m) \in W \cup \{m\} \tag{4}$$
$$\forall w \in W \qquad \mu(w) \in M \cup \{w\} \tag{5}$$
$$\forall i \in M \cup W \qquad \mu(\mu(i)) = i \tag{6}$$
$$\forall i \in M \cup W \qquad \mu(i) = j \text{ only if } \rho_k(i, j) = 1 \tag{7}$$

We use the convention that agents that remain unmarried are matched to themselves. We use $M^* = \{m \in M \mid \mu(m) \in W\}$ to denote the set of all married men.

[18] A path from node i to t is a set of edges $(ij), (jk), \ldots, (st)$. The length of the path is the number of such pairs.

Because realized romantic pairings are close to those predicted by stability (Hitsch et al., 2010; Banerjee et al., 2013), we assume that marriages that occur in each society are *stable*. A marriage μ is k-stable if there is no man-woman pair (m, w) who are not married to each other such that

$$\rho_k(m,w) = 1 \quad (8)$$
$$\delta(m,w) < \delta(m,\mu(m)) \quad (9)$$
$$\delta(w,m) < \delta(w,\mu(w)) \quad (10)$$

Condition (8) is the only non-standard one in the matching literature, that ensures that a pair of agents cannot block a direct marriage if they are not connected in the corresponding graph, even if they prefer each other to their respective partner. Given our assumptions regarding agents' preferences,

Proposition 1. *For any positive integer k, every Euclidean or assortative society has a unique k-stable marriage.*

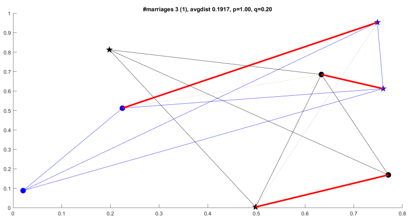
Proof. For the Euclidean society, a simple algorithm computes the unique k-stable marriage. Let every person point to their preferred partner to whom they are connected to by a path of length at most k. In case two people point to each other, marry them and remove them from the graph. Let everybody point to their new preferred partner to which they are connected to among those still left. Again, marry those that choose each other, and repeat the procedure until no mutual pointing occurs. The procedure ends after at most $\frac{rn}{2}$ iterations. This algorithm is a minor modification of the one suggested by Holroyd et al. (2009, Proposition 9) for 1-stable matchings.[19]

For the assortative society, assume by contradiction that there are two k-stable matchings μ and μ' such that for two men m_1 and m_2, and two women w_1 and w_2, $\mu(w_1) = w_1$ and $\mu(w_2) = w_2$, but $\mu'(w_1) = w_2$ and $\mu'(w_2) = w_1$.[20] The fact that both marriages are k-stable implies, without loss of generality, that for $i, j \in \{1, 2\}$ and $i \neq j$, $\delta(m_i, w_i) - \delta(m_i, w_j) < 0$ and $\delta(w_i, m_j) - \delta(w_i, m_i) < 0$. Adding up those four inequalities, one obtains $0 < 0$, a contradiction. ■ □

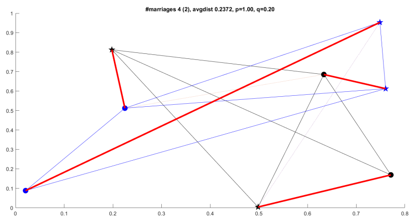
[19] Holroyd et al. (2009) require two additional properties: non-equidistance and no descending chains. The first one is equivalent to strict preferences, the second one is trivially satisfied. In their algorithm, agents point to the closest agent, independently if they are connected to them.

[20] It could be the case that in the two matchings there are no four people who change partner, but that the swap involves more agents. The argument readily generalizes.

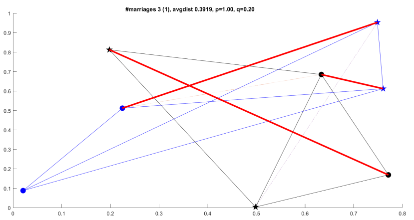
Figure 5 shows the direct and long stable marriages for the Euclidean and assortative societies depicted in Figure 3.



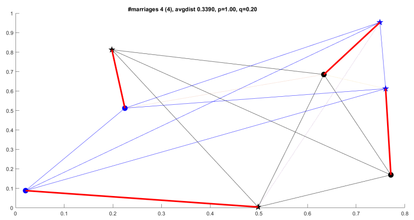
(a) Direct marriage, Euclidean pref.



(b) Long marriage, Euclidean pref.



(c) Direct marriage, assortative pref.



(d) Long marriage, assortative pref.

Figure 5: Direct and long stable marriages for the assortative society in Fig. 3.

2.5 Online Dating on Networks

We model online dating in a society S by increasing the number of interracial edges. Given the graph $S = (M \cup W; E)$, we create new interracial edges between every pair that is disconnected with a probability ϵ.[21,22]

S_ϵ denotes a society that results after online dating has occurred in society S. S_ϵ has exactly the same nodes as S, and all its edges, but potentially more.

[21] Online dating is likely to also increase the number of edges inside each race, but since we assume that each race is already fully connected, these new edges play no role. We perform robustness checks in Appendix B, increasing both p and q but keeping its ratio fixed.

[22] We could assume that particular persons are more likely than others to use online dating, e.g. younger people. Data shows that, from 2013 to 2015, the percentage of people who use online dating has increased for people of all ages. See: "5 facts about online dating", *Pew Research Center*, 29/2/2016. While this occurs at a different rate, to obtain our main result we only need a small increase in the probability of interconnection for each agent.

We say that the society S_ϵ is an *expansion* of the society S.

3 Welfare Measures

We want to understand how the welfare of a society changes after online dating becomes available, i.e. after a society becomes more interracially connected. We consider three welfare measures:

1. **Size**, i.e. the total number of marriages in a society. Formally,

$$sz(S) = |M^*| \tag{11}$$

2. **Diversity**, i.e. how many marriages are interracial. We normalize this measure so that 0 indicates a society with no interracial marriages, and 1 indicates a society in which a fraction $\frac{r-1}{r}$ of the marriages are interracial, i.e. the ratio that obtains in a complete graph in expectation. Note that diversity may be above 1.

Let $\mathcal{R}$ be a function that maps each agent to their race. Then

$$dv(S) = \frac{|\{m \in M^* \mid \mathcal{R}(m) \neq \mathcal{R}(\mu(m))\}|}{sz(S)} \cdot \frac{r}{r-1} \tag{12}$$

3. **Strength**, defined as $\sqrt{2}$ minus the average Euclidean distance between each married couple, denoted as $ds(S)$. A society is stronger whenever its marriages are between agents who are closer to each other. A marriage with a small distance is better than one with a large one because is less susceptible to break up when random agents appear on the unit square, provided that the new outcome is to be k-stable too. The previous observation holds for assortative societies as well.

The above indicator is divided by $\sqrt{2}$ (or the maximal distance possible) to normalize it between 0 and 1.

Formally,

$$ds(S) = \frac{\sum_{m \in M^*} \delta^E(m, \mu(m))}{sz(S)} \tag{13}$$

$$st(S) = \frac{\sqrt{2} - ds(S)}{\sqrt{2}} \tag{14}$$

If every married agent gets paired with her perfect match, then $st(S) = 1$.

4 Edge Monotonicity of Welfare Measures

Given a society S, the first question we ask is whether the welfare measures of a society always increase when its number of interracial edges grow, i.e. when online dating becomes available. We refer to this property as *edge monotonicity*.[23]

Definition 1. *A welfare measure w is edge monotonic if, for any society S, and any of its extensions S_ϵ, we have*

$$w(S_\epsilon) \geq w(S) \tag{15}$$

If a welfare measure is edge monotonic it means that a society unambiguously becomes better off after becoming more interracially connected. Unfortunately,

Proposition 2. *Diversity, strength, and size are all not edge monotonic.*

Before proving Proposition 2, let us build some intuition about it. It may be surprising that the number of interracial marriages can decrease when more interracial edges are formed. The intuition behind it is that a newly formed interracial edge may create one interracial marriage at the cost of destroying two existing ones, and the left-alone partners may now marry partners of their own race.

An interracial edge can similarly increase the average distance between couples if it provides a link between very desirable partners, i.e. those in the center for the case of Euclidean preferences. Those desirable partners are likely to drop their current spouses. The dropped agents now have to match with partners that have been dropped too, which are potentially further away from them. An interracial edge can similarly increase the average distance between couples if it provides a link between very desirable partners, i.e. those in the center for the case of Euclidean preferences. Those desirable partners are likely to drop their current spouses. The dropped agents now have to match with partners that have been dropped too, which are potentially further away from them.

[23] Edge monotonicity is different from node monotonicity, in which one node, with all its corresponding edges, is added to the matching problem. It is well-known that when a new man joins a stable matching problem, every woman weakly improves, while every man becomes weakly worse off (Theorems 5 in Kelso and Crawford, 1982, 2.25 and 2.26 in Roth and Sotomayor, 1992, and 1 and 2 in Crawford, 1991).

Finally, size can be reduced if the new interracial edge links people who were already highly connected in the society, making them leave partners who are poorly so. The left-alone partners may now become unable to find a partner.

We present now a formal proof for Euclidean societies with direct marriages.

Proof. To show that size is not edge monotonic, consider the society in Figure 3 and its direct stable matching in Figure 5a. Remove all interracial edges: it is immediate that in the unique stable matching there are 4 couples now, one more than when interracial edges are present.

For the case of strength, consider a simple society in which all nodes share the same y-coordinate, as the one depicted in Figure 6. There are two intraracial marriages and the average Euclidean distance is 0.35. When we add the interracial edge between the two central nodes, the closest nodes marry and the two far away nodes marry too. The average Euclidean distance in the expanded society increases to 0.45, hence reducing its strength.

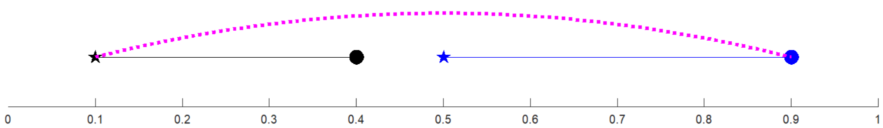
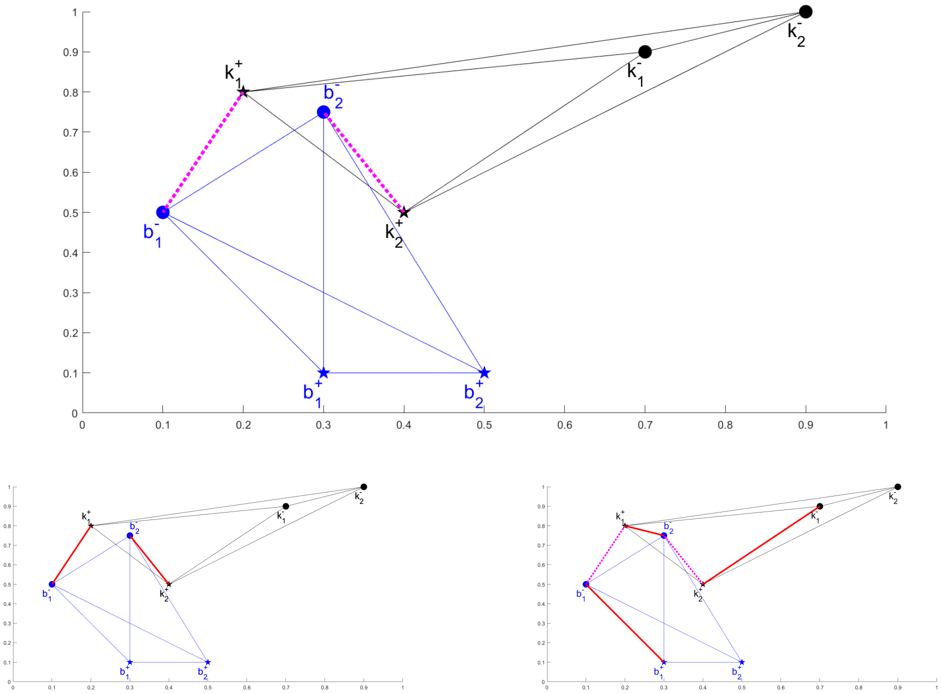


Figure 6: Strength is not edge monotonic.

To show that diversity is not edge monotonic, consider Figure 7. There are two men and two women of each of two races a and b. Each gender is represented with the superscript $^+$ or $^-$.

Stability requires that $\mu(b_1^-) = a_1^+$ and $\mu(b_2^+) = a_2^-$, and everyone else is unmarried. However, when we add the interracial edge $(a_1^+ b_2^-)$, the married couples become $\mu(b_1^-) = b_1^+$, $\mu(a_2^+) = a_1^-$, and $\mu(a_1^+) = b_2^-$. In this extended society, there is just one interracial marriage, out of a total of three, when before we had two out of two. Therefore diversity reduces after adding the edge $(a_1^+ b_2^-)$. ■ □

The failure of edge monotonicity by our three welfare measures makes



(a) $dv(S) = 2$ (b) $dv(S_\epsilon) = 1/3$

Figure 7: Diversity is not edge monotonic.

evident that to evaluate welfare changes in societies, we need to understand how welfare varies on an average society after introducing new interracial edges. We develop this comparison in the next Section.

A final comment on edge monotonicity. The fact that the size of a society is not edge monotonic, as shown in Proposition 2, implies that adding interracial edges may not lead to a Pareto improvement for the society. Some agents may become worse off after the society becomes more connected. Nevertheless, the fraction of agents that becomes worse off after adding an extra edge is never more than one-half of the society. Ortega (2017) discusses this phenomenon in detail and characterizes the associated welfare losses of those hurt by integration.

5 Average Welfare Measures

In the last Section we found that our three welfare measures may increase or decrease after adding interracial edges. Therefore, we need to analyze what happens to welfare in expectation when agents become more connected.

There are two ways to answer this question. The first one is to provide analytical expressions for the expected welfare measures as a function of the number of interracial edges. However, providing analytical solutions is incredibly complicated, if not impossible. Already solving the expected average distance in a toy society with just one race, containing only one man and one woman, requires a complicated computation (which eventually becomes $\frac{2+\sqrt{2}+5\ln(\sqrt{2}+1)}{15} \approx 0.52$).[24]

The second way to approach the problem is to simulate several random societies and observe how their average welfare change when they become more connected. This is the route we follow. We create ten thousand random societies, and increase the expected number of interracial edges by increasing the parameter q. In the following subsections, we describe the changes of our welfare measures for different values of q.

In all cases we fix $n = 50$ and $p = 1$.[25] We consider the following four scenarios:

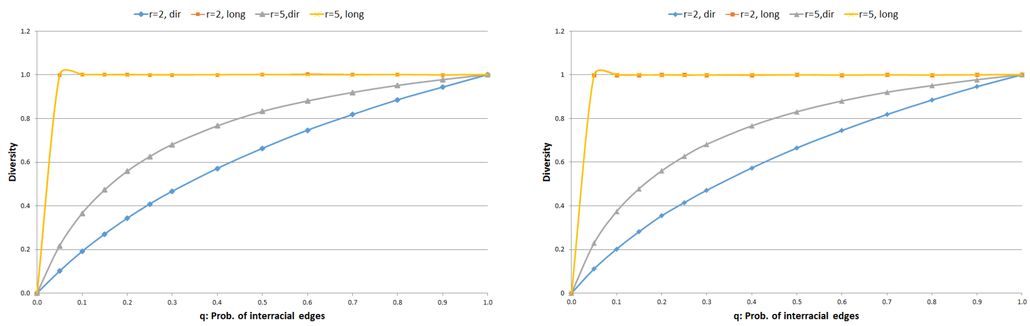
1. Two races and direct marriages, appears in blue with diamond markers ♦.
2. Five races and direct marriages, appears in grey with square markers ■.
3. Two races and long marriages, appears in orange with triangle markers ▲.
4. Five races and long marriages, appears in yellow with cross markers ×.

[24]The detailed computation appears in "Distance between two random points in a square", *Mind your Decisions*, 3/6/2016.

[25]We restrict to $n = 50$ and ten thousand replications because of computational limitations, even though we used the high performance computing facilities at the University of Glasgow. The results for other values of p are similar and we present them in Appendix B. The Matlab code is available at www.josueortega.com.

5.1 Diversity

In the case of long marriages, even the smallest increase in the probability of interracial connections (in this case of 0.05) achieves perfect social integration. With either two or five races we obtain that diversity is exactly one. For the cases with direct marriages, the increase in diversity is slower but still fast: an increase of q from 0 to 0.1 increases diversity to 0.19 for $r = 2$, and from 0 to 0.37 with $r = 5$.[26]



(a) Euclidean society.

(b) Assortative society

Figure 8: Average diversity (y-axis) of a random society for different values of q.

The yellow and orange curves are indistinguishable in this plot because they are identical. Exact values and standard errors (which are in the order of 1.0e-04) are provided in Appendix A.

Figure 8 summarizes our main result, namely

Result 1. *Diversity is fully achieved with long marriages, even if the increase in interracial connections is arbitrarily small.*

With direct marriages, diversity is achieved partially but still substantially, so that an increase in q always yields an increase in diversity of a larger size, i.e. diversity is a concave function of q.

The intuition behind full diversity for the case of long marriages is that, once an agent obtains just one edge to any other race, he gains $\frac{n}{2}$ potential

[26] Empirical evidence suggests that q is close to zero. Echenique and Fryer (2007) find that the typical American public school student has 0.7 friends of another race. It is also a sensible assumption that p is large, given the clear residential segregation patterns in the U.S. (Cutler et al., 1999) and that around 90% of people who attend religious services do so with others from their same race (Fryer, 2007).

partners. Just one edge to a person of different race gives access to that person's complete race.

The reader may think that the full diversity result heavily depends on each race being fully connected, i.e. $p = 1$. This is not the case. We obtain full diversity for many other values of p, as we discuss in Appendix B. When same-race agents are less interconnected within themselves, agents gain fewer connections once an interracial edge is created, but those fewer connections are relatively more valuable, because the agent had himself less potential partners before the creation of new interracial edges.

Result 1 implies that, assuming long marriages are formed, very few interracial links can lead a society to almost complete racial integration, and leads to very optimistic views on the role that dating platforms can play in the reduction of racial segregation in our society. Our result is in sharp contrast to the one of Schelling (1969, 1971) in its well-known models of residential segregation, in which a society always gets completely segregated.

We pose this finding as the first testable hypothesis of our model

Hypothesis 1. *The number of interracial marriages should increase after the popularization of online dating.*

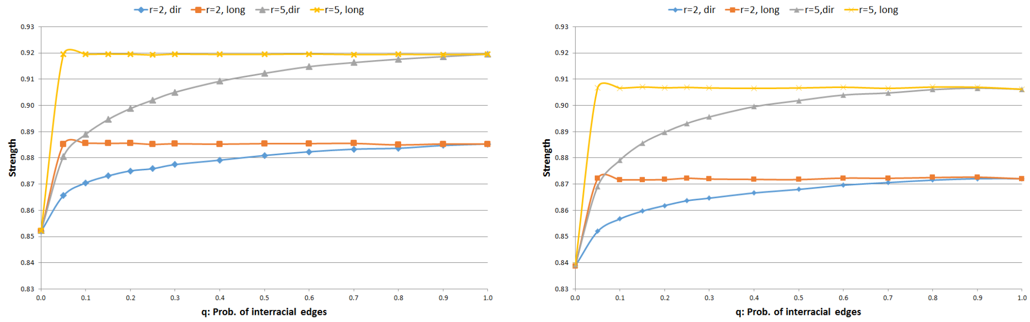
5.2 Strength

A second observation, less pronounced that the increase in diversity, is that the strength of the society goes up when increasing q. For an illustration, see Figure 9, which considers the same four cases as before in both Euclidean and assortative societies.

It is clear that, for all combinations of parameters (see Appendix B for further robustness checks), there is a consistent trend downwards in the average distance of partners after adding new interracial edges, and thus a consistent increase in strength of the societies. We present this observation as our second result.

Result 2. *Strength increases after the number of interracial edges increases. The increase is faster whenever the society has more races, and converges to a higher level with long marriages.*

Assuming that marriages with a higher average distance have a higher



(a) Euclidean society. (b) Assortative society

Figure 9: Average strength (y-axis) of a random society for different values of q.

Exact values and standard errors (which are in the order of 1.0e-04) provided in Appendix A.

chance to end up divorcing, because they are more susceptible to break up when new nodes are added to the society graph, we can reformulate our result as our second hypothesis.

Hypothesis 2. *Marriages created in societies with online dating should have a lower divorce rate.*

Finally, our last welfare measure, size, keeps constant for most of our simulations, so we do not discuss it further. The detailed data behind Figures 8 and 9 appear in Appendix A.

Our analysis of the expected changes in welfare gives us with two testable hypotheses. In the next Section, we contrast them against data on of interracial marriage in the U.S, and the quality of the marriages created through online dating.

6 Hypotheses and Data

6.1 Hypothesis 1: More Interracial Marriages

What does the data reveal? Is our model consistent with observed demographic trends? Figure 10 presents the evolution of interracial marriages among newlyweds in the U.S. from 1967 to 2015, based on the 2008-2015

American Community Survey and 1980, 1990 and 2000 decennial censuses (IPUMS). In this Figure, interracial marriages include those between White, Black, Hispanic, Asian, American Indian or multiracial persons.[27]

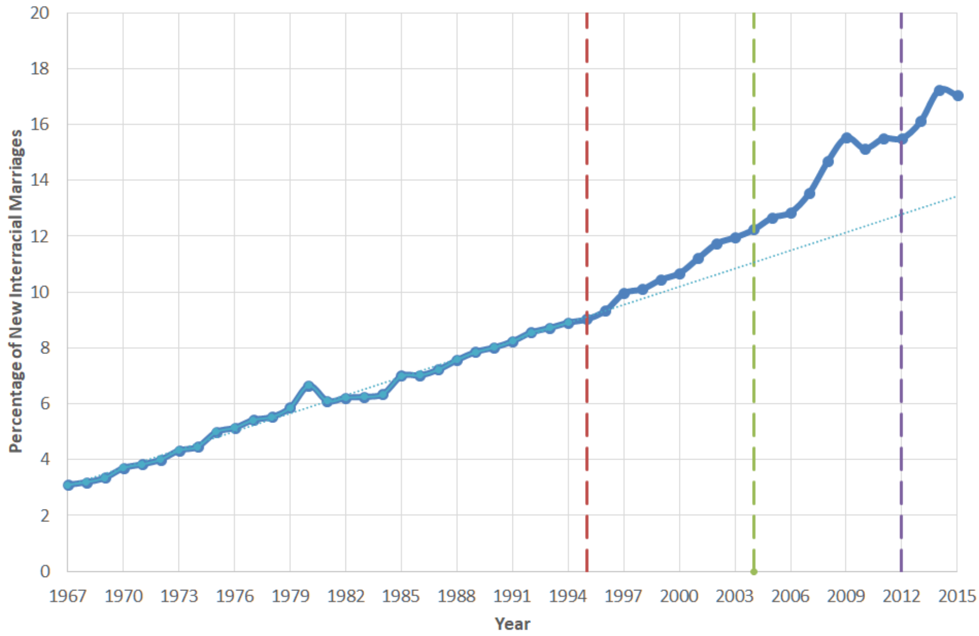


Figure 10: Percentage of interracial marriages among newlyweds in the U.S.

Source: Pew Research Center analysis of 2008-2015 American Community Survey and 1980, 1990 and 2000 decennial censuses (IPUMS). The red, green, and purple lines represent the creation of Match.com, OKCupid, and Tinder, three of the largest dating websites. The blue line represents a linear prediction for 1996 – 2015 using the data from 1967 to 1995.

We observe that the number of interracial marriages has consistently increased in the last 50 years, as it has been documented by several other authors (Kalmijn, 1998; Fryer, 2007; Furtado, 2015). However, it is intriguing that shortly after the introduction of the first dating websites in 1995, like Match.com, the percentage of new marriages created by interracial couples increased rapidly. The increase becomes steeper around 2004, when online dating became more popular: it is then when well-known platforms such like OKCupid emerged. During the 2000's decade, the percentage of new marriages that are interracial changed from 10.68% to 15.54%, a huge increase of nearly 5 percentage points, or 50%.

[27]We are grateful to Gretchen Livingston from the Pew Research Center for providing us with the data. Data prior to 1980 are estimates. The methodology on how the data was collected is described in Livingston and Brown (2017).

After the 2009 increase, the proportion of new interracial marriage jumps again in 2014 to 17.24%, remaining above 17% in 2015 too. Again, it is interesting that this increase occurs shortly after the creation of Tinder, considered the most popular online dating app. Tinder, created in 2012, has approximately 50 million users and produces more than 12 million matches per day.[28] Matches can be thought of newly established edges, in the language of our model.

We do not claim that the increase in the share of new marriages that are interracial in the last 20 years is caused by the emergence of online dating alone, but this finding is in line with Hypothesis 1 in our model.

Another cause for the steep increase described could be that the U.S. population is more interracial now than 20 years ago. The reduction of the percentage of Americans who are White, falling from 83.1% to 72.4% from 1980 to 2010, would yield an increase of the rate of interracial marriage, assuming random marriage. However, the change in the population composition in the U.S cannot explain the huge increase in intermarriage that we observe. In Appendix C we show that, even controlling for demographic change, we observe an increase of interracial marriages, although certainly smaller.

A more transparent way to see that the increase in the number of interracial marriages cannot be due to changes in population composition alone is too look at the growth of interracial marriages for Black Americans. Black Americans are the racial group whose rate of interracial marriage has increased the most, going from 5% in 1980 to 18% in 2015. However, the fraction of the U.S. population who is Black has changed very little in the last 40 years, remaining constant around 12% of the population. Random marriage accounting for population change would then predict that the rate of interracial marriages remains roughly constant, although in reality it has more than triplicate in the last 35 years.

[28] "Tinder, the fast-growing dating app, taps an age-old truth", *New York Times*, 29/10/2014. The company claims that 36% of Facebook users have had an account on their platform.

6.2 Hypothesis 2: Marriages Created Online Are Less Likely to Divorce

Cacioppo et al. (2013) find that marriages created online were less likely to break up and reveal a higher marital satisfaction, using a sample of 19,131 Americans who married between 2005 and 2012. They write: *"Meeting a spouse on-line is on average associated with slightly higher marital satisfaction and lower rates of marital break-up than meeting a spouse through traditional (off-line) venues"*.

The findings of Cacioppo and his coauthors show that our model predictions closely match the observed properties of marriages created online, and its strength compared to marriages created on other, more traditional venues.

Our model predicts that, on average, marriages created when online dating becomes available last longer than those created in societies without this technology. Yet, it is silent regarding comparisons between the strength of interracial and intraracial marriages. There is empirical evidence showing that interracial marriages are more likely to end up in divorce (Bratter and King, 2008; Zhang and Van Hook, 2009).

Our model is also silent on why some intraracial marriages from a particular race last longer than intraracial marriages from another race (e.g. Stevenson and Wolfers, 2007 show that Blacks who divorce spend more time in their marriage than their White counterparts).

7 Final Remarks

7.1 Further Applications

The theoretical model we present discusses a general matching problem under network constraints, and hence it can be useful to study other social phenomena besides interracial marriage. The races or communities in our model can be understood as arbitrary groups of highly clustered agents. Agents can be clustered by race, but also by ethnicity, education, socioeconomic status, religion, nationality, etcetera. Thus, our theoretical model can be also applied to study interfaith marriages, or marriages between people of different social status.

The role of connecting highly clustered groups is also not only linked to online dating. Another example is the European student exchange program "Erasmus", which helped more than 3 million students and over 350 thousand academics and staff members to spend time at a University abroad.[29]

The matching of agents also goes beyond marriage. Think of nodes being researchers at a University, races being academic departments, and edges representing who knows whom. Matchings indicate academic collaboration in articles or grants. The Euclidean distance interpretation makes sense, as a microeconomist in a business school may be better off partnering with a game theorist at the biology department rather than with an econometrician in his own business school. Diversity in a University would be then a measure for interdisciplinary research, often encouraged by higher education institutions and funding bodies. Interdiscplinary seminars, for example, could take the role of creating links between academics in different departments.

It would be interesting to test our model against in this other scenarios. We leave this task for further research.

7.2 Conclusion

We introduce a simple theoretical model which tries to explain the complex process of deciding whom to marry in the times of online dating. As any model, ours has limitations. It categorizes every individual with only two characteristics, it assumes a very simple structure inside each race, it poses restrictions on agents' preferences. Furthermore, it fails to capture many of the complex features of romance in social networks, like love. There are multiple ways to enrich and complicate the model with more parameters.

However, the simplicity of our model is its main strength. With a basic structure, it can generate very strong predictions. It suggests that the diversity of societies, measured by the number of interracial marriages in it, should increase drastically after the introduction of online dating. Societies where online dating is available should produce marriages that are less likely to break up. Both predictions are consistent with observed demographic

[29] "ERASMUS: Facts, figures and trends.", *European Commission*, 10/6/2014. Interestingly, Parey and Waldinger (2011) find that participating in ERASMUS to study abroad increases the probability of working abroad by 15 percentage points.

trends.

Simple models are great tools to convey an idea. Schelling's segregation model clearly does not capture many important components of how people decide where to live. It could have been enhanced by introducing thousands of parameters. Yet it has broadened the way how we understand racial segregation, and has been widely influential: according to Google Scholar, it has been quoted 3,258 times by articles coming from sociology to mathematics. It has provided us a way to think about an ubiquitous phenomenon.

Our model goes in the same direction.

References

Amir, G., O. Angel, and A. Holroyd (2016): "Multicolor Poisson matching," *unpublished.*

Arrow, K. (1998): "What has economics to say about racial discrimination?" *Journal of Economic Perspectives*, 12, 91–100.

Backstrom, L. and J. Kleinberg (2014): "Romantic partnerships and the dispersion of social ties: A network analysis of relationship status on Facebook," *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 831–841.

Banerjee, A., E. Duflo, M. Ghatak, and J. Lafortune (2013): "Marry for what? Caste and mate selection in modern India," *American Economic Journal: Microeconomics*, 5, 33–72.

Becker, G. (1973): "A theory of marriage: Part I," *Journal of Political Economy*, 81, 813–846.

——— (1974): "A theory of marriage: Part II," *Journal of Political Economy*, 82, S11–S26.

——— (1981): *A treatise on the family*, Harvard University Press.

Bogomolnaia, A. and J.-F. Laslier (2007): "Euclidean preferences," *Journal of Mathematical Economics*, 43, 87 – 98.

Bollobás, B. (2001): *Random graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2 ed.

Bratter, J. and R. King (2008): "But will it last?: Marital instability among interracial and same-race couples," *Family Relations*, 57, 160–171.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014): *Economics of the family*, Cambridge Surveys of Economic Literature, Cambridge University Press.

Cacioppo, J., S. Cacioppo, G. Gonzaga, E. Ogburn, and T. VanderWeele (2013): "Marital satisfaction and break-ups differ across online and off-line meeting venues," *Proceedings of the National Academy of Sciences*, 110, 10135–10140.

Chiappori, P.-A., R. McCann, and B. Pass (2016): "Multidimensional matching," Papers 1604.05771, arXiv.org.

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2012): "Fatter attraction: Anthropometric and socioeconomic matching on the marriage market," *Journal of Political Economy*, 120, 659–695.

Crawford, V. (1991): "Comparative statics in matching markets," *Journal of Economic Theory*, 54, 389 – 400.

Cutler, D., E. Glaeser, and J. Vigdor (1999): "The rise and decline of the American ghetto," *Journal of Political Economy*, 107, 455–506.

Duncan, B. and S. Trejo (2011): "Intermarriage and the intergenerational transmission of ethnic identity and human capital for Mexican Americans," *Journal of Labor Economics*, 29, 195–227.

Echenique, F. and R. Fryer (2007): "A measure of segregation based on social interactions," *Quarterly Journal of Economics*, 122, 441–485.

Erdős, P. and A. Rényi (1959): "On random graphs I." *Publicationes Mathematicae (Debrecen)*, 6, 290–297.

Fisman, R., S. Iyengar, E. Kamenica, and I. Simonson (2008): "Racial preferences in dating," *Review of Economic Studies*, 75, 117–132.

Fryer, R. (2007): "Guess who's been coming to dinner? Trends in interracial marriage over the 20th century," *Journal of Economic Perspectives*, 21, 71–90.

Fu, X. (2007): "Inter-racial marriage and family socio-economic status: A study among Whites, Filipinos, Japanese, and Hawaiians in Hawaii," *Journal of Comparative Family Studies*, 38, 533–554.

Furtado, D. (2012): "Human capital and interethnic marriage decisions," *Economic Inquiry*, 50, 82–93.

——— (2015): "Ethnic intermarriage," in *International Encyclopedia of the Social & Behavioral Sciences*, ed. by J. Wright, Oxford: Elsevier, 118 – 122.

Furtado, D. and N. Theodoropoulos (2010): "Why does intermarriage increase immigrant employment? The role of networks," *The B.E. Journal of Economic Analysis & Policy*, 10, 1–33.

Gale, D. and L. Shapley (1962): "College admissions and the stability of marriage," *American Mathematical Monthly*, 69, 9 – 15.

Gee, L., J. Jones, and M. Burke (2017): "Social networks and labor markets: How strong ties relate to job finding on Facebooks social network," *Journal of Labor Economics*, 35, 485–518.

Gilbert, E. (1959): "Random graphs," *Annals of Mathematical Statistics*, 30, 1141–1144.

Gius, M. (2013): "The effects of interracial marriage on individual-level earnings: An analysis using 2010 PUMS data," *Journal of Applied Business and Economics*, 14, 37–41.

Goel, D. and K. Lang (2017): "Social ties and the job search of recent immigrants," *ILR Review*, forthcoming.

Granovetter, M. (1973): "The strength of weak ties," *American Journal of Sociology*, 78, 1360–1380.

Hitsch, G., A. Hortacsu, and D. Ariely (2010): "Matching and sorting in online dating," *American Economic Review*, 100, 130 – 63.

Holroyd, A. (2011): "Geometric properties of Poisson matchings," *Probability Theory and Related Fields*, 150, 511–527.

Holroyd, A., R. Pemantle, Y. Peres, and O. Schramm (2009): "Poisson matching," *Annales de l'Institut Henri Poincar, Probabilits et Statistiques*, 45, 266–287.

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

Kantarevic, J. (2004): "Interethnic marriages and economic assimilation of immigrants," IZA Discussion Papers 1142, Institute for the Study of Labor.

Kelso, A. and V. Crawford (1982): "Job matching, coalition formation, and gross substitutes," *Econometrica*, 50, 1483–1504.

Kramarz, F. and O. Skans (2014): "When strong ties are strong: Networks and youth labour market entry," *Review of Economic Studies*, 81, 1164–1200.

Lee, S. and B. Edmonston (2005): "New marriages, new families: U.S. racial and Hispanic intermarriage," *Population Bulletin*, 60, 1–36.

Lewis, K. (2013): "The limits of racial prejudice," *Proceedings of the National Academy of Sciences*, 110, 18814–18819.

Livingston, G. and A. Brown (2017): "Intermarriage in the U.S. 50 years after Loving v. Virginia," *Pew Research Center Report*.

McGrath, A., G. Tsunokai, M. Schultz, J. Kavanagh, and J. Tarrence (2016): "Differing shades of colour: Online dating preferences of biracial individuals," *Ethnic and Racial Studies*, 39, 1920–1942.

Meng, X. and R. Gregory (2005): "Intermarriage and the economic assimilation of immigrants," *Journal of Labor Economics*, 23, 135–174.

Ortega, J. (2017): "Can everyone benefit from social integration?" Papers 1705.08033, arXiv.org.

Parey, M. and F. Waldinger (2011): "Studying abroad and the effect on international labour market mobility: Evidence from the introduction of ERASMUS," *Economic Journal*, 121, 194–222.

Pearce-Morris, J. and V. King (2012): "The well-being of children living with interethnic parents," *Journal of Family Issues*, 33, 898–919.

Potarca, G. and M. Mills (2015): "Racial preferences in online dating across European countries," *European Sociological Review*, 31, 326–341.

Qian, Z. (1997): "Breaking the racial barriers: Variations in interracial marriage between 1980 and 1990," *Demography*, 34, 263–276.

Rosenfeld, M. and R. Thomas (2012): "Searching for a mate: The rise of the internet as a social intermediary," *American Sociological Review*, 77, 523–547.

Roth, A. and M. Sotomayor (1992): *Two-sided matching: A study in game-theoretic modeling and analysis*, Econometric Society Monographs, Cambridge University Press.

Rudder, C. (2009): "How your race affects the messages you get," OK Cupid Blog.

——— (2014): *Dataclysm: Who we are (when we think no one's looking)*, New York, NY, USA: Crown Publishing Group.

Schelling, T. (1969): "Models of segregation," *American Economic Review*, 59, 488–493.

——— (1971): "Dynamic models of segregation," *Journal of Mathematical Sociology*, 1, 143–186.

——— (1972): "A process of residential segregation: Neighborhood tipping," in *Racial Discrimination in Economic Life*, ed. by A. Pascal, Lexington, Massachusetts: Lexington Books - RAND, chap. 5, 157–184.

Stevenson, B. and J. Wolfers (2007): "Marriage and divorce: Changes and their driving forces," *Journal of Economic Perspectives*, 21, 27–52.

Wang, W. (2012): "The rise of intermarriage," *Pew Research Center Report*.

Wong, L. (2003): "Why do only 5.5% of black men marry white women?" *International Economic Review*, 44, 803–826.

Zhang, Y. and J. Van Hook (2009): "Marital dissolution among interracial couples," *Journal of Marriage and Family*, 71, 95–107.

A Appendix A: Simulation Results

Table 1: Supporting data for Figures 8 and 9

q	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Panel A: Welfare on Euclidean societies														
$r = 2$, direct marriages														
Dv	0.00	0.10	0.19	0.27	0.34	0.41	0.47	0.57	0.66	0.75	0.82	0.89	0.94	1.00
St	0.85	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.89
Sz	1.00	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00
$r = 2$, long marriages														
Dv	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.85	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.88	0.89	0.89
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$,direct marriages														
Dv	0.00	0.22	0.37	0.47	0.56	0.62	0.68	0.77	0.83	0.88	0.92	0.95	0.98	1.00
St	0.85	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.91	0.91	0.92	0.92	0.92	0.92
Sz	1.00	0.97	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00
$r = 5$, long marriages														
Dv	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.85	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Panel B: Welfare on assortative societies														
$r = 2$, direct marriages														
Dv	0.00	0.11	0.20	0.28	0.35	0.41	0.47	0.57	0.66	0.75	0.82	0.88	0.95	1.00
St	0.84	0.85	0.86	0.86	0.86	0.86	0.86	0.87	0.87	0.87	0.87	0.87	0.87	0.87
Sz	1.00	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00
$r = 2$, long marriages														
Dv	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.84	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$,direct marriages														
Dv	0.00	0.23	0.37	0.48	0.56	0.63	0.68	0.77	0.83	0.88	0.92	0.95	0.98	1.00
St	0.84	0.87	0.88	0.89	0.89	0.89	0.90	0.90	0.90	0.90	0.90	0.91	0.91	0.91
Sz	1.00	0.97	0.98	0.98	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00
$r = 5$, long marriages														
Dv	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.84	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Average over 10,000 random simulations, $n = 50$, $p = 1$.

Sz equals the percentage of agents married.

Standard errors in the order of 1.0e-04, so we do not present them.

B Appendix B: Robustness Checks

In this Appendix we conduct several robustness checks to show that the fast increase in the diversity of societies, described in Result 1, occurs for many combinations of model parameters.

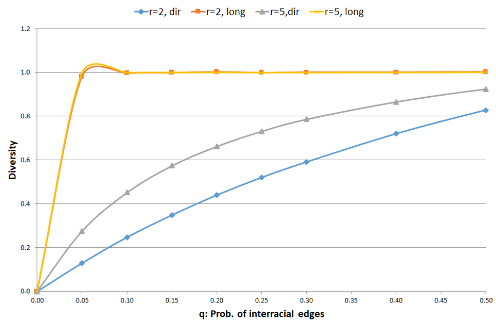
B.1 Different Values of p

The first exercise we conduct is to simulate the model again, but varying the probability of intraracial connections p to 0.7, 0.5 and 0.3. We allow q to vary between 0 and p, as we have explained in the text that $q \leq p$, because people tend to be more connected to people from their own race.

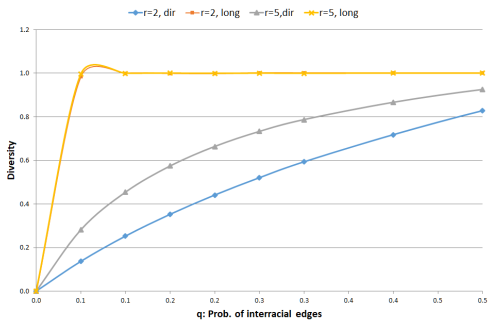
With respect to diversity, long marriages always lead to an almost immediate increase to 1, meaning complete social integration. This increase is shown in Figure 11. As expected, a society integrates faster when the value of p is higher.

With respect to strength, we also observe minor variations, which appear in Figure 12. A smaller p makes agents less connected to potential partners, and thus the strength of resulting marriages becomes weaker. With long marriages, strength converges very quickly to its optimal value, around 0.9, which again, is smaller in societies with low values of p and q.

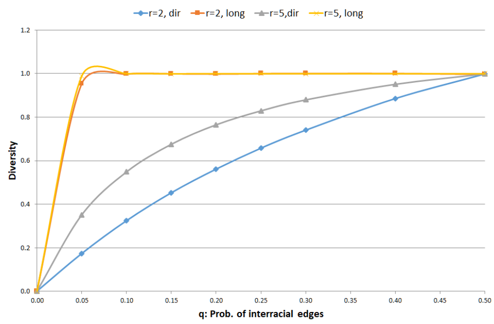
The detailed results of our simulations with p equal to 0.7, 0.5, and 0.3 appear in Tables 2, 3 and 4 at the end of this Appendix.



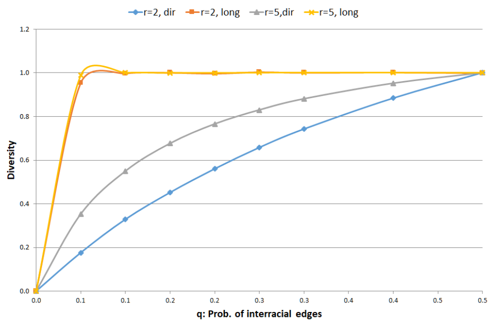
(a) Euclidean society, $p = .7$.



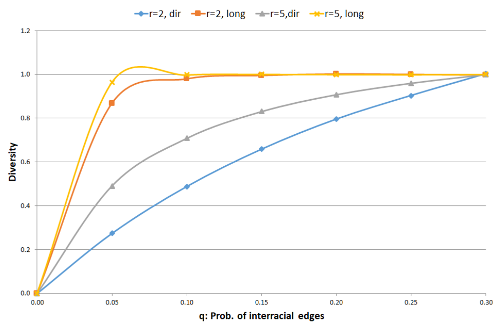
(b) Assortative society, $p = .7$



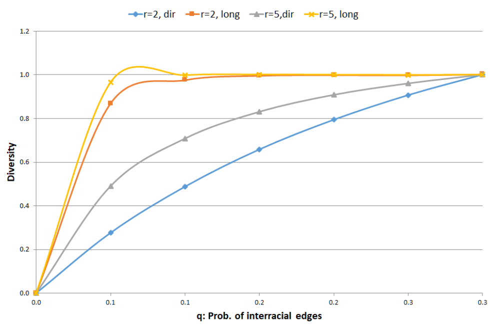
(c) Euclidean society, $p = .5$.



(d) Assortative society, $p = .5$

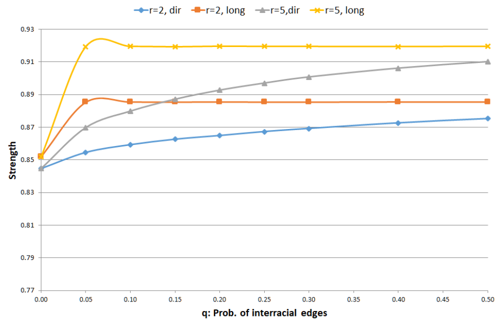


(e) Euclidean society, $p = .3$.

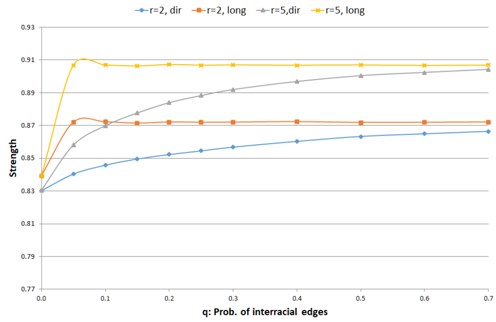


(f) Assortative society, $p = .3$

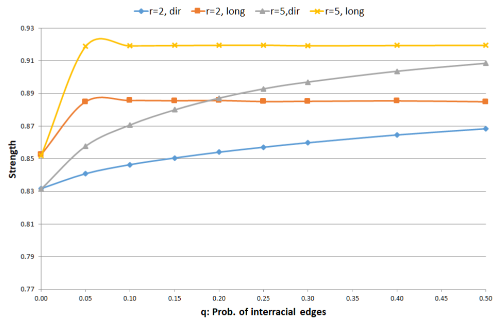
Figure 11: Average diversity (y-axis) of a random society for several values of p.



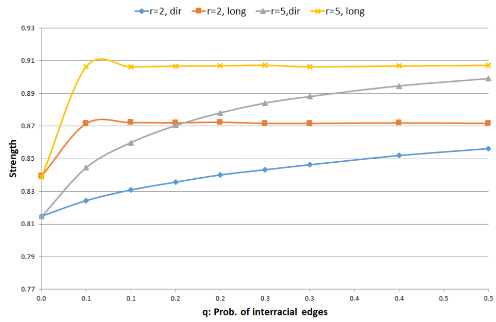
(a) Euclidean society, $p = .7$.



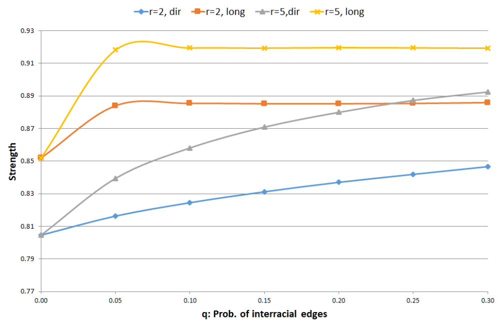
(b) Assortative society, $p = .7$



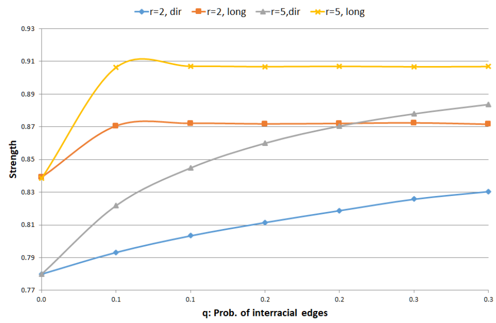
(c) Euclidean society, $p = .5$.



(d) Assortative society, $p = .5$



(e) Euclidean society, $p = .3$.



(f) Assortative society, $p = .3$

Figure 12: Average stregth (y-axis) of a random society for several values of p.

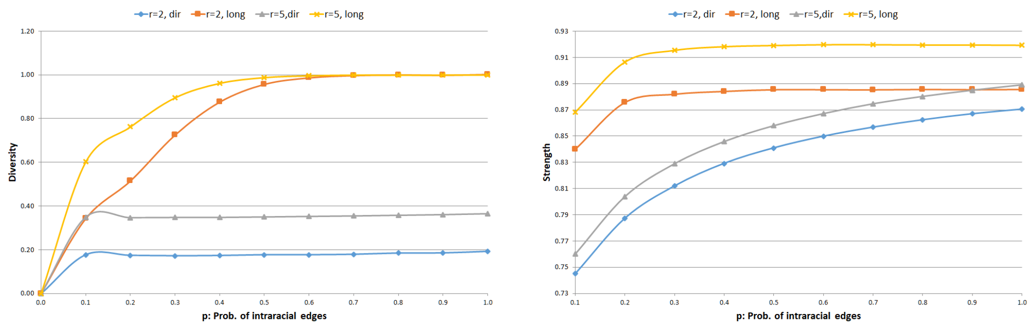
B.2 Varying p and q Simultaneously

The second robustness test we perform is to vary p and q simultaneously but keeping its ratio fixed. Both parameters indicate how connected a person is to people of his own race and to people of other races.

To find a good estimate of the ratio $\frac{p}{q}$, we use data from the American Values Survey by the Public Religion Research Institute (PRRI), a nonpartisan, independent research organization. The data is well described in the following article from the Washington Post: "Three quarters of Whites dont have any non-White friends", 25/8/2014.

The PRRI data shows that, if a White American has 100 friends, 91 are expected to be of his own race, and 1 Black, 1 Latino, and 1 Asian (the rest are multiracial or of unknown race). Black Americans are more interracially connected, with 83 friends expected to be of his own race, 8 Whites, 2 Latinos, and and no Asians.

Based on this data, we use the ratio $p/q = 10$, based on the ratio between the expected number of Black and White friends for Black people. This ratio implies that a person is 10 times more likely to be connected to a person from her own race. We vary p from 0 to 1. We present the results for Euclidean societies only (as we have seen that Euclidean and assortative societies produce almost identical results).



(a) Diversity. (b) Strength.

Figure 13: Average diversity and strength of a random society for $p \in [0, 1]$.

A first conclusion we obtain is that, with long marriages, we observe complete integration, just as we did when increasing q alone. However, this time it does not happen as quickly as when we increase only q. With direct marriages the increase is very fast but full integration is not obtained. It only reaches values of 20% and 40% in societies with 2 and 5 races, respectively.

We could say that the diversity achieved when agents' intra and interracial circles both expand is much lower, compared to the results shown in the main text. But this this conclusion is flawed, because we compare our diversity measure to one where agents were completely connected within their own race, i.e. a complete graph. Therefore, the diversity obtained already is 20% and 40% of the diversity in a complete graph. This is a very high percentage of interracial marriages, because we fix that agents are 10 times more connected withing their own race. Notice that the results (Table 6 are consistent with what is displayed in Figures 8 and 9, for the point $q = 0.1$.

Finally, the strength levels we observe with direct marriages are the lowest we have found so far, which is not surprising given the small number of potential partners that agents have. It is equally expected to observe that the strength of a society increases when p grows.

The detailed results of our simulations with $p/q = 10$ appear in Table 5 at the end of this Appendix.

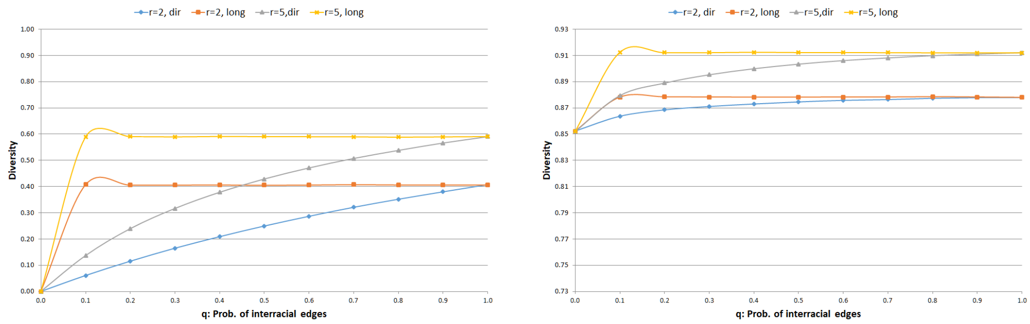
B.3 Homophily

The third robustness test we perform is to introduce intraracial preferences, as described in equation (3). We do this in the following intuitive way. Agents prefer marrying someone from their own race β times as much as marrying someone from another race. This is, for agents i, j, k, with agents i and j being from the same race, and agent k being from another race, i is indifferent between j and k only if $\delta(i,j) = \beta\, \delta(i,k)$, where $\beta \geq 1$. We still impose that marrying any potential partner is better than remaining alone for all agents.

There is evidence suggesting that persons substitute similarities in race for similarities in personality traits. See Furtado (2012) for evidence of tradeoffs in marriage choices between race and education.

Figure 14 presents the behavior of diversity and strength when agents prefer their own race twice and $\beta = 2$. With long marriages, we obtain a fast increase in diversity of our societies. However, only a diversity of 0.4 and 0.6 is achieved with 2 and 5 races, respectively. This is the diversity with respect to a society with $\beta = 1$. Therefore, the diversity achieved is large, even when agents have intraracial preferences.

With direct marriages, we observe that the increase converges to the same levels as with long marriages, but at a slower rate. The increase is a concave function of q, as documented in the main text when no homophily is present.



(a) Diversity. (b) Strength.

Figure 14: Average diversity and strength of a random society with $\beta = 2$.

The reader may wonder how large β needs to be so that no diversity occurs in the society. Figure 15 shows how diversity changes as a function of β, for a society with $p = q = 1$. What we find is that even when agents prefer their own race 3.5 times as much as any other race, the society achieves 20% of the integration it would achieve without any racial preferences.

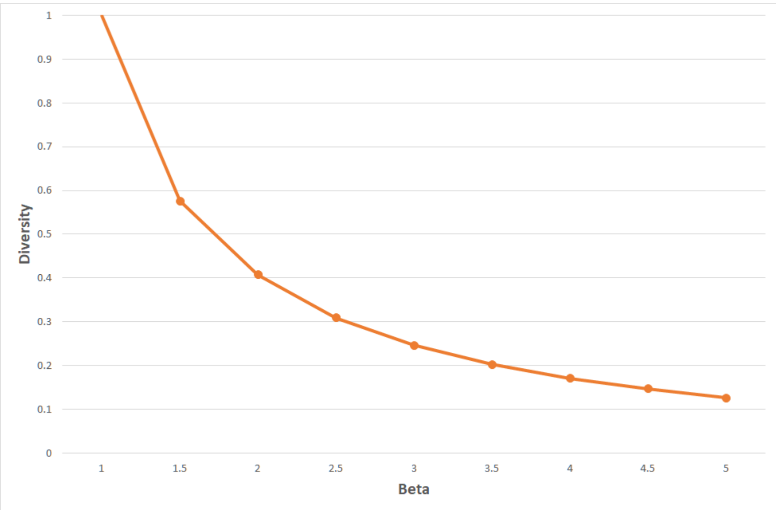


Figure 15: Relationship between β and diversity.

Parameters $n = 50$, $r = 2$, $p = q = 1$.

Table 2: Welfare with $p = 0.7$

q	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.60	0.70
Panel A: Welfare on Euclidean societies											
$r = 2$, direct marriages											
Dv	0.00	0.13	0.25	0.35	0.44	0.52	0.59	0.72	0.83	0.92	1.00
St	0.84	0.85	0.86	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88
Sz	0.98	0.97	0.97	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99
$r = 2$, long marriages											
Dv	0.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.85	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages											
Dv	0.00	0.28	0.45	0.57	0.66	0.73	0.79	0.87	0.92	0.97	1.00
St	0.84	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.91	0.92
Sz	0.98	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99	0.99	1.00
$r = 5$, long marriages											
Dv	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.85	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Panel B: Welfare on assortative societies											
$r = 2$, direct marriages											
Dv	0.00	0.14	0.25	0.35	0.44	0.52	0.59	0.72	0.83	0.92	1.00
St	0.83	0.84	0.85	0.85	0.85	0.85	0.86	0.86	0.86	0.86	0.87
Sz	0.98	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.99	0.99	0.99
$r = 2$, long marriages											
Dv	0.00	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.84	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages											
Dv	0.00	0.28	0.45	0.57	0.66	0.73	0.79	0.87	0.93	0.97	1.00
St	0.83	0.86	0.87	0.88	0.88	0.89	0.89	0.90	0.90	0.90	0.90
Sz	0.98	0.97	0.97	0.98	0.98	0.99	0.99	0.99	0.99	1.00	1.00
$r = 5$, long marriages											
Dv	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.84	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Average over 10,000 random simulations, $n = 50$.

Sz equals the percentage of agents married.

Standard errors in the order of 1.0e-04.

Table 3: Welfare with $p = 0.5$

q	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
Panel A: Welfare on Euclidean societies									
$r = 2$, direct marriages									
Dv	0.00	0.17	0.32	0.45	0.56	0.66	0.74	0.89	1.00
St	0.83	0.84	0.85	0.85	0.85	0.86	0.86	0.86	0.87
Sz	0.96	0.95	0.96	0.96	0.96	0.97	0.97	0.97	0.98
$r = 2$, long marriages									
Dv	0.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.85	0.88	0.89	0.89	0.89	0.89	0.89	0.89	0.88
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages									
Dv	0.00	0.35	0.55	0.68	0.77	0.83	0.88	0.95	1.00
St	0.83	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91
Sz	0.96	0.96	0.97	0.97	0.98	0.98	0.98	0.99	0.99
$r = 5$, long marriages									
Dv	0.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.85	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Panel B: Welfare on assortative societies									
$r = 2$, direct marriages									
Dv	0.00	0.18	0.33	0.45	0.56	0.66	0.74	0.88	1.00
St	0.58	0.58	0.59	0.59	0.59	0.60	0.60	0.60	0.61
Sz	0.96	0.96	0.96	0.96	0.97	0.97	0.97	0.98	0.98
$r = 2$, long marriages									
Dv	0.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.84	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages									
Dv	0.00	0.35	0.55	0.68	0.77	0.83	0.88	0.95	1.00
St	0.58	0.60	0.61	0.62	0.62	0.63	0.63	0.63	0.64
Sz	0.96	0.96	0.97	0.98	0.98	0.98	0.99	0.99	0.99
$r = 5$, long marriages									
Dv	0.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.84	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Average over 10,000 random simulations, $n = 50$.
Sz equals the percentage of agents married.
Standard errors in the order of 1.0e-04.

Table 4: Welfare with $p = 0.3$

q	0	0.05	1	0.15	0.2	0.25	0.3
Panel A: Welfare on Euclidean societies							
$r = 2$, direct marriages							
Dv	0.00	0.27	0.49	0.66	0.80	0.90	1.00
St	0.80	0.82	0.82	0.83	0.84	0.84	0.85
Sz	0.91	0.92	0.93	0.93	0.94	0.95	0.95
$r = 2$, long marriages							
Dv	0.00	0.87	0.98	1.00	1.00	1.00	1.00
St	0.85	0.88	0.89	0.89	0.89	0.89	0.89
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages							
Dv	0.00	0.49	0.71	0.83	0.91	0.96	1.00
St	0.00	0.05	0.10	0.15	0.20	0.25	0.30
Sz	0.91	0.94	0.95	0.97	0.97	0.98	0.98
$r = 5$, long marriages							
Dv	0.00	0.96	1.00	1.00	1.00	1.00	1.00
St	0.85	0.92	0.92	0.92	0.92	0.92	0.92
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Panel B: Welfare on assortative societies							
$r = 2$, direct marriages							
Dv	0.00	0.28	0.49	0.66	0.80	0.91	1.00
St	0.78	0.79	0.80	0.81	0.82	0.83	0.83
Sz	0.92	0.93	0.93	0.94	0.95	0.95	0.96
$r = 2$, long marriages							
Dv	0.00	0.87	0.98	1.00	1.00	1.00	1.00
St	0.84	0.87	0.87	0.87	0.87	0.87	0.87
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages							
Dv	0.00	0.49	0.71	0.83	0.91	0.96	1.00
St	0.78	0.82	0.84	0.86	0.87	0.88	0.88
Sz	0.92	0.94	0.96	0.97	0.98	0.98	0.98
$r = 5$, long marriages							
Dv	0.00	0.97	1.00	1.00	1.00	1.00	1.00
St	0.84	0.91	0.91	0.91	0.91	0.91	0.91
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Average over 10,000 random simulations, $n = 50$.
Sz equals the percentage of agents married.
Standard errors in the order of 1.0e-04.

Table 5: Welfare with $\frac{p}{q} = 10$

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Welfare on Euclidean societies										
$r = 2$, direct marriages										
Dv	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
St	0.18	0.17	0.17	0.17	0.18	0.18	0.18	0.18	0.19	0.19
Sz	0.75	0.79	0.81	0.83	0.84	0.85	0.86	0.86	0.87	0.87
$r = 2$, long marriages										
Dv	0.75	0.87	0.91	0.94	0.95	0.96	0.97	0.97	0.98	0.98
St	0.34	0.52	0.73	0.88	0.96	0.99	1.00	1.00	1.00	1.00
Sz	0.84	0.88	0.88	0.88	0.89	0.89	0.89	0.89	0.89	0.89
$r = 5$, direct marriages										
Dv	0.91	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
St	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.36	0.36	0.36
Sz	0.76	0.80	0.83	0.85	0.86	0.87	0.87	0.88	0.88	0.89
$r = 5$, long marriages										
Dv	0.60	0.76	0.90	0.96	0.99	1.00	1.00	1.00	1.00	1.00
St	0.87	0.91	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Sz	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Average over 10,000 random simulations, $n = 50$.

Sz equals the percentage of agents married.

Standard errors in the order of 1.0e-04.

Table 6: Welfare with $\beta = 2$

q	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Welfare on Euclidean societies											
$r = 2$, direct marriages											
Dv	0.00	0.06	0.12	0.16	0.21	0.25	0.29	0.32	0.35	0.38	0.41
St	0.85	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.88
Sz	1.00	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00
$r = 2$, long marriages											
Dv	0.00	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41
St	0.85	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$r = 5$, direct marriages											
Dv	0.00	0.14	0.24	0.32	0.38	0.43	0.47	0.51	0.54	0.57	0.59
St	0.85	0.88	0.89	0.90	0.90	0.90	0.91	0.91	0.91	0.91	0.91
Sz	1.00	0.98	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00
$r = 5$, long marriages											
Dv	0.00	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59
St	0.85	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Sz	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Average over 10,000 random simulations, $n = 50$, $p = 1$.
Sz equals the percentage of agents married.
Standard errors in the order of 1.0e-04.

C Appendix C: Interracial Marriages and Population Composition

In this final Appendix, we estimate the number of interracial marriages that would occur in 2015 if the racial composition of the U.S. population would have remained constant since 1980.

The main complication of estimating the adjusted rate of interracial marriage is that there is little data available regarding newlyweds. Only the 1980 U.S. Census and the American Community Survey (ACS) from 2008 to 2015 allow us to identify subjects who recently married. The new marriages in 1980 can be identified using the variables *age* and *age of marriage*. Whenever those two coincide, we know that a couple married within a year of the data collection. In the ACS 2008 – 2015, married subjects were asked directly whether they married within the last year. The data is available at https://usa.ipums.org/usa. We use the 1 percent samples.

We obtained the percentage of subjects that married interracially by race in 1980 and 2015. Hispanics were not recorded as a race in 1980, so we estimate which percentage of other races are Hispanics. The races we consider are White, Black, Native Americans, Asians, and Hispanics. Also from the Census and the ACS, we obtain the racial composition of the U.S. both in 1980 and in 2015 (Table 7), and estimate the interracial marriages that would occur with random marriage.[30]

Table 7: U.S. population composition by race, in percentage

Race	1980	2015
White	80	64
Black	11.6	12.2
Native	0.5	0.7
Asian	1.5	4.84
Hispanic	6.5	16.3
Multiracial	0	2

Source: Authors' analysis of 1980 decennial census and 2015 ACS.

Random marriage is easy to compute. If 80% of Americans were White in 1980, a White American had a 0.2 probability of intermarrying. This is

[30] The Stata code is available at www.josueortega.com.

5.33 times larger than the real intermarrying rate for Whites in 1980, which was 0.0375% only (Table 8). Our constructions of interracial rates by race are different from those by Lee and Edmonston (2005) as we estimate the race for Hispanics in 1980. Hispanic was not considered a race in the 1980 decennial census, despite the fact that Hispanics have problems identifying themselves with any of the major races.

Table 8: U.S. interracial marriage rate by race

Race	1980	2015
White	3.8	10.8
Black	5.6	20.0
Native	51.5	55.3
Asian	24.1	32.3
Hispanic	27.3	30.2

Source: Authors' analysis of 1980 decennial census and 2015 ACS.

We fix the 1980 ratio between actual and predicted interracial marriages. We use this ratio to compute the interracial marriage rate that would have occurred in 2015 with the population composition of 1980, using the prediction obtained from random marriage.[31]

Our estimates suggest that, even accounting for demographic change, the percentage of interracial marriages in our society increases by 30% to 37%. While this estimation needs to be taken with care, due to the limitations of the data available, it adds further evidence to our claim that increase in the number of interracial marriages in our societies cannot be due exclusively to changes in the composition of the U.S. population by race. Furthermore, if the increase in interracial marriage we observe in the data was due to changes in the population composition, the intermarriage rates for Black Americans should remain relatively constant over time, just like the fraction of the U.S. population that is Black. However, we observe that the intermarriage rates for Black Americans more than triplicate from 1980 to 2015, as described in the main text.

[31] A similar estimation appears on "Why is interracial marriage on the rise?", *Priceonomics*, 1/9/2016.