

# Can everyone benefit from social integration?\*

Josué Ortega

University of Glasgow

[j.ortega-sandoval.1@research.gla.ac.uk](mailto:j.ortega-sandoval.1@research.gla.ac.uk),

[www.josueortega.com](http://www.josueortega.com)

May 24, 2017

## Abstract

There is no matching mechanism that satisfies integration monotonicity and stability. If we insist on integration monotonicity, not even Pareto optimality can be achieved: the only option is to remain segregated.

A weaker monotonicity condition can be combined with Pareto optimality but not with path independence, which implies that the dynamics of social integration matter.

If the outcome of integration is stable, integration is always approved by majority voting, but a non-vanishing fraction of agents always oppose segregation. The side who receives the proposals in the deferred acceptance algorithm suffers significant welfare losses, which nevertheless become negligible when societies grow large.

**KEYWORDS:** integration, desegregation, integration monotonicity, interracial marriage, school desegregation.

**JEL CLASSIFICATION:** C78 (matching theory).

---

\*I have received helpful comments from Christopher Chambers, Julien Combe, Takashi Hayashi, David Manlove, Hervé Moulin, and two anonymous referees from the EC 2017 workshop on mechanism design for social good. I am also grateful to Mark Meenan for his help using the high performance computing facilities at the University of Glasgow. All errors are mine.

## 1. OUR PROBLEM AND ITS RELEVANCE

Imagine several completely isolated communities that match within themselves, but that could expand their boundaries to merge and match as a unified community instead. The question I ask is whether every person would prefer that all communities integrate as one, provided that the matching outcome is either stable or efficient. Some examples that motivate my research question are:

1. **Interracial marriage.** Three infamous cases of societies that banned interracial marriage are i) the U.S. before the *Virginia vs Loving* case in 1967 ([Arrow, 1998](#); [Fryer, 2007](#)), ii) Nazi Germany, where the marriage between arians and non-arians was forbidden ([Caestecker and Fraser, 2008](#)), and iii) South Africa during the apartheid era, when the *the Prohibition of Mixed Marriages Act* was established ([Hyslop, 1995](#)).

In all cases, social integration occurred only after complicated social movements, and without unanimous approval.

2. **Centralized Kidney Exchange.** After 2000, kidney exchanges began to take place internally in hospitals around the U.S. Few years after, centralized programs started to conduct regional kidney exchanges by asking hospitals to share their donor-patient pairs. Using a centralized procedure would always weakly increase the number of transplants, yet it has been noticed that some hospitals may not have incentives to integrate into to the central clearinghouse, preferring to conduct exchanges only internally. The aforementioned rejection to integrate to a centralized clearinghouse has been documented in practice ([Ashlagi and Roth, 2014](#)).
3. **School Desegregation** In 1954, school segregation was declared illegal in the U.S. following the *Brown vs Board of Education* case. Although the desegregation ruling was widely acknowledged as a major accomplishment, it was not well-received by some. A shameful example is the resistance by the governor of Arkansas, who tried to prevent a few Black students from attending a newly desegregated school. The students were able to enter the school only when they were escorted by federal forces ([U.S. Commission of Civil Rights, 1977](#)).

All these examples show instances on which social integration was complicated to achieve in matching environments. I formalize these environments with an extended version of the [Gale and Shapley \(1962\)](#) matching problem with non-transferable utility. The idea of the model is simple: we take several classical Gale-Shapley problems and find their segregated women-optimal stable partner. Then we put

them together and compute the integrated women-optimal stable matching, and compare who prefers the integrated matching to the segregated one. Stability is natural requirement to ask for, because decentralized matching environments produce outcomes close to those predicted by stability (Hitsch et al., 2010; Banerjee et al., 2013), and because centralized mechanisms that produce stable outcomes are widely regarded as successful (Roth, 2002).

I derive several impossibilities showing that social integration cannot benefit everyone whenever the matching outcome is stable or efficient. The impossibility disappears when only weak integration monotonicity and Pareto efficiency are required. I also prove that social integration is always approved by a weak majority of the population.

Interestingly, such majority rarely surpasses 80% of the population when considering random instances of matching problems, emphasizing the complicated dynamics of desegregation, and making evident that a non-vanishing minority may oppose social integration because they foresee that they will be worse off belonging to an integrated community. Also interestingly, the welfare losses for those who oppose integration become negligible with respect to the size of the grand society as communities grow large, but in small societies the side who receives the proposals in the deferred acceptance algorithm suffers significant welfare losses. Finally, it is also surprising that those who oppose integration are indistinguishable in terms of expected ranking from those who prefer integration.

I use a one-to-one matching framework, but all the impossibility results obviously extend to the many-to-one matching case. Since the interracial marriage example is the one closer to one-to-one matching, I will present the model in those terms, but the reader should keep in mind that the results apply to general matching problems.

After reviewing the related literature in Section 2, Section 3 introduces the model and defines the integration monotonicity property. Section 4 presents several impossibility theorems regarding the existence of integration monotonic matchings that have also any degree of efficiency, emphasizing the rigid structure that monotonicity imposes in matching problems.

Section 5 details the limits of our impossibility results, showing that although some agents oppose social integration, they are always a minority. It also describes the size of such group when the societies become large. Section 6 explores the properties of those who oppose integration and their welfare loss. Finally, section 7 concludes.

## 2. RELATED LITERATURE

### 2.1 Comparative Statics in Matching Problems

Within the matching literature, there is a body of work that studies how the set of stable outcomes changes when a new agent joins an existing society. The main result in this literature is that, when a new man joins a stable matching problem, every woman weakly improves, while every man becomes weakly worse off. This result is robust to various formulations of the problem such as many-to-one extensions and preferences determined by choice functions: see theorem 5 in [Kelso and Crawford \(1982\)](#), theorems 2.25 and 2.26 in [Roth and Sotomayor \(1992\)](#), theorems 1 and 2 in [Crawford \(1991\)](#), and theorem 2 in [Chambers and Yenmez \(2017\)](#).

The aforementioned welfare loss that men suffer when a new man joins the problem has been recently quantified, by assigning agents with independent random preferences over all their partners. [Pittel \(1989\)](#) shows that, in expectation, the side of the society that proposes in the deferred acceptance algorithm, say men, gets matched to a woman ranked  $\log(n)$  in his preference list, whereas women get in expectation a man in the  $\frac{n}{\log(n)}$  position of their rank. The convention is that the partner ranked as 1 is the best possible partner, and so on, and there are  $n$  potential partners for each gender.

Using the same probabilistic framework as [Pittel](#), [Ashlagi et al. \(2017\)](#) find that just by adding an additional man, men receive a partner ranked  $\frac{n}{\log(n)}$  with high probability (in the men-optimal matching), whereas women will receive someone close to  $\log(n)$ . Interestingly, all stable marriages are similar whenever societies are even slightly unbalanced in their ratio between men and women.

In all cases, the discussion centers on what happens when adding an individual alone to a society and not when merging isolated societies of same size. This is my main contribution with respect to the surveyed literature.

### 2.2 Integration and Population Monotonicity Elsewhere

[Chambers and Hayashi \(2017\)](#) introduced integration monotonicity and derived similar results for economic integration. They consider several exchange economies, in which each agent has an initial endowment, that integrate as one. They find that there is no path-independent exchange mechanism that is integration monotonic and Pareto efficient. If the integration mechanism is Pareto efficient and satisfies the additional property of equal treatment of equals, it must necessarily harm one third of all agents in the economy.

[Sprumont \(1990\)](#) considers population monotonic schemes in cooperative games with transferable utility. An allocation scheme is population monotonic if each time an agent joins an existing problem, the payoff for every existing member increases.

He shows that every convex game admits a population monotonic allocation scheme, and provides a tighter characterization using linear combinations of games with veto control. His work deals with transferable utility games only.

Related population monotonicity concepts in cooperative games are widely used in different environments, based on the seminal work of Moulin and Thomson (Thomson (1983), Moulin (1990, 1992); Moulin and Thomson (1988)). This paper's title is inspired on the last of those articles. As in the matching literature, all these monotonicity concepts deal with adding an agent to a problem instead of merging problems of the same size. Sprumont (2008) presents a detailed review of the work in this area.

### 3. MODEL

Let  $S^k$  be a **society** of race  $k$  that consist of  $n$  men  $M^k$  and  $n$  women  $W^k$ . I refer to man  $i$  (woman  $j$ ) that belongs to society  $k$  by  $m_i^k$  ( $w_j^k$ ). When I refer to an agent of arbitrary gender I use  $x_i^k$ ; I omit the subindices when  $n = 1$ . There are  $r \geq 2$  races and  $R = \{1, \dots, r\}$ . For any subset  $T \subseteq R$ , let  $M^T = \bigcup_{k \in T} M^k$ ,  $W^T = \bigcup_{k \in T} W^k$ , and  $S^T = M^T \cup W^T$ .  $S^R$  is called the grand society.

Each man (woman) has strict preferences over the entire set of women  $W^R$  (men  $M^R$ ) and not only over those belonging to her own race. I represent the preferences of an arbitrary person  $x_i^k$  by  $P(x_i^k)$ .  $F(x_i^k)$  denotes the weak preference relation associated to  $P(x_i^k)$  so that for any two agents  $y_j^l$  and  $z_g^e$ ,  $y_j^l F(x_i^k) z_g^e$  if and only if  $y_j^l P(x_i^k) z_g^e$  or  $y_j^l = z_g^e$ . I assume that every person prefers matching with any potential partner of the opposite gender than remaining alone.

$P^T$  will denote the preference lists of every person in  $S^T \subseteq S^R$ . The pair  $(S^R, P^R)$  is an **interracial matching problem** (IMP).

A **matching**  $\mu : M^R \cup W^R \times 2^R \rightarrow M^R \cup W^R$  is a mapping such that,  $\forall T \subseteq R$ ,

$$\forall m_i^k \in M^T, \quad \mu(m_i^k, T) \in W^T \tag{1}$$

$$\forall w_j^k \in W^T, \quad \mu(w_j^k, T) \in M^T \tag{2}$$

$$\forall x_i^k \in S^T, \quad \mu(\mu(x_i^k, T), T) = x_i^k \tag{3}$$

so that every man is married to a woman in the specified society  $S^T$  and vice versa. The function  $\mu$  indicates who marries whom under every union of races. Naturally,  $\mu(x_i^k, T)$  is only defined whenever  $x_i^k \in T$ .

In the majority of matching literature, a matching is defined instead as a mapping  $\mu' : M \cup W \rightarrow M \cup W$ . My definition of  $\mu$  corresponds to the one of an allocation scheme of the matching  $\mu'$ , as defined by Sprumont (1990), which specifies a matching

$\mu'$  for each subset of  $R$ . For convenience, I just refer to such allocation scheme as a matching.

Let  $T$  and  $Q$  be an arbitrary partition of  $R$ , with  $T = \{a, b, \dots, r'\}$ . Let the **colorblind equivalent** of  $S^T$  be denoted by  $S^{\bar{T}}$ , in which every agent  $x_i^k \in S^T$  becomes of a new race  $\overline{ab \dots r'}$ , and in which the preferences of each agent in  $S^R$  remain the same up to the renaming in agents' race.

We define the following properties of interest for an arbitrary matching  $\mu$ , given an IMP.

**Pareto Optimality** There is no different matching  $\mu'$  such that, for all  $T \subseteq R$  and all  $x_i^k \in S^T$

$$\mu'(x_i^k, T) \succ F(x_i^k) \succ \mu(x_i^k, T), \quad (4)$$

and for some  $Q \subseteq R$  and some  $y_j^l \in S^R$ ,

$$\mu'(y_j^l, Q) \succ P(y_j^l) \succ \mu(y_j^l, Q) \quad (5)$$

Pareto optimality is a classical requirement and a basic efficiency concern. It can be strengthened to the stronger efficiency concept of stability.

**Stability** For every subset  $T \subseteq R$ , and for every  $m_i^k, w_j^l \in S^T$ , such that

$$m_i^k \notin \mu(w_j^l, T) \quad \text{and} \quad w_j^l \notin \mu(m_i^k, T) \quad (6)$$

either

$$\mu(m_i^k, T) \succ P(m_i^k) \succ w_j^l \quad \text{or} \quad \mu(w_j^l, T) \succ P(w_j^l) \succ m_i^k \quad (7)$$

Stability is an important requirement because it closely predicts realized outcomes in decentralized environments and because if the final outcome was not stable, it would be unlikely that it lasted long from a game-theoretical perspective.<sup>1</sup> Now we turn to integration monotonicity.

**Integration Monotonicity** For all disjoint subsets  $T, Q \subseteq R$ , and for every  $x_i^k \in S^T$

$$\mu(x_i^k, T \cup Q) \succ F(x_i^k) \succ \mu(x_i^k, T) \quad (8)$$

Note that integration monotonicity not only requires that the matching obtained when all races have integrated is better than the one obtained with a society alone. It requires that anytime another race joins, it always benefits every agent in the existing societies. We will relax this requirement in Section 5. Finally, we define path independence.

---

<sup>1</sup>For example, realized romantic pairings are similar to those predicted by stability, see [Hitsch et al. \(2010\)](#) and [Banerjee et al. \(2013\)](#). Stability is also related to the successful operation of centralized matching mechanisms such as kidney exchanges programs and school choice ([Roth, 2002](#)).

**Path Independence** For all disjoint subsets  $T, Q \subseteq R$ , and for every  $x_i^k \in S^R$

$$\mu(x_i^k, T \cup Q) = \mu(x_i^k, \bar{T} \cup Q) \quad (9)$$

where  $\bar{T}$  denotes the colorblind equivalent of  $T$ .

Path independence is a more technical requirement, but nevertheless relevant because if a matching violates path independence, the dynamics of integration would play a role in determining the final pairings.

From our four properties, only stability and Pareto optimality are related.<sup>2</sup>

**Lemma 1.** *Every stable matching is Pareto optimal.*

*Proof.* Let  $\mu$  be stable. Therefore in any alternative matching  $\mu'$  that is better for person  $x_i^k$  at  $T \subseteq R$ , we have that  $\mu(\mu'(x_i^k, T), T) P(\mu'(x_i^k, T)) x_i^k$ , which means the new partner of  $x_i^k$  prefers matching  $\mu$  to  $\mu'$  at  $T$ , and hence  $\mu'$  is not a Pareto improvement.  $\square$

The converse statement is clearly not true.

#### 4. RESULTS

Unfortunately, stability and integration monotonicity are not compatible even with just two societies with two persons each.<sup>3</sup>

**Proposition 1.** *Not every IMP admits a matching that satisfies stability and integration monotonicity.*

*Proof.* (Example 1) Let  $R = \{a, b\}$  and  $n = 1$ , and let agents' preferences be

$$\begin{aligned} \mathbf{m}^a : w^a P(m^a) w^b & & \mathbf{w}^a : m^b P(w^a) m^a \\ \mathbf{m}^b : w^a P(m^b) w^b & & \mathbf{w}^b : m^b P(w^b) m^a \end{aligned}$$

The unique stable matching has  $\mu(m^k, \{k\}) = w^k$  for  $k \in \{a, b\}$  but  $\mu(w^a, R) = m^b$  and  $\mu(w^b, R) = m^a$ . Yet  $\mu$  violates integration monotonicity for  $m^a$  because  $\mu(m^a, \{a\}) P(m^a) \mu(m^a, R)$ . The same occurs for  $w^b$ .  $\square$

Given that stability and integration monotonicity are incompatible, an obvious question is whether we can weaken any of those two properties to avoid the impossibility. To address it, let us define a particular matching, called the **segregated matching**.

<sup>2</sup>In exchange economies integration monotonicity and efficiency imply core stability (Lemma 2 in [Chambers and Hayashi \(2017\)](#)). A similar conclusion applies in the housing model of [Shapley and Scarf \(1974\)](#). For two-sided matching that relationship does not hold.

<sup>3</sup>[Sprumont \(1990\)](#) proves a similar result: any assignment game with two men and two women lacks a population monotonic assignment scheme. His result does not imply any of mines because his works deals with transferable utility games only.

Let  $\lambda$  be a matching such that  $\lambda(x_i^k, T)$  assigns to each agent  $x_i^k$  the women-optimal stable matching<sup>4</sup> in the matching problem  $(M^T, W^T; P^T)$  for each  $T \subseteq R$ . The segregated matching  $\sigma$  is defined as

$$\forall T \subseteq R, \forall x_i^k \in S^R, \quad \sigma(x_i^k, T) = \sigma(x_i^k) = \lambda(x_i^k, \{k\}) \quad (10)$$

so that for any subset  $T$ , it assigns to each individual the women-optimal matching obtained when matching each race alone. The segregated matching is clearly integration monotonic, but it fails to be stable when aggregating the individual societies. The segregated matching even fails Pareto optimality, as the preferences in Example 2 shows.

*Example 2:* The segregated matching is not Pareto optimal.

$$\begin{array}{ll} \mathbf{m}^a : w^b P(m^a) w^a & \mathbf{w}^a : m^b P(w^a) m^a \\ \mathbf{m}^b : w^a P(m^b) w^b & \mathbf{w}^b : m^a P(w^b) m^b \end{array}$$

I start by weakening stability and requiring Pareto optimality only. Can we obtain always a matching that is Pareto optimal and integration monotonic? The answer is that not even such weakening of optimality is enough.

**Proposition 2.** *Not every IMP admits a Pareto optimal and integration monotonic matching.*

*Proof.* (Example 3) Let  $R = \{a, b, c\}$  and  $n = 3$ , and let agents' preferences be

$$\begin{array}{ll} \mathbf{m}^a : w^b P(m^a) w^c P(m^a) w^a & \mathbf{w}^a : m^b P(w^a) m^c P(w^a) m^a \\ \mathbf{m}^b : w^c P(m^b) w^a P(m^b) w^b & \mathbf{w}^b : m^c P(w^b) m^a P(w^b) m^b \\ \mathbf{m}^c : w^a P(m^c) w^b P(m^c) w^c & \mathbf{w}^c : m^a P(w^c) m^b P(w^c) m^c \end{array}$$

Any Pareto optimal matching  $\mu$  has  $\mu(w^a, \{a, b\}) = m^b$  and  $\mu(w^a, \{a, c\}) = m^c$ . Therefore  $\mu(w^a, R) = m^b$  by integration monotonicity. But exactly the same argument for  $m^b$  shows that he gets  $\mu(m^b, R) = w^c$ . Therefore, no Pareto optimal matching satisfies integration monotonicity.  $\square$

Note that [Chambers and Hayashi \(2017\)](#) are always able to find a mechanism that is Pareto optimal and integration monotonic, although not path-independent.

---

<sup>4</sup>I always pick the women-optimal stable matching to have a consistent selection from the set of stable matchings. We could consider the men-optimal one as well. The selection problem is not a big issue, as in large societies there is a unique stable matching whenever agents have short preferences or the societies are unbalanced in their ratio between men and women ([Immorlica and Mahdian, 2005](#); [Kojima and Pathak, 2009](#); [Ashlagi et al., 2017](#)).

Therefore, the impossibility we obtain is stronger. A first conclusion is that *achieving complete social integration is more difficult than obtaining complete economic integration*.

An immediate corollary follows, showing that if one is to pursue integration monotonicity, there is no room for efficiency even in its weakest form.

**Corollary 1.** *The only matching that satisfies integration monotonicity in every IMP is the segregated matching.*

Given the negative result obtained, let us focus on matchings that satisfy a more flexible monotonicity condition, defined below.

**Weak Integration Monotonicity** For any race  $k \in R$ , and for every  $x_i^k$ ,

$$\mu(x_i^k, R) \succ F(x_i^k) \succ \sigma(x_i^k) \quad (11)$$

Weak integration monotonicity only requires that the corresponding matching when all races have integrated is better than the segregated matching when all societies are segregated. It says nothing about the relationship between matchings obtained under partial integration.

This mild monotonicity is still inconsistent with stability, as our previous Example 1 shows. It can be combined with optimality, yet not without consistency problems.

**Proposition 3.** *Every IMP admits a matching that is weakly integration monotonic and Pareto efficient. If we add path-independence, we obtain an impossibility.*

*Proof.* For any  $T \subseteq R$ , implement the matching  $\sigma$ . If the segregated matching is not Pareto optimal, then implement a Pareto optimal matching  $\mu$  that dominates  $\sigma$ , and so on. Trivially, every agent is better off. Note that every agent has a veto power over stable matchings that benefits others but hurt her/him.

To show that path-independence cannot be added, consider the society in Example 3. Let us merge societies into their colorblind equivalents:  $a$  and  $b$  into  $\overline{ab}$ , and  $a$  and  $c$  into  $\overline{ac}$ . The unique Pareto optimal matching  $\mu$  is such that  $\mu(m_1^a, \{\overline{ab}, c\}) = w_1^b$ , but  $\mu(m_1^a, \{\overline{ac}, b\}) = w_1^c$ .  $\square$

## 5. THE LIMITS OF SEGREGATION

How many people prefer segregation over complete integration, provided that the (women-optimal) stable matching will realize when societies merge? If there will be a referendum asking whether all individual societies should merge, could it be that segregation would obtain a majority of votes?

Let us assume that everybody who does not get hurt by integration votes in favor of it. In Example 1, half of the society votes against integration. Can it be more? The answer is no.

**Proposition 4.** *For any IMP  $(S^R, P^R)$ , at most  $\lfloor rn \rfloor$  agents prefer segregation. The bound is tight.*

*Proof.* Let us partition  $S^R$  into three sets  $A$ ,  $B^+$  and  $B^-$ , defined as

$$A = \{x_i^k \in S^R \mid \lambda(x_i^k, R) = \sigma(x_i^k)\} \quad (12)$$

$$B^+ = \{x_i^k \in S^R \mid \lambda(x_i^k, R) P(x_i^k) \sigma(x_i^k)\} \quad (13)$$

So  $A$  is the set of people who keep the same partner after integration,  $B^+$  are those who prefer their “integrated” partner, and  $B^-$  are those who prefer the “segregated” partner.

Now consider the directed graph which contains all women from  $B^+$  and  $B^-$ , in which every woman points to the woman from whom she “stole” her new husband in the integrated society, i.e. this is each woman  $w_i^k$  points towards  $\sigma(\lambda(w_i^k, R))$ . A cycle always forms whenever  $A \neq S^R$ .

Now consider an arbitrary woman in  $B^+$ , which must exist if  $A \neq S^R$  because  $\lambda$  produces the women-optimal stable matching. She points to a woman  $w_j^l$ , who can either be in  $B^+$  or in  $B^-$ . If she is in the latter, it must be that  $\sigma(w_j^l)$  is in  $B^+$ , because she proposed to him at some point in the woman-proposing deferred acceptance algorithm but he rejected her. This goes on for any woman who is worst off after integration: her previous partner must necessarily be better off after integration, because he rejected her when she proposed to him in the deferred acceptance algorithm.

It follows that  $|B^+| \geq |B^-|$ , and thus  $|A| + |B^+| \geq |B^-|$ , which implies that always at least half of the society supports integration, completing the proof.  $\square$

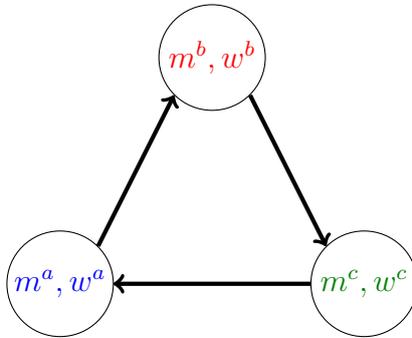


Figure 1: The procedure in the proof of Proposition 4 applied to Example 3.

Proposition 4 is interesting because it tells us that a referendum for integration will always be accepted by a *weak* majority. However, it could be that the voting rule we need is a *supermajority*, that implements integration only if the number of agents that get hurt from integration are at most a fraction  $\epsilon$  of the population.<sup>5</sup>

A natural conjecture is that, for large societies, integration is always approved in any  $\epsilon$ -supermajority, for any arbitrarily small  $\epsilon$ . The conjecture is natural because, when the number of agents grows, agents win a larger pool of potential partners when social integration realizes. Stated differently, that the fraction of people that reject social integration is vanishingly small when the societies become large. Yet, this conjecture is false for small values of  $r$ . Looking at what happens when  $r$  is small is particularly interesting because in reality we have only a few races.

Let  $(S^R, P^R)$  be an IMP in which each agents' preferences are chosen independently and uniformly at random from the set of possible strict preferences. Let  $\Omega_r(n)$  denote the expected number of agents who prefer the segregated matching over the women-optimal stable matching in the grand society.

**Proposition 5.** For  $r \leq 5$ ,

$$\lim_{n \rightarrow \infty} \frac{\Omega_r(n)}{2rn} \neq 0 \quad (14)$$

We know that  $\frac{\Omega_r(n)}{2rn}$  does not vanish from Monte Carlo simulations using Matlab presented in Table 1.<sup>6</sup> The code used is available from my webpage. I stopped the simulations at  $2n = 1000$  because it already took three days to run in a high performance computing facility ( $2n$  is the number of agents of each race). It is clear from Table 1 that convergence occurs in all cases.

Table 1: How many people (in percentage) prefer segregation?  
Average over a thousand simulations with preferences drawn uniformly at random.

$r \setminus 2n$	100	200	1000
2	25.42	25.76	25.47
3	25.43	25.83	25.99
4	24.84	25.14	25.57
5	24.30	24.60	25.05

Our  $\frac{\Omega_r(n)}{2rn} \approx .25$  is related to theorem 2 in [Chambers and Hayashi \(2017\)](#). Their result states that when societies merge, the fraction of people who oppose *economic*

<sup>5</sup>[Chambers and Hayashi \(2017\)](#) use the equivalent concept of *integration monotonicity under P-vetoes*, in which at most a group of people of size  $|P|$  may oppose integration. Both concepts are mild versions of integration monotonicity.

<sup>6</sup>The code was run on the high performance computing facilities of the University of Glasgow, and it uses the Matlab package to compute the women-optimal stable marriage, developed by S. Gopalakrishnan.

integration is always above one third under equal treatment of equals. Their result, looking at a worst case scenario, is obtained in a very different fashion. The comparison of our results suggest a second conclusion: *it is easier to achieve partial social integration than partial economic integration.*

Using the same probabilistic IMP with random preferences, we can find the expected welfare gains derived from integration. Applying the well-known result from [Pittel \(1989\)](#) about expected rankings of partners in random matching problems, it is easy to see that women get a higher ranked partner in expectation after integration occurs, because

$$\underbrace{\log(n) \left( \frac{rn+1}{n+1} \right)}_{\text{exp. ranking w. segregation}} - \underbrace{\log(rn)}_{\text{exp. ranking w. integration}} = \frac{n(r-1)}{n+1} \log(n) - \log(r) \quad (15)$$

which is positive for all sensible values of  $r$  and  $n$ , meaning women get a partner that appears earlier on their preference lists. Similarly, men get a better partner after integration for sensible values of  $r$  and  $n$ , because

$$\underbrace{\frac{n}{\log(n)} \left( \frac{rn+1}{n+1} \right)}_{\text{exp. ranking w. segregation}} - \underbrace{\frac{rn}{\log(rn)}}_{\text{exp. ranking w. integration}} = A[(rn+1) \log(r) - (r-1) \log(n)] \quad (16)$$

where  $A = n/[(n+1) \log(n) \log(rn)]$ . The gains from integration in an IMP with random uncorrelated preferences are given by the sum of the previous expressions multiplied by  $rn$ . The normalized gains from integration for a man and a woman are depicted in Figure 2.

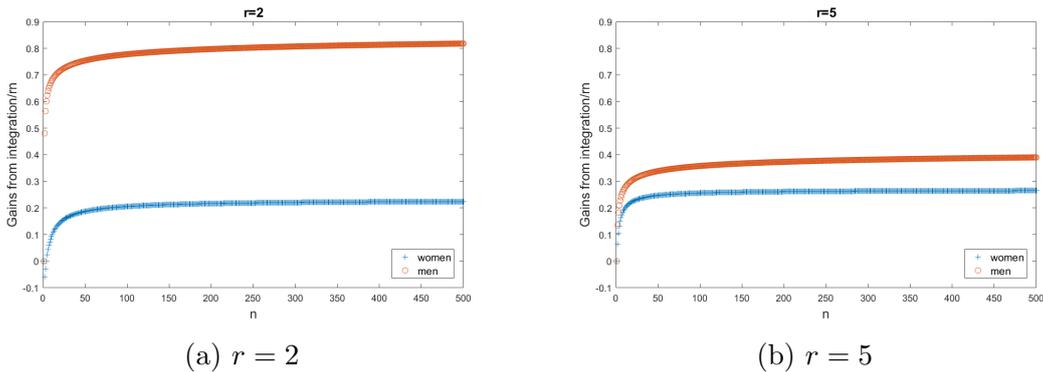


Figure 2: Individual gains from integration divided by  $rn$ , by gender

## 6. WHO PREFERS SEGREGATION?

First let us look at the expected relative number of people who keep the same partner after integration. Since the preferences are drawn uniformly, everybody has the same

probability of matching an agent from their own race: this is  $1/r$ . A natural guess is that, among those,  $1/2$  of them do not change their marriage, which provides a good intuition of the real numbers described in Table 2.

Table 2: How many people (in percentage) keep the same partner after integration?  
Average over a thousand simulations with preferences drawn uniformly at random.

$r \setminus 2n$	100	200	1000	$1/2r$
2	27.26	27.14	27.02	25
3	18.39	18.23	17.89	16.66
4	13.93	13.62	13.41	12.5
5	11.31	11.21	10.84	10

Table 2 shows that, as  $r$  grows, the number of people who are indifferent between integration and segregation becomes smaller. Since the proportion of people who oppose social integration keeps relatively constant as described in Table 1, the number of people who strongly prefer integration does grow, although as we saw it rarely goes over four-fifths of the entire society.

Another natural conjecture is that the people who oppose social integration have a lower expected desirability than those who do not. In other words, they are usually ranked lower in the preference lists of the potential partners. And this new conjecture is false too. Table 3 describes the expected rank of people who prefer segregation: it is immediate that those who prefer segregation have the same expected ranking as a random person, showing that people who prefer segregation are not particularly undesirable agents, they are just like anybody else.

Table 3: Average rank of people who prefer segregation, by gender.  
Average over a thousand simulations with preferences drawn uniformly at random.

$r \setminus 2n$	100		200		1000	
	women	men	women	men	women	men
2	50.7	50.5	100.6	100.5	500.5	500.5
3	75.7	75.5	150.7	150.5	750.6	750.5
4	100.6	100.5	200.6	200.5	1000.7	1000.5
5	125.6	125.5	250.6	250.5	1250.6	1250.5

Finally, we look at the welfare losses suffered by those who prefer segregation when integration realizes, in terms of ranking of their current partner. If their loss was relatively small it would be a strong argument for saying that the impossibilities described in Section 3 are basically irrelevant. Table 4 summarizes an interesting result: the side of the society who does not propose, in this case men, get severely hurt by integration for moderate values of  $n$ . Women, the proposing side, suffer a moderate hurt at most.

Table 4: Average welfare loss by people who prefer segregation, by gender.  
Average over a thousand simulations. Welfare loss measured in difference in ranking of partners.

$r \setminus 2n$	100		200		1000	
	women	men	women	men	women	men
2	4.9	19.7	5.7	34.9	7.4	136.6
3	5.4	27.4	6.2	49.2	7.9	193.8
4	5.7	35	6.5	62.1	8.1	250.8
5	6	41.8	6.8	74.9	8.4	303.3

Table 4 reveals that the welfare loss becomes smaller with respect to the size of the grand society as  $n$  grows, suggesting that they become negligible in the limit. This finding suggests that integration could be more easily implemented in large societies.

The numbers in Table 4 should be understood as a lower bound for the welfare losses, which would increase when agents are endowed with correlated preferences. Correlation in preferences is evident for certain matching environments, like school choice and marriage. We also assume that preferences are independent from race, an ideal scenario but probably not the case in reality for many matching environments, see [Fisman et al. \(2008\)](#) and [Garcia \(2008\)](#) for evidence of racial preferences for dating and school choice, respectively.

## 7. CONCLUSION

When two or more communities integrate to match, there are always some people that become worse off. If the final matching pattern is stable, integration is always approved by a majority of agents, but the fraction of those that oppose social integration does not vanish, even when communities grow large. The welfare losses of those hurt by integration become negligible with respect to the size of the grand society when communities grow large, suggesting that social integration is easier to achieve in sizable communities.

Two interesting questions remain open. The first one is studying the limits of social segregation in many to one matching. The impossibility results carry over, but the question on whether more or less people get hurt by integration, and the exact magnitude of the welfare losses, remains open.

Secondly, there is a recent literature that studies matching in the large using cardinal utilities: e.g. [Che and Tercieux \(2015\)](#) and [Lee \(2017\)](#). Their formulation of preferences makes it easier to introduce correlation, and can provide cardinal measures on the welfare loss of agents that get hurt by integration. Although I conjecture one would obtain similar results using their type of formulation, this remains to be shown formally.

## REFERENCES

- ARROW, K. J. (1998): “What has economics to say about racial discrimination?” *Journal of Economic Perspectives*, 12, 91–100.
- ASHLAGI, I., Y. KANORIA, AND J. D. LESHNO (2017): “Unbalanced random matching markets: The stark effect of competition,” *Journal of Political Economy*, 125, 69–98.
- ASHLAGI, I. AND A. E. ROTH (2014): “Free riding and participation in large scale, multi-hospital kidney exchange,” *Theoretical Economics*, 9, 817–863.
- 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.
- CAESTECKER, F. AND D. FRASER (2008): “The extraterritorial application of the Nuremberg laws. Rassenschande and mixed marriages in European liberal democracies,” *Journal of the History of International Law*, 10, 35.
- CHAMBERS, C. AND T. HAYASHI (2017): “Can everyone benefit from economic integration?” *unpublished*.
- CHAMBERS, C. AND M. B. YENMEZ (2017): “Choice and matching,” *American Economic Journal: Microeconomics*, forthcoming.
- CHE, Y.-K. AND O. TERCIEUX (2015): “Efficiency and stability in large matching markets,” *unpublished*.
- CRAWFORD, V. P. (1991): “Comparative statics in matching markets,” *Journal of Economic Theory*, 54, 389 – 400.
- FISMAN, R., S. S. IYENGAR, E. KAMENICA, AND I. SIMONSON (2008): “Racial preferences in dating,” *Review of Economic Studies*, 75, 117–132.
- FRYER, R. G. (2007): “Guess who’s been coming to dinner? Trends in interracial marriage over the 20th century,” *Journal of Economic Perspectives*, 21, 71–90.
- GALE, D. AND L. S. SHAPLEY (1962): “College admissions and the stability of marriage,” *American Mathematical Monthly*, 69, 9 – 15.
- GARCIA, D. R. (2008): “Academic and racial segregation in charter schools,” *Education and Urban Society*, 40, 590–612.
- HITSCH, G. J., A. HORTAÇSU, AND D. ARIELY (2010): “Matching and sorting in online dating,” *American Economic Review*, 100, 130–63.

- HYSLOP, J. (1995): “White working-class women and the invention of apartheid: Purified Afrikaner nationalist agitation for legislation against mixed marriages, 1934-9,” *Journal of African History*, 36, 57–81.
- IMMORLICA, N. AND M. MAHDIAN (2005): “Marriage, honesty, and stability,” in *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, Philadelphia, USA: Society for Industrial and Applied Mathematics, SODA '05, 53–62.
- KELSO, A. S. AND V. P. CRAWFORD (1982): “Job matching, coalition formation, and gross substitutes,” *Econometrica*, 50, 1483–1504.
- KOJIMA, F. AND P. A. PATHAK (2009): “Incentives and stability in large two-sided matching markets,” *American Economic Review*, 99, 608–27.
- LEE, S. (2017): “Incentive compatibility of large centralized matching markets,” *Review of Economic Studies*, 84, 444.
- MOULIN, H. (1990): “Fair division under joint ownership: Recent results and open problems,” *Social Choice and Welfare*, 7, 149–170.
- (1992): “An application of the Shapley value to fair division with money,” *Econometrica*, 60, 1331–1349.
- MOULIN, H. AND W. THOMSON (1988): “Can everyone benefit from growth?” *Journal of Mathematical Economics*, 17, 339 – 345.
- PITTEL, B. (1989): “The average number of stable matchings,” *SIAM Journal on Discrete Mathematics*, 2, 530–549.
- ROTH, A. E. (2002): “The economist as engineer: Game theory, experimentation, and computation as tools for design economics,” *Econometrica*, 70, 1341–1378.
- ROTH, A. E. AND M. A. O. SOTOMAYOR (1992): *Two-sided matching: A study in game-theoretic modeling and analysis*, Econometric Society Monographs, Cambridge University Press.
- SHAPLEY, L. AND H. SCARF (1974): “On cores and indivisibility,” *Journal of Mathematical Economics*, 1, 23 – 37.
- SPRUMONT, Y. (1990): “Population monotonic allocation schemes for cooperative games with transferable utility,” *Games and Economic Behavior*, 2, 378 – 394.
- (2008): “Monotonicity and solidarity axioms in economics and game theory,” in *Rational Choice and Social Welfare: Essays in Honor of Kotaro Suzumura*,

ed. by P. Pattanaik, T. Koichi, X. Yongsheng, and Y. Naoki, Berlin, Heidelberg: Springer, 71–94.

THOMSON, W. (1983): “The fair division of a fixed supply among a growing population,” *Mathematics of Operations Research*, 8, 319–326.

U.S. COMMISSION OF CIVIL RIGHTS (1977): “School Desegregation in Little Rock, Arkansas,” Tech. rep., Staff Report of the U.S. Commission on Civil Rights.