Endogenous Games, Multiculturalism and Assimilation†

John P. Conley*

and

William S. Neilson**

Revised: March 2009

† We wish to thank Ehud Kalai, Herve Moulin, Simon Wilkie, and the participants of PET07, Vanderbilt for their comments and suggestions. We take responsibility for all remaining errors.
Keywords: Culture, Groups, Language, Endogenous games, Bilateral bargaining, Coordination games, Social norms

* Department of Economics, Vanderbilt University Nashville, TN 37235, j.p.conley@vanderbilt.edu

** Department of Economics, University of Tennessee, Knoxville, TN 37996 wneilson@utk.edu

Abstract

A diversity of cultures does much to strengthen a society. At the same time, assimilating immigrants gives them better access to all the benefits a society offers. This paper explores the tension between diversity and assimilation in an endogenous games framework which formally models the notion that most interactions in work and personal life are voluntary in nature. People encounter each other randomly and decide whether they benefit more from continued search or from associating with one another. If agents benefit from interacting with similar agents, we find that there is a strong bias towards assimilation and the emergence of a single culture. This result can be overturned, however, by such factors as heritage, cultural affinities, caste, and government policies. Finally, we show that if agents benefit from interacting with agents dissimilar from themselves, not only will different cultures coexist, but also there must necessarily arise a class of cultural market makers who specialize in bridging the cultural divide.

1. Introduction

Empirical evidence suggests that peer effects, neighborhood effects, and social networks are important determinants of success in many aspects of life, including employment (Topa, 2001; Conley and Topa, 2002), education (Hoxby, 2000; Hanushek, Kain, Markman, and Rivkin, 2003; Ding and Lehrer, 2007; Friesen and Krauth, 2007), and avoiding prison (Case and Katz, 1991; Calvo-Armengol, Patacchini, and Zenou, 2005). Furthermore, theoretical models provide a link between peer and neighborhood effects and poverty traps (de Bartolome, 1990; Benabou, 1993; Durlauf, 1996; Sethi and Somanathan, 2004). One possible implication of these findings is that introducing higher-income peers into a neighborhood or social network can foster improvements for the less-fortunate incumbents.

Consistent with these studies, Card and Rothstein (2007) find that the black-white test score gap is larger in more racially-segregated cities. However, they also find that test score gaps are independent of school segregation once one accounts for neighborhood segregation. One explanation might be that students who attend an integrated school may nevertheless socially self-segregate once they arrive. The same reasoning about self-segregation might explain Hanushek, Kain, and Rivkin's (2002) finding that the negative effects of increasing the percentage of black schoolmates are restricted to other black students, especially those with high-ability.

Recent studies of college student social networks shed further light on the endogenous vs. exogenous nature of peer and neighborhood effects. Mayer and Puller (2008) use data from Facebook.com, a university social networking website, to examine factors affecting the likelihood that students of different racial or ethnic backgrounds list each other as "friends." They find that friendship is driven primarily by post-meeting preferences over personal attributes, and not by the probability that two people from different groups meet. Thus, while agents meet more or less randomly, deciding if they should continue the relationship is a choice made based on how well the characteristics of the two agents match. Marmaros and Sacerdote (2006), using data on email exchanges among students at Dartmouth College, find that race, geographic proximity (e.g. same dormitory), and entering class all greatly increase the probability of social interactions. Mayer and Puller find a large "friends-of-friends" effect, but Marmaros and Sacerdote do not.

Together, these studies raise the issue of whether individuals take advantage of positive (or negative) peer or neighborhood influences. In this paper we ask the following question: Can two geographically-mingled subgroups coexist, adopt different behavioral norms, and interact only with others in the same subgroup? If so, then integrating a neighborhood or school might not lead to positive outcomes for the less-advantaged subgroup. If not, then integration necessarily means a merging of cultures.

In general, partnering with someone with different cultural cues and expectations can introduce difficulties and reduce the value of a partnership. An Irishman might want to go to a pub after work to consider how to deal with a problem. A Muslim would not. "Yes" in Japan means something very different than it does in New York. French workers have different expectations about job security and working hours than workers in Vietnam.

While agents may be endowed with different cultures initially, agents can also choose to change their outlooks and become assimilated. The central theme of this paper is to show what factors push agents to adopt a more advantageous set of behavioral norms in hopes of finding better partners when interactions are voluntary.

More formally, we consider an *endogenous game* with two stages.[1] The first stage is a kind of pregame in which there are several modes of behavior (actions) and individuals choose to "learn" a subset of these behaviors. In the second stage, individuals are randomly matched, observe the behavioral modes available to their potential partners then choose either to play a coordination game against the matched partner or to pay a waiting cost and get rematched next period. If they both choose to play, they must

[1] An endogenous game is one in which players can restrict their actions sets and refuse to play against matched opponents. In contrast to other games where the set of players, their sets of actions, and the payoff function are all exogenous, only the payoff function is taken as given in an endogenous game. Conley and Neilson (2009) looks specifically at prisoners' dilemmas and shows that a commonly agreed upon level of cooperation (which can be viewed as a "social norm") is the only equilibrium outcome. Here we use the endogenous game structure to analyze behavior in a coordination game and look for conditions under which multiple social norms can coexist.

employ one of the actions they learned in the pregame. The coordination game has a positive payoff if both players play the same action and zero if they play different actions.

To make this concrete, consider a language example. Agents can learn English, Spanish, or both. Two agents can observe the set of languages their potential partners are capable of speaking. They both gain from trade only if they end up speaking the same language. If the waiting cost is small, someone who speaks only Spanish will choose to interact with a Spanish-only partner or a bilingual partner, but not with an English-only partner (and inversely). The question then is, given this ability to refuse interactions, can there be an equilibrium in which some agents speak only Spanish and some speak only English (as we see in reality), or does equilibrium dictate that all agents learn the same language? Our initial analysis of the symmetric game shows that everyone must learn the same language, although some may learn to be bilingual.

Since this is obviously counter-factual, we go on to explore how different social conventions might exist together in equilibrium. We note that the game may not be symmetric in the case of language. In particular, agents might prefer interactions made in their native language to those made in a non-native language. We check the robustness of our prior result that equilibrium prescribes a common language in this new setting. We find that the ability for different monolingual groups to coexist depends on the size of the waiting cost, the magnitude of the preference to interact in the native language, and the size of the native group within society. A society can consist of two monolingual groups that do not interact with each other only if the waiting cost is small, the groups are relatively equal in size, and the preference for speaking the native language is strong.

We also consider a discoordiation game in which agents benefit from interacting with others who have learned a different strategy than they have. For example, agents might learn how to be leaders or followers. Not surprisingly, we find that both these behavioral norms will coexist in equilibrium. What is surprising is that we also find that there must necessarily arise in equilibrium a class of "social market makers" who have learned to pay both social roles. This is true even if there is a cost to learning strategies. In contrast, there is no role for social market makers in coordination games and they disappear entirely in equilibrium if there is even a slight cost of learning new strategies.

The paper precedes as follows. Section 2 gives a formal definition of the model and defines an equilibrium concept called *Pregame Perfect Equilibrium*. Section 3 looks at symmetric coordination games and establishes that one behavioral norm must dominate. Section 4 looks at issues of heritage and preferences for one mode of conduct over another. Section 5 explores a discoordination game in which multiple conventions must coexist in equilibrium. Section 6 relates these results to the literature on neighborhood effects and evolution, and Section 7 offers some concluding remarks.

2. The Model

Consider an economy with I agents indexed $i \in \{1, \ldots, I\} \equiv \mathcal{I}$. Each agent has a set of strategies $\mathcal{X}$ available for him to learn. Agents receive payoffs by finding a partner with whom to play a bilateral game. The payoff to each agent is given by the function $F : \mathcal{X} \times \mathcal{X} \rightarrow \Re$. Thus, the payoff to agent i choosing strategy x_i playing agent j choosing strategy x_j is $F(x_i, x_j)$.

We assume a two-part game. The first part is the *pregame*. Here, each agent chooses to learn a subset of strategies $\ell_i \subseteq \mathcal{X}$. We refer to ℓ_i as agent i's *list*. Denote the set of all possible lists as $\mathcal{L}$ (that is, all possible subsets of $\mathcal{X}$).

The second part of the game is a multistage matching game similar to the one described by Rubinstein and Wolinsky (1985). In each round of the stage game, agents are randomly matched with another agent and can choose one of two options: play or not play.

If both agents decide to Play (P), they choose strategies from their list, settle on a Nash equilibrium and retire from the game. We assume that when an agent retires, he is replaced by an agent with the same name who has chosen to play in the same way. While this is a strong assumption from a theoretical standpoint, it will become clear that in the applications we consider, it has very little bite.

If at least one agent decides to Not play (N), they receive no payoff but pay a *delay cost* of $z > 0$ and search for a new partner in the next round of play.

Formally, an agent's strategy in the second part of the game is a mapping from the strategy list of the agent with whom he is matched to his own list plus N. Since matching is random and anonymous and the population of agent types is stable across periods, we assume that history does not influence strategic choice. Let agent i's second stage strategy be denoted $g_i : \mathcal{L} \rightarrow \ell_i \bigcup N$. Let $\mathcal{G}$ denote the set of all such mappings. We will refer to g_i as *agent i's stage game strategy.*

An agent's strategy therefore consists of a pair $(\ell_i, g_i) \in (\mathcal{L}, \mathcal{G})$. Collectively, we will denote this as $s_i \in \mathcal{S}$. A *strategy profile* for the game is denoted: $S \equiv (s_1, \ldots, s_I) \in \mathcal{S} \times \ldots \times \mathcal{S}$. It will be useful to refer to the profile of lists and stage game strategies separately on occasion. We denote these $L = (\ell_1, \ldots, \ell_I)$, and $G = (g_1, \ldots, g_I)$. We also use the notation s_{-i} to denote the strategy profile for all agents excluding agent i.

Next turn to defining a reasonable equilibrium concept for endogenous games. The essential idea is that agents play subgame perfect strategies based on a reasonable set of consistent beliefs. While straightforward from a conceptual standpoint, it is notationally dense to define formally. We therefore will follow the strategy of putting the detailed formal development of these concepts in the mathematical appendix and focus only on the points that are of real economic and theoretical interest.

One requirement of our equilibrium notion is that agents choose strategies in the pregame that maximize their expected payoffs. We will denote the *expected payoff that agent i receives under strategy profile s* as $EPO(s_i, s_{-i})$ (see the appendix for details).

We also require agents to follow optimal strategies in every subgame. Thus, we will need to know the expected continuation payoff from every subgame. That is, conditional on having been matched with an agent who has chosen list $\hat{\ell}$, we need to know the expected payoff from this point forward. Denote the set of lists that agent i

might encounter during equilibrium play given S by:

$$L_i \equiv \{\ell \in \mathcal{L} \mid \exists j \in \mathcal{I}, j \neq i \text{ s.t. } \ell_j = \ell\}$$

and the set of lists that should not be encountered by agent i given S by:

$$\bar{L}_i \equiv \{\ell \in \mathcal{L} \mid \nexists j \in \mathcal{I}, j \neq i \text{ s.t. } \ell_j = \ell\}$$

Similarly, define the set of lists that might be encountered by *any agent* in equilibrium play given S by

$$L \equiv \{\ell \in \mathcal{L} \mid \exists j \in \mathcal{I} \text{ s.t. } \ell_j = \ell\}$$

and the set of lists that should not be encountered by *any agent* in equilibrium given S by:

$$\bar{L} \equiv \{\ell \in \mathcal{L} \mid \nexists j \in \mathcal{I}, \text{ s.t. } \ell_j = \ell\}$$

The expected payoff for a subgame in which an agent faces an opponent with list $\hat{\ell} \in L_i$, that is, a list agent i expects to encounter with positive probability, is straightforward to compute and denoted by $EPO(s_i, s_{-i} \mid \hat{\ell})$. The expected payoff in a subgame in which agent i faces a list that should *not* be encountered on the equilibrium path, that is, a list $\ell \in \bar{L}_i$, is more problematic. Subgame perfection requires that if agent i and the agent with the deviant list $\ell \in \bar{L}_i$ agree to play, they must play a Nash equilibrium of the stage game. If the stage game has multiple equilibria agent i must act based on beliefs about which of those equilibria will ensue. If agent i chooses not to play against the agent with the deviant list, agent i must recompute the continuation payoffs based on the information that one agent has deviated from the equilibrium path. It could be that this is the only deviator, or it could be that there are multiple deviators. Agent i's choice between playing or passing against the deviant list depends on his beliefs about what will happen if he chooses to play and what will happen if he chooses to pass.

We now turn to the task of specifying a set of consistent beliefs regarding the stage game strategies of both equilibrium and disequilibrium agents as well as beliefs about expected payoffs when a match is refused. Some of these beliefs are pinned down by subgame perfection, but others, especially regarding the expected payoffs relevant when paired with an agent with an out-of-equilibrium list, are not. The definition below extends the notion of credible beliefs to endogenous games.

We use the notation $p(g \mid S, \ell)$ to indicate agents' beliefs regarding the probability that a member of the population who has chosen list ℓ uses stage game strategy g given the strategy profile S. Given a strategy profile S, $p : \mathcal{G} \times \mathcal{S} \times \ldots \times \mathcal{S} \times \mathcal{L} \to [0,1]$ is a *consistent belief system* held in common by all agents if:

1. For all $\ell \in \mathcal{L}$, $\sum_{g \in \mathcal{G}} p(g \mid S, \ell) = 1$ and $p(g \mid S, \ell) \geq 0$.
2. For all $\ell \in L$, $p(g \mid S, \ell) = \frac{\mid \{i \in \mathcal{I} \mid g_i = g \text{ and } \ell_i = \ell\} \mid}{\mid \{i \in \mathcal{I} \mid \ell_i = \ell\} \mid}$
3. For all $\ell \in \bar{L}$, $p(g \mid S, \ell) > 0$ only if there does not exist $\bar{g} \in \mathcal{G}$ such that for all $\hat{\ell}$ such that there exists $i \in \mathcal{I}$ with $\ell_i = \hat{\ell}$[2]

$$EPO(\bar{g}, \ell, S \mid \hat{\ell}) > EPO(g, \ell, S \mid \hat{\ell}).$$

Condition 1 says p is a probability distribution. Condition 2 says the beliefs are consistent with the equilibrium choice of lists by agents. Condition 3 says for lists not seen in equilibrium, the common belief system about agents with such lists places positive probability weight only on stage game strategies that are payoff maximizing.

Given these beliefs, we will finally need to express the continuation payoff for an agent who encounters a list that he should not see in equilibrium. Note that there is a small complication since the out-of-equilibrium lists may be different for different agents. In particular, if *only* a single agent i chooses a given list ℓ then of course agent i would expect never to see this list in an opponent. The remainder of the population, on the other hand, would just assume it had encountered agent i. We have embedded one assumption in this regard in the definition in consistent beliefs: if an agent i who happens to be the only member of the population who chooses a list ℓ encounters this list anyway, he assumes that that this disequilibrium opponent is just like him and

[2] This is a slight abuse of notation. Rather than define a new EPO for disequilibrium agents, we simply pretend the population expanded by 1 to $I + 1$ and use the old definition. Note in this case, S plays the role of s_{-i}

chooses the same stage game strategy, g_i, as he does. This keeps his beliefs consistent with the rest of the agents in the game. It would be more general to allow agent i's beliefs to be unconstrained, but it would complicate notation and would not change any of the conclusions in the next section. We therefore opt for our simpler approach.

Given that strategies are complete contingent plans, every agent must specify a continuation game strategy for each possible list he could choose. The probability an agent with an out-of-equilibrium list plays stage game strategy g can be inferred from the equilibrium strategy profile S. The expected payoff from choosing to play against an out-of-equilibrium list can then be computed in a straightforward manner (see the appendix for details). As for the expected payoff from choosing to not play, we make a "small deviation" assumption that the out-of-equilibrium list disappears after the current round and is replaced by the appropriate equilibrium list. In essence, this means that agents never expect to face disequilibrium lists in the future, regardless of whether they currently face one or not.

We denote the expected payoff agents receive when encountering out of equilibrium lists ($\ell \in \bar{L}_i$) by $\overline{EPO}(s_i, s_{-i} \mid \ell)$. We use a "bar" to remind us that this continuation payoff refers only to out of equilibrium play.

We are now ready to define our equilibrium concept.

Pregame perfection: A strategy profile $s \in S$ and consistent belief system p is a pregame perfect equilibrium (PPE) if

1. For all agents $i \in \mathcal{I}$, and all $\hat{s}_i \in \mathcal{S}$,

$$EPO(s_i, s_{-i}) \geq EPO(\hat{s}_i, s_{-i})$$

2. For all agents $i \in \mathcal{I}$, all $\ell \in L_i$ and for all $\hat{g}_i \in \mathcal{G}$

$$EPO(s_i, s_{-i} \mid \ell) \geq EPO(\hat{g}_i, \ell_i, s_{-i} \mid \ell)$$

3. For all agents $i \in \mathcal{I}$, all $\ell \in \bar{L}_i$ and for all $\hat{g}_i \in \mathcal{G}$

$$\overline{EPO}(s_i, s_{-i} \mid \ell) \geq \overline{EPO}(\hat{g}_i, \ell_i, s_{-i} \mid \ell)$$

4. For all agents $i \in \mathcal{I}$ and all $\ell \in \mathcal{L}$ if

$$\max_{x_i \in \ell_i} \min_{x \in \ell} F(x_i, x) \geq EPO(s_i, s_{-i}) - z,$$

then $g_i(\ell) \neq N$.

Condition 1 says that for all agents, the strategies chosen are maximize the expected payoffs. Condition 2 says for all agents and for every list it is possible to encounter in equilibrium, the strategy chosen conditioned on this encounter maximizes the expected payoffs. Condition 3 is the same as 2 except that it applies to out-of-equilibrium encounters. Condition 4 says an agent will always agree to play with potential partner when the worst that can happen is that he realizes a weakly higher payoff than he would expect from passing instead. This is necessary to prevent trivial outcomes where all agents pass in response to encountering all lists. Passing, of course, is a weak best response to passing, so this forms a degenerate equilibrium.

A very reasonable question is why don't we simply use subgame perfection or sequential equilibrium instead of pregame erfection? Let us be clear on this point: we do not see PPE as being of much economic or theoretic interest. We are, inf fact, frustrated at not being able to simply applying the existing equilibrium concepts directly. However, it turns out that the particular features of the endogenous game form we study make this impossible, and to be fully rigorous, we are forced to define PPE. The reasons are a bit subtle, so we refer the interested reader to the appendix.

In the next two sections, we maintain the two assumptions that population size, I, is even and $I \geq 4$. The first is needed because this is a matching game, and if the population were odd, we would have to include the possibility of not finding a potential match in any given period. This would needlessly complicate our model. The second is needed because if $I = 2$ agents would always be matched with the same player each period which would make the endogenous of the game degenerate.

3. The coexistence of conventions in a simple coordination game

Let Γ be a simple coordination game with $X = \{A, B\}$ and $F(A, A) = F(B, B) = 1$ and $F(A, B) = F(B, A) = 0$, as shown below.

Coordination Game		
	A	B
A	α, α	0, 0
B	0, 0	β, β

This game has the feature that the two agents receive gains from trade if they choose the same actions, and no gains from trade if they do not choose the same actions. The possible lists are $\{A\}$, $\{B\}$, and $\{A, B\}$. We will refer to an agent with list $\{A\}$ as an A-type individual, one with list $\{B\}$ as a B-type individual, and one with list $\{A, B\}$ as an AB-type individual. The issue we wish to address is the existence of a pregame perfect equilibrium in which some agents choose list $\{A\}$ and some choose list $\{B\}$, and A-types and B-types do not interact with each other. That is, when do agents choose not to interact with people who are different from them, and when can two conventions or social norms coexist?

We first address a narrower question: Can there be a pregame perfect equilibrium with both A-types and B-types but no AB-types, and in which A-types and B-types do not interact? We show below that the answer is "no".

Proposition 1. *There is no pregame perfect equilibrium with $0 < I_A < I$ A-type agents who do not play against B-type agents, $I - I_A$ B-type agents who do not play against A-type agents, and no AB-types at all.*

Proof/

See appendix.

Next, we consider whether this nonexistence result is robust to the introduction of AB-types as intermediaries. When an A-type is matched with an AB-type, the only Nash equilibrium of the restricted base game is for both agents to play A, and so an A-type is indifferent between interacting with another A-type and interacting with an AB-type. B-types are also indifferent between interacting with other B-types and interacting with AB-types. Can these intermediaries allow the different conventions to coexist? As the next proposition shows, the answer is also negative.

Proposition 2. *There is no pregame perfect equilibrium with $I_A > 0$ A-type agents who do not play against B-type agents and $I_B > 0$ B-type agents who do not play against A-type agents.*

Proof/

See appendix.

The intuition behind Propositions 1 and 2 is the same. If an A-type's expected payoff is at least as high as a B-type's, the B-type has an incentive to defect and choose the list $\{A\}$. Switching enlarges the population of A-types, increasing the A-type's payoff advantage. So, switching from list $\{B\}$ to list $\{A\}$ improves the agent's payoff in two ways, by moving him to a type that earns higher expected payoff in the first place, and by increasing the expected payoff of that type by making them more likely to match successfully. Intermediaries who have list $\{A, B\}$ play no role in this argument, and cannot help two conventions to coexist.

Having established that some classes of equilibria do not exist, it remains to establish the classes that do. The next proposition characterized the entire set of PPE equilibria.

Proposition 3. *There exist exactly two classes of pregame perfect equilibria: equilibria with no A-types and equilibria with no B-types.*

Proof/

See appendix.

Applying this analysis to the language example of the introduction, suppose that A-types speak only English, B-types speak only Spanish, and AB-types are bilingual. Propositions 1 and 2 state there is no equilibrium with both English-only speakers and Spanish-only speakers, regardless of whether or not the population contains bilingual people. Proposition 3 establishes the existence of equilibria with a combination of English-only and bilingual agents and equilibria with a combination of Spanish-only and bilingual agents. Moreover, these are the only equilibria, and in equilibrium a single language is sufficient for all interactions. Thus, it appears that cultural assimilation is inevitable.

4. Heritage

The previous section shows that when agents are indifferent between the behavioral norms they use, then as long as they can coordinate, everyone must share a common culture. This suggests that integration in schools or neighborhoods should be enough to generate the benefits of peer effects. However, casual inspection and the evidence presented in the introduction attests to the fact that we do see cultures coexisting within integrated societies. Clearly, the benchmark case considered in the previous section must be missing something. The reasoning from the preceding section suggests that when agents match randomly with the population at large, they benefit from sharing the same behavioral norm as the majority group. Several factors might make it more likely that minority cultures are able to continue to exist in light of the pressures towards assimilation.

- A group that wants to preserve its culture can improve its odds by increasing the likelihood that its members will interact with each other rather than the population at large. For example, members may go to the same church, engage in the same line of work or live in a ghetto.
- Policies may make it cheap or expensive to learn certain strategies. For example, teaching Welsh, Celtic or Catalan in schools, or forbidding Native American children to use anything but English in school will affect assimilation.
- Agents might have a strong preference for one of the coordinated outcomes over

the other. For example, assimilation is more likely if the "affinity" benefits are outweighed by the "large network" benefits of joining the larger community; conversely, cultural segregation might arise when the affinity benefits are large. If the size of a group falls below a critical mass, however, separation is hard to sustain. Given the search cost, it will not benefit agents to hold out in hopes of finding another agent of their affinity type, and agents will instead learn the strategies used by the larger society.

In this section we consider in more detail the case in which there are two types of players: a-natives and b-natives. The first step is to adapt the theory of endogenous games to allow for these "heritage-based" preferences.

To construct an endogenous game from the coordination game with heritage-based preferences, suppose that the numbers of a-natives and b-natives in the population are H^a and H^b, respectively, and these numbers are common knowledge. Set $I = H = H^a + H^b$. In the pregame agents choose lists, as before. In the continuation game, agents are randomly matched, and observe their potential opponents' lists *along with their heritage* (i.e. their opponents' payoff functions). The game proceeds as before, with agents deciding whether to play or pay a delay cost of z and be rematched. When a pair elects to play, they play actions from their lists, receive their payoffs, and exit the game to be replaced by agents with the same heritage and the same strategies. The definitions of consistent beliefs and pregame perfect equilibrium can be extended to this setting in a straightforward manner by replacing lists in the earlier definitions by list-heritage pairs.

First, consider the base game in which

$$F^a(A, A) = F^b(B, B) = \alpha + \eta$$

$$F^a(B, B) = F^b(A, A) = \alpha$$

$$F^h(A, B) = F^h(B, A) = 0$$

for $h = a, b$ where $\eta > 0$. The parameter η captures agents' preferences for their native actions, with increases in η corresponding to stronger preferences for the native action. Since we assume that both heritage and lists are observable, the strategies must be conditioned on both. Formally $g_{\ell}^{h}(h', \ell')$ is the strategy employed by an agent of affinity type h with list ℓ who encounters an agent with affinity type h' with list ℓ'. We are now able to address the question of whether there exists a pregame perfect equilibrium in which agents refuse to interact with agents of the other heritage.

Proposition 4. *If the delay cost z is small enough, there exists a pregame perfect equilibrium in which agents' lists do not contain the non-heritage action and agents play only against others from the same heritage.*

Proof/

See appendix.

Proposition 4 establishes that, if the delay cost z is small relative to the preference for the native action, η, there exists an equilibrium of cultural separation in which a-natives never learn action B and only interact with other a-natives, while b-natives never learn action A and only interact with other b-natives. When the delay cost is small, a-natives find it more beneficial to wait to meet another a -native in order to gain the benefit η than to learn action B and play the first opponent, and b-natives share the same incentives.

The actual condition for "small enough" z is worth a closer look. It reduces to

$$z < \frac{H^j - 1}{H - 1}\eta$$

for $j = a, b$. The restriction on z is looser, and therefore cultural separation is more likely to occur, when the population is more evenly split among the two types and the affinity for native behavior η is large. Conversely, when the population contains a large majority and a small minority, assimilation is more likely to occur.

This analysis highlights the reason why cultural separation can persist in schools or social networks, as discussed in the introduction. Even though agents with different preferences are intermingled and are sometimes matched with members of the other heritage group, if the native preference is strong, the heritage group is relatively large, or the cost of waiting for a match with someone from the same heritage group is small, cultural separation can persist. In the presence of these conditions, integrating two cultural groups is not enough, because while integration can lead to face-to-face encounters, individuals still can refuse to turn these into mutually beneficial relationships.[3]

5. Discoordination Games

In this section, we turn our attention to the "discoordination game" which is in a sense the opposite of the coordination games above. Here, agents want to find a partner who will play the opposite strategy as he does. For example, someone who knows how to cook meat would prefer to partner with someone who knows how to cook side dishes for a picnic. We will argue that not only will it necessarily be the case that multiple conventions coexist, but also that there must exist "social market makers" who can interact with either convention.

Consider the symmetric *discoordination game* in which $X = \{A, B\}$ and $F(A, A) = F(B, B) = 0$, and $F(A, B) = F(B, A) = \alpha$, as shown below.

Discoordination Game		
	A	B
A	0, 0	α, α
B	α, α	0, 0

A pregame perfect equilibrium exists for this game when I is divisible by 3. In the equilibrium $I/3$ agents choose list $\{A\}$, $I/3$ choose list $\{B\}$, and $I/3$ choose the full list $\{A, B\}$. All agents use the same continuation game strategy $g(\ell' \mid \ell)$ given by:

[3] Of course, the converse is also true. That is, if delay cost is large, native preference week, or the heritage group relative small, then cultural segregation becomes more difficult to sustain. At some point, all PPE involve cultural assimilation as in section 3.

$$g_A(\{A\}) = N \qquad g_A(\{B\}) = A \qquad g_A(\{AB\}) = A.$$

$$g_B(\{A\}) = A \qquad g_B(\{B\}) = N \qquad g_B(\{AB\}) = B.$$

$$g_{AB}(\{A\}) = A \qquad g_{AB}(\{B\}) = B \qquad g_{AB}(\{AB\}) = N.$$

It is straightforward to verify that the continuation strategies satisfy the conditions for pregame perfect equilibrium.[4] To see that the choices of lists are consistent with equilibrium, note that $EPO_A(S) = EPO_B(S) = EPO_{AB}(S) = \alpha - \frac{I-3}{2I}z$, and if any agent switches to a different list, his expected payoff changes to $\alpha - \frac{I}{2I-3}z$, which is worse for all $z > 0$.

In a coordination game, agents benefit from playing against others who behave the same as themselves, and pregame perfect equilibrium allows only one mode of behavior to exist in equilibrium. In a discoordination game, in contrast, two modes of behavior *must* coexist. Pregame perfect equilibrium requires the existence of intermediaries to serve as "social market makers," that is, agents with full lists who can interact with both single-list types. Thus, pregame perfect equilibrium in discoordination games precludes full specialization by type.

To understand the discoordination game more fully, consider the following classical interpretation. (We provide two interpretations that relate more directly to culture below.) Two agents are matched and, if they agree to play, receive positive payoffs if they find each other and zero payoffs if they do not. The two available actions are to "go looking" or "remain in one place." If they both remain in one place they fail to find each other, and if they both go looking they also fail to find each other. Successful outcomes occur only when one agent goes looking and the other remains in one place. Parents realize this and teach their children to remain in one place when they

[4] Because all possible lists occur in the equilibrium, the consistent belief requirement that governs beliefs when encountering an off-the-equilibrium-path list plays no part in the equilibrium.

become separated. However, that rule only works for asymmetric populations where parents are always matched with children. What would happen in a population with only adults? In the endogenous game the equilibrium specifies a group that only goes looking (finders), a group that only remains where they are (stayers), and a group that can do either (intermediaries), all of equal size. Finders refuse to play against other finders, stayers refuse to play against other stayers, and intermediaries refuse to play against other intermediaries. If a stayer, say, switches to become a finder, then he joins a now-larger group and has a higher probability of matching with another finder, which is unsatisfactory. The same thing occurs if he switches to become an intermediary.

The difference between the discoordination game and the coordination game, then, and what drives the existence of equilibria with distinct groups in the former but not in the latter, is the property that in a coordination game a player wants to match with someone like himself but in a discoordination game a player wants to match with someone who is different. This, then, is the key to the failure of the coexistence of conventions. Conventions cannot coexist when individuals with one convention benefit from interacting with others who share the same convention but not with agents who follow different conventions.

Discoordination games have interesting implications in the context of culture and behavior we are concerned with in this paper.

Alpha, Beta, and Gamma Males: Consider a situation in which men come together for a partnership to complete a project and share the profits. In each partnership, a leader and a follower are needed. Men develop skill at taking either role, or sometimes, both roles. Examples include things like hunting expeditions, preparing a presentation for the boss, or trying to meet women at a fraternity party. Two alpha-males means two leaders and a failed venture. Two beta-males means no leadership, and a failed venture. Two gamma-males leads to a coordination problem and only half the partnerships consisting of a leader and a follower. Thus, we find that all three types of males will exist in equilibrium and agents will spend time finding the compatible partner. [5]

Caste systems: A society needs people to specialize in, and perform all kinds of work. The analysis above offers a potential explanation for the apparent stability of caste-based societies. Suppose successful partnerships consist of a high-caste and a low-caste member (or even several low-caste members). The major difference here is that in this context search is not an entirely random process. Caste is usually a clearly identifiable characteristic, even from a distance. If a high-caste agent needs a low-caste partner, he can readily identify one and partner without search (and the same is true for a low-caste agent.) Thus, market makers do not arise in this context. More importantly, a low-caste agent would not have any incentive to learn high-caste skills. He would never be asked for a partnership by another low-caste agent. Thus, the endogeneity of the game serves to reinforce the caste-system. If an agent can “pass” for being a member of another type, however the story is different. A low-caste person who wants to provide a high-caste skill must appear outwardly to be high-caste as well in order to have success in finding partners. This might explain the relative persistence of caste distinctions based on race, gender, family connections, and even such characteristics as height and beauty. On the other hand, caste distinctions based on behaviors or appearances that *any* agent can learn and adopt are more difficult to maintain.

6. Cultural Integration and Evolution

Besides relating to peer and neighborhood effects, this paper can be considered as a model of cultural integration, as in Lazear (1999) and Kuran and Sandholm (2008).[6]

[5] Of course, in some (but not all) games like this, it might be possible for gamma-males to meet and then agree who will be the alpha and who will be the beta. In this case all agents would choose to be gamma-males. However, we can recover the result by adding a cost of learning a strategy (or mode of behavior). In this case, the higher the cost the fewer gamma-males we see in equilibrium, but we would still see all types existing together (at least until the cost of learning a second strategy got prohibitive.)

[6] Kuran and Sandholm (2008) use an evolutionary game theory framework in which preferences are

Interpreted in terms of culture, our results from the symmetric coordination game imply that when culture is a choice variable and subgroups are integrated, one culture must dominate in the sense that everyone must have a shared culture, although some individuals might learn more than one. Like Kuran and Sandholm (2008), we find that the twin social goals of social integration and multiculturalism are incompatible in that when societies are integrated one culture must dominate and the other becomes superfluous for use in interactions. If multiculturalism is valued, then the model suggests that the best way to obtain it is through social segregation.[7]

The paper can also be considered as a contribution on the coexistence of conventions. The existing literature relies on the evolutionary approach pioneered by Kandori, Mailath, and Rob (1993) and Young (1993). These papers explore 2×2 coordination games with a Pareto efficient outcome and a risk dominant strategy and find that evolutionary processes favor the risk dominant strategy. Ellison (1993) redirects attention to a matching protocol that increases the likelihood for a player to be matched with a "nearby" player and decreases the likelihood of being matched with a "distant" player. Morris (2000) generalizes the approach and uses it to model contagion. To address the issue of the coexistence of conventions, Sugden (1995) embeds the game in a location space, with different locations occurring with different probabilities and agents getting signals of their locations. Agents can then condition their behavior on their location signals, and under appropriate conditions on the location density function, different conventions can coexist in different localities. Oechssler (1997) and Ely (2002) allow for endogenous migration between locations, rather than Sugden's exogenous but imperfectly-observed assignment to locations, and find that evolution favors only the Pareto efficient convention. For their models, switching locations requires switching strategies, and Bhaskar and Vega-Redondo (2004) show that when these two switches

allowed to adjust over time, while Lazear (1999) uses a non-strategic model in which individual choices are best responses to the population choices.

[7] Lazear (1999) finds that the majority language is more likely to dominate the smaller the minority subgroup. Since our paper only looks at equilibria and not dynamics leading to them, we find conditions under which one language dominates, but not which one.

are decoupled different conventions can coexist.

In all of these evolutionary studies agents are randomly paired and paired agents are required to play the game. Agents' strategy choices are exogenous in each round of play. In contrast, in the endogenous game approach, paired agents have the option of refusing to play by paying a waiting cost and being randomly rematched the next period and agents choose to restrict their own strategy sets in a pregame in order to improve the value of the partnerships they are able to form in subsequent play. Thus, our approach is better suited to address questions of cultural integration and segregation when interactions between agents are both voluntary and endogenous. [8]

7. Conclusions

This paper addresses the question of whether two non-segregated groups with different modes of behavior or different cultural conventions can coexist while interacting only with others in their own group. The answer depends on the types of games being played. If everyone has the ability to make the same choices, that is, if there are no inherent differences between agents before the game starts, *de facto* cultural segregation is inconsistent with equilibrium. In particular, there must be one prevalent mode of behavior or cultural convention that everyone in the society shares. There may be subcultures to the extent that everyone in one group maintains both their own culture and the society-wide one, but these people are multicultural, and in equilibrium no one possesses only the non-dominant culture. On the other hand, heritage may lead to different groups of players having different different preferences over outcomes and in

[8] The evolutionary approach has some advantages over ours. In particular, our analysis is equilibrium analysis and has nothing to say about dynamics, while the evolutionary approach is dynamic and can address issues regarding which initial conditions lead to convergence to which stable equilibria. Shelling (1971) takes a different approach to the dynamics of the problem in his spatial proximity model. Two types of agents are randomly distributed in geographic space and have a preference to have neighbors who are the same type as they are. Every period agents can move to a new location that they expect to like better. He finds that neighborhoods become highly segregated by type, and Pancs and Vriend (2007) establish the robustness of his results in a variety of settings.

this setting cultural segregation can persist as long as the costs of avoiding cross-group interactions is small.

To relate these results to the real world, consider the case of Welsh speakers in Wales. The theory predicts that either everyone speaks English and some also speak Welsh, or everyone speaks Welsh and some also speak English. Nearly everyone in Wales can communicate in English. Furthermore, according to British census data, 54% of the population spoke Welsh in 1954, but by 1981 that number had fallen to 19%. This raises the issue of how a group can go about preserving its culture, given that equilibrium requires that everyone adopt the same dominant culture. Wales addressed the disappearance of its language with the Education Act of 1988, which required the teaching of Welsh to schoolchildren up to age 16. In part because of this effort, the fraction of bilinguals in the population rose to 21% in the 2001 census.

The Welsh institutional response essentially encourages the expansion of the bi-cultural group, but makes no attempt to create or preserve a mono-cultural, Welsh-only group. In fact, the theory implies that there is no way for "cultural intermediaries" who are able and willing to act with two different mono-cultural groups to preserve the multiculturalism. The reason is that members of the two mono-cultural groups are sometimes put in a position in which they can benefit from interacting with each other but prevented from doing so by cultural differences. If passing up interactions is costly, equilibrium prescribes one dominant culture, and cultural struggles are winner-take-all. The only way, then, to preserve mono-cultural societies is through institutional means that prevent opportunities for members of different cultures to meet in the first place.

Our model shows that self-segregation in a geographically-intermingled population is not enough. Rather, physical separations that keep members of different cultures from meeting are the only way to prevent such cultural struggles. Thus, separatist movements, such as those among the Basque population of Spain, are consistent with the model, while efforts to force cultural unification, such as English-only legislation in some U.S. states, are not. The model predicts that one culture will become universal throughout the population anyway, so attempts to force this are unnecessary, while efforts to keep different cultural groups separate are needed to preserve mono-cultural identities.

Appendix A - Justifying a new equilibrium concept

While Pregame Perfect Equilibrium is formally different from Sequential Equilibrium, we see PPE as simply the natural extension of subgame perfection to endogenous games. To see why we need to make this modification, consider the following simple example. Suppose there are four players, each of whom chooses a list from $\{a, b\}$. We will pay particular attention to player 1 who is going to be matched with player 2. Each individual has three possible lists: $\{a\}$, $\{b\}$, and $\{a, b\}$. For any matched pair, then, there are 9 possible pairs of lists:

$$\begin{aligned}
&\{a\}, \{a\},\\
&\{a\}, \{b\},\\
&\{a\}, \{a, b\},\\
&\{b\}, \{a\},\\
&\{b\}, \{b\},\\
&\{b\}, \{a, b\},\\
&\{a, b\}, \{a\},\\
&\{a, b\}, \{b\},\\
&\{a, b\}, \{a, b\}.
\end{aligned}$$

Thus, for each matched pair there are 9 possible nodes corresponding to the 9 possible pairs of lists.

Now suppose that player 1 is matched with player 2 and observes player 2's list. Player 1's knows her own list, and she also observes player 2's list. So, player 1 knows which node pertains from the set of 9 generated by the (player 1, player 2) pairing. However, player 1 does not know which of the 9 nodes from the (player 3, player 4) pairing is the actual decision set. Consequently, player 1's information set is not a singleton, and contains all of the possible nodes from the (player 3, player 4) pairing.

The definition of subgame perfection requires that play be a Nash equilibrium in every proper subgame, and a proper subgame is one that extends from a singleton information set. As described above, player 1's information set is not a singleton, and therefore there are no proper subgames extending from the first match. Consequently, subgame perfection has no bite. In essence, player 1 is missing the information about the lists of the players with whom he is matched with.

If player 1 and 2 agree to play, the information about the lists of players 3 and 4 becomes irrelevant. The information set is not a singleton, but every node in the information set yields the same set of branches with the same payoffs. We could just choose one of the nodes and work out the Nash equilibrium from there. But, if players 1 and 2 do not agree to play, the lists of players 3 and 4 become relevant. This is why subgame perfection will not work.

Alternative solution concepts, notably sequential equilibrium, govern behavior from non-singleton information sets. However, sequential equilibrium is constructed so that player 1 can infer things about player 2's type after observing player 2's action. This does not help in our game, because player 2's action, his choice of list, completely reveals his "type." We need an equilibrium concept that allows player 1, after observing player 2's action, to infer things about players 3 and 4. In essence, we want player 1 to observe player 2's list and then infer something about which node pertains for the (player 3, player 4) pairing so that player 1 can form expectations about what will happen if 1 and 2 decide not to play against each other.

Because neither subgame perfection nor sequential equilibrium tie down beliefs about the nodes reached by other pairings, we identify a new equilibrium concept, pregame perfect equilibrium. In essence, a combination of lists and subsequent actions is a pregame perfect equilibrium if every player believes that all other players choose lists according to the equilibrium profile, so that no player ever expects to see an out-of-equilibrium list.

To see how this works with the above example, suppose that the candidate equilibrium has all four players choosing list $\{a\}$, choosing to play against anyone with list $\{a\}$ and choosing not to play otherwise, and playing a when they play. When players 1 and 2 are matched, player 1 believes that players 3 and 4 have both chosen the list $\{a\}$, and so player 1 places zero probability on the other 8 nodes. This, in essence, makes player 1's information set a singleton and the structure of subgame perfection suffices for solving the game.

Appendix B - Formal definition of pregame perfection

In this appendix, we give formal definitions to some of the concepts used in the body of the paper. We begin by defining the various notions of expected payoffs.

The structure of the game is for agents to match randomly and decide either to play or continue searching. Given a pair of strategies, the following binary function $PL : \mathcal{S} \times \mathcal{S} \rightarrow \{0, 1\}$ indicates the outcome of any given random pairing:

$$PL(s_i, s_j) \equiv \{0 \text{ if } g_j(\ell_i) = N \text{ or } g_i(\ell_j) = N; \text{ and } 1 \text{ if } g_j(\ell_i) \neq N \text{ and } g_i(\ell_j) \neq N\}.$$

Given a particular agent i's choice of strategy s_i and a strategy profile for the remaining agents s_{-i}, we can divide the set of potential partners two sets: those who will accept a match with i and with whom i will also agree to play, and matches that will result in at least one of the two agents choosing not to play. Respectively, these may be defined formally as follows:

$$\mathcal{I}^P(s_i, s_{-i}) \equiv \{j \in \mathcal{I} \setminus i \mid PL(s_i, s_j) = 1\},$$

$$\mathcal{I}^N(s_i, s_{-i}) \equiv \{j \in \mathcal{I} \setminus i \mid PL(s_i, s_j) = 0\}.$$

Thus, given any strategy choice s_i for agent i and strategy profile s_{-i} for the remaining agents, probability that an agent does not match in any given period can be calculated as follows:

$$P^N(s_i, s_{-i}) \equiv \frac{\mid \mathcal{I}^N(s_i, s_{-i}) \mid}{(I-1)}$$

where $\mid \bullet \mid$ denotes the cardinality of a set.

Given this, the expected payoff that agent i receives under strategy profile s is:

$$EPO(s_i, s_{-i}) = \sum_{t=0}^{\infty} \left(P^N(s_i, s_{-i})\right)^t \left(\frac{1}{I-1} \sum_{j \in \mathcal{I}^P(s_i, s_j)} F(g_i(\ell_j), g_j(\ell_i)) - P_i^N(s_i, s_{-i}) z \right)$$

In order to test that agents are following optimal strategies in every subgame, we will also need to know the expected continuation payoff from every subgame. That is, conditional on having been matched with an agent who has chosen list $\hat{\ell}$, we need to know the expected payoff from this point forward.

To calculate this we need to restate the objects above to take account of the current matching realization. Thus:

$$\mathcal{I}^P(s_i, s_{-i} \mid \hat{\ell}) \equiv \{j \in \mathcal{I} \setminus i \mid PL(s_i, s_j) = 1 \text{ and } \ell_j = \hat{\ell}\},$$

$$\mathcal{I}^N(s_i, s_{-i} \mid \hat{\ell}) \equiv \{j \in \mathcal{I} \setminus i \mid PL(s_i, s_j) = 0 \text{ and } \ell_j = \hat{\ell}\}.$$

$$P^N(s_i, s_{-i} \mid \hat{\ell}) \equiv \frac{\mid \mathcal{I}^N(s_i, s_{-i} \mid \hat{\ell}) \mid}{\mid \mathcal{I}^N(s_i, s_{-i} \mid \hat{\ell}) \bigcup \mathcal{I}^P(s_i, s_{-i} \mid \hat{\ell}) \mid}$$

The *expected continuation payoff* can then be expressed as follows:

$$EPO(s_i, s_{-i} \mid \hat{\ell}) =$$

$$\frac{1}{\mid \{j \in \mathcal{I} \setminus i \mid \ell_j = \hat{\ell}\} \mid} \sum_{j \in \mathcal{I}^P(s_i, s_j \mid \hat{\ell})} F(g_i(\ell_j) g_j(\ell_i)) + P^N(s_i, s_{-i} \mid \hat{\ell}) \left(EPO(s_i, s_{-i}) - z\right)$$

Next, we define the consistent belief system for a particular agent as derived from a commonly held consistent belief system $p(g \mid S, \ell)$. Informally, the beliefs held by agent i, $p_i(g \mid S, \ell)$, regarding the probability distribution of strategies he himself will encounter is a modification of p that removes agent i from the population probabilities:

1. For all $g \in \mathcal{G}$ and $\ell \neq \ell_i$, $p_i(g \mid S, \ell) = p(g \mid S, \ell)$.
2. For all $g \in \mathcal{G}$ and $\ell \in L_i$, $p_i(g \mid S, \ell) = \frac{\mid \{j \in \mathcal{I} \setminus i \mid g_j = g \text{ and } \ell_j = \ell\} \mid}{\mid \{j \in \mathcal{I} \setminus i \mid \ell_j = \ell\} \mid}$
3. For all $g \in \mathcal{G}$ and $\ell = \bar{L}_i$, $p_i(g \mid S, \ell) = p(g \mid S, \ell)$.

We will place a bar over the next several objects we define to remind us that these refer to out of equilibrium play. Note that the following are only well defined for these out of equilibrium encounters. We begin by defining the analog of $\mathcal{I}^P$ and $\mathcal{I}^N$ for disequilibrium situations. In this case, we need to know the set of stage game strategies used by disequilibrium players that will or will not result in a match taking place.

$$\bar{\mathcal{G}}^P(s_i, s_{-i} \mid \hat{\ell}) \equiv \{g \in \mathcal{G} \mid PL(s_i, g, \hat{\ell}) = 1\},$$

$$\bar{\mathcal{G}}^N(s_i, s_{-i} \mid \hat{\ell}) \equiv \{g \in \mathcal{G} \mid PL(s_i, g, \hat{\ell}) = 0\}.$$

Given this and a set of consistent beliefs p, the probability of no match taking place in a given round is easily calculated:

$$\bar{P}^N(s_i, s_{-i} \mid \hat{\ell}) \equiv \sum_{g \in \bar{\mathcal{G}}^N} p_i(g \mid S, \ell)$$

Using this, we can finally state expected payoff in the event of an out of equilibrium encounter as follows:

$$\overline{EPO}(s_i, s_{-i} \mid \hat{\ell}) =$$

$$\left(\sum_{g \in \bar{\mathcal{G}}^P(s_i, s_j \mid \hat{\ell})} p_i(g \mid S, \hat{\ell}) F(g_i(\hat{\ell}), g(\ell_i)) + P^N(s_i, s_{-i} \mid \hat{\ell}) \left(EPO(s_i, s_{-i}) - z \right) \right).$$

Appendix B - Proofs

Proposition 1. *There is no pregame perfect equilibrium with $0 < I_A < I$ A-type agents who do not play against B-type agents, $I - I_A$ B-type agents who do not play against A-type agents, and no AB-types at all.*

Proof/

Let I_A, I_B, and I_{AB} denote the numbers of A-types, B-types, and AB-types, respectively. We want to establish the existence or nonexistence of a pregame perfect equilibrium in which $I_A > 0$, $I_B > 0$, and $I_{AB} = 0$. Note that in any PPE, agents will always agree to play with their own types since this give the highest payoff possible. Thus, the expected payoff to A-types in this case is:

$$EPO_A(S) = \frac{I_A - 1}{I - 1}\alpha + \frac{I_B}{I - 1}\left(-z + EPO_A(S)\right) = \alpha - \frac{I_B}{I_A - 1}z$$

and similarly the expected payoff to a B-type is

$$EPO_B(S) = \beta - \frac{I_A}{I_B - 1}z.$$

The key point to note is that $EPO_i(S)$ increases in I_i and decreases in I_j for $i, j \in \{A, B\}$ with $i \neq j$. Thus, if the payoff from being an B-type, for example, is higher than the payoff from being an A-type (either because β is higher or $I_B > I_A$), then A-types are better off becoming B-types. In addition, by making the switch, they further increase the payoff advantage of being an B-type. Even if the payoffs to each type are initially equal, an A-type increases his payoff by becoming a B-type. Consequently, for any possible combination of $I_A > 0$ and $I_B > 0$ with $I_A + I_B = I$ some type has an incentive to choose the other type's list, and so there is no PPE that satisfies the hypothesis.

■

Proposition 2. *There is no pregame perfect equilibrium with $I_A > 0$ A-type agents who do not play against B-type agents and $I_B > 0$ B-type agents who do not play against A-type agents.*

Proof/

Suppose there is such an equilibrium. The expected payoff to an A-type is

$$EPO_A(S) = \frac{I_A + I_{AB} - 1}{I - 1}\alpha + \frac{I_B}{I - 1}(-z + EPO_A(S)) = \alpha - \frac{I_B}{I_A + I_{AB} - 1}z$$

and similarly the expected payoff to a B-type is

$$EPO_B(S) = \beta - \frac{I_A}{I_B + I_{AB} - 1}z.$$

We can compute

$$EPO_A(S) - EPO_B(S) = \alpha - \beta + \frac{I_A - I_B}{(I_A + 1)(I_B + 1)}(I - 1)z,$$

which again is increasing in I_A and decreasing in I_B. The argument is therefore the same as for Proposition 1. In all cases either an A-type, a B-type, or both have an incentive to defect, contradicting the hypothesis that an equilibrium with both types exists.

■

Proposition 3. *There exist exactly two classes of pregame perfect equilibria: equilibria with no A-types and equilibria with no B-types.*

Proof/

First consider an economy in which I_A agents choose list $\{A\}$, $I - I_A$ agents choose list $\{A, B\}$. Two cases arise based on the value of the delay cost z.

Case 1: $z < \alpha$. Suppose that A-types choose the following strategy in the continuation game:

$$g_A(\{A\}) = A \qquad g_A(\{B\}) = N \qquad g_A(\{AB\}) = A$$

and AB-type agents choose the strategy given by:

$$g_{AB}(\{A\}) = A \qquad g_{AB}(\{B\}) = B \qquad g_{AB}(\{AB\}) = A.$$

In addition, suppose that both types of agents think that if they happen to encounter a B-type agent, his continuation strategy is:

$$\bar{g}_B(\{A\}) = N \qquad \bar{g}_B(\{B\}) = B \qquad \bar{g}_B(\{AB\}) = B.$$

We will leave aside for the moment the question of whether this generates a consistent belief system.

Observe that given these strategies, all agents match in the first round and get the maximum payoff. Thus, no strategy could yield a higher payoff from the standpoint of the pregame and so condition (1) of PPE is satisfied. This also means that in any matching that occurs after the pregame, the strategies for each type of player still yield the maximum payoff. Thus, in equilibrium encounters, the conditional expected payoff is maximized and condition (2) of the PPE is satisfied. Now suppose that an A or AB type happens to be matched with a B type. This is an out of equilibrium encounter. For AB-types, this presents no problem as both agents play "B" and get maximum payoffs. For A-types, the expected payoff to this strategy is to suffer a cost of $z < \alpha$ this period and then get the maximal payoff of 1 with certainty next period. Since $z < \alpha$, condition (3) of PPE is satisfied. Finally, from the above arguments, it is clear that when $z < \alpha$, it is optimal for A-types to pass in this case, and so condition (4) is satisfied.

To complete the proof for this case, we have to show that this belief system is consistent. Parts (1) and (2) are trivially satisfied, so only part (3) is in question. In particular, we have to check that the strategy ascribed to the (non-existent) B-types is, in fact, an optimal strategy. Clearly, B-types should play B against other B-types and against AB-types. The question is playing N against B-types credible? Since A-types play N when meeting B-types, however, it does not matter what B-types do. Thus, the no other strategy would yield a higher payoff, and so part (3) of the definition of a consistent belief system is satisfied as well.

Case 2: $z \geq \alpha$. Under this condition an A-type's best response when facing a B-type is to play, since the penalty for delay is too steep. Accordingly, suppose that A-types choose the following strategy in the continuation game:

$$g_A(\{A\}) = A \qquad g_A(\{B\}) = A \qquad g_A(\{AB\}) = A$$

and AB-type agents choose the strategy given by:

$$g_{AB}(\{A\}) = A \qquad g_{AB}(\{B\}) = B \qquad g_{AB}(\{AB\}) = A.$$

In addition, suppose that both types of agents think that if they happen to encounter a B-type agent, his continuation strategy is:

$$\bar{g}_B(\{A\}) = B \qquad \bar{g}_B(\{B\}) = B \qquad \bar{g}_B(\{AB\}) = B.$$

As with Case 1, this generates a consistent belief system, and the continuation game strategies are consistent with PPE.

It remains to show for this case that no agent wants to change his list in the pregame. An A-type has no incentive to change to a B-type or an AB-type because he receives the maximal payoff in the first period. For the same reason, an AB-type has no incentive to change to being an A-type or a B-type. So, even though the high delay cost makes it optimal for an A-type to play against a B-type, there is no incentive to become a B-type.

We conclude that a PPE exists with no B-types. A similar argument establishes the existence of a PPE with no A-types for $z \geq \beta$ and $z \leq \beta$. Proposition 2 then implies that the *only* pregame PPE have either no B-types or no A-types when AB-types can solve the coordination problem.

■

Proposition 4. *If z is small enough, there exists a pregame perfect equilibrium in which agents' lists do not contain the non-heritage action and agents play only against others from the same heritage.*

Proof/

We suppose that a-affinity agents learn list $\{A\}$, b-affinity agents learn list $\{B\}$, and these agents play the following strategies for $h = a, b$:

$$g^a(h, \{A\}) = A \qquad g^a(h, \{B\}) = N \qquad g^a(h, \{AB\}) = A$$
$$g^b(h, \{A\}) = N \qquad g^b(h, \{B\}) = B \qquad g^b(h, \{AB\}) = B.$$

We must also state a belief system for heritage/list combinations that do not appear in equilibrium. Here the "bar" reminds us that these are not equilibrium agents and the notation $\bar{g}^{h'}_{\ell'}(h, \ell)$ denotes the continuation game strategy of an h'-affinity player with list ℓ' matched against an h-affinity player with list ℓ. Thus, for $h = a, b$:

$$\bar{g}^a_B(h, \{A\}) = N \qquad \bar{g}^a_B(h, \{B\}) = B \qquad \bar{g}^a_B(h, \{AB\}) = B$$
$$\bar{g}^a_{AB}(h, \{A\}) = A \qquad \bar{g}^a_{AB}(h, \{B\}) = N \qquad \bar{g}^a_{AB}(a, \{AB\}) = A \qquad \bar{g}^a_{AB}(b, \{AB\}) = N$$
$$\bar{g}^b_A(h, \{A\}) = A \qquad \bar{g}^b_A(h, \{B\}) = N \qquad \bar{g}^b_A(h, \{AB\}) = A$$
$$\bar{g}^b_{AB}(h, \{A\}) = N \qquad \bar{g}^b_{AB}(h, \{B\}) = B \qquad \bar{g}^b_{AB}(a, \{AB\}) = N. \qquad \bar{g}^b_{AB}(b, \{AB\}) = B.$$

We begin by showing the definition of PPE is satisfied for a-natives. From the above, we can calculate an a-native's expected equilibrium payoff as:

$$EPO^a(S) = \frac{H^a - 1}{H - 1}(\alpha + \eta) + \frac{H^b}{H - 1}(-z + EPO^a(S)) = (\alpha + \eta) - \frac{H^b}{H^a - 1} z$$

There are a variety of alternative strategies that an a-native might employ. In principle, showing that the strategy we suggest produces the highest expected payoff starting from the pregame requires that we test them all. Fortunately, most of these are obviously suboptimal, so we will begin by focusing on the one that is non-trivial and potentially optimal. Suppose that an a-native considers an alternative strategy s' in which learns B as well as A, and plays B when encountering a b-native. In this case, his expected payoff given the new strategy profile S' is:

$$EPO^a(S') = \frac{H^a - 1}{H - 1}(\alpha + \eta) + \frac{H^b}{H - 1}\alpha = \frac{H^a - 1}{H - 1}(\alpha + \eta) + \frac{H - H^a}{H - 1}\alpha =$$

$$\frac{\alpha H^a + \eta H^a - \alpha - \eta + \alpha H - \alpha H^a}{H - 1} = \frac{(H - 1)\alpha + (H^a - 1)\eta}{H - 1}.$$

For this to be a best response for the deviating a-native, it must be the case that:

$$\frac{(H - 1)\alpha + (H^a - 1)\eta}{H - 1} \geq (\alpha + \eta) - \frac{H^b}{H^a - 1}z$$

Clearly, for small enough z, this will not be true. Thus, the strategy s' is not superior to s for this a-native.

The other alternative strategies are:

1. Learn only B and play this against all partners: clearly this is dominated by learning A and B and playing A if the agent happens to be partnered with an a-native
2. Learn A or A and B but play only against b-natives: obviously, this is a dominated strategy.
3. Learn anything but always pass: obviously dominated.
4. Learn A and B but play only against a-natives: this has the same payoff as s and results in the same play.

We conclude that s maximizes the expected payoff a-natives when the delay cost is small and so condition (1) of the definition of PPE is satisfied.

Now we need to show that these strategies continue to be optimal in any equilibrium subgame. Suppose an a-native meets an a-native. Clearly playing A maximizes the expected payoff. Suppose instead he meets a b-native. Since b-natives don't play in this case, no stage game strategy can give anything more than $-z$ plus the continuation payoff. It follows that Not playing is expected payoff maximizing. Thus, condition (2) of the definition of PPE is satisfied.

Next, consider out of equilibrium encounters.

1. If an a-native meets an a-native who only can play B, Not playing is optimal for small z.
2. If an a-native meets an a-native who only can play A and B, playing is A is optimal.
3. If an a-native meets a b-native who only can play A, playing A is optimal.

4. If an a-native meets a b-native who can play both A and B, not playing is optimal for small z.

Thus, condition (3) of the definition of PPE is satisfied. Finally, the only time a-natives choose to Not play is when they encounter a b-native who can only play B. We show above this is optimal for small z. Thus, condition (4) of the definition of PPE is satisfied.

Obviously, since the strategies are symmetrically inverted for b-natives, the arguments above also show that the strategy S with this belief system also satisfies the definition of PPE for b-natives as well.

It remains only to show that the strategies assigned to players not seen in equilibrium form a consistent belief system (that is, condition (3) of the definition of consistent beliefs if satisfied). Observe that the strategies, $\bar{g}$, given for these players fall into two categories: (1) For agents who have not learned their native strategies (a-natives who only know B. for example) $\bar{g}$ requires them to play the non-native strategy when matched with someone who also has that strategy in their list, but not against someone without that strategy in their list. They play in order to earn a payoff of α, and for sufficiently small z waiting for the payoff of α is optimal. (2) For agents who know both strategies, $\bar{g}$ requires them to play their native strategies against any equilibrium player who will also play this strategy and to Not play otherwise. Again, for small z, this is expected payoff maximizing. We conclude that this PPE is supported by a consistent belief system.

■

References

Benabou, Roland (1993): "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics,* **108** pp. 619-652.

Bhaskar, V. and Fernando Vega-Redondo (2004): "Migration and the Evolution of Conventions," *Journal of Economic Behavior and Organization,* **55** pp. 397-418.

Card, David and Jesse Rothstein (2007): "Racial Segregation and the Black-White Test Score Gap," *Journal of Public Economics,* **91** pp. 2158-2184.

Case, Anne C. and Lawrence F. Katz (1991): "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," NBER Working Paper 3705.

Conley, John P. and William S. Neilson (2008): "Endogenous Games and Equilibrium Adoption of Social Norms and Ethical Constraints," mimeo, Vanderbilt University.

Conley, Timothy G. and Giorgio Topa (2002): "Socio-economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics,* **17** pp. 303-327.

de Bartolome, Charles A. M. (1990): "Equilibrium and Inefficiency in a Community Model with Peer Group Effects," *Journal of Political Economy,* **98** pp. 110-133.

Ding, Weili and Steven F. Lehrer (2007): "Do Peers Affect Student Achievement in China's Secondary Schools?," *Review of Economics and Statistics,* **89** pp. 300-312.

Durlauf, Steven N. (1996): "A Theory of Persistent Income Inequality," *Journal of Economic Growth,* **1** pp. 75-93.

Ellison, Glenn (1993): "Learning, Local Interaction, and Coordination," *Econometrica,* **61** pp.1047-1071.

Ely, Jeffrey C. (2002): "Local Conventions," *Advances in Theoretical Economics,* Vol. 2, No. 1, Article 1.

Friesen, Jane and Brian Krauth (2007): "Sorting and Inequality in Canadian Schools," *Journal of Public Economics,* **91** pp. 2185-2212.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin (2003): "Does Peer Ability Affect Student Achievement?," *Journal of Applied Econometrics,* **18** pp. 527-544.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin (2002): "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement," NBER Working Paper 8741.

Hoxby, Caroline M. (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working Paper 7867.

Kandori, Michihiro, George J. Mailath, and Rafael Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica,* **61** pp. 29-56.

Kuran, Timur and William H. Sandholm (2008): "Cultural Integration and Its Discontents," *Review of Economic Studies,* **75** pp. 201-228.

Lazear, Edward P. (1999): "Culture and Language," *Journal of Political Economy,* **107** pp. S95-S126.

Marmaros, David and Bruce Sacerdote (2006): "How Do Friendships Form?," *Quarterly Journal of Economics,* **121** pp. 79-119.

Mayer, Adalbert and Steven L. Puller (2008): "The Old Boy (and Girl) Network: Social Network Formation on University Campuses," *Journal of Public Economics,* **92** pp. 329-347.

Morris, Stephen (2000): "Contagion," *Review of Economic Studies,* **67** pp. 57-78.

Oechssler, Jorg (1997): "Decentralization and the Coordination Problem," *Journal of Economic Behavior and Organization,* **32** pp. 119-135.

Pancs, Romans and Nicolaas J. Vriend (2007): "Schelling's Spatial Proximity Model of Segregation Revisited," *Journal of Public Economics,* **91** pp. 1-24.

Rubinstein, Ariel and Asher Wolinsky (1985): "Equilibrium in a Market with Sequential Bargaining," *Econometrica,* **53** pp. 1133-1150.

Schelling, Thomas C. (1971): "Dynamic Models of Segregation," *Journal of Mathematical Sociology,* **1** pp. 143-186.

Sethi, Rajiv and Rohini Somanathan (2004): "Inequality and Segregation," *Journal of Political Economy,* **112** pp. 1296-1321.

Sugden, Robert (2004): *The Economics of Rights, Co-operation and Welfare.* Palgrave Macmillan.

Topa, Giorgio (2001): "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies,* **68** pp. 261-295.

Young, H. Peyton (1993): "The Evolution of Conventions," *Econometrica,* **81** pp. 57-84.