No. 09-0005

Institutionalizing Cooperation:
Public Goods Experiments in the Aftermath of Civil War

by
Marcus Alexander and Fotini Christia

Working Paper Series

Institutionalizing Cooperation:
Public Goods Experiments in the Aftermath of Civil War

by

Marcus Alexander and Fotini Christia

Paper No. 2009-0005

March 2009

About the Authors:

Marcus Alexander, Ph.D. candidate in Government; Graduate Student Associate, Weatherhead Center and Institute for Quantitative Social Science; and Doctoral Fellow, Harvard Kennedy School. Email: malexand@fas.harvard.edu.
Fotini Christia, assistant professor, Department of Political Science, MIT. Starting in the fall of 2009, Harvard Academy Scholar, Weatherhead Center. Email: mch@stanford.edu.

Abstract

Fostering cooperation is one of the main tasks of state building in the aftermath of civil wars, yet little is known about the effects of institutions of integration in increasing inter-ethnic cooperation and facilitating peace. We conducted N-person public goods experiments with costly sanctions in the ethnically-divided city of Mostar in Bosnia-Herzegovina to examine whether and how the introduction of institutions of integration affects cooperation both within and across ethnic groups—in our case Catholic Bosnian Croats and Muslim Bosniacs. Our results indicate that even a limited policy intervention such as the creation of an integrated high school can offset the negative effects of ethnic heterogeneity, driving up peoples' willingness to contribute to public goods. We find that the introduction of institutions of integration is distinct from, and may be necessary for, the effectiveness of sanctions in driving up contributions. The results of this experiment suggest that the presence of integrative institutions can bring about cooperation even when increased heterogeneity diminishes it, thus introducing new ways of thinking about the role of institutions in post-conflict divided societies.

Acknowledgments

1. Introduction

Ethnic diversity has been associated with a wide range of coordination problems ranging from poor economic performance to violent conflict (Wade, 1994; Easterly and Levine, 1997; Alesina, Baqir and Easterly, 1999; Miguel, 2004; Miguel and Gugerty, 2005; Alesina and La Ferrara, 2000, 2005; Van Evera, 2001; Petersen, 2002). Recent evidence, however, suggests that different types of institutions may be critical in preventing civil war (Fearon and Laitin, 2003), fostering minority identification with the state (Elkins and Sides, 2007), increasing the level and access to public goods (Banerjee, Iyer and Somanathan, 2005; Tsai 2007a, 2007b) and promoting long-run development (Banerjee and Iyer, 2004; Acemoglu, Johnson and Robinson, 2005a, 2005b). Still, the discipline is only starting to address the open question of what type of institutions can improve ethnic cooperation in the aftermath of civil war — a moment in time when policy-makers often have extraordinary freedom in implementing new institutional designs. This study examines whether the introduction of an institution specifically designed to promote integration between two previously warring ethnic groups has the positive effect of increasing peoples' willingness to cooperate, as opposed to no effect, or even a negative effect as bringing groups together could reignite old animosities.

The primary hypothesis of whether institutions of integration increase cooperation was tested by conducting controlled public goods games in the city of Mostar in Bosnia-Herzegovina, a post-conflict setting that enabled the study of how real-life institutions interact with ethnic diversity to affect contribution to public goods. More specifically, the partial integration of Mostar's secondary school system, undertaken in a way that offers an exogenous variation in institutions of integration and segregation that resembles a natural experiment (the international community integrated one of the Bosniac and one of the Croat secondary schools into one (serving as treatment), leaving two other schools, one from each respective group, as segregated (serving as controls)), allowed us to bring N-person public goods games into a field setting and examine directly the effect of institutions on the participants' willingness to support public goods.

The major experimental finding is that institutions of integration can mitigate the negative effect that ethnic diversity has on agents' willingness to contribute to public goods, making sanctions effective and driving up contributions. Our results indicate (1) that institutions of integration drive up public goods contributions and (2) that the way through which institutions affect contributions is by reinforcing the use of costly sanctions. Though there are several methodological limitations in conducting as well as extrapolating from experiments to real-world institutional settings (as further discussed below), the findings

presented provide new evidence to a growing literature on ethnicity, conflict, and cooperation that argues for a more central role for institutions (Habyarymana et al. 2008, Whitt and Wilson 2007).

Our approach is innovative for several reasons. While the study of institutions currently occupies a central role in the political economy of developing countries, identifying the causal effect of institutions in contribution to public goods and beyond has emerged as a major challenge. Our design overcomes that challenge by going beyond observational studies (Acemoglu, Johnson and Robinson 2005a, 2005b; Banerjee, Iyer and Somanathan, 2005; Tsai 2007a, 2007b). In fact, it uses a combination of controlled experiments and a unique policy intervention in a post-civil war setting to make inferences on the causal effect of institutions. Second, our research complements the growing body of political science experimental literature that employs dictator and ultimatum games (Fowler 2006; Habyarimana, Humphreys, Posner, and Weinstein, 2007; Whitt and Wilson 2007); by conducting N-person public goods games with costly punishment, we offer a different way of discerning inter-ethnic and intra-ethnic cooperation strategies that go beyond individual self-interest and also account for altruism (Fehr and Schmidt, 1999). Third, we supplement our experimental data with participant surveys on a range of socio-economic as well as attitudinal indicators that allow us to evaluate the correspondence of peoples' self-perceived degree of ethnic and religious identification with their willingness to cooperate. Finally, our study bridges the traditionally macro-level question of institutions with the new body of micro-level behavioral research and policy prescriptions (e.g. Habyarimana et al. 2007, 2008; Whitt and Wilson 2007).

The paper proceeds as follows: section 2 reviews the literature and situates the study's hypotheses in existing findings; section 3 describes our methodology, discussing, among others, the selection of our field site and the set-up of the public goods experiments; section 4 presents the experimental results; and section 5 notes the implications, as well as the limitations, of our findings.

2. Ethnic Diversity and Public Goods Contribution

When addressing the role of institutions in fostering ethnic cooperation, a rich body of political science literature anchors itself, either directly or indirectly, to consociationalism — largely theorizing upon a context of a horizontally-segmented state authority with a number of formal and informal institutions set up along social divisions to ensure equal representation and access to a few major groups (Lijphart 1968, 1977). The empirical contributions to this debate tend to be contested interpretations of very few country cases,

where the role of institutions in promoting peace is often conflated with several observable and unobservable differences between countries (Lijphart 1985; Horowitz 1985; Reynolds 2002). Indeed, a series of studies, reviewed by Elkins and Side (2007), have provided mixed evidence on the relationship between formal state institutions (proportional representation and federalism) and ethnic cooperation or conflict — be it levels of violence, rebelion, or the status of minorities (Cohen 1997; Lijphart 1999; Saideman et al. 2002; Hechter 2000; Lustick, Miodownik and Eidelson 2004; Amoretti and Bermeo 2004). In the most extensive study to date of minority attitudes towards state authority across 51 multiethnic states, Elkins and Side (2007) find that the two institutions of proportional representation and federalism designed to distribute state power fail to promote a sense of attachment to the state for either the majority or the minority groups. Combined with other systematic cross-country evidence (Norris 2004), these findings present a new challenge in investigating what kind of institutions can promote ethnic cooperation.

Even if cross-country evidence were to yield consistent results on the association between institutions and ethnic cooperation, additional methodological obstacles remain. The establishment of new institutions is a rare event that usually takes place after wars and thus the stakes for randomized assignment of entire institutional designs prohibitively high, making it difficult to draw causal inference about the effect of institutions. Statistical solutions for isolating causal effect such as matching and instrumental variables (IV) regression have been difficult to implement empirically due to the small number of cases available and due to the challenges of finding valid instruments or other strategies to obtain identification of the unbiased effects particular institutional choices have on political and economic outcomes (Acemoglu 2005; Elikins and Side 2007).

Looking at the role of institutions in public goods provision, Tsai's (2007a, 2007b) study, situated in a context of an authoritarian regime such as that of China, finds that informal and normative institutions of accountability, manifested in solidary groups, result in higher levels of public goods provision. Looking specifically at ethnically divided societies, there is only one observational study that we are aware of, that of Banerjee, Iyer and Somanathan (2005), which examines the effects of institutionalized sources of social division and public goods contribution. Drawing upon public goods data from 1991, the authors find that areas with higher caste-determined social fragmentation, as well as areas with a colonial past and areas with high landownership rates have lower levels of public goods provision. They in turn underline the need to determine the institutional effects on social behavior as well as examine the possible interaction effects of various macro-historic processes. Indeed, existing research on the contribution to public goods in ethnically divided societies has not yet addressed the role of institutions and has primarily focused on establishing the

negative relationship between ethnic diversity and contribution to public goods. Though the body of observational work has confirmed this negative finding, (Wade, 1994; Alesina, Baqir and Easterly, 1999; Miguel and Gugerty, 2005; Alesina and La Ferrara, 2000 among others), researchers have posed conflicting explanations as to why that is the case. Some like Alesina, Baqir and Easterly (1999) and Alesina and La Ferrara (2000) suggest that heterogeneity of preferences among groups is what prompts lower contributions towards public goods. Others, like Wade (1994), argue that differences in returns to public goods will determine the interaction between ethnic diversity and collective action. Yet others, like Miguel and Gugerty (2005), attribute low contributions to public goods to the inability to impose social sanctions across ethnic groups, assuming that sanctioning mechanisms are stronger within co-ethnics rather than across non-co-ethnics.

Departing from the observational approach, recent experimental research has tried to identify the exact causal mechanism linking ethnic diversity and public goods. While this work does not address the role of institutions, it offers notable new insights on public goods provision in ethnically divided societies. Habyarimana, Humphreys, Posner, and Weinstein's research (2007) — the only experimental work to date that specifically examines why ethnically diverse communities witness lower contributions to local public goods — highlights three mechanisms in trying to identify the empirical relationship between ethnic heterogeneity and contribution to public goods. Drawing upon the rich existing literature on ethnicity and collective action, they parse out three distinct processes that point to a clear pattern of preference for intra-ethnic rather than inter-ethnic cooperation. The first mechanism is one of "preferences" explained by innate in group-favoritism versus out-group antipathies (Tajfel, Billig, and Bundy, 1971; van den Berghe 1981; Barkow, Cosmides and Tooby 1992). The second mechanism, coined "technology", pertains to efficiency concerns in terms of communication impediments (Bates 1973; Chandra 2004). The third and last mechanism, called "strategy selection" relies on the notion of reciprocity-triggered in-group cooperation (Platteau 1994; Fearon and Laitin 1996). Using a random sample of subjects from the Kampala slum area of Uganda and two classic games in the literature — the dictator game and the ultimatum game — the authors find no evidence for "preferences;" only weak evidence for "technology;" and strong evidence for "strategy selection."

Another recent experimental study contributes significantly to the research presented here, even though it does not directly discuss institutions or public goods. Whitt and Wilson's (2007) study, conducted in Bosnia-Herzegovina, found that players in dictator games exhibited strong in-group favoritism but failed to exhibit equally strong, either in magnitude or in incidence, out-group bias. Consequently, the authors argue that the failure of ethnic difference to induce significantly more selfish behavior suggests a potential for successful

post-conflict ethnic reconciliation. They nevertheless find that opportunism against non-co-ethnics is an issue in Bosnia-Herzegovina and suggest the need to identify institutions that can assist in overcoming this concern.

Our methodology builds on Habyarymana et al's (2007) and Whitt and Wilson's (2007) line of innovative experimental work by randomly assigning individuals to a range of strategic interactions with co-ethnics and non-co-ethnics as a way to gauge interethnic cooperation. While our findings are consistent with the "strategy selection" mechanism, this projects does not conclude that "preferences" or "technology" are not important, too. We rather focus on a unique natural experimental setup to gauge the effect of institutions in interethnic cooperation: we exploit an exogenous variation in institutions of integration resulting from the international community's move to integrate one of the Bosniac and one of the Croat secondary schools into one (serving as treatment), leaving two other schools, one from each respective group, as segregated (serving as controls). The experimental results we present extend Habyarimana et al.'s (2007) and Whitt and Wilson's (2007) landmark works by utilizing N-person public goods games with and without costly sanction (in integrated and in segregated institutions). Group interactions in these games are markedly different from dictator and ultimatum games. The costly sanction treatment they afford, following Fehr and Schmidt (1999) for example, offers a way to identify the role of both ethnic diversity and that of institutions on contributions, as distinct from the sanctioning mechanism itself. Notably, and despite the differences in scope and approach, our study largely confirms the aforementioned works' findings and policy prescriptions, by offering evidence that institutions of integration actually work, i.e. that they have a significant positive effect on interethnic cooperation in post-conflict divided societies.

The key questions relevant to this project that are raised by the works on ethnicity, public goods and institutions reviewed so far are as follows: (1) Does ethnic diversity lower public goods contribution? (2) Does the availability of sanctions enforce cooperation, even in ethnically heterogeneous groups? (3) Do institutions of integration restore cooperation in the presence of ethnic diversity? The experiments presented here address these questions by testing the following three hypotheses:

(1) *An increase in the ethnic diversity of a group will have a negative, systemic effect on contributions to a public good.*
(2) *The introduction of costly sanctions will promote cooperation, leading to an increase in public goods contributions.*
(3) *The presence of institutions of integration will increase contributions to a public good.*

3. Methodology

Since our research question was guided by the need to inform the contemporary debate on state building and interethnic cooperation in post-conflict divided societies, the focus of the research was on the first generation of individuals in a post-conflict state. After background research and repeated trips to the field, the city of Mostar in Bosnia-Herzegovina was identified as the field site of preference, since the secondary school set-up in that city offered a set of institutional controls rarely found in post-conflict states, therefore limiting possible biases that could affect the validity of an experiment with human subjects in a field setting. This section in turn describes the field site and the set-up of the public goods experiments carried out in the field.[4]

3.1. Site Selection

After the war, the city of Mostar had four general education high schools — called gimnazija— two of which were in the Croat-majority west part of town and the remaining two in the Bosniac-majority east part of town. By a February 2004 Cantonal Judicial Decision, which provided legal affirmation to work that had started in the summer of 2003,[5] one of the Croat secondary schools (Fra. Dominik Mandic Gymnasium) was administratively merged with one of the Bosniac secondary schools (First Gymnasium) in the former's premises. Indeed, the way the international community advised the undertaking of the merger of the two schools was unintentionally exogenous, precluding selection bias. More specifically, because the Croat school's premises could not afford the doubling of its students all at once, in September 2004, roughly 200 Bosniac students, constituting the entire second and third grade of the all-Bosniac (First Gymnasium) high school joined the 300 Croat students already attending the all-Croat (Fra. Dominik Mandic Gymnasium) high school. It was not until the academic year 2005-06, that a freshman and sophomore class from the all-Bosniac school was

---

[4]In the summer of 2005, while in the city of Mostar conducting interviews with local political and military war-time elites on an unrelated academic project, we discovered a case of institutional engineering that appeared to amount to a natural experiment.

[5]Relevant Cantonal Decisions: Odluka o brisanju Gimanazije fra Dominik Mandic iz Registra ustanova kod nadleznog suda, Gradska Opcina Mostar Jugozapad, Vjece Gradske Opstine , Feb 2, 2004; Odluka o administrativnom ujedinjenju/spajanju Gimnazije fra Dominik Mandic Mostar i Prve Gimnazije u Mostaru, Gradska Opcina Mostar Jugozapad, Vjece Gradske Opstine, Aug 28, 2003; Odluka o utvrdjivanju broja zlanova Privremenog skolskog odbora Srednje skole Gimnazija Mostar, Gradska Opcina Mostar Jugozapad, Vjece Gradske Opstine Mostar Jugozapad, Aug 28, 2003; Rjesenje o upisu promijena podataka o subjektu upisa u sudski registar, Vrhovni sud Federacije BiH, Sarajevo, Zupanski sud Mostar - Kantonalni sud Mostar, Feb. 27, 2004.

moved into the building. This merger created a Bosniac-Croat integrated school, leaving the city with two segregated high schools-one mono-ethnic Bosniac (Second Gymnasium) and one mono-ethnic Croat high school (Fra. Grge Martica Gymnasium).[6]

One could argue that this was as close to "randomized" as one could get in a real-world setting: There was no substantial possibility for selection bias as the first set of students who joined the school (sophomores and juniors) had no option to register or move to either of the other two mono-ethnic gymnasia. A possibility for selection bias existed with the incoming class of freshmen, who could choose to attend the integrated school or their mono-ethnic school of preference, and that is why they were left out of the sample of students we worked with. As highlighted above, there was no such option or choice for the Bosniac and Croat upperclassmen and the present experiment focused on this population of subjects. More specifically, we recruited individuals who were in the sophomore incoming class in September 2004 and rising seniors at the time of the experiment in late May/June 2006, and with whom there was no possibility of selection bias (i.e. these were students not attending the integrated school because they had a preference, or higher tolerance for interaction with students of other ethnic groups than their counterparts attending the mono-ethnic schools but rather because they had been assigned to that school).

The institutional setup of our experiment thus involves rising seniors from the three schools: the integrated gymnasium and the two remaining mono-ethnic gymnasia. In terms of location, while the integrated high school, which resumed its prewar name (Mostar Gymnasium), is situated on what used to be the center of the old city and the frontline of wartime hostilities, linking the west and east side of town, the two mono-ethnic schools, the all-Croat and all-Bosniac gymnasiums, are respectively situated well into the west and east side of the city. The all-Bosniac gymnasium is on the hill over the old town of Mostar, in the same building as the seventh primary school, working in alternate shifts. The all-Croat gymnasium is on the outer side of West Mostar, a 15 minute walk from the old gymnasium, in the same building that it occupied before the war.

<div align="center">Insert Figure 1 here.</div>

It should be made clear that though the students in the integrated school are housed in the same building, they study in separate classrooms. That is because the Bosnian educational system allows for three curricula taught in Bosnian, Croatian or Serbian. The difference between the integrated and mono-ethnic schools is that the former allows for shared facilities (such as the library, school yard, sports hall, IT lab); a joint student council of 8 Croat and

---

[6]Interview with Matthew Newton, OSCE Education Officer, Regional Office Mostar, February 2006.

8 Bosniac students that meets on a weekly basis to organize joint activities (art workshops, cleaning up the park, community service etc); a joint school board with a total of 3 Croat and 3 Bosniac Members (2 teachers, 2 parents and 2 school trustees); as well as joint school administration (one school director, one secretary, one accountant, one registry book, one school bank account).[7] Though this may not be considered a fully integrated school, as the students study in separate classrooms, it is the closest to an integrated institutional setting in education in Bosnia-Herzegovina. And since it poses a tougher test for integration, positive results would suggest that even limited amounts of integration can have effects on interethnic cooperation.

In order to ensure that students' allocation to the mono-ethnic and integrated schools–which we clearly did not supervise–indeed approximated random assignment, we had to look beyond the documented and declared lack of self-selection in joining the integrated school, and to also show that there were no relevant pre-existing differences between the students attending the two Bosniac high schools (as compared to each other) and those attending the two Croat schools (as compared to each other). By the time we got to the field the integrated school was already operational and we had no access to previous student records. Because the students had already been exposed to different institutional environments, any comparison of student characteristics (e.g. academic performance or inter-ethnic attitudes) would likely suffer from post-treatment bias. Instead, we focused on predetermined characteristics. Table 1 represents the comparison of parental income as the most important socio-economic characteristic (correlated with parental education and occupation) and parental frequency of worship measured by religious service attendance at churches or mosques, as the most relevant attitudinal proxy for the intensity of parents' ethno-religious identification. In addition, we present a comparison of students' aptitude as measured by mathematics grades, selected as the most valid measure available and least likely to change in a short period of time given the structure of the local curriculum.

Insert Table 1 here

The comparison of the relevant covariates[8] reveals no difference in socio-economic status or the intensity of ethnic identification among Croat students from the different schools. The

---

[7]Gymnasium Mostar: Mostars Other Landmark, Reconstruction and Revitalization Efforts Overview, OSCE Mission to BiH, January 2005.; Overview of Two Schools Under One Roof, OSCE Mission to BiH, May 2005. Report on Implementation of the Interim Agreement on Accommodation of Specific Needs and Rights of Returnee Children, Coordination Board for the Implementation of the Interim Agreement on Returnee Children, March 2005.

[8]We tested for the difference across Croat students attending the segregated versus the integrated school using the t-test for the comparison of means and the Kolmogorov-Smirnov nonparametric

difference in aptitude among Croat schools from the different schools is significant at the 10 % level. There was no statistically significant difference between the Bosniac students from the different schools with the notable exception of parental income levels, which were higher for students attending the integrated school. For this reason, in addition to comparison of means, our analysis of the experimental data also includes estimation of the treatment effects conditional on the distribution of income and other related variables both through regression analysis and matching. Finally, our comparison also revealed that as a whole, Croat and Bosniac students differ regardless of institutional variation. However, this was to be expected given the real-world setting of our experiment, where ethnic groups do not enjoy complete equality. Even if baseline, pre-treatment preferences for public goods contribution are a function of parental socio-eocnomic conditions, for example, our experimental design still affords accurate measurement of treatment effects of institutions, ethnic diversity and sanctions. In other words, even if Croats prefer to contribute more because they are on average better off, the random assignment of subjects ensures an accurate test of whether they will contribute more in ethnically homogenous groups, when sanctions are available, and in the presence of institutions of integration.

Given the setup of what approximated random assignment to different institutions — and our ability to statistically evaluate and account for any subject baseline differences, — if the results of the experiment were to prove that the students attending the integrated school are more cooperative than the students of the mono-ethnic schools, it would be plausible to ascertain that it was due to the institutional effect of the integrated school and not due to a pre-existing student predisposition to interact favorably with members of other ethnic groups.

3.2. **Recruitment and Assignment of Subjects**

This study's methodology and research instruments were approved by Harvard University's Committee on the Use of Human Subjects in Research. Since the participants were 17 years old, written consent was obtained from their guardians. The experiment was conducted in late May-early June 2006. Our sample of subjects consisted of 257 randomly selected students from the three participating high schools: the integrated school (Mostar Gimnazija) as well as the Bosniac and Croat segregated schools (Druga Gimnazija and Fra. Grge Martica Gimnazija respectively). Students were chosen using a random number generator from the respective school rosters and more than 78 % of selected students agreed to participate, with the remaining study sample consisting of randomly chosen alternates. There were no

---

test for the equality of distributions (to account for the possibility of non-normal distribution of characteristics). The same analysis was performed for the Bosniac students.

instances of attrition — all of the participants who consented to participating played the game to the end.

3.3. Description of the Experiment

We conducted standard N-person public goods experiments. These experiments have been previously carried out in many settings across the world and have been shown to be reliable in eliciting preferences for cooperation (for a review, see Fehr and Schmidt, 2000). In the public goods experiment, all players decide simultaneously on the contribution they want to make to the public good. The contribution of a player to the public good can be a minimum of zero and a maximum of the player's starting endowment. Everybody starts off with the same endowment. The contributions of all players in each game are summed up, the sum increases by a certain specified percentage and the public good is equally divided among the number of players in the game, regardless of their contribution. While the aggregate welfare of the group is maximized when each agent contributes their entire endowment, the dominant strategy for a purely selfish player is to contribute nothing to the public good.

In the second stage, the public goods game is extended to include costly sanction. In this stage, players receive information on each others' contribution and then decide whether to sanction other players. Sanctions are credible because they are costly to impose; rather than being just cheap talk, a player needs to sacrifice their own utility in order to induce others to cooperate. Game-theoretic analysis of public goods games with and without sanctions yields a number of different equilibria that can be tested empirically. In the absence of sanctions. there is a unique equilibrium with no contributions, i.e. everybody wants to free-ride. In this set up, the main obstacle to sustaining cooperation is self-interest, i.e. the individuals' aversion to being duped. However, this dynamic changes when costly sanctions become available. They are expected to promote higher contributions if there is sufficient preference for cooperation within the group. That is, cooperation can be a sustainable outcome if there is a critical mass of players willing to undertake personal loss in order to use sanctions that induce other more selfish players to cooperate as well.

The game was played through a computer interface and sessions were conducted in groups of four. Upon arrival, the proctors verified the participants' identity by their invitation letter and consent form and gave them a randomly assigned number. Then, in the participants' presence, the proctor destroyed any information that could be used to verify the participants' identity and handed the participant 10 KM (5 Euro) for agreeing to participate. At the beginning of each session, the game was explained in the respective local languages by our proctors, who read from a standard script, in the presence of one

of the authors. Two randomly selected students had to repeat the directions of the game back to the proctor and participants were encouraged to ask clarification questions during the trial runs. In addition to 10 KM for participating, money which was theirs to keep, subjects were granted a starting endowment of 10 KM to play the game. In stage one of the game, the subjects were asked to contribute a portion of their endowment to a common project. In stage two, the collected money was multiplied by 1.3 and distributed equally to the participants. Each player received a message on how much money had been raised and the sum of their final endowment for the period. The run was then repeated on average 20 times, each time with the same starting endowment and the subjects randomly re-assigned to a new group of four. The equal starting endowment is the strategy for accurately assessing propensity to cooperate while eliminating second-order effects that arise from repeated games.[9] There was an average of five trial runs to ensure that all subjects fully understand the game and the strategies involved. Three treatment conditions were randomly assigned in a fully-factorial design as described next.

After conducting the experiment, we administered a short survey instrument in the respective local languages, collecting information on each participant's age, gender, school performance, religiosity, ethnic and economic background among others. At the very end of the experiment and the survey, the subjects were debriefed.

3.4. Treatment Conditions

The experiment consisted of three different treatment conditions, each in turn manipulating: (a) ethnic diversity, (b) institutions of integration, and (c) the availability of sanctions. Figure 2 below provides a graphic summary of the treatment conditions.

Insert figure 2 here.

The first treatment was used to evaluate the effects of ethnic diversity. It consisted of two conditions: a mono-ethnic treatment condition (only Bosniac or only Croat participants in each group) and a multi-ethnic one (half Bosniac and half Croat participants in each group). The focus was on attaining a maximum level of polarization that would most closely mirror

---

[9]The money won or lost was deposited into the subjects' "account", and at the end of the last round of the experiment, a lottery was conducted and the winners received the amount that was in their "account." We thank Ernst Fehr for pointing out the importance of an equal starting endowment in each period and random re-assignment to groups in each period as the best available means of avoiding reputation and other repeated-game effects that could bias the results.

the challenges facing public goods contribution in highly-polarized societies, such as Mostar in Bosnia-Herzegovina.[10]

We determined the participants' ethnicity based on the choice of curriculum at the school they attended — be it Croat or Bosniac. We cross-referenced our coding with the post-experiment survey, suggesting a coding accuracy of 92.2 %. Following standard experimental protocols, while we ensured that the subjects knew how to play the game and were properly consented according to Human Subjects guidelines, our intent was to minimize their awareness that the experiment was testing the role of ethnicity. Hence, we conveyed information about the participants' ethnicity priming other subjects in the group without leading them to second-guess the socially-acceptable way to play the game. That is, our treatment aimed to elicit second-order rather than first-order consciousness of ethnicity.[11] In the context of these high schools, the city of Mostar affords notable simplicity in ethnic categorization: ethnicity is largely uni-dimensional, across the Croat/Bosniac divide. There is one measure of co-ethnicity, with first and last names serving as a primary identifier, with no real objective or subjective variance as people look identical (there is no racial dimension). This way we did not have to be concerned whether there were other types of ex ante salient dimensions of ethnicity (Chandra 2004, Posner 2005). The subjects were thus primed to the diversity treatment by using the first initial and an ethnically-identifiable last name for each player. The names were selected randomly, while ensuring that they did not correspond to the participants' real names. The names were pre-tested with local native speakers who gauged and confirmed their ethnic identifiability prior to the experiment. The first initial was used in order to avoid gender-bias, therefore minimizing the possible intervening effect of information and perceptions about gender. Because the game was conducted on the computer, the players' identity was kept anonymous. To minimize the effect of reputation and endowment in repeated interactions, the players were randomly re-assigned to groups with the same treatment assignment after each game, with the names getting re-assigned.[12]

The second treatment, consisting of two conditions, was used to evaluate the effects of the availability of an enforcement mechanism. The first no-enforcement treatment involved

---

[10]Future research should address how the results of this experiment would change if the ethnic makeup of the groups was manipulated to obtain majority-versus-minority ethnic group dynamics. Due to statistical power and degrees-of-freedom concerns related to the exogenously constrained number of subjects involved in the game, an application of such majority-versus-minority treatments was not feasible in the present experiment.

[11]We thank David Laitin for offering this insight while we were drafting the research design for this experiment.

[12]We thank Ernst Fehr for suggesting this design.

participants only deciding on how much they want to contribute and the game ending with everyone receiving their final endowment as adjusted by their share of group contributions. The second enforcement treatment consisted of the identical set-up with the addition of a final stage, where each player was also given information about everyone's contribution, and then given an option to use x amount to sanction any of the other players. The sanctions consisted of the punished player losing $2x$ amount from his endowment at the end of the game. For example, after the group made contributions, and the sum was divided, player A may observe that player B did not contribute any money to the public good, and player A may as a result spend 2 KM from her own endowment in order to sanction player B, who would lose 4 KM from his endowment. In setting up the sanction treatment, the experiment drew upon the broader literature on behavioral economics that implements games designed with the task of identifying the role of sanctions in one-shot public goods game settings, while at the same time using real money and costly punishment to elicit preferences for public goods as distinct from simple monetary self-interest (Fehr and Schmidt 2000).

The third treatment, consisting of two conditions, was used to evaluate the effects of institutions of integration. This treatment exploited the policy intervention of arbitrary assignment of students into the new integrated high school. Hence, the integrated institutions treatment condition was applied by carrying out the experiment with students in the Mostar Gymnasium. In turn, the segregated institutions treatment condition was applied by carrying out the experiment with students from the schools containing virtually all Bosniac or all Croat students, those being the Second Gymnasium and the Fra. Grge Martica Gymnasium respectively.

The resulting approach yielded a total of 12 experimental treatment groups in which the public goods experiments were conducted, allowing for the measurement and recording of all of the players' contributions. Individual subjects' contributions constitute the main outcome variable. We should note that each of the 12 treatment in this experiment is effectively an "information condition" in respect to ethnicity, where players receive ethnic information about whom they are playing. This is in contrast to experiments conducted with no-information conditions, i.e. where players receive no information about other players. Apart from providing information to determine the ethnicity of the other players, the game was anonymous. Because a parallel line of research in political science is concerned with the nature of ethnic identifiability (Habyarimana et al. 2005, 2007), it is also important to consider the advantages and limitations of both priming subjects to ethnicity and ensuring subject anonymity in the experiment.

While the last names unambiguously revealed information about the students' ethnic group affiliation — be it Croat or Bosniac, – they did not allow subjects to recognize each

other as they were chosen randomly and did not coincide with the participating students' names. Similarly, the use of pictures was purposefully avoided as there was high likelihood that students from the same school would know each other (i.e. be classmates or friends) thus introducing serious bias in the results. Moreover, in the context of Bosnia-Herzegovina, pictures would not have been a good tool for ethnic identifiability, as there are no observable physical differences between Bosniacs and Croats. On the contrary, they could have undermined the validity of the results by introducing biases generated by the role of gender (Eckel and Grossman 1998; Holm 2000); the shrinking of social distance (Bohnet and Frey 1995; Charness and Gneezy 2000); or other physical traits such as attractiveness (Bertrand et al. 2005).

4. Results

Group heterogeneity is found to have a significant and substantively large negative effect on levels of contribution (t-test, $p< 0.001$, two-sided, N=256).[13] Subjects in ethnically homogenous groups make on average 62% higher contributions to public goods. The magnitude of this difference suggests strong effects of ethnic diversity; with each subject having played an average of 20 rounds, the effect of diversity is to lower a person's total contribution by 19.2KM. In homogenous groups, per-period mean contribution is 2.51KM (s.e.=0.19), as compared to the mixed group mean of = 1.55KM (s.e.= 0.16). This finding confirms the first hypothesis that ethnic heterogeneity has deleterious effects on cooperation, and as such is consistent with the previous results from observational studies and experimental work employing ultimatum and dictator games.

To ensure that our diversity treatment was in fact a way of manipulating the role ethnic identity plays in public goods contributions, we investigated how the participants' survey responses correlated with their contributions.This approach followed that of Whitt and Wilson (2007), who stress the importance of comparing attitudinal survey data with behavioral experimental data. Respondents were asked to describe their identity by using three adjectives; and if any one of the adjectives referred to ethnic identity, the ethnic identity was coded as strong and otherwise as weak.

Our results indicate that those players who consider ethnicity as an important aspect of their identity contributed significantly more in homogenous groups (two-sided t-test, $p< 0.001$) compared to mixed groups. In the course of 20 rounds, players who identified strongly

---

[13]The non-parametic Kolmogorov-Smirnov test for equality of distributions (which unlike the t-test is not sensitive to parametric assumptions of the distribution of the mean) confirms the results (p= 0.003).

with their ethnic group contributed an average of 70.8KM in homogenous groups, compared to an average of 42.4KM in mixed groups. This amounted to a difference of 27.2KM, as compared to an 11.4KM difference when players had a weak ethnic identification. In homogenous groups, those who had strong ethnic identities contributed 17.2KM more than those with weak ethnic identities (p< 0.001). There was no statistically significant difference in contribution in mixed groups based on the strength of ethnic identification (p= 0.57). These results are consistent with previous findings that in-group contributions are higher than out-group contributions among those who feel strongly about their ethnic identity (Whitt and Wilson 2007). More importantly, it confirmed that our public goods experiments were able to elicit strategic behavior driven, in part, by the participants' responses to changes in ethnic diversity.

As far as sanctions are concerned, overall, the presence of costly sanctions results in higher average levels of contribution (t-test, p=0.056, two-sided, N=256).[14]. On average, subjects contributed 27% more when faced with sanctions (mean= 2.37KM, s.e.=0.12), than without sanctions (mean= 1.87, s.e.=0.23). Over an average of 20 rounds the use of sanctions amounts to a 10KM higher contribution per participant. What explains the uncertainty in the effect of sanctions is the difference in institutional context, to which we now turn.

We find that the integrated school participants contributed 59% more than those from segregated schools (t-test, p= 0.003, two-sided, N=256).[15] Over 20 rounds, the difference results in 19.8KM more in the integrated institutional context, the largest difference of the three treatments, with mean per-period contributions in the integrated context of 2.66KM (s.e. 0.23), compared to 1.67KM (s.e. 0.14) in the segregated one.

Two additional sets of analyses were conducted to confirm the robustness of our results. Nearest-neighbour matching was used to estimate the sample average treatment effect (SATE), accounting for heteroskedasticity and correcting for potential bias as suggested by Abadie and Imbens (2002) and Imbens (2003). Table 2 reports unmatched results as well as the SATE after matching. We matched on our three main pre-determined variables discussed above: income, parents' religious worship, and math ability. In addition, we report the results after matching on a more extensive set of controls that included both

---

[14]The Kolmogorov-Smirnov test for equality of distributions confirms the results (p= 0.022). A one-sided t-test rejected the hypothesis that the presence of sanctions lowers contributions (t-test, p=0.028, one-sided, N=256)

[15]The Kolmogorov-Smirnov test for equality of distributions confirms the results (p= 0.016).

parents' education, their occupation, the parents' club membership and the parents' volunteer experience. Our matching analysis supports the unmatched results reported above. To assess the balance of our covariates, we report the standard summary statistics for matched and unmatched data samples in table 3. According to Imai et al. (2008), who identify the "balance test fallacy," this is a more accurate way to ascertain balance than reporting hypothesis tests.[16]

Insert Table 2 here.

Insert Table 3 here.

Multivariate regression analysis (using four different models to improve robustness, including OLS, Tobit, random effects GLS, and panel Tobit) estimates the partial effects of the three different treatments, controlling for the predetermined participant characteristics discussed above (parental income, parental worship frequency, and the participants' math scores). Results in Table 4 support the conclusion that higher public goods contributions are associated with ethnically homogenous group composition, with the presence of institutions of integration, and, relatively less so, with the availability of sanctions.

Insert Table 4 here.

4.1. The Differential Effect of Institutions

Even though institutions of integration on average increase contributions, looking beyond average effects uncovers how they condition the effects of the other two treatments (i.e. ethnic heterogeneity and availability of sanctions). While ethnic heterogeneity in the segregated institutional context decreases the participants' contributions, the negative effect of heterogeneity entirely disappears in the integrated context. Indeed, the analysis of per-period contributions shows that institutions of integration are conducive to overcoming the parochialism observed in the segregated context; in fact, participants in mixed groups contributed 13% more than participants in ethnically homogenous groups (two-sided t-test, p= 0.011), suggesting that the institutional effect encouraging norms of inter-cooperation is strong at work.[17] We repeated the analysis using average participant contributions instead of per-period contributions. This analysis fully confirms the results in the segregated context (p< 0.001). In the integrated context, though the direction and the magnitude of

---

[16]Both the Kolmogorov-Smirnov test (bootstrap version that accounts for ties) and the Wilcoxon rank-sum test indicate no statistically significant difference between our control and treatment groups (p> 0.10 for all covariates.

[17]The Kolmogorov-Smirnov test confirms the distributions are not equal (p< 0.001).

the difference stay the same, there is no statistically significant difference between the two treatment conditions.

In regards to sanctions, our results show that they increase public goods contributions only in the integrated context, raising the average contribution by a notable 70% (two-sided t-test, p=0.002). The findings also hold if we consider per-period contributions. In the case of integrated institutions, the groups with sanctions have an average per-period contribution of 3.32KM, compared to 1.95KM in groups without sanctions (p< 0.001).[18]

Finally, Figure 3 makes a comparison by all possible combination of treatments. The two main findings, are that institutions of integration mitigate the negative influence of ethnic diversity and make sanctions effective. That is to say, in the presence of institutions of integration, costly sanctions have a large substantive effect, driving up contributions in both ethnically homogenous and ethnically heterogenous groups.

<center>Insert Figure 3 here.</center>

To confirm the robustness of these findings, we repeat the same regression analysis as above, partitioning our sample into integrated and segregated treatment conditions. The results indicate that sanctions have a positive and significant effect on contributions in the integrated context only; while group heterogeneity drives contributions down only in the segregated context.

<center>Insert Table 5 here.</center>

It is important to note that our results on the role of institutions relate directly to existing theories on public goods. Game theoretic models such as Fehr and Schmidt's (1999) predict that it is the introduction of credible, costly sanctions that shifts the equilibrium from a noncooperative one (dominated by selfish motives and characterized by free-riding) to a cooperative one (where players' preferences for more equal group outcomes are sufficient for at least one group player to undertake personal monetary cost in order to make costly sanctioning credible). The evidence presented in this section suggests that a reason why institutions of integration drive up contributions is because they can make the threat of punishment more credible and, consequently, render sanctions more effective.

4.2. Institutions and the Effectiveness of Costly Sanctions

Given that the effect of sanctions on contribution differed by institutional context, we next turn to examining whether and how the level of sanctioning varies across our treatment

---

[18]Kolmogorov-Smirnov test confirms the result (p< 0.001).

conditions. Instead of levels of contributions, the focus of the analysis now turns to the level of sanctioning as the main outcome variable. Since our results show that heterogeneity lowers contributions and integration increases them, and since the theory of public goods contributions suggests that costly sanctions are responsible for cooperation, we would expect these relationships to manifest themselves in the level of sanctioning. That is to say, we would expect that higher levels of ethnic diversity would lead to decreased sanctioning and that institutions of integration would lead to increased sanctioning. These hypotheses are consistent with the literature on in-group policing (Fearon and Laitin, 1996) as well as the mechanism of strategy selection identified as significant by Habyarimana et al. (2007).

First, the experimental results show that increasing ethnic heterogeneity significantly diminished the participants'willingness to sanction others in the group (two-sided t-test: $p <$ 0.001).[19] Increasing group heterogeneity reduces the average individual sanction amounts by 46%. In ethnically homogenous groups, over the course of an average of 20 periods, subjects spent an average total of 34.5KM sanctioning others, which was more than three times the starting endowment for each period. In contrast, in mixed groups, the average total sanctioning levels were only 18.54KM.This means that after receiving information about each of the players in the game (including their ethnic identity and their contribution in the first part of the game), individuals playing in homogeneous groups were willing to spend double the amount in order to punish low-level contributors and achieve a more egalitarian outcome in the second stage of the game. In other words, our results indicate that sanctioning works in homogenous groups, but that it may be an ineffective tool for promoting cooperation in ethnically mixed groups.

Second, the introduction of institutions of integration increases the willingness to sanction, but the size of this effect is relatively small and statistically significant only at 10% level (two-sided t-test: $p = 0.102$). More specifically, the introduction of institutions of integration increases the willingness to sanction by an average of 14.5%. Over a period of 20 rounds, players that were exposed to integration sanctioned for a total of 30.6KM, while the sanctioning levels in the presence of institutions of segregation were 26.8KM. Indeed, these results suggest that in comparison to ethnic diversity, the effect of institutions is relatively small; diversity had an effect that was almost twice as large as the effect of institutions. However, further investigation revealed that the reason behind the smaller average effect of institutions is that institutions have a different effect on homogenous and heterogenous groups. Figure 4 illustrates the differential effect of institutions on sanctioning levels. The

---

[19]All the results of the t-tests in this section are confirmed using the non-parametric Kolmogorov-Smirnov test.

introduction of institutions had a large and statistically significant positive effect on sanctioning in mixed groups (two-tailed t-test: p< 0.05). In fact, when considering only the subsample of players in mixed groups, institutions of integration drove up sanctioning levels more than four times. Over an average of 20 periods, the players who were in mixed groups in presence of integration sanctioned for a total of 33.5KM, compared to only 6.3KM in presence of institutions of segregation. As figure 4 indicates, institutions of integration mitigate the negative effect of diversity on sanctioning. Hence, the main result of our previous section — that integration mitigates the effects of diversity on *contributions* — is now strengthened; the analysis of the treatment effects on sanctions shows that institutions drive up sanctioning in mixed groups, expanding the concept of what constitutes an ingroup for effective ingroup policing and therefore ensuring that sanctioning is a credible instrument for promoting public goods contribution.

Insert Figure 4 here.

To further validate that the use of sanctions is strategic rather than retributional, we investigated how the participants sanctioning behavior correlated with the strength of their ethnic identification, as reflected in their survey responses. Among those with strong ethnic identities, sanctioning was actually lower in mixed groups than in homogenous groups, although the difference was not statistically significant (two-tailed t-test, p= 0.127). Over the course of 20 rounds, those with strong ethnic identities sanctioned an average 24.6KM in homogenous groups, compared to 19.2 in mixed groups. This supports the assertion that sanctioning is used strategically to enforce cooperation; if sanctions were used as retribution, we would have expected to see greater sanctioning when players with strong ethnic identities face non co-ethnics. Among those who have weaker ethnic identities, we observed twice as high sanctioning in homogenous groups than in mixed groups (p< 0.001), with an average sanctioning level of 32.4KM in homogenous groups compared to 17.6KM in mixed groups. The finding that the change in sanctioning is greater among players with weak ethnic identities than among those with strong identities further suggests that sanctions serve a strategic purpose and are used to induce cooperation. They are not tools of ethnic parochialism.

Insert Table 6 here.

To confirm the robustness of our findings, we estimated a series of regression models that in addition to the main two treatment effects (diversity and institutions) also include (1) the interaction of institutions and diversity, (2) the level of contribution in the first part of the

period, and (3) the predetermined controls used perviously: income, parents' worship frequency, and math ability level. Table 6 reports the results of OLS (with heteroskedasticity-robust standard errors), Tobit, random effects panel model (with heteroskedasticity-robust standard errors), and the Tobit panel model with random effects. The two panel models, random effects and Tobit random effects, are estimated with and without period effects. In sum, regression results confirm the above analysis: the effect of integrative institutions in mixed groups is to completely compensate for any negative effect that diversity has on sanctioning behavior. The conclusion of the analysis is that institutions of integration induce sanctioning behavior in mixed groups.

Furthermore, we evaluated whether punishment levels indeed serve the strategic purpose of credible sanctions. If punishments are employed as a sanction to induce cooperation, one would expect a statistically significant and positive relationship between punishment and contributions. The regression results, presented in table 6 show that the higher contribution levels are indeed associated with higher punishment levels by individuals in those groups.[20]

5. Conclusion

This study confirms previously reported evidence that, on average, ethnic diversity lowers contributions to public goods while costly sanctions increase such contributions. However, our findings also suggest three key effects resulting from the institutional context: (1) on average, institutions of integration increase contribution to public goods; (2) institutions of integration mediate the role of ethnic diversity, with the diversity of groups being positively correlated with lower contributions only in the case of segregation; and (3) institutions condition the role of costly sanctions, with institutions of integration allowing sanctions to drive up contributions, while institutions of segregation render sanctions ineffective.

In all, the experimental design of our study has a comparative advantage in testing causal relationships over existing observational studies on ethnicity and public goods as it affords random assignment of the various treatments, arguably alleviating internal validity concerns. Moreover, rather than providing attitudinal measures of inter-ethnic relations, which do not involve any costs and therefore not necessarily reliable, this study takes a behavioral approach evaluating actual actions that are in turn associated with non-trivial monetary returns thus adding a level of concreteness to our results. Our approach is also

---

[20]For the purpose of clarity, we limit our discussion of the role of sanctions here and we present further evidence on the dynamics of sanctioning and their role in promoting cooperation in our other work.

significantly different from classic field experiments or randomized policy interventions: Because the unit of analysis is individual behavior and because each of the students in the integrated school was effectively assigned the treatment of being in such an environment, the set up allowed us to conduct the identical laboratory experiment on contribution to public goods in both the segregated and the integrated institutional context. Such a laboratory setting provided a controlled environment that maximized the accuracy of estimated treatment effects, a setup that rarely exists in more realistic, but also more complicated, field experiments or random policy interventions where experimenters have no control over the unobservable intervening variables.

Nevertheless, our experiment is not immune to methodological limitations. By bringing the lab to the field a degree of internal validity was sacrificed to increase external validity. For instance, while all evidence suggests that students were arbitrarily assigned to different institutions (be they integrated or segregated schools), the estimated treatment effect remains less reliable than if it were to be obtained in a controlled laboratory setting. In addition, though there appears to be no self selection on the level of the teaching staff, as the teachers from the merged segregated schools became the teachers of the integrated school with no notable attrition in the process-, and interviews with school principals suggest no notable underlying competence or attitudinal differences between the teachers in the integrated and the segregated schools, resource constraints prevented us from confirming these statements through a survey of the teaching staff. Teachers and teaching attitudes are undoubtedly an integral part of what renders a school a community institution. Therefore, even if there were a degree of self selection or possible variation in the levels of teachers' competence or attitudes, our results on the positive effect of institutions of integration in promoting cooperation would not be undermined. After all, the objective of this project is to evaluate the differences, if any, between integrated and segregated schools on attitudes towards contribution to public goods. That said, our work is certainly limited in that it does not parse out the mechanisms of what makes an integrated institution more successful in changing these attitudes-be it the interaction of students from diverse ethnic groups in the same learning environment; the competence and attitudes of teachers staffing the school; or, most likely, an interaction of the two.

In addition, the international community's involvement in Bosnia and Herzegovina could arguably cast light to the characterization of the institution of integration under study suggesting that the institutional effects observed are somehow related to the students' efforts to please international donors. Interviews with teachers and administrators at all schools revealed that the concern for good relations with the international community did not differ between the segregated and the integrated schools. And the involvement of the

international community in implementing school integration helped increase confidence that the institutional variation was exogenous, preventing potential sources of bias that could have arisen if the process of school integration was endogenous to local government politics. Ultimately, this concern is another example of the trade-off the study faced in trying to simultaneously maximize internal and external validity. While the real-world setting in which the institutions are implemented make it difficult to perfectly isolate their treatment effect, this limitation comes in exchange for real-world relevance.

But even when confident that our treatments have caused the aforementioned effects in the context of Mostar, the extent to which we can generalize to other settings, populations or methods of measurement suggests strong caution. For instance, though the people who participated in our study were undoubtedly a representative sample of the secondary school (gimnazija) population of Mostar, they are not a random sample of the population of the city of Mostar or Bosnia and Herzegovina as a whole. Neither is Bosnia and Herzegovina, a country which received an extraordinary amount of post-conflict humanitarian and development assistance, the typical post-conflict environment. Further experiments in other post-conflict settings with different subjects and different types of integrative institutions would be necessary to test the replicability and applicability of our findings to other subject populations and to a wider range of real-world problems of ethnic cooperation.

In sum, drawing upon the experience of post-conflict Bosnia and Herzegovina, we find that institutions of integration can prove essential in fostering a cooperative society in the aftermath of violence. More specifically, our findings suggest that the presence of integrative institutions can bring about cooperation even when increased heterogeneity diminishes it, thus introducing new ways of thinking about the role of institutions in war-torn divided societies.

TABLE 1. Relevant covariates by institution and ethnicity

	Croat		Bosniac	
	Segregated	Integrated	Segregated	Integrated
Parental income				
Mean	1366	1297	716	1080
SD	939	742	402	673
[min, max]	[499, 5000]	[450, 4500]	[200, 2500]	[499, 3500]
N	63	44	67	52
T-test	$t = 0.671$		$t = 0.014$	
Kolmogorov-Smirnov test	$p = 1.000$		$p < 0.001$	
Parental worship frequency				
Mean	1.677	1.674	1.362	1.111
SD	0.784	0.837	1.014	1.058
[min, max]	[0, 3]	[0, 3]	[0, 3]	[0, 3]
N	62	43	69	54
T-test	$t = 0.985$		$t = 0.186$	
Kolmogorov-Smirnov test	$p = 1.000$		$p = 0.548$	
Math performance				
Mean	2.955	2.534	2.836	3.136
SD	1.010	1.133	0.920	1.099
[min, max]	[1, 5]	[1, 5]	[2, 5]	[2, 5]
N	55	44	61	59
T-test	$t = 0.058$		$t = 0.109$	
Kolmogorov-Smirnov test	$p = 0.075$		$p = 0.579$	

TABLE 2. Matching analysis of the average treatment effect of sanctions, diversity and institutions on public goods contribution.

	unmatched	matched I	matched II	unmatched	matched I	matched II
Average treatment effect of diversity:						
	-0.964**	-1.000**	-1.786**	-0.805**	-0.705**	-1.500**
	(0.262)	(0.264)	(0.316)	(0.090)	(0.077)	(0.083)
Average treatment effect of sanctions:						
	0.505†	0.444	0.637†	0.679**	0.367**	0.684**
	(0.263)	(0.303)	(0.382)	(0.083)	(0.079)	(0.076)
Average treatment effect of institutions of integration:						
	0.984**	1.357**	1.532**	0.862**	1.366**	1.563**
	(0.258)	(0.320)	(0.409)	(0.086)	(0.081)	(0.072)
Matching variables	income parents' worship math	income parents' worship math mother's occupation father's occupation mother's education father's education parents' club memb parents' volunteer		income parents' worship math	income parents' worship math mother's occupation father's occupation mother's education father's education parents' club memb parents' volunteer	
N	256	219	170	4360	3699	2819

Note: Treatment effects estimated individually for each treatment. For matched estimators, bias-adjusted SATE reported, standard errors are robust to heteroskedasticity. For unmached, difference in means reported. The dependent variable is either average contribution by participant (when $N \leq 229$), or a participant's contribution in each period (when ≤ 3892). Income is in 100s of KM. Significance levels: †10%, *5%, **1%.

TABLE 3. Covariate balance.

	All data				Matched data			
	Means treated	Means control	SD control	Mean diff	Means treated	Means control	SD control	Mean diff
Treatment: Diversity								
Income	10.933	10.126	5.437	0.807	10.933	10.495	5.999	0.438
Parents' worship	1.479	1.247	0.958	0.232	1.479	1.384	0.981	0.096
Math	3.045	2.794	0.999	0.251	3.045	2.904	1.040	0.141
Mother's education	3.151	3.021	0.946	0.130	3.151	3.137	0.855	0.014
Father's education	3.288	3.340	0.945	-0.053	3.288	3.288	0.905	0.000
Mother's occupation	2.027	2.206	1.314	-0.179	2.027	2.205	1.322	-0.178
Father's occupation	2.123	2.371	1.372	-0.248	2.123	2.178	1.358	-0.055
Parents' club membership	0.068	0.062	0.242	0.007	0.068	0.082	0.277	-0.014
Parents' volunteering	0.096	0.062	0.242	0.034	0.096	0.082	0.277	0.014
Treatment: Sanctions								
Income	11.061	9.910	5.291	1.151	11.061	10.123	5.308	0.938
Parents' worship	1.361	1.333	1.019	0.028	1.361	1.277	0.992	0.084
Math	2.813	2.986	1.093	-0.173	2.813	2.973	1.092	-0.160
Mother's education	3.145	3.011	0.869	0.133	3.145	3.072	0.823	0.072
Father's education	3.313	3.322	0.800	-0.009	3.313	3.337	0.816	-0.024
Mother's occupation	2.133	2.126	1.301	0.006	2.133	2.169	1.314	-0.036
Father's occupation	2.361	2.172	1.314	0.189	2.361	2.241	1.303	0.120
Parents' club membership	0.048	0.080	0.274	-0.032	0.048	0.060	0.239	-0.012
Parents' volunteering	0.084	0.069	0.255	0.015	0.084	0.060	0.239	0.024
Treatment: Institutions of Integration								
Income	11.564	9.670	6.560	1.894	11.564	10.453	7.177	1.111
Parents' worship	1.208	1.449	0.986	-0.241	1.208	1.375	0.971	-0.167
Math	2.983	2.842	0.975	0.141	2.983	2.944	1.002	0.039
Mother's education	3.264	2.939	0.929	0.325	3.264	3.264	0.628	0.000
Father's education	3.431	3.235	0.939	0.196	3.431	3.403	0.867	0.028
Mother's occupation	2.264	2.031	1.350	0.233	2.264	2.208	1.383	0.056
Father's occupation	2.486	2.102	1.366	0.384	2.486	2.333	1.404	0.153
Parents' club membership	0.056	0.071	0.259	-0.016	0.056	0.056	0.231	0.000
Parents' volunteering	0.083	0.071	0.259	0.012	0.083	0.069	0.256	0.014

Note: K-S test is the Kolmogorov-Smirnov test (bootstrap version that accounts for ties) and W test is the Wilcoxon (Man-Whitney) rank sum test.

TABLE 4. Regression analysis of treatment effects on public goods contributions

	OLS	Tobit	Random Effects	Random Effects	Panel Tobit	Panel Tobit
Treatments						
Mixed	-0.707*	-0.701†	-0.694†	-0.764*	-0.937†	-1.015*
	(0.254)	(0.294)	(0.279)	(0.276)	(0.382)	(0.384)
Sanctions	0.457	0.450	0.470	0.470	0.405	0.412
	(0.276)	(0.279)	(0.265)	(0.265)	(0.363)	(0.365)
Integrated	0.986**	1.000**	0.981**	1.106**	1.286**	1.442**
	(0.275)	(0.291)	(0.292)	(0.289)	(0.377)	(0.380)
Controls						
Income	0.036	0.038	0.036	0.036	0.050	0.049
	(0.022)	(0.021)	(0.020)	(0.020)	(0.027)	(0.027)
Parents' worship	0.268†	0.268	0.268†	0.276†	0.272	0.281
	(0.123)	(0.140)	(0.130)	(0.129)	(0.182)	(0.183)
Math	-0.046	-0.048	-0.050	-0.044	0.026	0.035
	(0.129)	(0.134)	(0.135)	(0.135)	(0.174)	(0.175)
Constant	1.166†	1.144†	1.170†	0.000	0.601	-0.194
	(0.481)	(0.524)	(0.523)	(0.000)	(0.681)	(0.722)
Period effects	—	—	No	Yes	No	Yes
N	219	219	3699	3699	3699	3699

Note: Reported coefficients and standard errors (robust for heteroskedasticity for OLS and R.E.) in parentheses.

The dependent variable for the first two column is average contribution by participant; the latter two columns it is a participant's contribution in each period. Income is in 100s of KM.

Significance levels: †10%, *5%, **1%.

TABLE 5. Regression analysis of treatment effects on public goods contributions in the context of segregated and integrated institutions

| | Segregated Institutions | | | | Integrated Institutions | | | |
	OLS	Tobit	Random Effects	Panel Tobit	OLS	Tobit	Random Effects	Panel Tobit
Treatments								
Mixed	-1.385**	-1.376**	-1.385**	-1.672**	0.434	0.421	0.434	0.391
	(0.248)	(0.242)	(0.239)	(0.298)	(0.447)	(0.577)	(0.597)	(0.777)
Sanctions	-0.365	-0.374	-0.366	-0.619†	1.524*	1.539*	1.524*	1.898*
	(0.246)	(0.245)	(0.243)	(0.303)	(0.480)	(0.506)	(0.491)	(0.682)
Controls								
Income	0.040†	0.041†	0.040	0.055†	0.036	0.040	0.036	0.050
	(0.020)	(0.019)	(0.024)	(0.023)	(0.035)	(0.036)	(0.030)	(0.049)
Parents' worship	0.182	0.179	0.183	0.130	0.322	0.335	0.322	0.417
	(0.108)	(0.125)	(0.120)	(0.154)	(0.219)	(0.252)	(0.250)	(0.340)
Math	-0.001	0.002	-0.007	0.055	-0.186	-0.195	-0.186	-0.139
	(0.105)	(0.131)	(0.124)	(0.161)	(0.195)	(0.222)	(0.226)	(0.299)
Constant	1.842**	1.830**	1.863**	1.615*	1.618	1.586	1.618	1.060
	(0.405)	(0.465)	(0.472)	(0.571)	(0.861)	(0.910)	(0.946)	(1.228)
N	118	118	1679	1679	101	101	2020	2020

Note: Reported coefficients and standard errors (robust for heteroskedasticity for OLS and R.E.) in parentheses. The dependent variable for the first two column is average contribution by participant, while for the latter two columns it is a participant's contribution in each period. Income is in 100s of KM. Significance levels: †10%, *5%, **1%.

TABLE 6. Regression analysis of treatment effects on sanctions by subjects in a public goods game, stratified by type of institutions.

	OLS	Tobit	Random Effects	Random Effects	Panel Tobit	Panel Tobit
Treatments						
Mixed	-1.275**	-1.299**	-1.292**	-1.338**	-3.167**	-3.167**
	(0.322)	(0.296)	(0.299)	(0.236)	(0.465)	(0.465)
Integrated	-0.494	-0.498	-0.462	-0.412	-0.822†	-0.822†
	(0.380)	(0.278)	(0.321)	(0.259)	(0.408)	(0.408)
Mixed and Integrated	1.763**	1.784**	1.925**	1.971**	3.584**	3.584**
	(0.509)	(0.455)	(0.474)	(0.376)	(0.685)	(0.685)
Controls						
Contribution	0.089†	0.090†	0.074*	0.071*	0.086*	0.086*
	(0.042)	(0.037)	(0.026)	(0.025)	(0.030)	(0.030)
Income	-0.005	-0.004	-0.004	-0.004	0.002	0.002
	(0.015)	(0.014)	(0.013)	(0.011)	(0.020)	(0.020)
Parents' worship	-0.051	-0.056	-0.032	-0.029	-0.165	-0.165
	(0.097)	(0.113)	(0.105)	(0.086)	(0.173)	(0.173)
Math	-0.068	-0.063	-0.071	-0.069	-0.075	-0.075
	(0.088)	(0.107)	(0.098)	(0.078)	(0.161)	(0.161)
Constant	1.922**	1.908**	1.855**	0.000	1.689*	1.689*
	(0.422)	(0.372)	(0.394)	(0.000)	(0.563)	(0.563)
Period effects	—	—	No	Yes	No	Yes
N	107	107	1841	1841	1841	1841

Note: Reported coefficients and standard errors (robust for heteroskedasticity for OLS and R.E.) in parentheses. The dependent variable for the first two column is average sanctioning by participant, while for the latter two columns it is a participant's sanction in each period. Income is in 100s of KM.

Significance levels: †10%, *5%, **1%.

1- Gimnazija Fra. Grge Martića
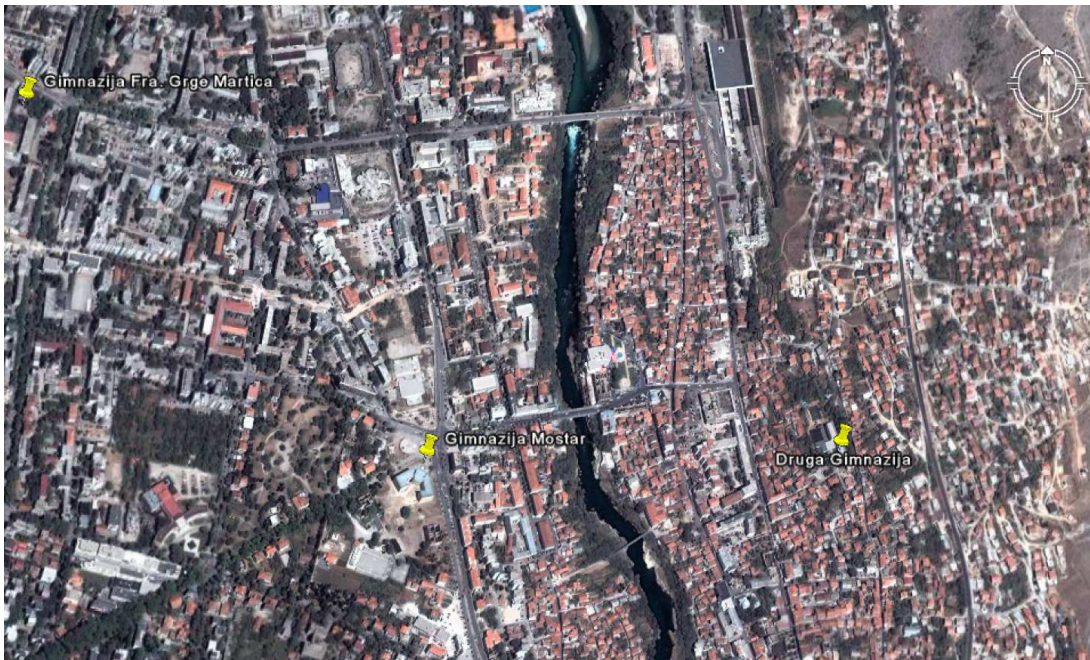2- Gimnazija Mostar
3- Druga Gimnazija

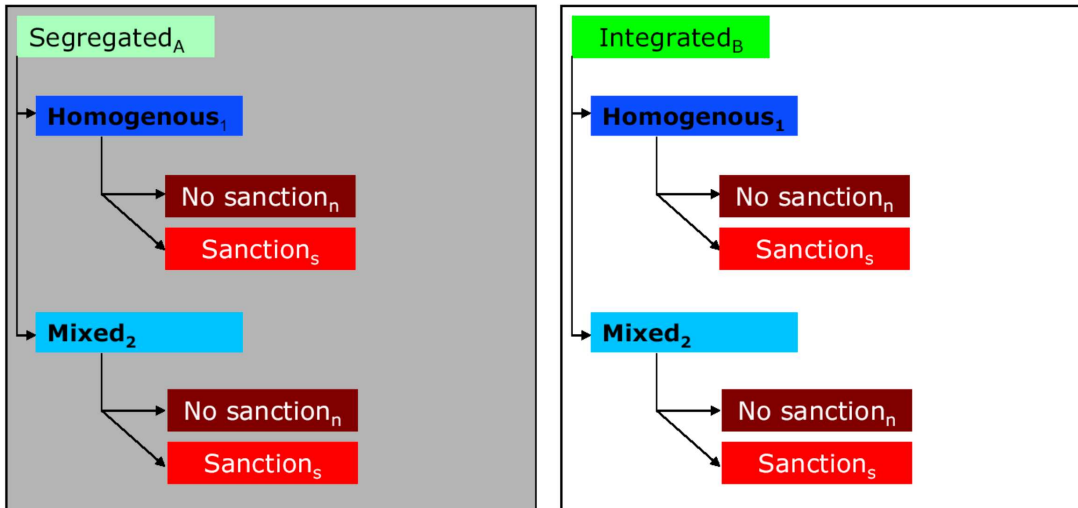FIGURE 1. Location of the three schools involved in the experiment.
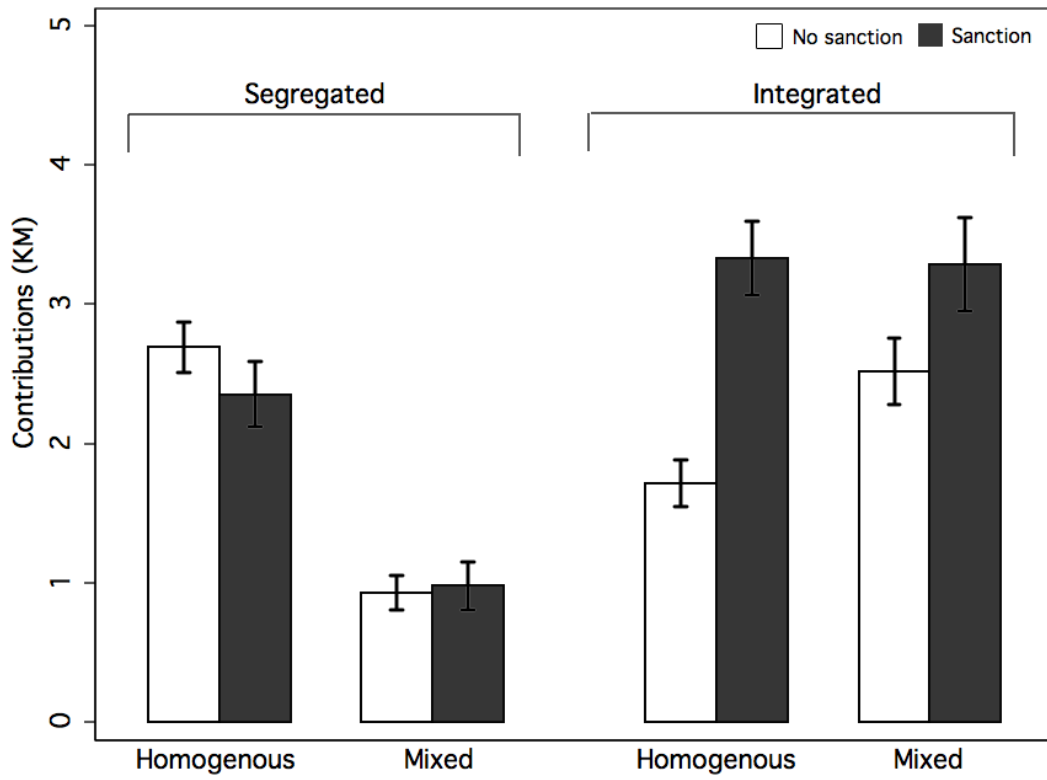
FIGURE 2. Description of treatments.

FIGURE 3. Summary of the mean levels of public good contributions by three experimental treatments (diversity, sanctions, and institutions)
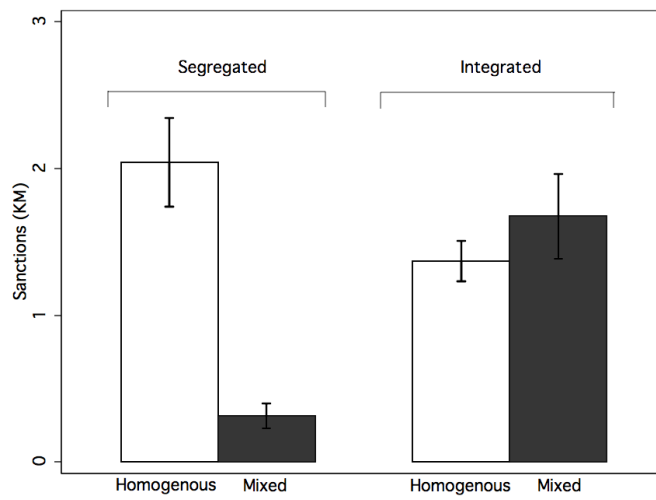
FIGURE 4. Summary of the mean levels of sanctioning levels by two experimental treatments (diversity and institutions)

Abadie, Alberto and Guido Imbens. 2002. "Simple and Bias-Corrected Matching Estimators." Technical report, Department of Economics, UC Berkeley.

Acemoglu, Daron. 2005. "Constitutions, Politics and Economics: A review Essay on Persson and Tabellini's 'The Economic Effect of Constitutions.' *Journal of Economic Literature* 43 (December):1025-48.

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2005. "The Rise of Europe; Atlantic Trade, Institutional Change and Economic Growth." *American Economic Review* 95 (June): 546-79.

Alesina, Alberto and Elianna La Ferrara. 2000. "Participation in Heterogeneous Communities." *Quarterly Journal of Economic* 115(August): 847-904.

Alesina, Alberto and George Marios Angeletos. 2005. "Fairness and Redistribution." *American Economic Review* 95 (September): 960-80.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114 (November): 1243–84.

Alesina, Alberto and Eliana LaFerrara. 2002. "Who Trusts Others?" *Journal of Public Economics* 85 (August): 207–234.

Alesina, Alberto and Ellana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." Journal of Economic Literature 43 (September): 762-800.

Amorettin, Ugo M. and Nancy Bermeo. 2003. *Federalism and Territorial Cleavages*. Baltimore: John Hopkins University Press.

Axelrod, Robert and WD Hamilton. 1981. *"*The Evolution of Cooperation." *Science* 211 (March): 1390-96.

Axtell, Robert, H. M. Epstein, and H. P. Young. 2001. "Emergence of Class Norms in a Multi-Agent Model of Bargaining." In *Social Dynamics*, ed. S. Durlauf and H. P. Young. Cambridge: MIT Press.

Bahry, Donna, and Rick Wilson. Forthcoming. "Confusion or Fairness in the Field? Rejections in the Ultimatum Game under the Strategy Method." *Journal of Economic Behavior and Organization*.

Banerjee, Abhijit and Lakshmi Iyer. 2004. "History, Institutions and Economic Theory and Policy: Essays in Honour of Dipak Banerjee, ed. Dutta Bhaskar. Oxford: Oxford University Press.

Banerjee, Abhijit, Lakshmi Iyer and Rohini Somanathan. 2005 "History, Social Divisions and Public Goods in Rural India." *Journal of the European Economic Association* 3 (April-May): 639–47.

Barkow, Jerome, Leda Cosmides and John Tobby, eds. 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press.

Bates, Robert. 1973. *Ethnicity in Contemporary Africa.* Syracuse: Program in East African Studies.

Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.

Besley, Timothy, Stephen Coate, and Glenn Loury. 1993. "The Economics of Rotating Savings and Credit Associations." *American Economic Review* 83 (September): 792–810.

Bernhard, Helen, Urs Fischbacher and Ernst Fehr. 2006. "Parochial Altruism in Humans". *Nature* 442 (August): 912-15.

Brown, Martin, Armin Fallk and Ernst Fehr. 2004. "Relational Contracts and the Nature of Market Interactions". *Econometrica* 72 (March): 817-69.

Bohnet, Iris, and Bruno Frey. 1999. "Social Distance and Other- Regarding Behavior in Dictator Games: Comment." *American Economic Review* 89 (March): 335–39.

Bolton, Gary E., and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90 (March): 166–93.

Bowles, Samuel, and Herbert Gintis. 2004a. "Persistent Parochialism: Trust and Exclusion in Ethnic Networks." *Journal of Economic Behavior and Organization* 55 (September): 1–23.

Bowles, Samuel, and Herbert Gintis. 2004b. "The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations." *Theoretical Population Biology* 65 (February): 17–28.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.

Cederman, Lars-Erik, and Luc Girardin. 2007. "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies." *American Political Science Review* 101 (February): 173–85.

Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. New York: Cambridge University Press.

Charness, Gary and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests". *Quarterly Journal of Economics* 117 (March): 817-69.

Charness, Gary, and Uri Gneezy. 2000. "What's in a Name? Anonymity and Social Distance in dictator and Ultimatum games." Unpublished manuscript, University of California, Santa Barbara.

Cohen, Thomas. 1997. "Proportional Versus Majoritarian Ethnic Conflict management in Democracies". *Comparative Political Studies* 30 (October): 607-31

Crawford, Vincent P., and Hans Haller. 1990. "Learning How to Cooperate: Optimal Play in Repeated Coordination Games." *Econometrica* 58 (May): 571–95.

Deutsch, Karl. 1966. *Nationalism and Social Communication*. Cambridge: The MIT Press.

de Quervain, Dominique J.F., Urs Fischbacker, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment". *Science* 305 (August): 1254-58.

Elkins, Zachery and John Sides. 2007. "Can Institutions Build Unity in Multiethnic States?" *American Political Science Review* 101 (November): 693-708.

Englemann, Dirk and Martin Strobel. 2004. "Inequality Aversion, Efficiency and Maximin Preferences in Simple Distribution Experiments". *American Economic Review* 94 (September): 857-69.

Fafchamps, Marcel. 2000. "Ethnicity and Credit in African Manufacturing." *Journal of Development Economics* 61 (February): 205–35.

Fearon, James and David Laitin. 2003. "Ethnicity, Insurgency, and Civil War". *American Political Science Review* 97 (February): 75-90.

Fearon, James D. 1999. "Why Ethnic Politics and 'Pork' Tend to Go Together?" Unpublished manuscript, Stanford University.

Fearon, James D., and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90 (December): 715–35.

Fehr, Ernst and Simon Gachter. 2000. Cooperation and Punishment in Public Goods Experiments. *American Economic Review 90: 4(September): 980-994.*

Fehr, Ernst, and Urs Fischbacher. 2004. "Third-Party Punishment and Social Norms." *Evolution and Human Behavior* 25 (March): 63–87.

Fehr, Ernst, and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (August): 817–68.

Fehr, Ernst and Armin Falk. 1999. "Wage rigidity in a competitive incomplete contract market." *Journal of Political Economy* 107 (February): 106-34.

Fehr, Ernst and Urs Fischbacker. 2003. "The nature of human altruism". *Nature* 425 (October): 785-91.

Fehr, Ernst and Urs Fischbacher. 2004. "Social Norms and human cooperation". *Trends in Cognitive Sciences* 8 (April): 185-90.

Fehr, Ernst, Urns Fischbacher and Simon Gatcher. 2002. " Strong reciprocity, human cooperation, and the enforcement of social norms." *Human Nature-An Interdisciplinary Biosocial Perspective* 13 (March): 1-25.

Fehr, Ernst and Bettina Rockenbach. 2003. "Detrimental effects of sanctions on human altruism." *Nature* 422 (March): 137-40.

Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economic*s 116 (February): 351–77.

Fowler, James H. 2006. "Altruism and Turnout." *Journal of Politics* 68 (August): 674–83.

Fowler, James H., Tim Johnson and Oleg Smirnov. 2005. "Human behaviour: Egalitarian motive and altruistic punishment." *Nature* 433(January) :1 p following 32; discussion following 32.

Frank, Robert H. 1985. *Choosing the right pond: human behavior and the quest for the status*. New York: Oxford University Press

Ghosh, Parkishit, and Debraj Ray. 1996. "Cooperation in Community Interaction without Information Flows." *Review of Economic Studies* 63 (July): 491–519.

Gil-White, Francisco. 2004. "Ultimatum Game with an Ethnicity Manipulation: Results from Kohvdiin Bulgan Sum, Mongolia." In *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.

Greif, Avner. 1989. "Reputations and Coalitions in Medieval Trade: Evidence on the Maghribi Traders." *Journal of Economic History* 49 (December): 857–82.

Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society." *Journal of Political Economy* 102 (October): 912–50.

Greig, Fiona, and Iris Bohnet. 2006. "Why Women Cooperate with Women and not Men: Evidence from a Slum in Nairobi, Kenya." Unpublished manuscript, Kennedy School of Government, Harvard University.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and JeremyWeinstein. 2004. "Ethnic Identifiability: An Experimental Approach." Unpublished manuscript, UCLA.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and JeremyWeinstein. 2008. "Is Ethnic Conflict Inevitable? Better Institutions, Not Partition." *Foreign Affairs* (July/August).

Haley, Kevin, and Daniel Fessler. 2005. "Nobody's Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game." *Evolution and Human Behavior* 26 (May): 245–56.

Hardin, Russell. 1995. *One for All: The Logic of Group Conflict*. Princeton: Princeton University Press.

Hechter, Michael. 2001. *Containing Nationalism*. Oxford. Oxford University Press.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. New York: Oxford University Press.

Henrich,Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr and Herbert Gintis, eds. New York: Oxford University Press. Golooba-Mutebi, Frederick. 2003. "Devolution and Outsourcing of Municipal Services in Kampala City, Uganda: An Early Assessment." *Public Administration and Development* 23 (December): 405–18.

Hoffman, Elizabeth, Kevin McCabe, and Vernon L. Smith, 1996, "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review* 86 (June): 653–60.

Horowitz, Donald L. 1985. *Ethnic groups in conflict.* Berkeley. University of California Press

Imai, Kosuke, Gary King and Elizabeth A. Stuart. 2008. "Misunderstandings Between Experimentalists and Observationalists about Causal Inference*," Journal of the Royal Statistical Society A* 171:2: 481-502.

Imbens, Guido 2003. Semiparametric Estimation of Average Treatment Effects under Exogeneity: A Review.bTechnical report, Department of Economics, UC Berkeley.

Kahneman, Daniel, Jack Knetsch, and Richard Thaler. 1986. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review* 76 (September): 728– 41.

Khwaja, Asim. 2002. "Can Good Projects Succeed in Bad Communities? Collective Action in the Himalayas." Unpublished manuscript, Harvard University.

King, Gary and Jonathan Wand. 2007. "Comparing Incomparable Survey Responses: New Tools for Anchoring Vignettes." *Political Analysis* 15 (Winter): 46-66.

King, Gary, Christopher J.L. Murray, Joshua A. Salomon and Ajay Tandon. 2003. "Enhancing the Validity of Cross- cultural Comparability of Survey Research." *American Political Science Review* 98 (February): 191-207.

Knaus, Gerald and Felix Martin. 2003. "Travails of the European Raj." *Journal of Democracy* 14 (July): 60-74.

Kuran, Timur. 1998. "Ethnic Norms and their Transformation through Reputational Cascades." *Journal of Legal Studies* 27 (June): 623–59

Lijphart, Arend. 1968. *The politics of accommodation: pluralism and democracy in the Netherlands*. Campus; 142 2d ed., rev ed. Berkeley: University of California Press.

Lijphart, Arend. 1977. *Democracy in plural societies: a comparative exploration*. New Haven: Yale University Press.

Lijphart, Arend. 1999. *Patterns of democracy: government forms and performance in thirty-six Countries*. New Haven: Yale University Press.

Lustick, Ian S., Dan Miodownik & Roy J. Eidelson. 2004. "Secessionism in Multicultural States: Does Sharing Power Prevent or Encourage It?" *American Political Science Review* 98(May):209–29.

Miguel, Edward, and Mary Kay Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89 (December): 2325–68.

Miguel, Edward. 1999. "Ethnic Diversity, Mobility, and School Funding: Theory and Evidence from Kenya." Unpublished manuscript, Harvard University.

Miguel, Edward. 2004. "Tribe or Nation? Nation-Building and Public Goods in Kenya versus Tanzania." World Politics 56 (3): 327-362.

Mozaffar, Shaheen, James R. Scarritt, andGlenGalaich. 2003. "Electoral Institutions, Ethnopolitical Cleavages, and Party Systems in Africa's Emerging Democracies" *American Political Science Review* 97 (August) 379–90.

Nordlinger, Eric A. 1972. *Conflict regulation in divided societies*. Center for International Affairs, Harvard University.

Norris, Pippa. 2004. *Electoral engineering: voting rules and political behavior.* Cambridge, UK: Cambridge University Press.

Onyach-Olaa, Martin. 2003. "The Challenges of Implementing Decentralization: Recent Experiences in Uganda." *Public Administration and Development* 23 (January): 105–13.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* New York: Cambridge University Press.

Ostrom, Elinor. 2000. "Collective Action and the Evolution of Social Norms." *Journal of Economic Perspectives* 14 (Summer): 137–58.

Petersen, Roger D. 2002. *Understanding Ethnic Violence: Fear, Hatred, and Resentment in Twentieth-Century Eastern Europe.* New York: Cambridge University Press.

Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa*. New York: Cambridge University Press.

Posner, Daniel N. 2004. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi." American Political Science Review, 98, 4 (November): 529-545.

Poterba, James. 1997. "Demographic Structure and the Political Economy of Public Education." *Journal of Fiscal Analysis and Management* 16 (Winter): 48–66.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83(December):1281–1302.

Reynolds, Andrew. 2002. *The architecture of democracy: constitutional design, conflict management, and democracy*. Oxford, UK: Oxford University Press

Saideman, Stephen M., Danid J. Lanoue, Michael Campenni and Samuel Stanton. 2002. "Democratization, Political Institutions and Ethnic Conflict: A Pooled Time-Series Analysis, 1985-1998."*Comparative Political Studies* 35(   ):103–29.

Spolaore, Enrico, and Romain Wacziarg. 2006. "The Diffusion of Development." Unpublished manuscript, Stanford University.

Tajfel, Henri, Billig M. G., R. P. Bundy and Claude Flament. 1971. "Social Categorization and Integroup Behavior*." European Journal of Social Psychology* 1(February):149–77.

Tajfel, Henri. 1974. "Intergroup Behavior, Social Comparison and Social Change." Katz-Newcomb Lectures, University of Michigan,
Ann Arbor.

Tajfel, Henry and Marlene E. Turner. 1986. *The Social Identity Theory of Intergroup Behavior*. In Psychology of Integroup Relations, ed. Stephen Worchel and William Austin. Chicago: Nelson Hall pp. 7–24.

Tsai, Lily. 2007a. "Solidary Groups, Informal Accountability, and Local Public Goods Provision in Rural China." *American Political Science Review*, 101: 2 (May): 355-372.

Tsai, Lily 2007b, *Accountability Without Democracy: Solidary Groups and Public Goods Provision in Rural China,* Cambridge: Cambridge Studies on Comparative Politics, Cambridge University Press.

Van Evera, Stephen. 2001. "Primordialism Lives!" APSA-CP. 12(Winter 2001) 20-2.

Van den Berghe, Pierre L. 1981. *The ethnic phenomenon*. New York: Elsevier.

Vigdor, Jacob. 2004. "Community Composition and Collective Action: Analyzing Initial Mail Response to the 2000 Census." *Review of Economics and Statistics* 86: 1: 303–12.

Wade, Robert. 1994. *Village republics: economic conditions for collective action in South India*. San Francisco, Calif.: ICS Press.

Whitt, Sam and Rick K. Wilson. 2007. "The Dictator Game, Fairness and Ethnicity in Postwar Bosnia." *American Journal of Political Science* 51(June):655–668.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan , Eldar Shafir and Jon Zinman. 2005. "What's psychology worth? A Field Experiment in the Consumer Credit Market." Unpublished manuscript, Department of Economics, Harvard University.

Yamagishi, Toshio. 2007. "The Social Exchange Heuristic:A Psychological Mechanism That Makes a System of Generalized Exchange Self-Sustaining." In *Cultural and Ecological Foundations of the Mind*. Mark Radford, Susumu Ohnuma, and Toshio Yamagishi, eds. Sapporro: Hokkaido University Press.

Gymnasium Mostar: Mostars Other Landmark, Reconstruction and Revitalization Efforts Overview, OSCE Mission to BiH, January 2005.

Overview of Two Schools Under One Roof, OSCE Mission to BiH, May 2005. Report on Implementation of the Interim Agreement on Accommodation of Specific Needs and Rights of Returnee Children, Coordination Board for the Implementation of the Interim Agreement on Returnee Children, March 2005.

Vallery Perry, Reading Writing and Reconciliation: Educational Reform in Bosnia and Herzegovina, ECMI Working Paper #18, European Center for Minority Issues, September 2003.

Statistical Yearbook of BiH for 1991, Central Statistical Office of BiH, September.