SCISPACE
formerly Typeset

Open access • Journal Article • DOI:10.1111/J.1467-9779.2012.01557.X

Reputation, Social Identity and Social Conflict — Source link

John Smith

Institutions: Rutgers University

Published on: 01 Aug 2012 - Journal of Public Economic Theory (New Brunswick, NJ: Dep. of Economics, Rutgers, the State Univ. of New Jersey)

Topics: Social group, Social identity theory, Social identity approach, Social conflict and Population

Related papers:

- Reputation, Social Identity and Social Conflict
- Managing diversity by creating team identity
- Decisiveness, peace, and inequality in games of conflict
- Reasons for Conflict : Lessons from Bargaining Experiments
- Social identity, competition, and finance: a laboratory experiment

Share this paper:

View more about this paper here: https://typeset.io/papers/reputation-social-identity-and-social-conflict-55jkzq1ffs

Reputation, Social Identity and Social Conflict[*]

John Smith[†]
Rutgers University-Camden

October 11, 2007

Abstract

We interpret the psychology literature on social identity and examine its implications in a population partially composed of such agents. We model a population of agents from two exogenous and well defined social groups. Agents are randomly matched to play a reduced form bargaining game. We show that this struggle for resources drives a conflict through the rational destruction of surplus. We assume that the population contains both rational players and behavioral players. Behavioral players aggressively discriminate against members of the other social group. The existence and specification of the behavioral player is motivated by the social identity literature. For rational players, group membership has no payoff relevant consequences. We show that rational players can contribute to the conflict by aggressively discriminating and that this behavior is consistent with existing empirical evidence. Our paper relates to the empirical literature which finds that our measure of social heterogeneity tends to be increasing in economic variables which we interpret as signifying inefficiency. We provide an explanation that, as social groups compete for the benefits of public goods, disagreement and inefficiency can result. Our work also relates to the social conflict literature, which examines the relationship between macro level factors such as unemployment and civil disturbances. This literature finds that the amount of social conflict tends to be increasing in what we refer to as the inequitability of the environment.

[*]The author would like to acknowledge helpful comments from Roland Benabou, Armin Falk, Faruk Gul, Jo Hertel, Wolfgang Pesendorfer, Jack Worrall and the participants of the Social Identity Theory Seminar in the Princeton Psychology Department organized by Debbie Prentice.

[†]Email: smithj@camden.rutgers.edu; Phone: (856) 225-6319.

1 Introduction

We claim that every population consisting of members of different social groups, will also contain some individuals with a preference for discrimination: favoring members of their own group at the expense of members of other groups. Indeed, this is the primary insight of the vast psychology literature on social identity, which we describe below. In this paper, we insert agents who behave as described by this literature into a heterogeneous population of agents. The interesting questions are then, what can we say about agents with no such preference for discrimination and what can we say about outcomes in such a society.

We present a model in which each player lives for two periods and in each period is matched to play a reduced form bargaining stage game. In each stage game, both players have a better material outcome by agreeing to a distribution than not agreeing. Also, in the stage game, each player has a better material outcome by securing the larger share of the surplus. We assume that every agent is a member of one of two social groups and that this status is observable.

Players are assumed to be of the following two types: rational and behavioral. Rational players are motivated entirely by material payoffs. Group membership contains no payoff relevant consequences for rational players. A behavioral player is an agent who has payoffs significantly different than rational players. Consistent with the social identity literature, we make the following assumptions regarding behavioral players. When matched with a member of their own group (an ingroup match), behavioral players are rational. When matched with a member of the other group (an outgroup match) behavioral players intransigently play the action which attains the largest difference between his own payoff and the payoff of his opponent. The important implication being that the behavioral players will destroy surplus rather than accept a payoff lower than the outgroup opponent.

The use of behavioral players is a standard technique in game theory, which was pioneered by Kreps and Wilson (1982). The novelty in our approach lies in merging this concept with our interpretation of the Social Identity literature. Roughly, psychologists find that the division of people into social groups is a sufficient condition for some to discriminate against those in a different group. Therefore, we view this social identity literature as providing justification for the existence and specification of our behavioral players.

What emerges is a conflict between agents over payoffs. We show that a social conflict need not require an entire population of agents with a preference for aggressive discrimination. Rather, we show that rational players contribute to the conflict through the destruction of surplus and it is the struggle for resources which drives the inefficiency arising from the conflict.

Our first main result (Proposition 3) shows that the inefficiency in a society tends to be increasing in our measure of heterogeneity of that society. Therefore, we expect more heterogenous societies to exhibit more social conflict. We also vary the bargaining specifics and we define less equitable contexts as an inequitable environment. Our second main result (Proposition 4) shows that inefficiency is increasing in the inequitability of the environment. These results relate to the following two strands of literature.

Researchers have examined the relationship between social heterogeneity and economic conditions. For instance, Mauro (1995) demonstrates that ethnic heterogeneity is related to poor government institutions and that poor government institutions are related to poor economic growth.[1] Moltalvo and Reynal-Querol (2005) show that measures of heterogenous populations are negatively related to economic development. Of specific interest here, Alesina, Baqir, and Easterly (1999) show that spending on productive public goods, such as education and roads, is inversely related to the ethnic heterogeneity in U.S. communities. We contend that our model provides an account for this finding. We argue that productive public goods differentially benefit various social groups. As individuals of different social groups compete for the larger share of these benefits, disagreement can result. If, as a consequence of the disagreement, the project is not undertaken, then we interpret this outcome as inefficient. We demonstrate the positive relationship between our measure of social heterogeneity and social conflict as measured by such inefficiency.

Additionally, researchers have noted the relationship between the level of social conflict and the inequitability of the environment. Falk and Zweimuller (2005) show a relationship between local economic conditions and aggressive behavior. Specifically, the authors show that higher local unemployment rates (and hence, larger probabilities of inequitable outcomes) lead to higher incidences of right-wing extremist crimes. It is important to note that the authors find that it is the threat of a worse economic position, and not the economic position per se, which induces this conflict. Therefore, we interpret these findings as evidence of a positive relationship between the inequitability of the environment[2] and social conflict. There is also a large sociological literature relating various forms of social conflict to the inequitability of the environment. For instance, Olzak (1992) examines US data between 1877 and 1914 and finds a positive relationship between the inequitability of the environment and ethnic conflict, as measured by violent events.[3] Our model also provides an explanation for these findings. Specifically we show that the amount of social conflict is increasing in the inequitability of the environment.

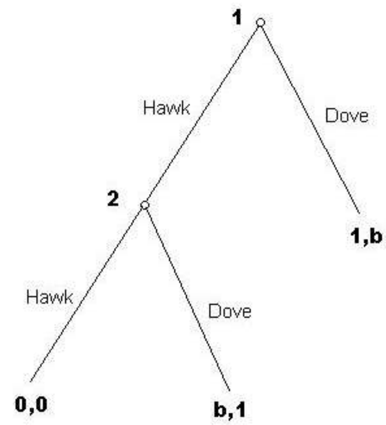
Although we use the same formalism throughout the paper, we wish to address both of the previously discussed literatures. Therefore, we offer two interpretations of the model: the legislator interpretation and the citizen interpretation. In the *legislator interpretation* we assume there is a population of legislators representing either green or red people. In each

[1] Also see Easterly and Levine (1997).

[2] What we refer to as "inequitability of the environment" sociologists refer to as "competition." Sociologists define competition to be the threat of a worse economic position. Here, we believe this term to be inappropriate as "competition" has a different meaning to economists.

[3] Lubbers and Scheepers (2001), Scheepers et. al. (2002), Quillian (1995, 1996) also find a positive relationship between the inequitability of the environment and social conflict, as measured by prejudiced beliefs. Olzak, Shanahan, and West (1994) find the relationship in the context of school busing in U.S. cities.

of two periods, two legislators are randomly matched in order to agree on the construction of a public good, say a university. We assume that each legislator gets a payoff of $b > 1$ if the university is built in his district, a payoff of 1 if the university is built in a different district and a payoff of 0 if the university is not built due to disagreement about its location. Such a bargaining situation can be represented by the following extensive form game:[4]



The action Hawk (H) is interpreted as demanding that the university will be built in the legislator's own district and the action Dove (D) is interpreted as conceding that the university can be built in the other district. Further, suppose that in the second period, each legislator can perfectly observe the first period behavior of the current opponent.

Suppose that every agent has the payoffs as described above. Then the unique subgame perfect outcome is for each player 1 to play H and each player 2 to play D in both repetitions of the stage game. Now suppose that there is some fraction of legislators who, when matched with a legislator of a different color, will always demand that the university be built in his district. However, if the matched opponent is the same color, these legislators have payoffs as described above. In this case, a rational legislator might destroy surplus in order to obtain a reputation for having a preference for aggressive discrimination. In other words, it could be optimal for the legislator to destroy surplus in the first period, in order to have the second period opponent believe the agent is likely to be behavioral. Under this legislator interpretation, we show that the amount of inefficiency, caused by disagreement over the allocation of the public good tends to be increasing in social heterogeneity.

We define $\theta > 0.5$ to be fraction of the society in the majority group and γ to be the fraction of behavioral players in each group. We define and apply the concept of a Symmetric Perfect Bayesian Equilibrium ($SPBE$). An $SPBE$ requires a Perfect Bayesian Equilibrium for every player in the game, that beliefs are updated in a reasonable manner and that every group member has identical strategies.

[4]The actions Hawk and Dove have been named as such because the stage game is an extensive form chicken game.

Next, we define a strongly symmetric game to be one in which the first period $SPBE$ strategies can be written without reference to group membership. In Proposition 1 we show that in every strongly symmetric game, the majority players have a higher ex-ante payoff than the minority players. In Proposition 2 we show that in every $SPBE$, minority players behave at least as aggressively as the majority players.

The reader more interested in the relationship between social conflict and social heterogeneity or the relationship between social conflict and inequitable environments can skip to section 4. As θ is the fraction of the majority group, we use $1 - \theta$ is a measure of the heterogeneity of society. Proposition 3 demonstrates that the inefficiencies in a society tend to be increasing in heterogeneity. This relationship between inefficiency and heterogeneity is consistent with the results in the empirical literature. Again, the interpretation we give to this result is as follows: public goods will differentially accrue benefits to members of social groups. A population of legislators matched in order to agree on an allocation of public goods, in the presence of behavioral legislators who intransigently play H in an outgroup match, will exhibit the observed relationship between inefficiency and heterogeneity.

Our final main result relates the inequitability of the environment to social conflict. We define b to be the inequitability of the environment. This should be uncontrovertial, as the difference between winning and losing is increasing in b. For this result, we refer to the second interpretation of the model: the *citizen interpretation*. The agents are interpreted as citizens of a society bargaining with other citizens for material payoffs. The citizens are to split a pie of size $b + 1$, into portions of $b > 1$ and 1. The action H represents demanding b and the action D represents accepting the payoff of 1. If agreement cannot be reached then both players get 0. Proposition 4 illustrates the positive relationship between social conflict and the inequitability of the environment.

Our specification of the behavioral players is motivated by the social identity literature. A very large and venerable psychology literature has found that placing people into groups is a sufficient condition for discriminating behavior.[5] Of particular importance is the finding that people tend to prefer better material outcomes for ingroup members than outgroup members and that they are also prepared to create inefficiencies (destroy surplus) to secure this outcome. For instance, the discriminating person would prefer to allocate \$6 to an ingroup member and \$2 to an outgroup member rather than \$5 to each. Tajfel et. al. (1971) find that these preferences imply the maximization of the payoff difference between the groups.[6] In other words, the discriminating person will accept some inefficiency in allocating resources in order to secure a better material outcome for the ingroup.

[5] A very small sample of this enormous literature would include Sumner 1906, Murdock (1949), Sherif et. al. (1961), Tajfel (1970), Tajfel et. al. (1971), Tajfel (1978), Tajfel and Turner (1979), Kramer and Brewer (1984), Tajfel and Turner (1986), Dawes, Van De Kragt, and Orbell (1988).

[6] There is, however, no consensus on this statement. Messick and Mackie (1984 pg. 64) point out that some authors find that discrimination can come in the form that the joint allocation is maximized "as long as the ingroup gets more than the outgroup." This perspective also suffices to justify our specification of behavioral players.

We view the social identity literature as providing specific justification for our model. The first assumption justified involves the formation of social groups based on some shared characteristic and that such social groups provide a basis upon which *some* will condition their behavior. In other words, we assume that the shared characteristic implies the existence of social groups and that this possibly superfluous identity can affect the behavior of some, but not all. The second justified assumption is our modeling choice that all players are "rational" in an ingroup match and in an outgroup match, some players will pick the action which maximizes the difference between the groups. In our setting, this means that behavioral players intransigently play action H in an outgroup match. The condition that some people prefer ingroup members to have better outcomes than outgroup members does not have bite in our ingroup matches. Therefore, we assume that behavioral players are rational in ingroup matches.

1.1 Related Literature

Discrimination poses a difficult problem to economists as discriminating behavior is usually suboptimal in material payoffs. It therefore follows that there should be no discrimination, as such agents will be driven from the market. To address this concern, economists have modeled discrimination by identifying particular situations which cause agents to act in a discriminatory manner. In this literature, no agent has a taste for discrimination but due to the specifics of the situation (for instance the information structure) the agent acts as if he possesses a preference for discrimination. This literature is referred to as statistical discrimination.[7]

A primary distinction between the statistical discrimination literature and the present paper is that here it is the strategic interaction, rather than the details of the information structure, which drives the discriminating behavior of rational players. In other words, here we acknowledge the fact that some people have a preference for discrimination and we examine its social implications.

Akerlof and Kranton (2000) present a general model of identity and economics. Akerlof and Kranton assume that an agent's identity related preferences are affected by actions of others. In other words, a member of a group has a behavioral prescription for certain actions. Thus, not performing the prescribed action will result in a loss of identity, which the agent prefers to avoid. The authors present this notion as a basis for group membership, therefore their notion of a social group is fluid. By contrast, we model a social conflict between well defined social groups which are not fluid and not defined by behavior. A primary application of Akerlof and Kranton is the explanation of the harmful behavior of members of less advantaged social groups. Similar to Akerlof and Kranton, the behavior in our model is optimal from the perspective of the agent. However, the behavior in both models can be suboptimal in other

[7] See Phelps (1972) and Arrow (1973) for the originators of this literature.

ways: in our model discrimination leads to inefficiencies and in Akerlof and Kranton agents can engage in destructive activities.[8]

There exists a literature which formally models social conflict, however each strand focuses on different issues than we do here. For instance, Fearon and Laitin (1996) and Nakao (2007) focus on the role in which ingroup policing helps to maintain social order by avoiding social conflict between groups. Specifically, it is assumed that information is differentially better for the histories of ingroup members than outgroup members and that no agents have a preference for discrimination. By contrast, we examine the implications of the preference for discrimination

Like Basu (2005), we model social conflict in a heterogenous society containing some members with a preference for discrimination. Additionally, we both show how the presence of these types can induce those without such a preference to discriminate. Unlike Basu, where the presence of special types of agents induces a more defensive posture in other agents, in our paper the resulting behavior is a more aggressive posture. In other words, the inefficiencies in Basu are driven by fear of aggressive behavior of the opponent. By contrast, inefficiencies in our model are driven by aggressive behavior of rational agents induced by material gains.

Rapoport and Weiss (2003) present a model in which group conflict arises due to the effects the groups have on market conditions. Specifically, the majority will resent the minority as the latter might not participate in, and thereby not contribute to the reduction of the inefficiency of, the market. The authors are not concerned with the conditions under which the "nice" outcomes occur in outgroup matches, but are rather concerned with the effect of the nice outcomes occurring in ingroup matches will have on the market. By contrast, in our model, the ingroup matches are always nice and we focus on the nature of outgroup matches.

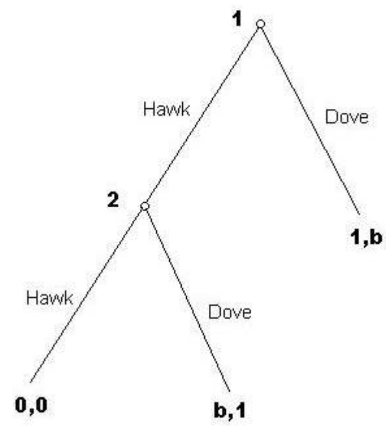
Our paper also relates to the work of Esteban and Ray (1994). These authors provide an axiomatization relating the amount of polarization (and hence potential for conflict) in a society to the distribution of characteristics of individuals in that society. Although the authors accommodate a more rich profile of characteristics than considered here, we focus on the individual behavior which might yield such a conflict.

Finally note that we share a similar methodology to Silverman (2004). Like this paper, Silverman uses matching in a two-sided reputation model to explore suboptimal outcomes not generated by a model with complete rationality.

2 The Model

We study a sequential stage game repeated for $T = 2$ periods. The stage game payoffs are described by the following game tree $\mathcal{T}$:

[8]For more on identity in economics, see Sobel (2004), Kirman and Teschl (2004), Davis (2005) and Shayo (2004).



where b is strictly larger than one.[9] In each repetition of the stage game, player 1 chooses an action of either Hawk (H) or Dove (D). In the event that player 1 selects H, player 2 chooses between H and D. We do not allow transfers between agents.

There is a continuum of players $i \in [0,1]$. Each player is a member of exactly one of two social groups. This group identity is described by the social identity parameter $\theta \in (0.5, 1)$. All agents such that $i \in [0,\theta] = M$ are in the majority group and all agents such that $j \in (\theta, 1] = m$ are in the minority group. In each period, agents are matched to play the stage game where the matching probability is uniform on the population. In each match, the probability of being a player 1 is identical to that of being a player 2. We assume that each player is matched only with players of the same age. If two players i, j such that $i \in M$ and $j \in m$ are matched, we refer to this as an outgroup match; otherwise it is an ingroup match.

In each group, there are two types of players: rational and behavioral. The rational players have their payoffs described by $\mathcal{T}$. Behavioral players always play H in an outgroup match and have payoffs as described by $\mathcal{T}$. The ex-ante fraction of behavioral players, in each group, is γ. The entire game Γ is therefore described by $\Gamma = (\mathcal{T},\, b,\, \theta,\, \gamma)$.

To simplify the subsequent analysis, note that in every ingroup match the subgame perfect equilibrium of the stage game is played: player 1 plays H and the player 2 plays D. No player has an incentive to deviate. Player 2 gains no future benefit by playing H. Player 1, knowing this, plays H. Therefore, we take the ingroup matches as given and focus exclusively on the behavior in the outgroup matches.

Player i's action is denoted $a \in \{H, D\} = A$. We define the condition of the match as $c \in \{1, H\} = C$. Here $c = 1$ indicates that i is a player 1. Likewise, $c = H$ indicates that i is a player 2 whose opponent played H. The history of the matched opponent is

[9] All of the following would hold if instead we exchanged b and 1 with x and $1-x$ respectively where $x = \frac{b}{b+1} > \frac{1}{2}$.

perfectly observed. We can write the relevant set of histories for player i in the first period as $h^i \in \mathcal{H}^i = \{I, H1, D1, HH, DH, E\}$. The first element refers to an ingroup match. The following two elements refer to playing H and D as a player 1. Likewise the next two refer to playing H and D as a player 2 against a player 1 who played H. The last element refers to a player 2 matched against a player 1 who played D. We define the set of player histories $\mathcal{H}_D$ in which the action of D has been observed in an outgroup match

$$\mathcal{H}_D = \{D1, DH\}$$

A first period strategy for player i is a mapping $\sigma_1^i : C \to \Delta A$ and the second period strategy as a player 1 is a mapping $\sigma_2^i : C \times \mathcal{H}^j \to \Delta A$. We define $\sigma^i = \sigma_1^i \times \sigma_2^i$. We also define $\sigma = \times_{i\in[0,1]}\sigma^i$. We denote $\sigma^i(\cdot)$ as the probability that H is played. After a history of h^i the posterior belief that player i is behavioral is denoted $p^i(h^i)$. Players maximize the sum of expected utility payoffs. We assume no discounting. In period 2, for a given history h_1^j and condition c, player i's expected payoff from the profile of strategies is defined to be $U_2^i(\sigma|c, h^j)$. In period 1, for a given c, player i's expected payoff from the profile of strategies in periods 1 and 2 is defined to be $U_1^i(\sigma|c)$.

The reader should recall that we intend to model a general conflict situation with as few asymmetries as possible. Specifically, we designed the model in such a way that the groups are as meaningless as possible. As such, we have assumed that each group has an identical fraction of behavioral players ($\gamma = \gamma_M = \gamma_m$). We have also assumed that the probabilities that an agent is designated as a player 1 and player 2 are equal for agents in each group. Despite these symmetry assumptions, we still observe the inefficiencies associated with a social conflict. Indeed our assumptions regarding γ are weaker than warranted by recent experimental evidence. For instance, Cho and Connelley (2002) find that the competitiveness of an outgroup setting is associated with a higher degree of identification of subjects. We interpret this as finding evidence of a positive relationship between γ and b. Although we do not assume such a relationship, if we did then the results in the paper would be stronger.

In our solution concept, we use the following definition:

Definition 1 *Beliefs $p^j(h^j)$ satisfy condition $(*)$ if $h^j \in \mathcal{H}_D$ then $p^j(h^j) = 0$.*

Condition $(*)$ requires beliefs to be updated in an intuitive manner. On or off-the-equilibrium path, it requires that if player j ever played D in an outgroup match, opponents ascribe probability 0 to j being behavioral.

Now we define the notion of equilibrium which we will use throughout the paper.

Definition 2 *A strategy profile σ is a Symmetric Perfect Bayesian Equilibrium (SPBE) if:*

$$(i)\ \ U_1^i(\sigma|c) \geq U_1^i(\widetilde{\sigma}^i, \sigma^{-i}|c) \text{ for every } i,\ \widetilde{\sigma}^i \neq \sigma^i \text{ and } c \in \{1, H\}$$

$$(ii)\ \ U_2^i(\sigma|c, h^j) \geq U_2^i(\widetilde{\sigma}^i, \sigma^{-i}|c, h^j) \text{ for every } i,\ \widetilde{\sigma}^i \neq \sigma^i,\ c \in \{1, H\} \text{ and } h^j \in \mathcal{H}^j$$

$$(iii)\ \ \text{for any } i, k \in M \text{ and any } j, l \in m,\ \sigma^i = \sigma^k \text{ and } \sigma^j = \sigma^l$$

Furthermore, beliefs $p^j(h^j)$ must satisfy condition $(*)$ and are updated using Bayes Rule wherever possible, for all j and $h \in \mathcal{H}$.

Definition 2 is a slightly more restrictive version of a Perfect Bayesian Equilibrium (PBE). Condition (i) requires that period 1 actions are optimal, as both a player 1 and 2, given any set of initial beliefs. Condition (ii) is the analogous requirement for period 2. Condition (iii) requires that every member of a group use the same strategy. Note that in equilibrium, this requirement only bites when players are indifferent between actions. In such a case, condition (iii) allows us to break ties in a manner consistent with a social identity interpretation. Condition (iii) also allows us to refer to strategies for the group rather than for the individual. For instance, $\sigma_1^M(1)$ refers to the strategy of the majority group as a first period player 1. Finally, we require that beliefs are updated using Bayes Rule, wherever possible and that a player, who selected D in the first period, is known with certainty to be rational.

We note Lemma 2, found in Appendix A, which shows that, all rationals always play D as a second player in period 2 ($\sigma_2^i(H, h^j) = 0$ for all $h^j \in \mathcal{H}$ and $i \in \{m, M\}$). As there is no confusion, we write $\sigma_2^i(1, h^j)$ as $\sigma_2^i(h^j)$ in order to conserve notation. It also turns out that the $SPBE$ is generically unique which Corollary 4, found in the Appendix B, demonstrates.

We speak of *aggressive discrimination* whenever the actions (H, H) are observed. This terminology is appropriate as the outcome (H, H) never occurs in equilibrium in an ingroup match. More generally we refer to a play of H (in any period) as aggressive play.

Again, note that in a game without behavioral players ($\gamma = 0$), the unique subgame perfect equilibrium is to play H as a player 1 and to play D against an H as a player 2. There are, however, conditions under which a rational player will optimally destroy surplus in order to secure a reputation for being a behavioral player. In our perturbed game ($\gamma > 0$), this rational destruction of surplus can take one of the two following forms:

Definition 3 *Agent i exhibits Reputation as a Player 2 ($P2$) if the $SPBE$ is such that:*

$$\sigma_1^i(H) > 0$$

If player i exhibits $P2$, he will play H with positive probability in response to a player 1 selecting H, even though playing H means forgoing a certain payoff of 1 in order to have more favorable future matches. However, another type of reputation can be observed when the agent is a player 1:

Definition 4 *Agent i exhibits Reputation as a Player 1 ($P1$) if the $SPBE$ is such that:*

$$\sigma_1^i(1) > 0$$
$$(1 - \gamma)(1 - \sigma_1^j(H))b < 1$$

If player i exhibits $P1$, he will play H as a player 1 with positive probability, even though it is sufficiently likely that the opponent will play H. In order to compare the two definitions, note that if an agent displays $P2$ then the player exchanges a first period stage game payoff of 1 for a payoff of 0. However, a player 1 selecting H could be myopically optimal if the matched opponent is sufficiently likely to play D. In this case, we cannot claim that the player is motivated by reputation concerns. Therefore, we require the second condition so that the stage game play is not myopically optimal.

The following proposition states that $P1$ and $P2$ will never both occur in any $SPBE$.

Lemma 1 *If one player exhibits $P1$ ($P2$) then there are no parameter values for which any player exhibits $P2$ ($P1$).*

Proof: See Appendix A.

Intuitively, if a player exhibits $P1$ then it is sufficiently likely that the current opponent is a behavioral player. The smallest such probability makes the exhibition of $P2$ by any player unprofitable. Similarly, if a player exhibits $P2$ then it is sufficiently likely that a future opponent is a rational player, otherwise $P2$ would not be profitable. However the smallest such probability renders playing H as a player 1 myopically optimal, thus the agent cannot exhibit $P1$.

3 Symmetric Perfect Bayesian Equilibrium

To begin this section, we provide the following example which illustrates the intuition for the model. While we vary b, we assume specific values for θ and γ. In the first case ($b = 3$) neither group displays $P2$, in the second case ($b = 5$) only the minority displays $P2$ and in the final case ($b = 7$) both groups display $P2$.

Example 1 *Consider an $SPBE$ where the majority group composes 60% of the population ($\theta = 0.6$), each group contains 10% behavioral players ($\gamma = 0.1$), and the prize b is either 3, 5, or 7:*

(i) *In the case that $b = 3$, the $SPBE$ strategies look similar to that of the unperturbed game.*[10] *The only difference being that those matched with a player who played H as a player 2 in the first period will play D as a player 1. The $SPBE$ strategies are:*

$$\begin{aligned}
\sigma_1^i(1) &= 1 \text{ and } \sigma_1^i(H) = 0 \text{ for } i \in \{m, M\} \\
\sigma_2^i(1, h^j) &= 0 \text{ if } h^j = HH \text{ for } i \in \{m, M\} \\
\sigma_2^i(1, h^j) &= 1 \text{ if } h^j \neq HH \text{ for } i \in \{m, M\}
\end{aligned}$$

[10] The interested reader is referred to Appendix B for Propositions 5 (i), 6 and 7 respectively for the proofs of the strategies given in parts (i), (ii) and (iii) of the example.

When $b < \frac{2}{\theta(1-\gamma)} + 1 \approx 4.7$ *(and thus* $b < \frac{2}{(1-\theta)(1-\gamma)} + 1 \approx 6.6$*) the minority (majority) has no incentive to deviate from* $\sigma_1^i(H) = 0$*. Here, in both majority and minority groups, only behavioral players destroy surplus. Also note that the* $SPBE$ *strategies are identical for majority and minority and can be written without regard to group membership. In what follows, we define such a game as strongly symmetric.*

(ii) In the case that $b = 5$ *the incentives (and therefore first period strategies) are identical to the* $b = 3$ *case for* M*, but not for* m*. Here* $\sigma_1^m(H) = 0$ *cannot be part of* $SPBE$*. However it also cannot be that* $\sigma_1^m(H) = 1$ *because this would imply* $p^m(HH) = \gamma$ *and thus* $\sigma_2^M(HH) = 0$ *for* M *as* $\gamma < \frac{b-1}{b}$*. Therefore* $\sigma_1^m(H)$ *must be such that* $p^m(HH) = \frac{b-1}{b} = \frac{4}{5}$*. This is the posterior which makes the agent as a player* 1 *indifferent between* H *and* D*. This mixing probability occurs at* $\sigma_1^m(H) = \frac{\gamma}{(1-\gamma)(b-1)} = 0.028$*.*

(iii) In the case that $b = 7$*, both* m *and* M *will mix such that* $p^i(h) = \frac{6}{7}$*. This mixing probability occurs at* $\sigma_1^i(H) = 0.0185$*. Similarly both groups must mix as a second period player* 2 *in order to keep the first period player* 2 *indifferent between playing* H *and* D *against an* H*.*■

In the balance of this section we characterize some basic properties of the $SPBE$. We illustrate the underlying asymmetry in payoffs by showing that the majority always does strictly better for parameter values such that both groups have identical equilibrium strategies. We also show that reputation is always more valuable for the minority players. Hence, we find that minority players will always exhibit weakly more aggressive behavior in the first period, than do majority players.[11]

Although the $SPBE$ is generically unique, depending on the particular parameters of the game, the equilibrium can have significantly different properties. For some parameter values, $SPBE$ strategies and therefore equilibrium payoffs can exhibit some asymmetry. However, there is also a basic asymmetry inherent in our model, which is best illustrated when attention is restricted to strongly symmetric strategies - that is, first period strategy profiles which are identical across groups. This motivates the following definition:

Definition 5 *Let* σ^Γ *be the* $SPBE$ *of* Γ*. Then* Γ *is strongly symmetric if the first period strategies in* σ^Γ *can be written without reference to group membership.*

We say that a game is strongly symmetric if its parameters are such that all players have identical equilibrium strategies. However, even in such a markedly symmetric environment, the majority does strictly better than the minority, as the next result shows.

Proposition 1 *If* Γ *is strongly symmetric, the majority has a strictly higher ex-ante payoff than the minority.*

[11] As this paper proposes a general model of social conflict, the only assumed asymmetry involves the probability of an outgroup match. The following results crucially depend on this symmetry. In modeling a particular situation, where the symmetry assumptions are not justified, a modified version of our model will suffice.

Proof: *See Appendix C.*

This result follows from the fact that majority group members are more likely to be in an ingroup match than minority group members. If Γ is strongly symmetric, an ingroup match is more profitable than an outgroup match. Additionally, the posteriors for a given history are identical across groups which implies that second period strategies are also identical. These facts combine to produce the result.

Note that this result crucially depends on the existence of the behavioral players ($\gamma > 0$). In the unperturbed game, members of both groups have an expected payoff of $b+1$. Therefore if there are no types with a preference for discrimination, then we observe no payoff differences based on group membership.

As Proposition 1 states for strongly symmetric Γ, the majority always does better than the minority. However, the majority can do worse if the equilibrium strategies across groups are sufficiently asymmetric. We now present an example of such an $SPBE$ where the minority has a larger expected payoff than the majority.

Example 2 *Suppose that $\theta = 0.6$, $b = 2$, and $\gamma = 0.55$. The $SPBE$ which corresponds to these parameter values is described by Proposition 5 (iii) in Appendix B. In this $SPBE$ the minority displays $P1$ and the majority does not. Therefore the $SPBE$ is not strongly symmetric. If we let E^i represent the ex-ante payoff of player i, then it follows that:*[12]

$$E^m = 2.825 > E^M = 2.687$$

■

The above example demonstrates the importance of the strong symmetry assumption in Proposition 1. The intuition behind Example 2 is that the majority does not obtain a reputation while the minority does. Hence, the minority does sufficiently better than the majority in outgroup matches and so the minority does better overall.

In Example 2, the minority displays more aggressive behavior in the first period than does the majority. This is a general feature of the $SPBE$, as we show in the next proposition. We show that the minority is always at least as likely as the majority to play H as a first period player 1 and player 2.

Proposition 2 *In every generic $SPBE$, a minority member m plays at least as aggressively as a majority member M :*

$$\sigma_1^M(1) \leq \sigma_1^m(1) \text{ and } \sigma_1^M(H) \leq \sigma_1^m(H).$$

[12] The interested reader is referred to Appendix C for a more complete treatment of this example.

Proof: *See appendix C.*

The intuition behind Proposition 2 is that reputation is more valuable to the minority player than the majority player, as the minority player is more likely to be in a second period outgroup match. Note that in our model we assume very little asymmetry between the groups. We assume uniform matching, an equal probability of being a player 1 and 2 in each period for both groups, and an equal fraction of behavioral players in each group. The only assumed asymmetry relates to the composition of society. One could imagine a situation where these symmetry assumptions are not appropriate. However, the purpose of this paper is to investigate social outcomes when assuming as little between group asymmetry as possible. Therefore, we do not explore these issues.

We interpret Proposition 2 to be consistent with psychology literature related to the group identity of majorities and minorities. Psychologists find that minorities have a stronger group identity than do majorities.[13] As a result of this stronger identity, we expect stronger behavior; and in the context of our model, stronger behavior means more aggressive play.

4 Comparative Statics: Social Fragmentation and Inequitable Environments

In this section we present our main results. We apply our model to situations which have been subject to empirical investigation. We examine the relationship between social conflict (as measured by inefficiency) and social heterogeneity. We also examine the relationship between social conflict, as measured by inefficiency, and inequitable environments. In the examination of these relationships we reference, respectively, the legislator interpretation and citizen interpretation as defined in the introduction.

Recall the discussion of the empirical relationship between inefficient outcomes in a society and social heterogeneity in that society. Alesina et. al. (1999) offers a theoretical explanation for these findings. The authors present a class of preferences which are decreasing in the distance between the agent's optimal amount of the public good and that actually chosen. These preferences are used to show that an increasing diversity of preferences will lead to smaller investment in the productive public good. By contrast, we explicitly model the schisms which lead to the conflict: competition over scarce resources, in the presence of behavioral agents, leads some rationals to aggressively discriminate. Our interpretation is not unique as, for instance, Vigdor (2002) points out that results such as Alesina et. al. (1999) can be explained by "differential altruism."

Many authors use fragmentation as a measure of social heterogeneity. Fragmentation is defined as the probability that two randomly selected people are from a different group. Here, this would mean that fragmentation is $2\theta(1-\theta)$. By contrast we use $1-\theta$ as a measure of

[13] See Gurin et. al. (1999).

social heterogeneity. Our measure is appropriate as both functions are maximized on $[0, 0.5]$ at $\theta = 0.5$ and are strictly decreasing in θ. Furthermore, nothing is gained by considering the more complicated measure of heterogeneity. Therefore, we use $1 - \theta$ as a measure of heterogeneity.

To formally state this result, we first define the total efficiency loss in the $SPBE$ as $\mathcal{I}(b, \theta)$. This quantity is the probability of (H, H) outcomes in either period multiplied by the total material surplus which could have been achieved in any other outcome, $b + 1$. We state $\mathcal{I}(b, \theta)$ as explicitly depending on θ and b but not on γ (fraction of behavioral players), as we will shortly explore the implications of varying the first two but not the last parameter. Furthermore, γ is hard to measure and to our knowledge, no empirical papers have studied the matter.

Definition 6 *$\mathcal{I}(\theta, b)$ is the total efficiency loss in the $SPBE$:*

$$\mathcal{I}(b, \theta) = (b + 1) \left[P((H, H) \ in \ t = 1) + P((H, H) \ in \ t = 2) \right],$$

Although the definition of $\mathcal{I}(b, \theta)$ is rather straightforward, its computation can be quite involved. The interested reader can turn to Expression (9) in Appendix C for more about $\mathcal{I}(b, \theta)$.

Note that $\mathcal{I}$ is not a measure of social welfare. Specifically, $\mathcal{I}$ is not the average of the utilities of the agents in the game. The value of $\mathcal{I}$ is intended to provide a measure material payoffs not captured through the bargaining between agents. We feel that this is the most appropriate criteria to consider. While it is often assumed that a social planner seeks to maximize the utility of every agent, with standard assumptions regarding utility, this condition is equivalent to maximizing the material surplus of each agent. However, in our case, these two notions are not identical. Indeed, to be consistent with the spirit of the social planner, would seek to maximize the volume of the trade rather than accommodating the discriminatory preferences of the behavioral players. The value of $\mathcal{I}$ provides a measure of the material outcomes in a population and we consider this to be the most appropriate criteria.

The next result shows that there exists a level of heterogeneity such that for every smaller value of heterogeneity, $\mathcal{I}$ is strictly increasing in heterogeneity. Although the statement of Proposition 3 is rather intricate, it roughly states that inefficiency tends to be increasing in heterogeneity. We now state this formally.

Proposition 3 *For all (b,γ), there exists a $1 - \theta^* > 0$ such that for all $1 - \theta < 1 - \theta^*$ inefficiency $\mathcal{I}$ is strictly increasing in $1 - \theta$.*

Proof: *See Appendix C.*

The intuition behind the proposition is as follows: when heterogeneity increases, the occurrence of outgroup matches also increases. Within these outgroup matches are matches involving only behavioral players and matches involving at least one rational player. Obviously, in the behavioral only matches, an increase in heterogeneity will, by assumption, imply a greater inefficiency. Also, matches involving exactly one rational player will imply a greater inefficiency unless every rational always plays D.

To better understand the nuanced statement of the proposition we look to the underlying $SPBE$. Appendix B characterizes six different qualitative $SPBE$ regimes. Within each of these $SPBE$, the inefficiency increases from 0 at no heterogeneity. One simple variation is illustrated by the Figure 1.

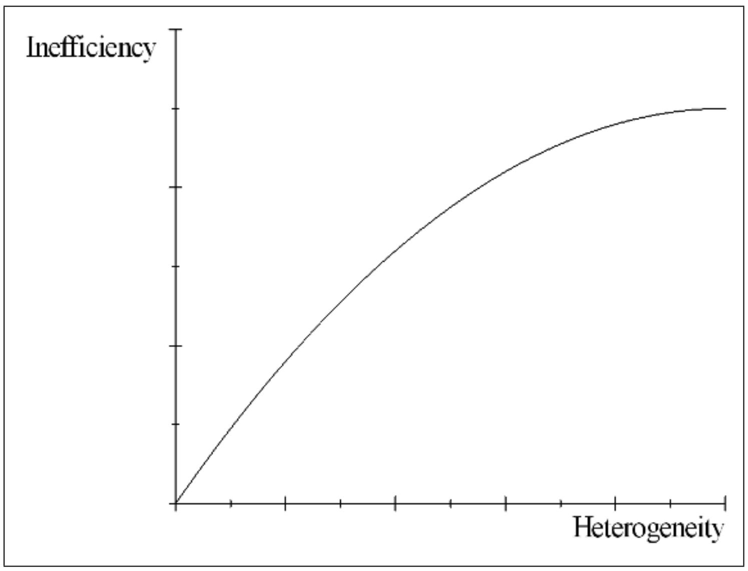


Figure 1-Inefficiency strictly increasing in heterogeneity.

Although there are parameters for which a single qualitative $SPBE$ describes the behavior for all values of heterogeneity, it could also be the case that, as heterogeneity increases, a qualitatively different $SPBE$ can occur. As $1-\theta$ gets larger, the minority reputation becomes less valuable and the majority reputation becomes more valuable. Only two types of such "jumps" can occur as $1-\theta$ becomes larger. Either the majority does not exhibit reputation for any heterogeneity whereas the minority exhibits reputation for small $1-\theta$ and for large values does not exhibit reputation (Figure 2). Or it can be that the minority always exhibits reputation and for small $1-\theta$ the majority does not display reputation and for large values, the majority does (Figure 3).

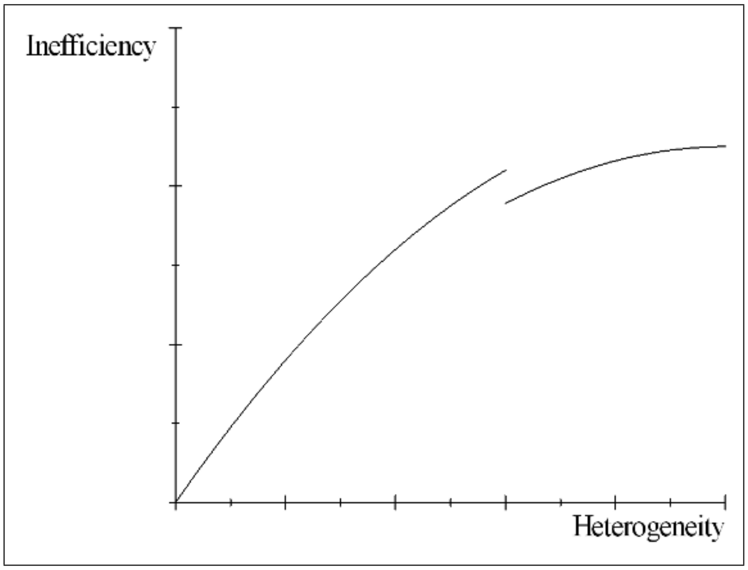


Figure 2-Inefficiency almost everywhere increasing in heterogeneity, with a single downward discontinuity.

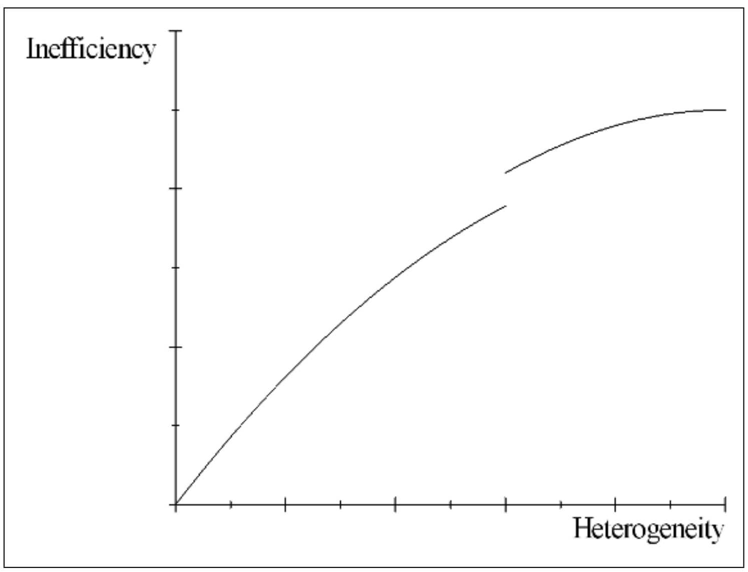


Figure 3-Inefficiency everywhere increasing in heterogeneity, with a single upward discontinuity.

In five of the six $SPBE$, inefficiency strictly increases almost everywhere from $1-\theta=0$ to 0.5 with at most one point of discontinuity. In other words, for these values there are no interior extrema. However, there exists one $SPBE$ where such an interior extrema exists. Figure 4 illustrates a possible relationship.

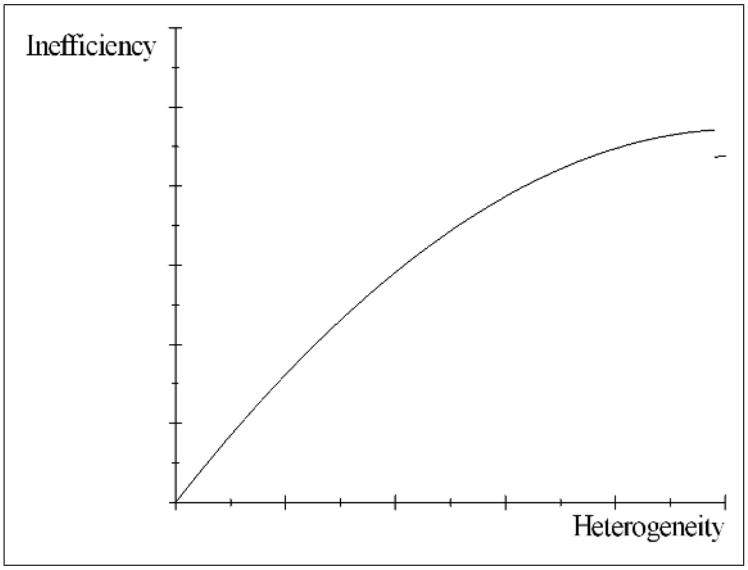


Figure 4-Inefficiency increasing in heterogeneity, with a maximum at 0.485 and a downward discontinuity at 0.49.

Here in Figure 4, for $1-\theta$ less than 0.49 the minority displays $P2$ and the majority does not. Howerver, for $1-\theta$ greater than 0.49 neither the majority nor the minority display $P2$. There is an interior maximum of inefficiency at 0.485. Therefore for such a case to hold we need the interior maximum on the inefficiency function where only m displays $P2$ to occur at a smaller degree of heterogeneity than the point of discontinuity. Although the extremum is always "close" to 0.5, it still remains that there is a small region for which inefficiency is decreasing in heterogeneity. In the cases of Figures 1 and 3 inefficiency is everywhere strictly increasing in heterogeneity, therefore $1-\theta^* = 0.5$. In the case of Figure 2, $1-\theta^*$ is at the point of downward continuity. And in Figure 4, $1-\theta^*$ is at the interior maximum[14] or the point of downward discontinuity.[15]

In summary, Proposition 3 characterizes the claim that inefficiency tends to be increasing in social heterogeneity. Indeed under the legislator interpretation, our model is consistent with the empirical evidence of a positive relationship between heterogeneity and inefficiency.

This completes our discussion of the relationship between social conflict and social heterogeneity. We now turn to the relationship between social conflict and the inequitability of the environment. Recall the discussion of the empirical relationship between inefficient outcomes in a society and inequitability in that society.

In what follows, we show that increasing the inequitability of the environment leads to an increase in inefficiency (Proposition 4). To relate the following proposition to the inequitability

[14] Note that this interior maximum only ranges from $1-\theta^* = 0.4833$ to 0.5.

[15] Here only m displays $P2$. The mixing probability of m is decreasing in heterogeneity and this effect dominates when inefficiency otherwise becomes nearly constant. When both m and M display $P2$, the probability mix of M increases in heterogeneity, and the changes in the mixing of m are offset by the mixing of M.

of the environment literature, we apply the citizen interpretation as defined in the introduction. In what follows, we show that an uneven distribution of rewards leads to social conflict. Social conflict as measured by inefficiency is positively associated with b as we show in the next proposition.

Proposition 4 *$\mathcal{I}$ is strictly increasing in b.*

Proof: See Appendix C.

The map $\frac{\mathcal{I}}{b+1}$ is a function in b with five points of upward discontinuity. The intuition behind the result is as follows: as b increases, playing H becomes more attractive. This leads to an increase in the frequency for which rationals play H and so an increase in inefficiency. Figure 5 illustrates a typical relationship between $\frac{\mathcal{I}}{b+1}$ and b.[16]

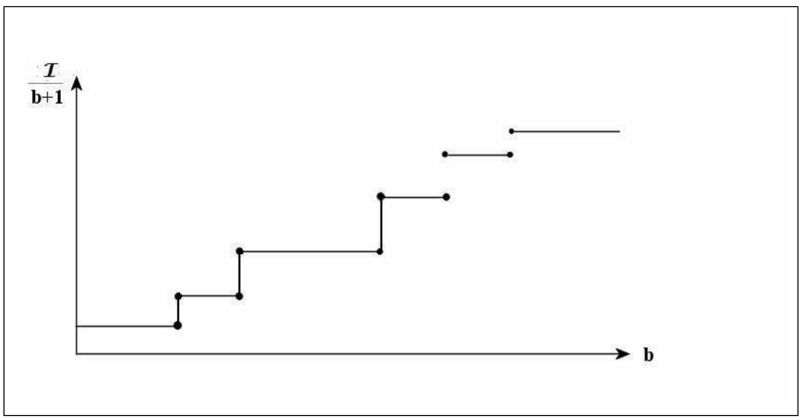


Figure 5-Probability of inefficient outcome and heterogeneity.

Our model provides an explicit account of the individual behavior which drives the social conflict. Specifically, the presence of behavioral players ($\gamma > 0$) means that inefficiency is increasing in the inequitability of the environment. Furthermore, Proposition 4 is free of the built-in inefficiency present in Proposition 3. The increases in inefficiency are completely driven by the behavior of rationals. In Figure 5, any increases beyond the smallest value of $\frac{\mathcal{I}}{b+1}$ are driven entirely by the behavior of the rationals.

One possible alternative explanation for the data is that every member of the society has a preference for better material outcomes for ingroup members, however the fraction of agents intransigently playing H in outgroup matches is increasing in b. We regard our explanation as superior to this alternate explanation, as the latter effectively assumes the result.

[16] To better understand the values for which the $SPBE$ is not unique, see Proposition 8 in Appendix B.

5 Concluding Remarks

We have modeled a social setting containing some agents as described by our interpretation of the social identity literature. We have demonstrated that the struggle for resources, in the presence of agents with a taste for discrimination, can induce rational agents (those without such a taste) to aggressively discriminate. We showed that our model is consistent with empirical papers which claim a relationship between social conflict and a measure of the social heterogeneity. The results in our model are also consistent with the literature identifying a relationship between social conflict and the inequitability of the environment.

The paper also includes the following results. For games which induce a sufficiently symmetric equilibrium, the majority has a greater ex-ante payoff than the minority. We showed that the minority always plays the game at least as aggressively as the majority. We interpret this result as consistent with the experimental findings that minorities have stronger group identities than do majorities.

Although the model is quite complicated there remain interesting questions. For instance, it would be interesting to investigate a model in which information is less than perfect. Obviously some information is required for the results to hold, however it might prove fruitful to investigate weaker assumptions. It would also be interesting to model the presence of three of more groups. It could be the case that there is be an interaction among the groups which is not present with only two groups.

In light of the recent interest in fairness, it is useful to note that there exist aspects of every society which could be described as unfair. In every society, economic inequalities persist on the basis of race, religion and gender. We argue that, in economic situations, *unfairness* is at least as important than *fairness*. It is also our opinion that the social identity literature is useful in providing direction for the study of unfairness.

6 Appendix

The appendix is arranged as follows. Appendix A contains some background, technical aspects of the model which we use subsequently. Appendix B contains a complete characterization of the $SPBE$ for all parameter values. We characterize the $SPBE$ where it is unique (Propositions 5, 6 and 7) then we characterize the $SPBE$ where it is not unique (Proposition 8). Appendix C contains the proofs of the propositions presented in the body of the paper. These proofs are quite involved, as they often require verification across the various $SPBE$ specified in Appendix B, calculation of expectations across every possible history or both.

6.1 Appendix A

Lemma 2 *In every $SPBE$ a second period player 2 plays D with probability 1.*

Proof: In the second period, no rational agent will play H against an H because there is no future reputation to protect and H gives a strictly worse stage game payoff than D. Therefore in outgroup matches it must be that H is played with zero probability for second period player 2.■

Lemma 2 allows us to focus our entire attention towards characterizing second period strategies as a player 1. Recall from the body of the paper that since $\sigma_2^i(H, h^j) = 0$ for all $h^j \in \mathcal{H}$ and $i \in \{m, M\}$ we write $\sigma_2^i(1, h^j)$ as $\sigma_2^i(h^j)$.

Lemma 3 *In every $SPBE$, for all h^i if $(1-p^j(h^j))b < 1$ then $\sigma_2^i(h^j) = 0$, if $(1-p^j(h^j))b > 1$ then $\sigma_2^i(h^j) = 1$ and if $(1-p^j(h^j))b = 1$ then $\sigma_2^i(h^j) \in [0, 1]$.*

Proof: Given Lemma 2, the player will receive a second period material payoff of 1 for selecting D and $(1 - p^j(h^j))b$ for selecting H. The strategy employed by the player 2 does not depend on the history of the player 1. Therefore, the second period player 1 strategy is completely determined by the opponent's posterior probability of being behavioral.■

Lemma 3 states that σ_2^i is completely determined by the posterior probability of the opponent. Since we have already determined the strategy of a second period player 2 is independent of history, it follows that the entire relevant history in the second period can be summarized by the opponent's history in the first period. An implication of the above lemmas, is that characterizing the $SPBE$ boils down to characterizing $\sigma_1^i(1)$, $\sigma_1^i(H)$ and $\sigma_2^i(h^j)$ for all $i \in M(m)$, $j \in m(M)$ and all $h^j \in \mathcal{H}$.

We define $v_i(h^i)$ as the expected payoff of i entering period 2 with a history of h^i. An implication of Lemmas 2 and 3 is that the difference in continuation payoffs can be summarized by the difference in expected payoffs as a second period player 2 as strategy for an ingroup and outgroup as a player 1 are independent of the player's own history.

Corollary 1 *The term* $v_i(Hc) - v_i(Dc)$ *for* $c \in \{1, H\}$ *is equivalent to the difference in expected payoffs as a player* 2.

Another implication of Lemma 3 and the Condition (∗) is stated below as a corollary.

Corollary 2 *If* $h^i \in \mathcal{H}_D$ *then* $\sigma^j(h^i) = 1$

Corollary 2 simply says that if an agent ever plays D in the first period, then the second period opponent will play H as a player 1. This follows directly from Lemma 3 and Condition (∗).

The following two lemmas provide useful technical results.

Lemma 4 *If* $\gamma \geq \left(\frac{b-1}{b}\right)^2$ *then* $b < \frac{2}{\theta(1-\gamma)} + 1$

Proof: Note that $b < \frac{2}{\theta(1-\gamma)} + 1$ is equivalent to

$$\gamma > \frac{\theta(b-1)-2}{\theta(b-1)} \tag{1}$$

With a domain of $\theta \in [0.5, 1]$, the right hand side of expression (1) attains a maximum at $\theta = 1$. Therefore

$$\frac{b-3}{b-1} \geq \frac{\theta(b-1)-2}{\theta(b-1)}$$

Notice that for all $b > \frac{1}{3}$

$$\left(\frac{b-1}{b}\right)^2 > \frac{b-3}{b-1} \tag{2}$$

Expression (2) then implies that if $\gamma \geq \left(\frac{b-1}{b}\right)^2$ then it must be that $\gamma > \frac{\theta(b-1)-2}{\theta(b-1)}$ and so the lemma is proved.■

Lemma 5 $b < \frac{2}{(1-\theta)(1-\gamma)} + 1$ *(*$b < \frac{2}{\theta(1-\gamma)} + 1$*) if and only if* M *(*m*) does not exhibit P2.*

Proof: According to Corollary 1, $\sigma_1^M(H) > 0$ if and only if

$$1 + \left(\frac{1-\theta}{2}\right) \geq 0 + \left(\frac{1-\theta}{2}\right)(b(1-\gamma)+\gamma)$$

The left side represents the expected utility heading into the second period with a posterior of 1 and the right side represents the expected utility entering the second period known to be a rational. The analogous reasoning holds for m.■

Corollary 3 *P*2 *cannot occur in any SPBE if* $\gamma \geq \left(\frac{b-1}{b}\right)^2$

This corollary follows from Lemmas 4 and 5 since $b \geq \frac{2}{\theta(1-\gamma)} + 1$ ($b \geq \frac{2}{(1-\theta)(1-\gamma)} + 1$) is a necessary condition for m (M) to display $P2$. This is the lower bound of b for which a player would sacrifice an immediate payoff of 1 in order to find entering the second period with a posterior of 1. This allows us to restrict attention to the $SPBE$ which contains $P2$ to $\gamma < \left(\frac{b-1}{b}\right)^2$. Furthermore, note that the second condition for $P1$ requires that $(1-\gamma)(1-\sigma_1^j(H))b < 1$. This implies that $P1$ only occurs when $\gamma \geq \frac{b-1}{b}$ as $\frac{b-1}{b} > \left(\frac{b-1}{b}\right)^2$. In other words, there are no parameter values for which the $SPBE$ exhibits both $P1$ and $P2$, as the following lemma states.

Lemma 1 *If any player exhibits* $P1$ *(*$P2$*) then there are no parameter values for which any other player exhibits* $P2$ *(*$P1$*).*

6.2 Appendix B

In the next four subsections we completely characterize the $SPBE$. Although we do not explicitly use any of the results, they form the basis for the propositions given in the body of the paper.

6.2.1 $SPBE$ for Small b, ($P2$) does not exist for either group

In this class of $SPBE$ neither group will display $P2$ because it will not be profitable to play H as a player 2 in order to enter the second period with a posterior even as high as 1. The proposition below characterizes the $SPBE$ for small b.

Descriptively, for small γ, (case (i)) both groups play aggressively as a player 1. The only situation where the strategy of playing H as a player 1 is not $SPBE$ is when the second period opponent played H in the first period. This is because a history of HH is the only history leading to a posterior greater than $\frac{b-1}{b}$. In both periods, the optimal strategy turns out to be the one which myopically maximizes payoffs.

For case (ii), both groups display $P1$. In the first period, both groups play H as a player 1 rather than D, despite the fact that the latter yields a higher stage game payoff. Here, D is myopically superior to H despite the fact that first period player 2 does not play H. The myopic action is not selected because the first period player 1 selecting D forfeits reputation in the second period and it is sufficiently valuable.

For case (iii), only m displays $P1$. This asymmetry arises because M does not find it profitable to maintain its reputation.

For case (iv), neither player selects H in the first period as a player 1 because of the high likelihood of being matched with a behavioral. No rational agent plays H as a second period player 1 unless the opponent has played D in the first period.

Proposition 5 *If* $\frac{2}{(1-\theta)(1-\gamma)} + 1 > \frac{2}{\theta(1-\gamma)} + 1 > b$ *then the unique* $SPBE$ *is such that* $\sigma_1^i(H) = 0$*,* $\sigma_2^i(h_1^j) = 0$ *if* $h_1^j = HH$ *and* $\sigma_2^i(h_1^j) = 1$ *if* $h_1^j \in \mathcal{H}_D$ *for all* $i \in \{m, M\}$*. Furthermore, if in addition to* $\frac{2}{\theta(1-\gamma)} + 1 > b$ *it is also the case that*

(i) $\gamma < \frac{b-1}{b}$ *then the unique* $SPBE$ *is such that* $\sigma_1^i(1) = 1$*,* $\sigma_2^i(h_1^j) = 0$ *if* $h_1^j \neq HH$ *for all* $i \in \{m, M\}$*.*

(ii) $\gamma \in (\frac{b-1}{b}, \overline{\gamma}_M)$ *where* $\overline{\gamma}_M = \frac{b-1+\left(\frac{1-\theta}{2}\right)(b-1)}{b+\left(\frac{1-\theta}{2}\right)(b-1)} > \frac{b-1}{b}$ *then the unique* $SPBE$ *is such that for* $i \in N$*,* $\sigma_1^i(1) = 1$*,* $\sigma_2^i(h_1^j) = 0$ *if* $h_1^j = \{I, H1, E\}$*.*

(iii) $\gamma \in (\overline{\gamma}_M, \overline{\gamma}_m)$ *where* $\frac{b-1+\left(\frac{\theta}{2}\right)(b-1)}{b+\left(\frac{\theta}{2}\right)(b-1)} = \overline{\gamma}_m > \frac{b-1+\left(\frac{1-\theta}{2}\right)(b-1)}{b+\left(\frac{1-\theta}{2}\right)(b-1)} = \overline{\gamma}_M > \frac{b-1}{b}$ *then the unique* $SPBE$ *is such that* $\sigma_1^m(1) = 1$*,* $\sigma_1^M(1) = 0$*,* $\sigma_2^i(h_1^j) = 0$ *if* $h_1^j \in \{I, H1, E\}$*.*

(iv) $\gamma > \overline{\gamma}_m$ *then the unique* $SPBE$ *is such that for all* $i \in \{m, M\}$*,* $\sigma_1^i(1) = 0$*,* $\sigma_2^i(h_1^j) = 0$ *for* $h_1^j \notin \mathcal{H}_D$

Proof of Proposition 5: In any $SPBE$ with

$$\frac{2}{(1-\theta)(1-\gamma)} + 1 > \frac{2}{\theta(1-\gamma)} + 1 > b$$

it must be that $\sigma_1^i(H) = 0$, by Lemma 5. This implies posteriors of $p^i(h^i) = 1$ for $h^i = HH$ and $p^i(h^i) = 0$ for $h^i = DH$ and strategies $\sigma^j(h^i) = 0$ for $h^i = HH$. If $\sigma_1^i(H) = 0$ then $p(HH) = 1$ and therefore by Lemma 3, $\sigma_2^i(HH) = 0$. And by Corollary 2, $\sigma_2^i(h_1^j) = 1$ if $h_1^j \in \mathcal{H}_D$. Furthermore, there can be no other $SPBE$ strategies.

(i) It will be that $\sigma_2^i(h^j) = 1$ if $h^j \in \{I, E\}$ because $p^j(h^j) = \gamma < \frac{b-1}{b}$. It remains to determine $\sigma_1^i(1)$ and $\sigma_2^i(H1)$. It cannot be that $\sigma_1^i(1) = 0$ as this would imply that $p^i(H1) = 1$ and $\sigma_2^j(H1) = 0$. However, a deviation is easy to find as both the first period stage game payoffs are higher for H:

$$b(1-\gamma) > 1 \tag{3}$$

and

$$v_i(H1) > v_i(D1) \tag{4}$$

because $p^i(H1) = 1 > \frac{b-1}{b} > p^i(D1) = 0$. Therefore $\sigma_1^i(1) \neq 0$. It cannot be that $\sigma_1^i(1) = \alpha^* \in (0,1)$ because the first period player 1 cannot be indifferent between playing H and D as a player 1. Therefore $\sigma_1^i(1) = 1$ and $p^i(H1) = \gamma$ so that $\sigma_2^i(h^j) = 1$. Furthermore, there can be no other $SPBE$ strategies.

(ii) Here it cannot be that $\sigma_1^i(1) = 0$ as this would imply $p^i(H1) = 1$, $\sigma_2^i(h^j) = 0$ for

$h^j = H1$. However, a deviation exists for M:

$$
\begin{aligned}
b(1-\gamma) + v_M(H1) &> 1 + v_M(D1) \\
b(1-\gamma) + \left(\frac{1-\theta}{2}\right)(b-1)(1-\gamma) &> 1 \qquad (5)\\
\frac{b-1+\left(\frac{1-\theta}{2}\right)(b-1)}{b+\left(\frac{1-\theta}{2}\right)(b-1)} = \overline{\gamma}_M &> \gamma
\end{aligned}
$$

And similarly for m:

$$
\frac{b-1+\left(\frac{\theta}{2}\right)(b-1)}{b+\left(\frac{\theta}{2}\right)(b-1)} = \overline{\gamma}_m
$$

where $\overline{\gamma}_m > \overline{\gamma}_M > \frac{b-1}{b}$. Therefore, $\sigma_1^i(1) > 0$ despite the fact that the first period stage game payoff for D is greater than that of H for a player 1 of both groups. Hence, we say that both m and M display $P1$. It also cannot be that $\sigma_1^i(1) \in (0,1)$. In order for the first period player 1 to mix, it would require:

$$
b(1-\gamma) + v_i(H1) = 1 + v_i(D1) \qquad (6)
$$

Since $\gamma > \frac{b-1}{b}$, (or $b(1-\gamma) < 1$), expression (6) will only hold if $v_i(H1) > v_i(D1)$. According to Lemma 3, expression (6) only holds when $\sigma_1^i(1)$ is such that $p^i(H1) \geq \frac{b-1}{b}$. Since $\overline{\gamma}_M > \gamma$

$$
b(1-\gamma) + v_M(H1) > 1 + v_M(D1)
$$

if $p^i(H1) > \frac{b-1}{b}$. Therefore, the only way to satisfy (6) is to select $\sigma_2^j(h^i)$ for $h^i = H1$ such that $p^i(H1) = \frac{b-1}{b}$ and this is impossible given that the prior γ is strictly greater than $\frac{b-1}{b}$. If $\sigma_1^i(1)$ is such that $p^i(H1) > \frac{b-1}{b}$ then $\sigma_2^j(h^i) = 0$ for $h^i = H1$. Therefore, the optimal choice is $\sigma_1^i(1) = 1$ and as a consequence $\sigma_2^j(h^i) = 0$ for $h^i = H1$. It also follows that since $\gamma > \frac{b-1}{b}$ that $\sigma_2^i(h^j) = 0$ for $h^j \in \{I, E\}$. Indeed, this last fact holds for the final three sections of the proof. Furthermore, there can be no other $SPBE$ strategies.

(iii) Since $\gamma \in (\overline{\gamma}_M, \overline{\gamma}_m)$ we can make identical arguments as those given in part (ii) only for m and not M. Therefore $\sigma_1^m(1) = 1$ and $\sigma^M(h^m) = 0$ such that $h^m = H1$. In the case of M, it cannot be that $\sigma_1^M(1) = 1$ because expression (5) no longer holds. It cannot be that $\sigma_1^M(1) \in (0,1)$ because expression (6) cannot be satisfied by any value in this range. Therefore, $\sigma_1^M(1) = 0$ and $\sigma_2^m(h^M) = 0$ for $h^M = H1$ as $p^M(H1) = 1$ as it is no longer for worthwhile for M to display $P1$. Furthermore, there can be no other $SPBE$ strategies.

(iv) Now the arguments supporting $\sigma_1^i(1) \in (0,1]$ in cases (ii) and (iii) do not hold for either group. Therefore, $\sigma_1^i(1) = 0$ and $\sigma_2^i(h^j) = 0$ for $h^j = H1$ as $p^i(H1) = 1$. It is no longer for either group to display $P1$. Furthermore, there can be no other $SPBE$ strategies.■

6.2.2 $SPBE$ for Medium b, $(P2)$ exists for m but not M

Here, m finds it profitable to play H as a player 2 with probability strictly between 0 and 1 in order to enter the next period with a posterior of $\frac{b-1}{b}$. Unlike m, M never finds it profitable to play H as a first period player 2 even if it secures a posterior of 1 in the second period. Therefore, m displays $P2$ and M does not. Note that by Lemma 4, we can restrict attention to $\gamma < \left(\frac{b-1}{b}\right)^2$ and therefore every agent plays H as a first period player 1. By being able to restrict attention to $\gamma < \left(\frac{b-1}{b}\right)^2$ we do not have the multitude of cases that we had in the Proposition 5. Every first period player demands H and in the second period demands H in response to a history of $H1$ as the probability of a behavioral is sufficiently low.

Proposition 6 *If* $\frac{2}{(1-\theta)(1-\gamma)} + 1 > b > \frac{2}{\theta(1-\gamma)} + 1$ *then the unique* $SPBE$ *must be that* $\sigma_1^M(H) = 0$, $\sigma_2^m(h^M) = 0$ *for* $h^M = HH$, $\sigma_2^M(h^m) = 1 - \frac{1}{\left(\frac{\theta}{2}\right)(b-1)(1-\gamma)}$ *for* $h^m = HH$, $\sigma_1^m(H) = \beta^* \in (0,1)$ *such that* $p^m(HH) = \frac{b-1}{b}$ *where* $\beta^* = \left(\frac{\gamma}{1-\gamma}\right)\left(\frac{1}{b-1}\right)$ *and for* $i \in \{m, M\}$, $\sigma_1^i(1) = 1$, $\sigma_2^i(h^j) = 1$ *for* $h^j \notin HH$.

Proof of Proposition 6: In any $SPBE$ with $\frac{2}{(1-\theta)(1-\gamma)} + 1 > b > \frac{2}{\theta(1-\gamma)} + 1$, it must be that $\sigma_1^m(H) = \beta^* \in (0,1)$ such that $p^m(HH) = \frac{b-1}{b}$ and $\sigma_1^M(H) = 0$. By Lemma 5, it cannot be that $\sigma_1^M(H) > 0$. Therefore, $\sigma_1^M(H) = 0$ and $\sigma_2^m(h^M) = 0$ when $h^M = HH$. In the case of m, it cannot be that $\sigma_1^m(H) = 0$. It also cannot be that $\sigma_1^m(H) = 1$ as this implies that $p^m(HH) = \gamma < \frac{b-1}{b}$ and so $v_m(HH) = v_m(DH)$. Therefore, $\sigma_1^m(H) = 0$ is a profitable deviation. It must be that $\sigma_1^m(H) = \beta^*$ such that

$$\begin{aligned} p^m(HH) &= \frac{b-1}{b} = \frac{\gamma}{\gamma + (1-\gamma)\beta^*} \\ \beta^* &= \left(\frac{\gamma}{1-\gamma}\right)\left(\frac{1}{b-1}\right) \end{aligned}$$

If $\sigma_1^m(H) > \beta^*$ then $p^m(HH) < \frac{b-1}{b}$ which would imply $\sigma_2^M(h^m) = 1$ where $h^m = HH$. There would be no benefit for $\sigma_1^m(H) > 0$, and so it must be that $\sigma_1^m(H) \leq \beta^*$. If $\sigma_1^m(H) < \beta^*$ then $p^m(HH) > \frac{b-1}{b}$ which would imply that $\sigma_2^M(h^m) = 0$ where $h^m = HH$. However, if $\sigma_2^M(h^m) = 0$ where $h^m = HH$ then $\sigma_1^m(H) = 1$ is optimal. By the above argument this cannot be the case, therefore $\sigma_1^m(H) = \beta^*$. The $SPBE$ requires

$$0 + v_m(HH) = 1 + v_m(DH)$$

$$0 + \left(\frac{\theta}{2}\right)[b(1-\gamma)(1-\sigma_2^M(HH)) + (1-\gamma)\sigma_2^M(HH) + \gamma] = 1 + \left(\frac{\theta}{2}\right)$$

so that

$$\sigma_2^M(HH) = \frac{\left(\frac{\theta}{2}\right)(b-1)(1-\gamma) - 1}{\left(\frac{\theta}{2}\right)(b-1)(1-\gamma)}$$

Therefore $\sigma_1^m(H) = \beta^*$ such that $p^m(HH) = \frac{b-1}{b}$. Additionally, since $\gamma < \left(\frac{b-1}{b}\right)^2 < \frac{b-1}{b}$, it must be that $\sigma_2^i(h^j) = 1$ for $h^j \in \{I, E, D1, DH\}$. Since $\gamma \leq \left(\frac{b-1}{b}\right)^2$ the $SPBE$ must be that

$\sigma_1^M(1) = 1$ because

$$b(1-\gamma)(1-\beta^*) + v_M(H1) \geq 1 + v_M(D1) \tag{7}$$

$v_M(H1) = v_M(D1)$ as $p^M(H1) < \frac{b-1}{b}$. Expression (7) holds when $\gamma \leq \left(\frac{b-1}{b}\right)^2$. Furthermore, $\sigma_1^m(1) = 1$ and $\sigma_2^M(h^m) = 1$ for $h^m = H1$. This is true as $v_m(H1) = v_m(D1)$ and $b(1-\gamma) > 1$. Furthermore, there can be no other $SPBE$ strategies.

■

6.2.3 $SPBE$ for Large a, $(P2)$ exists for both groups

Here, both groups display $P2$. Much of the reasoning above involving m now holds for both groups. Again, by Lemma 4, we restrict attention to $\gamma < \left(\frac{b-1}{b}\right)^2$. Both groups play H as a player 1 in the first period and play H as a second period player 1 against a player with a history of $H1$.

Proposition 7 *If $b > \frac{2}{(1-\theta)(1-\gamma)} + 1 > \frac{2}{\theta(1-\gamma)} + 1$ then the $SPBE$ must be that $\sigma_2^m(HH) = 1 - \frac{1}{\left(\frac{1-\theta}{2}\right)(b-1)(1-\gamma)}$, $\sigma_2^M(HH) = 1 - \frac{1}{\left(\frac{\theta}{2}\right)(b-1)(1-\gamma)}$, $\sigma_1^i(H) = \beta^* \in (0,1)$ such that $p^i(HH) = \frac{b-1}{b}$ and $\sigma_2^i(h^j) = 1$ where $h^j \in \{I, E, H1\}$ and $\sigma_1^i(1) = 1$.*

Proof of Proposition 7: In any $SPBE$ with $b \geq \frac{2}{(1-\theta)(1-\gamma)} + 1 > \frac{2}{\theta(1-\gamma)} + 1$, it must be that $\sigma_1^i(H) = \beta^* \in (0,1)$ such that $p^i(HH) = \frac{b-1}{b}$. Here, the argument presented in the proof of Proposition 6 goes through for both M and m. It also must be that $\sigma_2^i(h^j) \in (0,1)$ where $h^j = HH$. Just as in Proposition 6, in order to determine $\sigma_2^M(HH)$ it must be that

$$0 + \left(\frac{\theta}{2}\right)(b(1-\gamma)(1-\sigma_2^M(HH)) + (1-\gamma)\sigma_2^M(HH) + \gamma) = 1 + \left(\frac{\theta}{2}\right)$$

and similarly for $\sigma_2^m(HH)$. Additionally, Lemma 4 allows us to restrict attention to $\gamma < \left(\frac{b-1}{b}\right)^2 < \frac{b-1}{b}$. This allows us to determine that $\sigma_2^i(h^j) = 1$ for $h^j \in \{I, E\}$. Since $\gamma \leq \left(\frac{b-1}{b}\right)^2$ arguments in the proof of Proposition 6 apply to both M and m therefore $\sigma_1^i(1) = 1$ and $\sigma_2^i(h^j) = 1$ for $h^j = H1$. Furthermore, there can be no other $SPBE$ strategies.■

6.2.4 Non-generic parameter values

The $SPBE$ is generically unique, as the following corollary shows. Following the corollary, is a result which describes the $SPBE$ for generic parameter values. There exists a set Ψ, of measure zero, in the parameter space for which the $SPBE$ is not unique. For parameter values not contained in Ψ, the $SPBE$ is unique. We explicitly define Ψ as

$$\begin{aligned} \Psi &= \{(b,\theta,\gamma) : b \in \{\frac{2}{\theta(1-\gamma)} + 1, \frac{2}{(1-\theta)(1-\gamma)} + 1\} \\ \text{or } \gamma &\in \{\frac{b-1}{b}, \frac{b-1+\left(\frac{1-\theta}{2}\right)(b-1)}{b+\left(\frac{1-\theta}{2}\right)(b-1)}, \frac{b-1+\left(\frac{\theta}{2}\right)(b-1)}{b+\left(\frac{\theta}{2}\right)(b-1)}\}\} \end{aligned}$$

The following corollary follows from Propositions 5 (i), 5 (ii), 5 (iii), 5 (iv), 6 and 7.

Corollary 4 *If parameters (b, θ, γ) are not contained in the set Ψ then the σ satisfying the conditions for $SPBE$ will be unique.*

Lemma 4 demonstrates that either a condition for b can be satisfied or a condition for γ can be satisfied, but not both. The values of b given above are the values for which the minority (respectively majority) will be indifferent between displaying $P2$ or not. The first value of γ represents the value for which a second period player 2 will be indifferent between playing H and D against an opponent with a history h such that $p^i(h) = \gamma$. The second (and third) value(s) of γ denotes the parameter for which the majority (minority) is indifferent between displaying $P1$ and not.

Now, we characterize the $SPBE$ for each element of Ψ.

Proposition 8 *(a) If $b < \frac{2}{\theta(1-\gamma)} + 1$ and $\gamma = \frac{b-1}{b}$ then the $SPBE$ is not unique as the strategies specified in Proposition 5 (i) or (ii) or any mixture will suffice.*

(b) If $b < \frac{2}{\theta(1-\gamma)} + 1$ and $\gamma = \frac{b-1+\left(\frac{1-\theta}{2}\right)(b-1)}{b+\left(\frac{1-\theta}{2}\right)(b-1)}$ then the $SPBE$ is not unique as the strategies specified for M in Proposition 5 (ii) or (iii) or any mixture will suffice.

(c) If $b < \frac{2}{\theta(1-\gamma)} + 1$ and $\gamma = \frac{b-1+\left(\frac{\theta}{2}\right)(b-1)}{b+\left(\frac{\theta}{2}\right)(b-1)}$ then the $SPBE$ is not unique as the strategies specified for m in Proposition 5 (iii) or (iv) or any mixture will suffice.

(d) If $b = \frac{2}{\theta(1-\gamma)} + 1$ then the $SPBE$ is not unique as the strategies specified for m in Proposition 6 and those specified in Proposition 5 (i), however no mixture between them will suffice.

(e) If $b = \frac{2}{(1-\theta)(1-\gamma)} + 1$ then the $SPBE$ is not unique as the strategies specified for M in Proposition 7 and those specified in Proposition 6, however no mixture between them will suffice.

Proof: In the case of (a), any $\sigma_2^i(h) \in [0,1]$ where h such that $p^j(h) = \gamma$ is an $SPBE$. For such histories, according to Lemma 3, the second period player 2 is indifferent between actions. For histories I, $H1$ and E any second period strategies will suffice. In the case of (b), the majority is indifferent between displaying $P1$ or not. Any $\sigma_1^M(1) \in [0,1]$ will constitute an $SPBE$. These first period player 1 strategies will induce posteriors strictly between γ and 1. Therefore, the second period strategies are unchanged. In the case of (c), the minority is indifferent between displaying $P1$ or not. Reasoning similar to case (b) applies to m. In the case of (d), the minority is indifferent between displaying $P2$ or not. However, unlike the previous cases, the $SPBE$ cannot contain any mixture between the equilibria will not form a $SPBE$. Given condition (iii) of the definition of $SPBE$ it must be either $\sigma_1^m(H) \in \{0, \left(\frac{\gamma}{(1-\gamma)(b-1)}\right)\}$. Any other value would imply $p^m(HH) \neq \frac{b-1}{b}$. Unlike the cases of (a), (b), and (c), the first period strategy nontrivially affects the second period posteriors, as $\gamma < \frac{b-1}{b}$. For the parameter values given, there is no deviation from the m strategy given in Proposition 6. Likewise, there is no deviation from the strategy given in Proposition 5(i). In the case of (e), the majority is indifferent between displaying $P2$ or not. Reasoning similar to case (d) applies to M.■

The statement of Proposition 8 elucidates Figure 5 in the body of the paper. In this figure, the relationship between I and b is connected at 3 points of discontinuity ((a), (b) and (c)) and not connected at two points of discontinuity ((d) and (e)).

6.3 Appendix C

Here, we prove the results presented in the body of the paper. First, we give a more detailed account of Example 2. Then, we prove Propositions 1, 2, 4, and 3. Note that we prove Proposition 4 before Proposition 3, although in the body of the paper they are presented in the opposite order.

Example 2 Suppose that $\theta = 0.6$, $b = 2$, and $\gamma = 0.55$. The $SPBE$ which corresponds to these parameter values is described by Proposition 5 (iii) in Appendix B. In this $SPBE$ the minority displays $P1$ and the majority does not. Therefore, the $SPBE$ is not strongly symmetric although any aspects of the $SPBE$ do exhibit symmetry. For instance the first period strategy as a player 2 plays H with probability 0. In the second period each player selects H with probability 1 if the opponent played D in an outgroup match in period 1 and plays H with probability 0 otherwise:

$$\sigma_1^m(H) = \sigma_1^M(H) = 0$$
$$\sigma_2^m(h \in \mathcal{H}_D) = \sigma_2^M(h \in \mathcal{H}_D) = 1$$
$$\sigma_2^m(h \notin \mathcal{H}_D) = \sigma_2^M(h \notin \mathcal{H}_D) = 0$$

The $SPBE$ is not strongly symmetric because in the first period as a player 1 the minority group plays H with probability 1 and the majority group plays H with probability 0:

$$\sigma_1^m(1) = 1 \text{ and } \sigma_1^M(1) = 0$$

These strategies induce a distribution of histories of

$$\begin{aligned} \mu_m(I) &= 1 - \theta \\ \mu_m(H1) &= \frac{\theta}{2} \\ \mu_m(E) &= \frac{\theta}{2}(1-\gamma) \\ \mu_m(DH) &= \frac{\theta}{2}\gamma \end{aligned}$$

and

$$\begin{aligned} \mu_M(I) &= \theta \\ \mu_M(H1) &= \mu_M(HH) = \left(\frac{1-\theta}{2}\right)\gamma \\ \mu_M(D1) &= \mu_M(DH) = \left(\frac{1-\theta}{2}\right)(1-\gamma) \end{aligned}$$

The expected payoff for m:

$$E^m = (1-\theta)\left(\frac{b+1}{2}\right) + \theta b(1-\gamma) + \theta\gamma + \left(\frac{\theta}{2}\gamma\right) v_m(h \in \mathcal{H}_D) + \left(1 - \frac{\theta\gamma}{2}\right) v_m(h \notin \mathcal{H}_D)$$

where

$$v_m(h \in \mathcal{H}_D) = (1-\theta)\left(\frac{b+1}{2}\right) + \frac{\theta}{2}[1 + (1-\theta)(1-\gamma)b + \theta + (1-\theta)\gamma]$$
$$v_m(h \notin \mathcal{H}_D) = (1-\theta)\left(\frac{b+1}{2}\right) + \frac{\theta}{2}[(1-\gamma)b + \gamma + (1-\theta)(1-\gamma)b + \theta + (1-\theta)\gamma]$$

The expected payoffs for M:

$$E^M = \theta\left(\frac{b+1}{2}\right) + (1-\theta) + \theta v_M(h \notin \mathcal{H}_D) + (1-\theta) v_M(h \in \mathcal{H}_D)$$

where

$$v_M(h \in \mathcal{H}_D) = \theta\left(\frac{b+1}{2}\right) + \left(\frac{1-\theta}{2}\right)[1 + \left(1 - \frac{\theta\gamma}{2}\right) + \frac{\theta\gamma}{2}b]$$
$$v_M(h \notin \mathcal{H}_D) = \theta\left(\frac{b+1}{2}\right) + \left(\frac{1-\theta}{2}\right)[(1-\gamma)b + \gamma + \left(1 - \frac{\theta\gamma}{2}\right) + \frac{\theta\gamma}{2}b]$$

These imply $v_m(h \in \mathcal{H}_D) = 1.242$, $v_m(h \notin \mathcal{H}_D) = 1.377$, $v_M(h \in \mathcal{H}_D) = 1.333$ and $v_M(h \notin \mathcal{H}_D) = 1.423$. Therefore

$$E^m = 2.825 > E^M = 2.687$$

■

Proof of Proposition 1: This proof consists in showing that $E^M - E^m > 0$ for the $SPBE$ described in Propositions 5 (i), (ii), (iv) and 7. Consider the $SPBE$ as specified by Proposition 5 (i). The expected values of the majority and the minority are:

$$E^M = \theta[\left(\frac{b+1}{2}\right) + v_M(I)] + \left(\frac{1-\theta}{2}\right)[b(1-\gamma) + v_M(H1)] + \left(\frac{1-\theta}{2}\right)[1 + v_M(DH)]$$

and

$$E^m = (1-\theta)[\left(\frac{b+1}{2}\right) + v_m(I)] + \frac{\theta}{2}[b(1-\gamma) + v_m(H1)] + \frac{\theta}{2}[1 + v_m(DH)]$$

Further simplification can be done as $v_i(I) = v_i(H1) = v_i(DH) = v_i$ for both groups $i \in \{M, m\}$ as $\gamma < \frac{b-1}{b}$. This implies that

$$E^M - E^m = \left(\frac{2\theta - 1}{2}\right) b\gamma + v_M - v_m$$

where

$$v_M = \frac{\theta}{2}(b+1) + \left(\frac{1-\theta}{2}\right)[1 + \mu_m(HH) + b\mu_m(DH) + b(1-\gamma)\mu_m(H1) + b(1-\gamma)\mu_m(I)]$$
$$= \frac{1}{2} + \frac{\theta b}{2} + \left(\frac{1-\theta}{2}\right)\left(b(1-\gamma) + \frac{\theta}{2}\gamma\right)$$

and likewise

$$v_m = \left(\frac{1-\theta}{2}\right)(b+1) + \frac{\theta}{2}[1 + \mu_M(HH) + b\mu_M(DH) + b(1-\gamma)\mu_M(H1) + b(1-\gamma)\mu_M(I)]$$
$$= \frac{1}{2} + \frac{(1-\theta)b}{2} + \frac{\theta}{2}\left(b(1-\gamma) + \left(\frac{1-\theta}{2}\right)\gamma\right)$$

so that

$$E^M - E^m = (2\theta - 1)b\gamma$$

which is always positive.

Consider the $SPBE$ as specified by Proposition 5 (ii). In this case both groups display $P1$, therefore it will not be the case that every continuation value will be equal. The expected values of the majority and the minority are:

$$E^M = \theta[\left(\frac{b+1}{2}\right) + v_M(I)] + \left(\frac{1-\theta}{2}\right)[b(1-\gamma) + v_M(H1)] + \left(\frac{1-\theta}{2}\right)[1 + v_M(DH)]$$

and

$$E^m = (1-\theta)[\left(\frac{b+1}{2}\right) + v_m(I)] + \frac{\theta}{2}[b(1-\gamma) + v_m(H1)] + \frac{\theta}{2}[1 + v_m(DH)]$$

Note that $v_i(I) = v_i(H1) > v_i(DH)$ for both groups. This implies that

$$E^M - E^m = \left(\frac{2\theta-1}{2}\right)b\gamma + \left(\frac{1-\theta}{2}\right)v_M(DH) + \left(\frac{1+\theta}{2}\right)v_M(I) - \frac{\theta}{2}v_m(DH) - \left(1 - \frac{\theta}{2}\right)v_m(I)$$

where

$$v_M(I) = \frac{\theta}{2}(b+1) + \left(\frac{1-\theta}{2}\right)(b(1-\gamma) + \gamma) + \left(\frac{1-\theta}{2}\right)[\mu_m(HH) + b\mu_m(DH) + \mu_m(H1) + \mu_m(I)]$$

$$v_M(DH) = \frac{\theta}{2}(b+1) + \left(\frac{1-\theta}{2}\right) + \left(\frac{1-\theta}{2}\right)[\mu_m(HH) + b\mu_m(DH) + \mu_m(H1) + \mu_m(I)]$$
$$= \frac{1}{2} + \frac{\theta b}{2} + \left(\frac{1-\theta}{2}\right)\left(\frac{\theta}{2}(1-\gamma)b + \left(1 - (1-\gamma)\frac{\theta}{2}\right)\right)$$

and likewise

$$v_m(I) = \left(\frac{1-\theta}{2}\right)(b+1) + \frac{\theta}{2}\left(b(1-\gamma)+\gamma\right) + \frac{\theta}{2}[\mu_M(HH) + b\mu_M(DH) + \mu_M(H1) + \mu_M(I)]$$

$$\begin{aligned} v_m(DH) &= \left(\frac{1-\theta}{2}\right)(b+1) + \frac{\theta}{2} + \frac{\theta}{2}[\mu_M(HH) + b\mu_M(DH) + \mu_M(H1) + \mu_M(I)] \\ &= \frac{1}{2} + \left(\frac{1-\theta}{2}\right)b + \frac{\theta}{2}\left(\left(\frac{1-\theta}{2}\right)(1-\gamma)b + \left(1-(1-\gamma)\left(\frac{1-\theta}{2}\right)\right)\right) \end{aligned}$$

so that

$$\begin{aligned} E^M - E^m &= \left(\frac{2\theta-1}{2}\right)b\gamma + v_M(DH) - v_m(DH) - \left(\frac{2\theta-1}{4}\right)(b-1)(1-\gamma) \\ &= \left(\frac{2\theta-1}{2}\right)\left(b\gamma + (b-1) - \frac{1}{2}(b-1)(1-\gamma)\right) \end{aligned}$$

which is always positive.

Consider the $SPBE$ as specified by Proposition 5 (iv). The expected values of the majority and the minority are:

$$\begin{aligned} E^M &= \theta[\left(\frac{b+1}{2}\right) + v_M(I)] + \left(\frac{1-\theta}{2}\right)[1 + v_M(D1)] \\ &\quad + \left(\frac{1-\theta}{2}\right)\gamma[1 + v_M(DH)] + \left(\frac{1-\theta}{2}\right)(1-\gamma)[b + v_M(E)] \end{aligned}$$

and

$$E^m = (1-\theta)[\left(\frac{b+1}{2}\right) + v_m(I)] + \frac{\theta}{2}[1 + v_m(D1)] + \frac{\theta}{2}\gamma[1 + v_m(DH)] + \frac{\theta}{2}(1-\gamma)[b + v_m(E)]$$

Further simplification can be done as $v_i(I) = v_i(E) > v_i(DH) = v_i(D1)$ for both groups $i \in \{M, m\}$ as $\gamma < \frac{b-1}{b}$. This implies that

$$\begin{aligned} & E^M - E^m \\ = & \left(\frac{2\theta-1}{2}\right)(b-1)\gamma + \left(\theta + \left(\frac{1-\theta}{2}\right)(1-\gamma)\right)v_M(I) + \left(\left(\frac{1-\theta}{2}\right)(1+\gamma)\,v_M(D1)\right) \\ & - \left(1-\theta+\frac{\theta}{2}(1-\gamma)\right)v_m(I) - \left(\frac{\theta}{2}(1+\gamma)\right)v_m(D1) \end{aligned}$$

where

$$\begin{aligned} v_M(I) &= \frac{\theta}{2}(b+1) + \left(\frac{1-\theta}{2}\right)\left(b(1-\gamma)+\gamma\right) \\ &\quad + \left(\frac{1-\theta}{2}\right)[b\left(\mu_m(D1) + \mu_m(DH)\right) + \left(1 - \mu_m(D1) - \mu_m(DH)\right)] \end{aligned}$$

$$\begin{aligned} v_M(D1) &= \frac{\theta}{2}(b+1) + \left(\frac{1-\theta}{2}\right) \\ &+ \left(\frac{1-\theta}{2}\right)\left[b\left(\mu_m(D1)+\mu_m(DH)\right) + (1-\mu_m(D1)-\mu_m(DH))\right] \\ &= \frac{1}{2} + \frac{\theta b}{2} + \left(\frac{1-\theta}{2}\right)(b\theta(1-\gamma) + (1-\theta(1-\gamma))) \end{aligned}$$

and likewise

$$\begin{aligned} v_m(I) &= \left(\frac{1-\theta}{2}\right)(b+1) + \frac{\theta}{2}(b(1-\gamma)+\gamma) \\ &+ \frac{\theta}{2}\left[b\left(\mu_M(D1)+\mu_M(DH)\right) + (1-\mu_M(D1)-\mu_M(DH))\right] \end{aligned}$$

$$\begin{aligned} v_m(D1) &= \left(\frac{1-\theta}{2}\right)(b+1) + \frac{\theta}{2} + \frac{\theta}{2}\left[b\left(\mu_M(D1)+\mu_M(DH)\right) + (1-\mu_M(D1)-\mu_M(DH))\right] \\ &= \frac{1}{2} + \left(\frac{1-\theta}{2}\right)b + \frac{\theta}{2}(b(1-\theta)(1-\gamma) + (1-(1-\theta)(1-\gamma))) \end{aligned}$$

so that

$$\begin{aligned} E^M - E^m &= \left(\frac{2\theta-1}{2}\right)(b-1)\gamma + v_M(D1) - v_m(D1) - \left(\frac{2\theta-1}{4}\right)(b-1)(1-\gamma) \\ &= \left(\frac{2\theta-1}{2}\right)\left((b-1)(1+\gamma) - \frac{1}{2}(b-1)(1-\gamma)\right) \end{aligned}$$

which is always positive.

Consider the $SPBE$ as specified by Proposition 7. The expected values of the majority and minority are:

$$E^M = \theta[\left(\frac{b+1}{2}\right) + v_M(I)] + \left(\frac{1-\theta}{2}\right)[(1-\sigma_1(H))(1-\gamma)b + v_M(H1)] + \left(\frac{1-\theta}{2}\right)[1+v_M(DH)]$$

and

$$E^m = (1-\theta)[\left(\frac{b+1}{2}\right) + v_m(I)] + \frac{\theta}{2}[(1-\sigma_1(H))(1-\gamma)b + v_m(H1)] + \frac{\theta}{2}[1+v_m(DH)]$$

Note that we have already accounted for the necessary condition for mixing: $1+v_i(DH) = 0+v_i(HH)$ for both groups. Further simplification can be done as $v_i(I) = v_i(H1) = v_i(DH) = v_i$ for both groups $i \in \{M, m\}$ where $\gamma < \frac{b-1}{b}$. This implies that

$$E^M - E^m = \left(\frac{2\theta-1}{2}\right)\left[\frac{b^2\gamma}{b-1}\right] + v_M - v_m$$

where

$$\begin{aligned} v_M &= \frac{\theta}{2}(b+1) + \left(\frac{1-\theta}{2}\right)[1+\mu_m(HH)+b\mu_m(DH)+b(1-\gamma)\mu_m(H1)+b(1-\gamma)\mu_m(I)] \\ &= \frac{1}{2}+\frac{\theta b}{2}+\left(\frac{1-\theta}{2}\right)b(1-\gamma) \end{aligned}$$

and

$$\begin{aligned} v_m &= \left(\frac{1-\theta}{2}\right)(b+1)+\frac{\theta}{2}[1+\mu_M(HH)+b\mu_M(DH)+b(1-\gamma)\mu_M(H1)+b(1-\gamma)\mu_M(I)] \\ &= \frac{1}{2}+\frac{(1-\theta)b}{2}+\frac{\theta}{2}b(1-\gamma) \end{aligned}$$

so that

$$E^M - E^m = \left(\frac{2\theta-1}{2}\right)\left(\left(\frac{b^2\gamma}{b-1}\right)+b\gamma\right)$$

which is always positive.

■

Proof of Proposition 2: We begin by showing that $\sigma_1^M(H) \leq \sigma_1^m(H)$. Suppose there was an $SPBE$ such that

$$\sigma_1^M(H) > \sigma_1^m(H)$$

First note that by Lemma 4, if $\sigma_1^i(H) > 0$ then $\gamma < \left(\frac{b-1}{b}\right)^2$. If $\sigma_1^i(H) = 1$ and $\gamma < \frac{b-1}{b}$ then there is no benefit to foregoing payment in the first period because $p^i(HH) = \gamma < \frac{b-1}{b}$. Furthermore, arguments advanced in the Proof of Proposition 6 show that if $\sigma_1^i(H) \in (0,1)$ then it must be that $\sigma^i(H) = \beta^*$ such that $p^i(HH) = \frac{b-1}{b}$. Therefore, $\sigma_1^i(H) \in \{0, \beta^*\}$. To satisfy the inequality it must be that $\sigma_1^M(H) = \beta^* > \sigma_1^m(H) = 0$. In order to support this $SPBE$ it must be that

$$1 = v_M(HH) - v_M(DH)$$

and by Corollary 1

$$1 = \left(\frac{1-\theta}{2}\right)(b-1)(1-\gamma)(1-\sigma_2^m(HH))$$

It must also be that

$$\begin{aligned} 1 &> v_m(HH) - v_m(DH) \\ 1 &> \left(\frac{\theta}{2}\right)(b-1)(1-\gamma) \end{aligned}$$

This is a contradiction as

$$\left(\frac{1-\theta}{2}\right)(b-1)(1-\gamma)(1-\sigma_2^m(HH)) < \frac{\theta}{2}(b-1)(1-\gamma)$$

and so it is proved that $\sigma_1^M(H) \leq \sigma_1^m(H)$.

Now we show that $\sigma_1^M(1) \leq \sigma_1^m(1)$. By way of contradiction, suppose that:

$$\sigma_1^M(1) > \sigma_1^m(1)$$

In the case that $\gamma > \frac{b-1}{b}$, for all $\sigma_1^i(H) \in [0,1]$, it must be that $p^i(HH) > \frac{b-1}{b}$ and so $\sigma_2^j(HH) = 1$. Therefore, in order for $\sigma_1^i(H) \in (0,1)$, it must be that[17]

$$b(1-\gamma) + v_i(H1) = 1 + v_i(D1)$$

This condition only holds for $\overline{\gamma}_M$ in the case of the majority and $\overline{\gamma}_m$ in the case of the minority. Since we are restricting attention to generic parameters, we can exclude $\sigma_1^i(H) \in (0,1)$. Therefore the only remaining case for $\gamma > \frac{b-1}{b}$ is: $1 = \sigma_1^M(1) > \sigma_1^m(1) = 0$. This implies that

$$\begin{aligned} b(1-\gamma) + v_M(H1) &> 1 + v_M(D1) \\ b(1-\gamma) + v_m(H1) &< 1 + v_m(D1) \end{aligned}$$

and so

$$\left(\frac{\theta}{2}\right)(b-1)(1-\gamma) < 1 - b(1-\gamma) < \left(\frac{1-\theta}{2}\right)(b-1)(1-\gamma)$$

This is a contradiction and so for $\gamma > \frac{b-1}{b}$, it must be that $\sigma_1^M(1) \leq \sigma_1^m(1)$.

In the case that $\gamma < \frac{b-1}{b}$, $\sigma_1^i(H)$ will affect $\sigma_2^j(HH)$. We investigate $\sigma_1^i(H) \in (0,\alpha^*) \cup \{\alpha^*\} \cup (\alpha^*,1)$ where $\alpha^* = \left(\frac{\gamma}{(1-\gamma)(b-1)}\right)$ which implies $p^i(HH) = \frac{b-1}{b}$. In order for i to mix, it must be that:

$$b(1-\gamma)(1-\sigma_1^j(H)) + v_i(H1) = 1 + v_i(D1) \tag{8}$$

It must be that $v_i(H1) \geq v_i(D1)$. Since $\gamma < \frac{b-1}{b}$, expression (8) only holds when $\sigma_1^j(H) > 0$. However, since $\sigma_1^j(H)$ only takes one nonzero value: $\frac{\gamma}{(1-\gamma)(b-1)}$. Since a player is displaying $P2$, by Lemma 4 it must be that $\gamma < \left(\frac{b-1}{b}\right)^2$. However, $b\left(1 - \frac{b\gamma}{b-1}\right) = 1$ is not satisfied by any $\gamma < \left(\frac{b-1}{b}\right)^2$ therefore $b\left(1 - \frac{b\gamma}{b-1}\right) + v_i(H1) = 1 + v_i(D1)$ cannot be satisfied by any $\gamma < \left(\frac{b-1}{b}\right)^2$. Therefore, the only remaining case for $\gamma < \frac{b-1}{b}$ is: $1 = \sigma_1^M(1) > \sigma_1^m(1) = 0$. In this case, $v_M(H1) = v_M(D1)$ as $p^M(H1) = \gamma < \frac{b-1}{b}$. A deviation of m would imply $p^m(H1) = 1$ and therefore, $v_m(H1) > v_m(D1)$. This leads to a contradiction as it cannot be that

$$b(1-\gamma) > 1$$

and

$$b(1-\gamma) + v_m(H1) < 1 + v_m(D1)$$

Therefore, $\sigma_1^M(1) \leq \sigma_1^m(1)$ for generic parameter values.■

Proof of Proposition 4: The proof entails calculating the value of expression (9) for

[17] Note that since $\gamma > \frac{b-1}{b}$ no player displays $P2$.

every possible value of b.

$$\frac{\mathcal{I}}{b+1} = \theta(1-\theta)[\gamma+(1-\gamma)\sigma_1^M(1)][\gamma+(1-\gamma)\sigma_1^m(H)] \tag{9}$$
$$+\theta(1-\theta)[\gamma+(1-\gamma)\sigma_1^m(1)][\gamma+(1-\gamma)\sigma_1^M(H)]$$
$$+\theta(1-\theta)\sum_{h^m\in\mathcal{H}}[\gamma+(1-\gamma)\sigma_2^M(h^m)]p^m(h^m)\mu^m(h^m)$$
$$+\theta(1-\theta)\sum_{h^M\in\mathcal{H}}[\gamma+(1-\gamma)\sigma_2^m(h^M)]p^M(h^M)\mu^M(h^M)$$

Given any values of θ and γ, there exists a b which induces an $SPBE$ in each of the regions specified in Appendix B. Therefore, we calculate $\mathcal{I}$ in each possible $SPBE$. Furthermore, as b increases, the $SPBE$ will pass through regions specified by Proposition 5 (iv), (iii), (ii), (i), Proposition 6 and Proposition 7. Recall that $\mu(h)$ denotes the distribution of histories.

For the $SPBE$ as in Proposition 5 (iv):

$$\mu^m(I) = 1-\theta \text{ and } p^m(I)=\gamma$$
$$\mu^m(H1) = \frac{\theta}{2}\gamma \text{ and } p^m(H1)=1$$
$$\mu^m(D1) = \frac{\theta}{2}(1-\gamma) \text{ and } p^m(D1)=0$$
$$\mu^m(HH) = \frac{\theta}{2}\gamma^2 \text{ and } p^m(HH)=1$$
$$\mu^m(DH) = \frac{\theta}{2}\gamma(1-\gamma) \text{ and } p^m(DH)=0$$
$$\mu^m(E) = \frac{\theta}{2}(1-\gamma) \text{ and } p^m(E)=\gamma$$

with similar values for M with θ exchanged for $1-\theta$. Expression (9) implies

$$\frac{\mathcal{I}}{b+1} = 2\theta(1-\theta)\gamma^2+\theta(1-\theta)[(1-\theta)+\frac{\theta}{2}(1-\gamma)]\gamma^2+\theta(1-\theta)[\frac{\theta}{2}\gamma+\frac{\theta}{2}\gamma^2]\gamma$$
$$+\theta(1-\theta)[\theta+\left(\frac{1-\theta}{2}\right)(1-\gamma)]\gamma^2+\theta(1-\theta)[\left(\frac{1-\theta}{2}\right)\gamma+\left(\frac{1-\theta}{2}\right)\gamma^2]\gamma$$

Therefore

$$\frac{\mathcal{I}}{b+1} = 4\theta(1-\theta)\gamma^2 \text{ for } b\in(1,\frac{1+\frac{\theta}{2}(1-\gamma)}{1-\gamma+\frac{\theta}{2}(1-\gamma)})$$

For the $SPBE$ as in Proposition 5 (iii):

$$\begin{aligned}
\mu^m(I) &= 1-\theta \text{ and } p^m(I)=\gamma \\
\mu^m(H1) &= \frac{\theta}{2} \text{ and } p^m(H1)=\gamma \\
\mu^m(D1) &= 0 \text{ and } p^m(D1)=0 \\
\mu^m(HH) &= \frac{\theta}{2}\gamma^2 \text{ and } p^m(HH)=1 \\
\mu^m(DH) &= \frac{\theta}{2}\gamma(1-\gamma) \text{ and } p^m(DH)=0 \\
\mu^m(E) &= \frac{\theta}{2}(1-\gamma) \text{ and } p^m(E)=\gamma
\end{aligned}$$

and:

$$\begin{aligned}
\mu^M(I) &= \theta \text{ and } p^M(I)=\gamma \\
\mu^M(H1) &= \left(\frac{1-\theta}{2}\right)\gamma \text{ and } p^M(H1)=1 \\
\mu^M(D1) &= \left(\frac{1-\theta}{2}\right)(1-\gamma) \text{ and } p^M(D1)=0 \\
\mu^M(HH) &= \left(\frac{1-\theta}{2}\right)\gamma \text{ and } p^M(HH)=1 \\
\mu^M(DH) &= \left(\frac{1-\theta}{2}\right)(1-\gamma) \text{ and } p^M(DH)=0 \\
\mu^M(E) &= 0 \text{ and } p^M(E)=\gamma
\end{aligned}$$

Expression (9) implies:

$$\begin{aligned}
\frac{\mathcal{I}}{b+1} &= \theta(1-\theta)\gamma^2+\theta(1-\theta)\gamma+\theta(1-\theta)[\theta]\gamma^2+\theta(1-\theta)[\left(\frac{1-\theta}{2}\right)\gamma+\left(\frac{1-\theta}{2}\right)\gamma]\gamma \\
&\quad +\theta(1-\theta)[(1-\theta)+\frac{\theta}{2}+\frac{\theta}{2}(1-\gamma)]\gamma^2+\theta(1-\theta)[\frac{\theta}{2}\gamma^2]\gamma
\end{aligned}$$

Therefore:

$$\frac{\mathcal{I}}{b+1}=\theta(1-\theta)\gamma(1+3\gamma) \text{ for } b\in(\frac{1+\frac{\theta}{2}(1-\gamma)}{1-\gamma+\frac{\theta}{2}(1-\gamma)},\frac{1+\left(\frac{1-\theta}{2}\right)(1-\gamma)}{1-\gamma+\left(\frac{1-\theta}{2}\right)(1-\gamma)})$$

and by Proposition 8 at $b=\frac{1+\frac{\theta}{2}(1-\gamma)}{1-\gamma+\frac{\theta}{2}(1-\gamma)}$:

$$\mathcal{I}\in[(b+1)4\theta(1-\theta)\gamma^2,(b+1)\theta(1-\theta)\gamma(1+3\gamma)]$$

For the $SPBE$ as in Proposition 5 (ii):

$$\begin{aligned}
\mu^m(I) &= 1-\theta \text{ and } p^m(I)=\gamma \\
\mu^m(H1) &= \frac{\theta}{2} \text{ and } p^m(H1)=\gamma \\
\mu^m(D1) &= 0 \text{ and } p^m(D1)=0 \\
\mu^m(HH) &= \frac{\theta}{2}\gamma \text{ and } p^m(HH)=1 \\
\mu^m(DH) &= \frac{\theta}{2}(1-\gamma) \text{ and } p^m(DH)=0 \\
\mu^m(E) &= 0 \text{ and } p^m(E)=\gamma
\end{aligned}$$

with similar values for M with θ exchanged for $1-\theta$. Expression (9) then implies:

$$\begin{aligned}
\frac{\mathcal{I}}{b+1} &= 2\theta(1-\theta)\gamma+\theta(1-\theta)[(1-\theta)+\frac{\theta}{2}]\gamma^2+\theta(1-\theta)[\frac{\theta}{2}\gamma]\gamma \\
&\quad +\theta(1-\theta)[\theta+\left(\frac{1-\theta}{2}\right)]\gamma^2+\theta(1-\theta)[\left(\frac{1-\theta}{2}\right)\gamma]\gamma
\end{aligned}$$

Therefore:

$$\frac{\mathcal{I}}{b+1}=2\theta(1-\theta)\gamma(1+\gamma) \text{ for } b\in(\frac{1+\left(\frac{1-\theta}{2}\right)(1-\gamma)}{1-\gamma+\left(\frac{1-\theta}{2}\right)(1-\gamma)},\frac{1}{1-\gamma})$$

and by Proposition 8 at $b=\frac{1+\left(\frac{1-\theta}{2}\right)(1-\gamma)}{1-\gamma+\left(\frac{1-\theta}{2}\right)(1-\gamma)}$:

$$\mathcal{I}\in[(b+1)\theta(1-\theta)\gamma(1+3\gamma),(b+1)2\theta(1-\theta)\gamma(1+\gamma)]$$

For the $SPBE$ as in Proposition 5 (i):

$$\begin{aligned}
\mu^m(I) &= 1-\theta \text{ and } p^m(I)=\gamma \\
\mu^m(H1) &= \frac{\theta}{2} \text{ and } p^m(H1)=\gamma \\
\mu^m(D1) &= 0 \text{ and } p^m(D1)=0 \\
\mu^m(HH) &= \frac{\theta}{2}\gamma \text{ and } p^m(HH)=1 \\
\mu^m(DH) &= \frac{\theta}{2}(1-\gamma) \text{ and } p^m(DH)=0 \\
\mu^m(E) &= 0 \text{ and } p^m(E)=\gamma
\end{aligned}$$

with similar values for M with θ exchanged for $1-\theta$. Expression (9) then implies:

$$\begin{aligned}
\frac{\mathcal{I}}{b+1} &= 2\theta(1-\theta)\gamma+\theta(1-\theta)[(1-\theta)+\frac{\theta}{2}]\gamma+\theta(1-\theta)[\frac{\theta}{2}\gamma]\gamma \\
&\quad +\theta(1-\theta)[\theta+\left(\frac{1-\theta}{2}\right)]\gamma+\theta(1-\theta)[\left(\frac{1-\theta}{2}\right)\gamma]\gamma
\end{aligned}$$

Therefore:
$$\frac{\mathcal{I}}{b+1} = \theta(1-\theta)\gamma(3.5+0.5\gamma) \text{ for } b \in (\frac{1}{1-\gamma}, \frac{2}{\theta(1-\gamma)}+1]$$

and by Proposition 8 at $b = \frac{1}{1-\gamma}$:

$$\mathcal{I} \in [(b+1)2\theta(1-\theta)\gamma(1+\gamma), (b+1)\theta(1-\theta)\gamma(3.5+0.5\gamma)]$$

For the $SPBE$ as in Proposition 6:

$$\begin{aligned}
\mu^m(I) &= 1-\theta \text{ and } p^m(I) = \gamma \\
\mu^m(H1) &= \frac{\theta}{2} \text{ and } p^m(H1) = \gamma \\
\mu^m(D1) &= 0 \text{ and } p^m(D1) = 0 \\
\mu^m(HH) &= \frac{\theta}{2}\left(\frac{b}{b-1}\gamma\right) \text{ and } p^m(HH) = \frac{b-1}{b} \\
\mu^m(DH) &= \frac{\theta}{2}\left(\frac{b-1-b\gamma}{b-1}\right) \text{ and } p^m(DH) = 0 \\
\mu^m(E) &= 0 \text{ and } p^m(E) = \gamma
\end{aligned}$$

and:

$$\begin{aligned}
\mu^M(I) &= \theta \text{ and } p^M(I) = \gamma \\
\mu^M(H1) &= \left(\frac{1-\theta}{2}\right) \text{ and } p^M(H1) = \gamma \\
\mu^M(D1) &= 0 \text{ and } p^M(D1) = 0 \\
\mu^M(HH) &= \left(\frac{1-\theta}{2}\right)\gamma \text{ and } p^M(HH) = 1 \\
\mu^M(DH) &= \left(\frac{1-\theta}{2}\right)(1-\gamma) \text{ and } p^M(DH) = 0 \\
\mu^M(E) &= 0 \text{ and } p^M(E) = \gamma
\end{aligned}$$

Expression (9) implies:

$$\begin{aligned}
\frac{\mathcal{I}}{b+1} &= \theta(1-\theta)[\gamma+(1-\gamma)\left(\frac{\gamma}{(1-\gamma)(b-1)}\right)] + \theta(1-\theta)\gamma + \theta(1-\theta)[(1-\theta)+\frac{\theta}{2}]\gamma \\
&+\theta(1-\theta)[\frac{\theta}{2}\left(\frac{b}{b-1}\gamma\right)]\left(\frac{\frac{\theta}{2}(b-1)-1}{\frac{\theta}{2}(b-1)}\right)\left(\frac{b-1}{b}\right) + \theta(1-\theta)[\theta+\left(\frac{1-\theta}{2}\right)]\gamma \\
&+\theta(1-\theta)[\left(\frac{1-\theta}{2}\right)\gamma]\gamma
\end{aligned}$$

Therefore

$$\frac{\mathcal{I}}{b+1} = \theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma) \text{ for } b \in (\frac{2}{\theta(1-\gamma)}+1, \frac{2}{(1-\theta)(1-\gamma)}+1)$$

and by Proposition 8 at $b = \frac{2}{\theta(1-\gamma)} + 1:$

$$\mathcal{I} \in \{(b+1)\theta(1-\theta)\gamma(3.5+0.5\gamma), (b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma)\}$$

For the $SPBE$ as in Proposition 7:

$$\begin{aligned}
\mu^m(I) &= 1-\theta \text{ and } p^m(I) = \gamma \\
\mu^m(H1) &= \frac{\theta}{2} \text{ and } p^m(H1) = \gamma \\
\mu^m(D1) &= 0 \text{ and } p^m(D1) = 0 \\
\mu^m(HH) &= \frac{\theta}{2}\left(\frac{b}{b-1}\gamma\right) \text{ and } p^m(HH) = \frac{b-1}{b} \\
\mu^m(DH) &= \frac{\theta}{2}\left(\frac{b-1-b\gamma}{b-1}\right) \text{ and } p^m(DH) = 0 \\
\mu^m(E) &= 0 \text{ and } p^m(E) = \gamma
\end{aligned}$$

with similar values for M with θ exchanged for $1-\theta$. Expression (9) then implies:

$$\begin{aligned}
\frac{\mathcal{I}}{b+1} &= 2\theta(1-\theta)\left(\frac{b}{b-1}\gamma\right) + \theta(1-\theta)[(1-\theta)+\frac{\theta}{2}]\gamma \\
&\quad +\theta(1-\theta)[\frac{\theta}{2}\left(\frac{b}{b-1}\gamma\right)]\left(\frac{\frac{\theta}{2}(b-1)-1}{\frac{\theta}{2}(b-1)}\right)\left(\frac{b-1}{b}\right) + \theta(1-\theta)[\theta+\left(\frac{1-\theta}{2}\right)]\gamma \\
&\quad +\theta(1-\theta)[\left(\frac{1-\theta}{2}\right)\left(\frac{b}{b-1}\gamma\right)]\left(\frac{\left(\frac{1-\theta}{2}\right)(b-1)-1}{\left(\frac{1-\theta}{2}\right)(b-1)}\right)\left(\frac{b-1}{b}\right)
\end{aligned}$$

Therefore

$$\frac{\mathcal{I}}{b+1} = 4\theta(1-\theta)\gamma \text{ for } b > \frac{2}{(1-\theta)(1-\gamma)} + 1$$

and by Proposition 8 at $b = \frac{2}{(1-\theta)(1-\gamma)} + 1$:

$$\mathcal{I} \in \{(b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma), (b+1)4\theta(1-\theta)\gamma\}$$

Therefore $\frac{\mathcal{I}}{b+1}$ is increasing in b. It follows that $\mathcal{I}$ is strictly increasing in b and so the proposition is proved.■

Proof of Proposition 3: For every set of parameter values (b, θ, γ), the statement of Proposition 4 maps to the corresponding values of $\mathcal{I}$. Therefore in the proof of Proposition 3, we note the trajectory of $\mathcal{I}$, given b and γ, as θ varies. As $1-\theta$ changes, the incentives for each group changes. Specifically, as $1-\theta$ gets larger, the minority reputation becomes less valuable and the majority reputation becomes more valuable. As $1-\theta$ becomes large one of the following three possibilities occur. In the first case, no qualitative change occurs in the $SPBE$. In the second case, the majority does not exhibit reputation whereas the minority exhibits reputation for small $1-\theta$ and for large values does not exhibit reputation. In the third case, the minority always exhibits reputation and for small $1-\theta$ the majority does not display reputation and for large values, the majority does display reputation.

Now we characterize the relationship between $\mathcal{I}$ and $1-\theta$ for every pair of (b,γ). If $b \leq \frac{2+(1-\gamma)}{3(1-\gamma)}$, then for all values of $1-\theta$, it will be that $\mathcal{I} = (b+1)\theta(1-\theta)[4\gamma^2]$. This implies that for values of (b,γ) in this region $\mathcal{I}$ is strictly increasing and continuous in $1-\theta$. Therefore $1-\theta^* = 0.5$.

If $b \in (\frac{2+(1-\gamma)}{3(1-\gamma)}, \frac{4+(1-\gamma)}{5(1-\gamma)})$ then for small values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)[\gamma(1+3\gamma)]$ and for large values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)[4\gamma^2]$. Intuitively, for small $1-\theta$ the minority exhibits $P1$. However, for large $1-\theta$, it is no longer profitable for the minority to exhibit $P1$. This downward discontinuity occurs at $1-\theta$ such that $b = \frac{2+\theta(1-\gamma)}{(2+\theta)(1-\gamma)}$. Note that at this downward discontinuity the minority is indifferent between displaying $P1$ or not. Therefore, $\mathcal{I} \in [(b+1)\theta(1-\theta)[4\gamma^2], (b+1)\theta(1-\theta)[\gamma(1+3\gamma)]]$ at $1-\theta$ where $b = \frac{2+\theta(1-\gamma)}{(2+\theta)(1-\gamma)}$. Hence, $1-\theta^*$ is where $b = \frac{2+\theta(1-\gamma)}{(2+\theta)(1-\gamma)}$ and this is strictly larger than zero.

If $b = \frac{4+(1-\gamma)}{5(1-\gamma)}$ then $\mathcal{I} = (b+1)\theta(1-\theta)[\gamma(1+3\gamma)]$ for all values of $1-\theta$. This implies that for values of (b,γ) such that $b = \frac{4+(1-\gamma)}{5(1-\gamma)}$ then $\mathcal{I}$ is strictly increasing and continuous in $1-\theta$. Therefore, $1-\theta^* = 0.5$.

If $b \in (\frac{4+(1-\gamma)}{5(1-\gamma)}, \frac{1}{1-\gamma})$ then for small values $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)[\gamma(1+3\gamma)]$ and for large values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)2\gamma(1+\gamma)$. Intuitively, for small $1-\theta$ the majority does not exhibit $P1$ however for large $1-\theta$ the reputation of the majority becomes sufficiently profitable to display $P1$. This upward discontinuity occurs at $1-\theta$ such that $b = \frac{2+(1-\theta)(1-\gamma)}{(3-\theta)(1-\gamma)}$. Note that at this discontinuity, the majority is indifferent between displaying $P1$ or not. Thus, $\mathcal{I} \in [(b+1)\theta(1-\theta)[\gamma(1+3\gamma)], (b+1)\theta(1-\theta)2\gamma(1+\gamma)]$ at $1-\theta$ where $b = \frac{2+(1-\theta)(1-\gamma)}{(3-\theta)(1-\gamma)}$. As there is a single upward discontinuity and is increasing at every point of continuity therefore $1-\theta^* = 0.5$.

If $b = \frac{1}{1-\gamma}$ then for all values of $1-\theta$ it will be that $\mathcal{I} \in [(b+1)\theta(1-\theta)[\gamma(1+3\gamma)], (b+1)\theta(1-\theta)\gamma(3.5+0.5\gamma)]$. Note that for these particular values of b and γ any value of $\mathcal{I}$ in the above specified region will suffice. However, given any second period strategies for the histories I, $H1$ or E, inefficiency is increasing and continuous in $1-\theta$. Therefore, $1-\theta^* = 0.5$.

If $b \in (\frac{1}{1-\gamma}, \frac{2}{1-\gamma}+1]$ then for all values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)\gamma(3.5+0.5\gamma)$. This implies that for values of (b,γ) in this region $\mathcal{I}$ is strictly increasing and continuous in $1-\theta$. Therefore, $1-\theta^* = 0.5$.

If $b \in (\frac{2}{1-\gamma}+1, \frac{4}{1-\gamma}+1)$ then for small values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma)$ and for large values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)\gamma(3.5+0.5\gamma)$. Intuitively, for small $1-\theta$ the minority exhibits $P2$ and for large $1-\theta$ the minority does not exhibit $P2$. This boundary occurs at $1-\theta \in (0, 0.5)$ such that $b = \frac{2}{\theta(1-\gamma)}+1$. Although the minority is indifferent between exhibiting $P2$ or not, it is not the case that any combination will suffice. Therefore, at $1-\theta''$ where $b = \frac{2}{\theta(1-\gamma)} + 1$, the minority either exhibits $P2$ or not: $\mathcal{I} \in \{(b+1)\theta(1-\theta)\gamma(3.5+0.5\gamma), (b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma)\}$. Due to the particular behavior of $(b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma)$ we denote its interior maximum as $1-\theta' = \frac{9-\gamma-\sqrt{\gamma^2+6\gamma+57}}{3(1-\gamma)}$. The quantity $1-\theta'$ is increasing from 0.4833 when $\gamma = 0$ to 0.5 when $\gamma = 1$. Therefore, $1-\theta^* = \min\{1-\theta', 1-\theta''\}$ and this is bounded away from zero.

If $b = \frac{4}{1-\gamma}+1$ then for all values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma)$. This implies that for values of b and γ in this region $\mathcal{I}$ is strictly increasing and continuous in $1-\theta$. Therefore, $1-\theta^* = 1-\theta'$

If $b \in (\frac{4}{1-\gamma}+1, \infty)$ then for small values of $1-\theta$ it will be that $\mathcal{I} = (b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}+\left(\frac{1-\theta}{2}\right)\gamma)$ and for large values of $1-\theta$ it will be that $(b+1)\theta(1-\theta)4\gamma$. Intuitively, for small $1-\theta$ the majority does not find it profitable to exhibit $P2$ however for large $1-\theta$ the reputation of the majority becomes sufficiently profitable. This upward discontinuity occurs at $1-\theta$ such that $b = \frac{2}{(1-\theta)(1-\gamma)}+1$. Although the majority is indifferent between exhibiting $P2$ or not, it is not the case that any combination will suffice. Therefore, the majority either exhibits $P2$ or not: $\mathcal{I} \in \{(b+1)\theta(1-\theta)\gamma(3.5+\frac{\theta}{2}), (b+1)\theta(1-\theta)4\gamma\}$ at $1-\theta$ where $b = \frac{2}{(1-\theta)(1-\gamma)}+1$. Therefore, $1-\theta^* = 1-\theta'$.

Therefore, for every value of (b,γ) there exists $1-\theta^* > 0$ such that for all $1-\theta < 1-\theta^*$, inefficiency $\mathcal{I}$ is increasing in $1-\theta$.

■

7 References

Akerlof, George and Kranton, Rachel (2000): "Economics and Identity," *Quarterly Journal of Economics*, 115(3), 715-753.

Alesina, Alberto, Baqir, Reza and Easterly, William (1999): "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114(4), 1243-1284.

Arrow, Kenneth (1973): "The Theory of Discrimination," In Orley Ashenfelter and Albert Rees (Eds.) *Discrimination in Labor Markets*, Princeton, NJ, Princeton University Press, 3-33.

Basu, Kaushik (2005): "Racial Conflict and the Malignancy of Identity," *Journal of Economic Inequality*, 3, 221-241.

Cho, Bongsoon and Connelley, Debra (2002): "The Effect of Conflict and Power Differentials on Social Identity and Intergroup Discrimination," AoM Conflict Management Division 2002 Mtgs. No. 12522. Available at SSRN: http://ssrn.com/abstract=320286.

Davis, John (2005): "Social Identity Strategies in Recent Economics," Tinbergen Institute Discussion Paper TI 2005-078/2.

Dawes, Robyn, Van De Kragt, Alphons, and Orbell, John (1988): "Not Me or Thee But We: The Importance of Group Identity in Eliciting Cooperation in Dilemma Situations: Experimental Manipulations," *Acta Psychologica*, 68, 83-97.

Easterly, William and Levine, Ross (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 112(4), 1203-1250.

Esteban, Joan-Maria and Ray, Debraj (1994): "On the Measurement of Polarization," *Econometrica*, 62(4), 819-851.

Falk, Armin and Zweimuller, Josef (2005): "Unemployment and Right-Wing Extremist Crime," unpublished IZA and Bonn.

Fearon, James and Laitin, David (1996): "Explaining Interethnic Cooperation," *American Political Science Review*, 90(4), 715-735.

Green, Donald and Seher, Rachel (2003): "What Role Does Prejudice Play in Ethnic Conflict?" *Annual Review of Political Science*, 6, 509-531.

Gurin, Patricia, Peng, Timothy, Lopez, Gretchen and Nagda, Biren (1999): "Context, Identity and Intergroup Relations," in Deborah Prentice and Dale Miller (Eds.), *Cultural*

Divides: Understanding and Overcoming Group Conflict, New York, Russel Sage Foundation, 133-170.

Kirman, Alan and Teschl, Miriam (2004): "On the Emergence of Economic Identity," *Revue De Philosophie Economique*, 9, 59-86.

Kramer, Roderick and Brewer, Marilynn (1984): "Effects of Group Identity on Resource Utilization in a Simulated Commons Dilemma," *Journal of Personality and Social Psychology*, 46, 1044-1057.

Kreps, David and Wilson, Robert (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

Lubbers, Marcel and Scheepers, Peer (2001): "Explaining the Trend in Extreme Right-Wing Voting: Germany 1989-1998," *European Sociological Review*, 17(4), 431-449.

Mauro, Paulo (1995): "Corruption and Growth," *Quarterly Journal of Economics*, 110(3), 681-712.

Messick, David and Mackie, Diane (1989): "Intergroup Relations," *Annual Review of Psychology*, 40, 45-81.

Murdock, George (1949): *Social Structure*. New York, Macmillan.

Nakao, Keisuke (2007): "Creation of Social Order in Ethnic Conflict," unpublished Boston University.

Olzak, Susan (1992): *Dynamics of Ethnic Competition and Conflict*, Palo Alto, CA, Stanford University Press.

Olzak, Susan, Shanahan, Suzanne and West, Elizabeth (1994): "School Desegregation, Interracial Exposure, and Antibusing Activity in Contemporary Urban America," *American Journal of Sociology*, 100(1), 196-241.

Phelps, Edmund (1972): "The Statistical Discrimination of Racism and Sexism," *American Economic Review,* 62, 659-661.

Quillian, Lincoln (1995): "Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-Immigrant and Racial Prejudice in Europe," *American Sociological Review*, 60(4), 586-611.

Quillian, Lincoln (1996): "Group Threat and Regional Change in Attitudes Toward African-Americans," *American Journal of Sociology*, 102(3), 816-860.

Rapoport, Hillel and Weiss, Avi (2003): "The optimal size for a minority," *Journal of Economic Behavior and Organization*, 52, 27-45.

Scheepers, Peer, Gijberts, Merove and Coenders, Marcel (2002): Ethnic Exclusion in European Countries: Public Opposition to Civil Rights for Legal Migrants as a Response to Perceived Ethnic Threat," *European Sociological Review*, 18(1), 17-34.

Shayo, Moses (2004): "A Theory of Social Identity with an Application to Redistribution," unpublished Hebrew University of Jerusalem.

Sherif, Muzafer, Harvey, O.J., White, B.Jack, Hood, William, and Sherif, Carolyn (1961): *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*, Norman, University of Oklahoma Book Exchange.

Silverman, Dan (2004): "Street Crime and Street Culture," *International Economic Review*, 45(3), 761-786.

Sobel, Joel (2005): "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, 43(2), 392-436.

Sumner, William (1906): *Folkways: A Study of the Sociological Importance of usages, manners, customs mores and morals.* Boston, Ginn.

Tajfel, Henri (1970): "Experiments in Intergroup Discrimination," *Scientific American*, 223, 96-102.

Tajfel, Henri (1978) *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations*, London, Academic Press.

Tajfel, Henri, Flament, Claude, Billig, Michael, and Bundy, R. (1971): "Social Categorization and Intergroup Behavior," *European Journal of Social Psychology*, 1, 149-178.

Tajfel, Henri and Turner, John (1979): "An Integrative Theory of Intergroup Conflict," in W.G.Austin and S. Worchel (Eds.) *The Social Psychology of Intergroup Relations*, Monterey, CA, Brooks/Cole, 33-48. Reprinted (2001) in Michael Hogg and Dominic Abrams (Eds.) *Intergroup Relations*, Ann Arbor, MI, Psychology Press, 94-109.

Tajfel, Henri and Turner, John (1986): "The Social Identity Theory of Intergroup Behavior," In Steven Worchel and William Austin (Eds.), *Psychology of Intergroup Relations*, Chicago, Nelson-Hall, 7-25.

Vigdor, Jacob (2002): "Interpreting Ethnic Fragmentation Effects," *Economics Letters*, 75, 271-276.