

Why are heterogenous communities inefficient?

Theory, history and an experiment ^{*,†}

David Hugh-Jones

Carlo Perroni

University of East Anglia

University of Warwick and CAGE

26th February 2015

Abstract

We examine why heterogenous communities may fail to provide public goods. Current work characterizes sanctioning free-riders as an under-supplied public good. We argue that often free-riders can be punished by the coordinated action of a group. This punishment can be profitable, and need not be undersupplied. But the power to expropriate defectors can also be used to expropriate outgroups. Heterogenous societies may be inefficient because minorities, rather than free-riders, are expropriated. Even if this is not so, groups' different beliefs about the reasons for expropriation may make the threat of punishment less effective at preventing free-riding. We illustrate our theory with evidence from California mining camps, contemporary India, and US schools. In a public goods experiment using minimal groups and a profitable punishment institution, outgroups were more likely to be punished, and reacted differently to punishment than ingroup members.

KEY WORDS: Group Coercion, Social Heterogeneity

JEL CLASSIFICATION: H1, H4, N4, D02

*We are grateful to CAGE for financial support.

†Comments and correspondence should be addressed to David Hugh-Jones, School of Economics, University of East Anglia, D.Hugh-Jones@uea.ac.uk

1 Introduction

Ethnically diverse communities appear to be less successful at providing public goods (Alesina, Baqir and Easterly, 1999, Alesina and La Ferrara, 2000, Easterly 2001*b*). It is important to understand why, since many societies have increasing levels of ethnic diversity. Existing work suggests that in these societies, sanctioning free-riders is harder, and therefore more free-riding takes place. Proposed mechanisms include differing expectations of being sanctioned, and a social network structure that makes it harder to identify and punish free-riders from out-groups (Miguel and Gugerty, 2005; Habyarimana et al., 2007; Habyarimana, Humphreys and Posner, 2009).

This paper proposes a new mechanism to explain the phenomenon. In diverse societies, collective sanctioning easily shades into expropriation of vulnerable minorities. This makes it less effective and more dangerous. It is less effective because targeting minorities redirects coercive power away from the enforcement of cooperation; also, even the perception of ethnic bias in sanctioning weakens both the sanction's deterrent power and its legitimacy. It is more dangerous because misperceived sanctions may lead to counter-mobilization by the minority, and to ethnic conflict. Below (in Section 2) we give several examples, but the following dialogue between a white teacher and a black student illustrates our basic mechanism.

Teacher: (to student): Please sit down and stop talking.

Student: I was only seeing if I could borrow a pencil and a piece of paper for that quiz you were talking about.

Teacher: You know you're supposed to be in your seat.

Student: But you'll give me a zero if I don't have a quiz paper.

Teacher: (slightly exasperated): Sit down. You're supposed to bring those things to class or to borrow them before class.

Student: (voice rising): Why you picking on me? *You don't pick on white kids who borrow a piece of paper.*

(Grant, 1988; emphasis added.)

In this episode, conflict arises from a difference in beliefs about an unobservable counterfactual: would the teacher act the same if the student were white? We conjecture that situations like this are widespread.

Our argument is based on a new analysis of the problem of collective action. The standard view is that the key problem for social groups is to prevent individual free-riding, which can be done by sanctioning free-riders; because sanctioning free-riders is costly to the sanctioner but benefits the community as a whole, enforcement is itself a public good. This is known as the “second-order collective action problem”.¹

We start from the observation that while the setup costs of institutions to monitor free-riding must be borne by the community, punishing free-riders need not itself be costly.² For, a credible threat of punishment can be converted to a transfer from the guilty party which benefits the punisher; and, when enough people coordinate to punish somebody, they run few risks and may benefit materially by expropriating the person.³ In these cases, punishment is individually rational for the punishers and the threat of punishment is credible. Then, sanctioning free-riders is simply a coordination problem: the threat of expropriation of defectors is credible (as coordinated expropriation is gainful), and can act as a deterrent if defection triggers a switch from non-expropriation to coordinated expropriation – both of which are equilibrium outcomes of the ensuing coordination game. Indeed, a recent review of the anthropological literature shows that much free-riding is punished costlessly (Guala, 2012; see also Boehm, 2001).

The ease of punishment by groups, however, can itself create inefficiencies. Majorities may target individuals inappropriately, falsely identifying unusual, unpopular or vulnerable people as free-riders. This has happened across many social settings and historical periods.⁴ Furthermore, the real or perceived threat of expropriation by a majority can make individuals unwilling to

¹Ostrom (1990) describes the “general problem” of common pool resource situations as “how to organize to avoid the adverse outcomes of *independent action* [our italics]” (p. 29) and argues “[s]ubstantial benefits have to be obtained to make costly monitoring and sanctioning activities worthwhile” (p. 36). See also Frohlich and Oppenheimer (1970); Oliver (1980); Yamagishi (1986).

²Recent critiques of the importance of costly punishment have focused on the evolutionary difficulty of sustaining costly punishment motivations, and the inefficiency of costly punishment institutions in laboratory games (Dreber et al., 2008). We agree, but would add that punishment need not be costly, and that the most powerful forms of punishment (those administered by groups) will typically be gainful to the punishers. Also, the costly punishment behavior observed in laboratory experiments, where it may be the only punishment option available, may be a carry-over from habits formed in real-world settings where punishment is not costly.

³Lanchester’s (1916) “square law” for physical conflict between an individual and a group states that the expected *total* harm to group members decreases with the group’s size, and the expected harm to the individual increases with the group’s size. This is also consistent with theoretical predictions in e.g. trade wars, where larger countries or a large coalition of small countries can experience a net gain from retaliation against a small country.

⁴Examples include English popular culture (Thompson, 1992), witchcraft accusations in Europe, Africa and America (Horsley, 1979; Evans-Pritchard and Gillies, 1976; Boyer and Nissenbaum, 1974), and anti-semitism.

take risks or stand out, for example by individual enterprise (Platteau, 1996). Thus, although individual free-riding may arise due to failure to coordinate on punishment, societies also face the problem of expropriation by majorities causing deadweight losses and weakening individual incentives. A society's ability to provide public goods – its “state capacity” – ultimately hinges on its ability to harness the coercive power of groups and ensure that this is directed towards defectors.⁵ By focussing on individual free-riding to the exclusion of collective expropriation, social scientists may have misread the problems of self-governing communities.

The problems of collective expropriation are particularly severe in ethnically diverse communities. Ethnic markers provide a natural coordination device for picking a target. Furthermore, ethnicity has the advantage for expropriators that it limits the targeted group to a known set of individuals. As Pastor Niemoller famously pointed out, those who fail to protect the vulnerable from oppression risk becoming victims themselves next time; this risk may discourage arbitrary expropriation. But the risk is minimised when victimizers are always from one ethnic group and victims from another.

Even without actual expropriation of minorities, ethnic diversity can cause inefficiencies: if minority markers act as salient labels for coordinated expropriation, *or are perceived to do so* by minorities, then group expropriation loses its power to deter defection. As the dialogue quoted earlier suggests, minority targets of sanctioning face a difficult problem of causal identification: are they targeted for breaking the rules, or because of their identity? Not surprisingly, different communities often disagree radically about the interpretation of particular events. In turn, punishers who believe that they are being targeted unfairly will perceive few incentives, material or moral, not to free-ride. Minority communities may even counter-mobilize in response to the targeting. These possibilities make cross-ethnic punishment of free-riders more risky and less effective. In settings where a small disagreement can spark a riot, the price of ethnic peace is tolerating bad behavior. It is therefore natural that, as previous work has found, technologies of community punishment remain bounded by ethnicity.

Some ingredients of our argument are echoed in the existing literature on community sanctioning. The successful CPR systems analysed in *Governing The Commons* (Ostrom, 1990) often involved fines and redistribution of “ill-gotten gains” as sanctions; Weissing and Ostrom (1991)

⁵There is a small but growing literature on the determinants of state capacity (Besley and Persson, 2010, 2009), but this abstracts from the question of how enforcement can be sustained in groups in the absence of a binding social contract, and thus neglects the mechanisms we highlight here.

examine a game in which individuals monitor each other for rule violations, and receive rewards when they observe them. Other authors have examined group coordination as a means of punishing free-riders. Weingast (1997) models how citizens can coordinate to preserve the rule of law against infringements by rulers. Experiments on coordinated punishment in the context of public goods games include Casari and Luini (2009). Recent work on “mobbing” shows that majorities are willing and able to coordinate on vulnerable individuals (Abbink and Doğan, N.d.). Our experiment below allows “mobbing” to punish free-riders.

Some experimental work has examined public good provision in different societies. Habyarimana et al. (2007; 2009) conduct experiments in a Ugandan slum, examining the causal mechanism behind diversity’s effect on public good provision. They classify possible mechanisms into preference, technology and strategy selection mechanisms, and argue that network links between coethnics facilitate punishing defectors. Herrmann, Thöni and Gächter (2008) show that in public goods games with punishment, societies with weak norms of participation have more “antisocial punishment”, i.e. punishment of those who contribute to the public good.

Finally, there is a large literature on violent ethnic conflict (Horowitz, 1985, 2001; Posen 1993; Fearon and Laitin 2003; Caselli and Coleman, 2006; Esteban and Ray, 2008). However, few papers draw the link between non-violent, peacetime inefficiency and the shadow of conflict. Below, we discuss how different groups may come into conflict as each group expropriates individuals from the other.

We stress three key empirical implications of our theory. First, the problems of ethnically divided societies are often thought to lie in deep cultural differences. However, even when there is a strong shared culture, or where different cultural norms are mutually well-understood, the mere fact of visible difference may be enough to make one group a potential target. For example, in Bosnia, Muslims, Serbs and Croats lived in mixed villages and, while maintaining cultural differences, also shared identities as villagers, and were knowledgeable about each others’ customs (Bringa, 1993); yet, during the Yugoslav conflict, different identities swiftly became a basis for conflict. Second, the central mechanism making heterogenous communities less efficient is the threat and reality of coordinated expropriation. If so, then removing the possibility of expropriation should remove the difference. Third, part of this mechanism is factual and normative ambiguity. Where free-riding is easily observable and clearly defined, minority free-riders are unlikely to feel unfairly victimized if they are targeted, but as the scope for different interpretations increases, the problems are likely to grow.

We use a laboratory experiment to test these three implications. We implement heterogeneity in a “minimal” way, as a publicly visible label. We vary the availability of coordinated punishment. Additional treatments vary the ambiguity of the situation. Before describing our experiment, we develop our argument in a series of examples.

2 Informal enforcement and ethnic heterogeneity

2.1 Diversity, expropriation and conflict in California mining camps

In 1849, shortly after California was ceded to the United States, gold was discovered in the territory. The legal status of gold claims remained unclear until 1866. Furthermore, state authority was distant or absent in the California territory: there was no police force, and in the first years even the identity of the legal authorities was unclear. As a result, miners had to make and enforce their own rules.

As early histories of the gold rush emphasized, they did so successfully. American miners, with experience of United States institutions like the town meeting, set up “miners’ meetings” to allocate property rights and make other rules. The camps were not a Hobbesian war of all against all, but an ordered anarchy (Umbeck, 1977). Property rights – “claims” to a particular area – were enforced by the majority. A miner who violated others’ claims could expect to have his own claim considered nonexclusive and open to jumpers, in an example of punishment as profitable expropriation. Other crimes were also collectively punished by the population after short and informal trials (Howe, 1923).

Not all camps were equal, though. Stewart (2009) gathers evidence on twenty-five early mining camps, and shows that ethnically homogenous camps were more likely to be successful at enforcing property rights. His explanation is that members of different ethnic groups were less willing to cooperate, due to different norms and lower altruism towards outgroup members, or less able to cooperate and coordinate, due to cultural and linguistic differences. This is the standard explanation described above. However, more recent historical work suggests a further reason. There was a dark side to mining camp law. Since rule violators could have their own claims expropriated, groups of miners faced a temptation to find others guilty of violations. Thus, punishment of rule-breakers shaded into arbitrary expropriation. Non-Anglo-Americans were particularly often targets: “[t]he French, the Spanish Americans, and the Chinese each tended

to be separate elements in the population.... from an early date the majority displayed toward these nationalities a persistent antagonism that was easily triggered into bullying, persecution and ostracism” (Paul, 1980). A classic example is the “Chilean war” of 1849:

Anglos had begun mining some dry diggings that Chileans had worked earlier, and when the latter returned for the winter the two groups soon were at odds. In a called meeting the Anglos formed a mining district and drafted regulations, beginning with “No foreigners shall be permitted to work at these mines.”... (Paul 1980, p. 241).

Ramon Gil Navarro, an Argentinian miner, described the affair: “Twenty-two Americans came over to the camp of the Chileans and had coffee with them.... While they were doing this, they went about taking the weapons of the Chileans without their noticing.... They were all taken to the house of the judge, who extorted 150 ounces of gold from them.” Later there was a pitched battle between Chileans and Anglo-Americans; Chileans were arrested and brought before a judge, three were hanged and others possibly had their ears cut off. Each of these sides had their own “judge” to legitimize their actions (Navarro, 2000). The failure of ethnically diverse mining camps was thus due not only to individual free-riding, but also to rampant expropriation of minority group members by the majority.

Why should ethnic minorities in particular be targeted? Coethnics’ greater ability to coordinate, or their in-group loyalty and outgroup dislike, are possible reasons. Another is to do with dynamics. If a coalition tries to target another person in the group for expropriation, then other group members, including members of the coalition, may fear that if the attempt succeeds, they will be next in line. This reduces the willingness of coalition members to take part in the first place, and increases the willingness of others to defend the target. However, when the target is in a clearly defined minority group, then future expropriation attempts may be expected also to fall on that group. In other words, group membership provides a coordination device which safely separates persecutors from victims.

2.2 Ambiguity, conflict and sanctioning in India

When punishment shades into expropriation, ethnic divisions cause a further problem. In ethnically polarized communities, punishment institutions are plagued by the ambiguity between targeting a deviator and targeting a minority. Many ethnic riots are sparked by the violation of

one community's norms by someone from the other community (Horowitz, 2001). These incidents are commonly seen very differently by different sides.

Brass (1997) describes an incident in 1982 in the Indian town of Daphnala. A father reported that his daughter had been kidnapped. Three suspects, two Muslims and a Hindu, were arrested, subsequently released, then rearrested. Locals became convinced that the police were not taking the investigation seriously and an angry crowd assembled at the police station, led by a prominent local member of the Bharatiya Janata Party (BJP). Brickbats were hurled and a *lathi*-charge was made by the police. Later police investigation revealed that the girl was a prostitute and her "father" was in fact her pimp.

This story was interpreted differently by different participants. BJP politicians described the police as protecting Muslim rapists. According to them, both Hindus and Muslims had been involved in the anti-police disturbance; they describe the protesters as motivated by "civic pride". The police, the District Magistrate and Congress politicians saw the disturbance as an attempt to stir up Hindu-Muslim tensions, with only Hindus involved, motivated by "communal passion". In theory this is an empirical matter. In practice, figuring out the truth in such situations can be very difficult. As Brass puts it, "it is out of such ambiguous materials that mob violence and Hindu-Muslim riots are created" (see also Brass, 2004). The issue in game-theoretic terms is that different actors have conflicting beliefs. A crowd's collective action is a protest against police injustice; or it is a communal riot. In India, these differing beliefs are encouraged by political actors with an interest in ethnic tensions. At the same time, each side insists that the other side, not it, is motivated by communalism.⁶

There are several possible outcomes. Attempts to sanction minority members can lead to violent resistance from the minority. Or expropriation may succeed under the guise of punishment. In some cases, inter-ethnic civic associations may be able to promote peace by scotching rumors, thus harmonizing different groups' beliefs (Varshney, 2003). What matters for us is not only the possibility of ethnic conflict *per se*, but its effect on peacetime relations. The ambiguity between punishment and expropriation makes it risky to impose punishment on those from other communities. As a result, sanctions can only be imposed within separate communities. Hence the commonly-expressed attitude, among ethnic groups at peace, that "we mind our own business".

⁶In California mining camps also, the line between punishment and exploitative expropriation was contested by different groups, who indeed possessed their own sources of legitimacy, the judges.

2.3 Ambiguity and legitimacy in US schools

Even when the threat of violence is absent, the ambiguity of punishment may cause inefficiency. If some groups perceive that they are targeted based on their identity and not their actions, this will have two effects. First, group members have little incentive to behave well, since they believe targeting cannot be affected by their behavior. Second, they are likely to resent the punishment as illegitimate and unfair, and this may further increase their willingness to free-ride (Rabin, 1993).

After desegregation in “Hamilton High” , the school described by Grant, 1988, when white teachers punished black pupils, each side interpreted this differently: the teachers thought that impartial rules were being fairly applied, while the pupils saw the rules as biased towards white standards and applied unfairly. This is the context of the dialogue quoted in the introduction. An anecdote from another school shows how difficult it can be to choose between such interpretations. A black pupil doesn’t stop talking as she passes the vice-principal on the stairs. He makes her go back and come up quietly. She does so, but in a slightly mocking way: “there is the barest hint of parody in her primness, which Russell does not miss”. The encounter ends with her being sent to the teacher’s office for the day. Would a white child have been perceived or treated differently? (Ferguson, 2001, cited in Akerlof and Kranton, 2010).

Such differences in perceptions led to problems with discipline:

Some teachers associated black children with classroom disruption. They sent from the room children who had no conflict at all with other teachers. On the other side, some black children were particularly sensitive to slights because of their experience of being criticized for their race. They might take offense at a teacher’s action and enter into conflict with him when similar white children would not have.

(Metz, 1979)

These differences of perception may be due to in-group biases in information processing (Hilton and Von Hippel, 1996), but they can also be exacerbated by the un informativeness of equilibrium behavior. Two relevant quantities over which beliefs may differ are:

$$Pr(\text{punished} \mid \text{black, behaves badly}) - Pr(\text{punished} \mid \text{white, behaves badly}) \quad (1)$$

and

$$Pr(\text{punished} \mid \text{black, behaves badly}) - Pr(\text{punished} \mid \text{black, behaves well}) \quad (2)$$

Whites and/or teachers may believe (1) is zero and (2) is positive: punishment comes from bad behavior. Black pupils may believe (1) is positive and (2) is zero: one is punished because of one's minority status. To estimate these accurately, many episodes of black and white pupils misbehaving, and of black pupils behaving well, must be observed. Now suppose, to take an extreme case, that black pupils believe they will be punished irrespective of their behavior; as a result they always behave badly. White pupils believe they will only be punished for bad behavior, and therefore always behave well. It will then be impossible to estimate either of the above quantities, since one term in each of them is never observed. Both parties have legitimately differing beliefs about a counterfactual.⁷ Of course, more realistically there will be varying degrees of misbehavior and punishment, but it is still true that equilibrium behavior can make it hard for involved actors to reach common beliefs.

Similar patterns arise in other areas. For instance, different ethnic group members often have different perceptions of police "stop and search" policies. As this example suggests, the pattern is not limited to community punishment, but may occur when punishment is delegated to state agents. However, informal community enforcement is particularly likely to suffer from these problems: norms are often unwritten and ambiguous; there is no due process to provide assurances of equal treatment; and the coordination problems of decentralized enforcement create especially strong incentives to target salient minorities.

3 Experiment

The above discussion points to an alternative explanation for the well-known link between heterogeneity and failure to provide public goods. This explanation does not invoke different preferences or norms, in-group bias, or failure to cooperate because of cultural differences. Instead, the bare presence of visible minorities provides an alternative target for coordinated expropriation; this creates an ambiguity in the targeting rule; and punishment of out-groups may then lose its

⁷We are appealing here to an idea of self-confirming equilibrium (Fudenberg and Levine, 1993).

legitimacy and deterrent power, and potentially cause conflict. Our historical examples showed these problems. However, it is hard to make clear-cut inferences about causality from observations in the field. So, we also ran a laboratory experiment, using artificial, minimal groups, without cultural differences, to reproduce our mechanism in a controlled environment.

3.1 Experimental design

Subjects were randomly allocated into groups of four, and stayed in the same group throughout the experiment. They played twenty rounds under one of three institutions.

In Treatment P, each round was a standard public goods game. Each player was endowed with 50 Experimental Currency Units (ECU) and could contribute them to a common fund. Total contributions were multiplied by 1.4 before being divided equally among the group. Feedback on individual contributions was given after each round.

In each round of treatment PX, feedback from the public goods game was followed by an expropriation game. In this game, subjects were endowed with 50 ECU each. Each subject could *target* either a single other group member, or nobody. Targeting was costless. If three subjects targeted the same person, the target lost 50 ECU and the three each gained 15 ECU. After the game, subjects received feedback on how many people targeted each group member.

In treatment X, each round consisted of only the expropriation game, without the public goods game. This allows us to examine how coordinated expropriation is affected by the preceding public goods game.

Player i 's payoff for a given round can be written:

$$\pi_i = I_P(50 - p_i + 1.4 \sum_{j=1}^4 p_j) + I_X(50 - 50x_i + \sum_{j \neq i} 15x_j) \quad (3)$$

where I_P and I_X are indicator variables for the presence of the public goods game and the expropriation game; p_j is player j 's contribution to the public goods game; and x_j is an indicator variable taking the value 1 if player j was targeted by the three other players, otherwise 0. In treatment P, the unique equilibrium has zero contributions. In treatments X and PX, there are many equilibria, some of which support positive contributions in PX. For example, it is an equilibrium for all group members to contribute 50 (or any other number) and for deviators to be expropriated.

The three “institution” treatments varied between sessions. They were crossed with within-session treatments which manipulated group heterogeneity. Every subject was given a color label, *purple* or *green*, which stayed the same throughout the experiment and was visible to other group members. Instructions made clear that the label assignment was random. In homogenous treatments, all subjects within a group had the same color. In heterogenous treatments, one subject within each group had a different label from the other three subjects. We call such subjects *minorities* and the other subjects the *majority*. Thus, these treatments introduce a minimal form of heterogeneity: a public, randomly assigned label.

Table 1 summarizes the treatments. After the experiment finished, subjects answered a short questionnaire and were paid privately for one randomly chosen round, at a rate of 1 ECU = 10 UK pence, plus a £2.50 showup fee.

	Homogenous	Heterogeneous
Expropriation game only (X)	X-Hom	X-Het
Public goods + expropriation game (PX)	PX-Hom	PX-Het
Public goods game only (P)	P-Hom	P-Het

Table 1: Treatments

Our first two hypotheses are:

1. Groups which target expropriation on low contributors will achieve higher contribution levels.
2. In PX-Het, groups will expropriate fewer low contributors than in PX-Hom.

Hypothesis 1 simply tests that participants best respond to the threat of punishment. Hypothesis 2 will hold if expropriators in heterogenous groups coordinate on minority group members, rather than on the lowest contributor, or if they fail to coordinate at all. If both hypotheses hold, we expect:

3. Contributions will be higher in PX-Hom than in PX-Het.

We include treatment P to test that our proposed mechanism drives our results. We expect inefficiency to be caused by less effective expropriation. Removing the expropriation stage should remove any difference between homogenous and heterogenous treatments:

4. Heterogeneity will not affect contributions in P treatments.

The last hypothesis is motivated by the idea that minorities may interpret being expropriated differently. If they believe they are likely to be expropriated regardless of their contribution, they will have less incentive to contribute more. The same effect could occur if they view their expropriation as unfair and react resentfully.

5. After experiencing expropriation, minorities will increase their contributions less than majorities.

3.2 Results

Sessions were run at two UK universities from June 2013 to October 2014, using zTree (Fischbacher, 2007). Subjects were invited using the hroot subject pool management software (Bock, Nicklisch and Baetge, 2012). In total 532 subjects took part in 35 sessions. Sessions lasted about one hour. Average earnings were £12.85.

Figure 1 shows mean public goods contributions over time in treatments P and PX. There is a clear difference between P and PX treatments, and a smaller difference between homogenous and heterogenous treatments.

We first test hypothesis 1 formally. If low contributors were expropriated, then contributions should increase, as in public goods games with costly punishment (Fehr and Gächter, 2000). Indeed, mean contributions per group are significantly higher in PX-Het than P-Het (Mann-Whitney test, $p < 0.001$) and higher in PX-Hom than P-Hom ($p < 0.05$). Table 2 regresses groups' average contributions on the number of rounds in which a lowest contributor was expropriated.⁸ Column 1 shows results for all PX groups. The coefficient on number of expropriations is highly significant. There is potential endogeneity: if lower (higher) average contributions led to more expropriations, this would probably bias coefficients towards (away from) zero. To mitigate this problem, we regressed average contributions from rounds 11-20 on total expropriations in rounds 1-10, controlling for average contributions in rounds 1-10 (*Early contributions*); results are shown in column 2. The effect of expropriations increases and remains significant. In column 3, we look at PX-Het groups, and add the number of expropriations of the minority member.⁹ The coefficient

⁸Results would be almost unchanged if we used all expropriations.

⁹This category can overlap with expropriations of a lowest contributor.

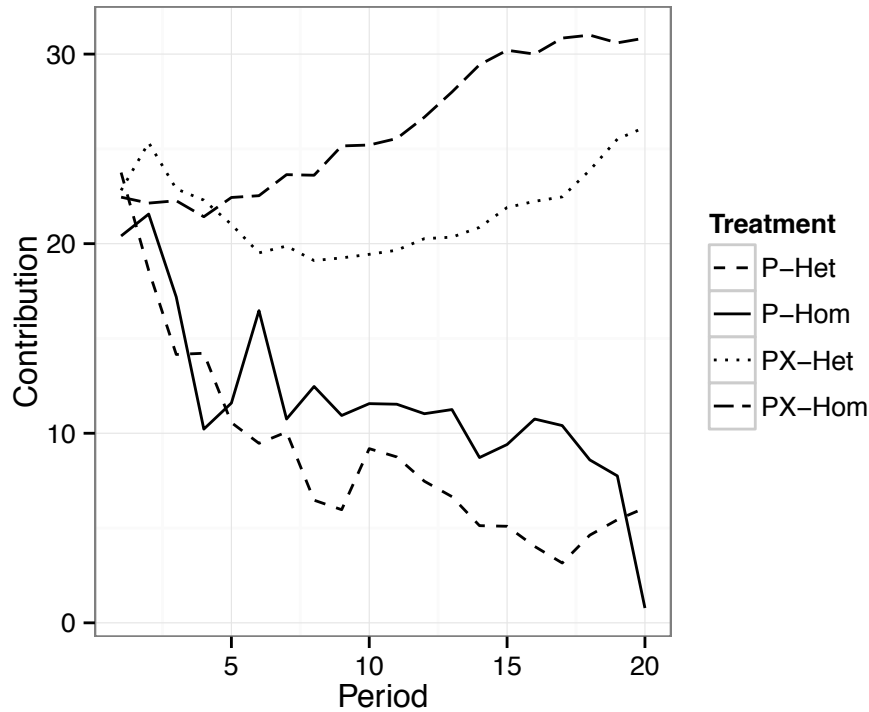


Figure 1: Mean contributions over time by treatment

on this is negative and weakly significant, suggesting that expropriations of minorities decrease, rather than increase, efficiency.

To examine hypothesis 2, Figure 2 shows the proportion of subjects expropriated per round, grouped by both contribution and minority status. Expropriation was overwhelmingly targeted on lowest contributors in all treatments.¹⁰ A *unique* lowest contributor was especially likely to be expropriated. At group level, there were significantly more expropriations per group in PX-Hom than in PX-Het (9.62 vs. 6.56, $p < 0.05$),¹¹ and also more expropriations of a lowest contributor (Mann-Whitney $p < 0.10$), confirming hypothesis 2. This difference has two causes. First, there are significantly more expropriation *attempts*, i.e. cases where a subject targeted another subject, in PX-Hom than in PX-Het (Mann-Whitney $p < 0.05$). Second, PX-Hom had more successful expropriations per expropriation attempt (Mann-Whitney $p < 0.05$). Thus, in the

¹⁰About 32% of subjects were lowest contributors in a given round (16% unique lowest, 16% non-unique lowest). If all group members contributed the same amount, we do not classify them as “lowest contributors”.

¹¹There were also significantly more expropriations in PX-Hom than in X-Hom (9.62 vs. 2.67, Mann-Whitney $p < 0.01$). No other differences between treatment pairs are significant.

	All PX	All PX, rounds 11-20	PX-Het
Intercept	10.41** (2.94)	-3.58 (3.37)	15.36** (3.66)
Exprop. lowest	1.74*** (0.33)	3.65*** (0.49)	1.75** (0.48)
Early contributions		0.76*** (0.16)	
Exprop. minority			-2.91+ (1.45)
R ²	0.48	0.82	0.52
Adj. R ²	0.46	0.81	0.44
Num. obs.	32	32	16

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Table 2: OLS regressions of average contributions per group

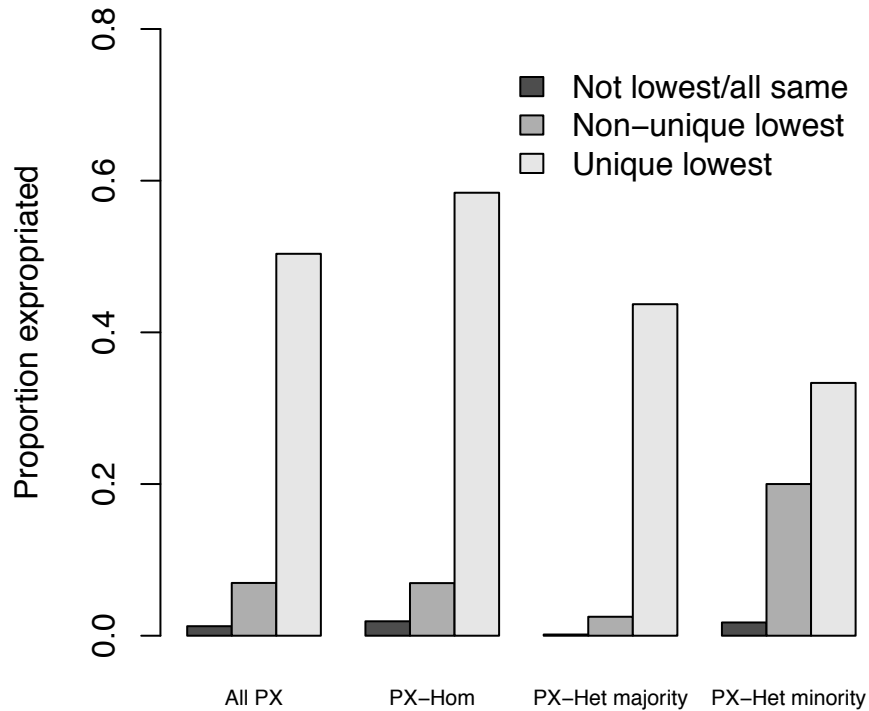


Figure 2: Proportion expropriated by minority status and contribution

PX treatment, homogenous groups seem to have been both more willing to expropriate, and more likely to coordinate on a particular target. Within PX-Het, minorities were significantly more likely to be targeted than majorities, but only when they were non-unique lowest contributors.¹² Thus, heterogeneous groups appear to have only targeted minorities so as to coordinate on one of several equal lowest contributors.

We next test for differences in contributions to the public good. Mean contributions are higher in homogenous than heterogeneous groups, but these differences are not significant (PX treatments: $p = 0.27$, P treatments: $p = 1$). Thus, hypothesis 4 is supported, but hypothesis 3 is not.

Finally, Table 3 shows how individual contribution decisions responded to being expropriated. In column 1, contribution decisions are regressed on lagged contribution decisions, and lagged expropriations interacted with minority status. In column 2 the regression is repeated using the (lagged) number of other group members targeting the subject (*Targeters*). The idea is that being targeted by 1 or 2 others increases the perceived risk of future expropriation. Subjects react to being expropriated, or to being picked by 2 other group members, by increasing their contributions. There is no evidence for Hypothesis 5, that minorities react differently than majorities to being expropriated or targeted: all interaction terms are insignificant, although most have the expected negative sign.

3.3 Additional treatments

Our initial sessions supported only parts of our hypothesized mechanism. Experiment participants did coordinate on minority group members, but only when they were one of two or more equal lowest contributors. Minority participants did not react differently to punishment than majorities, perhaps because they did not face different risks of punishment unless they were equal lowest contributors. We ran three additional treatments to test the robustness of these results.

All treatments were variations of the PX institution. *History treatments* consisted of ten rounds of the expropriation game only, followed by ten rounds of the public goods game plus expropriation game. We expected that majorities might coordinate on expropriating minorities in the first ten rounds, as they had in the original X treatments, and that this history might affect coordination strategies, and minorities' expectations, in the last ten rounds.

¹²See the logit regression in the Appendix.

	PX-Het	PX-Het
Intercept	2.95*** (0.83)	2.27** (0.79)
Lag contrib.	0.84*** (0.04)	0.85*** (0.04)
Lag expropriated	8.00*** (1.58)	
Minority	0.34 (0.26)	0.27 (0.92)
Minority * lag exprop.	-1.48 (2.50)	
Targeters=1		0.48 (1.13)
Targeters=2		2.80* (1.13)
Targeters=3		8.51*** (1.66)
Minority * Targeters=1		0.82 (1.48)
Minority * Targeters=2		-2.68 (2.07)
Minority * Targeters=3		-1.37 (2.64)
Num. obs.	1216	1216
R ²	0.67	0.67
Adj. R ²	0.66	0.67
L.R.	1332.61	1343.71

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Robust s.e.s clustered by group

Table 3: OLS regressions of individual contributions

In *noise treatments*, we added an independent, uniformly distributed shock (-5 to +5) to each player's report of other players' contributions. Thus, different players would see different contributions and might find it harder to coordinate on expropriating free-riders. At the same time, the noise made it harder to interpret the cause of any player's expropriation. Our original treatments had made contribution levels common knowledge. We aimed to mimic some of the real-world situations described above, where the level of free-riding, and hence the reasons for expropriation, are uncertain and interpreted differently across groups.

In *stranger treatments*, subjects were rematched in new groups of 4 between rounds. They always stayed within the same Heterogenous/Homogenous treatment and kept the same color label. The rematching approximates the effect of having a larger community where individuals interact within a larger group. This might make it harder to coordinate on relatively complex strategies like targeting the lowest contributor, and might therefore make minorities more salient.

Results from these extra treatments were as follows. First, Figure 3 shows expropriation rates for the heterogenous groups only. As in the main treatments, minorities were significantly more likely to be expropriated than majorities. In history and stranger treatments, this was true for both unique and non-unique lowest contributors.¹³

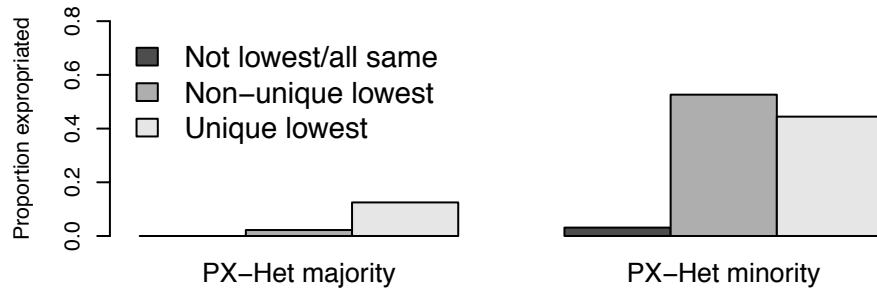
Second, minorities reacted differently to expropriation than majorities (see Table 4). As in the main treatments, majority subjects increased their contribution after being targeted by 2 group members or expropriated by 3 group members. However, minority subjects increased their contributions less, or not at all: the interaction of *Targeters* with minority status is consistently negative.

Lastly, as in the main treatments, these differences in behavior did not lead to lower overall contribution levels in heterogenous treatments (Mann-Whitney tests on group averages: history $p = 0.24$; noise $p = 0.52$; session-treatment averages in stranger treatment, $p = 0.93$).

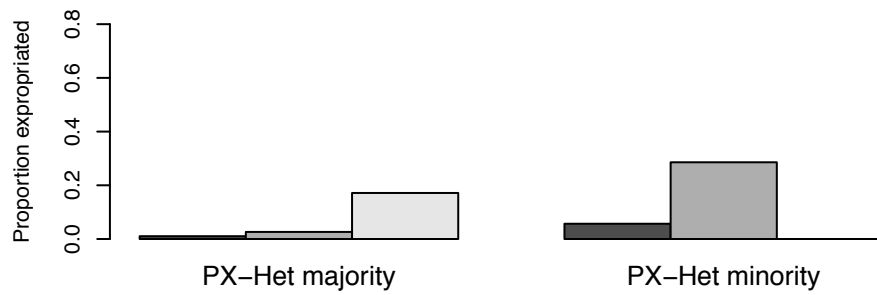
3.4 Discussion

Our results show our proposed mechanisms at work. In our main treatments, minorities were targeted more for expropriation: groups appeared to use the minority label as a coordination

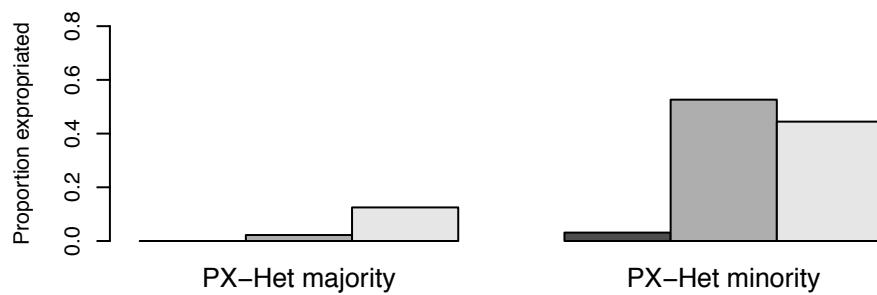
¹³See regressions in the Appendix. In the noise treatment, minority unique lowest contributors were not expropriated. This may have happened because the noise made it hard to identify the true lowest contributor; only 3 minority subjects were ever unique lowest contributors in this treatment.



History



Noise



Stranger

Figure 3: Proportion expropriated by minority status and contribution, supplementary Het treatments

	History, rounds 12-20	Noise	Stranger
Intercept	2.40* (1.16)	2.05* (0.90)	3.38*** (0.83)
Lag contrib.	0.74*** (0.07)	0.80*** (0.09)	0.74*** (0.05)
Targeters=1	-0.93 (1.07)	0.32 (0.65)	-0.24 (0.54)
Targeters=2	3.88* (1.59)	1.74* (0.81)	1.40 (1.08)
Targeters=3	12.26*** (3.03)	10.70** (3.74)	5.41** (1.77)
Minority	0.87 (0.96)	1.47 (1.05)	-0.27 (0.64)
Minority * Targeters=1	-0.01 (1.87)	-0.43 (1.58)	-1.09 (0.74)
Minority * Targeters=2	-5.61** (1.88)	-0.65 (2.23)	-1.57 (1.63)
Minority * Targeters=3	-10.54* (5.08)	-14.31*** (3.45)	-2.37 (1.43)
Num. obs.	396	380	1368
R ²	0.57	0.65	0.53
Adj. R ²	0.56	0.64	0.53
L.R.	331.62	399.29	1032.50

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Robust s.e.s clustered by group/matching group.

Table 4: OLS regressions: individual contributions, supplementary treatments

device when there were two or more lowest contributors. However, minority individuals did not react differently to being targeted. In our additional treatments, minorities were targeted more than majorities even when they were unique lowest contributors. They also reacted differently: unlike majority group members, they did not increase their contributions after being targeted or expropriated.

These mechanisms did not lead to significant unconditional differences in contributions between homogenous and heterogenous treatments, so it could be argued that the mechanisms above are not having a strong impact. However, the high variance in contributions within treatments may also make it hard to detect differences. Because expropriation must be coordinated, there are multiple equilibria with different levels of contributions. Among heterogenous groups, those which expropriated minorities more often saw lower contributions, suggesting that heterogeneity may, but need not always, lead to lower efficiency. Overall, we were able to reproduce our mechanisms of interest even using a rather minimal group manipulation – different colour labels. This supports our claim that under coordinated punishment institutions, small, seemingly irrelevant differences can become salient.

4 Conclusion

Punishing free-riders need not always be costly. It can be too profitable. Existing theory conceptualizes social order as preventing individual free-riding, and downplays the simultaneous problem of controlling expropriation by majorities. The two problems are related in complex ways, and this relationship suggests a new explanation for the inefficiency of heterogenous communities. Our cases show that different perceptions of episodes of “punishment” or “expropriation” can spark conflict and free-riding. Our laboratory experiment imposed a minimal form of heterogeneity: different, randomly assigned, public color labels. Nevertheless, this was enough for majority groups to use as a coordination device for expropriation. In some treatments, minorities reacted to expropriation differently from majority group members, refusing to increase their contributions. At the same time, there were fewer expropriations overall, and fewer expropriations of a lowest contributor, in heterogenous groups, suggesting that heterogeneity can impair people’s willingness and ability to punish free-riding.

In one sense, our argument’s policy implications are pessimistic. “Nation-building” is often proposed as a solution to ethnic tensions (Miguel, 2004). However, if even minimal differences

between groups can cause problems, nation-building must wholly erase these differences to succeed. This process is likely either to be slow, or, as in many cases of attempted nation-building, to be experienced as highly oppressive by the groups involved. Similar remarks apply to “integration” policies for ethnic minorities in developed societies.

However, an alternative approach is available. If ethnic conflict and inefficiency are linked to group differences in beliefs about punishment, a possible solution is to create institutions that reduce ambiguity and legitimize the punishment of free-riders. In our experiment, minorities only reacted differently to punishment when a history of previous expropriation, or an ambiguous situation, made it hard to know the reason for being punished. In the main treatment where contributions were common knowledge, they contributed more after being expropriated just as majority group members did. In the real world, liberal institutions such as trials with due process may reassure all groups that they are being treated fairly. Even when the state is not involved, intergroup forums may defuse conflict by squashing rumors and misperceptions (Varshney, 2003). Broadly speaking, institutional quality can help mitigate the problems of ethnic heterogeneity.¹⁴

A part of our argument which we could not test explicitly is that the willingness to punish outgroup members may be affected by the risk of sparking counter-punishment and intergroup conflict. In turn the lower ability to punish may lead to less intergroup trust and affect the network of relationships linking different groups. Thus, the potential for ethnic conflict and inefficiency in public good provision may be closely linked. These links could be explored in laboratory and field settings.

¹⁴The cross-country regressions in Easterly (2001a) provide some support for this view.

Appendix

	Main PX	History, rounds 11-20	Noise	Stranger
Intercept	-2.60*** (0.36)	-3.08*** (0.49)	-2.01*** (0.39)	-2.90*** (0.56)
Minority	2.28*** (0.55)	3.89*** (1.01)	2.69 (1.79)	2.35*** (0.60)
Heterog.	-1.07 (0.55)	-0.70 (1.06)	-1.60 (1.26)	-0.12 (0.69)
Unique lowest	2.94*** (0.43)	1.53*** (0.34)	0.86 (0.49)	2.47*** (0.51)
Minority * unique lowest	-2.72*** (0.78)	-2.17 (1.12)	-9.03*** (1.78)	-1.63** (0.56)
Heterog. * unique lowest	0.47 (0.65)	0.31 (1.24)	1.17 (1.31)	0.06 (0.70)
Num. obs.	805	322	326	912
Pseudo R ²	0.35	0.27	0.08	0.24
L.R.	230.83	52.21	16.08	168.25

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Robust s.e.s clustered by group (matching group for stranger treatments). Non-lowest contributors excluded.

Table 5: Logit regressions: logged odds of being expropriated, supplementary treatments

Table 5 shows logit regressions where the dependent variable is being expropriated. Non-lowest contributors are excluded, since they were almost never expropriated. In PX treatments, the coefficient on *Minority* is significant, but the summed coefficient *Minority* + *Minority* * *Unique lowest* is not. Thus, minorities were only targeted more when they were one of two or three equal lowest contributors in the group. In the additional treatments, summed coefficients *Minority* + *Minority* * *Unique lowest* are significantly positive in the history and stranger treatments ($p = 0.08$; and $p = 0.02$ respectively), significantly negative in the noise treatment.

References

- Abbink, Klaus and Gönül Doğan. N.d. “How to Choose Your Victim.” . Forthcoming.
- Akerlof, G. A. and R. E. Kranton. 2010. “Identity economics: how our identities shape our work, wages, and well-being.”
URL: <http://library.wur.nl/WebQuery/clc/1974543>
- Alesina, Alberto and Eliana La Ferrara. 2000. “Participation in heterogeneous communities.” *The quarterly journal of economics* 115(3):847–904.
- Alesina, Alberto, Reza Baqir and William Easterly. 1999. “Public goods and ethnic divisions.” *The Quarterly Journal of Economics* 114(4):1243–1284.
- Besley, Timothy and Torsten Persson. 2009. “The Origins of State Capacity: Property Rights, Taxation, and Politics.” *American Economic Review* 99(4):1218–44.
- Besley, Timothy and Torsten Persson. 2010. “State capacity, conflict, and development.” *Econometrica* 78(1):1–34.
- Bock, Olaf, Andreas Nicklisch and Ingmar Baetge. 2012. “hroot: Hamburg registration and organization online tool.” *H-Lab Working Paper (1)* .
- Boehm, Christopher. 2001. *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Harvard University Press.
- Boyer, P. and S. Nissenbaum. 1974. “Salem possessed: the social origins of witchcraft.”
URL: <http://library.wur.nl/WebQuery/clc/348151>
- Brass, Paul R. 1997. *Theft of an idol: text and context in the representation of collective violence*. Princeton, N.J.: Princeton University Press.
- Brass, Paul R. 2004. “Development of an institutionalised riot system in Meerut City, 1961 to 1982.” *Economic and Political Weekly* pp. 4839–4848.
- Bringa, Tone R. 1993. “Nationality categories, national identification and identity formation in" multinational" Bosnia.” *Anthropology of East Europe Review* 11(1 & 2):80–89.
URL: <https://www.scholarworks.iu.edu/journals/index.php/aeer/article/viewFile/591/693>

- Casari, Marco and Luigi Luini. 2009. "Cooperation under alternative punishment institutions: An experiment." *Journal of Economic Behavior & Organization* 71(2):273–282.
- Caselli, F. and W. J Coleman. 2006. *On the Theory of Ethnic Conflict*. Cambridge, Mass.: National Bureau of Economic Research.
- Dreber, Anna, David G. Rand, Drew Fudenberg and Martin A. Nowak. 2008. "Winners don't punish." *Nature* 452(7185):348–351.
- Easterly, William. 2001a. "Can Institutions Resolve Ethnic Conflict?*" *Economic Development and Cultural Change* 49(4):687–706.
- Easterly, William. 2001b. "The Middle Class Consensus and Economic Development." *Journal of Economic Growth* 6(4):317–335.
URL: <http://www.springerlink.com/content/k22j04j648150615/abstract/>
- Esteban, J. and D. Ray. 2008. "On the salience of ethnic conflict." *American Economic Review* 98(5):2185–2202.
- Evans-Pritchard, Edward Evan and Eva Gillies. 1976. *Witchcraft, oracles and magic among the Azande*. Clarendon Press Oxford.
- Fearon, J. D. and D. D. Laitin. 2003. "Violence and the social construction of ethnic identity." *International Organization* 54(04):845–877.
- Fehr, E. and S. Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *The American Economic Review* 90(4):980–994.
- Ferguson, Ann Arnett. 2001. *Bad Boys: Public Schools in the Making of Black Masculinity*. University of Michigan Press.
- Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental economics* 10(2):171–178.
- Frohlich, Norman and Joe A. Oppenheimer. 1970. "I Get By With a Little Help From My Friends." *World Politics* 23(01):104–120.
- Fudenberg, Drew and David K. Levine. 1993. "Self-confirming equilibrium." *Econometrica: Journal of the Econometric Society* pp. 523–545.

- Grant, Gerald. 1988. *The World We Created at Hamilton High 1953-1987*. Harvard University Press.
- Guala, Francesco. 2012. "Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate." *Behavioral and Brain Sciences* 35(01):1–15.
- Habyarimana, J., M. Humphreys, D. N Posner and J. Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision? An Experimental Approach." *American Political Science Review* 101(4):709–725.
- Habyarimana, James, Macartan Humphreys and Daniel N. Posner. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action (Russell Sage Foundation Series on Trust)*. Russell Sage Foundation Publications.
- Herrmann, Benedikt, Christian Thöni and Simon Gächter. 2008. "Antisocial punishment across societies." *Science* 319(5868):1362–1367.
- Hilton, James L. and William Von Hippel. 1996. "Stereotypes." *Annual review of psychology* 47(1):237–271.
- Horowitz, D. L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Horowitz, D. L. 2001. *The deadly ethnic riot*. University of California Press.
- Horsley, Richard A. 1979. "Who Were the Witches? The Social Roles of the Accused in the European Witch Trials." *The Journal of Interdisciplinary History* 9(4):689–715.
URL: <http://www.jstor.org/stable/203380>
- Howe, Octavius Thorndike. 1923. *Argonauts of '49: History and Adventures of the Emigrant Companies from Massachusetts, 1849-1850*. Harvard University Press.
- Lanchester, Frederick William. 1916. *Aircraft in warfare: The dawn of the fourth arm*. Constable limited.
- Metz, Mary Haywood. 1979. *Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools*. University of California Press.
- Miguel, E. 2004. "Tribe or nation? Nation building and public goods in Kenya versus Tanzania." *World Politics* 56(3):327–362.

- Miguel, E. and M. K Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of Public Economics* 89(11-12):2325–2368.
- Navarro, Ramón Gil. 2000. *The gold rush diary of Ramón Gil Navarro*. Lincoln: University of Nebraska Press.
- Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *American Journal of Sociology* 85(6):1356–1375. Article-Type: research-article / Full publication date: May, 1980 / Copyright © 1980 The University of Chicago Press.
URL: <http://www.jstor.org/stable/2778382>
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.
- Paul, Rodman W. 1980. *Mining frontiers of the Far West, 1848-1880 ...* Albuquerque: University of New Mexico Press.
- Platteau, Jean-Philippe. 1996. Traditional sharing norms as an obstacle to economic growth in tribal societies. Technical report.
- Posen, B. R. 1993. "The security dilemma and ethnic conflict." *Survival* 35(1):27–47.
- Rabin, M. 1993. "Incorporating fairness into game theory and economics." *The American Economic Review* 83(5):1281–1302.
- Stewart, James I. 2009. "Cooperation when N is large: Evidence from the mining camps of the American West." *Journal of Economic Behavior & Organization* 69(3):213–225.
URL: <http://www.sciencedirect.com/science/article/pii/S0167268108001984>
- Thompson, Edward Palmer. 1992. "Rough music reconsidered." *Folklore* 103(1):3–26.
- Umbeck, John. 1977. "The California gold rush: a study of emerging property rights." *Explorations in Economic History* 14(3):197–226.
- Varshney, Ashutosh. 2003. *Ethnic conflict and civic life: Hindus and Muslims in India*. Yale University Press.

Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *The American Political Science Review* 91(2):245–263. ArticleType: research-article / Full publication date: Jun., 1997 / Copyright © 1997 American Political Science Association.

URL: <http://www.jstor.org/stable/2952354>

Weissing, Franz and Elinor Ostrom. 1991. *Irrigation institutions and the games irrigators play: rule enforcement without guards*. Springer.

Yamagishi, Toshio. 1986. "The provision of a sanctioning system as a public good." *Journal of Personality and Social Psychology* 51(1):110–116.