Social contagion of ethnic hostility

Michal Bauer[a,b,1], Jana Cahlíková[c], Julie Chytilová[a,b], and Tomáš Želinský[b,d]

[a]CERGE-EI, A joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, 111 21 Prague, Czech Republic; [b]Institute of Economic Studies, Faculty of Social Sciences, Charles University, 110 00 Prague, Czech Republic; [c]Max Planck Institute for Tax Law and Public Finance, 80539 Munich, Germany; and [d]Faculty of Economics, Technical University of Košice, 040 01 Košice, Slovakia

Edited by George A. Akerlof, University of California, Berkeley, CA, and approved March 30, 2018 (received for review November 22, 2017)

Interethnic conflicts often escalate rapidly. Why does the behavior of masses easily change from cooperation to aggression? This paper provides an experimental test of whether ethnic hostility is contagious. Using incentivized tasks, we measured willingness to sacrifice one's own resources to harm others among adolescents from a region with a history of animosities toward the Roma people, the largest ethnic minority in Europe. To identify the influence of peers, subjects made choices after observing either destructive or peaceful behavior of peers in the same task. We found that susceptibility to follow destructive behavior more than doubled when harm was targeted against Roma rather than against coethnics. When peers were peaceful, subjects did not discriminate. We observed very similar patterns in a norms-elicitation experiment: destructive behavior toward Roma was not generally rated as more socially appropriate than when directed at coethnics, but the ratings were more sensitive to social contexts. The findings may illuminate why ethnic hostilities can spread quickly, even in societies with few visible signs of interethnic hatred.

ethnic conflict | discrimination | hostile behavior | contagion | peer effects

Intergroup conflict remains one of the most pressing problems facing human society because it can give rise to civil wars, ethnic cleansing, and discrimination. Many violent conflicts escalate quickly and are characterized by relatively sudden changes in the behavior of masses, from cooperating with individuals across ethnic lines to taking an active part in ethnic aggression (1–6). However, it remains unclear what triggers this change in willingness to cause harm. Here, we explore the influence of peers and provide a study that investigates experimentally whether intergroup hostility is contagious. The aim was to capture the following situation in a controlled environment: Imagine a person living in a community in which other people suddenly start doing harm to members of a different ethnic group. Will this change in social environment make the person more hostile toward outsiders as well, thus causing harmful behavior to spread?

Our work is motivated by evidence suggesting that individuals within own social network may drive the diffusion of hostility toward other ethnic groups and nationalities. In Germany, the rise of the Nazi movement was much faster in areas with a high density of civic associations (7). During the Rwandan genocide, "hate-radio" fueled participation in violence not only in villages with radio access, but its influence was magnified via spillover effects into neighboring villages (6). In terms of policy, concerns about the spreading of ethnic hostility, and thus greater social costs, are embedded in the legal codes of many countries, which impose greater penalties for racially or ethnically motivated crimes. Despite the importance of this issue, causal evidence is still lacking on whether hostile attitudes and behaviors spread more easily when they target outgroup instead of ingroup members.

Social scientists have long studied the prevalence of intergroup discrimination by measuring how behavior in experimental tasks is affected by the identity of a counterpart. These efforts have been, in large part, motivated by conceptual work describing how narrow group identities may cement ingroup cooperation but also create a fertile ground for intergroup hostility (8–10). A unifying feature of the existing experiments has been the focus on individuals making anonymous decisions when isolated from others, largely abstracting from the influence of people from own social network, and thus leaving open the question how social context can activate discrimination.

Studies which employ real-world groups often find evidence that natural group attributes, including ethnicity, influence behavior in experiments. Punishers have been found to protect ingroup members in a study focusing on indigenous tribes in Papua New Guinea (11). Male subjects in Israeli-Jewish society were found to systematically discriminate based on ethnicity in the Trust game (12), and exogenous allocation of individuals to real-life social groups in the Swiss Army led to increased cooperation rates among ingroup members in the Prisoner's Dilemma game (13). At the same time, little evidence of discrimination based on shared language was detected among adolescents in Italy (14). Also, a large-scale study from Kenya, a setting with a recent history of interethnic violence, found no evidence of coethnic bias (15), suggesting that intergroup discrimination among natural groups may not necessarily manifest in everyday decisions but may be triggered by situational factors.

Another approach is to create social categories in the laboratory (16) by dividing participants based on an irrelevant characteristic, such as preference for artwork. The groups lack social content, but the experimenter has more control over the identity formation process. Starting with the classical experiments of Tajfel et al. (16), studies using such "minimal groups" mostly find that people placed in groups favor income allocations to members of own group (16–19). A study focusing on a wide range of tasks that pit self-interest against the welfare of others showed that shared group identity increases altruistic sharing and reduces punishment of norm violators (17). Nevertheless, recent experiments have documented substantial heterogeneity in the

Significance

We provide experimental evidence on peer effects and show that behavior that harms members of a different ethnic group is twice as contagious as behavior that harms coethnics. The findings may help to explain why ethnic hostilities can spread quickly (even in societies with few visible signs of interethnic hatred) and why many countries have adopted hate crime laws, and illustrate the need to study not only the existence of discrimination, but also the stability of attitudes and behaviors toward outgroup members.

Author contributions: M.B., J. Cahlíková, J. Chytilová, and T.Ž. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data and statistical code allowing replication of reported figures and tables are available at Harvard Dataverse (doi:10.7910/DVN/G651WB).

[1]To whom correspondence should be addressed. Email: bauer@cerge-ei.cz.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1720317115/-/DCSupplemental.

Published online April 23, 2018.

extent of intergroup discrimination, suggesting that it may be driven by a minority of individuals (20).

The unique element of our experiment is the manipulation of social context, which allowed us to center the attention on dynamics of ethnic hostility caused by social influences. We tested the hypothesis that susceptibility to follow peers becomes magnified when harm is done to ethnic outgroup members compared with coethnics. Put differently, we estimated the extent of discrimination in "normal" circumstances (i.e., in a decision situation similar to previous studies) and then explored whether the prevalence of intergroup discrimination increases when an individual happens to observe hostility of peers. We also tested one potential mechanism: that social norms regulating destructive behavior toward outgroup members are more fragile and contingent on social context. In particular, we estimated how being allocated into an environment with hostile peers influences individual willingness to pay to harm others in an incentivized task (study 1) and how it affects perceptions about whether a harmful choice constitutes a normatively right way to behave (study 2).

The focus on peer influence relates our work to experiments which demonstrate that people often follow what others are doing. Building on the work of Asch (21) that documents high levels of conformity in people's (mis)judgements, researchers have shown that peers also affect the prevalence of prosocial (22) as well as unethical behaviors, including littering and cheating on examinations (23, 24). Because this evidence indicates that people may have a general tendency to conform, to cleanly identify the contagion of group-based hostility, we compared the spreading of hostile behavior toward outgroup members and toward coethnics.

Our sample consists of adolescents from a majority population (the Slavic ethnolinguistic group) in Eastern Slovakia. We study behavior toward Roma people (also known as Gypsies), the largest ethnic minority group in Europe. Since World War II, when Roma were targets of similar policies and persecution as Jews, there has not been any systematic violent conflict involving Roma. Nevertheless, the frequency of anti-Roma violence has been increasing in the last decade, especially in Central and Eastern Europe, making this region an apt natural setting for studying factors which facilitate the spread of hostility against a dissimilar ethnic group.

Study 1

Study 1 (n = 327) was implemented in 2013 in 13 schools (*SI Appendix*, Fig. S1), allowing us to study the influence of real-life peers. We administered a Joy of Destruction game, a money-burning task designed to identify hostile behavior (25, 26). Two players received an endowment of €2 each and simultaneously chose whether to pay €0.20 to reduce their counterpart's income by €1 or to keep the payoffs unchanged. Since the interaction is one-shot and anonymous and the destruction is costly for the decision maker, selfishness cannot justify destruction. We denoted the choice to reduce the other's payoff as hostile or destructive.

To measure discrimination, we implemented a SAME condition, in which an anonymous counterpart was a coethnic, and an OTHER condition, in which a counterpart was a member of the Roma minority. Specifically, subjects were provided a list of 20 typical majority or Roma names of potential counterparts from an unspecified distant school in Eastern Slovakia, chosen to reliably signal ethnicity. To uncover the dynamics of hostility and estimate the susceptibility to follow peers, subjects were matched with two classmates and all three sequentially made the decision whether to destroy the resources of a counterpart from the same name list. Given that the matching of peers and the order of choices were randomly determined, this design allowed us to identify the influence of peers among those who made their choice second or third and thus observed either hostile peers (DESTRUCTIVE PEER) or nonhostile peers (PEACEFUL PEER) before making their own decision. *SI Appendix*, Table S1 shows that the observable characteristics of subjects who made choices in DESTRUCTIVE PEER and PEACEFUL PEER conditions are statistically indistinguishable, indicating that the allocation to social environment was indeed exogenous. The greater the influence of being in an environment with hostile peers on individual hostility, the greater the extent of spreading of hostile behavior. Ultimately, we aimed to explore interaction effects—that is, whether hostile peers are more influential in OTHER than in SAME.

The decision environment in our field laboratories aimed to mimic real-life situations in which hostility may spread within cohesive social groups. We studied the influence of peers in a context where the three matched individuals knew each other well, directly observed each other's decision, and faced a common fate in the sense that one of the three individual decisions was randomly chosen to be payoff relevant for all three matched peers. Therefore, the estimated influence of peers on choices may capture multiple channels: subjects may follow peers because they learn about the prevailing social norm, update their beliefs about the characteristics of counterparts, or take into account peer preferences. Details about the setting and the experimental design are provided in *Materials and Methods* and *SI Appendix*.

We found that peers were influential in shaping individual willingness to destroy. Strikingly, the influence more than doubled when the choices impacted Roma (Fig. 1*A* and Fig. S2). We started by analyzing the choices of individuals who observed one peer before making their own decision. When the subjects observed a peaceful peer, the prevalence of destructive choices in OTHER was 19%. When the preceding player was destructive, the prevalence sharply increased to 77% (*SI Appendix*, Table S2). This difference is large in magnitude and highly significant statistically (χ^2 test, $P < 0.001$). Peers also affected behavior in SAME, but to a lesser extent: The prevalence of destructive behavior increased from 23% when a peer was nondestructive to 51% when he was destructive ($P = 0.01$). In a regression analysis, we found a strong positive interaction effect between DESTRUCTIVE PEER and OTHER on the likelihood of choosing a destructive action (Table 1; $P = 0.02$). As a result, discrimination against the ethnic minority arose when participants observed classmates' hostility, and the gap is large in magnitude (26 percentage points, $P = 0.03$). In contrast, when participants observed peaceful peers, we found virtually no differences in behavior across OTHER and SAME conditions ($P = 0.62$).

Similar patterns emerge for subjects who observed the choices of two peers. We compared the behavior of individuals who observed destructive behavior of both classmates vs. individuals who did not (those who observed one or both classmates being peaceful). The difference is 70 percentage points (88% vs. 18%, respectively; $P < 0.001$) in OTHER, and 38 percentage points (67% vs. 29%, respectively; $P < 0.001$) in SAME. The difference in the estimated influence of destructive peer behavior is statistically significant across conditions (OTHER*DESTRUCTIVE PEER; Table 1; $P = 0.04$). Interestingly, observing one destructive and one peaceful peer is not enough to trigger ethnic discrimination (*SI Appendix*, Table S3). Lastly, we pooled the choices of subjects who observed one or two peers and found that the interaction effect of observing destructive peers and OTHER ("Second and third decision makers" column in Table 1) is again large and statistically significant (35 percentage points, $P = 0.001$). Further results reported in *SI Appendix* show that the main pattern is robust. The findings hold across different locations in Eastern Slovakia; are robust to using alternative estimators, controlling for various design features and school fixed effects; and are not driven by a lack of understanding or by

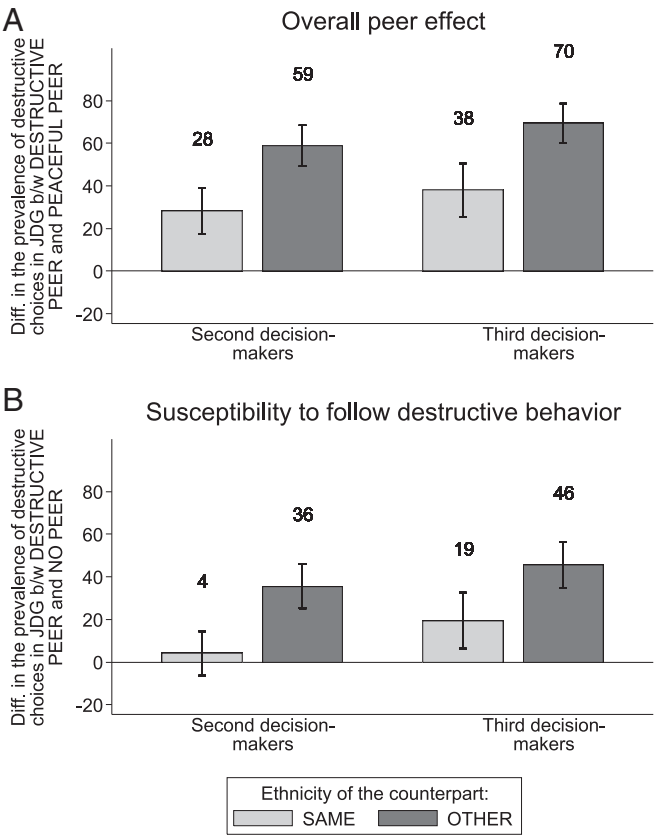

Fig. 1. Susceptibility to follow destructive behavior of peers magnifies when harm targets Roma instead of coethnics. The figure displays the difference (Diff.) in the prevalence of destructive choices in the Joy of Destruction game (JDG) between (b/w) DESTRUCTIVE PEER and PEACEFUL PEER (*A*), and between DESTRUCTIVE PEER and NO PEER (*B*). DESTRUCTIVE PEER indicates that all peers (one or two), whose choices a subject observed before making own decision, made a destructive choice. PEACEFUL PEER indicates that a decision maker observed at least one nondestructive peer before deciding. NO PEER indicates the choice was made by a first decision maker who did not observe anyone's choice before deciding. OTHER indicates that the experimental counterpart was a member of a different ethnic group (Roma), and SAME indicates that she/he was a coethnic. Bars indicate mean ± SEM.

individuals from families with low socioeconomic status (*SI Appendix*, Tables S4 and S5).

Next, we explored whether the greater influence of peers in OTHER is due to a greater susceptibility to follow destructive peer behavior, peaceful behavior, or both (i.e., general conformism). To analyze this (Fig. 1*B* and *SI Appendix*, Table S6), we used (as the benchmark) the choices of subjects who decided first among the three matched classmates and thus did not observe peer behavior before own choice (NO PEER). Relative to the NO PEER condition, subjects in the PEACEFUL PEER condition were less likely to be destructive, and the magnitude of this effect is very similar in SAME and OTHER conditions. In contrast, susceptibility to follow destructive peer behavior was strong in OTHER, while we found virtually no effect in SAME. The interaction effect is large in magnitude and highly significant statistically (Table 1, $P < 0.01$). These results, as well as additional evidence described in *SI Appendix*, reveal that the greater influence of peers in OTHER originates in greater readiness to follow unambiguously hostile peers, but not in differences in following peaceful or cooperative peers (*SI Appendix*, Tables S6–S8). It is also noteworthy that in NO PEER, the prevalence of destruction is very similar in SAME and OTHER. Put differently, subjects do not discriminate unless they happen to be in an environment with hostile peers (*SI Appendix*, Table S9).

Study 2

In study 2, we focused on one plausible mechanism that could explain the observed pattern of the spreading of destruction: whether individual perceptions of social norms are more fragile in OTHER than in SAME. The study was conducted in 2016 among adolescents ($n = 204$) from the same region. We implemented a coordination game (22, 27), which elicits perceived social (in)appropriateness of destructive behavior. The participants were described the Joy of Destruction game implemented in study 1 in four situations: the destructive choice was made either in SAME or in OTHER, and it was made either without observing peers (NO PEER) or after observing a destructive peer (DESTRUCTIVE PEER). Instead of making own choices, subjects rated the social appropriateness of a destructive choice on a six-point scale. In task 1, as in previous work (27), subjects were incentivized to match the modal response provided by others (from their school) rating the same choice environment. In addition, we implemented task 2, in which subjects were incentivized to estimate the rating in task 1 of 10 randomly selected subjects from their school. The data from task 2 provided information about the distribution of appropriateness ratings at an individual level, which helped us to assess whether individuals believed there was a social norm.

Note that coordination on an action does not necessarily measure social norms, since coordination games have multiple equilibria and subjects may coordinate in ways that have nothing to do with norms. This elicitation method yields a representation of a social norm if there is general social agreement that some actions are (in)appropriate, constituting the social norm, and if such shared perceptions create a focal point which can help subjects to match others' responses.

To gauge whether the participants believed a social norm existed, we first analyzed overall distributions of ratings in the NO PEER condition (*SI Appendix*, Table S10). The ratings in both tasks follow a positively skewed mound-shaped distribution, with a single peak reaching "quite socially inappropriate" (in one case, there is a second mode at a neighboring option, "somewhat socially inappropriate"). However, the peaks are not very high. The modal rating was chosen by 27 to 32% of the participants. Nevertheless, a destructive choice is rated as socially inappropriate (very, quite, or somewhat) in the majority of cases (75 to 83%). When looking at individual-level distributions of ratings from task 2, we also found that in most cases, they have a single mode (65%) or two modes represented by two neighboring options (21%).

Overall, these patterns indicate that most participants believed there was a social norm that the destructive choice is socially inappropriate, but the judgements varied about the precise extent of inappropriateness. Therefore, we used two measures of social norms in the analysis: rating of the appropriateness of the destructive choice on a six-point scale and a binary variable indicating whether the destructive choice was rated as socially appropriate or socially inappropriate.

The results reinforce the findings from study 1 (Table 2 and *SI Appendix*, Table S11). When evaluating behavior in an environment without hostile peers, the appropriateness ratings are similar in OTHER and in SAME; if anything, subjects perceived harming in OTHER as more inappropriate than in SAME. Importantly, however, in OTHER, the destructive action becomes perceived as less socially inappropriate when it follows peer behavior, whereas the normative judgments are quite stable in SAME (*SI Appendix*, Table S11). In task 1, an environment with hostile peers made destructive behavior 11 percentage points more likely to be rated as socially appropriate in OTHER ($P < 0.01$),

Table 1. The influence of peers on hostile behavior (study 1)

Condition	Destructive choice in the Joy of Destruction game[a]		
	Second and third decision makers	Second decision makers	Third decision makers
OTHER*DESTRUCTIVE PEER	0.35***	0.32**	0.32**
	[0.00]	[0.02]	[0.04]
OTHER[b]	−0.08	−0.04	−0.11
	[0.19]	[0.66]	[0.19]
DESTRUCTIVE PEER[c]	0.29***	0.28**	0.37***
	[0.00]	[0.01]	[0.00]
No. of observations	294	146	148
DESTRUCTIVE PEER in OTHER	0.64***	0.60***	0.70***
	[0.00]	[0.00]	[0.00]

We controlled for gender and school grade. SEs are robust. Because interaction effects cannot be readily interpreted in Probit models, we used the ordinary least-squares estimator in this table, and report the robustness of the results to using the Probit estimator in *SI Appendix*, Table S4. *P* values appear in brackets.

[a]The dependent variable is equal to +1 if a subject chose the destructive option in the Joy of Destruction game.

[b]OTHER indicates that the experimental counterpart was a member of a different ethnic group (Roma), rather than a coethnic.

[c]DESTRUCTIVE PEER indicates that all peers (one or two), whose choices a subject observed before making own decision, made a destructive choice.

***$P < 0.01$.

**$P < 0.05$.

while the shift was only 5 percentage points in SAME ($P = 0.24$). In task 2, the differences were 8 percentage points in OTHER ($P < 0.001$) and 4 percentage points in SAME ($P = 0.01$). The results are qualitatively similar when we use the measure of appropriateness rating on a scale. Consequently, destructive choices become perceived as less socially inappropriate in OTHER compared with SAME when made after a hostile peer, although for some of the measures, the difference is not statistically significant.

Lastly, in a regression analysis (Table 2), we tested whether the sensitivity to social context is significantly higher in OTHER than in SAME. We found that there is a positive interaction effect between OTHER and DESTRUCTIVE PEER on perceived appropriateness of destructive behavior measured on a scale ($P = 0.07$ for task 1 and $P = 0.02$ for task 2), as well as on the probability that a destructive choice is rated as appropriate ($P = 0.16$ for task 1 and $P = 0.04$ for task 2). The results do not change or, if anything, they become stronger when we restrict the sample to subjects who were more likely to believe a social norm existed, in the sense that their distribution of ratings in task 2 had a single mode or two modes represented by two neighboring options (*SI Appendix*, Table S12).

Discussion

Our findings provide reasons for optimism and caution at the same time. On the optimistic side, subjects do not discriminate when making choices in an environment in which peers are peaceful, which is an encouraging finding in light of the widespread concern about the pervasive nature of ethnic discrimination. On the pessimistic side, however, hostile behavior toward Roma is much more socially contagious than toward coethnics, in line with the view that contextual features may magnify or attenuate biases against outgroup members (19, 28–30).

This experiment raises as many questions as it answers. Below, we discuss limitations of our findings and offer some directions for future research. First, this paper studies behavior toward one ethnic group (Roma) among Slavic adolescents in one country. Clearly, more research is needed to explore whether our findings generalize to other settings, especially those with more salient interethnic conflicts.

Second, several plausible mechanisms can explain, in principle, the documented differential response to observing peers when the hostile action harms a member of the Roma minority compared with a coethnic. We provide suggestive evidence supporting the interpretation that peer behavior influences perceptions of social norms regulating hostile behavior toward Roma. However, subjects could also take peers' preferences more into account when the counterpart was Roma, perhaps because of concerns about future out-of-laboratory punishment for not conforming to peer behavior. An interesting open question is whether patterns of spreading of hostility would be similar in an environment in which the subjects receive information about the choices of anonymous individuals from own ethnic group, instead of directly observing peers with whom they share common fate and with whom they will interact in the future, as in our experiment.

Our findings may also be affected by participant's personal experiences. Ethnic groups often live segregated, as is the case in the setting we study, and interact relatively less frequently with individuals across ethnic lines, potentially leading to less knowledge about the outgroup members, or greater uncertainty about the attitudes of peers toward outgroup members. In such an environment, subjects could learn more from the actions of their peers toward outgroup members than from the same actions toward ingroup members, leading to a differential behavioral response. In line with this mechanism, our analysis of beliefs about the behavior of counterparts suggests that subjects held greater suspicion about hostility of Roma (*SI Appendix*). Future experiments could explore whether our findings also hold in settings in which people live in communities that are more ethnically mixed. Or, researchers could use a similar design to study behavior in minimal groups created in a laboratory, building on refs. 16 and 17, thus eliminating the role of beliefs caused by out-of-laboratory experiences. Ideally, one could leverage the relative advantages of both approaches by studying natural as well as minimal groups within a single sample (20) to assess whether the susceptibility to follow hostile behavior targeting outgroup members is a deep, generalizable response.

Taken at face value, our findings can shed light on several important phenomena and can have policy implications. First, they may illuminate why the ethnic hostilities of masses can spread quickly, even in societies in which there are few visible signs of systematic interethnic hatred (2, 3). Second, they may help to explain why "entrepreneurs of hatred" (i.e., individuals who

Table 2. The influence of peers on the perceptions of social norms (study 2)

	Rating of appropriateness of destructive choices in the Joy of Destruction game[a]			Destructive choices rated as appropriate (binary)[b]		
	All subjects			All subjects		
Condition	All estimates	Estimates of the modal rating (task 1)	Estimates of ratings of 10 other subjects (task 2)	All estimates	Estimates of the modal rating (task 1)	Estimates of ratings of 10 other subjects (task 2)
OTHER*DESTRUCTIVE PEER	0.07**	0.10*	0.07**	0.04**	0.06	0.04**
	[0.01]	[0.07]	[0.02]	[0.03]	[0.16]	[0.04]
OTHER[c]	−0.03**	−0.04	−0.03*	−0.02*	−0.03	−0.02
	[0.05]	[0.26]	[0.05]	[0.10]	[0.26]	[0.11]
DESTRUCTIVE PEER[d]	0.03	0.00	0.03	0.04*	0.05	0.03*
	[0.29]	[1.00]	[0.23]	[0.07]	[0.14]	[0.07]
No. of observations	8,806	816	7,990	8,806	816	7,990
DESTRUCTIVE PEER in OTHER	0.10***	0.10**	0.10***	0.08***	0.11***	0.08***
	[0.00]	[0.03]	[0.00]	[0.00]	[0.00]	[0.00]

We controlled for gender and school grade. SEs are clustered on a subject level. Ordinary least-squares; *P* values appear in brackets.
[a]The dependent variable is the appropriateness rating of destructive choice in the Joy of Destruction game ranging from −1 (very socially inappropriate) to +1 (very socially appropriate).
[b]The dependent variable is equal to +1 if subjects rated a destructive choice as socially appropriate (somewhat, quite, or very).
[c]OTHER indicates that the experimental counterpart was a member of a different ethnic group (Roma), rather than a coethnic.
[d]DESTRUCTIVE PEER indicates that the destructive choice was made after observing a destructive peer instead of in isolation.
***$P < 0.01$.
**$P < 0.05$.
*$P < 0.1$.

may benefit from causing social unrest and conflict)—perhaps aware of the inflammability of ethnic hostility—often choose minority groups as targets of their aggressive political campaigns. Third, since ethnic hostility has larger social costs that go beyond the immediate victim by potentially threatening social cohesion, the results can help to explain why many societies have found it desirable to institute hate crime laws. Last, our results suggest that although the readiness to be hostile may be quite latent and thus harmless in peaceful times, it may gain importance when tensions arise. Thus, early diagnoses seem of vital importance when fighting ethnic hostility.

Materials and Methods

Study 1.

Sample selection. Study 1 was conducted in Eastern Slovakia in 13 schools in villages and small towns attended predominantly by majority students but located within 5 km from a Roma neighborhood or settlement. The sample consists of 327 majority-ethnicity adolescents, aged 13 to 15 y.

Experimental tasks. In the Joy of Destruction game, two players received €2 each and simultaneously chose whether to pay €0.20 to reduce their counterpart's income by €1 or to keep the payoffs as they were. Participants first made the choice without knowing what the other player did. Subsequently, they were also asked to state their beliefs about the decision of the counterpart (using hypothetical questions) and, using the strategy method, to make two conditional decisions—for both possible actions of the counterpart. Subjects also participated in the Prisoner's Dilemma game, described in *SI Appendix*.

Manipulation 1: Identity of the counterpart. In the SAME condition, the counterpart came from the majority group (i.e., was a coethnic), whereas in the OTHER condition, the counterpart was Roma. The two conditions were implemented "within-subject", in a random order. In both conditions, the counterpart was unknown to the decision maker and came from an unspecified distant school in Eastern Slovakia. To signal ethnicity, we revealed the list of 20 names of potential counterparts (10 male and 10 female names). In SAME, the list contained typical majority names (e.g., Macejko, Poláková) and in OTHER, it contained typical Roma names (e.g., Horváth, Pačanová). To mitigate the experimenter-demand effect that can arise if subjects realize that the aim of the research is to study discrimination, we followed the approach of studies that provided subjects information about the name which strongly signals a given group attribute (12, 31). Nevertheless, such an indirect way of signaling a group attribute may lead to a conservative estimate of discrimination, as compared with using more salient signals, such as direct information about a group attribute (32) or personal contact (33). The names were actual names of students who made choices in the same tasks at the beginning of the data collection. The name lists were read aloud by the experimenter and included on the top of all answer sheets.

Manipulation 2: The social context of the decision. Subjects were randomly matched with two other peers from their class and sequentially made the same decision in a randomly determined order. The matching and the order remained the same for both tasks and across SAME and OTHER. Each individual decision was observed by the other two peers, and all subjects knew this a priori. No communication was allowed. Subjects in the DESTRUCTIVE PEER condition observed one peer who was destructive or two peers who were both destructive, while subjects in PEACEFUL PEER observed one peer who was nondestructive or two peers, one or both of whom were nondestructive. Subjects in NO PEER made their decisions first among the three matched peers. In addition, subjects in the INDIVIDUAL condition decided in isolation, without being observed by classmates and with no information regarding their classmates' choices (see *SI Appendix* for analysis). The social context of the decision was thus manipulated "between-subject" and orthogonally to the manipulation of coun-terpart's ethnicity.

Payoffs. In each identity condition (SAME or OTHER), one choice was randomly selected to be payoff relevant. Subjects received rewards in the form of credit to order items from an experimental store, which contained 48 items, ranging from snacks and drinks to stationery. In the conditions in which subjects were matched with two peers, one of the three individual decisions was randomly chosen to be payoff relevant for all three matched peers. Each subject's payoff was further determined by the choice of an experimental counterpart—that is, a randomly selected person from the name list (SAME or OTHER). For each counterpart from the name lists, only a choice of one subject from the first two schools was consequential for the payoff.

Study 2.

Sample selection. Social norms were elicited among a different sample of subjects ($n = 204$) who were in the same age range and came from the same area as the participants in study 1. The data collection took place in three schools.

Experimental tasks. The participants were first described the Joy of Destruction game, using the original instructions. Each subject was asked to rate the social appropriateness of a destructive choice on the following scale (converted into

numerical scores for the analysis): "very socially inappropriate" (−1), "quite socially inappropriate" (−0.6), "somewhat socially inappropriate" (−0.2), "somewhat socially appropriate" (0.2), "quite socially appropriate (0.6), or "very socially appropriate" (+1). In task 1, subjects estimated the modal appropriateness rating of other subjects in their school. In task 2, they estimated the appropriateness rating in task 1 of 10 randomly selected subjects from their school. Subjects were paid according to either task 1 (they received €2.00 if their response matched the modal response in their school) or task 2 (they received €0.40 per each correct guess out of 10).

Experimental manipulations. Subjects rated the social appropriateness of a destructive choice in four situations: the destructive choice was made either in SAME or in OTHER and it was made either in isolation from others without observing a peer (NO PEER) or after observing a destructive peer (DESTRUCTIVE PEER). One randomly chosen situation was payoff relevant.

Further information on the methods and the full experimental protocol for study 1 and study 2 are provided in *SI Appendix*.

Research approval. The experiments were approved by the Director of the Institute of Economic Studies at Charles University and by the Dean of the Faculty of Economics, Technical University of Košice. Since the participants were adolescents, they could not provide informed consent themselves. Detailed information about the experiments was provided to the headmasters of all participating schools who gave their approval. In all cases, they also decided that this data collection satisfied the conditions of general informed consent regulating participation in research-related data collections, signed by the parents at the beginning of the school year, and thus decided not to seek additional parental consent for this particular data collection. Participation was voluntary and the subjects could leave at any time. All students decided to complete the tasks.

Data and code availability. The data and statistical code allowing replication of reported figures and tables are available at Harvard Dataverse (doi:10.7910/DVN/G651WB).

ACKNOWLEDGMENTS. The data collection was funded by the Czech Science Foundation Grant P402/12/G130. Further support was provided by the Czech Science Foundation Grant 17-13869S (to M.B. and J. Chytilová) and the Slovak Research and Development Agency Grant APVV-0125-12 (to T.Ž.).

1. Esteban J-M, Ray D (2008) On the salience of ethnic conflict. *Am Econ Rev* 98: 2185–2202.
2. Basu K (2005) Racial conflict and the malignancy of identity. *J Econ Inequal* 3:221–241.
3. Fearon JD, Laitin DD (2000) Violence and the social construction of ethnic identity. *Int Organ* 54:845–877.
4. Bardhan P (2005) *Scarcity, Conflicts and Cooperation: Essays in Political and Institutional Economics of Development* (MIT Press, Cambridge, MA).
5. Glaeser EL (2005) The political economy of hatred. *Q J Econ* 120:45–86.
6. Yanagizawa-Drott D (2014) Propaganda and conflict: Evidence from the Rwandan genocide. *Q J Econ* 129:1947–1994.
7. Satyanath S, Voigtländer N, Voth H-J (2016) Bowling for fascism: Social capital and the rise of the Nazi Party. *J Polit Econ* 125:478–526.
8. Tajfel H, Turner J (1979) An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations*, eds Austin W, Worchel S (Brooks/Cole, Monterey, CA).
9. Akerlof GA, Kranton RE (2000) Economics and identity. *Q J Econ* 115:715–753.
10. Choi J-K, Bowles S (2007) The coevolution of parochial altruism and war. *Science* 318: 636–640.
11. Bernhard H, Fischbacher U, Fehr E (2006) Parochial altruism in humans. *Nature* 442: 912–915.
12. Fershtman C, Gneezy U (2001) Discrimination in a segmented society: An experimental approach. *Q J Econ* 116:351–377.
13. Goette L, Huffman D, Meier S (2006) The impact of group membership on cooperation and norm enforcement. *Am Econ Rev* 96:212–216.
14. Angerer S, Glätzle-Rützler D, Lergetporer P, Sutter M (2016) Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *Eur Econ Rev* 90:254–264.
15. Berge LIO, et al. (2015) How strong are ethnic preferences? NBER Working Paper No. 21715. National Bureau of Economic Research, Cambridge, MA).
16. Tajfel H, Billing MG, Bundy RP, Flament C (1971) Social categorization and intergroup behavior. *Eur J Soc Psychol* 1:149–178.
17. Chen Y, Li SX (2009) Group identity and social preferences. *Am Econ Rev* 99:431–457.
18. Chen R, Chen Y (2011) The potential of social identity for equilibrium selection. *Am Econ Rev* 101:2562–2589.
19. Pechar E, Kranton R (2017) Moderators of intergroup discrimination in the minimal group paradigm: A meta-analysis. Available at https://sites.duke.edu/rachelkranton/files/2016/09/Moderators-of-Intergroup-Discrimination.pdf. Accessed April 6, 2018.
20. Kranton R, Pease M, Sanders S, Huettel S (2018) Groupy and not groupy behavior: Deconstructing bias in social preferences. Available at https://sites.duke.edu/rachelkranton/files/2018/03/groupnotgroupy-march2018final.pdf. Accessed April 6, 2018.
21. Asch SE (1951) Effects of group pressure upon the modification and distortion of judgment. *Groups, Leadership and Men*, ed Gutzkow H (Carnegie Press, Pittsburgh).
22. Gächter S, Nosenzo D, Sefton M (2013) Peer effects in pro-social behavior: Social norms or social preferences? *J Eur Econ Assoc* 11:548–573.
23. Keizer K, Lindenberg S, Steg L (2008) The spreading of disorder. *Science* 322: 1681–1685.
24. Gino F, Ayal S, Ariely D (2009) Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychol Sci* 20:393–398.
25. Abbink K, Sadrieh A (2009) The pleasure of being nasty. *Econ Lett* 105:306–308.
26. Prediger S, Vollan B, Herrman B (2014) Resource scarcity and antisocial behavior. *J Public Econ* 119:1–9.
27. Krupka EL, Weber R (2013) Identifying social norms using coordination games: Why does dictator game sharing vary? *J Eur Econ Assoc* 11:495–524.
28. Hjort J (2014) Ethnic divisions and production in firms. *Q J Econ* 129:1899–1946.
29. Shayo M, Zussman A (2011) Judicial ingroup bias in the shadow of terrorism. *Q J Econ* 126:1447–1484.
30. Alexander M, Christia F (2011) Context modularity of human altruism. *Science* 334: 1392–1394.
31. Bertrand M, Duflo E (2017) Field experiments on discrimination. *Handbook of Economic Field Experiments*, eds Banerjee A, Duflo E (Elsevier, North-Holland, Amsterdam), Vol 1, pp 309–393.
32. Hoff K, Pandey P (2006) Discrimination, social identity, and durable inequalities. *Am Econ Rev* 96:206–212.
33. Riach PA, Rich J (2002) Field experiments of discrimination in the market place. *Econ J* 112:F480–F518.