Interethnic conflict, cooperation, and oppression: A model of the effect of ingroup conformity on intergroup interactions

Joshua Gubler and Abe Gong[1]
Department of Political Science
University of Michigan, Ann Arbor
jgub@umich.edu / agong@umich.edu
24 March 2009

**Draft prepared for presentation at the Annual Conference of the Midwest Political Science Association, 2-5 April 2009. Please do not reference or cite without express written consent from the authors.

MPSA Participants: Thank you for reading and commenting on our paper. This is a rough first pass at a formal and agent-based model of group-level interethnic interactions. As you will see, both our formal and agent-based models are still in need of some refinement. We would greatly appreciate your feedback on how to move both of them forward. Thanks, Josh and Abe.

Abstract
Despite an explosion of research on interethnic cooperation and conflict, scholars and policy-makers still cannot answer two basic questions: 1) How and why do ethnic groups fall into states of conflict, cooperation, and oppression?, and 2) What policy interventions might move these groups out of conflict and oppression and back towards cooperation? In this article, we present a formal model that allows us to identify conditions under which social interventions aimed at moving groups from one of these conditions to another will work. The key, we argue, lies in understanding the interaction between institutional incentives and ingroup conformity. Among other things, our results cast serious doubt on the efficacy of approaches to conflict resolution that are based solely on changing outgroup attitudes through increased interethnic interaction (i.e. the contact hypothesis). Moreover, we illustrate how due to path dependence, the order of the changes made to these factors has a large impact on the outcomes of these changes.

1. The authors wish to thank Dan Magleby for helpful comments on this draft.

I. Introduction

Under Ottoman rule, Arabs (Muslims and Christians) and Jews living in and around the Sanjak of Jerusalem lived side-by-side in a state of general cooperation. Although small-scale conflict certainly occurred at times, "day-to-day relations were generally close and good" (Tessler, 127). This period of cooperation gradually came to an end, however, at the start of the 20th Century. In the decades that followed, the relationship between these groups in this region has alternated between periods (and areas) of oppression and conflict, interspersed only rarely by periods of intergroup cooperation.

Unfortunately, this pattern is not unique to this region of the world: interethnic tensions, marked by oppression and conflict, now dot every region of the globe. In the period since 1945, interethnic conflict within states has taken more lives than any other form of conflict, including inter-state war (Kalyvas 2007). In response, recent years have witnessed an explosion of research seeking to understand the causes for this intergroup conflict and proposing solutions for its resolution or prevention (for an excellent overview of this research, see Varshney 2007). Most of this research has focused on the impact of *interethnic interaction* or *institutional incentives* on determining whether groups will fall into cooperation, oppression, or conflict equilibria. Not surprisingly, conflict resolution groups and policies have followed this lead, designing policy interventions based on promoting greater interthnic contact (the most common of these are based on some variation of Allport's "Contact Hypothesis") or changing institutional structure in an effort to move groups away from oppression and conflict and towards cooperation. These interventions have seen mixed success, at best (Pettigrew and Tropp 2006; Forbes 1997).

As the mixed nature of this success indicates, despite the explosion of research on this topic, scholars and policy-makers still cannot answer two basic questions: 1) How and why do ethnic groups fall into states of conflict, cooperation, and oppression?, and 2) What policy interventions might move these groups out of conflict and oppression and back towards cooperation? In this article, we present a formal model to provide a first pass at answering these questions.

Our models build upon current research to model the impact of the two variables highlighted above--*interethnic contact* and *intstitutional incentives*--on interthnic cooperation, conflict, and oppression, while drawing on research in social psychology to model the impact of a third variable as well, what we call *ingroup conformity*. The models illustrate how these three variables interact to motivate individual behavior in such a way that makes certain group equilibria (conflict, cooperation, or oppression) possible or not. In particular, we identify five "interethnic relations regimes" that can result from these interactions. Perhaps more importantly, the models also provide a framework that allows us to identify conditions under which social interventions aimed at moving groups from one of the equilibria to another will work. The key, we will argue, lies in understanding in which of the five regimes current interethnic relations are locked. Among other things, our results cast serious doubt on the efficacy of approaches to conflict resolution that are based solely on changing outgroup attitudes through increased interethnic interaction (i.e. the contact hypothesis). In some regimes, as we will show, even positive changes in outgroup expectations that might occur based on greater interethnic contact--however dramatic--cannot move two groups out of oppression or conflict without accompanying changes in the other two factors. Moreover, we will illustrate how due to path dependence, the order of the changes made to these factors has a large impact on the outcomes of these changes.

In what follows, we present a brief overview of the literature on interethnic cooperation and conflict and present our theory of why ingroup conformity should be an essential, interactive component in any model of intergroup interactions. We draw from a history of the Arab-Israeli conflict and literature in social psychology to support this theory. We then present a formal model that highlights how ingroup conformity interacts with interethnic contact and institutional incentives to place intergroup interactions into one of three equilibria: cooperation, oppression, or conflict. In conclusion, we present a dynamic (agent-based) model that captures how changes in the values of each of these three factors might interact to tip intergroup relations from one state to the next and briefly discuss possible extensions of our model.

II. Interethnic cooperation and conflict

Before detailing current theories that seek to explain interethnic cooperation and conflict, we first define the terms of our inquiry. We begin by defining what we mean by ethnic group. Borrowing from Varshney (2007), we define an ethnic group as any group that shares a "sense of collective belonging, which could be based on common descent, language, history, culture, race, or religion (or some combination of these)." Under this definition, Jews and Arabs in Israel along with Sunnis and Shiites in Iraq are considered ethnic groups. Catholics and Protestants in Ireland would also be considered different ethnic groups, although members of these groups are hardly distinct along linguistic and racial lines.

With this definition in hand, we define the three states that inform our model: interethnic cooperation, conflict, and oppression. Informally, these states are easy to understand. A state of *interethnic cooperation* is one in which members of different ethnic groups cooperate to reach an outcome (in our case typically a social outcome) that is desirable for all. *Interethnic conflict* is a state where members of the different groups refuse to cooperate or coordinate on reaching an outcome that is desirable by all, but instead choose behavior that will specifically harm those in the outgroup. A state of *interethnic oppression* is where members of one ethnic group group engage in behavior that is designed to harm members of the other ethnic group but members of the other group do not respond in kind. These general definitions, which will be formalized later using game theory, serve as our starting point. They are intentionally vague, as the words "outcome" and "harm" in the preceding definitions are purposefully left undefined to allow the model to be applied to a wide array of conditions and contexts.

At present, theories seeking to explain how and why ethnic groups fall into either conflict, cooperation, or oppressive relations generally focus (with few exceptions) either on the impact of interethnic contact or on institutional incentives. Recent work by Varshney (2003) argues the first point. After conducting an in-depth study of Muslim-Hindu relations within India, Varshney concludes that greater interethnic association (both at the informal and associational levels) engenders greater interethnic cooperation. Neighbors don't kill neighbors, Varshney argues, because they know each other -- they know each others children, their dreams, and their values -- and in "knowing" they find that the "other" is not as different as they originally thought. In other words, when members of different ethnic groups intermingle, their attitudes about the ethnic "other" change. And not just attitudes about what the "other" values and believes, but also about how the other *behaves*. This leads to a change in *expectations* regarding the other. In additional work on interethnic relations in India, the psychologist Kakar (1996) makes a similar argument. Other scholars within this group write of the potentially large negative impact ideas like nationalism and ethnocentrism can have on interethnic relations, focusing on how they shape negative expectations about the outgroup that often are made into reality (Steele 1997; Hardin 1997; Petersen 2002).

Not surprisingly, many proponents of this first view argue the best way to move ethnic groups from a state of ethnic oppression or conflict to cooperation lies in promoting greater interethnic interaction at the individual level. This approach—based on Allport's (1954) "contact hypothesis"—takes many forms, but most common are the "ethnic encounter" groups currently promoted most widely in Israel (Halabi 2004). Recent work in Israel provides strong support for the idea that contact indeed changes attitudes towards the outgroup (Schwarzwald et. al., 1992; Maoz, 2000a; Bargal, 2004; Maoz, 2004; Halabi, 2004; Bargal 2008), although certain caveats and problems exist (Maoz, 2000b; Maoz, 2002). A recent meta-analysis of contact theory research studies conducted around the world during the last 50 years also provides strong support for this idea (Pettigrew and Tropp, 2006). However, this research provides little or no support for the idea that these attitudes actually translate into behavior (Spencer et al., 2008).

A second set of scholars argues that institutional incentives best explain why ethnic groups fall into cooperation, conflict, or oppressive equilibria. Wilkinson (2004) argues that certain types of electoral institutions provide incentives for groups to fall into conflict. Drawing again from the Indian case, he argues that shifts into ethnic conflict in many cases in India are motivated by the desire for electoral gain. He supports this argument by analyzing the actions of the BJP. In addition to focusing on electoral institutions, scholars in this category also focus on the impact of factors like ethnic income and educational disparities that generate grievances to motivate one group to dislike another. This approach serves as the common assumption underlying explanations for a number of conflict worldwide, from interethnic conflict in Southeast Asia (Tambiah 1996) to Israel (Lustick 1980). Fearon and Laitin's (1996) work on interethnic relations falls in this second group. However, one aspect of their work makes them unique: they focus on the role of *ingroup policing* in this process, illustrating how this process can engender greater intergroup cooperation. In each of the cases cited above, institutions serve two purposes related to interethnic conflict: 1) they provide either an incentive towards intergroup conflict (providing either high or low "returns to extortion") or 2) they contribute to social structure in such a way that generates grievances within one of the ethnic groups.

The foregoing review details current theories seeking to explain why ethnic groups fall into conflict, cooperation, or oppressive equilibria and current policy attempts aimed at moving groups towards cooperation. In what follows, we argue that any model seeking to explain this phenomenon should also include one more factor: ingroup conformity.

III. The effect of ingroup conformity on interethnic cooperation and conflict

As the foregoing literature review illustrates, the potential impact of ingroup conformity on behavior towards the outgroup has received scant attention (at best) by scholars of interethnic cooperation and conflict. We argue that this is a large omission. In an oft-ignored section of his landmark monograph *The Nature of Prejudice* (1954), Allport argues that, "although we could not perceive our own ingroups excepting as they contrast to outgroups, still the ingroups are psychologically primary. We live in them, by them, and sometimes, for them. Hostility towards outgroups helps strengthen our sense of belonging, but it is not required" (41). Since Allport published this argument, a long line of convincing research within social psychology suggests that the desire to conform to the actions and preferences of one's ingroup often impacts individual behavior *more* than outgroup attitudes shaped by greater interethnic interactions or than institutional incentives.

At present, researchers posit varied mechanisms behind this impact. One set of scholars argue persuasively that we see this ingroup conformity effect as a result of the impact ingroups have in shaping how their members perceive others and events. This line of argument has a long history, beginning with experiments by Asch in the 1950s and the furthered by the group-based work of Sherif and others that followed (Asch 1951; Sherif 1951). In more recent work, Brewer (1999) and Yzerbyt et al. (2000) provide experiments that illustrate how ethnocentrism, or "ingroup bias," as Brewer calls it, shape individual behavior by impacting perception. Although the measures they employ are different, the results are similar.

Another set of scholars argue that ingroups engender conformity by facilitating a process of what Takacs has called "social control" (Takacs 2006). Here, the mechanism behind behavioral conformity has nothing to do with changing attitudes, but rather with the ability of the ingroup to incentivize behavioral conformity, regardless of what individual members think. The incentive might come from fear of ingroup social sanctioning should behavior not conform to the group norm (Fearon and Laitin 1996; Strauss 2006), or from the dense personal ties that constitute ingroups--ties that individuals are loathe to break by behavioral deviance (Takacs 2006).

Whatever the mechanism behind it, levels of ingroup conformity certainly impact whether ethnic groups fall into oppression, conflict, or cooperation. This is perhaps nowhere more clear than in the case of the Rwandan genocide. In early 2000, political scientist Scott Strauss conducted numerous interviews with individuals now in prison for killing their neighbors from the other ethnic group. What he found was striking: many of those he interviewed did not have negative attitudes towards the ethnic outgroup, neither at the time of the interview nor at the time they killed members of the outgroup. In fact, many of them had a large degree of interethnic contact and liked members of the ethnic outgroup. Why kill, then? The common answer was some variation of ingroup pressure (Strauss 2006). Ingroup conformity can take what would be individual-level conflict and, through the mechanisms described above, transform it into a conflict at the group level.

Although the mechanisms behind ingroup conformity are different, the result is the same: individuals within an ingroup conforming to the behavioral norm of others within their group. In our models, we focus on the impact of ingroup conformity *behavior* towards members of the outgroup. As such, the exact mechanisms behind this behavior have no bearing on our results. An individual might conform in her behavior towards the outgroup because she has truly adopted the attitudes of her ingroup (the first mechanism) or even if she has not, but just fears ingroup social sanctioning (the second mechanism).

With this theoretical justification for the inclusion of each of the three factors that comprise our models of interethnic relations, we now turn to the models themselves for a presentation of how these variables interact and the outcomes they produce. The formal model illustrates how these three variables interact to motivate individual behavior in such a way that makes certain group equilibria (conflict, cooperation, or oppression) possible or not. The dynamic model builds on this framework to allow us to identify conditions under which particular social interventions aimed at moving groups from one of the equilibria to another will work.

IV. A formal model of intergroup interactions

From a technical perspective, our model is a combination of Schelling's (1973) multiplayer prisoners' dilemma (MPD) and Fearon and Laitin's (1996) in-group policing model (IPM). Like Schelling's MPD, ours is a one-shot game played simultaneously among multiple players. Both the MPD and IPM assume that players face a binary decision, a modeling convention we adopt. This simplification helps keep our analysis tractable and makes our results more intuitive. It also allows us to use Schelling's excellent notation and graphs to describe the intuition behind our model.

Terminology differs here. To denote the actions individuals can take, Fearon and Laitin use labels from the prisoners dilemma: "cooperate" and "defect." Schelling uses the neutral labels "left" and "right." To make the normative implications of individual decisions clear, but emphasize the difference between our game and the prisoners' dilemma, we use the labels "nice" and "mean." To "play nice" is to cooperate (as defined earlier in this paper); to "play mean" means to choose otherwise.

Like the IPM, players in our game come in two types, denoted generically as A and B. Each player belongs to one type, his *ingroup*. The other group is the *outgroup*. When deciding how to act, players must consider how their actions will be rewarded by their in- and outgroups. Two micromotives come into play here. First, players want to conform to their in groups. Second, players want to derive maximum utility from a coordination game with the outgroup.

The remainder of this section is organized as follows. First, we describe each the in and out group sub-games in detail. Next, we explain how these sub-games are combined. Finally, we provide a full formal description of our model.

The in group sub-game

Consider a typical player, *a*, of type A. (Since the game is symmetric, conclusions for B-type players are analogous.) For convenience in exposition, assume that we can ignore *a*'s behavior when computing the average behavior of A-type players (i.e. *a*'s impact on his group norms is "small"). Let *p* denote the proportion of A-type players who play nice. In this case, *a*'s utility functions for the *ingroup sub-game* are as follows.[2]

---

2. This assumption makes many of our later derivations and clearer by making functions simpler and differentiable. It is easily justified when both groups are composed of many players making independent decisions. We leave descriptions of small-n scenarios for later work.

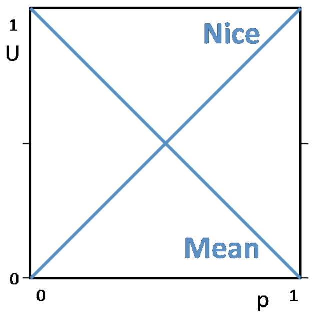
U(Nice | p, q) = p
U(Mean | p, q) = (1-p)

In this payoff structure, *a*'s utility is a function of the degree to which his behavior conforms to that of his ingroup. Decisions by members of the outgroup do not matter. If we treat group A as a single player adopting a mixed strategy of p, then the payoffs to *a* are identical to the classic left side/right side traffic coordination game (Table 1 and Fig 1).

Table 1. Ingroup payoffs for *a* given a uniform strategy within A

	A=N	A=M
***a*=N**	1	0
***a*=M**	0	1

Fig 1. Ingroup payoffs for *a* for mixed strategies within A



If group A and player *a* uniformly play nice (or mean), *a* receives his maximum ingroup payoff of one unit of utility. If group A uniformly plays nice (or mean), but *a* chooses the opposite action, *a* receives his minimum ingroup payoff of zero. For *p* between zero and one, *a*'s conditional payoffs vary smoothly between these extremes.[3]

These outcomes are consistent with the ingroup conformity behavior described above. As we argued in the literature review, this behavioral assumption allows us to capture the outcomes of ingroup conformity without explicitly describing the mechanisms behind it (we have identified two common mechanisms; there are certainly more). As such, our approach puts the *causes* of ingroup conformity into the proverbial black box. This means that our model will not be able to explain how ingroup conformity is maintained, but it will be able to shed light on the logical implications of this conformity, regardless of the social, political, or economic mechanisms maintaining it.

The outgroup sub-game
We now turn our attention to the impact of a player's outgroup on her payoffs. Using the same notation as above, we add a new parameter we call q: the proportion of B-type players playing nice. We include this factor to to model the impact of *interethnic interaction* on individual play. The theoretical justification for this inclusion can be found in our theory section earlier in the article. With the addition of this new parameter, *a*'s utility functions for the *outgroup sub-game* are as follows.

3. This is the simplest case of a broad class of functions that might describe conformity motives. Logistic functions symmetric about the point p=1/2 would be a good generalization.

U(Nice | p, q) = 2q - 1
U(Mean | p, q) = qx

If we treat group B as a single player adopting a mixed strategy of q, then a little algebra shows that *a*'s payoffs are equivalent to the following tables.

Table 2. Outgroup payoffs for *a* given a uniform strategy within B

	B=N	B=M
a=N	1	-1
a=M	x	0

Fig 2. Outgroup payoffs for *a* for mixed strategies within B when x > 1

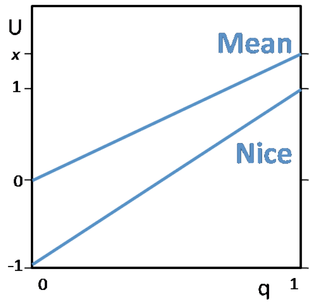
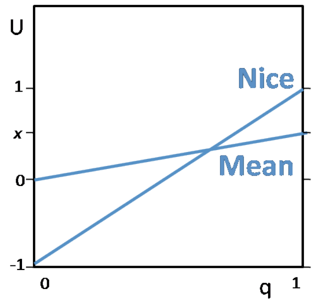


Fig 3. Outgroup payoffs for *a* for mixed strategies within B when x < 1



Within this sub-game, we introduce another new parameter as well, x, to capture a subset of *institutional incentives* that motivate interethnic interactions. We call this parameter *returns to extortion*: *a*'s incentive to play mean against a uniformly nice out group. As we described earlier in this article, current theories suggest a number of potential different institutional arrangements that can (and do) incentive this type of behavior. If x > 1, the sub-game is a prisoners' dilemma and each player has an incentive to play mean (i.e. "defect") regardless of the behavior of members of the outgroup. For 1 > x > 0, the sub-game is like a stag hunt. If possible, players prefer to coordinate on the nice-nice global maximum, but if the out-group is "mean" they will reciprocate with meanness. For x < 0, the sub-game becomes one of pure coordination. Players will tend to respond with whatever action they think the out group is most likely to choose, but the nice outcome is more rewarding than the mean.[4]

This use of games like the prisoners dilemma, stag-hunt, and coordination to describe interactions follows a long tradition in institutional analysis and recent use by Fearon and Laitin to describe interethnic interactions. By using x to parameterize the returns to extortion, we make our model flexible enough to account for all of these scenarios.[5]

Group cohesion

---

4. The difference between stag-hunt and pure coordination is mainly a function of risk aversion. Since our model does not include risk aversion, the difference is immaterial, but we include the terms anyway for reference.

5. Our model could be made even more general by parameterizing the value of the "sucker's payoff" (currently -1). For the general results discussed in this paper, the exact value of the sucker's payoff does not matter as long as it is less than zero.

Having examined the sub-games, we can now discuss how these games are combined. We accomplish this by introducing two additional parameters, α and β, to describe *group cohesion* within the model. Both α and β are bounded on the interval [0,1].

$$U(\text{Nice} \mid A, p, q) = \alpha\, p + (1-\alpha)(2q - 1)$$
$$U(\text{Mean} \mid A, p, q) = \alpha\, (1-p) + (1-\alpha)(qx)$$

These payoffs combine the two sub-games, using the parameter α capture the strength of group A's level of ingroup conformity. When α is close to zero, A-type players feel little pressure to conform, and their behavior will be mainly determined by outgroup incentives. In contrast, for α close to 1, the conformity motive dominates. The parameter β plays the same role for B-type players.

By varying α and β between 0 to 1, we can describe any level of relevance to ingroup behavior. When α (β) is high, we will say that group A (B) has high group cohesion. Note that group cohesion is not necessarily the same thing as uniformity of action. Members of a group with low cohesion (i.e. α close to 0) might still make uniform choices (i.e. p close to 1 or 0). Note also that group cohesion does not necessarily imply the potential for collective action. In our model, members of high cohesion groups exhibit a strong tendency to conform, but that tendency does not translate into ability to coordinate changes in behavior. We leave this promising extension for future work.

One possible objection to this scheme is that we may end up comparing apples and oranges when in- and outgroup payoffs are in different "currencies." For example, if ingroup payoffs come in the form of social acceptance, but outgroup payoffs are monetary, we end up with a nonsense equation summing social acceptance with money. Our response to this critique is that the "exchange rate" between in- and outgroup payoffs can be built into α and β. If we imagine that people are willing to make trade offs between acceptance and money, then the cohesion parameters can be thought of as capturing the terms of that trade off. If we don't believe that people make such trade offs, then the parameters can reflect that inelasticity by taking values of 0 and 1. This abstraction provides fertile ground for future work attempting to operationalize these values.[6]

Full game specification

Following is a full, formal specification of the game.

Players

Our model includes n players of two types, A and B, with group counts of *n*_A and *n*_B, respectively. From the perspective of an A-type player, A is the *ingroup* and B is the *outgroup*.

Strategies

Each player makes a binary decision to be *nice* or *mean* to the out-group. These strategies are denoted N and M, respectively.

---

6. Framing effects and non-constant exchange rates would be good candidates for extensions of this model.

Payoffs

Individual payoffs are a function of individual choice and the proportions of the in- and outgroup that decide to play nice. Let p and q ε [0,1] be the proportion of the groups A and B that play nice. α and β ε [0,1] are parameters representing the *group cohesion* of players of type A and B, respectively. The variable x represents the *returns to extortion* of the out group.

$$U(N \mid A, p, q) = \alpha\, p + (1-\alpha)(2q - 1)$$
$$U(M \mid A, p, q) = \alpha\, (1-p) + (1-\alpha)qx$$
$$U(N \mid B, p, q) = \beta\, q + (1-\beta)(2p - 1)$$
$$U(M \mid B, p, q) = \beta\, (1-q) + (1-\beta)px$$

Information

We assume that players' types and payoffs are common knowledge. As described so far, the model is a one-shot simultaneous game, so no other specification of information is necessary. Later, we will extend this to a dynamic model. In the dynamic model, we will use p' and q' denote the proportion of A- and B-type players who played nice in the previous round.

In summary, we now have a model of interethnic interactions where individual behavior is influenced by ingroup conformity norms (the in group subgame) and intergroup interactions (the outgroup subgame). Depending on the cohesion parameters, α and β, and the returns to extortion, x, the model can describe a variety of cases, including the special case where ingroup cohesion plays no role ($\alpha = \beta = 0$). This zero-cohesion case can serve as a baseline for comparison to institutional models without a group conformity motive.

V. Model equilibria and five "interethnic relations regimes"

The previous section established the underlying assumptions for our model. In this section, we build on this foundation by identifying sets of self-reinforcing behavior within the model. To do so, we use the concept of trembling-hand Nash equilibrium. To qualify as an equilibrium under this definition, a system of behavior must satisfy two criteria. First, it must be self-reinforcing: no player has anything to gain by switching his behavior. Second, it must be robust to perturbations: a small number of "mistakes" cannot make the system unstable. This equilibrium concept is appropriate for social systems where mistakes and miscommunications are uncommon but possible (Selten, 1975).[7]

The remainder of this section is organized as follows. Using this solution concept, we first derive and classify all equilibria for the game. Next, we classify the relative utility provided to individuals in each equilibrium. Finally, we describe the five interethnic relations regimes

---

7. We use this equilibrium concept as our starting point. However, in future iterations of this model we intend to explore evolutionary stable strategies as well.

that meet existence criteria.

Equilibrium Classification

Because this is an *n*-player game, the set of potential equilibria is large and potentially complex. We begin by showing that the set of potential equilibrium conditions can be drastically reduced.

Note: for the same reasons given above, we will continue to ignore the effect of an individual's behavior when computing the average behavior of his ingroup.

> *Proposition 1: In equilibrium, all players of a given type share the same best response function.*
> *This conclusion follows directly from the requirement that expectations about in- and outgroup behavior must match actual behavior in equilibrium. A player's best response is a function of his type, expectations, and payoffs. Payoffs and types do not vary within types. Consequently, equilibrium best responses must also be uniform within types.*

This result implies that from the perspective of an individual player with small impact on ingroup norms, equilibrium play is analytically equivalent to playing against one opponent of each type choosing a mixed strategy. Within these group aggregates, individual players may adopt any personal strategy as long as the average for the group stays at the equilibrium point.

Next, we simplify further by eliminating the possibility of mixed-strategy equilibria to the game.

> *Proposition 2: No mixed strategy can satisfy the trembling hand condition.*
>
> *(See appendix A for the proof.)*

The corollary to proposition 2 is that pure-strategy equilibria are the *only* viable equilibrium candidates for this game. There are only four such equilibria. Using ordered pairs to denote the A- and B-type strategies, they can be expressed as (M,M), (N,N), (N,M), and (M,N). Here we describe necessary conditions and payoffs for each. (Proofs for these existence conditions are given in appendices.)

> *Conflict (M,M)*
> Pure conflict occurs when all players of both groups choose to play mean. This situation is always a potential equilibrium of our game. In a conflict equilibrium, each player receives a payoff of 0.
>
> *Cooperation (N,N)*

Cooperative equilibrium occurs when "nice" is a self-sustaining choice for all players of both groups. If $x \le 1$, cooperation is always a possibility. For $x > 1$, cooperation is a possibility if and only if $\alpha, \beta > (x-1)/x$. In a cooperative equilibrium, each player receives a payoff of 1.

Oppressive (M,N), (N,M)
If all of one team plays nice while the other team unanimously plays mean, an oppressive equilibrium results. We will call the mean group the *oppressing* or *dominant* group, and the nice group the *oppressed* or *submissive* group.

The existence conditions for this equilibrium are the most complex. We describe half of them here, putting the A group in the dominant position. (Conditions for B are analogous and symmetric.) For all of the following conditions, β must be greater than 1/2. If $x \ge 1$, A-over-B oppression is always a possibility. For $x < 1$, A-over-B oppression is a possibility if and only if $\alpha > (x-1)/(x-2)$. In this equilibrium, each oppressing player receives a payoff of x. Oppressed players get -1.

Equilibrium Outcomes
Having categorized the possible equilibria for the game, we can compare them with respect to the relative utility they provide to their players. Table 3 reports group and average payoffs, as well as payoff differences.

Table 3. Equilibrium outcomes

Equilibrium Type	**A Payoff**	**B Payoff**	**Avg Payoff**	**Payoff Difference**
(M,M) Conflict	0	0	0	0
(N,N) Cooperation	1	1	1	0
(M,N) A-B Oppression	x	-1	(x-1)/2	x-1
(N,M) B-A Oppression	-1	x	(x-1)/2	-1-x

Not surprisingly, conflict yields an average payoff of zero. This should not be interpreted to mean that conflict generates no payoffs for some individuals, however. Since payoffs are ordinal measures, the important thing is the ordering of the regimes. Cooperation yields an average payoff of one, so cooperation is always a more desirable outcome than conflict. The utility of oppression for individuals in our game depends on the value of x. For x greater than one, oppression is the most efficient regime. For x less than one, oppression is less efficient than cooperation.

As suggested by our choice of the term "oppressive," outcomes under oppressive equilibria are not equal between groups, except in the special case $x=1$. We interpret this result loosely by saying that under oppressive equilibrium, the benefits of group interaction accrue unevenly across groups.

Interethnic Relations Regimes
Thus far we have shown that the existence of equilibria depends on the parameters x, α, and

β: we have derived the *existence* conditions for each type of equilibrium to exist. One reasonable question is whether we can derive *uniqueness* conditions for any of the equilibrium types. That is, given a specific configuration of x, α, and β, can we predict which equilibrium will result?[8]

For the static game, the answer is usually no. The simplest way to illustrate why is through a map of the five possible *equilibrium regimes* that result from our model over different parameterizations, shown in Fig 4. We call these regimes "interethnic relations regimes," as they constitute a complete set of all possible regimes that occur from the interactions generated by our model.

Figure 4. Equilibrium regimes over returns to extortion and group cohesion

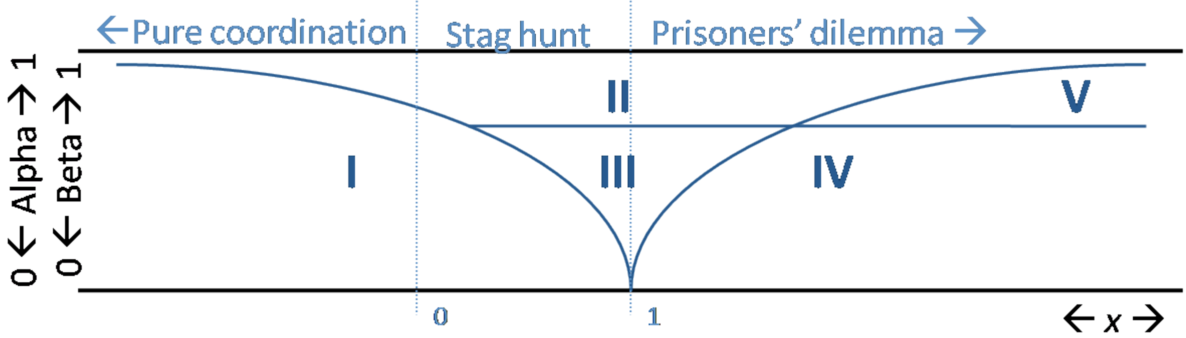


Table 4. Equilibrium regimes and existence

Regime Number	x	α β	Conflict (M,M)	Cooperation (N,N)	Oppression (M,N), (N,M)
I	less than 1	less than (x-1)/(x-2)	yes	yes	-
II	any	greater than 1/2 and (x-1)/x for x > 1, and (x-1)/(x-2) for x < 1	yes	yes	yes
III	between 0 and 2	less than 1/2 and greater than (x-1)/x for x > 1, and (x-1)/(x-2) for x < 1	yes	yes	-
IV	greater than 1	less than 1/2 and (x-1)/x	yes	-	-
V	greater than 2	greater than 1/2 and less than (x-1)/x	yes	-	yes

This map is a kind of Venn diagram of equilibrium types over x, α, and β. We assume that α = β, and project returns to extortion onto the x-axis and group cohesion onto the y-axis. This allows us to graph the boundaries of the existence conditions for different equilibrium types within this parameter space. These conditions divide the space into five equilibrium regimes. The patterns of behavior that are sustainable depends on the regime.

8. Note that the simple averages used here do not weight groups by size, effectively assuming that both groups are the same size. The general formula for average payoff is (n_A * U(A) + n_B * U(B))/n.

This map also gives us visualizations of several other cases of interest. An imaginary vertical line to the right of x = 1 represents the cases in the canonical prisoners' dilemma model, plus cohesion. Note that this line cuts through regimes where all three outcomes are possible. The horizontal line, $\alpha = 0$, at the bottom of the graph is the "atomistic" world where group cohesion plays no role. Note that oppressive equilibria are impossible without group cohesion.

As the foregoing diagrams and tables illustrate, we can only derive uniqueness conditions for conflict (regime IV). In all other regimes, more than one equilibrium is possible. Although we cannot derive uniqueness conditions for most of these regimes, however, the foregoing regime map does allow us to make the following three substantively important conclusions. First, in all five regime types, conflict is always a possibility. In other words, there does not exist a certain set of values for the three parameters that constitute our model that produce a regime in which interethnic conflict cannot occur. In the dynamic model that follows, we delve into this finding more to investigate which ordering and levels of factors will make conflict or the other equilibria more likely in each regime.

Second, for oppressive equilibria to exist, high ingroup conformity is a *necessary condition*. Our model suggests that ingroup conformity not only plays a role in cooperation (Fearon and Laitin 1996) and conflict (Strauss 2006), but also the key role in oppression. While this result is not necessarily surprising (given our literature review earlier), it underscores the vital impact of addressing the processes generating ingroup conformity in all policy efforts to improve interethnic cooperation. Indeed, this result should be of particular interest to policy makers, for it suggests that changing ingroup conformity levels (which in many instances might prove easier than changing patterns of interethnic interaction and outgroup attitudes) might prove key to moving groups out of oppressive equilibria towards cooperation. Our dynamic model addresses this possibility.

On a related note, there is a strong relationship between Fearon and Laitin's (1996) ingroup policing game and our model of interethnic interactions. For one thing, their game appears to support oppressive equilibria between groups (see appendix E). Moreover, careful inspection will show that the incentive structure of our game bears a strong resemblance to the expected utility stream of equilibrium play in Fearon and Laitin's model. Our parameters α and β depend on the proportion of interactions within groups and the presence of in-group policing. Our parameter x depends on the relative returns to defection and cooperation (labeled α and $-\beta$ in the other model). The values of other payoffs are the same in order, if not in magnitude, suggesting a strong correspondence between their modeling of ingroup policing and our approach.

Our third conclusion is perhaps the most important. The foregoing regime map provides a rough heuristic for what policy interventions will likely NOT work in given conditions. For example, our analysis suggests that if interethnic relations within a given state at a given time were located in quadrants IV and V, then no changes in p and q--no matter how dramatic--could push relations back towards cooperation or oppression. Moreover, if interethnic relations were located in quadrants II or V (oppressive equilibria), changes in p and q also will not move relations towards cooperation. In both cases, changes in ingroup cohesion and/or changes in institutional incentives (i.e. returns to extortion) are the policy options with the highest potential for success.

This result casts significant doubt on the potential positive impact of conflict resolution strategies that focus their efforts on promoting greater interethnic interaction to change outgroup expectations or attitudes (i.e. contact hypothesis approaches). As these strategies currently dominate other strategies in ethnically diverse countries around the word (Halabi 2004), the practical impact of this result is not trivial. Why do we arrive at this conclusion when others do not? For one, rather than focusing solely on the impact of one of the three main factors we identify at the start of this article (the usual approach), our model provides a framework for assessing how these factors interact. And this interaction provides counter-intuitive results. For example, the model highlights how high levels of ingroup conformity can cancel the effects of positive interactions or positive outgroup attitudes held by a few members of an ingroup. Theoretically, this makes complete sense. The same ingroup policing that can enforce conformity to cooperative intergroup relations can--and often does--play the opposite role as well.

This is perhaps nowhere more clear than in assessing the state of Arab-Jewish relations within Israel. In a series of recent interviews with proponents of "interethnic encounters" within Israel, Gubler often heard the lament that although these encounters worked in changing attitudes towards the ethnic outgroup, some of their former participants could still be found participating in acts of violence against the outgroup. The stories they told about former students they had found doing this were all strikingly similar: a former Arab-Israeli (you could plug in Jewish-Israeli here just as well) student who no longer has negative attitudes about Jewish-Israelis is approached by a group of friends (who had not participated in the "interethnic encounter" and thus did not share his positive attitudes) and asked to join them in committing some act of violence against their Jewish-Israeli neighbors. The student feels a need to conform to the behavior of his ingroup (for a myriad of reasons). As such, he goes with his friends.

Thus far, our model has provided a foundation for understanding the conditions under which certain policy interventions will not work--a highly unsatisfactory conclusion. And these are the limits of a static model such as ours. In what follows, however, we employ an agent-based model that adds just one additional dynamic to our formal model to explore policy prescriptions might work. We explore the conditions under which ethnic groups in certain regimes might tip into others.

VI. A dynamic model of intergroup interactions

In order to approach this problem, we invoke one additional dynamic assumption: changes in strategy don't happen instantly, but diffuse through the population over time. To incorporate this dynamic assumption, we reframe the basic model presented above as an iterated model. Players, strategies, payoffs, and information are the same as in the static model, with the addition of p' and q'. These two information parameters denote the group proportions of nice players in the previous round. At each time step, some proportion, $\delta \varepsilon [0,1]$, of players is selected at random. The selected players all consider the current parameters of the model, including p' and q', and adopt the strategy that maximizes their own utility. This iterative decision making process provides a simple way to model the conditions under which relations will tip from one equilibrium to another.

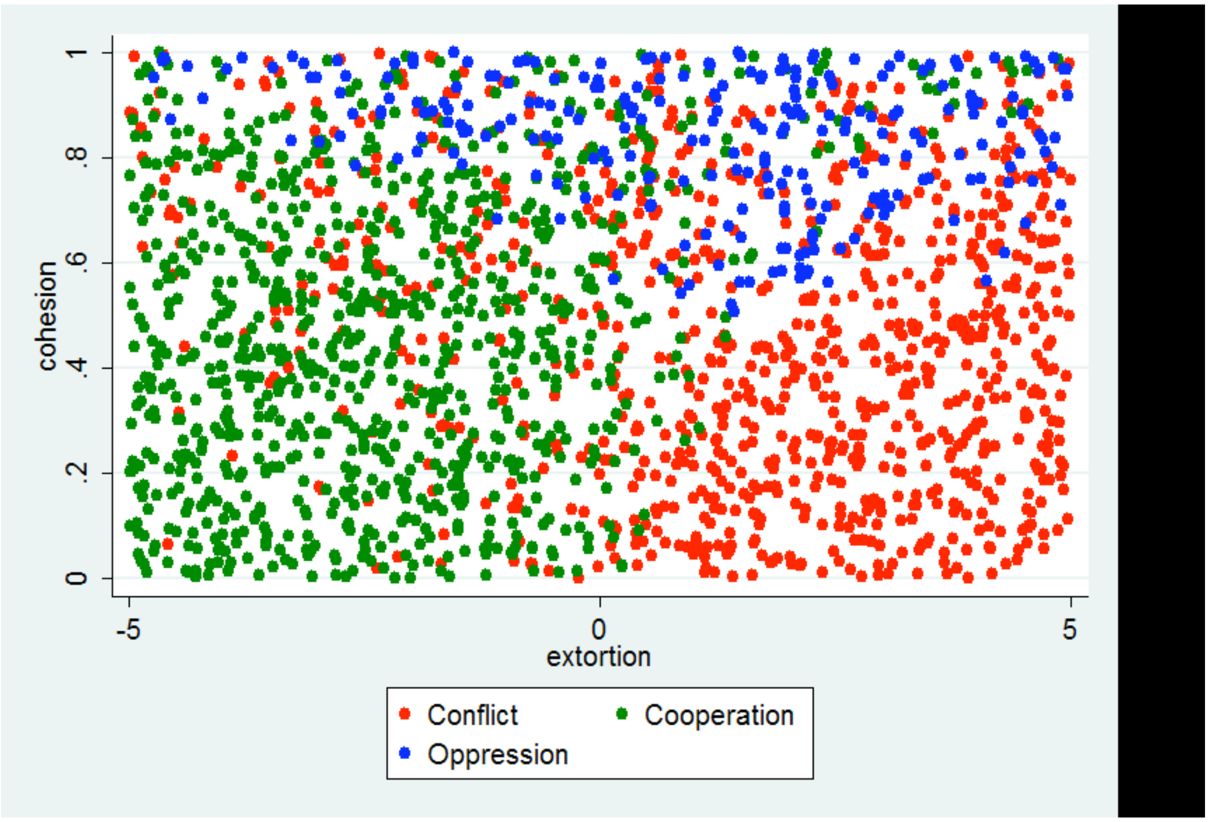
In this model, players are rational in the sense that they choose the optimal outcome for the current iteration. However, they are myopic in the sense that they do not consider trends in behavior, the long-term impact of their actions, or any kind of collective action. In other words, they work on local knowledge. This means that the set of equilibria in the dynamic model is the same as the set previously discussed: conflict, cooperation, and oppression. The difference is that we can now investigate the dynamics of the model far from equilibrium.

Agent-based model
To get a feel for these dynamics, we created an agent-based model of the dynamic game just described. This model encodes the game exactly as described above, allowing us to simulate the results of different parameterizations (regime starting points) and histories.

The following scatter plot is similar to the regime map in the previous section. It depicts the results of 2,000 simulations (a.k.a trials, or trial runs), each with 100 agents, split evenly between types A and B, with $\delta = .15$. In each trial we drew a random number from a uniform distribution on [0,1] and set α and β equal to it. We drew values of p' and q' independently from the same distribution. We also drew x from a uniform distribution on [-5, 5]. In other words, we set up the model such that each trial would be randomly assigned to start in one of the five regime types we identified earlier. With these parameters set, we ran each trial for 100 iterations. This turned out to be an ample amount of time to converge to equilibrium. We classified each run based on the equilibrium it generated (conflict, cooperative, or oppressive), and plotted the color-coded results on a scatter plot, with x as the x-axis and α and β on the y-axis.[9][10]

Figure 5. Trials starting with random expectations

9. For sufficiently high δ, the possibility of conditional strategies (e.g. tit for tat, grim trigger, alternating defection) re-emerges. However, a high δ value presupposes a high level of intergroup communication. This rather strong assumption goes against the intuition we are trying to capture in this model. Moreover, most conditional strategies do not seem to be robust to the trembling hand equilibrium condition. Consequently, we ignore them in this paper.
10. Meaning that on each iteration, roughly 1/6 of the population is given the opportunity to update their player strategies based on current local information.



This set of simulations captures a simple thought experiment: if 100 players were to start with shared but random expectation about group play, what outcome would they arrive at? In each trial, the result is partly stochastic and conditional on group cohesion and returns to extortion. These simulations serve to verify that our agent-based model does indeed mirror our original static game: for each kind of equilibrium, patterns of existence perfectly match those of our earlier regime map. In addition, we can infer that convergence to equilibrium is usually fast, and that the likelihood (not just existence) of a given outcome is conditional on the parameterization of the model.

This set of simulations is instructive, but has some unrealistic aspects. In particular, it is hard to imagine a situation in which a set of players would start with shared, but random expectations about future play. In most situations, we would expect players' expectations to be conditioned on past play. And given the model's rapid convergence to equilibrium, we would expect most instances of past play to be equilibria.

Following this intuition, we set up another set of simulations. Trial runs were identical to those above except that players' starting expectations were constrained to start at equilibrium. Starting with equilibrium conditions of cooperation (N,N), oppression (N,M), and conflict (M,M), we conducted a thousand trials each. This set of simulations allows us to begin to answer our earlier question: if the game is in equilibrium and parameters are suddenly changed, which equilibrium will the players converge on?

Results for trials starting in cooperative equilibrium are shown in Fig. 6. Trials starting in oppressive equilibrium are shown in Fig. 7. Trials starting in conflict are not shown, because the results were entirely uniform: every run ended in conflict.

Figure 6. Trials starting in cooperative equilibrium

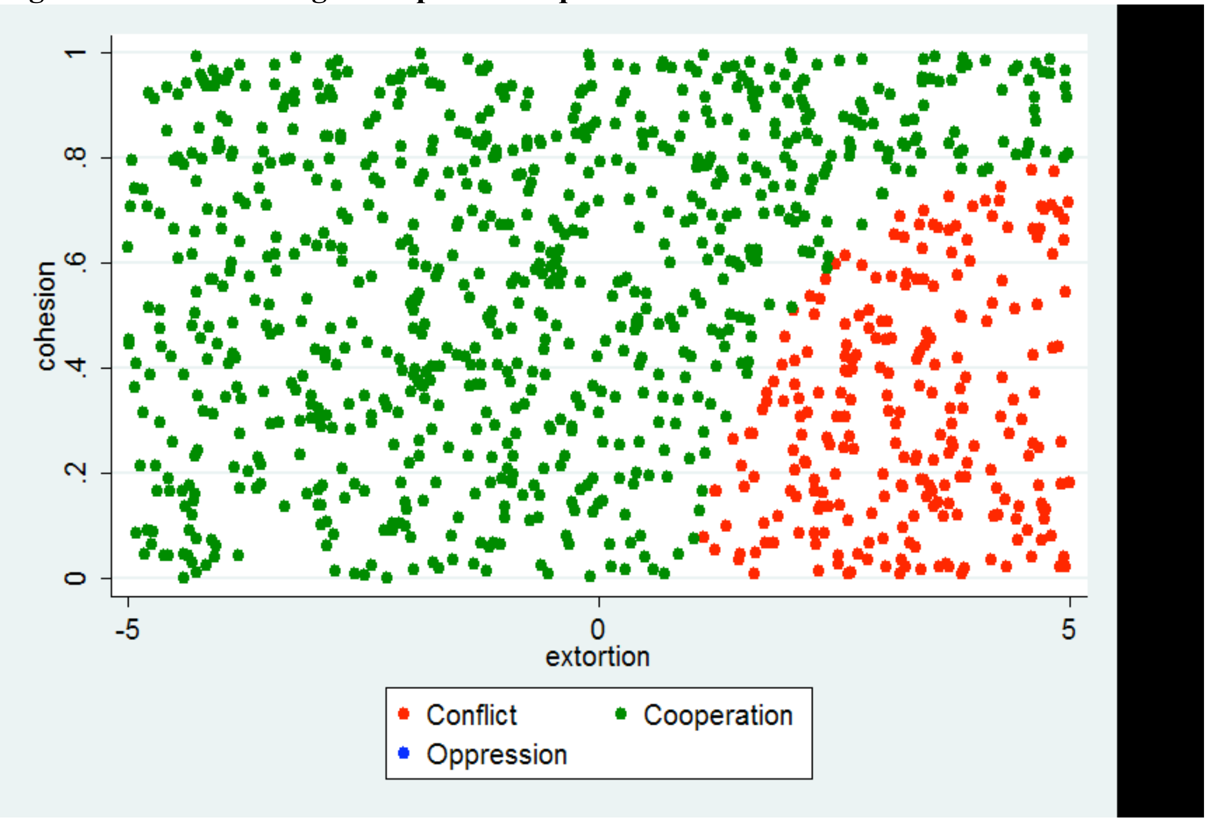
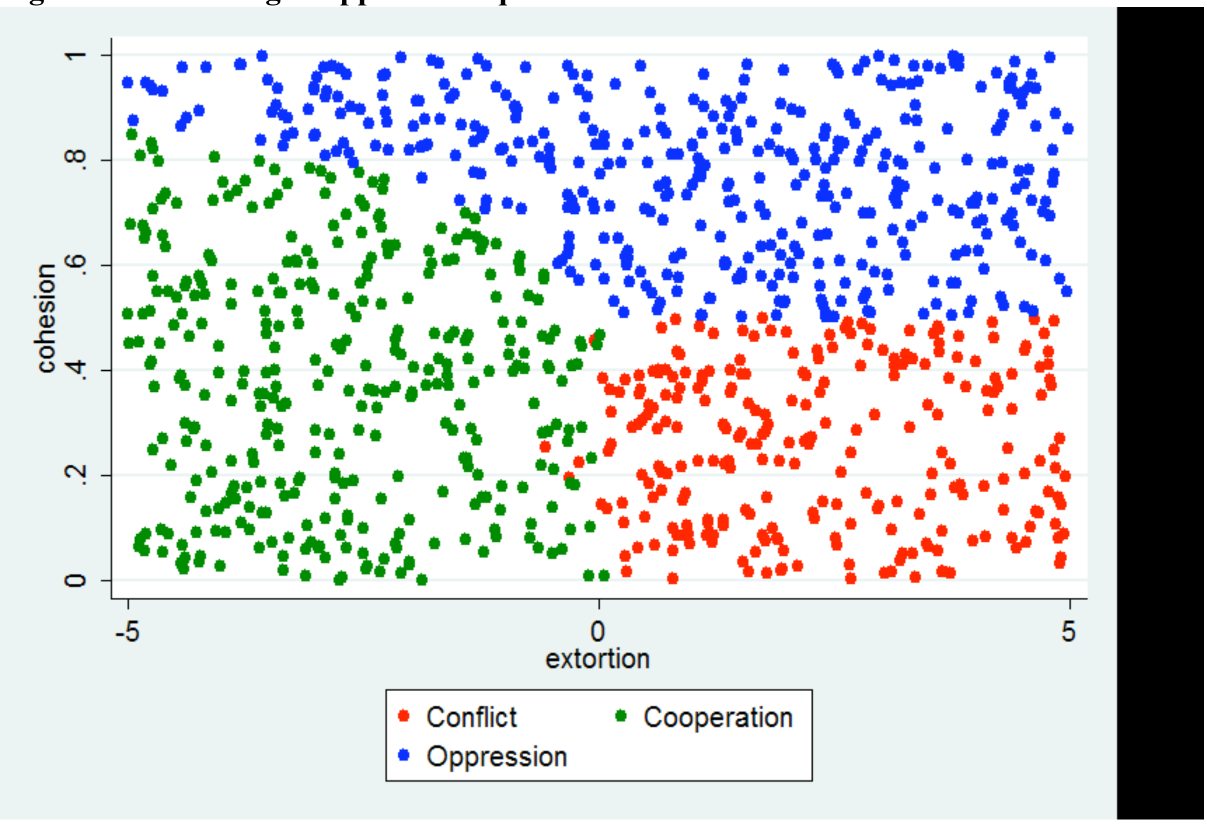


Fig. 7 Trials starting in oppressive equilibrium



Taken together, the results in these last two figures tell an intriguing story about the interaction of our three factors in the model and path dependence. These graphs show that as parameters change, the regime boundaries create tipping points between various equilibria. Thus, depending on where ethnic groups start, we can understand a little about how changes in our factors might predict where they end.

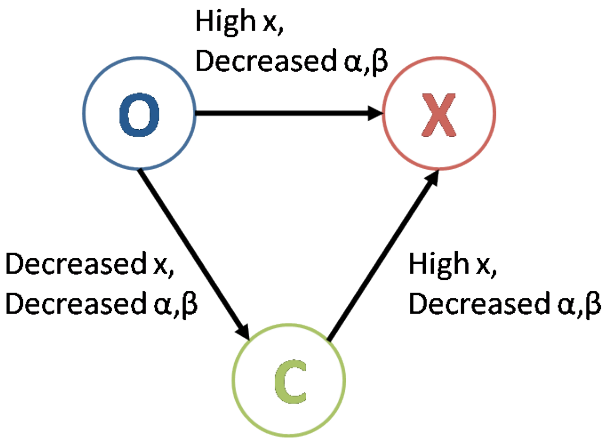
Three things are worth noting in the first figure. First, when ethnic groups start within a regime that can support cooperation, the system stays in cooperative equilibrium. A moment's reflection will reveal why: by definition, equilibrium expectations are self-reinforcing. Second, in our simulations, cooperation never tips into oppression, despite the fact that oppression is a viable equilibrium in many cases. (All cases in regime V are candidates to tip from cooperation to oppression.) Third, we might imagine that if an incremental increase in x tips the system from cooperation to conflict, a corresponding decrease would tip the system back. This turns out not to be the case. Conflict is a viable equilibrium regardless of the parameterization of the model. Therefore, no change in parameters is sufficient to tip the system out of conflict.

The last figure, describing simulations starting in oppressive equilibrium, provides the most substantively interesting results. It first suggests that the most efficient way to move groups from an oppressive equilibria is to lower ingroup cohesion, for in conditions of low cohesion, this equilibrium becomes unsustainable. However, what comes next depends on the values of x -- institutional incentives, in this case returns to extortion -- when the oppressive equilibrium falls apart. If x is low, the system shifts into a cooperative equilibrium. However, if x is high, the system degenerates into conflict. For values of x near 0, the results appear to depend on chance. Since neither cooperation nor conflict ever tips back to oppression, the value of x when group cohesion decreases could have an enormous impact on the long-term outcomes of group interaction.

Figure 8 is a state transition model describing the tipping points among conflict, cooperation, and oppression. Note that these results do not necessarily hold when groups have differing levels of cohesion, that is, when α and β are unequal .[11]

Fig. 8 State transitions between conflict (X), oppression (O), and cooperation (C)

11. The dynamics of intergroup norms get more complicated when α and β to take on different values. Part of the problem is that visualizing regimes becomes much more difficult. It would be interesting to ask under what conditions low or high coherence improve the dynamic outcomes of a group, but we leave that problem for a later version of this paper.



With these results in hand, we can return to our earlier question about existence and uniqueness of equilibria. In section five, we showed that the existence conditions of equilibria could be categorized into five regimes. However, in most of these regimes, multiple equilibria were possible. Therefore, knowledge of the *current* values of x, α, and β was insufficient to identify a unique equilibrium. With an understanding of the changes in parameters that lead to tipping points between regimes, we can now use information about *previous* values of the parameters to make deductions about the sequence of equilibrium states in the game.

For instance, consider two alternative four-step histories. Both sequences start in oppressive equilibrium with high returns to extortion and high group cohesion. In the first history, (1) cohesion for both groups drops to near zero, (2) then x decreases to well below zero, (3) then cohesion slowly increases back to its previous level, and (4) finally, x slowly increases. The second history is identical, except that the first two steps are reversed. *The difference in outcomes is substantial*. Following the first step, the groups in the first history fall into conflict and never recover. In the second history, the groups tip into cooperation after step two and spend the remainder of the sequence in that equilibrium.

This example demonstrates how the path dependent properties of this model can help us make inferences about historical questions where ingroup conformity norms are at play. The same logic can be used to make inferences about the likely outcomes of policy.

VII. Discussion and Conclusions

We began this article with two basic questions: 1) How and why do ethnic fall into states of conflict, cooperation, and oppression?, and 2) What policy interventions might move these groups out of conflict and oppression and back towards cooperation? The formal and agent-based models provided in this paper provide first-pass answers to both questions. Our formal model illustrated the necessary conditions for ethnic groups to fall into one of these three states and the five regime types it identified laid the groundwork for understanding how these groups might tip from one state to another. The agent-based model highlighted these tipping dynamics.

We believe that these results carry both theoretical and practical import. Theoretically, they highlight the importance of ingroup conformity on intergroup interactions. This parameter, previously omitted from models of intergroup interactions, interacts with the other two variables in counter-intuitive ways. Mapping this interaction is a small step forward in understanding interethnic relations in general. The practical import of these results trumps the theoretical. If the assumptions of our model are indeed correct, the results suggest the need for a major policy shift in efforts aimed at conflict resolution. In particular, we question the efficacy of approaches based solely on the contact hypothesis.

The results apply to other policy areas as well. In the previous section, we illustrated how the order in which changes are made in the three model parameters matters immensely. In the example we provided of a situation starting in an oppressive equilibrium, switching the order of the first two parameter changes made all the difference. In this example, the first history could easily represent the chain of events that began with the recent American invasion of Iraq. (We will fill in the story here post-MPSA).

These models illustrate how understanding the various values of these parameters and how they interact is essential to understanding the impact any possible policy intervention might have. Our models provide this understanding. However, the models we have presented are just models. We feel certain, however, that they conform to real-world situations. Our next steps will be to operationalize each of the factors in our model and test the model and its predictions empirically.

Possible Model Extensions
Throughout this paper, we have noted potential extensions of this model. Of these, the most promising is the incorporation of collective action. Many of our results depend on the assumption that individual actors have little impact on their group norms. For instance, the inability of players to escape from conflict can be ascribed to their inability to coordinate. Building in the potential for collective action would complicate analysis dramatically, but could greatly improve the realism and explanatory power of the model. To one way of thinking, our model already describes social and economic motives; collective action would add political motives to the mix.

Other modeling variations would allow for heterogeneous players, perhaps with a Gaussian distribution of individual x parameters. This would likely make the boundaries between regimes fuzzy, giving equilibrium transitions more of a stochastic quality. Alternatively, players might be arrayed spatially or on a network, with their preferences biased to reflect local interactions. (See Granovetter (1985) for a good justification of this approach.) In this case, equilibrium change would probably become a network percolation problem as described in Watts (2002), and might exhibit "local convergence and global polarization" (Axelrod, 1997). Additionally, we might give players the ability to reason more deeply from trends or heuristics, allow for the presence of more than two groups, or make endogenous the evolution of group cohesion and returns to extortion.

Appendix A: Proof of proposition two
Proposition 2: No mixed strategy can satisfy the trembling hand condition.

Let p, q ε [0,1] with either p or q, or both as mixed strategies. We will assume that p ε (0,1), that is, p is not 0 or 1. (The proof for mixed q is symmetric.) Consider a representative player of type A, choosing p* from P, the set of best responses given p and q. From proposition one, we know that p ε P in equilibrium, so we will consider p* = p. In order for p,q to define a trembling-hand equilibrium, U(p*|p,q) must be greater than U(p*|p+ε,q+ε) for any small ε. Expanding, this implies that $p^*(\alpha p + (1-\alpha)(2q - 1)) + (1-p^*)(\alpha (1-p) + (1-\alpha)qx) > p^*(\alpha (p+\varepsilon) + (1-\alpha)(2q - 1)) + (1-p^*)(\alpha (1-(p+\varepsilon)) + (1-\alpha)qx)$. Simplifying by subtracting terms, this reduces to $0 > \alpha\varepsilon$. Although this expression holds for negative ε, it does not hold for all ε. Therefore, we have a contradiction. We conclude that p* is not a best response, which means that p cannot exist in P, which means that p,q is not a trembling hand equilibrium.

Appendix B: Existence conditions for conflict equilibrium

For given values of α, β, and x, a conflict equilibrium exists if U(M|A,p,q) >= U(N|A,p,q) and U(M|B,p,q) >= U(N|B,p,q) when p = q = 0. We will show algebraically that these conditions are satisfied for A-type players regardless of α and x. The proof for B-types is symmetric.

$$
\begin{aligned}
&U(M|A,0,0) >= U(N|A,0,0) \\
&\alpha (1-0) + (1-\alpha)(0*x) >= \alpha * 0 + (1-\alpha)(2 * 0 - 1) \\
&\alpha >= (1-\alpha) * -1 \\
&\alpha >= \alpha-1 \\
&0 >= -1
\end{aligned}
$$

Since zero is always greater than -1, we conclude that conflict equilibrium is a possibility regardless of α, β, and x.

Appendix C: Existence conditions for cooperative equilibrium

For given values of α, β, and x, a cooperative equilibrium exists if U(N|A,p,q) >= U(M|A,p,q) and U(N|B,p,q) >= U(M|B,p,q) when p = q = 1. We will solve algebraically for values of α and x under which these conditions are satisfied for A-type players. Conditions for B-types are symmetric.

$$
\begin{aligned}
&U(N|A,1,1) >= U(M|A,1,1) \\
&\alpha * 1 + (1-\alpha)(2 * 1 - 1) >= \alpha (1-1) + (1-\alpha)(1*x) \\
&\alpha + (1-\alpha) >= (1-\alpha) x \\
&1 >= x-\alpha x \\
&\alpha x >= x-1
\end{aligned}
$$

At this point, we divide by x, which gives us two cases, conditional on the value of x.

If x is negative, we get: $\alpha <= (x-1)/x$. However, this condition is non-binding, since (x-1)/x is always greater than one, the maximum possible value of α, when x is negative. Thus, when x is negative, cooperation is always a viable equilibrium.

If x is positive, we get: $\alpha >= (x-1)/x$. For x greater than one, this condition is binding. Thus, when x is less than one, cooperation is always a viable equilibrium. For x greater than one, cooperation is a viable equilibrium if $\alpha, \beta >= (x-1)/x$.

Appendix D: Existence conditions for oppressive equilibrium

For given values of α, β, and x, an A-over-B oppressive equilibrium exists if $U(M|A,p,q) >= U(N|A,p,q)$ and $U(N|B,p,q) >= U(M|B,p,q)$ when $p = 0$ and $q = 1$. The conditions for B-over-A oppression are symmetric.

We begin by solving for equilibrium conditions for the oppressed group, B.

$$U(N|B,0,1) >= U(M|B,0,1)$$
$$\beta * 1 + (1-\beta)(2 * 0 - 1) >= \beta (1-1) + (1-\beta)*0*x$$
$$\beta + (1-\beta) * -1 >= 0$$
$$\beta >= 1/2$$

For the oppressing group, conditions are different.

$$U(M|A,0,1) >= U(N|A,0,1)$$
$$\alpha (1-0) + (1-\alpha) * 1 * x >= \alpha * 0 + (1-\alpha)(2 * 1 - 1)$$
$$\alpha + (1-\alpha) x >= (1-\alpha) * - 1$$
$$\alpha + x - \alpha x >= \alpha - 1$$
$$x+1 >= \alpha x$$

Once again we have two cases, conditional on the value of x.

For positive x, the condition $\alpha <= (x+1)/x$ is non-binding. That is, for $x > 0$, the oppressing group can always sustain an oppressive equilibrium.

For negative x, α must be greater than or equal to $(x+1)/x$ in order to sustain the equilibrium.

Appendix E: Sketch of an existence proof for oppressive equilibrium in Fearon and Laitin's in-group policing game

Here we describe an equilibrium corresponding to our notion of oppressive equilibrium, and sketch a proof to demonstrate that it can indeed be in Fearon and Laitin's inter-ethnic matching game. Unlike the spiraling and in-group policing equilibria described in the original paper, oppressive equilibria are conditional on player's ethnic types as well as the matches between them.

We define oppressive an equilibrium as follows. Members of the oppressing group cooperate with members of their own group who are not in punishment phase and defect against everyone else. A member of the oppressing group who defects against a cooperating player of his own group enters the punishment phase for some set number of rounds. Note that members of the oppressing group are not punished for defecting against members of the oppressed group.

Members of the oppressed group defect against members of their own group who are in punishment phase and defect against everyone else. A member of the oppressed group who defects against *any* cooperating player enters the punishment phase for some set number of rounds. This number might be different for offenses against the in- and out-group.

This asymmetric punishment mechanism enables the oppressing group to defect against members of the oppressed group without penalty. The more surprising aspect of this equilibrium is that the oppressed group continues to cooperate across ethnic lines in spite of receiving the sucker's payoff when playing against the oppressing group. Like the other equilibria outlined by Fearon and Laitin, oppressive equilibrium is only viable if in-group interactions make up a large enough portion of matches and it punishment phases are all long enough. The exact specification of these conditions can be derived by the usual method of comparing current payoffs to discounted future payoffs. From a technical perspective, the proof is a straightforward folk theorem.

Bibliography

Allport, Gordon. 1954. *The Nature of Prejudice*. Massachusetts: Addison-Wesley Press.

Asch, S.E. “Effects of group pressure upon the modification and distortion of judgment,” in H. Geutzknow, ed. Group, leadership, and men. 1951.

Axelrod, R. (1997). The Dissemination of Culture: A Model with Local Convergence and Global Polarization. The Journal of Conflict Resolution, 41(2), 203-226.

Bargal, David. 2004. “Structure and Process in Reconciliation-Transformation Workshops: Encounters Between Israeli and Palestinian Youth.” Small Group Research, Vol. 35(5): 596-616.

--------- 2008. “Group Processes to Reduce Intergroup Conflict: An Additional Example of a Workshop for Arab and Jewish Youth.” Small Group Research, Vol. 39(1): 42-59.

Fearon, J., & Laitin, D. (1996). Explaining Interethnic Cooperation. The American Political Science Review, 90(4), 715-735.

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. The American Journal of Sociology, 91(3), 481-510.

Halabi, Rabah. ed. 2004. Israeli and Palestinian Identities in Dialogue. New Jersey: Rutgers University Press.

Hardin, Russell. 1997. One For All: The Logic of Group Conflict. Princeton: Princeton University Press.

Kakar, Sudhir. 1996. The Colours of Violence. Chicago: University of Chicago Press.

Kalyvas, Stathis. 2007. "Civil Wars." In The Oxford Handbook of Comparative Politics, ed. C. Boix. Oxford: Oxford University Press.

Lustick, Ian. 1980. Arabs in the Jewish State: Israel's Control of a National Minority. Austin: University of Texas Press.

Maoz, Ifat. 2000a. “An Experiment in Peace: Reconciliation-Aimed Workshops of Jewish-Israeli and Palestinian Youth.” Journal of Peace Research. Vol. 37(6): 721-36.

-------- 2000b. “Power relations in intergroup encounters: a case study in Jewish-Arab encounters in Israel.” International Journal of Intercultural Relations. Vol. 24: 259-77.

---------- 2002. “Is there contact at all? Intergroup interaction in planned contact interventions between Jews and Arabs in Israel.” International Journal of Intercultural Relations. Vol. 26: 185-97.

--------- 2004. “Peace Building in Violent Conflict: Israeli-Palestinian Post-Oslo People-to-People Activities.” International Journal of Politics, Culture, and Society. Vol. 17(3): 563-74.

Peterson, Roger. 2002. Understanding Ethnic Violence. Cambridge University Press, 2002.

Pettigrew, T.F. and L.R. Tropp. 2006. “A Meta-Analytic Test of Intergroup Contact Theory.” Journal of Personality and Social Psychology, Vol. 90(5): 751-83.

Schelling, T. (1973). Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities. The Journal of Conflict Resolution, 17(3), 381-428.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory, 4(1), 25-55.

Sherif, M. (1951). Experimental study of intergroup relations. In J.H. Rohrer & M. Sherif (Eds.), Social Psychology at the Crossroads. New York: Harper and Brothers.

Spencer, Michael S., Mikel Brown, Shayla Griffin, and Shaban Abdullah. 2008. “Outcome Evaluation of the Intergroup Project.” Small Group Research, Vol. 39(1): 82-103.

Strauss, Scott. *The Order of Genocide*. Ithaca, New York: Cornell University Press, 2006.

Steele, C.M. A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist 1997(52): 613-29.

Takacs, Karoly. 2007. “Effects of Network Segregation in Intergroup Conflict: An Experimental Analysis.” Connections 27 (2): 59-76.

Tambiah, Stanley. 1996. Leveling Crowds: Ethnonationalist Conflicts and Collective Violence in South Asia. Berkely, CA: University of California Press.

Tessler, Mark. A History of the Israeli-Palestinian Conflict. Bloomington: Indiana University Press, 1994.

Watts, D. (2002). A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99(9), 5766-5771.

Varshney, Ashutosh. 2003. Ethnic Conflict and Civic Life: Muslim-Hindu violence in India.

New Haven: Yale University Press.

---------2007. “Ethnicity and Ethnic Conflict”, in Carles Boix and Susan Stokes, eds. Oxford Handbook of Comparative Politics, Oxford University Press.

Wilkinson, Steven. 2004. Votes and Violence: Ethnic Riots in India. New York: Cambridge University Press.