Coordination and Culture

Jean-Paul Carvalho*

Department of Economics
University of California, Irvine

Abstract

Culture constrains individual choice, rendering certain actions impermissible or taboo. While cultural constraints may regulate behavior within a group, they can have a pernicious effect in multicultural societies, inhibiting the emergence of unified social conventions. We analyze interactions between members of two cultural groups who are matched to play a coordination game with an arbitrary number of actions. Due to cultural constraints, miscoordination prevails despite strong incentives to coordinate behavior. In an application to identity-based conflict, exclusive ethnic and religious identities persist in poorer and more unequal societies. Occasional violation of cultural constraints can make miscoordination even more stable.

This version: 19 February 2016

JEL Classification: C72, C73, Z1

Keywords: Culture, Conflict, Coordination Failure, Stochastic Stability

*This paper has benefitted greatly from advice by Peyton Young, as well as comments by the Associate Editor, two anonymous referees, Blake Allison, Ken Binmore, Rob Boyd, Michael Caldara, David Myatt, Tom Norman, Michael Sacks, Stergios Skaperdas, Christopher Wallace and seminar participants at UCLA Anderson, Claremont Graduate University, the Institute for Mathematical Behavioral Sciences, UC Irvine and the University of Western Australia. All errors are mine. Financial support from the Commonwealth Bank Foundation in the form of a John Monash scholarship is gratefully acknowledged.

1 Introduction

Culture constrains individual choice. In matters of diet, dress, language, manners and much else, one's culture renders certain choices impermissible or taboo.[1] There are many well known taboos. Vegetarians do not consider switching to meat as circumstances dictate. It remains taboo for the Amish to use modern technology, even as it becomes more attractive. Christians in the Middle Ages observed usury prohibitions. Moreover, there are many less obvious ways in which culture can restrict an individual's choice set by making certain types of behavior unthinkable, unpalatable or prohibitively costly to learn. For example, suppose that members of two different religious groups can form new business contacts by interacting socially. Despite gains from social coordination, no individual would consider meeting at the other group's place of worship. To interact, they must coordinate on a neutral venue.

This paper proposes a particularly simple notion of culture. Let X be the global choice set. A member of cultural group k chooses from the set $X_k \subseteq X$. This is what Sen (1977) refers to as "commitments", Harsanyi (1982) as "moral values" and Rao & Walton (2004) as "constraining preferences".[2] It then explores the consequences of this notion of culture. While cultural constraints can help to coordinate and regulate behavior within a group, it is possible that they have a pernicious effect in multicultural societies, inhibiting the emergence of unified social conventions. This could occur, we suggest, even when coordination is attainable, that is when $X_A \cap X_B$ is nonempty for some groups A and B.

Consider a set of boundedly rational agents partitioned into two cultural groups, each with its own cultural constraint. Individuals are matched recurrently with members of the *other* group to play a coordination game with an arbitrary number of actions. They adapt their choices over time and occasionally make mistakes, à la Kandori, Mailath & Rob (1993) and Young (1993*a*). We are interested in whether they can learn to coordinate with each other within cultural constraints. The results are rather pessimistic. Certain forms of cultural

[1] For example, Tetlock et al. (2000) find that experimental subjects express moral outrage at even contemplating certain taboo transactions, including buying and selling of human body parts for medical transplant operations, surrogate motherhood contracts, adoption rights for orphans, votes in elections for political office, the right to become a U.S. citizen and sexual favors. See Roth (2007) for an account of how repugnance constrains markets.

[2] See also Henrich et al. (2001) who suggest that culture "limits choice sets" [p. 357]. Conventional definitions in economics equate culture with shared preferences (e.g. Bisin & Verdier 2000) and strategic beliefs (e.g. Greif 1993).

diversity—combinations of cultural constraints and preferences—can lead to the breakdown of coordination across groups, even where coordination is attainable and Pareto dominates miscoordination. Nevertheless, there are multiple equilibria and the emergence of a unified social convention across groups is always possible in the long run (as long as $X_A \cap X_B$ is nonempty).

To select among equilibria, we study the evolution of play when the error rate is positive but vanishingly small (Foster & Young 1990). Since every state of coordination Pareto dominates every state of miscoordination, one might expect miscoordination to be a tenuous phenomenon. On the contrary, miscoordination is stochastically stable for an open set of parameters.[3] In particular, we employ a graph-theoretic argument to show that if miscoordination pairwise (strictly) risk dominates all other equilibria, then it is the unique stochastically stable class. Because we study large coordination games, this does not follow from existing results, nor is the proof trivial. In addition, we derive a condition under which pairwise risk dominance is necessary and sufficient for miscoordination to be stochastically stable.

The analysis is applied to an example of identity-based conflict in which it is Pareto efficient for members of two ethnic groups to coordinate on an inclusive (e.g. national) identity. Sen (2006) argues that the roots of identity-based violence lie in the emphasis on an exclusive sense of self.[4] He claims that through reason one can choose to arrive at a more inclusive social identity. As we show, however, even when strong incentives to coordinate on an inclusive identity exist, exclusive (e.g. ethnic) identities may persist due to evolutionary pressures. In particular, we demonstrate that exclusive ethnic and religious identities are more likely to persist in poorer and more unequal societies.

A question remains as to whether occasional violation of cultural constraints could destabilize

[3] This phenomenon is qualitatively different to the selection of a Pareto-inefficient coordination equilibrium when agents use the same action set (Kandori, Mailath & Rob 1993, Young 1993*a*). Here, play might not settle into any coordination equilibrium, even though every coordination equilibrium Pareto dominates miscoordination.

[4] Important contributions to understanding identity-based conflict and cooperation are also made by Fearon & Laitin (1996), Laitin (2007), Esteban & Ray (2008), McBride et al. (2011), Greif & Tabellini (2012), Jha (2013) and Sambanis & Shayo (2013). Sambanis and Shayo present a formal theory in which individuals either identify with their ethnic group or the nation. Under certain conditions, multiple equilibria exist and cascades of ethnic identification can occur (see also Kuran 1998). Such miscoordination between cultural groups, as we show, is a more general phenomenon, not limited to ethnic conflict and independent of specific feedback mechanisms. By analyzing an explicit out-of-equilibrium process, we are also able to select among different equilibria and show that miscoordination is a surprisingly stable outcome.

miscoordination. Since miscoordination is supported as an equilibrium only when cultural constraints differ in a particular way, one might expect violation of cultural constraints to weaken the stability of miscoordination or at least limit what can be said. We find otherwise. Not only are the results sharper, but moreover small-scale violations can strengthen miscoordination, expanding the set of parameters for which it is stochastically stable. Specifically, we derive a necessary and sufficient condition for miscoordination to be stochastically stable, which is weaker than risk dominance.

This paper is related to several different lines of work. Firstly, our results contribute to the study of stochastic stability in general asymmetric coordination games. Early work analyzed stochastic stability in symmetric 2×2 coordination games (e.g. Young 1993*a*, 1996, Ellison 1993),[5] showing that risk dominance is necessary and sufficient for stochastic stability. Prior work on asymmetric coordination games has largely focussed on deterministic dynamics (Samuelson & Zhang 1992) and two-action games (Staudigl 2012).

Young (1993*a*, p. 73) shows by example that beyond 2×2 coordination games, risk dominance need not be sufficient for stochastic stability. Indeed risk dominance is a *pairwise* concept, determining the likelihood of *direct* transitions between equilibria. Hence the logic from 2×2 coordination games does not naturally extend to larger games. However, Young's example does not match our payoff structure. Kandori & Rob (1993) show that, under certain conditions, an equilibrium is stochastically stable if it pairwise risk dominates every other. Their result relies upon a 'total bandwagon property', the violation of which is critical to our analysis (as demonstrated later). Hence our results do not follow from prior work. Ellison (2000) shows, for symmetric $L \times L$ coordination games, that when a half-dominant equilibrium exists, it is stochastically stable. Our sufficient condition for stochastic stability is substantially weaker than half dominance. We also identify a condition under which risk dominance is necessary and sufficient for stochastic stability. In addition, when individuals occasionally violate cultural constraints, we derive a necessary and sufficient condition that is weaker than risk dominance.

Miscoordination has been studied from a different perspective in 2×2 coordination games. Myatt & Wallace (2004) analyze 2×2 coordination games in which agents' payoffs are determined by random draws from a normal distribution. Miscoordination occurs in certain

[5] A notable exception is Young (1993*b*) who examines bargaining between members of heterogeneous groups.

matches. As payoff trembles vanish, however, a single pure equilibrium is selected and coordination is almost always achieved. Quilter (2007) and Neary (2012) study 2×2 coordination games with nonvanishing differences in payoffs. Players belong to one of two groups and play a single action against all members of the population. When group A members prefer to coordinate on action 1 and group B members on action 2, there is a tradeoff between ingroup coordination on one's preferred action and coordination with outgroup members. As a result, multiple conventions can coexist.[6] Because our focus is on between-group interactions, the type of miscoordination that we uncover can occur where there is no tradeoff between ingroup and outgroup coordination. As such, our results also apply when individuals are *flexible*, viz. able to choose one action when interacting with ingroup members and another action with outgroup members.[7]

Among other related work, Kuran & Sandholm (2008) study culture as an evolving distribution of preferences and behaviors, not a cultural constraint on behavior. There is no breakdown of Pareto-efficient coordination in their work and cultural convergence eventually occurs among different groups. Bénabou & Tirole (2011) and Fershtman et al. (2011) analyze the conditions under which taboos emerge. Bénabou and Tirole present a theory of self-signaling in which taboos protect one's sense of self or identity. Fershtman et al. propose that taboos provide public benefits and examine the conditions under which they can unravel through a series of deviations. Our focus is on whether individuals from different cultural groups can learn to coordinate within a fixed set of cultural constraints or taboos.

The paper is structured in the following manner. The model is introduced and analyzed in section 2. Section 3 studies an application to identity choice and an extension to small-scale violations of cultural constraints. Section 4 concludes. All proofs and technical lemmas are located in the Appendix.

2 The Model

In this section, we analyze a two-population model of social coordination. Interactions occur between groups and individuals are bound by cultural constraints.

[6] In addition, when agents are differentially located in space, different groups can settle on different conventions (e.g. Young & Burke 2001).

[7] Members of ethnic or religious minorities often adopt different modes of behavior when interacting outside of their communities (see Akerlof & Kranton 2000, p. 738-9).

2.1 The Underlying Game

Consider a society composed of a finite set of roles N, with typical members i and j. A (possibly) changing cast of players fill these roles.[8] For convenience, we refer to elements of N as players (rather than roles).

Each player belongs to one of two equally sized cultural groups $k \in \{A, B\}$. The set of players comprising group k is denoted by N_k, where $|N_k| = n$ is finite. Each group has a different culture, which proscribes certain kinds of behavior.

Individuals are matched with a player drawn uniformly at random from the *other* group to play a coordination game.[9] The finite set X is the set of all actions that can potentially be chosen in the game. For example, X can represent a menu of diets, languages or social identities.

In addition to coordination, individuals have two considerations when choosing an action. The first consideration is respect for cultural constraints. This is the key assumption:

Condition 1. *(Cultural Constraints)* A group k member chooses from the set of actions that are culturally permissible, $X_k \subseteq X$.

Such a restriction on an agent's strategy set might come about through the internalization of culturally accepted standards of behavior during socialization or as a result of sanctions imposed by the group, though we do not model this process here (see Frank 1988, Elster 1989, Akerlof & Kranton 2000). For example, a Turkish woman raised in a secular household may not consider wearing a headscarf, whereas a Turkish woman raised in a religious household may not consider going out in public without one.[10] In section 3.2, we consider a setting in which all actions are available to all players but each action in $X \backslash X_k$ is strictly dominated for k members by some action in X_k.

Secondly, they have culturally defined preferences over actions that are independent of their partner's action. The payoff to a group k member who is matched with player j and plays

[8] Over time, we can think of players "dying" and their roles being filled by incoming players who inherit their predecessor's culture. One such example is the vertical transmission of culture from parent to child.

[9] The results in this section hold if we assume that in every period each agent is matched with every agent from the other group, with a set R_t of individuals selected to revise their strategies as above. The application analyzed in section 3.1, however, is developed for pairwise interactions.

[10] For theories of veiling see Patel (2012) and Carvalho (2013).

x is given by

$$u_k(x, x_j) = \mathbb{I}(x, x_j)\theta_k + \delta_{kx}. \tag{1}$$

The first term on the right-hand side (RHS) of (1) is defined as follows: $\mathbb{I}(x, x_j) = 1$ if $x = x_j$ and zero otherwise. Hence $\theta_k > 0$ is the payoff to a group k member from coordination. The second term, δ_{kx}, represents the cultural payoff from choosing action x for a group k member. Myatt & Wallace (2004) were the first to introduce such payoff disturbances in the context of coordination games. We refer to an action $x \in argmax_{x \in X_k} \delta_{kx}$ as group k's *cultural ideal.* Let $d_k \equiv max_{x \in X_k \setminus X_{k'}} \delta_{kx}$, $k \neq k'$, denote the maximal payoff to a group k member from choosing a mutually impermissible action. We shall refer to $d_k - \delta_{kx}$ as group k's *cultural bias* against x.

We impose the following non-degeneracy condition on payoffs:

Condition 2. *(ND)* $\delta_{kx'} - \delta_{kx} < \theta_k$ for all $(x, x') \in X^2$ and $k = A, B$.

In this paper, we are interested in social coordination. When ND is satisfied, coordination on any mutually permissible action Pareto dominates miscoordination despite conflicting cultural ideals. When ND is not satisfied, some mutually permissible actions are weakly dominated.

2.2 Adaptive Choice and Population Dynamics

Interactions take place in discrete time, indexed by $t = 0, 1, 2, \ldots$ Let $z_k^t \equiv \left(x_i^t\right)_{i \in N_k}$ be the group-k action profile at the beginning of time t. The state in period t is defined by the action profiles of the two groups, $z^t \equiv \left(z_A^t, z_B^t\right)$. The associated (finite) state space is $Z = X_A^n \times X_B^n$. The process begins in an arbitrary state $z^0 \in Z$.

In each period $t \geq 1$, a set of players R_t is selected at random. Each player $i \in R_t$ gets the opportunity to revise his strategy in period t. Let $f(R)$ denote the probability that set R is selected in a given period. We assume that $f(R) > 0$ for all $R \in \mathcal{P}(N)$, the power set of N. (Hence, among other things, simultaneous revisions occur with positive probability.)

Each revising k member at time t plays a myopic and noisy best response to $z_{k'}^t$. Let $p_{k'x}$ be the proportion of group k' members playing x. With high probability $1 - \varepsilon$, a revising player chooses a constrained best response by maximizing $p_{k'x}\theta_k + \delta_{kx}$, subject to $x \in X_k$. (When there are ties, each best response is played with equal probability.) With low probability ε, a revising group k member instead selects an action from X_k at random. Nonbest responses represent mistakes and/or random experimentation.

The best response protocol with uniform errors (see Kandori, Mailath and Rob 1993, Young 1993a) implies myopia on the part of agents in that choices are based solely on the current state, a standard assumption in evolutionary game theory (see Young 1998, Sandholm 2010). Computing the effect of current choices on the future trajectory of play is a complex and costly exercise. Myopic behavior is thus natural in large populations of boundedly rational individuals.

It is well known that this revision protocol produces a particular kind of evolutionary dynamic at the population level. There exist time-homogeneous transition probabilities between all pairs of states z, z', denoted by $P_{zz'}$. They define a finite Markov chain with a $|Z| \times |Z|$ transition probability matrix which depends on the noise level ε. For $\varepsilon > 0$, all pairs of states communicate, so the Markov chain is irreducible and thus ergodic—it has a unique stationary distribution μ^ε that is independent of the initial state z^0.[11] The process is also aperiodic when $\varepsilon > 0$, since there is a positive probability of remaining in any given state. As such, the stationary distribution tells us a lot about the asymptotic behavior of the process. Not only does the relative frequency of visits to state z *up through* time t converge to the frequency given by the unique stationary distribution μ^ε, but so does the probability of being in state z *at* time t.

Typically, we can greatly reduce the set of states in the support of μ^ε by taking the noise level to zero. We rely on the following equilibrium concept due to Foster & Young (1990):

Stochastic Stability. A state is *stochastically stable* if it is in the support of $\mu = lim_{\varepsilon \to 0}\mu^\varepsilon$.

A stochastically stable class is a recurrent class composed of stochastically stable states.[12] As the noise level becomes arbitrarily small, the process spends virtually all of the time as $t \to \infty$ in the stochastically stable classes.

[11] Recall that state z' is accessible from z if there exists a positive probability path from z to z'. The states communicate if they are each accessible from the other.

[12] A set of states E is closed if for all $x \in E$ and $y \notin E$, $P_{xy} = 0$. A recurrent class is a closed communication class.

2.3 Coordination and Miscoordination

Let us begin by studying the asymptotic behavior of the best response dynamic with $\varepsilon = 0$, before proceeding to the stochastic stability analysis. We would like to know if members of different cultural groups can learn to coordinate with each other. There are two types of recurrent states (i.e. members of a recurrent class):

Definition 1. *(Coordination)* A state of coordination is one in which $x_i = x$ for all $i \in N$.

Definition 2. *(Miscoordination)* A state of miscoordination is one in which $x_i \neq x_j$ whenever $i \in N_k$ and $j \in N_{k'}$, $k \neq k'$.

In a moment, we will show that social coordination can emerge even when payoffs differ across groups. Differing cultural constraints, however, create a new possibility in which social coordination permanently breaks down. Let us define the following relation:

Definition 3. *(Compatibility)* Cultures A and B are *compatible* if $X_A \cap X_B$ is nonempty. They are incompatible otherwise.

Clearly coordination between incompatible cultures is impossible. It turns out, however, that miscoordination can arise even among compatible cultures, when cultural ideals differ in a particular way. Let $\widetilde{X}_k \equiv argmax_{x \in X_k} \delta_{kx}$ be the set of cultural ideals for group k.

Definition 4. *(Misalignment)* Cultures A and B are *misaligned* if $\widetilde{X}_A \cap X_B = \emptyset$ and $\widetilde{X}_B \cap X_A = \emptyset$. They are aligned otherwise.

Two cultures are misaligned when every ideal action for one group is impermissible for the other. If the cultures are incompatible, then they are also misaligned, since $\widetilde{X}_A \subseteq X_A$ and $\widetilde{X}_B \subseteq X_B$. The converse is not true; unlike compatibility, alignment depends on the difference between the cultural ideals of each group. Compatible cultures are misaligned if and only if they both have a positive cultural bias against all mutually permissible actions, i.e. $d_k - \delta_{kx} > 0$ for all $x \in X_A \cap X_B$ and $k = A, B$.

Define $\mathcal{M} = \widetilde{X}_A^n \times \widetilde{X}_B^n$ as the set of states in which everyone plays a culturally ideal action. When the cultures are misaligned each $z \in \mathcal{M}$ is a state of miscoordination. In this case, we simply refer to $\mathcal{M}$ as miscoordination. We can state the following result:

Proposition 1 *The best response dynamic ($\varepsilon = 0$) converges almost surely to:*

(i) *a state of coordination if the cultures are aligned,*

(ii) *either a state of coordination or miscoordination $\mathcal{M}$, depending on the initial state z^0, if the cultures are compatible but misaligned,*

(iii) *miscoordination $\mathcal{M}$ if the cultures are incompatible.*

To prove Proposition 1, we identify the conditions under which the states of coordination and $\mathcal{M}$ are recurrent classes of the best response dynamic operating on this game and hence correspond to its static Nash equilibria. In addition, we show that these are the only recurrent classes. As any finite Markov chain converges almost surely to one of its recurrent classes, this suffices to establish the proposition. All proofs are in the Appendix.

When cultures are aligned, individuals inevitably learn to coordinate with each other in cross-cultural interactions. When cultures are incompatible, they cannot. The interesting case is when cultures are compatible but misaligned. In this case, coordination between different cultural groups can *permanently* break down even though coordination Pareto dominates miscoordination. Whether players can learn to coordinate depends on initial conditions. If play begins with a sufficiently large proportion of individuals choosing a mutually permissible action, $x \in X_A \cap X_B$, then the prospect of coordination will induce revising agents to depart from their cultural ideals. Otherwise, miscoordination can arise.[13] Suppose that the process begins in a state of miscoordination. A revising player believes he has no chance of coordinating with members of the other group, so he chooses one of his ideal actions, say x. If the cultures are misaligned then $x \notin X_A \cap X_B$, so the dynamic remains in a state of miscoordination. If however the cultures are aligned, then revising players from at least one group will respond with a mutually permissible action. Thus the dynamic can exit from a state of miscoordination and social coordination may be achieved.

It is cultural diversity that hinders social coordination, not any other aspect of our setup.[14]

[13]This does not mean that the state space can be partitioned into a basin of attraction for each state of coordination and a basin of attraction for $\mathcal{M}$. There are some states from which both a state of coordination and miscoordination can be reached with positive probability, depending on which players are chosen to revise their strategy.

[14]This accords with empirical evidence. For example, interracial, interethnic and interreligious marriages are more likely to end in divorce than homogamous marriages (see Fu (2006) and references therein) and

When players are homogeneous (a single-population model), standard results apply (e.g. Young 1998, p. 68): the best response dynamic converges almost surely to a state of coordination from any initial state. In our two-population model, the difference in cultural constraints is critical to social miscoordination. Notice that for two cultures to be compatible and misaligned, there must be an action that group A members can take but group B members cannot and an action that group B members can take but group A members cannot. Suppose instead that $X_A \subseteq X_B$ or $X_B \subseteq X_A$. In this case, the cultures are aligned and, by Proposition 1(i), social coordination emerges in the long run between the two groups regardless of differences in their cultural ideals.

2.4 Stochastically Stable Miscoordination

Henceforth we shall examine interactions between groups that are compatible but misaligned. When the cultures are incompatible, we know that $\mathcal{M}$ is the unique recurrent class of the best response dynamic ($\varepsilon = 0$). Since the set of stochastically stable states is a non-empty subset of the set of recurrent classes of the unperturbed dynamic, $\mathcal{M}$ is the unique stochastically stable set of states.

As long as the cultures are compatible, the states of coordination are recurrent classes of the evolutionary dynamic. Moreover, each state of coordination Pareto dominates every state of miscoordination. Hence one might expect social miscoordination to be a tenuous phenomenon, accessible from only a small set of initial conditions. In this section, we show that, on the contrary, miscoordination is the most likely outcome of play in the long run for an open set of parameters. In fact, we can be more precise.

Consider the perturbed best response dynamic with error rate $\varepsilon > 0$. Again, studying stochastic stability by taking the limit $\varepsilon \to 0$ will allow us to make sharp statements about the asymptotic behavior of the evolutionary dynamic, independently of initial conditions.

As discussed in the introduction, not much is known about stochastic stability in asymmetric coordination games with more than two actions. In our setting, individuals from two different populations play a coordination game with an arbitrary number of actions. Interactions occur

cultural differences reduce the volume of and returns from cross-border mergers (Ahern et al. forthcoming). The subdiscipline of cross-cultural management is devoted to such issues. See Akerlof & Kranton (2010) for numerous examples of organizational conflict along gender, class and racial lines. On the other hand, Page (2007) shows how organizations can benefit from diversity.

between populations, with each population having different idiosyncratic payoffs over actions and different action sets. We shall now show that risk dominance is sufficient for $\mathcal{M}$ to be stochastically stable. We also derive a condition under which it is necessary and sufficient.

Define group k's *opposition to coordination on* x by $D_{kx} \equiv (d_k - \delta_{kx})/\theta_k$. This is the group's cultural bias against x scaled by its coordination payoff.

According to Harsanyi & Selten (1988), an equilibrium z^* strictly risk dominates another z^{**} if the product of (unilateral) deviation losses from z^* is greater than that from z^{**}. An equilibrium is strictly risk dominant if it pairwise strictly risk dominates all other equilibria. In our setting, miscoordination $\mathcal{M}$ strictly risk dominates a state of coordination on $x \in X_A \cap X_B$ if

$$\begin{aligned} (d_A - \delta_{Ax})(d_B - \delta_{Bx}) &> (\theta_A + \delta_{Ax} - d_A)(\theta_B + \delta_{Bx} - d_B) \\ \theta_B(d_A - \delta_{Ax}) + \theta_A(d_B - \delta_{Bx}) &> \theta_A\theta_B \\ D_{Ax} + D_{Bx} &> 1. \end{aligned}$$

Hence:

Definition 5. Miscoordination $\mathcal{M}$ is *strictly risk dominant* if

$$\min_{x \in X_A \cap X_B} \left(D_{Ax} + D_{Bx}\right) > 1. \tag{2}$$

Let us also define the notion of simple payoffs, in which cultural payoffs are uniform over mutually permissible actions:

Definition 6. Payoffs are *simple* if $\delta_{kx} = \delta_k$ for all $x \in X_A \cap X_B$ and $k = A, B$.

We can now state the result.

Proposition 2 *Suppose the groups are compatible but misaligned. Consider the perturbed best response dynamic in the small noise limit $\varepsilon \to 0$, for a sufficiently large population size n.*

(i) Miscoordination $\mathcal{M}$ is the unique stochastically stable class if it is strictly risk dominant.

(ii) When payoffs are simple, $\mathcal{M}$ is the unique stochastically stable class if and only if it is strictly risk dominant.

The proof of Proposition 2 employs a graph-theoretic method developed by Young (1993*a*). It relies on several technical lemmas (stated and proved in the appendix), as well as two tree-surgery arguments. Proposition 2(i) states that strict risk dominance is sufficient for $\mathcal{M}$ to be stochastically stable.[15] Proposition 2(ii) states that when cultural payoffs are uniform over mutually permissible actions, this condition is necessary and sufficient.

Before relating this result to prior work, several further remarks are in order.

Remark 1. The condition for strict risk dominance of $\mathcal{M}$, (2), is stronger than misalignment.

Recall that $\mathcal{M}$ is a recurrent class of the unperturbed best response dynamic if and only if the cultures are misaligned. Misalignment is equivalent to

$$\min_{x \in X_A \cap X_B} \min_{k \in \{A,B\}} D_{kx} > 0. \tag{3}$$

This does not preclude

$$\min_{x \in X_A \cap X_B} \max_{k \in \{A,B\}} D_{kx} < \frac{1}{2},$$

in which case condition (2) is violated.

In addition, (2) can only be satisfied when the cultures are misaligned. Suppose contrary to (3) that $D_{kx} \leq 0$ for some $k \in \{A, B\}$ and $x \in X_A \cap X_B$, so that the cultures are aligned. In this case,

$$D_{Ax} + D_{Bx} \leq \max_{k \in \{A,B\}} D_{kx}. \tag{4}$$

Recall that condition ND requires $\theta_k > d_k - \delta_{kx}$ for all $x \in X_A \cap X_B$ and $k \in \{A, B\}$. Hence the RHS of (4) is less than one, which in turn means that (2) is violated.

[15] The population size n needs to be sufficiently large only to break ties in resistances between recurrent classes induced by rounding up to the nearest integer. When n is small, multiple classes can be selected. But as long as $\mathcal{M}$ is stochastically stable for some population size n it remains stochastically stable for all smaller n.

How much stronger than misalignment is the condition for stochastic stability (2)? To illustrate, consider the following symmetric example.

Example 1. $\theta_k = \theta$, $d_k = d$ and $\delta_{kx} = \delta$ for all $x \in X_A \cap X_B$ and $k = A, B$.

By ND, $d - \delta < \theta$. Misalignment requires that $d - \delta > 0$. The condition for stochastic stability reduces to $d - \delta > \frac{1}{2}\theta$. For $\mathcal{M}$ to be uniquely stochastically stable then, the cultural bias need only be greater than half the coordination payoff. *Miscoordination can prevail even when incentives for coordination are much stronger than cultural incentives.*

The intuition is as follows. The mutually permissible actions yield high payoffs when everyone coordinates on them. Whether the gain from coordination is realized, however, depends on the actions of the other group. Cultural payoffs, in contrast, are payoffs to taking actions *per se*, independently of actions taken by the other group. If cultural payoffs favor mutually impermissible actions, then states of miscoordination may be more robust to random shocks, yielding higher payoffs out of equilibrium than the states of coordination. Out-of-equilibrium behavior is important in our context because the perturbed dynamic can transit between recurrent classes due to a sequence of errors. In the limit as $\varepsilon \to 0$, the dynamic spends virtually all of the time in the classes that require few errors to get into and many errors to get out of. Thus, when the cultural payoffs to mutually impermissible actions are relatively high, coordination is easily destabilized and miscoordination takes place virtually all of the time.

This mirrors the logic behind selection of risk-dominant equilibria in 2×2 coordination games (Kandori et al. 1993, Young 1993*a*) and larger coordination games satisfying certain properties (Kandori & Rob 1993), as well as half-dominant equilibria in symmetric coordination games (Ellison 2000). Nevertheless, we shall now establish that Proposition 2 does not follow from existing results. (Nor is the proof trivial.) Hence it is not obvious that the intuition derived from other settings should apply here.[16]

Young (1993*a*, p. 73) shows by example that beyond 2×2 coordination games, risk dominance need not be sufficient for stochastic stability. His example does not, however, match our payoff structure. In fact, Kandori & Rob (1993) show that, under certain assumptions, an

[16]Naturally, the results coincide when $|X_A \cap X_B| = 1$, but our analysis is more general.

equilibrium is stochastically stable if it pairwise risk dominates every other. Their result relies on a 'total bandwagon property' being satisfied, which in our setting means:

$$p\theta_k + \delta_{kx} > \delta_{kx'} \text{ for all } p > 0,\ (x, x') \in X^2 \text{ and } k = A, B,$$

or equivalently $\delta_{kx} = \delta_{kx'}$ for all $(x, x') \in X^2$ and $k = A, B$. This in turn implies that the cultures are aligned, so that $\mathcal{M}$ is not a candidate for stochastic stability. Hence the violation of one of their key assumptions is critical to our analysis.

The other result that we have to consider is the selection of half-dominant equilibria in symmetric coordination games by Ellison (2000). Our condition for $\mathcal{M}$ to be stochastically stable is weaker than half-dominance. According to Morris et al. (1995), a strategy x is half-dominant if it is a strict best response to any state in which at least $\frac{1}{2}$ of the population play x. This is a symmetric notion, that does not account for the potential asymmetries in multi-population games. In our asymmetric setting, with between-group interactions, half-dominance can be defined as follows (see Kajii & Morris 1997):[17]

Definition 7. (*Half-Dominance*) Consider a state $\widetilde{z}$ in which, for each group $k \in \{A, B\}$, at least $\frac{1}{2}n$ players choose a mutually impermissible action $x \in X_k \backslash X_{k'}$, $k \neq k'$. $\mathcal{M}$ is half-dominant if for all such states $\widetilde{z}$ and each $k = A, B$, there exists some action $x' \in X_k \backslash X_{k'}$ which is a strictly better response to $\widetilde{z}$ for k members than any action $x \in X_A \cap X_B$.

In any state $\widetilde{z}$, as in the hypothesis of half-dominance, at most half of k' members play action $x \in X_A \cap X_B$. Hence the expected payoff to a k member from playing x is at most $\frac{1}{2}\theta_k + \delta_{kx}$. Some action $x' \in X_k \backslash X_{k'}$ will be a strictly better response if $d_k > \frac{1}{2}\theta_k + \delta_{kx}$ or $D_{kx} > \frac{1}{2}$. This holds for all $x \in X_A \cap X_B$ and $k \in \{A, B\}$ if and only if

$$\min_{x \in X_A \cap X_B} \min_{k \in \{A,B\}} D_{kx} > \frac{1}{2}. \tag{5}$$

Remark 2. The condition under which miscoordination is stochastically stable, the strict risk dominance condition (2), is weaker than half-dominance, defined by condition (5).

To see this, first note that:

$$D_{Ax} + D_{Bx} \geq 2 \min_{x \in X_A \cap X_B} \min_{k \in \{A,B\}} D_{kx},$$

[17] In Kajii and Morris' terminology, what we define is a $(\frac{1}{2}, \frac{1}{2})$-dominant equilibrium.

which is greater than one by (5). Hence half-dominance implies strict risk dominance.

The converse is not true. Consider the following asymmetric example.

Example 2. Payoffs are $\theta_A = \theta_B = \theta$, $d_A = d$, $d_B = \alpha d$ where $\alpha > 1$, and $\delta_{Ax} = \delta_{Bx} = 0$ for all $x \in X_A \cap X_B$.

We require $\theta < 2d$ for (5) to be satisfied and miscoordination to be half-dominant. However, we only require $\theta < (1+\alpha)d$ for (2) to be satisfied and miscoordination to be stochastically stable. To highlight the role of asymmetry parameterized here by α, recall that condition ND requires that $\theta > \alpha d$. As $\alpha \to \frac{\theta}{d}$, $(1+\alpha)d \to d + \theta$. Hence when α is very large, (2) is satisfied for almost all $d > 0$, whereas half-dominance is only satisfied for $d > \frac{\theta}{2}$.

While this section has focused on miscoordination, Proposition 2(ii) enables us to say something about the conditions under which a state of coordination is stochastically stable. Notice that payoffs are simple in Example 2. Hence the states of coordination are stochastically stable if and only if $\theta \geq (1+\alpha)d$. That is, coordination is the long run outcome when each group's coordination payoff is greater than or equal to the sum of miscoordination payoffs across groups. More generally, one would expect that coordination payoffs need to be high relative to the cultural biases against coordination for a state of coordination to prevail in the long run.

Now it could be that when payoffs are not simple, $\mathcal{M}$ is uniquely stochastically stable despite not being strictly risk dominant. This remains an open question. We will, however, derive a (weaker) necessary and sufficient condition for $\mathcal{M}$ to be stochastically stable in section 3.2 under a different error structure, without requiring simple payoffs.

3 An Application and Extension

This section analyzes an application of our model to identity-based conflict and subsequently extends the model to allow for occasional violation of cultural constraints.

3.1 Identity-Based Conflict

To illustrate how our model can be applied, consider the following application to identity-based conflict inspired by Sen (2006). The setting will be rather stylized to keep the focus on evolutionary pressures and how they shape identity formation rather than on precise institutional details.

Suppose that individuals belonging to different cultural groups, A and B, interact in pairs. They each choose one of $L \geq 3$ identities $x \in X$ and in doing so are bound by cultural constraints. Let identities 1 and L be *exclusive*, clearly defining an ingroup and outgroup along ethnic or religious lines. Identity 1 is an exclusive group A identity and identity L is an exclusive group B identity. All other identities are *inclusive*; they are not associated with an ethnicity or religion and may even emphasize membership in some common social category (e.g. nation).[18]

Identity choice is constrained in the following manner: a member of group A would not consider adopting the exclusive group B identity, $X_A = X \backslash \{L\}$. Likewise, a group B member would not consider adopting the exclusive group A identity, $X_B = X \backslash \{1\}$. To be concrete, an Egyptian Muslim may grow a beard and become an active member of the Muslim Brotherhood—an exclusive Muslim identity. An Egyptian Christian may tattoo a Coptic cross on his wrist and become an active member of the Coptic church—an exclusive Christian identity. Alternatively, they could both choose to adopt a neutral Egyptian identity without any clear religious markers. What an Egyptian Muslim would not ordinarily do is consider adopting an exclusive Christian identity and *vice versa*.

Given these cultural constraints, we shall show that exclusive ethnic or religious identities can persist even when they are Pareto dominated by coordination on an inclusive identity.

Let us specify payoffs. Suppose that players obtain privileged access to group resources worth β if they adopt an exclusive identity in intergroup interactions. Iannaccone (1992, 1994) demonstrate how religious groups that require stigmatizing modes of dress and behavior can more efficiently produce religious club goods, including social welfare services. For example, growing a beard and joining the Muslim Brotherhood can give one preferential access to group-provided healthcare services and employment opportunities; it can also confer higher

[18] Note that individuals are born into either group A or B. This is not a choice. However, they can choose whether to emphasize their exclusive group identity or adopt an inclusive identity.

status within one's group (e.g. Wickham 2002). There are, however, even stronger incentives to coordinate in our model, as coordination on an inclusive identity limits the chance of conflict between members of different cultural groups.

Each individual has property worth w. If the two players in an interaction choose the same identity, then no conflict takes place and each individual retains w. If their identities differ, then a violent winner-takes-all contest ensues in which each individual's property is up for grabs. The total prize in this contest, worth $2w$, is allocated based on a standard Tullock contest.[19] The probability that the group A member wins the prize when exerting e_A units of effort in the contest is $e_A/(e_A + e_B)$, while the cost of exerting e_A units of effort is simply e_A. The expected payoff to the group A member is thus:

$$\frac{e_A}{e_A + e_B} 2w - e_A. \tag{6}$$

The expected payoff to individual B is given by switching the A and B subscripts. The unique Nash equilibrium contest payoff for each individual is $\frac{1}{2}w$.

Thus the payoff from choosing an inclusive identity $x \in \{2, \ldots L-1\}$ is $\mathbb{I}(x, x_j)w + (1 - \mathbb{I}(x, x_j))\frac{1}{2}w$, where x_j is the action chosen by the player with whom the individual is matched. That is, if coordination on an inclusive identity occurs, $\mathbb{I}(x, x) = 1$, individuals retain their existing wealth w. Otherwise, conflict ensues and each receives an expected contest payoff of $\frac{1}{2}w$. The payoff from choosing an exclusive identity is $\frac{1}{2}w + \beta$, since conflict occurs for sure. This is equivalent to our model of coordination with coordination payoffs $\theta_A = \theta_B = \frac{1}{2}w$ and cultural biases $d_A - \delta_{Ax} = d_B - \delta_{Bx} = \beta$.

Notice that payoffs are simple, meaning uniform over all inclusive (mutually permissible) identities $x \in X_A \cap X_B$. According to Proposition 2(ii) then, condition (2) is necessary and sufficient for miscoordination to be stochastically stable. That is, $\mathcal{M}$ is stochastically stable if and only if $\beta/\theta_A + \beta/\theta_B > 1$. Define the critical level of cultural bias, denoted by $\overline{\beta}$, as follows:

$$\overline{\beta}\left(\frac{1}{\theta_A} + \frac{1}{\theta_B}\right) = 1. \tag{7}$$

Miscoordination is uniquely stochastically stable (for n sufficiently large) if and only if $\beta > \overline{\beta}$. Substituting $\theta_A = \theta_B = \frac{1}{2}w$ and manipulating (7), we find that $\overline{\beta} = \frac{1}{4}w$. That is to say,

[19] This contest function was first applied to conflict by Skaperdas (1992) and intergroup conflict by McBride et al. (2011) and Sambanis & Shayo (2013). Sambanis and Shayo suggest that lower levels of ethnic conflict are observed when a common national identity is adopted, because individuals who adopt an ethnic identity aim to maximize the share of national resources captured by their ethnic group.

exclusive identities will be adopted virtually all of the time if the value of access to cultural group resources, β, is greater than one quarter of an individual's wealth. Holding the value of group resources constant, one could predict on this basis that poorer societies are 'more likely' to end up miscoordinating.[20,21]

Now consider a mean-preserving wealth spread, where $w_A = w + \Delta$ and $w_B = w - \Delta$. Obviously, this can be generated by transferring resources from one group to the other. Coordination payoffs are now:

$$\begin{aligned} \theta_A &= w + \Delta - \tfrac{1}{2}w = \tfrac{1}{2}w + \Delta \\ \theta_B &= w - \Delta - \tfrac{1}{2}w = \tfrac{1}{2}w - \Delta. \end{aligned} \tag{8}$$

Note that in the event of miscoordination, the expected payoff in the contest is still $\frac{1}{2}w$ for each individual, since the total prize $(2w)$ remains the same.

We show that the critical level of cultural bias is decreasing in Δ. That is, more unequal societies are 'more likely' to end up miscoordinating. The reason for this is not obvious. Notice that inequality does not alter either agent's equilibrium contest payoff of $\frac{1}{2}w$. Hence this is not a matter of the poorer group having an incentive to exert greater expropriation effort. The mean-preserving wealth spread does however translate into a mean-preserving spread in coordination payoffs [see (8)]. It turns out that the effect of this change is asymmetric, in the following sense. Fewer errors by A members are now required for an exclusive identity to be a best response for poorer B members. On the other hand, a larger number of errors by B members is required for an exclusive identity to be a best response for wealthier A members. That is, poorer individuals become more inclined to choose exclusive identities, while wealthier individuals persist more strongly in coordination. The former dominates, however, and the net effect is to increase the stability of miscoordination. *Hence inequality can translate into ethnic/religious conflict, not for the usual reasons related to incentives for expropriation, but due to out-of-equilibrium selection pressures.*

[20] Of course, the value of access to group resources β may depend on w. But we have reason to think that this effect is weak. Production of club goods typically involves labor-intensive modes of production (Iannaccone 1992, Berman 2000), especially in poorer societies (see Wickham 2002).

[21] Fearon & Laitin (2003) report a large and significant negative relationship between income levels and the incidence of civil war. Their explanation centers on the link between higher income levels and the opportunity cost of insurgency. More closely related to our explanation is work by McBride et al. (2011). They propose that conflict destroys resources. Hence in a repeated games setting, higher income levels provide greater incentive for groups to strike a peaceful agreement.

We can state the following proposition:

Proposition 3 *Consider the example of identity-based conflict with wealth levels* $w_A = w+\Delta$ *and* $w_B = w - \Delta$*, where* $\Delta \in [0, \frac{1}{2}w)$*. The critical level of cultural bias is:*

(*i*) *strictly increasing in* w*,*

(*ii*) *strictly decreasing in* Δ*.*

The conclusion from this stylized example is that exclusive ethnic and religious identities are more persistent in poorer, more unequal societies.

In related work on 2×2 coordination games by Quilter (2007) and Neary (2012), members of two groups play a single action against all members of the population. A different form of miscoordination can arise in their setting, due to a tradeoff between ingroup coordination on one's preferred action and coordination with outgroup members. Our analysis points to a different, but not mutually exclusive, source of miscoordination based on interactions between groups. Indeed there is reason to believe that identity choice may be driven primarily by its effect on outgroup interactions. A central theme in the economics of religion literature is that exclusive identities are designed as a 'tax' on outgroup activity (see Iannaccone 1992, Berman 2000). For example, Aksoy & Gambetta (2015) show that contact with natives in Belgium increases the propensity of highly-religious Muslim women to veil. See also Douglas (1966) on how taboos maintain community boundaries. In addition, our analysis links miscoordination to economically relevant variables such as poverty and inequality, and applies when individuals can choose different actions when interacting with ingroup and outgroup members.

3.2 Violating Cultural Constraints

Consider two compatible but misaligned groups. We have established that differing cultural constraints can lead to miscoordination among such groups. Does this result hinge upon strict observance of cultural constraints? Does occasional violation of cultural constraints destabilize miscoordination?

We shall address this question in the following manner. Recall that introducing small errors into the strategy revision process enables us to make sharp predictions about the evolution of play in the long run. We have hitherto assumed that individuals do not violate cultural constraints when making errors. An error by a revising group k member has so far consisted of a strategy chosen at random from X_k. We shall now allow errors to be *unconstrained*. With probability ε, a revising player chooses an action at random from the global action set X. (Note that (1) specifies payoffs even for $x \notin X_k$.) This means that individuals can violate cultural constraints, but the likelihood of doing so vanishes at the same rate as regular nonbest responses.[22] It turns out that allowing players to violate cultural constraints in this way leads to sharper results and strengthens the stability of miscoordination $\mathcal{M}$.

Another interpretation of this specification is that all players can choose all actions in X, but actions in $X \backslash X_k$ are strictly dominated for k members: $d_k > \theta_k + \delta_{kx}$ for $x \in X \backslash X_k$. Hence such actions are only chosen erroneously. Kim & Wong (2010) show that any recurrent class can be made stochastically stable by introducing some dominated strategies. Adding dominated strategies in the manner described here, however, only strengthens the stability of miscoordination.

We can now state a necessary and sufficient condition for $\mathcal{M}$ to be stochastically stable, without imposing restrictions such as simple payoffs.

Proposition 4 *Suppose the groups are compatible but misaligned. Consider the perturbed best response dynamic with unconstrained errors. For n sufficiently large, the unique stochastically stable class is miscoordination $\mathcal{M}$ if and only if*

$$\min_{x \in X_A \cap X_B} \left(2 \min_{k \in \{A,B\}} D_{kx} + \max_{k \in \{A,B\}} D_{kx} \right) > 1. \tag{9}$$

By inspection, condition (9) is weaker than strict risk dominance (2). Let us reconsider Example 2 for a moment. Under constrained errors, we learned that the states of coordination are stochastically stable if and only if $\theta \geq (1+\alpha)d$. Under unconstrained errors, (9) implies that the states of coordination are stochastically stable if and only if $\theta \geq (2+\alpha)d$. For

[22] A non-vanishing likelihood of violating cultural constraints is analytically equivalent to allowing individuals to revise their cultural constraints.

coordination to be stochastically stable, coordination payoffs still need to be high relative to cultural biases against coordination. Occasional violation of cultural constraints raises the threshold, making coordination harder to achieve.

The intuition for this result is as follows. When errors are unconstrained, members of one group can make a sequence of errors in which they violate cultural constraints and choose a cultural ideal of the other group. At this point, not only can a revising player gain from choosing his cultural ideal as before, but there is also some chance that he coordinates on his own cultural ideal with a deviant player from the other group. Hence introducing such errors can make it easier to break from a state of coordination and can thereby strengthen miscoordination. Of course, large-scale violations of cultural constraints may ultimately undermine the force of such constraints, a situation examined elsewhere by Fershtman et al. (2011).

4 Concluding Comments

This paper studies interactions between members of two cultural groups who are matched to play a coordination game with an arbitrary number of actions. When individuals observe cultural constraints, miscoordination between cultural groups is a surprisingly stable phenomenon. Occasional violation of cultural constraints can further strengthen miscoordination. Unlike prior work, the type of miscoordination uncovered here occurs even when there is no tradeoff between ingroup and outgroup coordination. Yet, since externalities are limited in our model, it is possible that we have understated the likelihood of miscoordination. For example, we have only considered payoffs from adopting an inclusive identity which are linear in the number of outgroup members who adopt an inclusive identity. There are, however, cases in which only a few salient deviants are needed to create general distrust between ethnic groups, communal riots and civil war (e.g. de Figueiredo & Weingast 1999). Such externalities would compound the inefficiencies uncovered in this paper.

References

Ahern, K. R., Daminelli, D. & Fracassi, C. (forthcoming), 'Lost in translation? The effect of cultural values on mergers around the world', *Journal of Financial Economics*.

Akerlof, G. A. & Kranton, R. E. (2000), 'Economics and identity', *Quarterly Journal of Economics* **415**(3), 715–753.

Akerlof, G. A. & Kranton, R. E. (2010), *Identity Economics: How our Identities Shape our Work, Wages and Wellbeing*, Princeton University Press, Princeton, NJ.

Aksoy, O. & Gambetta, D. (2015), 'Behind the veil: The strategic use of religious garb', working paper, Nuffield College, University of Oxford.

Bénabou, R. & Tirole, J. (2011), 'Identity, morals, and taboos: Beliefs as assets', *Quarterly Journal of Economics* **126**, 805–855.

Berman, E. (2000), 'Sect, subsidy, and sacrifice: an economist's view of ultra-orthodox Jews', *Quarterly Journal of Economics* **115**(3), 905–953.

Bisin, A. & Verdier, T. (2000), 'Beyond the melting pot: cultural transmission, marriage, and the evolution of ethnic and religious traits', *Quarterly Journal of Economics* **115**(3), 955–988.

Carvalho, J.-P. (2013), 'Veiling', *Quarterly Journal of Economics* **128**(1), 337–370.

de Figueiredo, R. & Weingast, B. (1999), The rationality of fear: Political opportunism and ethnic conflict, *in* B. Walter & J. Snyder, eds, 'Civil Wars, Insecurity, and Intervention', Columbia University Press, New York, NY, pp. 261–302.

Douglas, M. (1966), *Purity and Danger: An Analysis of the Concepts of Pollution and Taboo.*, Routledge, London.

Ellison, G. (1993), 'Learning, local interaction and coordination', *Econometrica* **61**(5), 1047–1071.

Ellison, G. (2000), 'Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution', *Review of Economic Studies* **67**, 17–45.

Elster, J. (1989), 'Social norms and economic theory', *Journal of Economic Perspectives* **4**, 99–117.

Esteban, J. & Ray, D. (2008), 'On the salience of ethnic conflict', *American Economic Review* **98**(5), 2185–2202.

Fearon, J. D. & Laitin, D. D. (1996), 'Explaining interethnic cooperation', *American Political Science Review* **90**(4), 715–735.

Fearon, J. D. & Laitin, D. D. (2003), 'Ethnicity, insurgency, and civil war', *American Political Science Review* **97**(1), 75–90.

Fershtman, C., Gneezy, U. & Hoffman, M. (2011), 'Taboos and identity: Considering the unthinkable', *American Economic Journal: Microeconomics* **3**(2), 139–164.

Foster, D. & Young, H. P. (1990), 'Stochastic evolutionary game dynamics', *Theoretical Population Biology* **38**, 219–232.

Frank, R. H. (1988), *Passions within Reason*, Norton, New York, NY.

Fu, X. (2006), 'Impact of socioeconomic status on inter-racial mate selection and divorce', *Social Science Journal* **43**(2), 239–258.

Greif, A. (1993), 'Contract enforceability and economic institutions in early trade: the Maghribi trader's coalition', *American Economic Review* **83**(3), 525–548.

Greif, A. & Tabellini, G. (2012), 'The clan and the city: Sustaining cooperation in China and Europe', working paper, Stanford University.

Harsanyi, J. (1982), Morality and the theory of rational behavior, *in* A. Sen & B. Williams, eds, 'Utilitarianism and Beyond', Cambridge University Press, Cambridge, pp. 39–62.

Harsanyi, J. & Selten, R. (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA.

Henrich, J., Albers, R. W., Boyd, R., Gigerenzer, G., McCabe, K. A., Ockenfels, A. & Young, H. P. (2001), What is the role of culture in bounded rationality?, *in* G. Gigerenzer & R. Selten, eds, 'Bounded Rationality: The Adaptive Toolbox', MIT Press, Cambridge, MA, pp. 343–359.

Iannaccone, L. R. (1992), 'Sacrifice and stigma: Reducing free-riding in cults, communes, and other collectives', *Journal of Political Economy* **100**(2), 271–291.

Iannaccone, L. R. (1994), 'Why strict churches are strong', *American Journal of Sociology* **99**(5), 1180–1211.

Jha, S. (2013), 'Trade, institutions and ethnic tolerance: Evidence from South Asia', *American Political Science Review* **107**(4).

Kajii, A. & Morris, S. (1997), 'The robustness of equilibria to incomplete information', *Econometrica* **65**(6), 1283–1309.

Kandori, M., Mailath, G. & Rob, R. (1993), 'Learning, mutation, and long-run equilibria in games', *Econometrica* **61**(1), 29–56.

Kandori, M. & Rob, R. (1993), 'Bandwagon effects and long run technology choice', discussion paper 93-F-2, University of Tokyo.

Kim, C. & Wong, K.-C. (2010), 'Long-run equilibria with dominated strategies', *Games and Economic Behavior* **68**, 242–254.

Kuran, T. (1998), 'Ethnic norms and their transformation through reputational cascades', *Journal of Legal Studies* **27**(2), 623–659.

Kuran, T. & Sandholm, W. H. (2008), ‘Cultural integration and its discontents’, *Review of Economic Studies* **75**, 201–228.

Laitin, D. D. (2007), *Nations, States and Violence*, Oxford University Press, New York, NY.

McBride, M., Milante, G. & Skaperdas, S. (2011), ‘Peace and war with endogenous state capacity’, *Journal of Conflict Resolution* **55**(3), 446–468.

Morris, S., Rob, R. & Shin, H. S. (1995), ‘p-Dominance and belief potential’, *Econometrica* **63**(1), 145–157.

Myatt, D. P. & Wallace, C. C. (2004), ‘Adaptive play by idiosyncratic agents’, *Games and Economic Behavior* **48**, 124–138.

Neary, P. R. (2012), ‘Competing conventions’, *Games and Economic Behavior* **76**(1), 301–328.

Page, S. E. (2007), *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton University Press, Princeton, NJ.

Patel, D. S. (2012), ‘Concealing to reveal: The informational role of Islamic dress in Muslim societies’, *Rationality and Society* **24**(3), 295–323.

Quilter, T. (2007), ‘Noise matters in heterogeneous populations’, working paper, Department of Economics, University of Edinburgh.

Rao, V. & Walton, M. (2004), Culture and public action: relationality, equality of agency, and development, *in* V. Rao & M. Walton, eds, ‘Culture and Public Action’, Stanford University Press, Stanford, CA, pp. 2–36.

Roth, A. E. (2007), ‘Repugnance as a constraint on markets’, *Journal of Economic Perspectives* **21**(3), 37–58.

Sambanis, N. & Shayo, M. (2013), ‘Social identification and ethnic conflict’, *American Political Science Review* **107**(2), 294–325.

Samuelson, L. & Zhang, J. (1992), ‘Evolutionary stability in asymmetric games’, *Journal of Economic Theory* **57**, 363–391.

Sandholm, W. H., ed. (2010), *Population Games and Evolutionary Dynamics*, MIT Press, Cambridge, MA.

Sen, A. (1977), ‘Rational fools: a critique of the behavioural foundations of economic theory’, *Philosophy and Public Affairs* **6**, 317–344.

Sen, A. (2006), *Identity and Violence: The Illusion of Destiny*, W.W. Norton, New York, NY.

Skaperdas, S. (1992), 'Cooperation, conflict, and power in the absence of property rights', *American Economic Review* **82**(4), 720–739.

Staudigl, M. (2012), 'Stochastic stability in asymmetric binary choice coordination games', *Games and Economic Behavior* **75**, 372–401.

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C. & Lerner, J. S. (2000), 'The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals', *Journal of Personality and Social Psychology* **78**(5), 853–870.

Wickham, C. R. (2002), *Mobilizing Islam: Religion, Activism and Social Change in Egypt*, Columbia University Press, New York, NY.

Young, H. P. (1993*a*), 'The evolution of conventions', *Econometrica* **61**(1), 57–84.

Young, H. P. (1993*b*), 'An evolutionary model of bargaining', *Journal of Economic Theory* **59**, 145–168.

Young, H. P. (1996), 'The economics of convention', *Journal of Economic Perspectives* **10**(2), 105–122.

Young, H. P. (1998), *Individual Strategy and Social Structure*, Princeton University Press, Princeton, NJ.

Young, H. P. & Burke, M. A. (2001), 'Competition and custom in economic contracts: A case study of Illinois agriculture', *American Economic Review* **91(3)**, 559–573.

Appendix

Proof of Proposition 1. We shall characterize the set of recurrent classes of the best response dynamic. As any finite Markov chain converges almost surely to one of its recurrent classes, this establishes the proposition.

Firstly, we claim that there are no recurrent classes of the best response dynamic other than the states of coordination and $\mathcal{M}$. To establish the claim it suffices to show that there is a positive probability of transiting from any state to a state of coordination or a state in $\mathcal{M}$ in a finite number of periods.

Consider an arbitrary time t and state z^t. Let x be a best response for group A members. Suppose that in period t, the set of revising players is $R_t = N_A$ and that each of them chooses x. In addition, suppose that in period $t+1$, $R_{t+1} = N_B$. All of this occurs with positive probability.

Case 1. $x \in X_A \cap X_B$. (By definition, this is impossible if the cultures are incompatible.) Then the expected payoff to each revising group B member from playing x is $\theta_{Bx} + \delta_{Bx}$ for all samples. The payoff from choosing $x' \neq x$ is at most d_B, which is less than the payoff from x by ND. Hence each revising group B member chooses x and the dynamic transits with probability to a state of coordination in two periods.

Case 2. $x \notin X_A \cap X_B$. Then the payoff to each revising group B member from playing any $x' \in X_B$ is $\delta_{Bx'}$. Hence each revising group B member chooses an action in $\widetilde{X}_B \equiv argmax_{x \in X_B} \delta_{Bx}$. When the cultures are misaligned, every such action lies outside $X_A \cap X_B$, as does x. Hence the dynamic transits with positive probability to a state in $\mathcal{M}$ in two periods. When the cultures are aligned, we have either $\tilde{X}_A \cap X_B$ is nonempty or $\tilde{X}_B \cap X_A$ is nonempty. Thus the dynamic transits with positive probability in at most two periods to a state in which each group B member plays an action $x' \in X_A \cap X_B$. From there, the argument used in case 1 establishes that the dynamic transits to a state of coordination in finite time.

This establishes the claim.

The fact that the states of coordination are recurrent classes of the best response dynamic if and only if the cultures are compatible follows immediately from the argument in case 1. That $\mathcal{M}$ is a recurrent class if and only if the cultures are misaligned follows immediately from the argument in case 2. This establishes the proposition. □

Proposition 2. The proof of Proposition 2 employs two tree-surgery arguments. These arguments make use of the following three lemmas and corollary, which compare the resistance of transitions between recurrent classes, defined as follows. The *cost* of a path $(z^1, z^2, \ldots z^H)$ is the minimum number of errors required for the perturbed dynamic ($\varepsilon > 0$) to transit from state z^1 to z^H along this path. The *resistance* $r(z^1, z^H)$ is the minimum cost of a path starting at z^1 and ending in z^H. The definition can be extended so that for sets E and E', the resistance $r(E, E')$ is the minimum cost of a path starting in some state in E and ending in some state in E'.

In addition, let $x_i(z)$ be the action played by i in state z. A *direct path* $(z^1, z^2, \ldots z^H)$ from state z^1 to z^H is one in which $x_i(z^h) \in \{x_i(z^1), x_i(z^H)\}$ for all $h = 1, 2, \ldots H$ and $i \in N$.

Define $X_i(E) \equiv \{x_i(z) : z \in E\}$. A direct path $(z^1, z^2, \ldots z^H)$ from set E to E' is one in which $z^1 \in E$, $z^H \in E'$ and $x_i(z^h) \in X_i(E) \cup X_i(E')$ for all $h = 1, 2, \ldots H$ and $i \in N$.

Lemma 1 *Let z_x denote the state of coordination on x. For all distinct pairs $(x, x') \in (X_A \cap X_B)^2$,*

$$r(\mathcal{M}, z_x) \leq r(z_{x'}, z_x).$$

Proof. Let $r(z_{x'}, z_x) = a$. This means that for some $k \in \{A, B\}$, after a errors by k' members, x is a best response for k members, and hence a weakly better response than a mutually impermissible action. That is:

$$d_k \leq \frac{a}{n}\theta_k + \delta_{kx}. \tag{10}$$

Let α_k be the minimum number of errors such that (10) holds:

$$\alpha_k = \left\lceil \frac{d_k - \delta_{kx}}{\theta_k} n \right\rceil \equiv \lceil D_{kx} n \rceil. \tag{11}$$

We know $r(z_{x'}, z_x) = a \geq \min_{k \in \{A,B\}} \alpha_k$. By inspection of (10), $\min_{k \in \{A,B\}} \alpha_k$ is the minimum cost of a direct path from any state in $\mathcal{M}$ to z_x. Hence $r(\mathcal{M}, z_x) \leq \min_{k \in \{A,B\}} \alpha_k$. This establishes the Lemma. □

Lemma 2 *Suppose that in state z, n_a members of group A and n_b members of group B use strategy x. Define z' as the state in which the same is true, with the remaining $n - n_a$ members of group A and $n - n_b$ members of group B using mutually impermissible actions. Then $r(z', z_x) \leq r(z, z_x)$.*

Proof. The argument in the proof of Lemma 1 establishes the result. □

Lemma 3 *If miscoordination $\mathcal{M}$ is strictly risk dominant, then for n sufficiently large*

$$r(z_x, \mathcal{M}) < r(\mathcal{M}, z_x)$$

for all $x \in X_A \cap X_B$.

Proof. We claim that a direct path from $\mathcal{M}$ to z_x has minimum cost among all paths from $\mathcal{M}$ to z_x. Consider an arbitrary indirect path from $\mathcal{M}$ to z_x. Since the path is indirect, it runs through a state z in which some mutually permissible action $x' \neq x$ is played. The minimum cost of such a path is $r(\mathcal{M}, z) + r(z, z_x)$.

In state z, replace every action by a k member not equal to x with an action in $X_k \backslash X_{k'}$, $k' \neq k$. Label this state z'. There is a direct path from $\mathcal{M}$ to z_x through state z'. The minimum cost of such a path is $r(\mathcal{M}, z') + r(z', z_x)$. Clearly, $r(\mathcal{M}, z') < r(\mathcal{M}, z)$. In addition, $r(z', z_x) \leq r(z, z_x)$ by Lemma 2. Hence the direct path has lower cost than the indirect path, establishing the claim.

By (11), the minimum cost of a direct path is $\min_{k\in\{A,B\}} \lceil D_{kx} n \rceil$. Hence we need only find a path from z_x to $\mathcal{M}$ with lower cost. Consider a direct path. Starting from z_x, k' members make a errors. For $x'' \in \widetilde{X}_k \backslash X_{k'}$ to be a best response for k members and to thereby reach a state in $\mathcal{M}$ without further errors, a must satisfy

$$\left(1 - \frac{a}{n}\right)\theta_k + \delta_{kx} \leq d_k.$$

That is,

$$a \geq \left(1 - \frac{d_k - \delta_{kx}}{\theta_k}\right) n \equiv (1 - D_{kx}) n.$$

Thus the minimum cost of such a transition is

$$\min_{k\in\{A,B\}} \lceil (1 - D_{kx}) n \rceil.$$

For n sufficiently large, this is less than $\min_{k\in\{A,B\}} \lceil D_{kx} n \rceil$ if and only if

$$\begin{aligned} \min_{k\in\{A,B\}} D_{kx} &> \min_{k\in\{A,B\}} \big(1 - D_{kx}\big) \\ \min_{k\in\{A,B\}} D_{kx} + \max_{k\in\{A,B\}} D_{kx} &> 1 \\ D_{Ax} + D_{Bx} &> 1. \end{aligned} \tag{12}$$

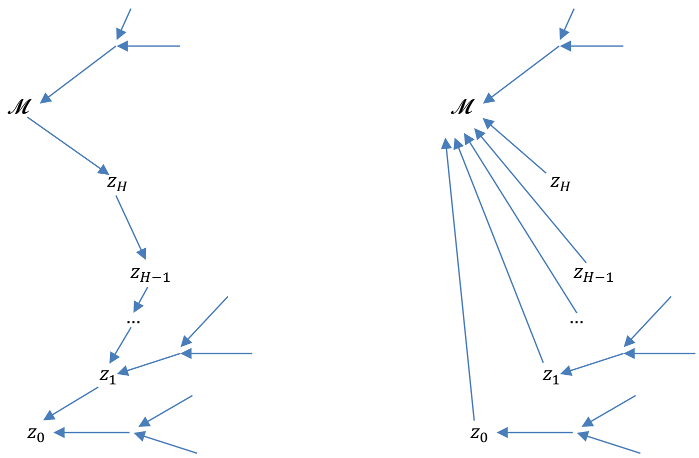
The last inequality holds, because $\mathcal{M}$ strictly risk dominates z_x [see (2)]. $\square$

The next result follows immediately from Lemmas 1 and 3:

Corollary 1 *If miscoordination $\mathcal{M}$ is strictly risk dominant, then for n sufficiently large*

$$r(z_x, \mathcal{M}) < r(z_{x'}, z_x)$$

for all distinct pairs $(x, x') \in (X_A \cap X_B)^2$.



(a) 0-tree (b) $\mathcal{M}$-tree

Figure 1: Beginning with the 0-tree in (a), steps 1-2 produce the $\mathcal{M}$-tree in (b).

Proof of Proposition 2. (i) Suppose that $\mathcal{M}$ is strictly risk dominant. We shall show that for n sufficiently large, $\mathcal{M}$ is the unique stochastically stable class using the spanning-tree method of Young (1993a).

The following tree-surgery argument is used. An x-tree is a tree consisting of a set of $|X_A \cap X_B|$ directed edges, one emanating from each recurrent class (node) other than z_x. Consider a tree, say rooted at 0 and denoted by T_0, that has minimum resistance among x-trees. Through a series of operations on T_0 we will produce an $\mathcal{M}$-tree with lower resistance than T_0. Hence all trees with minimum resistance are $\mathcal{M}$-trees. By Young (1993a, Theorem 2) then, $\mathcal{M}$ is the unique stochastically stable class.

Again let T_0 have minimum resistance among x-trees. By definition, there is a unique route from $\mathcal{M}$ to z_0 through T_0. (See Figure 1(a) for example.) Denote the nodes on this route by $z_H, z_{H-1}, \ldots z_1, z_0$. The edges composing this route are $(\mathcal{M}, z_H), (z_H, z_{H-1}), \ldots (z_1, z_0)$.

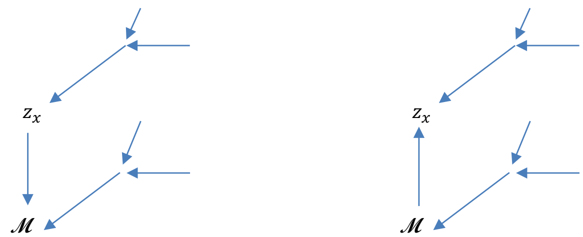
Step 1. Replace edge $(\mathcal{M}, z_H)$ with $(z_H, \mathcal{M})$.

Step 2. Replace each edge (z_h, z_{h-1}) with $(z_{h-1}, \mathcal{M})$ for $h = 1, \ldots H$.

Steps 1-2 produce an $\mathcal{M}$-tree. (See Figure 1(b) for example.) The difference in resistance between the original 0-tree and this $\mathcal{M}$-tree is:

$$\big[r(\mathcal{M}, z_H) - r(z_H, \mathcal{M})\big] + \sum_{h=1}^{H} \big[r(z_h, z_{h-1}) - r(z_{h-1}, \mathcal{M})\big]. \tag{13}$$

For n sufficiently large, the first term is positive by Lemma 3 and the second term, if it exists, is positive by Corollary 1. Hence the $\mathcal{M}$-tree produced by steps 1-2 has lower



(a) $\mathcal{M}$-tree (b) x-tree

Figure 2: Beginning with the $\mathcal{M}$-tree in (a), replace edge $(z_x, \mathcal{M})$ with $(\mathcal{M}, z_x)$ to produce the x-tree in (b).

resistance than the original x-tree, T_0. Since the original x-tree was chosen arbitrarily, all trees with minimum resistance are $\mathcal{M}$-trees, establishing the result.

(ii) Again a tree-surgery argument is employed. Consider an arbitrary $\mathcal{M}$-tree. There exists at least one edge $(z_x, \mathcal{M})$, for some $x \in X_A \cap X_B$. Replace this edge with $(\mathcal{M}, z_x)$. This transforms the original $\mathcal{M}$-tree into an x-tree. See Figure 2 for an illustration.

The difference in resistance between the original $\mathcal{M}$-tree and this x-tree is:

$$r(z_x, \mathcal{M}) - r(\mathcal{M}, z_x).$$

We claim that this is nonnegative when payoffs are simple and $\mathcal{M}$ is not strictly risk dominant. Hence for every $\mathcal{M}$-tree there exists an x-tree with no greater resistance. By Young (1993*a*, Theorem 2) then, $\mathcal{M}$ is not the unique stochastically stable class.

Let us now establish the claim. In the proof of Lemma 3, we showed that $r(\mathcal{M}, z_x) = \min_{k\in\{A,B\}}\lceil D_{kx}n\rceil$ and the cost of a direct path from z_x to $\mathcal{M}$ equals $\min_{k\in\{A,B\}}\lceil(1-D_{kx})n\rceil$.

An indirect path from z_x to $\mathcal{M}$ involves at least b erroneous plays of some mutually permissible action $x' \neq x$ by k' members, so that x' is a best response for k members. Hence b satisfies:

$$\begin{aligned}\left(1-\frac{b}{n}\right)\theta_k + \delta_{kx} &\leq \frac{b}{n}\theta_k + \delta_{kx'} \\ 2\frac{b}{n}\theta_k &\geq \theta_k + \delta_{kx} - \delta_{kx'} \\ b &\geq \frac{1}{2}n.\end{aligned}$$

The last line utilizes the fact that $\delta_{kx} = \delta_{kx'}$ because payoffs are simple. Thus the cost of an indirect path from z_x to $\mathcal{M}$ is at least $\lceil\frac{1}{2}n\rceil$.

Hence when payoffs are simple, $r(z_x, \mathcal{M}) \geq r(\mathcal{M}, z_x)$ if n is sufficiently large and

$$\min\left\{\min_{k\in\{A,B\}}(1-D_{kx}), \frac{1}{2}\right\} \geq \min_{k\in\{A,B\}} D_{kx}. \tag{14}$$

We shall now show that (14) is satisfied. By hypothesis, $\mathcal{M}$ is not strictly risk dominant. As payoffs are simple, this means that $\mathcal{M}$ is risk dominated by all states of coordination. By (12) this is equivalent to

$$\min_{k\in\{A,B\}} D_{kx} \leq \min_{k\in\{A,B\}} \big(1-D_{kx}\big), \tag{15}$$

which implies

$$\begin{aligned} \min_{k\in\{A,B\}} D_{kx} &\leq \max_{k\in\{A,B\}} \big(1-D_{kx}\big) \\ \min_{k\in\{A,B\}} D_{kx} &\leq 1-\min_{k\in\{A,B\}} D_{kx} \\ \min_{k\in\{A,B\}} D_{kx} &\leq \tfrac{1}{2}. \end{aligned} \tag{16}$$

Together (15) and (16) imply that (14) holds. This establishes the claim and indeed part (ii) of the proposition. □

Proof of Proposition 3. The result follows immediately from the fact that $1/x$ is strictly positive, strictly decreasing, and strictly convex. □

Proposition 4. When $\varepsilon = 0$, there is no violation of cultural constraints. Hence the recurrent classes of the best response dynamic ($\varepsilon = 0$) are still given by Proposition 1. These are the candidates for stochastic stability.

The only change in computation of resistances is to transitions from states of coordination to $\mathcal{M}$. Hence Lemmas 1 and 2 still apply.

We shall now introduce two new lemmas and a corollary. Lemma 4 and Corollary 2 are analogs of Lemma 3 and Corollary 1 respectively for the case of unconstrained errors. We shall then employ the tree-surgery arguments used in Proposition 2 to establish the result.

Lemma 4 *If (9) is satisfied, then for n sufficiently large*

$$r(z_x, \mathcal{M}) < r(\mathcal{M}, z_x)$$

for all $x \in X_A \cap X_B$.

Proof. Recall that $r(\mathcal{M}, z_x) = \min_{k\in\{A,B\}} \lceil D_{kx} n \rceil$.

Let us now compute $r(z_x, \mathcal{M})$. On the minimum cost path from z_x to $\mathcal{M}$ one of two things must occur. Starting from z_x, through a series of errors, either (i) a mutually impermissible action becomes a best response, in which case the process transits to $\mathcal{M}$ without any further errors, or (ii) an action $x' \in X_A \cap X_B$ becomes a best response.

The minimum cost path conforming to case (i) is a direct path as follows. As errors are unconstrained, a transition from z_x to $\mathcal{M}$ can occur with a erroneous plays of an action in $\widetilde{X}_k \setminus X_{k'}$ by k' members, where a satisfies

$$\left(1 - \frac{a}{n}\right)\theta_k + \delta_{kx} \leq \frac{a}{n}\theta_k + d_k. \tag{17}$$

The value of a that equates (17) is $(1 - D_{kx})\frac{n}{2}$. Hence the cost of such a path from z_x to $\mathcal{M}$ is

$$\min_{k\in\{A,B\}} \left\lceil (1 - D_{kx})\frac{n}{2} \right\rceil = \min_{k\in\{A,B\}} \left\lceil \left(1 - \frac{d_k - \delta_{kx}}{\theta_k}\right)\frac{n}{2} \right\rceil. \tag{18}$$

The minimum cost path conforming to case (ii) involves at least b erroneous plays of some mutually permissible action $x' \neq x$ by k' members, so that x' is a best response for k members. Hence b satisfies

$$\left(1 - \frac{b}{n}\right)\theta_k + \delta_{kx} \leq \frac{b}{n}\theta_k + \delta_{kx'}.$$

Thus the cost of such a path from z_x to $\mathcal{M}$ is at least

$$\min_{k\in\{A,B\}} \left\lceil \left(1 - \frac{\delta_{kx'} - \delta_{kx}}{\theta_k}\right)\frac{n}{2} \right\rceil. \tag{19}$$

This is greater than (18) because $d_k > \delta_{kx'}$ by misalignment. Hence the minimum cost path from z_x to $\mathcal{M}$ is direct and $r(z_x, \mathcal{M})$ equals (18).

Therefore, when errors are unconstrained and n is sufficiently large the following statements are equivalent:

$$\begin{aligned} r(\mathcal{M}, z_x) &> r(z_x, \mathcal{M}), \\ \min_{k\in\{A,B\}} D_{kx} &> \min_{k\in\{A,B\}} \tfrac{1}{2}(1 - D_{kx}), \\ 2\min_{k\in\{A,B\}} D_{kx} + \max_{k\in\{A,B\}} D_{kx} &> 1. \end{aligned} \tag{20}$$

The last line holds by hypothesis (9). □

The next result follows immediately from Lemmas 1 and 4.

Corollary 2 *If (9) is satisfied, then for n sufficiently large*

$$r(z_x, \mathcal{M}) < r(z_{x'}, z_x)$$

for all distinct pairs $(x, x') \in (X_A \cap X_B)^2$.

Lemma 5 $r(z_x, \mathcal{M}) \leq r(z_x, z_{x'})$ *for all distinct* $(x, x') \in (X_A \cap X_B)^2$.

Recall that $r(z_x, \mathcal{M})$ equals (18).

Now consider $r(z_x, z_{x'})$. On the minimum cost path from z_x to $z_{x'}$ one of two things must occur. Starting from z_x, through a series of errors, either (i) a mutually impermissible action becomes a best response, in which case $r(z_x, z_{x'}) \geq r(z_x, \mathcal{M})$, or (ii) an action $x'' \in X_A \cap X_B$ (possibly x') becomes a best response. The path with minimum cost that conforms to the case (ii) involves β erroneous plays of x'' by k' members, where β equals (19).

We established in the proof of Lemma 4 that (19) is greater than (18). This establishes the Lemma. □

Proof of Proposition 4. To prove that (9) is sufficient for $\mathcal{M}$ to be stochastically stable we use the same tree-surgery argument as in Proposition 2(i). Starting with an arbitrary x-tree, we transform it through steps 1-2 into an $\mathcal{M}$-tree. Again the difference in resistance between the original x-tree and the resultant $\mathcal{M}$-tree is given by (13). If (9) holds, then the first term is positive by Lemma 4 and the second term, if it exists, is positive by Corollary 2. Hence (9) is sufficient.

To establish that (9) is also necessary, note that Lemma 5 implies that the minimum cost $\mathcal{M}$-tree is composed entirely of direct links, i.e. edges $(z_x, \mathcal{M})$ for each $x \in X_A \cap X_B$.

Suppose that (9) were violated, so that $2\min_{k\in\{A,B\}} D_{kx} + \max_{k\in\{A,B\}} D_{kx} \leq 1$ for some $x \in X_A \cap X_B$. Recall that the minimum cost $\mathcal{M}$-tree contains edge $(z_x, \mathcal{M})$. Replace this edge with $(\mathcal{M}, z_x)$. This operation transforms the minimum cost $\mathcal{M}$-tree into an x-tree. The difference in resistance between the original $\mathcal{M}$-tree and the resultant x-tree is $r(z_x, \mathcal{M}) - r(\mathcal{M}, z_x)$ which is non-negative as (9) is violated [see (20)]. Hence $\mathcal{M}$ is not the unique stochastically stable class. □