

Ethnic Norms and Interethnic Violence: Accounting for Mass Participation in the Rwandan Genocide*

RAVI BHAVNANI

Department of Political Science, Michigan State University

This article specifies a simple mechanism – a *behavioral* norm defined in *ethnic* terms – to understand the dynamics of mass participation by reluctant Hutu in the 1994 Rwandan genocide. The mechanism, which has not been analyzed systematically in the scholarly literature, is specified using an agent-based model of *within*-group interaction that captures the role of intra-Hutu coercion in precipitating genocidal violence, yet is general enough to be applied to other group conflicts and contexts. The model highlights the role of individual heterogeneity, within-group punishment, behavioral adaptation, group networks, and ethnic entrepreneurs, and generates a set of results on norm formation and change, a number of which are *not* intuitive. These findings suggest that (1) norms are not equally likely to form in groups with similar aggregate preference; (2) a violence-promoting norm can emerge in a group dominated by moderates, and violence is not the inevitable outcome in a group dominated by extremists; (3) strong punishments are a prerequisite for the emergence of norms that promote interethnic violence or cooperation; (4) interaction patterns matter – for example, the segregation of an ethnic group clearly inhibits norm formation; and (5) an ethnic entrepreneur can effectively amplify norm formation within a group in the absence of strong punishment.

Introduction

In his work on the 1994 genocide in Rwanda, Mamdani (2001: 224) relates how

a respondent named Mectilde described mass Hutu participation in the violence: ‘She gave me a rough count: ten percent helped; 30 percent were forced to kill; 20 percent killed reluctantly; 40 percent killed enthusiastically.’¹ The response illustrates that the Hutu population was not monolithic in its propensity to engage in Tutsi-directed

* I am indebted to Robert Axelrod, Rick Riolo, and Jonas Nart for their comments on this article, as well as to Michael Findley for research assistance. David Backer, Zeynep Bulutgil, Pradeep Chhibber, Raymond Duvall, Lynne Eden, Stathis Kalyvas, David Laitin, Scott Page, Elizabeth Radziszewski, Michael Ross, Phil Shively, Carl Simon, Richard Snyder, Scott Straus, Elizabeth Wood, participants at Stanford University’s *Center for International Security and Cooperation* lecture series, and participants at Yale University’s *Order, Conflict, and Violence* lecture series all provided valuable suggestions, as did two anonymous reviewers. This article makes extensive use of the secondary source literature on Rwanda – Des Forges (1999) and Mamdani (2001) in particular. The simulation data used in this article can be found at <http://www.prio.no/jpr/datasets>. All faults remain my responsibility alone. Please address correspondence to bhavnani@msu.edu.

¹ The quote from Mamdani resonates with the distribution of behavior identified in Browning’s (1993: 168) study of Reserve Police Battalion 101 during the Holocaust: ‘a nucleus of increasingly enthusiastic killers who volunteered for the firing squads and “Jew hunts”; a larger group of policemen who performed as shooters and ghetto clearers when assigned but who did not seek opportunities to kill (and in some cases refrained from killing, contrary to standing orders, when no one was monitoring their actions); and a small group (less than 20 percent) of refusers and evaders.’

violence, in spite of which, hundreds of thousands of Hutu did participate in the genocide.² So whereas for Mamdani (2001: 224), 'It is the 40 percent, those who "killed enthusiastically," who represent the real moral and political dilemma of the Rwandan genocide', I am concerned with explaining why another 50% of Hutus participated, albeit reluctantly or under duress. What, in other words, explains mass participation by reluctant Hutu in violence against Tutsis? A related puzzle concerns the scale and intensity of violence, which was simply unprecedented. Why were previous episodes of Hutu-Tutsi violence localized and contained, while the episode in 1994 consumed the entire country?

I argue that the compulsion and resultant participation of the reluctant in the killing constituted a dramatic behavior shift, with related consequences for the scale and intensity of anti-Tutsi violence. As such, the unprecedented magnitude of violence in 1994 can, to a large extent, be attributed to the emergence of a violence-promoting norm among Rwandan Hutu, a norm that compelled all Hutu – reluctant or otherwise – to participate in the violence. Simply put, Hutu who opposed the genocide and were reluctant to participate in the killing, killed

in the hundreds and thousands because they were left with no other choice: 'kill or be killed'.³ If, as I argue, a violence-promoting norm did emerge in Rwanda and engender mass participation, this raises a number of related questions: Why, for instance, do norms that engender mass participation in violence emerge within some ethnic groups and not others? Can these behavioral norms vary in strength within a given episode of violence? Can these norms also promote nonviolent outcomes such as intergroup cooperation? Is norm formation equally likely in ethnic groups with similar aggregate characteristics? It is equally likely with weak or no punishments?

In this article, I analyze the emergence of *ethnic norms* – rules instituted and enforced *within* an ethnic group to shape the *behavior* of its members toward rivals.⁴ By delineating behavioral expectations in times of conflict and cooperation, norms constitute *one* mechanism to shape group behavior by increasing cohesiveness among co-ethnics and enlarging the set of participants in group action. Therefore, I do not claim that ethnic norms are the only mechanism that matters in episodes of mass violence. Rather, I seek to determine the *conditions* under which expectations of mass participation will emerge by analyzing the dynamics of within-group interaction. Using the Rwandan genocide as a foil, my framework underscores (1) the initial proclivity of group

² While no conclusive estimates of the number of participants have been established, the highest estimates cited by Kagame's former advisor Dusaidi in Gourevitch (1998: 244) suggest that anywhere between 1 and 3 million Hutu participated in the genocide. Mamdani (2001: 7) and Des Forges (1999: 2, 260) estimate the figure to be in the hundreds of thousands (Des Forges also cites a lower figure, referring to tens of thousands of participants), while Straus's (2004a: 172; 2004b: 93) estimate, supported by empirical evidence, places the number of perpetrators between 175,000 and 210,000. One reason for this variance concerns precisely what it means to be a 'participant'. Does it require that one was an actual perpetrator and directly complicit in the violence, or that one helped and was indirectly complicit? For work that addresses the question of why ordinary people *willingly* participate in violence, see Goldhagen (1996), Petersen (2002), De Figueiredo & Weingast (1999), Staub (1989), and Waller (2002). For work that addresses non-participation by ordinary people – the deliberate choice to *not* participate in the killing – see Davenport (2003).

³ Prunier (1995) notes that the compulsion to kill clearly varied across groups and localities in Rwanda, with the government account of spontaneous participation holding true in certain cases, and individual- or community-wide coercion required to elicit participation in other cases.

⁴ Note that ethnic norms, as defined, differ from social norms (Elster, 1989) in that they *are* outcome oriented (if you want to achieve *Y* then do *X*, versus do or don't do *X*), yet resemble social norms in that they apply to oneself and to others, are sustained by sanctions, and offer some scope for choice and manipulation. Note also that this definition differs from cultural or attitudinal norms used to explain the existence of deeply ingrained animosity towards ethnic rivals and explicitly excludes any notion of internalization as a result of the norm itself.

members to engage in interethnic violence and punish co-ethnics who fail to engage in violence; (2) the structure of social networks within the group; and (3) the strength of punishments imposed on individuals whose behavior deviates from emergent group behavior.

My usage follows Kuran (1998: 1), who defines 'ethnic norms' as 'the behavioral codes that individuals must follow to retain the acceptance of their ethnic groups . . . sustained partly by sanctions'. It also follows Axelrod (1986: 1097), insofar as 'a norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way'. Both definitions make it possible to study the growth and decay of norms based on how often a particular course of action is followed and the frequency of punishment for those who do not follow it.

To emphasize, my immediate interest is in norms that engender mass participation, rather than outcomes – whether an episode of interethnic conflict yielded a particular level of violence – though the latter are obviously a latent concern.⁵ In order, then, to explore variation in the strength of ethnic norms both within and across episodes of interethnic violence – and thus extend my analysis beyond Rwanda to derive a set of general propositions about the emergence of ethnic norms – I utilize complex systems theory and its simulation by agent-based modeling (ABM). This methodology, in particular, lends itself well to the study of emergent phenomena (Axelrod, 1997; Holland, 1999; Page, 1999). By repeating simulations and observing trajectories of participation, one can learn about the distribution of outcomes associated with different initial conditions, as well as about processes of norm formation and change within ethnic

groups. Moreover, ABM makes it possible to tie variation in individual characteristics, interaction, and behavior to variation in the strength of norms that engender mass participation in ethnic violence, thereby revealing how micro-level dynamics generate macro-level events.

The article is organized as follows. In the next section, I analyze the puzzle of mass participation in the Rwandan genocide in greater detail. I then discuss the norm-based framework in relation to other intragroup mechanisms used to engender collective behavior, before specifying an agent-based model of norm formation and change. I then present the model's results, and conclude with a discussion of my findings.

Conventional Explanations for Mass Participation

Explanations for mass participation by the large 'machete-wielding . . . social underbelly of the genocide' (Mamdani, 2001: 225) in Rwanda are diverse, and rightly so, given the heterogeneity of participants.⁶

One common set of explanations highlights the salience of *ethnic identity* to explain the existence of deeply ingrained animosity between Hutu and Tutsi dating back, in some instances, prior to the colonial period (Braeckman, 1994; Chrétien, 1995; Mamdani, 2001; Uvin, 1996). One variant of the identity-based explanation emphasizes the particularity of Rwandan culture, arguing that a *culture of conformity* and

⁵ For an analysis that explicitly explores variation in the scale and duration of ethnic violence, see Bhavnani & Backer (2000).

⁶ I draw upon a more detailed review of the literature by Straus (2004a), who parses theories for mass participation in violence into the following categories: (1) abnormality, deviancy, and dispositional theories; (2) frustration-aggression and deprivation theories; (3) intercommunal enmity and identity theories; (4) collective behavior theories; (5) 'ordinary' evil, social influence, and obedience theories; (6) security fears and dilemmas; and (7) theories of material opportunism. With respect to mass participation, Straus suggests that the motivation to participate is likely to have been heterogeneous, and that several theories are probably right.

unquestioning obedience to authority existed amongst the Hutu, facilitating mass participation in the genocide (Gourevitch, 1998; Khan, 2001; Prunier, 1995; Scherrer, 2002).

Another set of explanations highlights *structural violence* (Uvin, 1998) including *discontent* driven by relative deprivation. Exclusion, oppression, and land and resource scarcity are all abetted by a culture of animosity, generating violence (Des Forges, 1999; Khan, 2001; Mamdani, 2001; Prunier, 1995; Staub, 1989). Indeed, Mamdani (2001: 231) goes so far as to say that mass participation by hundreds of thousands of Hutu who had never killed before or been previously affected by ethnic violence can be explained by a desire to both maintain and expand the ranks of the Hutu middle class.

Psychological explanations suggest that 'deviant' individuals may be predisposed to engage in violent behavior, or that social-psychological forces bring out latent group identities which are conditioned by early childhood experiences and hateful representations of rivals (Kakar, 1996). The result is that fear, anxiety, and panic all serve to heighten the salience of group identity, making the crowd more susceptible to manipulation. Rumors, propaganda, and a catalyst can then set the crowd off, resulting in 'collective bloodlust' (Scherrer, 2002: 113). Thus, specific *triggering events* – such as the Tutsi-led Rwandese Patriotic Front (RPF) invasion and subsequent assassination of President Habyarimana – culminate in a situation of collective psychosis and begin to explain mass Hutu participation in violence.

Still other explanations emphasize the role of *institutional structure* in facilitating mass participation – state capacity and penetration in particular – given the highly rigid and hierarchical organization of Rwandan society into *préfectures*, *communes*, *collines*, *secteurs*, and *cellules* (Des Forges, 1999; Prunier, 1995; Verwimp, 2006). Extending explanations that underscore strict hierarchical

implementation of the center's genocidal project, Straus (2004a) argues that there is significant variation in *when* the genocide began in various communes and in *how* the violence started. Specifically, he identifies four main patterns: administratively led genocide; internal challenge to existing officials; military-led mobilization; and external incursion. Like Straus, Davenport (2003) adopts a micro-level approach and suggests that there are *préfectures* in which little killing took place, whereas the scale of violence and extent of complicity in others was high. His explanation focuses on structural factors, rationalist-political factors, and individual survival factors, yet stands in contrast to Straus (2004a), who reports little regional variation in the overall *level* of violence with the exceptions of communes liberated at an early stage by the RPF and protected by Operation Turquoise.

The Limits of Conventional Explanations

In sum, explanations for mass participation by Hutu in violence against Tutsi are heterogeneous, albeit suitably, given the complexity of the event. Nonetheless, most acknowledge that for a core group of Rwanda's Hutu elite the need to implement a *final solution* to eradicate Rwanda's Tutsi population can be understood in terms of the threat posed to the Hutu regime by the 1990 RPF invasion and subsequent compromise negotiated at Arusha in 1993 (Des Forges, 1999; Kuperman, 2004; Lemarchand, 2002; Mamdani, 2001; Prunier, 1995).⁷ Responding to an imminent loss in power, the Hutu *génocidaires* – drawn primarily from the army, social, and political elite – framed the collective Hutu interest in terms of the Tutsi threat and ensuing need for a final solution to eliminate this threat.

⁷ Kuperman (2004) argues that the RPF expected their invasion to provoke genocidal retaliation, but either grossly underestimated this cost or viewed the cost as acceptable for their return to power in Rwanda.

The explanations presented above nonetheless make a less than convincing argument for participation by reluctant Hutu in the genocidal violence. Ascribing mass participation to a deeply ingrained animosity between Hutu and Tutsi, for instance, fails to discern that anti-Tutsi sentiment has never been a constant among Hutu in Rwanda, and that even in 1994 – when as a group they had the most to fear from the impending Tutsi invasion – Rwanda's Hutu population was *not* monolithic in its propensity to engage in violence against Tutsi. In a similar vein, characterizing Rwanda as a culture of conformity and obedience fails to account for selective Hutu defiance of the genocidal regime, or the need for advance planning, propaganda, and persuasion to generate mass participation among the Hutu.⁸

Moreover, while there is little doubt that 'structural' factors that pertain to economic and social conditions have important implications for participation, the conventional preference for tracking structural factors – which either tend to remain constant or are replicated to some degree in most episodes of conflict – appears to be overstated. Similarly, while there is little doubt that Rwanda's hierarchical institutional system was designed to maximize social control, it did not change in any significant manner prior to the genocide.⁹

Finally, the contention that Hutu stood to gain economically from a mass Tutsi exodus fails to explain why participation extended

beyond criminal or disadvantaged classes (Mamdani, 2001: 202) or why Hutu in Butare – one of the poorest and most overpopulated *préfectures* in Rwanda – were the last to participate in the killing (Des Forges, 1999: 262; Straus, 2004a: 381). And the large number of *willing* participants in the genocide ostensibly undermines theories of deviant behavior since, for the majority of perpetrators, participation – far from being considered deviant behavior – was, in fact, sanctioned by the state.

Clearly, members of an ethnic group share affinities with and commitments to one another that should facilitate collective action. Yet, while group consciousness – abetted by fear, customary obligation, and, arguably, a culture of obedience – is *necessary* for collective action, it proves *insufficient* when the individual costs of participation are high. Mamdani (2001: 219) underscores this same point, observing that fear of Tutsi domination effectively eroded opposition by Hutu moderates but failed to generate enthusiasm for killing. Thus, while the explanations discussed above plausibly account for the behavior of more enthusiastic or willing perpetrators, they are less concerned with addressing *why* or *how* participation extended beyond this group to ultimately involve hundreds of thousands of reluctant or unwilling Hutu.¹⁰ For the latter, I argue, participation was to a significant

⁸ Mamdani (2001: 200) takes issue with this explanation as well: 'To believe that ordinary Rwandans killed, in their hundreds and thousands, and perhaps more, because of a congenital transhistorical condition – "a culture of fear" or of "deep conformity" – would require stretching one's sense of credibility.'

⁹ As noted by an anonymous reviewer, 'structural' conditions *did* change in Rwanda between 1989 and 1994, as a result of falling coffee prices as well as the devastating consequences of war. Bhavnani & Backer (2000) nonetheless argue that even if structural factors were to remain constant, variation in individual-level factors could still generate variation in the scale and duration of violence.

¹⁰ An exception is Straus (2004a: 227–232), who finds that, in his sample, intra-Hutu coercion explained 64% of the stated motivation for Hutu participation in the genocide, while obedience accounted for 12.9% of the responses. He notes that in-group pressure is more likely to explain participation by 'less violent' individuals, whereas fear or anger is more likely to explain participation by 'more violent' individuals. Straus also stresses the importance of war (legitimization, military thinking and involvement, and large-scale reprisal killing) and intra-Hutu political competition as central causes of the genocide and argues that the effectiveness of in-group sanctioning is related to Rwanda's history of state institutions (dating back to the pre-colonial period) and geography (settlement and topography). A revised version of his argument is forthcoming (Straus, 2006).

extent coerced, an explanation I turn to in the next section.

Using Norms To Explain Group Behavior

In the context of violence, norms that shape group behavior have emerged in settings as diverse as Cambodia, Guatemala, and former Yugoslavia. In Cambodia, the Khmer Rouge manipulated norms of 'face' and 'honor' in an effort to shame those who refused to kill, with the result that 'defectors' or 'traitors' were often killed themselves (Fein, 1993; Hinton, 1998).¹¹ During *La Violencia* in Guatemala, the 'guilty' were massacred while commanders and patrollers who failed to obey orders were threatened or killed, all in an effort to engender collective silence and compliance (Zur, 1994). In Croatia, the Serbian population was terrorized into submission and ethnically mobilized by Serbian guerrillas, police, and army units, while in Serbia, the SPS (Socialist Party Serbia) accused those who questioned the war of treason, sent reservists from opposition strongholds to the front first, and tortured or killed Serbs identified as 'disloyal' to the Serbian cause (Gagnon, 1995; Kuran, 1998; Rieff, 1995; Vulliamy, 1994). In each instance, conformity and participation increased as a result of compulsion, with grave consequences for those who failed to comply. This is illustrated most vividly in the case of the Rwandan genocide.

Accounts of the Rwandan genocide provide clear evidence that mass participation was by no means spontaneous or forthcoming. According to Des Forges (1999: 201), between 6,000 and 7,000 Hutu responded to the initial call for violence, casting doubt on the level of popular par-

ticipation in the genocide.¹² In an attempt to extend the scope of participation to all Hutu, moderate or otherwise – albeit as part of a deliberate effort to shift the blame for the killings away from themselves – the *génocidaires* implemented a calculated strategy.¹³

Appeals were made for Hutu unity against a common Tutsi enemy, and Hutu were imparted a moral and social obligation to eliminate the *inyenzi* (literally 'cockroaches', a derogatory name for the Tutsi rebels) and the *ibyitso* (their accomplices within Rwanda).¹⁴ Uncooperative officials were rapidly eliminated. Hutu who sheltered or hid Tutsi were punished, with sanctions that ranged from fines and beatings to rape and death at the hands of fellow Hutu. The reluctant were taught how to kill and compelled to participate in the killing – whether in their own locales or in locales where the killing of strangers purportedly facilitated the task (Des Forges, 1999: 262, 313).

In conjunction with punishment, a parallel system of rewards for participation in the killing emerged. Des Forges's (1999:

¹² The limited scope of participation was driven in some measure by an intra-Hutu rivalry in which Hutu moderates – adversaries of the Coalition pour le Défense de la République (CDR) and Mouvement Révolutionnaire Nationale pour le Développement (MRND) became the first victims of the genocide. Prunier (1995) puts the number of moderate Hutu killed during the genocide at 50,000.

¹³ In support of this view, Des Forges (1999: 231) reports that: 'Executing an extermination campaign rapidly required the mobilization of hundreds of thousands of ordinary people, tens of thousands to actually slaughter and others to spy, search, guard, burn and pillage. In some situations, crowds were needed immediately and for only a few days to participate in a massacre; in others, a reliable supply of long-term "workers" was required to do patrols, man the barriers and track survivors. Bagosora, the AMASASU, the CDR and Kangura had foreseen that turning out large numbers of civilians was the only way to attack an "enemy" dispersed in the population. As Karanira had said in his radio speech of April 12, this "war" had to become everyone's responsibility.'

¹⁴ Gourevitch (1998: 83) notes that 'Following the logic of state ideology – that identity equals politics and politics equals identity – all Tutsis were considered to be RPF "accomplices," and Hutus who failed to subscribe to this view were counted as Tutsi-loving traitors.'

¹¹ A parallel argument in the Rwandan case is that the genocidal regime used customary norms or obligations, referring to the killing as work or *umuganda* (communal work), in an apparent effort to distinguish 'indigenous' Hutu from 'non-indigenous' Tutsi. For a more detailed analysis, see Mamdani (2001: 193–194).

300) description of the reward structure indicates that it was nationally imposed and (1) assigned pillaged goods to the looter, with the exception of valuables that went to local authorities; (2) allocated land to the commune for rent or reallocation; and (3) allocated crops to individuals or local authorities.

Despite the documented use of rewards to promote cooperation with the genocidal agenda, the argument presented in this article remains punishment-centered and rests upon the premise that rewards are more effective in promoting limited compliance, whereas punishment is more effective in promoting universal compliance with expected behavior.¹⁵ Thus, I would argue that punishment was instrumental in mobilizing the mass of reluctant or unwilling Hutu to participate in the genocidal violence.

Needless to say, the demonstration effects of punishment for the reluctant (and rewards for the more willing set of participants) were powerful and led to a steep decline in the readiness of Hutu to protect Tutsi, as well as a steep rise in the killing (Des Forges, 1999: 221). Thus, where fear of Tutsi driven by the RPF invasion may have induced some Hutu to participate in early April, *fear of fellow Hutu* was decisive in extending participation over a wider range of perpetrators (Des Forges, 1999: 322).

In short, mass participation by Rwanda's Hutu population was essential for the genocide. Yet, it is important to note that behavioral conformity among the Hutu was by no means assured. Rather, conformity was cultivated through compulsion over the course of the genocide. Not only were behavioral expectations unambiguous, in that they were repeatedly broadcast by the *génocidaires*

over the private radio station (Radio Télévision Libre des Mille Collines) and appeared in the print media (*Kangura*), but these expectations were also summarily enforced by the Hutu Power movement and *interahamwe*. Individuals, in turn, adapted to these expectations in different ways, with the end result that hundreds of thousands of Hutu participated in the genocide, effectively lending social support to the state's genocidal agenda.

Modeling Norm Formation and Change

In contrast to normative behavior that is internalized over time through socialization, individual adherence to a behavioral norm, defined in ethnic terms, is driven in large measure by the threat of sanctions for non-compliance. By defining a norm as comprising two basic components – expected behavior and punishment for deviation from this behavior – my definition shifts the emphasis from legitimacy to sanctioning systems used to shape behavior. Indeed, given the fact that norms can rationalize violence, one would be hard pressed to argue that such norms are self-reinforcing or acquire 'legitimacy' in the absence of punishment. As defined, ethnic norms resemble other intragroup mechanisms – most notably in-group policing, identified by Fearon & Laitin (1996) as an effective means of promoting interethnic cooperation. Yet the norm-based framework I develop also differs from this mechanism in some notable ways.

For interethnic cooperation to constitute an equilibrium it is reasonable to assume that most members of an ethnic group cooperate, with the exception of a few defectors or spoilers who are identified and sanctioned by co-ethnics. With respect to interethnic violence, Fearon & Laitin acknowledge that in-group policing could backfire, or be turned on its head, if a group employs

¹⁵ Oliver (1984: 125–126) states this succinctly: 'Positive incentives are given to cooperators: if everyone cooperates in collective action everyone must be rewarded; if everyone defects, no one is rewarded. Negative incentives are just the opposite: if everyone cooperates no one gets punished, but if everyone defects, everyone must be punished.'

sanctions to induce its members to participate in attacks against rivals (rather than punishing them for doing so). However, the assumption that most members of an ethnic group collectively 'cooperate' or participate in interethnic violence is implausible, since the cost of partaking in violence is high, if not prohibitive (Gould, 1999), which thereby serves as a deterrent to all but the most extreme group members.

Thus, where mass participation in violence occurs under compulsion, the analogy to in-group policing – albeit the 'dark side' of in-group policing – weakens, since the cost associated with monitoring and enforcing compliance by the n th group member is also likely to be prohibitive.¹⁶ By specifically delineating which actions are 'right' in periods of heightened group conflict (and, by implication, which actions are 'wrong' when groups cooperate), ethnic norms can effectively reduce these costs and generate collective compliance. By implication, the strength of a norm would dictate how typical this behavior is of the group as a whole.

In Figure 1, distributions a and b violate our intuition about an ethnic norm. In a , behavioral expectations are normally distributed among group members, with no clear consensus for violence or cooperation vis-à-vis nominal rivals. In b , one faction within the group strongly supports interethnic violence while another faction supports interethnic cooperation. As a result, neither of these distributions meets the criterion of near universal support for one particular outcome as depicted by n – where a majority of group members support neutrality, indicating the *anticipated creation* of a norm

¹⁶ Fearon & Laitin (1996: 726) duly note that the costs associated with in-group policing are likely to increase with group size. Building on this work, Bhavnani & Backer (2000) specify a computational model that incorporates genocidal or 'dark side' norms; however, the authors do not explicitly address the collective-action problem groups face in mobilizing moderates (or restraining extremists) and, thus, the possibility that norms do not effectively take root within a given group.

where consensus had previously not existed. In contrast, the move from a to n signifies a change in expected behavior from absence of a clear consensus to violence – the *unanticipated creation of a norm* that supports violence.¹⁷

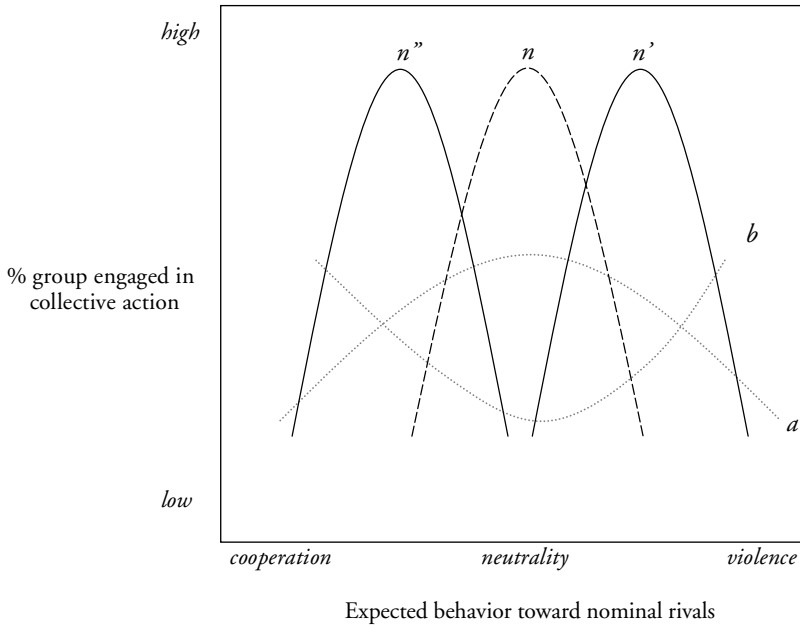
In developing a model to capture norm formation and change, existing analyses of both inter- and intragroup dynamics – Axelrod's (1986) norms game, Kuran's (1998) analysis of ethnic norms and ethnification cascades, Bhavnani & Backer's (2000) use of genocidal norms to explain variation in interethnic violence, Fearon & Laitin's (1996) examination of in-group policing and interethnic cooperation – all serve as useful starting points.

An Agent-Based Model of Ethnic Norms

Agent Characteristics The model consists of a finite population of agents from the same ethnic group. Each agent is characterized by a level of *animosity* towards ethnic rivals (i) and *tolerance* for co-ethnics who do not share their animosity (p), both of which are defined on the unit interval. I refer to agents with high animosity as extremists (i.e. those with a high propensity to engage in violence), and to agents with low animosity as moderates (i.e. those with a high propensity to prevent violence and cooperate with nominal rivals). Agents also vary in terms of their *influence* over co-ethnics, defined as the number of opportunities (r) each agent has to interact with and influence other members of the ethnic group. An agent's level of influence is set to a high value ($r = 5$) if they are both extremist and intolerant ($i > 0.75$ and $p < 0.25$), otherwise this is set to a low value ($r = 1$). High influence, therefore, affords extremists more opportunities to interact with others, and captures the role of ethnic entrepreneurs. Finally, the intensity of *punishment* administered to co-ethnics – a

¹⁷ I am indebted to David Laitin for a similar example.

Figure 1. Defining an Ethnic Norm



Distributions *a* and *b* violate our intuition about an ethnic norm, as neither of these meets the criterion of near universal support for one particular outcome. In contrast, the move from *a* to *n* represents *anticipated norm creation*; the move from *a* to *n'* (or *n''*) represents *unanticipated norm creation*; and the move from *n* to *n'* (or *n''*) represents *aggregate change*.

group attribute that is set to a single value for all agents – is given by $q \in (0, 1)$.

Group Characteristics Since punishment is used to bring agent behavior into conformity with group practice, group *networks* that connect agents (co-ethnics) to one another are instrumental in determining behavioral trajectories and the outcome of efforts to achieve collective compliance.¹⁸ Group networks, in effect, determine *how* and *how often* like-minded agents observe and sanction the behavior of agents with con-

trasting or opposing views. I consider two network structures in which agents are paired randomly with partners. The first structure places no restriction on interaction, such that any agent may interact with any other agent in the group. The second structure restricts interaction by dividing the group into two equally sized neighborhoods and by confining interaction to each neighborhood.¹⁹

Agent Interaction and Adaptation In the model, time is comprised of discrete events

¹⁸ Despite their prominence for recruiting participants for protest or rebellion in the social movement literature, networks have received limited attention in the context of ethnic violence. Brass (1997) notes that all riot-prone towns do have – to a greater or lesser degree – informal organizational networks that serve to mobilize members, whereas Varshney (2003) bases his argument on the existence of interethnic networks that promote civic engagement and reduce conflict. Both nonetheless fail to specify the structure of these networks and how this structure may vary.

¹⁹ Two additional network structures, one dividing the population of agents into four neighborhoods and the other permitting networks to form endogenously – with agents seeking out others least like themselves to interact with – were considered. Results from model runs indicated that there was little difference between the two-neighborhood and four-neighborhood networks, and little difference between the random-unrestricted and endogenous networks. As a result, these networks were dropped from the analysis. Note that a variety of additional network structures may be used in the model.

and 1,500 events constitute an episode (i.e. one run of the model). Each event presents some agent A with the opportunity to contact some other agent B through a social network connecting members of the group to one another. If the difference in animosity between A and B exceeds A 's threshold of tolerance (if $|i_A - i_B| > p_A$), then A punishes B , and B updates her animosity (i_B) and tolerance (p_B). As such, agents *adapt* their behavior in response to interaction in a variety of ways. For instance, agents can update their behavior to conform to the expectations of other agents they encounter, to adhere more closely to perceptions of local behavior, or to conform to expectations broadcast to the group as a whole.

To capture these responses, I specify four update rules. Using Rule 1, agents update animosity and tolerance to reflect the type of the punisher. Using Rule 2, agents update animosity and tolerance to reflect aggregate values for these characteristics in their neighborhood. And using Rules 3 and 4, all agents update animosity and tolerance to a global, exogenously determined level of animosity and tolerance. Equations (1) and (2) specify the general form of the update rules for animosity and tolerance respectively:

$$i_B^{t+1} = i_B^t (1 - q) + \delta q \quad (1)$$

$$p_B^{t+1} = p_B^t (1 - q) + \phi q \quad (2)$$

The values of δ and ϕ are then set to distinguish between the four specific update rules with the value of q set to denote the intensity of punishment administered to group co-ethnics, as noted above. For Rule 1, δ and ϕ reflect the punisher (A 's) type ($\delta = i_A$ and $\phi = p_A$). For Rule 2, δ and ϕ reflect the mean level of extremism and the mean level of tolerance in the agents' neighborhood ($\delta = i_N$, $\phi = p_N$). For rules 3 and 4, δ and ϕ reflect global, exogenously determined levels of animosity and tolerance ($\delta = i_G$ and $\phi =$

p_G),²⁰ such that Rule 3 promotes cooperation and tolerance (i_G is low and p_G is high) while Rule 4 promotes violence and intolerance (i_G is high and p_G is low). To reiterate, I vary each of these rules as parameters in the model.

To summarize, based on the nature of the encounter (i.e. whom one encounters and what their preferences are), behavioral expectations, the strength of punishment administered to co-ethnics for deviant behavior, and the structure of social relations that characterize the group, agents adapt their behavior to a greater or lesser extent.

Analyzing Norm Formation and Change I use the model described above to explore the link between group composition, punishment regimes, update rules, and network structure, on the one hand, and norm formation and change on the other hand. In all, each of the 80 distinct model variants was run 100 times with different random seeds. During each run, 100 agents interacted over the course of 1,500 events – with a single event providing one agent the opportunity to influence the behavior of another agent. Table I describes the range of model parameters considered in this study and provides measures for the dependent variables.²¹

To measure norm formation and change, I track changes in group animosity ($\Delta \bar{i}$), as well changes in group uniformity ($\Delta \sigma_i$). These two measures enable me to distinguish between four outcomes: (1) *no change* (NC) – the absence of norm formation, as depicted in Figure 2A ($\Delta \bar{i}$ is low, $\Delta \sigma_i$ is low); (2) *aggregate change* (AC) – a significant change in aggregate group animosity with no commensurate increase in group uniformity, as depicted in Figure 2B ($\Delta \bar{i}$ is high, $\Delta \sigma_i$ is low); (3) *anticipated norm creation* (ANC) – a large increase in group uniformity with no

²⁰ Note that Rules 3 and 4 cannot be used when the group is divided into two neighborhoods.

²¹ These parameter settings were selected to analyze only the most distinctive model variants.

Table I. Model Parameters and Output Variables

<i>Independent variables (parameter settings)</i>	<i>Dependent variables (measures)</i>
Update rule	
Rule 1: update to type of punisher A s.t. $\delta = i_A, \phi = p_A$	
Rule 2: update to local mean s.t. $\delta = \bar{i}, \phi = \bar{p}$	Change in group animosity
Rule 3: update to global value s.t. $\delta = i_G = 0.1, \phi = p_G = 0.9$	$\Delta \bar{i} = \bar{i}_{t_{1500}} - \bar{i}_{t_0}$
Rule 4: update to global value s.t. $\delta = i_G = 0.9, \phi = p_G = 0.1$	
Network	Change in group uniformity
1: no neighborhoods, unrestricted interaction	$\Delta \sigma_i = \sigma_{i,t_{1500}} - \sigma_{i,t_0}$
2: two neighborhoods, restricted interaction	
Animosity	Change in group tolerance
low: $i \sim N(0.1, 0.3)$	$\Delta \bar{p} = \bar{p}_{t_{1500}} - \bar{p}_{t_0}$
high: $i \sim N(0.9, 0.3)$	
Tolerance	Punishment rate
low: $p \sim N(0.1, 0.3)$	$\hat{q} = \frac{\sum q}{1500}$
high: $p \sim N(0.9, 0.3)$	
Punishment	
weak: $q = 0.1$	
strong: $q = 0.9$	
Entrepreneurs	
0: absent, $r = 1$	
1: present, $r = 5$ if $i > 0.75$ and $p < 0.25$, else $r = 1$	

This table depicts the values that each of six model parameters can take, resulting in a total of 80 model variants. When the network structure divides the group into two neighborhoods, only Rules 1 and 2 are used, whereas Rule 2 is not used in conjunction with unrestricted interaction. Individual animosity (i) and tolerance (p) are drawn from a normal distribution with low (0.1) or high (0.9) means and fixed variance (0.3), and with values of i and p truncated at 0 (lower bound) and 1 (upper bound), resulting in empirical means of 0.25 (low) and 0.75 (high).

commensurate change in group animosity, as depicted in Figure 2C ($\Delta \bar{i}$ is low, $\Delta \sigma_i$ is high); and (4) *unanticipated norm creation* (UNC) – a large change in group animosity with a commensurate increase in group uniformity, as depicted in Figure 2D ($\Delta \bar{i}$ is high, $\Delta \sigma_i$ is high). The measures – presented in Table II – are used to distinguish between low and high levels of group animosity, on the one hand, and low and high levels of group uniformity, on the other hand.²²

²² Note that uniformity must increase for ANC to occur, whereas the change in mean animosity for UNC can either be positive or negative depending on the group's initial position and growing support for interethnic violence or cooperation. Also note that the results of a simple sensitivity analysis, in which the thresholds for coding the

For purposes of analysis, I group the 80 model variants into 8 'cases' (ten variants in each case) to reflect similarities in the initial level of animosity, tolerance, and the strength of punishment. This division is reflected in Table III, where I present data from the simulation. Given the sizeable number of model variants, the data for variants in which no change occurred are omitted.

change in group animosity and uniformity as 'high' were changed, suggest that (1) both lowering and increasing the threshold for group animosity has little effect; (2) lowering the threshold for group uniformity results in a large shift from the NC to ANC outcome; and (3) increasing the threshold for group uniformity results in a large shift from ANC, UNC, and AC outcomes to the NC outcome, effectively eliminating UNC as a possible outcome.

Figure 2. Variation in Norm Formation

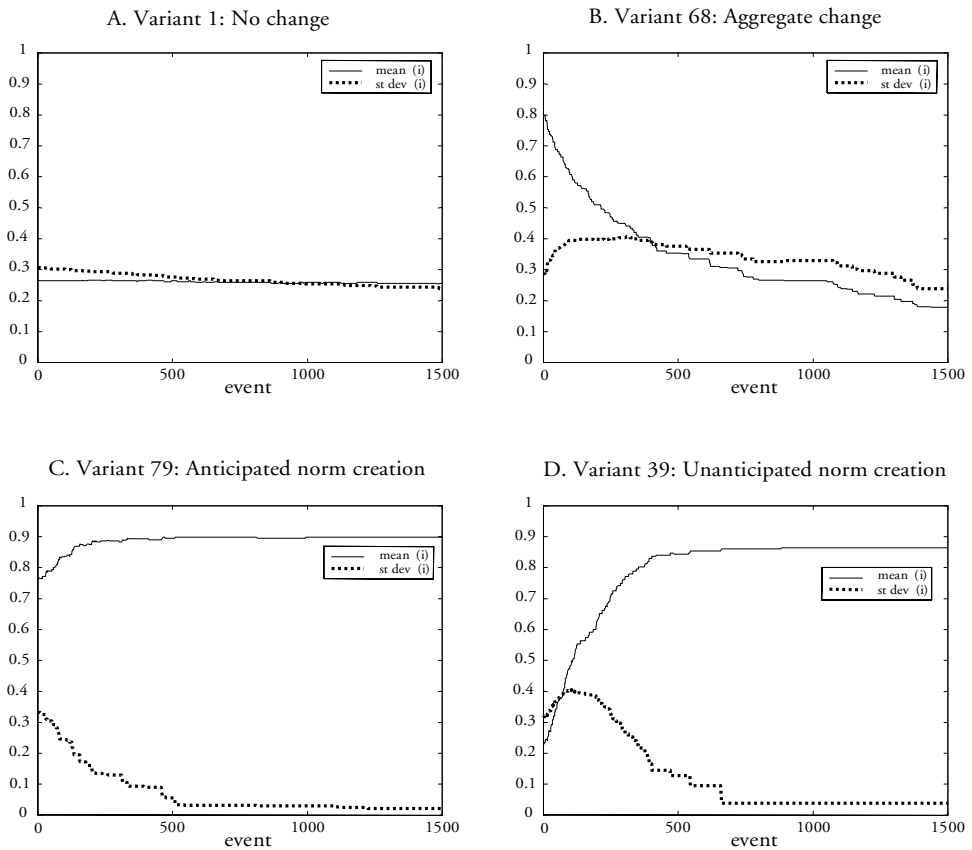


Figure 2 (A–D) depicts levels of aggregate group animosity \bar{z}_t and group uniformity σ_t over the course of single runs of specific model variants.

Results

Results from the runs of these different model variants indicate that the formation of violence-promoting or violence-preventing norms within an ethnic group (cases of ANC and UNC), while the exception, is not a low-probability event ($p = .375$). In groups with low initial levels of animosity (Cases 1–4), anticipated norm creation (ANC) occurred nine times out of 40, whereas in groups with high initial levels of animosity (Cases 5–8) ANC occurred 16 times out of 40, which effectively suggests that high (low) levels of group animosity are not sufficient for the formation of violence-promoting (violence-

preventing) norms.²³ It follows that norms are not equally likely to form in groups with similar aggregate preferences.²⁴

²³ Note that, in the model runs analyzed here, the lack of symmetry in ANC in groups characterized by low/high initial animosity is due to the presence of violence-promoting entrepreneurs. As such, it would have been possible to run the model with no entrepreneurs to generate symmetric results, with violence-preventing entrepreneurs who are equally intolerant, or with a balanced mix of both types of entrepreneurs. Removing violence-promoting entrepreneurs from the runs would have decreased the total number of reported ANC outcomes by approximately seven and the total number of UNC outcomes by approximately four.

²⁴ Kuran (1998) reaches a similar conclusion, noting that the incidence of ethnic activity is likely to be poorly correlated with aggregate group characteristics, yet his model does not analyze how 'ethnification' gets started within a group, relying instead on exogenous shocks to activate the process.

Table II. Rules for Coding Norm Formation

		<i>Change in group uniformity</i>	
		high $\Delta\sigma_i \leq \sqrt[3]{.05}$	low $\Delta\sigma_i > \sqrt[3]{.05}$
<i>Change in group animosity</i>	high $ \Delta\bar{i} \geq 0.5$	Unanticipated norm creation (UNC)	Aggregate change (AC)
	low $ \Delta\bar{i} < 0.5$	Anticipated norm creation (ANC)	No change (NC)

Values for $\Delta\bar{i}$ and $\Delta\sigma_i$ consist of averages derived from 100 runs of each model variant. The thresholds (0.5 and $\sqrt[3]{.05}$) were selected to generate a reasonable distribution of model outcomes, as well as to serve as a baseline for sensitivity tests with different thresholds. Results from this analysis are reported in note 22.

That said, while groups are more likely to generate norms that reflect their aggregate preferences, violence-promoting norms can emerge in groups dominated by moderates (low animosity) and violence is not the inevitable outcome in groups dominated by extremists (high animosity). Looking at Case 4 in Table III, where group members are initially opposed to violence, one finds the emergence of violence-preventing norms in variants 31–38, as well as the emergence of a violence-promoting norm in variants 39 and 40. Likewise, under Case 8 where group members are initially predisposed toward violence, variants 71–76, 79, and 80 depict the emergence of a violence-promoting norm, whereas in variant 78 a violence-preventing norm emerges.

In sum, where aggregate initial levels of animosity were low, unexpected norm creation (UNC) – represented by the formation of a violence-promoting norm – occurred four times out of 40, whereas in groups with high initial aggregate levels of animosity, UNC – represented by the formation of a violence-preventing norm – occurred only once. The initial composition of an ethnic group, therefore, appears to be an inadequate predictor of the creation of violence-promoting or preventing norms, given the incidence of UNC in the model in groups characterized by both high and low aggregate levels of animosity.

Second, the simulation results suggest that strong punishments are a prerequisite for the emergence of norms that promote

Table III. Simulation Results

<i>Model variant</i>	<i>Network</i>	<i>Update rule</i>	<i>Entrepreneur</i>	<i>Punishment rate</i> \dot{q}	<i>Change in tolerance</i> $\Delta\hat{p}$	<i>Change in uniformity</i> $\Delta\sigma_i$	<i>Change in animosity</i> $\Delta\hat{i}$	<i>Outcome</i>
Case 2: low animosity, low tolerance, weak punishment								
12	1	1	1	0.61	-0.13	-0.25	0.14	ANC
Case 3: low animosity, high tolerance, strong punishment								
21	1	1	0	0.33	-0.64	-0.25	0.05	ANC
22	1	1	1	0.40	-0.63	-0.25	0.13	ANC
29	1	4	0	0.09	-0.58	-0.29	0.58	UNC
30	1	4	1	0.09	-0.59	-0.28	0.60	UNC
Case 4: low animosity, low tolerance, strong punishment								
31	1	1	0	0.76	-0.23	-0.25	0.02	ANC
32	1	1	1	0.72	-0.23	-0.28	0.14	ANC
35	2	2	0	0.14	-0.01	-0.25	-0.06	ANC
36	2	2	1	0.10	-0.01	-0.26	-0.03	ANC
37	1	3	0	0.15	0.52	-0.26	-0.17	ANC
38	1	3	1	0.13	0.54	-0.25	-0.16	ANC
39	1	4	0	0.11	-0.15	-0.28	0.60	UNC
40	1	4	1	0.10	-0.15	-0.30	0.61	UNC
Case 6: high animosity, low tolerance, weak punishment								
52	1	1	1	0.55	-0.20	-0.30	0.14	ANC
54	2	1	1	0.55	-0.20	-0.29	0.13	ANC
56	2	2	1	0.35	0.00	-0.24	0.09	ANC
60	1	4	1	0.42	-0.12	-0.28	0.16	ANC
Case 7: high animosity, high tolerance, strong punishment								
62	1	1	1	0.38	-0.69	-0.29	0.12	ANC
66	2	2	1	0.11	-0.02	-0.23	0.14	ANC
68	1	3	1	0.14	0.14	-0.09	-0.54	AC
69	1	4	0	0.08	-0.52	-0.29	0.16	ANC
70	1	4	1	0.09	-0.56	-0.29	0.16	ANC
Case 8: high animosity, low tolerance, strong punishment								
71	1	1	0	0.77	-0.23	-0.25	0.00	ANC
72	1	1	1	0.47	-0.24	-0.32	0.15	ANC
73	2	1	0	0.71	-0.23	-0.23	-0.01	ANC
74	2	1	1	0.37	-0.25	-0.27	0.15	ANC
75	2	2	0	0.14	-0.01	-0.24	0.07	ANC
76	2	2	1	0.15	-0.01	-0.24	0.13	ANC
77	1	3	0	0.17	0.57	-0.18	-0.58	AC
78	1	3	1	0.09	0.60	-0.23	-0.61	UNC
79	1	4	0	0.12	-0.15	-0.30	0.16	ANC
80	1	4	1	0.07	-0.16	-0.31	0.16	ANC

In the column labeled *Outcome*, ANC stands for *anticipated norm creation*, UNC for *unanticipated norm creation*, and AC for *aggregate change*. Outcomes involving *no change* are not reported. In particular, Case 1 (low animosity, high tolerance, weak punishment) and Case 5 (high animosity, high tolerance, weak punishment) both generated *no change* for all model variants tested.

interethnic violence or cooperation. This finding holds with respect to both ANC and UNC. Of the 25 instances of ANC, 20 occurred with strong punishment, as did every instance of UNC. Nonetheless, while the strength of punishment clearly matters, the results indicate that the rate of punishment varies between these two outcomes, with the average punishment rate for ANC (34.6%) far exceeding the average for UNC (9.7%). What explains this apparently counter-intuitive result? One possibility is that the rate of punishment is likely to be higher when agents update their behavior to the type of the punisher, given that more agent interaction is required to achieve uniformity within the group, as opposed to when agents update to conform to some common (global) behavioral expectation.²⁵ For example, in variant 31 (ANC), agents update to the punisher's type and the punishment rate is 76%, whereas in variant 39 (UNC), agents update globally and the punishment rate is close to 12%. Since every instance of UNC occurred when agents updated to a global rule (Rule 3 or 4) – in contrast to 7 of 25 instances of ANC – the rate of punishment in the set of ANC outcomes exceeded that in the set of UNC outcomes.

Third, the findings indicate that interaction patterns matter: given restrictions on interaction among members of a group, ANC is the more likely outcome, whereas UNC depends entirely on the absence of such restrictions. All five instances of UNC in the variants examined occurred when members of the group were *not* segregated, when agents updated their behavior globally, and when intragroup punishments were strong. Group segregation – the division of the group into neighborhoods – therefore appears to inhibit UNC, a finding of some practical import.

²⁵ In general, the lower the mean level of tolerance within a group and the weaker the strength of punishment, the higher the rate of punishment is likely to be.

Fourth, the model suggests that entrepreneurs – i.e. influential members of an ethnic group – amplify both ANC and UNC by changing group animosity and increasing group uniformity. More specifically, the simulation results suggest that the role of entrepreneurs in generating norms that shape group behavior is critical precisely when within-group punishment – punishment by co-ethics – is weak. Under these circumstances, entrepreneurs were *necessary* for the emergence of norms. Their influence is attenuated when punishment is strong; they affected the emergence of norms in a small fraction (1/9) of the variants run under these conditions.

Discussion

This article specifies one mechanism – in the form of a behavioral norm defined in ethnic terms – to understand the dynamics of mass participation by reluctant Hutu in genocidal violence. The mechanism – specified in the form of an agent-based model of within-group interaction – highlights the salience of individual heterogeneity, within-group punishment, behavioral adaptation, the structure of group networks, and ethnic entrepreneurs. Each of these factors featured prominently in the Rwandan episode.

To begin with, Rwandan society was *not* strictly segregated along ethnic lines. Above and beyond sharing a common history, language, culture, and religion, Hutu and Tutsi lived in mixed communities, interacted socially and professionally, and, perhaps most notably, even married each other.²⁶ As

²⁶ According to Newbury (1998), in the pre-colonial period a family's ethnic identity could change over time through social recognition, rather than individual choice, thereby permitting the children of Hutu who accumulated wealth and married Tutsi to be recognized, in subsequent generations, as Tutsi and those Tutsi who suffered economically to be recognized as Hutu. While Newbury notes that this incentive structure changed after independence when it became advantageous to be recognized as Hutu, the end result is that an estimated 25–50% of Rwandans can claim Hutu and Tutsi among their great-grandparents.

a result, it is difficult to attribute hatred or fear of Tutsi uniformly to all Hutu.

In order to compel reluctant Hutu to participate in violence directed at Tutsis, the extremist regime in Rwanda initially resorted to the use of punishments – ranging in severity from the destruction of personal property to incarceration and death. Calls to violence and associated threats against those who remained inactive were broadcast openly over the radio and communicated repeatedly in consciousness-raising meetings, speeches, songs, and slogans, all couched in terms of a communal or national duty and amplified by the pervasiveness of radio ownership and use in Rwanda.

The result was that behavioral expectations were common knowledge among the Hutu, moderate or otherwise, and that punishments administered to those reluctant or unwilling to participate in the killing had strong demonstration effects, making repeated punishment unnecessary as the genocide progressed. As Des Forges (1999: 221) notes, cases in which Hutu protectors of Tutsi ‘were raped, beaten or killed . . . were widely known in local communities and often led other Hutu to refuse or end assistance to Tutsi’. The pressure to participate was reinforced by entrepreneurs, in most cases *burgomasters* or commune leaders who could directly recruit participants for the killing.

During the genocide, Hutu extremists interacted with Hutu moderates in a purposive effort to influence the latter’s behavior leaving the reluctant with a stark choice: conform to expected behavior or suffer the consequences. Mobile militias that visited communes in which the killing was less than enthusiastic, reluctant Hutu sent to kill Tutsi in adjoining communes based on the calculation that anonymity would facilitate participation, and radio broadcasts that reached every corner of the country, all suggest that group interaction was not bounded, with the

result that no Hutu could be sure of whom they would encounter or whom they could trust. Many Hutu were turned in by relatives, friends, or neighbors who bore grudges against them, sought favors, or coveted property, with the result that the prospect of being labeled a Tutsi sympathizer extended the violence over a wider range of targets. Whereas some Hutu may have willingly redressed their grievances against particular Tutsi, under these conditions they were pressed to be indiscriminate. Escaping, as a result, became more difficult for members of the targeted group – Hutu moderates and all Tutsi alike.

The agent-based model developed in this article captures some of the salient features of intra-Hutu mobilization during the Rwandan genocide, yet moves beyond the particularities of this case to determine the general conditions under which violence-promoting (or violence-preventing) norms will emerge within ethnic groups. The simulation results indicate that (1) ethnic norms are not equally likely to form in groups with similar aggregate preferences and that violence-promoting norms can emerge in groups initially opposed to violence (and, conversely, that violence-preventing norms can emerge in groups initially predisposed towards violence); (2) ethnic norms can promote or prevent intergroup violence; (3) ethnic norms are more likely to emerge with strong punishments, unrestricted interaction among group members, and in the presence of ethnic entrepreneurs; (4) anticipated norm creation (which occurs in 25 of the 80 variants examined) is more common than unanticipated norm creation (which occurs in only 5 of the 80 model variants examined).²⁷

The framework developed in this article may be extended in a number of ways. One

²⁷ In the remaining variants examined, there are 2 instances of ‘aggregate change’ and 48 instances of ‘no change’.

possibility is to seed the model with aggregate data from different *préfectures* or *communes* in Rwanda (Davenport, 2003), as well as with micro-level data on perpetrators (Straus, 2004a), with the aim of explaining regional variation in the onset, duration, and scale of Tutsi-directed violence. Alternatively, the model may be modified to examine the *strategic* interaction between moderates and extremists, to consider what separates the participation histories of neutral individuals from those of the numerous unwilling (and willing) executioners, and to examine the dynamics of norm formation *across* rival ethnic groups.

To conclude, the Rwandan genocide suggests that social norms are indeed reversible – that, under ruthless compulsion, ordinary people can engage in violence *en masse*, an improbable but nonetheless daunting prospect for multi-ethnic societies. While the available evidence seems to support the contention that a strong violence-promoting norm emerged during the genocide, the extent to which members of the Hutu community internalized this norm remains unclear. In certain instances, the killings persisted despite the official ‘pacification’ campaign and orders to end the violence, whereas in other instances individual Hutu and entire communities ceased killing Tutsi (Des Forges, 1999: 291–297). Contrary to expectations, Hutu militia then began to turn on fellow Hutu, with the end result that ‘a system dedicated to the destruction of Tutsi provided no security for Hutu either’ (Des Forges, 1999: 301).

Moving beyond Rwanda, events in India – where communal violence between Hindus and Muslims is endemic – demonstrate that it is also possible to grow norms that support interethnic cooperation in conflict-ridden societies. The severe riots that took place in the power-loom town of Bhiwandi in 1967, 1970, and 1984 established the city’s reputation as a hotbed of communal violence

between Hindus and Muslims (Madon, n.d.). Yet this same city, renowned for its history of communal violence, remained calm as Hindu–Muslim violence spread throughout India in 1992.

Bhiwandi’s shift from high levels of violence to cooperation can be traced to the period between 1988 and 1991, when 70 *mohalla*, or neighborhood committees, were established in an effort to stem Hindu–Muslim violence (Khopade, 1998; Engineer, 1995). These committees explicitly sought to exclude extremists, thus avoiding a mistake made by the conventional peace committees that exist in most Indian cities.²⁸ Through weekly meetings, the creation of a robust network to monitor troublemakers, preventive arrests, riot-control training exercises, and social events, a violence-preventing norm was painstakingly established over the course of three years, leading to a dramatic decline in Hindu–Muslim violence.²⁹ Bhiwandi’s success in preventing violence was touted nationally and followed by attempts to repeat the experiment in numerous other cities across India.

References

- Axelrod, Robert, 1986. ‘An Evolutionary Approach to Norms’, *American Political Science Review* 80(4): 1095–1111.
- Axelrod, Robert, 1997. *The Complexity of Cooperation*. Princeton, NJ: Princeton University Press.

²⁸ The flawed belief behind the Indian government’s experiment with peace committees was that by placing group leaders and other miscreants in positions of authority, these troublemakers would stem rather than participate in violence. The plan backfired, as members from rival groups used the committees as a forum to increase their leverage at the bargaining table, by demonstrating their ability to incite violence. This has been the experience in many Indian cities since then.

²⁹ In all, approximately 250 Hindus and 150 Muslims were arrested in Bhiwandi to prevent communal disturbances from escalating, with the arrests carried out in a decidedly non-partisan manner. See Khopade (1998).

- Bhavnani, Ravi & David Backer, 2000. 'Localized Ethnic Conflict and Genocide: Accounting for Differences in Rwanda and Burundi', *Journal of Conflict Resolution* 44(3): 283–307.
- Braeckman, Collette, 1994. *Rwanda: Histoire d'un Génocide* [Rwanda: History of a Genocide]. Paris: Fayard.
- Brass, Paul, 1997. *Theft of an Idol: Text and Context in the Representation of Collective Violence*. Princeton, NJ: Princeton University Press.
- Browning, Christopher, 1993. *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York: HarperCollins.
- Chrétien, Jean-Pierre, 1995. 'Rwanda, les Médias du Génocide' [The Media of the Genocide], Collection 'Hommes et Sociétés'. Paris: Karthala.
- Davenport, Christian, 2003. 'Mass Killing and the Oases of Humanity: Understanding the Rwandan Genocide and Resistance', paper presented at LiCEP 7, Laboratory in Comparative Ethnic Processes, Yale University, 9–11 May (<http://yale.edu/ycias/ocvprogram/licep/7/davenport/davenport.pdf>).
- De Figueiredo, Rui & Barry Weingast, 1999. 'The Rationality of Fear: Political Opportunism and Ethnic Conflict', in Barbara Walter & Jack Snyder, eds, *Civil Wars, Insecurity, and Intervention*. New York: Columbia University Press (261–302).
- Des Forges, Allison, 1999. *Leave None To Tell the Story: Genocide in Rwanda*. New York: Human Rights Watch.
- Elster, Jon, 1989. 'Social Norms and Economic Theory', *Journal of Economic Perspectives* 3(4): 99–117.
- Engineer, Ashghar Ali, 1995. *Lifting the Veil: Communal Violence and Communal Harmony in Contemporary India*. Hyderabad: Sangam.
- Fearon, James & David Laitin, 1996. 'Explaining Interethnic Cooperation', *American Political Science Review* 90(4): 715–735.
- Fein, Helen, 1993. 'Revolutionary and Antirevolutionary Genocides: A Comparison of State Murders in Democratic Kampuchea, 1975–1979, and in Indonesia, 1965–1966', *Comparative Studies in Society and History* 35(4): 796–823.
- Gagnon, V. P., 1995. 'Ethnic Nationalism and International Conflict: The Case of Serbia', *International Security* 19(3): 130–166.
- Goldhagen, Daniel, 1996. *Hitler's Willing Executioners: Ordinary Germans and the Holocaust*. New York: Knopf.
- Gould, Roger, 1999. 'Collective Violence and Group Solidarity: Evidence from a Feuding Society', *American Sociological Review* 64(3): 356–380.
- Gourevitch, Philip, 1998. *We Want To Inform You that Tomorrow We Will Be Killed with Our Families: Stories from Rwanda*. New York: Farrar, Straus and Giroux.
- Hinton, Alexander, 1998. 'Why Did You Kill? The Cambodian Genocide and the Dark Side of Face and Honor', *Journal of Asian Studies* 57(1): 93–122.
- Holland, John, 1999. *Emergence: From Chaos to Order*. New York: Persea.
- Kakar, Sudhir, 1996. *The Colors of Violence: Cultural Identities, Religion, and Conflict*. Chicago, IL: University of Chicago Press.
- Khan, Shaharyar, 2001. *The Shallow Graves of Rwanda*. London: I.B. Tauris.
- Khopade, Suresh, 1998. 'Bhiwandi Riots 1984', unpublished manuscript.
- Kuran, Timur, 1998. 'Ethnic Norms and Their Transformation Through Reputational Cascades', *Journal of Legal Studies* 27(2): 623–659.
- Kuperman, Alan, 2004. 'Provoking Genocide: A Revised History of the Rwandan Patriotic Front', *Journal of Genocide Research* 6(1): 61–84.
- Lemarchand, René, 2002. 'Disconnecting the Threads: Rwanda and the Holocaust Reconsidered', *Journal of Genocide Research* 4(4): 499–518.
- Madon, D. P. (Judge, High Court, Bombay), n.d. *Report of the Commission of Inquiry into the Communal Disturbances at Bhiwandi, Jalgaon and Mahad in May 1970*.
- Mamdani, Mahmood, 2001. *When Victims Become Killers: Colonialism, Nativism, and the Genocide in Rwanda*. Princeton, NJ: Princeton University Press.
- Newbury, David, 1998. 'Understanding Genocide', *African Studies Review* 41(1): 73–97.

- Oliver, Pamela, 1984. 'Rewards and Punishments as Selective Incentives: An Apex Game', *Journal of Conflict Resolution* 28(1): 123–148.
- Page, Scott, 1999. 'Computational Models from A to Z', *Complexity* 5(1): 35–41.
- Petersen, Roger, 2002. *Understanding Ethnic Violence: Fear, Hatred and Resentment in 20th Century Eastern Europe*. Cambridge: Cambridge University Press.
- Prunier, Gérard, 1995. *The Rwandan Crisis: History of a Genocide*. New York: Columbia University Press.
- Rieff, David, 1995. *Slaughterhouse: Bosnia and the Failure of the West*. New York: Simon & Schuster.
- Scherrer, Christian, 2002. *Genocide and Crisis in Central Africa: Conflict Roots, Mass Violence, and Regional War*. Westport, CT: Praeger.
- Staub, Ervin, 1989. *The Roots of Evil: The Origins of Genocide and Other Group Violence*. Cambridge: Cambridge University Press.
- Straus, Scott, 2004a. *The Order of Genocide: Race, Power, and War in Rwanda*. Doctoral dissertation, Department of Political Science, University of California, Berkeley.
- Straus, Scott, 2004b. 'How Many Perpetrators Were There in the Rwandan Genocide?' *Journal of Genocide Research* 6(1): 85–98.
- Straus, Scott, 2006. *The Order of Genocide: Race, Power, and War in Rwanda*. Ithaca, NY: Cornell University Press.
- Uvin, Peter, 1996. 'Tragedy in Rwanda', *Environment* 38(3): 6–29.
- Uvin, Peter, 1998. *Aiding Violence: The Development Enterprise in Rwanda*. West Hartford, CT: Kumarian.
- Varshney, Ashutosh, 2003. *Ethnic Conflict and Civic Life: Hindus and Muslims in India*. New Haven, CT: Yale University Press.
- Verwimp, Philip, 2006. 'Machetes and Firearms: The Organization of Massacres in Rwanda', *Journal of Peace Research* 43(1): 5–22.
- Vulliamy, Ed, 1994. *Seasons in Hell: Understanding Bosnia's War*. London: Simon & Schuster.
- Waller, James, 2002. *Becoming Evil: How Ordinary People Commit Genocide and Mass Killing*. Oxford: Oxford University Press.
- Zur, Judith, 1994. 'The Psychological Impact of Impunity', *Anthropology Today* 10(3): 12–17.

RAVI BHAVNANI, b. 1969, PhD in Political Science (University of Michigan–Ann Arbor, 2003); Assistant Professor, Michigan State University (2005–); doctoral training in comparative politics and methodology with emphasis on agent-based modeling; research interests: ethnic violence and civil war in sub-Saharan Africa and South Asia.