

Working Paper Series
(ISSN 1211-3298)

697

**Shifting Punishment on Minorities:
Experimental Evidence of Scapegoating**

**Michal Bauer
Jana Cahlíková
Julie Chytilová
Gérard Roland
Tomáš Želinský**

CERGE-EI
Prague, July 2021

ISBN 978-80-7343-504-2 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-599-7 (Národohospodářský ústav AV ČR, v. v. i.)

Shifting Punishment on Minorities: Experimental Evidence of Scapegoating

Michal Bauer
Jana Cahlíková
Julie Chytilová
Gérard Roland
Tomáš Želinský*

[This version: June 2021]

Abstract

This paper provides experimental evidence showing that members of a majority group systematically shift punishment on innocent members of an ethnic minority. We develop a new incentivized task, the Punishing the Scapegoat Game, to measure how injustice affecting a member of one's own group shapes punishment of an unrelated bystander ("a scapegoat"). We manipulate the ethnic identity of the scapegoats and study interactions between the majority group and the Roma minority in Slovakia. We find that when no harm is done, there is no evidence of discrimination against the ethnic minority. In contrast, when a member of one's own group is harmed, the punishment "passed" on innocent individuals more than doubles when they are from the minority, as compared to when they are from the dominant group. These results illuminate how individualized tensions can be transformed into a group conflict, dragging minorities into conflicts in a way that is completely unrelated to their behavior.

Key words: punishment, minority groups, inter-group conflict, discrimination, scapegoating, lab-in-field experiments

JEL Codes: C93, D74, D91, J15

* Bauer: CERGE-EI, a joint workplace of Center for Economic Research and Graduate Education, Charles University and the Economics Institute of the Czech Academy of Sciences, Politických veznu 7, P.O. Box 882, 111 21 Prague 1, Czech Republic and Institute of Economic Studies, Faculty of Social Sciences, Charles University, phone: 00420224005130, fax: 00420224005333, e-mail: bauer@cerge-ei.cz; Cahlíková: Max Planck Institute for Tax Law and Public Finance, 80539 Munich, Germany; Chytilová: Institute of Economic Studies, Faculty of Social Sciences, Charles University and CERGE-EI, 110 00 Prague, Czech Republic; Roland: UC Berkeley, CEPR and NBER, 94720 Berkeley CA, USA; Želinský: Technical University of Košice, Némcevej 32, 040 01 Košice, Slovakia and Durham University, School of Education, Leazes Road, Durham DH1 1TA, United Kingdom.

Abstrakt

Tato studie ukazuje, že lidé z majoritní etnické skupiny mají tendenci trestat nevinné členy minoritní skupiny. Popisujeme nový ekonomický experiment, který nazýváme "Trestání obětních beránků". Umožňuje nám měřit, jak pozorování anti-sociálního chování, jehož obětí je někdo z vlastní skupiny, ovlivňuje chování k nezúčastněným lidem (obětním beránkům), kteří se na bezpráví nijak nepodíleli. Zajímá nás, zda je chování v jinak totožné situaci ovlivněno tím, jakou má „obětní beránek“ etnickou příslušnost, konkrétně zda je to člen romské minority či člen majoritní populace na východním Slovensku. V situaci, kdy členové majoritní skupiny nečelí anti-sociálnímu chování, nevidíme diskriminační chování proti Romům. Pokud je ovšem někdo z majoritní skupiny poškozen a není možné přímo potrestat viníka, rozhodne se část lidí potrestat nezúčastněného „obětního beránka“. Naše hlavní zjištění je, že tato ochota přesměrovat trest na nevinné lidi je dvakrát větší, pokud je obětním beránkem Rom, než pokud je jím člen majoritní etnické skupiny. Tyto výsledky dokumentují mechanismus, který může přetvářet lokalizované tenze mezi jednotlivci ve skupinový konflikt tím, že vtáhne minority do rozepří, na jejichž vzniku se žádným způsobem nepodílely.

1. Introduction

How can localized social ills spiral into widespread aggressive behavior, often with tragic consequences for whole societies? Although it is widely acknowledged that aggressive behavior against ethnic minorities, including pogroms, often arise unexpectedly and spread quickly even in previously peaceful communities (Fearon and Laitin 2000; Bardhan 2005; Esteban and Ray 2008)¹, little is known empirically about the mechanisms that underlie these *changes* in behavior.

In this paper, we study people's biases in punishment behavior, in response to wrongdoing targeted towards members of one's own group. A novel aspect is that we focus on situations when punishers cannot punish the wrongdoer but only "bystanders", as in many real-life situations. Inspired by work in social psychology and political science (Doob et al. 1939; Allport 1954; Girard 1979; Staub 1992; Glick 2005), we take seriously the idea that people may have a tendency to pass revenge on innocent individuals from negatively-stereotyped minority groups. We refer to such behavior as minority scapegoating.² To study this behavioral phenomenon, we develop a new experimental task and document that a desire to "pass" the punishment on innocent individuals more than doubles when such persons are from an ethnic minority, as compared to when bystanders are from the dominant group. Such shifting of punishment on minorities violates a fundamental fairness principle (Kant 1965 and classical philosophers) embedded in the legal codes of most modern societies, i.e. that people should be punished only for wrongs they are responsible for and that they intentionally committed. Demonstrating such bias in punishment preferences of dominant groups is important, because it may drag minorities into conflicts that are completely unrelated to their behavior and transform individualized tensions into group conflicts.

¹ Minority groups are targets of violence in many parts of the world (Horowitz 1985; Yanagizawa-Drott 2014). Historical evidence suggests that aggressive attitudes and behaviors, including pogroms and attempted genocides, increase during periods of social unrest and economic problems within the majority group (Anderson, Johnson, and Koyama 2017; Voigtlander and Voth 2012; Grosfeld, Sakalli, and Zhuravskaya 2020). Survey evidence from the US and Europe indicates that hostility towards 'other' groups, especially weaker groups like migrants or sexual minorities, increases when social cohesion within the majority group starts to fall apart, creating a sense of normative threat (Stenner 2005; Stenner and Haidt 2018).

² The term "scapegoat" has biblical origins, coming from the Book of Leviticus—it refers to a sacrificial goat to which the ancient Israelites symbolically transferred their sins. A scapegoat was originally understood as a person or an animal that symbolically absorbed sins of others and took them away from those who committed them (Crossman 2019). According to social psychology, scapegoating occurs when punishment of the true source of the anger is inhibited and people shift their aggression towards other individuals (Newman and Caldwell 2005; Allport 1954; Glick 2005). Staub (1992, p. 48) describes this behavior as follows: "*When there is no aggressor or the aggressor is too powerful or the source of responsibility cannot be identified or the responsibility is one's own (or one's group), identifying a scapegoat will have "beneficial" psychological effects*". Scholars have also suggested that such displaced punishment typically targets individuals or groups who are weak and vulnerable, as well as already deprecated minority groups, perhaps because it is easier to *ex post* (irrationally) blame them for one's own misfortunes and to attribute negative imputed characteristics to justify aggression against them (Bettelheim and Janowitz 1950).

Identifying people's desire to pass punishment on innocent individuals, and its determinants, is empirically challenging with observational data. First, it is nearly impossible to rule out the role of the standard economic incentives to harm innocent individuals, such as simple self-interested grabbing of resources from them. In addition, in most real-life situations there is an element of uncertainty about who originated the harm. Members of the dominant group may punish innocent individuals from minority groups, because they attribute responsibility for misfortunes to actions of minority groups. A controlled experimental environment allows eliminating these confounding mechanisms, and focusing on whether shifts in punishment onto minorities are embedded in preferences or decision heuristics. Specifically, our aim is to identify an individual desire to engage in harmful behavior towards unrelated parties, in response to experienced hostility against someone one cares about. In light of the literature on scapegoating, the main question is whether indirect revenge becomes stronger when the unrelated party (bystander) is a member of a minority group. Answering this question requires an experimental setup that allows a researcher to (i) measure how people behave when punishment of wrongdoers is inhibited and people can punish only individuals who could not have causally contributed to the original harmful act, (ii) measure punishment responses in one-shot anonymous interactions that are costly for the punisher, and that provide no scope for material benefits of punishment, and (iii) compare behavior towards a weaker minority group and towards the own group.

To address these challenges, we employ a novel experimental paradigm, the *Punishing the Scapegoat Game*. In this game, impartial spectators from the majority ethnic group can impose a monetary punishment on others at own cost, after observing that someone malevolently destroyed the earned income of an individual from their own group. The existing incentivized experiments on punishment of socially undesirable behavior focus exclusively on direct punishment of individuals who make active decisions whether or not to violate a social norm, including the Third Party Punishment Game³ on which we build. In contrast, a key element of the Punishing the Scapegoat Game is that we add a fourth person, the passive Scapegoat who does not know the wrongdoer and who is not involved in any way in the original wrongdoing. This feature allows us to separate the person who commits a harmful act and a person whom the impartial spectator can punish. We exogenously manipulate information about the ethnicity of the Scapegoat to test whether people are more prone to redirect the punishment onto minority group members, if they cannot punish the Wrongdoer.

³ The Third Party Punishment game has been originally developed by Fehr and Fischbacher (2004) and later became one of the prominent experimental tasks to study direct punishment of norm-violators, see, for example, Bernhard, Fischbacher, and Fehr (2006); Goette, Huffman, and Meier (2006); Henrich et al. (2010).

We are interested in how behavior of spectators (Punishers) from the majority ethnic group responds to information on whether a Wrongdoer caused harm to a Victim from the majority ethnic group and how serious the harm was.⁴ We measure whether the gap in punishment of the Roma Scapegoat as compared to the majority Scapegoat is larger in a situation when the Punishers respond to observing harmful behavior targeting a member of their group than in a situation when no such harm was done. An important feature of using such difference-in-difference approach to identify scapegoating is that we can test whether out-group bias in punishment happens above and beyond out-group bias in circumstances when people do not respond to harm happening to their own group.

We study the interaction between real ethnic majority and minority group members. The setting is Eastern Slovakia, a region experiencing high unemployment, where the majority population lives side-by-side with an economically disadvantaged, segregated, and negatively stereotyped Roma ethnic minority, which represents approximately 15% of the local population. Scattered across many European countries, including Slovakia, the Czech Republic, France, Italy, and Hungary, Roma people are the largest ethnic minority in Europe. Thus, Eastern Slovakia is an apt natural setting to test whether aggression can be easily triggered towards members of a minority group that is relatively weak and negatively stereotyped.

We show that a non-negligible fraction of Punishers (23%) redirect punishment onto Scapegoats (bystanders) when they cannot punish the Wrongdoer. Importantly, the identity of the Scapegoat matters a lot. We find that the interaction between the intensity of harm committed by the Wrongdoer and the Scapegoats' identity has a large and robust effect on the punishment of the Scapegoat. When Wrongdoers harm the Victim, the punishment is twice as severe when the Scapegoat is from the Roma minority than when the Scapegoat is from the majority group. When Wrongdoers do not harm the Victim, ethnic majority Punishers do not discriminate against Scapegoats from the ethnic minority. Therefore, discrimination against the Roma minority Scapegoats *arises* only when Wrongdoers harm the Victim.

We consider several alternative mechanisms for why people exhibit a magnified tendency to retaliate against the minority. First, we show that it is not driven by collective punishment, i.e. by the case in which both the Wrongdoer and the Scapegoat are from the minority group. We find that the minority Scapegoats are punished more often and more severely than majority Scapegoats even when the original harm is committed by a member of the majority group, and thus when Wrongdoers and Scapegoats do not share the same ethnicity. Next, we argue that over-attribution of responsibility stemming from uncertainty about who committed the wrongdoing or statistical discrimination due to beliefs about how punishment

⁴ The harm committed by the Wrongdoer to the Victim was real. In a supplementary activity, the Wrongdoers made a decision whether and by how much to reduce earnings of a Victim earned for a tedious work assignment (8 euro).

may affect future interactions cannot explain our findings either, by virtue of the experimental design. Further, the patterns we observe also cannot be simply an outcome of stable unconditional out-group hostility against the minority that would manifest itself under any circumstances. Our results are consistent with discriminatory preferences being latent in “peaceful” times, but *activated* in environments when decision-makers respond to harm done to someone from their own group. There are two plausible psychological channels for why latent discriminatory preferences are predicted to more easily manifest themselves when people engage in taking revenge. Negative emotions associated with observing harm affecting someone from the own group may lower self-control and deep urges to discriminate may appear more easily. Also, the salience of conflict may provide a larger scope for excuses for individuals who desire to treat members of a certain group unfavorably while aiming to retain a self-image of being a moral person.

Finally, we study whether shifting of punishment on another group characterizes decisions of the dominant group towards a weaker minority, or also vice versa. To do so, we measure the behavior of Roma Punishers in the same experimental set up. We do not find evidence of biases in punishment of a Scapegoat bystander from the majority group among Roma decision-makers. This is in line with the view that shifting of punishment onto innocent individuals is psychologically easier when the target is a member of a negatively stereotyped, vulnerable, and smaller group (Bettelheim and Janowitz 1950; Allport 1954).

In addition to making decisions whether or not to lower the payoff of the Scapegoat, decision-makers also made choices whether to punish the Wrongdoer directly. In this task also, we observe a magnified tendency to engage in revenge: punishers from the majority group punish wrongdoers from the minority more harshly than members of their own group. Together, we establish that the biases in punishment preferences of the dominant group against the minority group we study are not limited to situations when members of the minority are responsible for harmful actions, but, importantly, mark behavior also in situations when it is clear that minority members have nothing to do with the harm done.

Our experimental design is most closely related to existing lab-in-field experiments on punishment of norm violations across group boundaries (Goette, Huffman, and Meier 2006; Bernhard, Fischbacher, and Fehr 2006). Earlier work has made progress in studying group biases in the direct punishment of active norm violators. In particular, among indigenous tribes in Papua New Guinea, Bernhard et al. (2006) document that third parties who share group membership with victims punish out-group members more than in-group members, which is similar to the patterns of direct Wrongdoer punishment that we find.⁵ We contribute by developing an experimental task that sheds light on determinants of the punishment of passive

⁵ Focusing on a different type of real-life social groups (based on poticial party affiliation, soccer club support and membership to different platoons in Swiss Army), a similar pattern has been found in Schiller, Baumgartner, and Knoch (2014) and Baumgartner et al. (2012), but not in Goette, Huffman, and Meier (2006) and Goette et al. (2012).

individuals. In this way, we can show that revenge may spill-over and affect behavior not only towards a wrongdoer but also towards unrelated individuals from an ethnic minority.^{6,7}

More broadly, the paper contributes to experiments that study existence of in-group vs. out-group biases based on real-life group attributes, such as ethnicity (Fershtman and Gneezy 2001; Goette, Huffman, and Meier 2006; Falk and Zehnder 2013; Bauer et al. 2018; Berge et al. 2020; Kranton et al. 2020). In-group favoritism and out-group discrimination have also been measured using a minimal-group experimental paradigm, by creating artificial group boundaries in the laboratory, based on, for example, having T-shirts of the same color or sharing preferences for art (Tajfel 1981; Chen & Li 2009). Here we show that out-group biases depend on the decision-environment – they are more pronounced when decisions happen in an environment characterized by wrongdoing against someone from the dominant group, as compared to biases revealed in decisions made in “peaceful” circumstances.

Further, our results speak to the ongoing debate about whether giving people opportunities to punish leads to desirable social outcomes or not. Many studies have documented that punishment can have a pro-social norm enforcement function (Fehr and Gächter 2002; Gächter, Renner, and Sefton 2008; Henrich et al. 2010). Recent work has started to explore the conditions in which providing an opportunity to punish may not have positive social effects (Herrmann, Thoni, and Gächter 2008; Abbink et al. 2010; Goette et al. 2012). For instance, Herrmann et al. (2008) find that, in some societies, a non-negligible fraction of people punish individuals who are relatively cooperative (a behavior called ‘anti-social punishment’), and that such behavior eliminates the pro-social effects of punishment. We show that anti-social punishment and scapegoating are two distinct behavioral phenomena: while we find no group bias in anti-social punishment (i.e., people do not punish cooperators from the other group more than their in-group members), we find systematic biases against the minority in the punishment of the Scapegoat.

Finally, the paper is related to recent experiments documenting the role of contextual factors on social behavior. It has been shown that people are less prone to behave morally when placed in decision-environments which provide “moral wiggle room” and a larger scope for excuses (Bénabou and Tirole 2011) by, for example, obscuring the role of the decision-maker in determining an outcome (Dana, Weber,

⁶ Another interesting line of experiments with a punishment option studies the attribution of responsibility and shifts in blame caused by the diffusion of responsibility for unkind actions across multiple individuals (Bartling and Fischbacher 2012). Here, we study redirecting punishment towards individuals who are completely passive. Responsibility for unkind behavior can thus not be attributed to them. Also, our focus is on the role of ethnicity of the passive person.

⁷ Interestingly, both Bernhard, Fischbacher, and Fehr (2006) and Goette, Huffman, and Meier (2006) find that people punish more severely wrongdoing targeting members of their own group, as compared to members of an out-group. This result motivated us to keep the identity of the victim fixed, and to focus on punishment responses for wrongdoing against own ethnic group.

and Kuang 2006; Falk, Neuber, and Szech 2020; Bauer et al. 2021). Our results suggest that when people engage in punishment, they may have less scruples to act on their latent discriminatory preferences.

2. Experimental Design

The experimental design is presented in six subsections. First, we provide a short background on the Roma ethnic minority. Second, we describe the sample selection. Third, we present the *Punishing the Scapegoat Game* we developed to measure punishment of the Scapegoat – an uninvolved bystander – in response to harmful behavior targeting a member of Punisher’s group. We also describe the *Punishing the Wrongdoer Game*, in which Punishers could directly punish the Wrongdoer. Fourth, we describe how we manipulate information about the identity of the Scapegoat and the Wrongdoer. Fifth, we describe how we elicited real wrongdoing from a supplementary sample of individuals who were given the option to malevolently reduce the earned income of other individuals. Last, we provide further details about our experimental procedures.

2.1 Background on the Roma ethnic minority in Europe

Eastern Slovakia represents an apt natural setting to study behavior of the dominant group towards a negatively stereotyped and weaker ethnic minority. The Roma people, a minority of Northern Indian origin, constitute the largest ethnic minority in Europe, estimated at 10-12 million persons.⁸ They live in generally poor socio-economic conditions and experience social exclusion all over Europe.⁹ The average education levels of Roma are low (only 20% finish upper-secondary education), they are poorly integrated into labor markets (less than one third are in paid employment), they generally live in substandard housing, and they have worse health and lower life expectancy than the majority populations. It is estimated that 85% of Roma in Europe live below national poverty lines. In Eastern Slovakia, the setting we study, the Roma represent around 15% of the local population. Around 65% of Slovakian Roma live segregated from the majority population, often in isolated settlements or on the edges of villages and towns. Previous research shows that the Roma are subject to prejudice and face discrimination in labor and housing markets (Bartoš et al. 2016). According to reports by the European Commission, almost one quarter of Europeans (38% of Slovaks) state that they would be uncomfortable having a Roma neighbor and 34% of Europeans (60% of

⁸ The background section draws heavily on our earlier paper (Bauer et al. 2018) about the Roma, which uses a different data set collected in the same setting.

⁹ The most significant populations are in Central and Eastern European countries (Bulgaria, Romania, Slovakia, Hungary and the Czech Republic), but Roma also live in France, Greece and Italy.

Slovaks) think citizens in their country would feel uncomfortable about their children having Roma classmates.

Since World War II, when the Roma were persecuted along with the Jews, there has not been any systematic violent conflict involving the Roma. Nevertheless, the frequency of anti-Roma violence has been increasing in recent decades, especially in Central and Eastern Europe (Council of Europe 2012). Anti-Roma marches have been staged in dozens of towns and cities across the region, and have commonly escalated in property damage and/or violence. Walls separating the majority population from their Roma neighbors have been built in numerous cities in Slovakia, Romania and Bulgaria. The most recent wave of hate speech and violence against Roma occurred during the Covid-19 pandemic and was interpreted by several commentators, including the United Nations Special Rapporteur on Minority Issues (UN News 2021; openDemocracy 2021), as “scapegoating” of Roma for the coronavirus pandemic.

2.2 Sample selection

The data collection took place between May and September 2017, across various localities in Eastern Slovakia. We study punishment behavior among a diverse (though not representative) sample of young adults of the majority Slovak ethnicity (N=337). We focus on young adults, because earlier research indicates that perpetrators of aggressive behavior against other ethnic groups, including hate crimes, are typically young (e.g., Levin and McDevitt 2002).

The subjects are students aged 18-23 from the last two grades of secondary school (57% of the sample) and a local university (43%). The secondary school students come from seven different schools, located in two regions—Košice and Prešov—and include students from the three most common types of high schools in Slovakia: general, technical, and business/commerce. In each region, we randomly selected at least one school of each type; all schools selected agreed to participate. The school headmasters determined which classes would take part. We randomly selected participants from all interested adult students (aged 18 and above) in those classes and the experiments were implemented in their respective schools. University students were recruited at the campus of the Technical University of Košice, and therefore the vast majority were enrolled in technical or economic majors. The sample characteristics are presented in Table A1 (in Supplementary Online Appendix A): the subjects are on average 19.3 years old, and 42% are female. Their parents completed at least secondary education in almost all cases; 25% of fathers and 32% of mothers also hold a university degree.

Further, we study the behavior of a sample of young adults from the Roma minority (N=484)¹⁰. We randomly selected 21 villages and towns with an estimated Roma population of over 1,000 (the list of eligible communities was based on the *Atlas of Roma Communities 2013*), across the same two regions. We contacted a village representative, typically a mayor, asking for permission to conduct research and to appoint a local contact person who was responsible for inviting participants (of 42 villages contacted, 21 agreed to take part). The local coordinators were instructed to select approximately 25 participants per village, Slovak-speaking, 18-24 years of age, literate, ideally living in different parts of the village and belonging to different social groups within the community. The experiments took place in local community centers. As expected, the Roma subjects had less education and came from a poorer socio-economic background (Table A2) compared to the subjects from the majority group: 15% were students, 49% were unemployed and among those no longer studying, 55% had finished only primary school.

2.3 Experimental tasks

The *Punishing the Scapegoat Game* is illustrated in Panel A of Figure 1. A Punisher is matched with three people – a Wrongdoer, a Victim and a Scapegoat, who are neutrally labeled Person A, Person B and Person C when described to the Punisher. Specifically, on a tablet computer, the Punisher observes three pictures: the first one with twenty potential Wrongdoers, the second one with twenty potential Victims, and the third one with twenty potential Scapegoats. Each picture displays twenty passport-style photos, homogenous in terms of ethnicity, and taken against a neutral background. The Punisher knows that s/he is matched with one person from each set of twenty photographs but does not know with whom specifically. The Punisher is informed that each of these three people (Wrongdoer, Victim, and Scapegoat) completed a work assignment and earned 8 euro for their work. Further, the Punisher learns that after completing their work, the Wrongdoer had an option to reduce the earnings of the Victim by 0, 2, 4, 6 or 8 euro, and that the Scapegoat was utterly passive. Punishers learn that the Wrongdoer, the Victim, and the Scapegoat are each from a different location, do not know each other, and that only the Wrongdoer had the option to reduce the earnings of someone else, and only the Victim's earnings could have been reduced. The task of the Punisher is to decide whether and by how much to reduce Scapegoat's payment. Thus, the decisions aim to capture scapegoating behavior: Punishers face a situation in which their ability to address the true source

¹⁰ We have intentionally selected a larger sample than from the majority group, because piloting of the experiment revealed that around 25% of the Roma sample did not properly understand the experimental instructions, based on crosscheck questions on understanding. In the analysis, we report results for the whole sample, as well as for a sub-sample that excludes individuals whose responses indicated limited understanding of the instructions. The results are robust.

of the social problem faced by the own group is inhibited – they cannot punish the Wrongdoer – but they can instead shift punishment onto a different person and treat him as a Scapegoat.

[Figure 1 around here]

Punishment is costly: reduction of each euro costs the Punisher 0.10 euro. Punishers' decisions are elicited for all five possible actions of the Wrongdoer towards the Victim, using a strategy method. When making the decision, in a simple tablet interface, the Punishers choose a preferred punishment level for each possible action of the Wrongdoer by pushing a “minus” or “plus” button—each tap on the tablet leads to a 2.00 euro decrease/increase of the earnings of the Wrongdoer or the Scapegoat. After each tap, the payoffs of the Wrongdoer, the Victim, the Scapegoat, and Punisher's own payoff are updated and clearly specified on the screen (sample decision screens are depicted in Figure A1). Importantly, the decision about the Scapegoat's payoff was framed neutrally, without referring to “punishment” or “scapegoat”. Specifically, the Punishers were asked to decide whether to “reduce the Person C's money by X euro by paying X*10 cents”. In order to avoid any scope for instrumental punishment of the Scapegoat, the Punisher knew that the Wrongdoer would not be informed of the decision affecting the Scapegoat. The full experimental protocol is provided in the Supplementary Online Appendix B.

In addition, each Punisher made a decision in an additional task which we refer to as the *Punishing the Wrongdoer Game*. The features of the *Punishing the Wrongdoer Game* were identical with the *Punishing the Scapegoat Game* (e.g., structure of payoffs, cost of punishment, visual design, strategy method), except that the Punishers were asked to decide whether and by how much to reduce Wrongdoer's payoff. The two tasks described above were conducted in random order, and the participants did not know about the existence of the second task until after they finished the first one. Each task was payoff-relevant at 10% probability (but the tasks could never both be payoff-relevant)¹¹, in which case the real decision of the Wrongdoer together with the Punisher's choice in that scenario determined the final payoffs of the Wrongdoer, the Victim, the Scapegoat and the Punisher. In order to avoid aversion to punishment motivated by disadvantageous inequality between the Punisher and Scapegoat, we set the Punisher's endowment at 9 euro and thus even if s/he chose the maximum punishment level, his/her final payoff was 8.20 euro, i.e. higher than the payoff of all other players.

¹¹ To make both tasks incentive-compatible in this way, at the end of the experiment, one ball was picked at random from a bag containing ten balls. Before starting Task 1, the Punishers knew that there was a star on one of the balls and Task 1 would be payoff-relevant if this ball was selected. After learning about the existence of Task 2, subjects were told that there was a triangle on another of the ten balls, and if that ball was selected, Task 2 would be payoff-relevant.

2.4 Manipulating the identity of Scapegoat and Wrongdoer

The Victim is always of the same ethnicity as the Punisher. Importantly, we exogenously manipulate signals of ethnicity of the Scapegoat and of the Wrongdoer, in order to identify how ethnicity affects the decisions of the Punishers, as illustrated in Figure 1. In the Scapegoat SAME condition, the Scapegoat has the same ethnicity as the Punisher, whereas in the Scapegoat OTHER condition, the Scapegoat comes from the other ethnic group. Similarly, in the Wrongdoer SAME condition, the Wrongdoer is of the same ethnicity as the Punisher, while in the Wrongdoer OTHER condition, the Wrongdoer comes from the other ethnic group. The signals of ethnicity of the Scapegoat and of the Wrongdoer are manipulated orthogonally. In this 2×2 “between-subject” design, each Punisher is randomly allocated to one of the four possible combinations of Wrongdoer SAME/OTHER and the Scapegoat SAME/OTHER conditions. Randomization checks indicate that the randomization was successful (Column 7 of Tables A1 and A2).

We use photographs to signal ethnicity. Photographs provide a clear signal of a group attribute, because Roma people (who are of Indian origins and have a darker skin color) are visually distinct from the Slovak majority (white). For each group of workers, we displayed pictures of twenty individuals instead of one, in order to avoid a case in which sympathies/antipathies towards a specific person in a picture might drive decisions. Also, the individuals displayed in the pictures were homogenous in terms of gender and age (all were young male, 18-23 years old), in order to avoid differential treatment based on these attributes.

Since we study differential behavior towards real groups, note that the pictures in SAME and OTHER may signal characteristics other than ethnicity, particularly socioeconomic status. Thus, a natural question is whether our design measures discrimination against other ethnic groups or against poorer or less educated individuals. To address this question, we asked each Punisher, after making their experimental choices, about their perceptions of socioeconomic status (employment status, father’s education) of individuals displayed as Scapegoats and Wrongdoers. We show that our results are robust to controlling for differences in perceptions across the SAME and OTHER conditions.

2.5 Elicitation of wrongdoing

Prior to the punishment experiment (May and early June 2017), we organized the following supplementary work activity among a sample of different individuals, in order to make real the situation a Punisher was confronted with, including the harm committed by the Wrongdoer to the Victim, and also to make the Punisher’s choices consequential.

We recruited workers from the majority group and workers from the Roma ethnic group. Since the experimental design requires six distinct groups of individuals who do not know each other (majority

Wrongdoers, Roma Wrongdoers, majority Victims, Roma Victims, majority Scapegoats and Roma Scapegoats), we recruited workers in six different localities. In each locality, we hired 20-23 workers, all young men, aged 18-23, and a professional photographer took a passport photo of each worker.¹² We selected the photographs of twenty workers and composed a picture consisting of these twenty photos. The pictures were used to signal the ethnicity of the Wrongdoer, the Victim and the Scapegoat to the Punishers in the *Punishing the Scapegoat Game* and in the *Punishing the Wrongdoer Game*, and to signal the ethnicity of the Victim to the Wrongdoers.

Workers performed eight hours of a simple and useful task, aimed at improving the local environment (e.g., cleaning; painting desks; painting a fence; collecting and sorting tree boughs in the woods) and could earn up to 8 euros per hour of work. In the experiment, we treat each hour of work as a separate work assignment. Thus, eight different work hours of the same worker entered the decisions of eight different Punishers. This procedure ensures that each Punisher observes a real situation in which the Wrongdoer, the Victim and the Scapegoat displayed in the pictures worked for one hour and could earn up to 8 euros for this work assignment. The Punishers were not informed that the workers worked for more than one hour.¹³

Upon completion of the work assignment, each Wrongdoer made a decision whether to reduce up to 8 euro of the potential earnings of the Victim. Specifically, they could leave Victim's earnings for an hour of work as they were, or lower them by 2, 4, 6, or 8 euro. Since we wanted to create a strong urge to punish among Punishers, we deliberately elicited a particularly malevolent form of harmful behavior: it reduces the earned income of the Victim and does not create a pecuniary benefit for anyone, including the Wrongdoer.¹⁴ We find that a non-negligible fraction of Wrongdoers chose to do harm (see the histogram in Figure A2), in line with beliefs about prevalence of harming, elicited among Punishers, as described below.

¹² The workers signed a consent form, agreeing to their participation in research and the use of their photograph for the purposes of the research.

¹³ Workers from two localities (one majority and one Roma) were assigned to the position of Victims for all eight work hours. Pictures from these two localities were always used to display a potential Victim. Workers from the remaining four localities were assigned to the position of Wrongdoers for four work hours (and made four decisions whether to reduce earning of a Victim) and to the position of Scapegoats for the remaining four work hours. Pictures from these localities were randomly assigned to display either the potential Wrongdoer or the potential Scapegoat, in order to ensure that specific photographs in each of the pictures do not drive the differences between the punishment of the Wrongdoer and the Scapegoat.

¹⁴ Note that in the standard Third Party Punishment game, the players receive resources as “manna from heaven”, while in our experiment, inspired by recent experiments on redistribution of earned income (Cappelen et al. 2007; Almas, Cappelen, and Tungodden 2020), we add the work stage, which creates a sense of entitlement, after which the Wrongdoer destroyed earned resources of the Victim. Also, while prior experiments study punishment of violators of norms that govern the positive side of social behavior (Goette, Huffman, and Meier 2006; Goette et al. 2012), we focus on the responses of Punishers to unambiguously nasty behavior—destruction of the Victim's earnings.

This finding is also consistent with the results of recent studies that show that anti-social behavior is not an anomaly, when tasks are implemented in which the decision-maker has to sacrifice some of his/her own money to reduce the other's payoff and where this destruction cannot be justified by standard fairness considerations.¹⁵

The Wrongdoers knew that their decisions would have real payoff consequences, and would be anonymous: the Victim would be given the remaining amount without any additional information. They also knew that the Victim could not reduce their (the Wrongdoer's) wage. Further, they were informed (as all other workers) that their own earnings could be affected by decisions of other people but did not know any details about the Punisher's task.

Given the expected destruction rates among the Wrongdoers and punishment rates among the Punishers, we could guarantee each worker the payment of 20 euros shortly after they finished the work. This amount is equivalent to the minimum wage in Slovakia at the time of the data collection (2.50 euro per hour). The workers were further informed that they would receive up to 44 additional euro in three months, based on the decisions of other people. Thus, the maximum payment they could receive was 64 euro, equivalent to 8 euro per hour of work. The average amount paid to workers was 45.3 euros and the minimum amount was 40 euros.

2.6 Procedures

We paid particular attention to maximize a correct understanding of the tasks. First, the experimenters explained the instructions one-on-one.¹⁶ Second, the working/wrongdoing stage of the experiment and the decisions were explained in detail using a simple tablet interface with the photos of Wrongdoers, Victims, and Scapegoats (Figure A1).¹⁷ Third, before making decisions, the Punisher had to answer six comprehension questions, demonstrating understanding of the roles of the three types of workers. If any of the answers were not correct, the experimenter explained the whole setup once again, and the

¹⁵ Among adult non-student samples, the proportion of harmful choices was 20% in Kenya (Bauer, Chytilová, and Miguel 2020), 19% in Slovakia (Bauer et al. 2021), 30% in India (Fehr, Hoff, and Kshetramade 2008) and 23% in Namibia (Prediger, Volland, and Herrmann 2014). It is also worth noting that subjects in the supplementary activity in our study were young male from socially and economically disadvantaged groups, and the existing evidence suggests that people with low SES are more prone to engage in anti-social behavior (Bauer, Chytilová, and Pertold-Gebicka 2014; Prediger, Volland, and Herrmann 2014; Bauer et al. 2021).

¹⁶ On a typical day, three experimenters implemented the experiments in one location. Each subject was randomly allocated to one of the experimenters. All (ten) experimenters were from the majority ethnic group, and they were randomly allocated to conditions within each location. The results are robust to controlling for experimenter fixed effects.

¹⁷ The app was programmed using z-Tree (Fischbacher 2007).

comprehension questions were asked a second time. As a robustness check, we exclude from the sample individuals who did not provide correct answers when the questions were asked for the second time.

To avoid perception of social pressure, the experimenters could not observe the subjects' decisions. For each possible action of the Wrongdoer, the experimenter described the situation, and then gave the subject privacy to anonymously make decisions on the tablet computer whether and how much to lower the payoffs of the Scapegoat (or the Wrongdoer). The experiments were conducted in private; other participants could not hear the instructions or observe the decisions that the subjects made on the tablets.

After completing the main tasks, we elicited beliefs about the action of the Wrongdoer, asking the Punishers to guess which of the five options the Wrongdoer chose (subjects were rewarded 1 euro for a correct guess). We also elicited beliefs in more detail, asking how many of the twenty potential Wrongdoers they thought chose each action, to gauge beliefs about the distribution of responses.¹⁸ A short post-experiment questionnaire followed. To elicit the perceptions of the Wrongdoer's and Scapegoat's social status, the Punishers were asked to guess their employment status (student/employed/unemployed) and the educational level of their father. Subjects also filled in a short questionnaire on their own demographics (gender, education, family size, marital status, parental education and employment, religiosity, and household amenities).

To determine final payoffs, at the end of the experiment, each Punisher chose one ball out of ten at random from a bag, which determined whether their decision in the *Punishing the Scapegoat Game* and in the *Punishing the Wrongdoer Game* was payoff-relevant or not. An actual decision of an anonymously matched Wrongdoer determined which decision situation in a given task was payoff-relevant (out of five).

3. Results

In the main analysis presented in Section 3.1, we focus on choices of decision-makers from the majority group in the *Punishing the Scapegoat Game*. In Section 3.2, we analyze choices in the *Punishing the Wrongdoer Game*. In Section 3.3, we provide additional results, including the analysis of decision-makers from the Roma minority, and discuss how our findings relate to minority scapegoating and other existing theories of discrimination (collective responsibility, statistical discrimination, and taste-based discrimination).

¹⁸ We did not elicit the more detailed beliefs for the Roma minority Punishers, because the explanation was too time-consuming and a large percentage of subjects did not understand this aspect of the experiment properly during the pilot.

3.1 Punishment of the Scapegoat

In this sub-section, we explore whether the initial harm committed by the Wrongdoer gives rise to harmful behavior towards the Scapegoat and whether the identity of the Scapegoat matters. Figure 2 displays the results for the punishment of the Scapegoat across the five specific amounts of Victim's earnings that the Wrongdoer could decide to destroy (0, 2, 4, 6 and 8 euros). Panel A shows the average amount of euros that the Punishers decided to destroy in the SAME condition, when the Scapegoat is also from the majority population, and in the OTHER condition, when the Scapegoat is a member of the Roma ethnic minority. Panel B depicts the proportion of the Punishers who decide to punish in each case. Table 1 presents the results of the regression analysis.

[Figure 2 around here]

We find several interesting patterns. First, a substantial fraction of Punishers responds to observing the Wrongdoer's misbehavior by lowering the payoff of the Scapegoat (Panel B of Figure 2) and the Punishers are sensitive to the intensity of the harm done by the Wrongdoer to the Victim. Only 4% of subjects reduce the payoff of the Scapegoat when the Wrongdoer did zero harm, but the number increases to 24% for the situations when the Wrongdoer destroyed all 8 euro of Victim's earnings. 23% of subjects reduce the payoff of the Scapegoat by a larger amount in a situation when maximum harm is committed, as compared to when no harm is done by the Wrongdoer. In terms of the size of the punishment, when the Wrongdoer does no harm, we observe only a small reduction of the Scapegoat's payoff (0.12 euro on average), whereas in a situation with maximum harm, it increases to 1.10 euro (Panel A of Figure 2).

[Table 1 around here]

Next, we study whether the ethnic identity of the Scapegoat matters. In circumstances when no harm is done by the Wrongdoer, there is no evidence of discrimination against the Roma ethnic minority—on average their payoff is lowered by 0.11 euro, while the payoff of the Scapegoat from the majority population is lowered by an average of 0.13 euro. Importantly, when analyzing situations in which the Wrongdoer harmed the Victim, we find a systematic difference in responses between the Scapegoat SAME and Scapegoat OTHER conditions—punishment of Scapegoats is twice as severe when the Scapegoat is from the Roma minority than when the Scapegoat is from the majority population. Specifically, in Scapegoat SAME, an increase in harm intensity by one additional euro motivates Punishers to lower the Scapegoat's earnings by an additional 0.08 euro. In Scapegoat OTHER, the effect doubles to 0.16 euro, and the difference between SAME and OTHER is statistically significant at the 1% level ($p=0.002$, Column 1 of Table 1). Due to such magnified punishment of the Scapegoat in OTHER, discrimination against the

ethnic minority gradually rises with greater harm intensity and becomes statistically significant for situations when the Wrongdoer destroyed 4, 6 or 8 euros of the Victim's earnings (Figure 3, p-values = 0.008, 0.002, 0.011, respectively).

[Figure 3 around here]

The interaction effect of the Wrongdoer's level of harm committed and OTHER on punishment of the Scapegoat is driven by the extensive as well as the intensive margin (Columns 2 and 3 of Table 1), and in both cases it is statistically significant at the 5% level. Each additional euro destroyed by the Wrongdoer leads to an increase in the proportion of those who decide to punish the Scapegoat, by 2 percentage points in Scapegoat SAME and by 3 percentage points in Scapegoat OTHER. Among those who decide to scapegoat, the amount destroyed from the Scapegoat's earnings increases by 0.20 euro in Scapegoat SAME and by 0.39 euro in Scapegoat OTHER.

The results are robust to controlling for various additional variables (Table A3). Columns 1-7 document that the coefficients for harm intensity and for the interaction between harm intensity and OTHER hardly change when we control for design features, including experimenter fixed effects, the subject's observable characteristics, location fixed effects and parental education level. The coefficients remain statistically significant at the 1% level, and the magnitude stays the same. The results are similar if we use a non-linear specification for the harm caused by the Wrongdoer (Columns 1-2 of Table A4).

The use of the strategy method to elicit choices for various possible actions of the Wrongdoer has the advantage of providing a rich picture of the Punisher's behavior. On the other side, this approach may, in principle, induce Punishers towards greater differentiation in behavior in different situations and thus lead to greater observed sensitivity of Punishers to intensity of the harm done by the Wrongdoer. Importantly, however, our main focus is on estimating the *differences* in the Punisher's sensitivity to harm intensity across the Scapegoat's ethnicity (SAME vs. OTHER). Since the decision environment is the same across these conditions, there is no obvious reason why any experimenter demand effects should matter more in OTHER than in SAME. This concern is further alleviated by the fact that we implemented the SAME and OTHER conditions using a between-subject design, and thus subjects could not be induced to differentiate the punishment of Scapegoats of the majority vs. minority ethnicity.

In addition, many existing experiments studied differences in choices when the strategy method versus the direct-response method is used. Jordan, McAuliffe, and Rand (2016) focus specifically on the Third Party Punishment Game and find that the use of the strategy method does not influence punishment decisions. Brandts and Charness (2011) provide an overview of 29 studies focusing on various experimental tasks. They find that although the use of the strategy method affected the levels of behavior in some of the

tasks, it is to be noted that in all of the experiments, a treatment effect identified with the strategy method was also observed with the direct-response method. Thus, we believe it is unlikely that any experimenter demand effect could explain the main pattern we observe -- magnified punishment of the Scapegoat in OTHER as compared to SAME.

Finally, we study heterogeneity and explore whether the observed biases in the punishment of the Scapegoat are driven by an easily identifiable subgroup of individuals (Table A5). We focus on three characteristics of majority Punishers, for which there is enough variation in our sample, and divide the sample based on gender, student status (secondary school vs. university students) and parental education (as a proxy for socio-economic status). The results are qualitatively similar across parental education and, interestingly, they also hold for university students. This suggests that even the future ‘elites’ are not immune to group biases in punishment. In terms of gender differences, the interaction effect is primarily driven by men.

To summarize, we find that members of the minority group receive systematically harsher punishment than members of the majority group for actions of unrelated individuals. In subsection 3.3, we present additional results and consider potential alternative explanations.

3.2 Punishment of the Wrongdoer

In terms of direct punishment of the Wrongdoer, we also find that the Punishers are sensitive to the intensity of harm done. The greater the harm caused by the Wrongdoer, the stronger the punishment (Figure A3). The punishment response is stronger, approximately around five times, when the subjects can directly punish the Wrongdoer, as compared to punishment of the Scapegoat.¹⁹

When no harm is done by the Wrongdoer, we do not detect any discrimination against the Roma minority. Importantly, the sensitivity to harm intensity is again systematically larger in OTHER as compared to SAME, both in terms of the extent of the punishment (Panel A of Figure A3) and of the likelihood that subjects punish the Wrongdoer (Panel B). In the regression analysis (Table A6), the coefficients for an interaction term between harm intensity and OTHER are statistically significant at the 1% level, both for the extensive and intensive margins. The results are robust to using various specifications and control variables, including order of the task, experimenter and location fixed effects (Table A7) and

¹⁹ As expected, punishment of the Scapegoat and punishment of the Wrongdoer are strongly positively correlated. Specifically, for the maximum harm committed by the Wrongdoer, 21% of Punishers do not punish anyone, 54% punish only the Wrongdoer, 23% punish both the Wrongdoer and the Scapegoat, while only 2% punish only the Scapegoat.

are qualitatively similar for different sub-samples based on gender, student status and parental education (Table A8).

Finally, we show that the magnified revenge in OTHER as compared to SAME gives rise to discrimination against the Roma minority (Figure A4). When the harm is large (4, 6 or 8 euros), Wrongdoers from the minority group are punished more harshly, by 13-20%, than Wrongdoers from the decision-makers' own group (p-values = 0.051, 0.095, 0.013 respectively). In the situation in which the Wrongdoer committed maximum harm (destroyed all 8 euros), the punishment gap is the most profound (4.75 euros in SAME vs. 5.69 euros in OTHER, on average). To sum up, when no harm is done to a member of the majority group, the decision-makers from the majority group do not discriminate against the ethnic minority. At the same time, Wrongdoers from the Roma ethnic minority are punished more severely than Wrongdoers from the majority population, for the same harmful actions.

3.3 Mechanisms behind the shifting of punishment on minorities

In this sub-section, we discuss and present additional results that speak to the underlying mechanisms of the main effects. We argue that the results support the interpretation that the observed greater inclination to shift punishment on the Roma minority originates in latent discriminatory preferences. We also show that the main effects are unlikely to be driven by statistical discrimination, over-attribution of responsibility, or due to "punishment by association" (collective punishment). We also consider whether the effects can be explained by differences in perceptions about the socio-economic status of Roma and non-Roma. Finally, we discuss whether the shifting of punishment on an out-group is specific to behavior of a dominant group towards ethnic minorities, or whether it extends to any interactions across ethnic boundaries.

3.3.1 Why is shifting of punishment larger when the Roma minority is a Scapegoat?

Statistical discrimination and over-attribution of responsibility. By virtue of the experimental design, beliefs and statistical discrimination are unlikely to explain our findings. First, subjects were randomly allocated into SAME and OTHER conditions and were provided the same information about the nature and the extent of the harm before making their punishment decisions. This essentially rules out the possibility that greater punishment of Roma minority members in our experiment can be explained by subjects thinking that Roma Wrongdoers committed more serious harm than Wrongdoers from the majority group. Also, we directly elicited beliefs about the destructiveness of Wrongdoers in SAME and in OTHER conditions, and

do not find evidence supporting the idea that subjects believed Roma Wrongdoers to be more likely to harm Victims than majority Wrongdoers.²⁰

Moreover, the observed bias in punishment is unlikely to be driven by differences in beliefs about future interactions – for example, by the possibility that Punishers would punish the majority Scapegoats (or Wrongdoers) less due to greater fear that they would in the future face revenge from them, either within or outside of the (field) lab. Punishers knew Scapegoats and Wrongdoers would not have any opportunity to take revenge after their punishment decision, the interactions were one-shot and anonymous, with Scapegoats and Wrongdoers coming from different locations.

Finally, Punishers faced no uncertainty about who was responsible for the wrongdoing, since the experimental protocol made it clear that Wrongdoers caused the harm to the Victim, while Scapegoats did not. Thus, shifting of the punishment in our experiment is unlikely to arise because of over-attribution of responsibility, and subsequent blaming of minority groups for injustices done to the majority group members. While these can be relevant drivers of punishment when the source of harm is unclear, they are ruled out here by experimental design.

Collective responsibility. A plausible mechanism for why people may punish innocent individuals from other ethnic groups is the notion of collective responsibility. “Punishment by association” is described as retaliation directed not only against the wrongdoer, but also against other members of his group who have no direct association with the perpetrator or direct control over his actions (Lickel, Schmader, and Hamilton 2003; Cushman, Durwin, and Lively 2012). Thus, in our experiment, collective responsibility would predict that punishment of minority Scapegoats is triggered specifically in a situation where the Wrongdoer was also from the minority group. To test this, we take advantage of the orthogonal experimental variation of identity of the Wrongdoer and of the Scapegoat.

Figure A5 displays the results graphically for all four possible combinations of ethnic identity of the Wrongdoer and of the Scapegoat, where the outcome of interest is the punishment of the Scapegoat. In all four situations, subjects respond to greater harm done with greater punishment of the Scapegoat. This response is, however, not stronger in situations when the Scapegoat has a shared ethnicity with the

²⁰ Panel A of Table A9 shows the distributions of beliefs, which are not significantly different for the Wrongdoer from the majority population and for the Wrongdoer from the Roma ethnic minority. Most subjects believe that the Wrongdoer did not commit any harm (61% in SAME and 64% in OTHER). Yet, a non-negligible number believe that the Wrongdoer destroyed some part of the earnings of the Victim and 6% of participants in SAME and 8% in OTHER believe that maximum harm was committed. We arrive to the same conclusion with our second, more detailed, measure of beliefs (Panel B).

Wrongdoer (solid lines), providing initial indication that punishment by association does not drive our results. Further, in a regression analysis, we restrict the sample to subjects who were informed that the Wrongdoer was from the majority group. Thus, magnified punishment of the minority Scapegoat cannot be explained by the desire to punish collectively based on the shared ethnicity of Wrongdoer and Scapegoat. Yet, we still find a strong bias in punishment, as indicated by the positive and statistically significant coefficient for the interaction term between harm intensity inflicted by the majority Wrongdoer and the Scapegoat OTHER condition ($p=0.006$, Column 1 of Table A10).²¹ In addition, in Column 4, we hold the Roma identity of the Scapegoat constant and estimate whether scapegoating becomes stronger in a situation where the Wrongdoer is also from the Roma minority, as compared to when the Wrongdoer is from the majority group. We do not find evidence for this—when Scapegoats are OTHER, the coefficient for the interaction term between harm intensity and Wrongdoer OTHER is in fact negative and not statistically significant.²²

Altogether, this set of results does not support the interpretation that the observed magnified punishment of Scapegoats from the Roma minority is driven by collective responsibility.²³ In fact, we find that the minority Scapegoats are more harshly punished regardless of the ethnic identity of the Wrongdoer. In other words, the minority is punished for any injustice done to the member of the majority group, *including those that originate from within the majority group*.

Latent discriminatory preferences. Plain unconditional outgroup hate, a strong form of taste-based discrimination, implies that decision-makers may treat the ethnic minority systematically more harshly than members of their own group, independently of the social context. Thus, differential treatment of the discriminated group should be relatively stable. Yet, we find that this is not the case, since we observe that

²¹ This result also indicates that our findings on differential treatment of minority Scapegoat is unlikely to be explained by pure conformism – Punishers simply copying the behavior of Wrongdoers. Assuming that conformism increases with a similarity of the decision, this explanation would imply larger punishment of majority as compared to minority Scapegoats, if both the Wrongdoer and Victim are from the majority group. This is an opposite pattern to what we find.

²² In Column 3 of Table A10, we study the role of collective responsibility for the Scapegoat from the majority population. In this case, we also find no evidence that collective responsibility plays a role, since the sensitivity to harm intensity is not greater if the Wrongdoer is of the same ethnicity as the Scapegoat, i.e. from the majority population.

²³ Note that, by design, we focus on the pure form of enforcement of collective responsibility, but at the same time we close some of the mechanisms behind collective punishment that might be relevant in real life, in particular the role of instrumental punishment. In our experiment, the Punisher knows that the Scapegoat will not be informed about the Wrongdoer's action (and thus what action triggered his punishment), and that the Wrongdoer will not be informed about any punishment, and thus the Punisher cannot send a moral message to the Wrongdoer by punishing the Scapegoat.

when there is no harm committed against a Victim, and thus no injustice, there is no evidence of discrimination against the minority. Discrimination against the Roma minority is *activated* when subjects observe harm committed against a member of their own group and have an opportunity to punish an unrelated party or the Wrongdoer, suggesting that people harbor latent discriminatory preferences that are manifested only in some situations.

Several features of the decision environment associated with making decisions after observing harm committed against a member from the own group make it psychologically easier for people who harbor latent discriminatory preference to act on them. First, discriminatory urges may be controlled at normal emotional states, while witnessing blatant injustice may trigger anger and reduce such self-control. Observing conflict and unethical behavior can reduce self-image costs from acting based on a latent animus. Finally, reduced earnings of the Victim from the own group can provide an excuse for destroying income of the ethnic minority member, justified by not letting the Victim lag behind the minority Scapegoat in terms of relative income. Empirically separating between these plausible channels that may activate latent discriminatory preferences and lead to minority scapegoating is beyond the scope of this paper.

3.3.2 Group attribute

We first aim to gauge whether the observed bias in punishment can be explained by differences in perceived socio-economic status, rather than ethnic minority status per se, since the pictures in SAME and OTHER may signal not only ethnicity, but also socioeconomic status (SES). To do so, we elicited subjects' perceptions of the socio-economic status of the individuals in the photos, after completing their experimental choices. Specifically, the participants reported their estimates of the employment status and the father's education of the pictured individuals. As expected, the perceptions of individuals in SAME and OTHER indeed differ widely (Table A11). People in photos signaling Roma minority are significantly more likely to be perceived as unemployed, less likely to be students, and less likely to have educated fathers.

We test whether the coefficient for the interaction effect between harm intensity and OTHER is robust to controlling for perceptions of unemployment and its interaction with harm intensity, and to controlling for perceptions of the father's education and its interaction with extent of harm inflicted. Table A12 shows the results for punishment of the Scapegoat and Table A13 for punishment of the Wrongdoer. In both tasks, we find that the estimated interaction effects between harm intensity and OTHER on punishment are similar to the baseline estimates. Also, we observe that Punishers who believe that OTHER Scapegoats or Wrongdoers are from a low SES background are not more likely to punish, compared to Punishers who believe Scapegoats or Wrongdoers are not from low SES, once we control for SAME/OTHER conditions.

These patterns suggest that differences in perceptions of socio-economic status, which are associated with SAME and OTHER conditions, cannot fully explain our findings. That said, note that the allocation to OTHER and SAME conditions is exogenous since it was randomly allocated, while the perceptions of parental education and employment status of the people in the photos are endogenous, and thus this analysis needs to be taken with a grain of salt.

Next, the literature on scapegoating suggests that such behavior is specific to decision-makers from the dominant group when interacting with members of negatively stereotyped, vulnerable and smaller groups (Allport 1954; Bettelheim and Janowitz 1950), rather than vice versa. To test this idea, we measured the behavior of Roma punishers in the same experimental set up and compared their punishment behavior with punishers from the majority group. Indeed, we find weaker evidence of biases in punishments by the minority group against the majority group, as compared to biases of the majority group against the minority group.

Specifically, we find that Roma subjects punish the Scapegoat, and the extent of scapegoating increases with harm intensity. Importantly, when we compare choices in Scapegoat SAME and Scapegoat OTHER, we do not find evidence of co-ethnic bias in any of the five situations with different harm intensity (Figure A6). Also, in the regression analysis, the coefficient for the interaction between harm intensity and Scapegoat OTHER is small in magnitude and not statistically significant ($p=0.737$, Column 1 of Table A14). In terms of direct punishment of the Wrongdoer, we find a weaker but qualitatively similar pattern as we did for the majority population. When no harm is done by the Wrongdoer, we find no evidence of discrimination against the majority population, but harmful actions of the Wrongdoer against a Roma subject trigger magnified revenge towards OTHER (majority) Wrongdoers as compared to SAME (Roma) Wrongdoers. The coefficient for an interaction term between the harm intensity and OTHER is statistically significant at the 10% level ($p=0.097$, Column 1 of Table A15).²⁴ Consequently, a gap in the punishment of majority vs. Roma Wrongdoers opens up with greater harm intensity and becomes marginally statistically significant for maximum harm ($p\text{-value}=0.099$, Figure A7). Note that among Roma Punishers, incomplete understanding is a potential concern, since only 77% of the subjects answered all comprehension questions

²⁴ While the patterns are qualitatively similar, the strength of the effects and the extent of punishment differ across the samples of decision-makers from the Roma and from the ethnic majority population. First, the extent of punishment in a situation where no harm is committed by the Wrongdoer is higher among Roma than among ethnic majority decision-makers (0.91 vs. 0.24 euro, respectively). Second, the extent of punishment in the situation of maximum harm is lower among Roma than majority decision-makers (2.71 vs. 5.23 euro, respectively). Thus, Roma decision-makers were less sensitive to harm intensity. Specifically, in the SAME condition, each additional euro destroyed by the Wrongdoer leads to an increase in punishment by 0.19 euro by Roma and by 0.56 euro by ethnic majority decision-makers.

correctly. It is reassuring that when we exclude the 23% of subjects who demonstrated imperfect understanding, the estimates are similar to the original results (Column 5 of Tables A14 and A15).

4. Conclusion

Social scientists have long speculated that members of dominant social groups tend to shift punishment for social ills originating within their own group onto innocent members of other, weaker groups. This behavior, termed scapegoating, is sometimes considered to be an important psychological mechanism in the emergence of pogroms, witch-hunts, and large-scale violence. However, anecdotal and historical evidence cannot rule out the role of standard economic motives in harming innocent members of minority groups, and an experimental test of this behavioral phenomenon in controlled environments had been missing so far. This is what we provided in this paper. We have developed a new incentivized task, the *Punishing the Scapegoat Game*, to uncover how observing harmful actions against members of one's own group shapes the punishment of innocent individuals. We study the behavior of young adults in Eastern Slovakia, a region experiencing intergroup tensions. In line with the idea that minorities are often treated as scapegoats by dominant groups, we show that Punishers from the ethnic majority group systematically redirect punishment onto innocent members of the Roma minority group for harmful actions committed by other people. Further, when they can punish actual Wrongdoers, they punish Wrongdoers from the minority group more harshly than Wrongdoers from the majority group, for the same wrongdoing. Importantly, the observed out-group bias in punishment happens above and beyond out-group bias in circumstances when people do not respond to harm done to their own group.

This paper leaves several questions open for future research. We study interactions between one majority group and one minority group (Roma people), in one setting in which the minority is largely segregated, economically disadvantaged, and ostracized. Although anecdotal evidence suggests that scapegoating is a relatively broad phenomenon, clearly, more research is needed to assess whether the patterns we observe are generalizable to other settings, and also to pin down which specific group attributes make certain individuals or groups convenient scapegoats. In this paper, we also study the influence of observing social ills—someone doing harm to a member of one's own group for no apparent reason. Although it is beyond the scope of our study, an interesting direction is to explore whether similar aggressive discriminatory responses are activated when experiencing personal ills, such as stress or an important income shock. We believe the experimental paradigm illustrated in this paper is a well-suited tool to address some of these questions.

At face value, our findings have several potentially important implications. First, we show that pure observation of injustice and wrongdoing against the individual's own group activates latent discriminatory preferences, both when treating innocent individuals as well as wrongdoers. This indicates that courts, and other settings in which people make punishment choices, are particularly discrimination-prone environments, in line with evidence of strong biases against minorities in judicial decisions (Alesina and Ferrara 2014; Rehavi and Starr 2014; Shayo and Zussman 2011). Second, the results suggest that ethnic minorities are at greater risk when social problems and unfair behavior become salient features of the societal environment. So far, economists have typically attributed sudden spikes in aggressive behavior towards weaker groups to changes in economic incentives or beliefs about the likelihood of facing a penalty for aggressive behavior (Blattman and Miguel 2010; Grosfeld, Sakalli, and Zhuravskaya 2020; Miguel 2005), assuming that revealed (anti-) social preferences towards other groups are stable. In our experiment, economic incentives are held constant and thus cannot explain the scapegoating behavior observed. Of course, this does not imply that economic incentives do not play an important role in real-life aggression towards minority groups. However, our evidence suggests that may not be the complete picture. It strengthens the case for taking seriously "behavioral" channels through which deterioration of the social environment may fuel inter-group conflicts.

Acknowledgements

We thank Alexander Cappelen, Ernst Fehr, Michael Kosfeld, Filip Matějka, Matthias Sutter, Bertil Tungodden, seminar participants at the University of Munich, Max Planck Institute in Bonn, UC Berkeley, University of Edinburgh, Goethe University in Frankfurt, Norwegian School of Economics, and participants at the ECBE conference 2020, EEA Virtual Congress 2020, Symposium for Social Cohesion in Ethnically Diverse Societies in Tilburg, and ERINN Annual Conference 2021 for many helpful comments. The data collection was supported by a grant from the Czech Science Foundation (17-13869S), and by the Slovak Research and Development Agency (APVV-0125-12 and APVV-19-0329). Bauer and Chytilová acknowledge support from ERC-CZ/AV-B (ERC300851901).

The research was approved by the Ethical committee at the Technical University of Košice (12/7/2016) and by the Director of the Institute of Economic Studies at the Faculty of Social Sciences, Charles University.

References

- Abbink, Klaus, Jordi Brandts, Benedikt Herrmann, and Henrik Orzen. 2010. "Intergroup Conflict and Intra-Group Punishment in an Experimental Contest Game." *American Economic Review* 100 (1): 420–47.
- Alesina, Alberto, and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." *American Economic Review* 104 (11): 3397–3433.
- Allport, Gordon W. 1954. *The Nature of Prejudice*. New York: Addison-Wesley.
- Almas, Ingvid, Alexander Cappelen, and Bertil Tungodden. 2020. "Cutthroat Capitalism versus Cuddly Socialism : Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5): 1753–88.
- Anderson, Robert Warren, Noel D. Johnson, and Mark Koyama. 2017. "Jewish Persecutions and Weather Shocks: 1100–1800." *Economic Journal* 127 (602): 924–58.
- Bardhan, Pranab. 2005. *Scarcity, Conflict and Cooperation*. Cambridge, MA: The MIT Press.
- Bartling, Björn, and Urs Fischbacher. 2012. "Shifting the Blame : On Delegation and Responsibility." *Review of Economic Studies* 79 (1): 67–87.
- Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka. 2016. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *American Economic Review* 106 (6): 1437–75.
- Bauer, Michal, Jana Cahlíková, Dagmara Celik-Katreniak, Julie Chytilová, Lubomír Cingl, and Tomáš Želinský. 2021. "Nastiness in Groups." Mimeo.
- Bauer, Michal, Jana Cahlíková, Julie Chytilová, and Tomáš Želinský. 2018. "Social Contagion of Ethnic Hostility." *Proceedings of the National Academy of Sciences* 115 (19): 4881–86.
- Bauer, Michal, Julie Chytilová, and Edward Miguel. 2020. "Using Survey Questions to Measure Preferences: Lessons from an Experimental Validation in Kenya." *European Economic Review*, 103493.
- Bauer, Michal, Julie Chytilová, and Barbara Pertold-Gebicka. 2014. "Parental Background and Other-Regarding Preferences in Children." *Experimental Economics* 17 (1): 24–46.
- Baumgartner, Thomas, Lorenz Goette, Rahel Guegler, and Ernst Fehr. 2012. "The Mentalizing Network Orchestrates the Impact of Parochial Altruism on Social Norm Enforcement." *Human Brain Mapping* 33 (6): 1452–69.
- Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126 (2): 805–55.
- Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Simon Galle, Edward Miguel, Daniel Posner, Bertil Tungodden, and Kelly Zhang. 2020. "Ethnically Biased? Experimental Evidence from Kenya." *Journal of the European Economic Association* 18 (1): 134–64.
- Bernhard, Helen, Urs Fischbacher, and Ernst Fehr. 2006. "Parochial Altruism in Humans." *Nature* 442 (7105): 912–915.
- Bettelheim, B., and M. Janowitz. 1950. *Dynamics of Prejudice: A Sociological Study of Veterans*. New York: Harper and Brothers.

- Blattman, Christopher, and Edward Miguel. 2010. "Civil War." *Journal of Economic Literature* 48 (1): 3–57.
- Brandts, Jordi, and Gary Charness. 2011. "The Strategy versus the Direct-Response Method: A First Survey of Experimental Comparisons." *Experimental Economics* 14 (3): 375–98.
- Cappelen, Alexander W, Astri Drange Hole, Erik Sorensen, and Bertil Tungodden. 2007. "The Pluralism of Fairness Ideals : An Experimental Approach." *American Economic Review* 97 (3): 818–27.
- Chen, Yan, and Sherry Xi Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99 (1): 431–457.
- Council of Europe. 2012. *Human Rights of Roma and Travellers in Europe*. Strasbourg: Council of Europe Publishing.
- Crossman, Ashley. 2019. "Definition of Scapegoat, Scapegoating, and Scapegoat Theory." *ThoughtCo.*, 2019.
- Cushman, Fiery, A. J. Durwin, and Chaz Lively. 2012. "Revenge without Responsibility? Judgments about Collective Punishment in Baseball." *Journal of Experimental Social Psychology* 48 (5): 1106–10.
- Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2006. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory* 33 (1): 67–80.
- Doob, Leonard W., Neal E. Miller, Orval Hobart Mowrer, Robert Sears, and John Dollard. 1939. *Frustration and Aggression*. New Haven, CT: Yale University Press.
- Esteban, Joan-Maria, and Debraj Ray. 2008. "On the Salience of Ethnic Conflict." *American Economic Review* 98 (5): 2185–2202.
- Falk, Armin, Thomas Neuber, and Nora Szech. 2020. "Diffusion of Being Pivotal and Immoral Outcomes." *Review of Economic Studies* 87 (5): 2205–29.
- Falk, Armin, and Christian Zehnder. 2013. "A City-Wide Experiment on Trust Discrimination." *Journal of Public Economics* 100: 15–27.
- Fearon, James D., and David D. Laitin. 2000. "Violence and the Social Construction of Ethnic Identity." *International Organization* 54 (4): 845–77.
- Fehr, Ernst, and Urs Fischbacher. 2004. "Third-Party Punishment and Social Norms." *Evolution and Human Behavior* 25 (2): 63–87.
- Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415 (6868): 137–40.
- Fehr, Ernst, Karla Hoff, and Mayursh Kshetramade. 2008. "Spite and Development." *American Economic Review: Papers and Proceedings* 98 (2): 494–499.
- Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economics* 116 (1): 351–377.
- Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2): 171–77.
- Gächter, Simon, Elke Renner, and Martin Sefton. 2008. "The Long-Run Benefits of Punishment." *Science* 322 (5907): 1510–1510.
- Girard, Rene. 1979. *Things Hidden Since the Foundation Of the World*. Stanford, CA: Stanford University

Press.

- Glick, Peter. 2005. "Choice of Scapegoats." In *On the Nature of Prejudice*, edited by John F Dovidio, Peter Glick, and Laurie Budman, 244–61. Blackwell Publishing.
- Goette, Lorenz, David Huffman, and Stephan Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement." *American Economic Review* 96 (2): 212–16.
- Goette, Lorenz, David Huffman, Stephan Meier, and Matthias Sutter. 2012. "Competition Between Organizational Groups: Its Impact on Altruistic and Antisocial Motivations." *Management Science* 58 (5): 948–60.
- Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya. 2020. "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire." *Review of Economic Studies* 87 (1): 289–342.
- Henrich, J., S.J. Heine, A. Norenzayan, and others. 2010. "The Weirdest People in the World." *Behavioral and Brain Sciences* 33 (2–3): 61–83.
- Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, and Edwina Gwako. 2010. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment." *Science* 1480 (5972): 1480–84.
- Herrmann, Benedikt, Christian Thoni, and Simon Gächter. 2008. "Antisocial Punishment across Societies." *Science* 319 (5868): 1362–67.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. London: University of California Press.
- Jordan, Jillian, Katherine McAuliffe, and David Rand. 2016. "The Effects of Endowment Size and Strategy Method on Third Party Punishment." *Experimental Economics* 19 (4): 741–63.
- Kant, Immanuel. 1965. *The Metaphysical Elements of Justice*. New York: Bobbs-Merrill Company.
- Kranton, Rachel, Matthew Pease, Seth Sanders, and Scott Huettel. 2020. "Deconstructing Bias in Social Preferences Reveals Groupy and Not-Groupy Behavior." *Proceedings of the National Academy of Sciences* 117 (35): 21185–93.
- Levin, Jack, and Jack McDevitt. 2002. *Hate Crimes Revisited: America's War Against Those Who Are Different*. Basic Books.
- Lickel, Brian, Toni Schmader, and David L. Hamilton. 2003. "A Case of Collective Responsibility: Who Else Was to Blame for the Columbine High School Shootings?" *Personality and Social Psychology Bulletin* 29 (2): 194–204.
- Miguel, Edward. 2005. "Poverty and Witch Killing." *Review of Economic Studies* 72 (4): 1153–72.
- Newman, Leonard S., and Tracy L. Caldwell. 2005. "Allport's 'Living Inkblots': The Role of Defensive Projection in Stereotyping and Prejudice." In *On the Nature of Prejudice: Fifty Years after Allport*, edited by John F. Dovidio, Peter Glick, and Laurie Rudman, 377–93. Oxford: Blackwell Publishing.
- openDemocracy. 2021. "The Shameful Resurgence of Violent Scapegoating in a Time of Crisis." <https://www.opendemocracy.net/en/can-europe-make-it/shameful-resurgence-violent-scapegoating-time-crisis/>.
- Prediger, Sebastian, Björn Vollan, and Benedikt Herrmann. 2014. "Resource Scarcity, Spite and Cooperation." *Journal of Public Economics* 119: 1–9.

- Rehavi, M. Marit, and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Charging and Its Sentencing Consequences." *Journal of Political Economy* 122 (6): 1320–54.
- Schiller, Bastian, Thomas Baumgartner, and Daria Knoch. 2014. "Intergroup Bias in Third-Party Punishment Stems from Both Ingroup Favoritism and Outgroup Discrimination." *Evolution and Human Behavior* 35 (3): 383–93.
- Shayo, Moses, and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics* 126 (3): 1447–84.
- Staub, Ervin. 1992. *The Roots of Evil: The Origins of Genocide and Other Group Violence*. Cambridge, MA: Cambridge University Press.
- Stenner, Karren. 2005. *The Authoritarian Dynamic*. New York: Cambridge University Press.
- Stenner, Karren, and Jonathan Haidt. 2018. "Authoritarianism Is Not a Momentary Madness, but an Eternal Dynamic with Liberal Democracies." In *Can It Happen Here? Authoritarianism in America*, edited by Cass R. Sunstein, 174–93. New York: HarperCollins.
- Tajfel, Henry. 1981. *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge: Cambridge University Press.
- UN News. 2021. "Hate Speech and Violence Still Rising against Roma despite Recognition Gains, Says Rights Expert." <https://News.Un.Org/En/Story/2021/04/1089232>.
- Voigtlander, Nico, and Hans-Joachim Voth. 2012. "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany." *The Quarterly Journal of Economics* 127 (3): 1339–92.
- Yanagizawa-Drott, David. 2014. "Propaganda and Conflict: Evidence from the Rwandan Genocide." *Quarterly Journal of Economics* 129 (4): 1947–94.

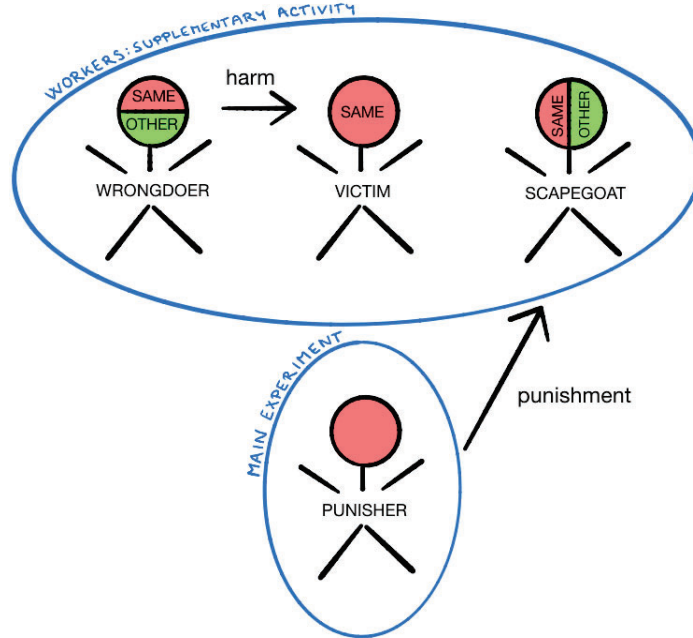
TABLE 1: PUNISHMENT OF THE SCAPEGOAT

	Punishment of the Scapegoat (intensity)	Punishment of the Scapegoat (yes)	Punishment of the Scapegoat (intensity)	Punishment of the Scapegoat (intensity)
Dependent variable			Punishment of the Scapegoat =	
Sample	All (1)	All (2)	yes (3)	All (4)
Harm intensity	0.08*** (0.02)	0.02*** (0.00)	0.20*** (0.07)	0.08*** (0.02)
Scapegoat OTHER	-0.00 (0.08)	0.03 (0.03)	-0.57 (0.37)	-0.04 (0.09)
Harm intensity*Scapegoat OTHER	0.09*** (0.03)	0.01** (0.01)	0.19** (0.08)	0.09*** (0.03)
Wrongdoer OTHER	-0.14 (0.13)	-0.03 (0.03)	-0.32 (0.24)	-0.14 (0.13)
Controls	baseline	baseline	baseline	full
Mean baseline (Scapegoat SAME, 0 harm)	0.13	0.04	3.14	0.13
Observations	1,685	1,685	329	1,685
R-squared	0.072	0.050	0.240	0.133

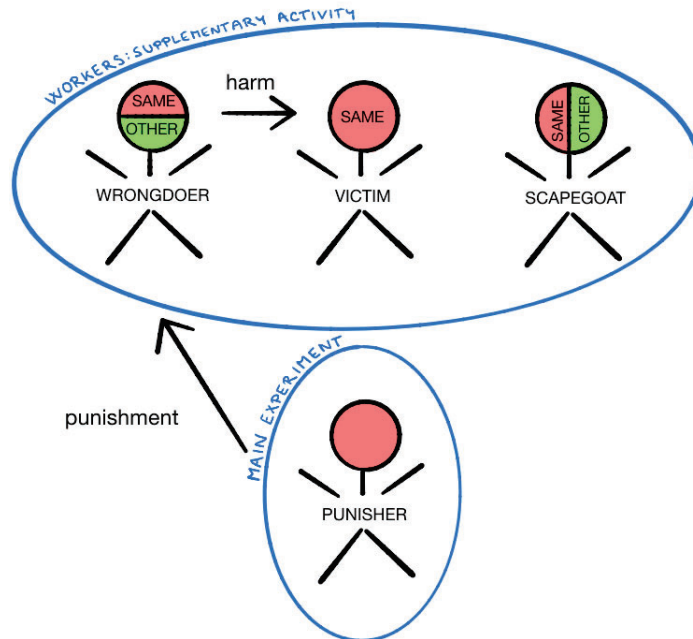
Notes: OLS, standard errors (in parentheses) are clustered at the Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable in Columns 1,3,4 is the extent of punishment of the Scapegoat (EUR 0-8). In Column 2, the dependent variable indicates that the Punisher chose non-zero punishment of the Scapegoat. "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Scapegoat OTHER" indicates that the Scapegoat comes from a different ethnic group (Roma minority) than the Punisher. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group than the Punisher. Baseline controls include the gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. Full controls also include experimenter fixed effects, a dummy variable indicating that the Punisher is a university student (vs. a secondary school student), location fixed effects, education of parents (dummy variables for mother/father with a university degree, dummy variables for education unknown), and a dummy variable indicating that the subject answered all control questions correctly on the first or second attempt. The sample is composed of Punishers from the majority group.

FIGURE 1: ILLUSTRATION OF THE EXPERIMENTAL TASKS

(A) PUNISHING THE SCAPEGOAT GAME



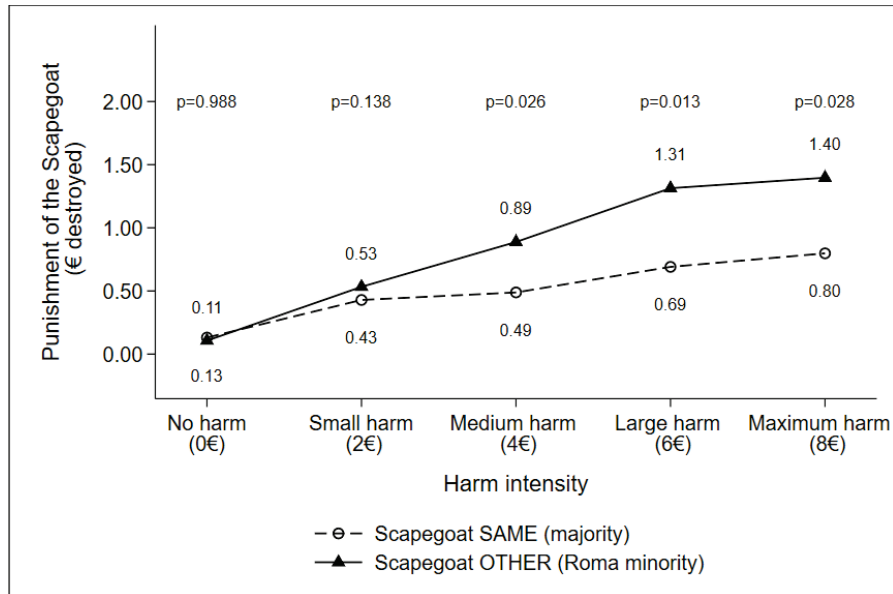
(B) PUNISHING THE WRONGDOER GAME



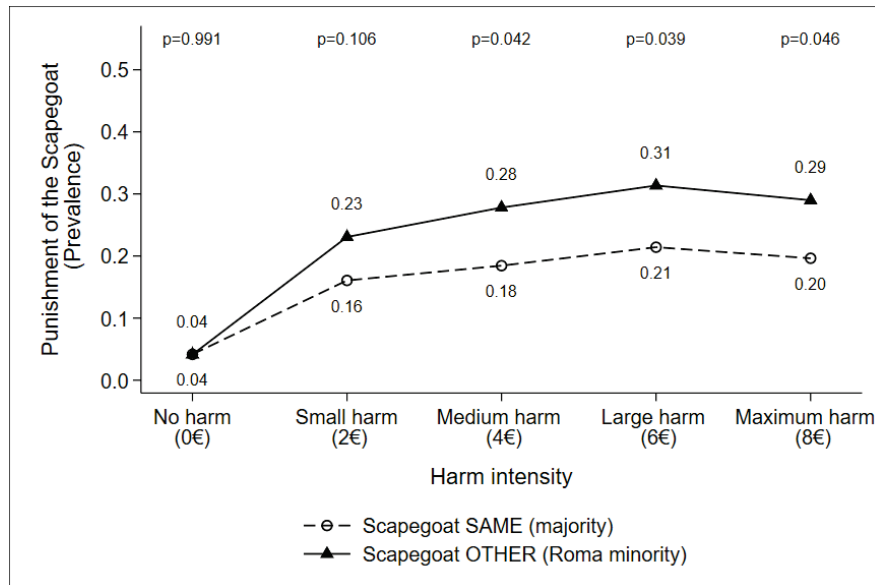
Notes: The main sample are Punishers, who come either from the majority group or from the Roma ethnic minority. Punishers learn about the harm committed by the Wrongdoers towards the Victims and can punish an innocent bystander—the Scapegoat (Panel A), or can directly punish the Wrongdoer (Panel B). The ethnic identity of the Scapegoat and the Wrongdoer is manipulated orthogonally: "SAME" indicates that the player (Wrongdoer/Victim/Scapegoat) is of the same ethnicity as the Punisher, while "OTHER" indicates that he is from the other ethnic group.

FIGURE 2: PUNISHMENT OF THE SCAPEGOAT, BY SCAPEGOAT'S ETHNICITY

(A) INTENSITY OF PUNISHMENT

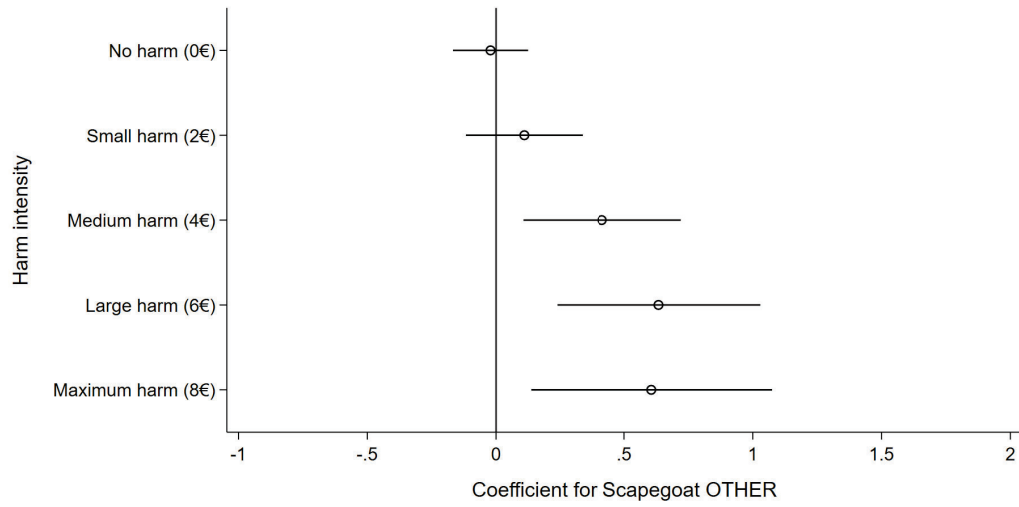


(B) PREVALENCE OF PUNISHMENT



Notes: Mean punishment of the Scapegoat (Panel A) and the share of Punishers who choose non-zero punishment of the Scapegoat (Panel B), by the ethnicity of the Scapegoat and the harm caused by the Wrongdoer to the Victim. Punishers (and Victims) are from the majority ethnic group. "Scapegoat SAME" indicates that the Scapegoat also comes from the majority ethnic group, while "Scapegoat OTHER" indicates that Scapegoat is ethnic Roma. Differences between the conditions are tested using the Wilcoxon rank-sum test; p-values are presented at the top. The sample is composed of Punishers from the majority group.

FIGURE 3: DISCRIMINATION AGAINST OTHER SCAPEGOATS (MAJORITY SAMPLE)



Notes: Estimated coefficients for "Scapegoat OTHER", with 95% confidence intervals. The dependent variable is the extent of punishment of the Scapegoat (EUR 0-8). The coefficients are estimated by separate OLS regressions for the five possible levels of harm caused by the Wrongdoer to the Victim. We control for the gender and age of the Punisher, and for a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. Punishers (and Victims) are from the majority ethnic group. "Scapegoat OTHER" indicates that the Wrongdoer comes from a different ethnic group (Roma minority) than the Punisher. The sample is composed of Punishers from the majority group.

[FOR ONLINE PUBLICATION]

Scapegoating: Experimental Evidence

Michal Bauer, Jana Cahlíková, Julie Chytilová, Gérard Roland, Tomáš Želinský

Supplementary Online Appendix A

This file contains appendix figure and tables

Appendix Figures A1-A7

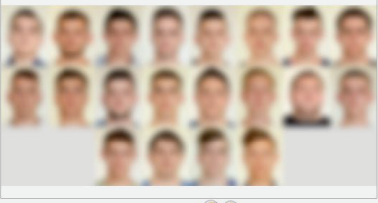
Appendix Tables A1-A15

FIGURE A1: EXAMPLES OF DECISION SCREENS

Panel A: Punishing the Scapegoat Game

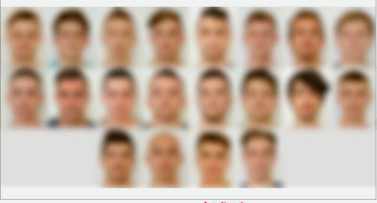
Situation 4 - Decision:

Person A



8 € 

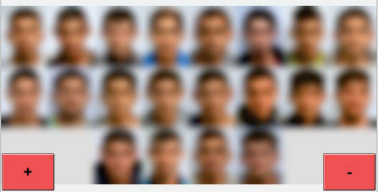
Person B



2 € 

Person A destroyed €6 of Person B

Person C



+ 6 €  -

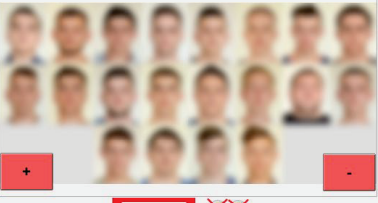
Your payoff: 8.80 €

OK

Panel B: Punishing the Wrongdoer Game

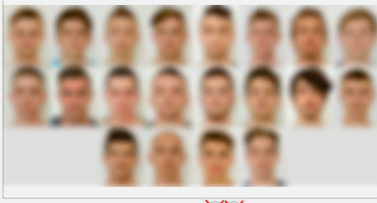
Situation 5 - Decision:

Person A



+ 4 €  -

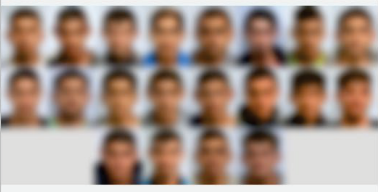
Person B



0 € 

Person A destroyed all €8 of Person B

Person C



8 € 

Your payoff: 8.60 €

OK

Notes: Panel A shows the decision screen for the Punishing the Scapegoat Game, for the situation in which the Wrongdoer (Person A) destroyed €6 of the earnings of the Victim (Person B). In this example, the Punisher is choosing to reduce the earnings of the Scapegoat by €2. Panel B shows the decision screen for the Punishing the Wrongdoer Game, for the situation in which the Wrongdoer (Person A) destroyed all €8 of the Victim's earnings (Person B). In this example, the Punisher is choosing to reduce the earnings of the Wrongdoer (Person C) by €4. The default/starting point in all situations is always Person A and Person C getting all €8. In the experiment, each picture shows 20 different passport-style photographs, which vary based on the condition. Here we show an example for the Wrongdoer SAME and Scapegoat OTHER condition. For privacy reasons, we blur the actual pictures in publicly available materials.

FIGURE A2: HISTOGRAM OF THE HARM IMPOSED ON THE VICTIMS BY THE WRONGDOERS

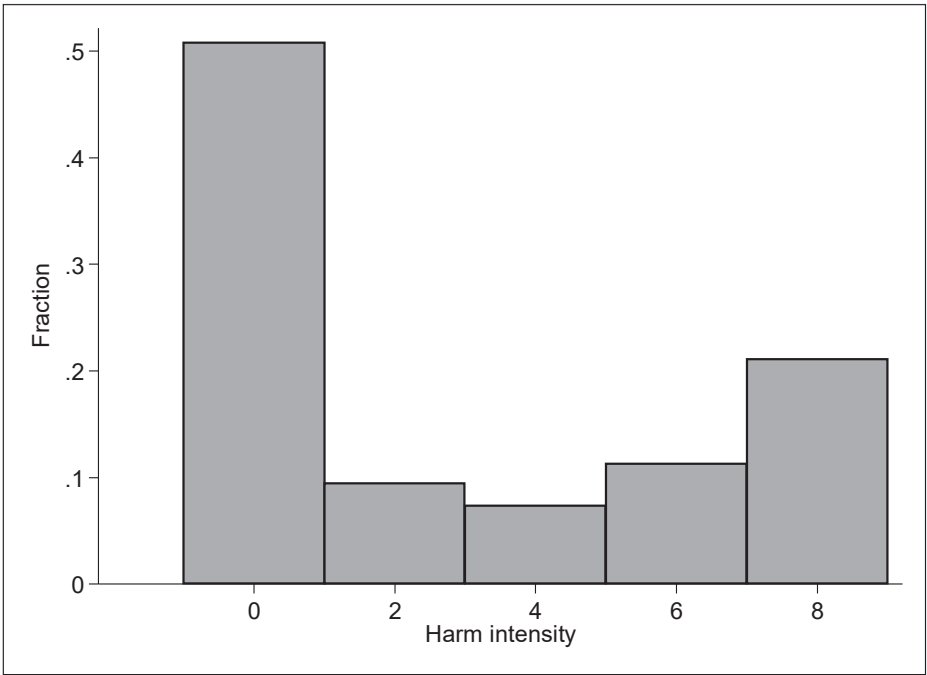
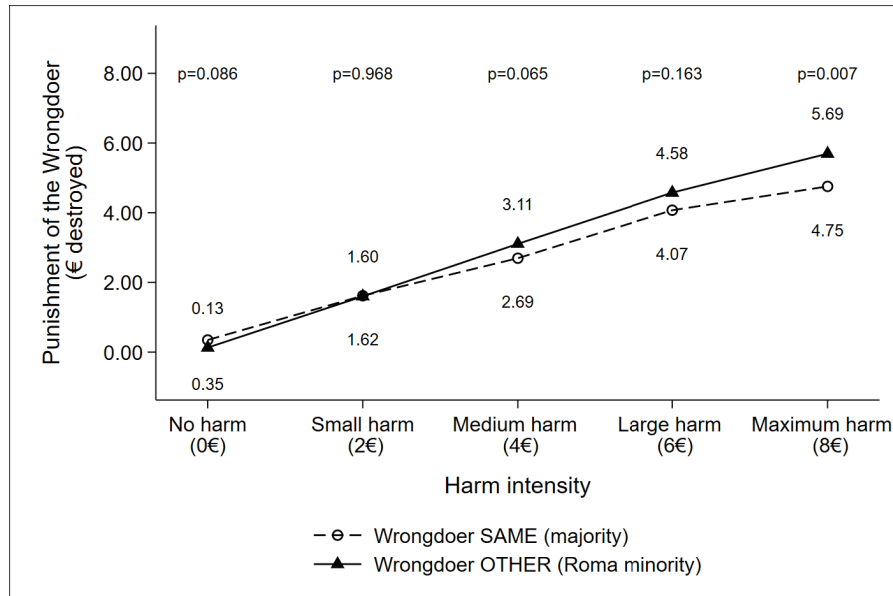
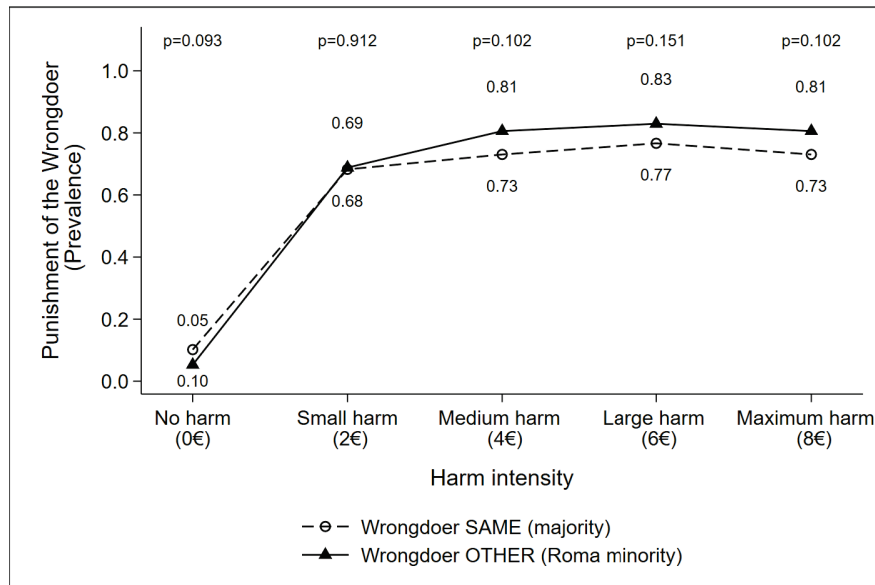


FIGURE A3: PUNISHMENT OF THE WRONGDOER, BY WRONGDOER'S ETHNICITY (MAJORITY SAMPLE)

(A) INTENSITY OF PUNISHMENT

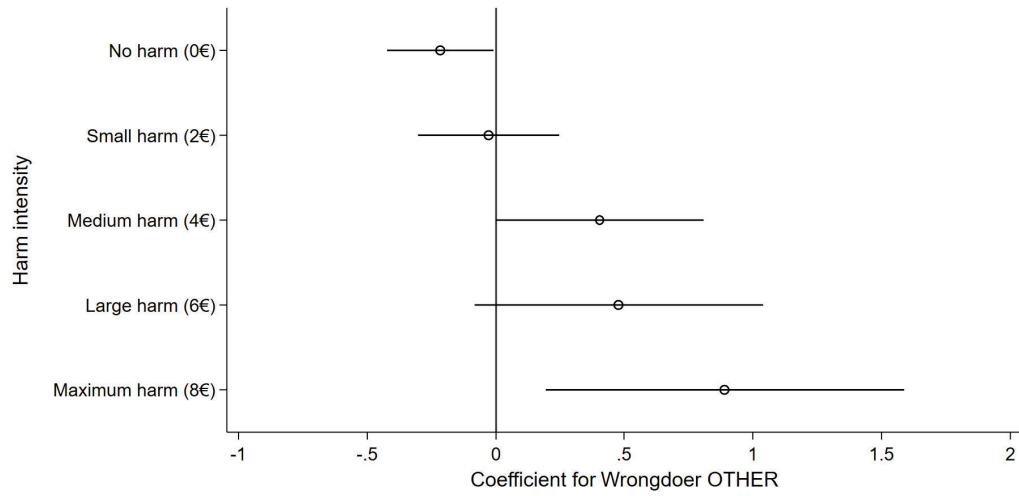


(B) PREVALENCE OF PUNISHMENT



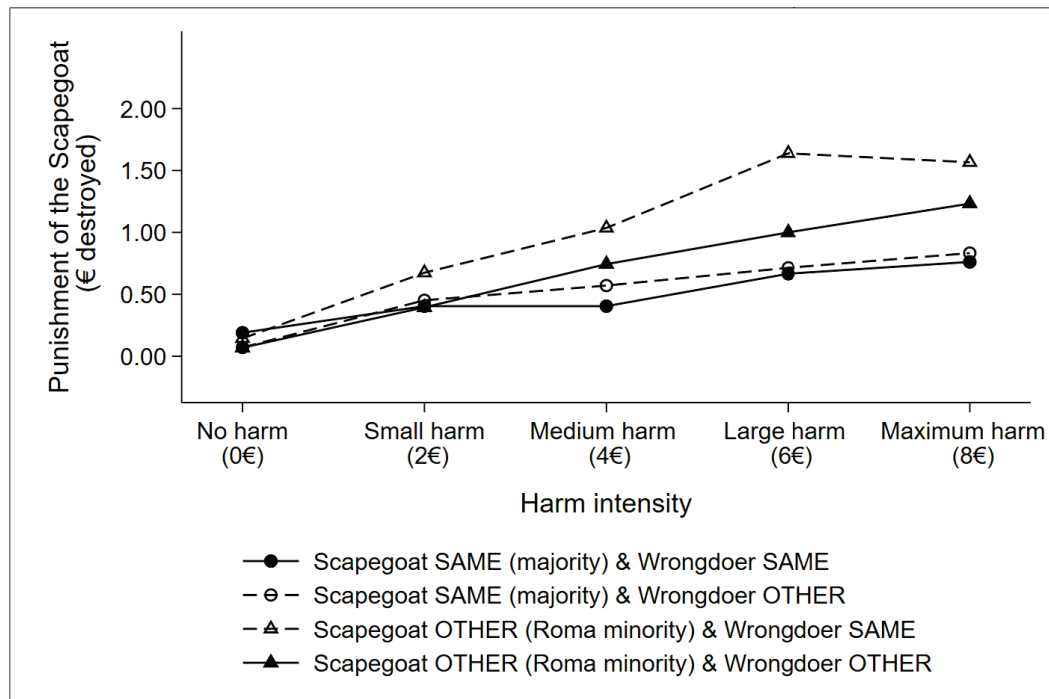
Notes: Mean punishment of the Wrongdoer (Panel A) and the share of Punishers who choose non-zero punishment of the Wrongdoer (Panel B), by the ethnicity of the Wrongdoer and the harm caused by the Wrongdoer to the Victim. Punishers (and Victims) are from the majority ethnic group. "Wrongdoer SAME" indicates that the Wrongdoer also comes from the majority ethnic group, while "Wrongdoer OTHER" indicates that the Wrongdoer is ethnic Roma. Differences between the conditions are tested using the Wilcoxon rank-sum test; p-values are presented at the top. The sample is composed of Punishers from the majority group.

FIGURE A4: DISCRIMINATION AGAINST OTHER WRONGDOERS (MAJORITY SAMPLE)



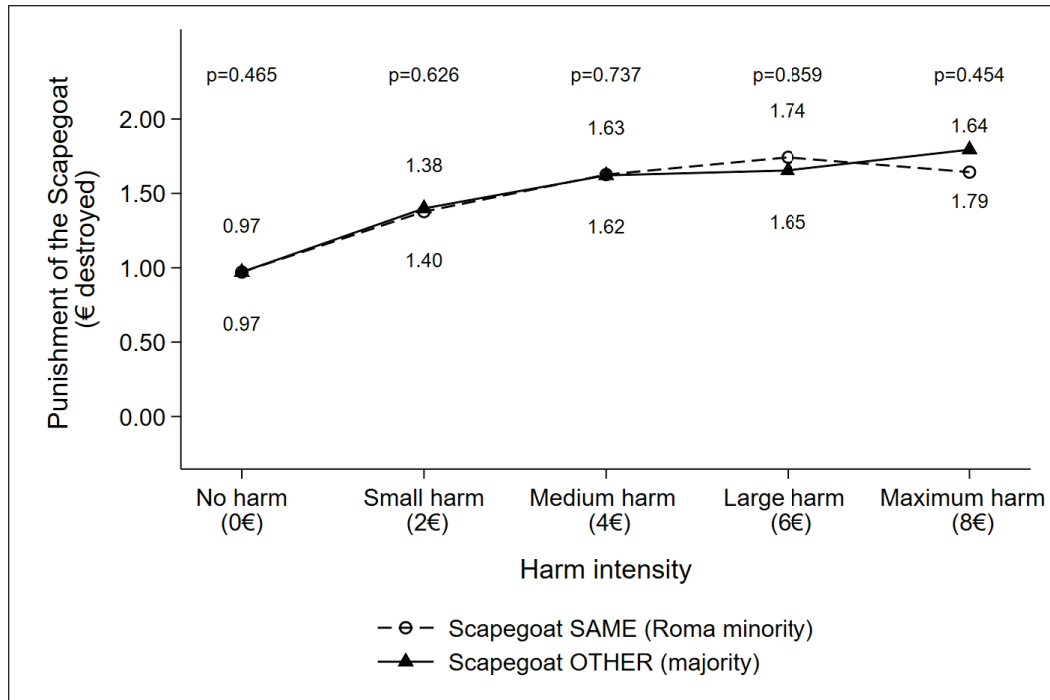
Notes: Estimated coefficients for "Wrongdoer OTHER", with 95% confidence intervals. The dependent variable is the extent of punishment of the Wrongdoer (EUR 0-8). The coefficients are estimated by separate OLS regressions for the five possible levels of harm caused by the Wrongdoer to the Victim. We control for the gender and age of the Punisher, and for a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. Punishers (and Victims) are from the majority ethnic group. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group (Roma minority) than the Punisher. The sample is composed of Punishers from the majority group.

FIGURE A5: PUNISHMENT OF THE SCAPEGOAT, BY WRONGDOER'S AND SCAPEGOAT'S ETHNICITY (MAJORITY SAMPLE)



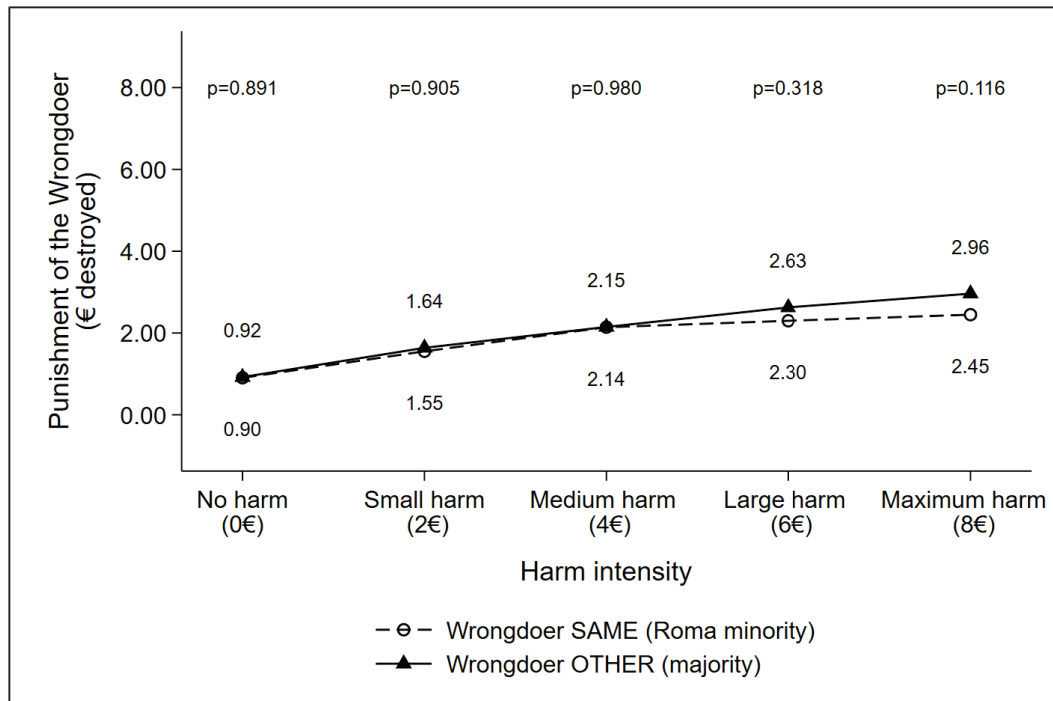
Notes: Mean punishment of the Scapegoat, by the ethnicity of the Scapegoat and the Wrongdoer and by the extent of harm caused by the Wrongdoer to the Victim. Punishers (and Victims) are from the majority ethnic group. "Scapegoat SAME" and "Wrongdoer SAME" indicate that the Scapegoat/the Wrongdoer also comes from the majority ethnic group, while "Scapegoat OTHER" and "Wrongdoer OTHER" indicate that the Scapegoat/the Wrongdoer is ethnic Roma. The sample is composed of Punishers from the majority group.

FIGURE A6: PUNISHMENT OF THE SCAPEGOAT, BY SCAPEGOAT'S ETHNICITY (ROMA MINORITY SAMPLE)



Notes: Mean punishment of the Scapegoat, by the ethnicity of the Scapegoat and the harm caused by the Wrongdoer to the Victim. Punishers (and Victims) are from the Roma ethnic group. "Scapegoat SAME" indicates that the Scapegoat also comes from the Roma ethnic group, while "Scapegoat OTHER" indicates that Scapegoat has majority ethnicity. Differences between the conditions are tested using the Wilcoxon rank-sum test; p-values are presented on top. The sample is composed of Punishers from the Roma ethnic minority group.

FIGURE A7: PUNISHMENT OF THE WRONGDOER, BY WRONGDOER'S ETHNICITY (ROMA MINORITY SAMPLE)



Notes: Mean punishment of the Wrongdoer, by the ethnicity of the Wrongdoer and the harm caused by the Wrongdoer to the Victim. Punishers (and Victims) are from the Roma ethnic group. "Wrongdoer SAME" indicates that the Wrongdoer also comes from the Roma ethnic group, while "Wrongdoer OTHER" indicates that Wrongdoer has majority ethnicity. Differences between the conditions are tested using the Wilcoxon rank-sum test; p-values are presented on top. The sample is composed of Punishers from the Roma ethnic minority group.

TABLE A1: DESCRIPTIVE STATISTICS AND RANDOMIZATION CHECKS (MAJORITY SAMPLE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Means across treatments				
Sample	Whole sample	N	Wrongdoer SAME	Wrongdoer OTHER	Scapegoat SAME	Scapegoat OTHER	Diff (F-test)
Female	0.42	337	0.45	0.39	0.45	0.39	0.31
Age	19.31	337	19.30	19.33	19.35	19.28	0.91
Student (high school)	0.57	337	0.57	0.56	0.57	0.57	0.97
Student (university)	0.43	337	0.43	0.44	0.43	0.43	0.97
Primary school degree	0.54	337	0.54	0.53	0.55	0.53	0.89
Secondary school degree	0.40	337	0.38	0.41	0.38	0.42	0.61
University degree	0.07	337	0.07	0.06	0.08	0.05	0.60
Child	0.00	337	0.01	0.00	0.01	0.00	0.37
Household size	4.24	337	4.32	4.17	4.17	4.31	0.31
Regularly goes to church	0.32	337	0.35	0.29	0.33	0.31	0.48
Father with secondary school	0.74	330	0.75	0.73	0.77	0.72	0.57
Father with university	0.25	330	0.25	0.26	0.23	0.28	0.66
Mother with secondary school	0.68	336	0.69	0.66	0.69	0.66	0.65
Mother with university	0.32	336	0.30	0.34	0.30	0.33	0.64
Father employed	0.96	309	0.96	0.97	0.97	0.96	0.65
Mother employed	0.88	320	0.87	0.89	0.88	0.88	0.81
Family owns a car	0.91	337	0.93	0.89	0.90	0.92	0.31
Family owns a computer	0.99	336	0.99	1.00	1.00	0.99	0.37
Family owns a smartphone	0.96	337	0.96	0.96	0.96	0.96	0.94
Family owns a tablet	0.64	337	0.67	0.60	0.71	0.56	0.01
Control questions first attempt (max=6)	5.93	337	5.96	5.89	5.95	5.90	0.11
Control questions second attempt (max=6)	5.99	337	6.00	5.97	5.99	5.98	0.22
All control questions correct	0.99	337	1.00	0.98	0.99	0.99	0.20
N	337	337	167	170	168	169	

Notes: Descriptive statistics of the majority sample (Column 1). Columns 3-6 present means for subjects in the specified experimental conditions. Experimental balance is tested in Column 7 using an F-test, showing that dummy variables “Wrongdoer OTHER” and “Scapegoat OTHER” do not significantly predict the specified descriptive variable. Variable “All control questions correct” indicates that the subject answered all control questions correctly at first or second attempt. For all variables, the values are missing for unspecified answers and for “I do not know” answers; Column 2 gives the number of non-missing values.

TABLE A2: DESCRIPTIVE STATISTICS AND RANDOMIZATION CHECKS (ROMA MINORITY SAMPLE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Means across treatments						
Sample	Whole sample	N	Wrongdoer SAME	Wrongdoer OTHER	Scapegoat SAME	Scapegoat OTHER	Diff (F-test)
Female	0.55	484	0.54	0.57	0.56	0.55	0.76
Age	20.23	484	20.36	20.10	20.17	20.28	0.25
Student (secondary school)	0.14	484	0.12	0.16	0.16	0.12	0.12
Student (university)	0.01	484	0.02	0.01	0.02	0.01	0.48
Employed	0.17	484	0.18	0.16	0.18	0.16	0.66
Unemployed	0.49	484	0.49	0.48	0.46	0.51	0.42
At home	0.19	484	0.19	0.19	0.18	0.20	0.90
Primary school degree	0.57	484	0.56	0.58	0.56	0.58	0.79
Secondary school degree	0.43	484	0.44	0.42	0.44	0.42	0.79
Married	0.31	484	0.29	0.33	0.32	0.31	0.59
Child	0.48	484	0.50	0.45	0.48	0.47	0.61
Household size	6.02	484	6.00	6.05	6.09	5.95	0.88
Regularly goes to church	0.27	484	0.26	0.28	0.24	0.29	0.43
Father with secondary school	0.39	401	0.37	0.42	0.39	0.39	0.60
Father with university	0.04	401	0.03	0.05	0.04	0.04	0.47
Mother with secondary school	0.32	414	0.32	0.33	0.33	0.32	0.97
Mother with university	0.03	414	0.03	0.02	0.02	0.03	0.67
Father employed	0.55	392	0.55	0.55	0.56	0.53	0.87
Mother employed	0.38	436	0.41	0.35	0.39	0.38	0.37
Family owns a car	0.32	484	0.32	0.31	0.32	0.31	0.93
Family owns a computer	0.52	484	0.50	0.54	0.53	0.51	0.69
Family owns a smartphone	0.85	484	1.06	0.65	1.06	0.65	0.37
Family owns a tablet	0.39	484	0.36	0.42	0.42	0.37	0.18
Control questions first attempt (max=6)	4.27	483	4.18	4.36	4.16	4.38	0.19
Control questions second attempt (max=6)	5.53	482	5.47	5.58	5.53	5.52	0.50
All control questions correct	0.77	484	0.76	0.78	0.77	0.77	0.88
N	484	484	241	243	241	243	

Notes: Descriptive statistics of the Roma minority sample (Column 1). Columns 3-6 present means for subjects in the specified experimental conditions. Experimental balance is tested in Column 7 using an F-test, showing that dummy variables “Wrongdoer OTHER” and “Scapegoat OTHER” do not significantly predict the specified descriptive variable. Variable “All control questions correct” indicates that the subject answered all control questions correctly at first or second attempt. For all variables, the values are missing for unspecified answers and for “I do not know” answers; Column 2 gives the number of non-missing values.

TABLE A3: PUNISHMENT OF THE SCAPEGOAT, ROBUSTNESS CHECKS (MAJORITY SAMPLE)

Dependent variable	Punishment of the Scapegoat (intensity)								Subjects who answered all control questions correctly	Excluding potentially hypothetical observations
Sample	All (1)	All (2)	All (3)	All (4)	All (5)	All (6)	All (7)	All (8)	All (9)	
Harm intensity	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.11*** (0.02)	
Scapegoat OTHER	-0.00 (0.08)	-0.01 (0.08)	-0.01 (0.08)	-0.00 (0.08)	-0.01 (0.08)	-0.01 (0.08)	-0.03 (0.09)	-0.02 (0.08)	0.03 (0.09)	
Harm intensity*Scapegoat OTHER	0.09*** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.08*** (0.03)	0.08*** (0.04)	
Wrongdoer OTHER	-0.14 (0.13)	-0.14 (0.13)	-0.15 (0.13)	-0.14 (0.13)	-0.14 (0.13)	-0.12 (0.13)	-0.12 (0.13)	-0.16 (0.13)	-0.13 (0.15)	
Controls										
Baseline controls	x		x	x	x	x	x	x	x	
Design features			x				x			
Sample University (vs. High school)				x			x			
Location fixed effects					x		x			
Education of parents						x	x			
Mean baseline (Scapegoat SAME, 0 harm)	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	
Observations	1,685	1,685	1,685	1,685	1,685	1,685	1,685	1,670	1,241	
R-squared	0.072	0.068	0.100	0.072	0.091	0.085	0.126	0.070	0.092	

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment towards the Scapegoat (EUR 0-8). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Scapegoat OTHER" indicates that the Scapegoat comes from a different ethnic group (Roma minority) than the Punisher. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group than the Punisher. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. Design features include experimenter fixed effects. Education of parents includes dummy variables for mother/father with a university degree, and dummy variables for unknown education. In Column 8, we exclude all subjects who did not answer all control questions correctly at first or second attempt. In Column 9, we exclude all observations for harm levels, which a decision-maker believed were chosen by none of the twenty potential Wrongdoers. The sample is composed of Punishers from the majority group.

TABLE A4: PUNISHMENT OF THE SCAPEGOAT AND THE WRONGDOER, NON-LINEAR SPECIFICATIONS
(MAJORITY SAMPLE AND ROMA MINORITY SAMPLE)

Dependent variable	Punishment of the Scapegoat (intensity)	Punishment of the Wrongdoer (intensity)	Punishment of the Scapegoat (intensity)	Punishment of the Wrongdoer (intensity)
Sample	Majority		Roma minority	
	(1)	(2)	(3)	(4)
Small harm (2 euro)	0.30*** -0.08	1.27*** -0.12	0.41*** -0.11	0.65*** -0.11
Medium harm (4 euro)	0.36*** -0.09	2.35*** -0.17	0.66*** -0.13	1.24*** -0.15
Large harm (6 euro)	0.56*** -0.12	3.72*** -0.23	0.77*** -0.15	1.39*** -0.18
Maximum harm (8 euro)	0.67*** -0.15	4.41*** -0.29	0.67*** -0.16	1.54*** -0.2
Wrongdoer OTHER	-0.14 -0.13	-0.24** -0.11	0.04 -0.17	-0.02 -0.16
Wrongdoer OTHER*Small harm (2 euro)		0.2 -0.15		0.07 -0.17
Wrongdoer OTHER*Medium harm (4 euro)		0.63*** -0.22		-0.01 -0.22
Wrongdoer OTHER*Large harm (6 euro)		0.72** -0.3		0.31 -0.26
Wrongdoer OTHER*Maximum harm (8 euro)		1.16*** -0.38		0.5 -0.31
Scapegoat OTHER	-0.02 -0.07		0 -0.17	
Scapegoat OTHER*Small harm (2 euro)	0.13 -0.11		0.02 -0.16	
Scapegoat OTHER*Medium harm (4 euro)	0.42*** -0.15		-0.01 -0.18	
Scapegoat OTHER*Large harm (6 euro)	0.65*** -0.2		-0.09 -0.21	
Scapegoat OTHER*Maximum harm (8 euro)	0.62*** -0.23		0.15 -0.24	
Controls	baseline	baseline	baseline	baseline
Mean baseline	0.13	0.35	0.97	0.9
Observations	1,685	1,685	2,420	2,420
R-squared	0.074	0.419	0.019	0.076

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable in Columns 1 and 3 is the extent of punishment towards the Scapegoat (EUR 0-8), the dependent variable in Columns 2 and 4 is the extent of punishment towards the Wrongdoer (EUR 0-8). Punishers (and Victims) in Columns 1 and 2 are from the majority ethnic group, Punishers (and Victims) in Columns 3 and 4 are from the Roma ethnic group. "Small harm (2 euro)", "Medium harm (4 euro)", "Large harm (6 euro)", and "Maximum harm (8 euro)" are dummy variables specifying the level of harm caused by the Wrongdoer to the Victim; zero harm (0 euro) is the omitted category. "Scapegoat OTHER"/"Wrongdoer OTHER" indicate that the Scapegoat/the Wrongdoer comes from a different ethnic group than the Punisher. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game.

TABLE A5: PUNISHMENT OF THE SCAPEGOAT, HETEROGENEITY ANALYSIS (MAJORITY SAMPLE)

Dependent variable	Punishment of the Scapegoat (intensity)					
	Men	Women	High school	University	Parents with at most secondary education	At least one parent with at university education
Sample	(1)	(2)	(3)	(4)	(5)	(6)
Harm intensity	0.05** (0.02)	0.11*** (0.03)	0.09*** (0.02)	0.06** (0.03)	0.08*** (0.02)	0.07** (0.03)
Scapegoat OTHER	-0.10 (0.12)	0.09 (0.10)	-0.04 (0.09)	-0.00 (0.15)	0.05 (0.08)	-0.09 (0.16)
Harm intensity*Scapegoat OTHER	0.13*** (0.04)	0.03 (0.04)	0.08** (0.04)	0.10** (0.04)	0.09** (0.04)	0.09* (0.05)
Wrongdoer OTHER	-0.01 (0.19)	-0.35* (0.19)	-0.15 (0.15)	-0.15 (0.22)	0.01 (0.17)	-0.32 (0.22)
Controls	baseline	baseline	baseline	baseline	baseline	baseline
Mean baseline (Scapegoat SAME, 0 harm)	0.19	0.05	0.08	0.19	0.06	0.24
Observations	980	705	960	725	975	685
R-squared	0.074	0.094	0.101	0.084	0.092	0.063

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment towards the Scapegoat (EUR 0-8). Columns 1 and 2 report results by the gender of the Punisher. Columns 3 and 4 report results for Punishers who are secondary school students and university students, respectively. Columns 5 and 6 report results for Punishers who have both parents with at most secondary school education and those who have at least one parent with a university degree, respectively (self-reported, omitting subjects with missing answers). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Scapegoat OTHER" indicates that the Scapegoat comes from a different ethnic group (Roma minority) than the Punisher. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group than the Punisher. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. The sample is composed of Punishers from the majority group.

TABLE A6: PUNISHMENT OF THE WRONGDOER

	Punishment of the Wrongdoer (intensity)	Punishment of the Wrongdoer (yes)	Punishment of the Wrongdoer (intensity) Punishment of the Wrongdoer	Punishment of the Wrongdoer (intensity)
Dependent variable				
Sample	All (1)	All (2)	= yes (3)	All (4)
Harm intensity	0.56*** (0.03)	0.07*** (0.00)	0.63*** (0.04)	0.56*** (0.04)
Wrongdoer OTHER	-0.26** (0.11)	-0.03 (0.03)	-0.53** (0.23)	-0.32** (0.13)
Harm intensity*Wrongdoer OTHER	0.14*** (0.05)	0.02*** (0.01)	0.14*** (0.05)	0.14*** (0.05)
Controls	baseline	baseline	baseline	full
Mean baseline (Wrongdoer SAME, 0 harm)	0.35	0.10	3.41	0.35
Observations	1,685	1,685	1,044	1,685
R-squared	0.417	0.201	0.540	0.453

Notes: OLS, standard errors (in parentheses) are clustered at the Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable in Columns 1,3,4 is the extent of punishment of the Wrongdoer (EUR 0-8). In Column 2, the dependent variable indicates that the Punisher chose non-zero punishment of the Wrongdoer. "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group (Roma minority) than the Punisher. Baseline and full controls are defined in Table 1. Full controls also include a dummy variable "Scapegoat OTHER" (equal to one if the Scapegoat comes from a different ethnic group than the Punisher). The sample is composed of Punishers from the majority group.

TABLE A7: PUNISHMENT OF THE WRONGDOER, ROBUSTNESS CHECKS (MAJORITY SAMPLE)

Dependent variable	Punishment of the Wrongdoer (intensity)							Subjects who answered all control questions correctly	Excluding potentially hypothetical observations
	All	All	All	All	All	All	All		
Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Harm intensity	0.56*** (0.03)	0.56*** (0.03)	0.56*** (0.04)	0.56*** (0.03)	0.56*** (0.04)	0.56*** (0.03)	0.56*** (0.04)	0.56*** (0.03)	0.62*** (0.04)
Wrongdoer OTHER	-0.26*** (0.11)	-0.24*** (0.11)	-0.28*** (0.12)	-0.26*** (0.11)	-0.27*** (0.12)	-0.28*** (0.11)	-0.31*** (0.13)	-0.28*** (0.11)	-0.35*** (0.13)
Harm intensity*Wrongdoer OTHER	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)
Controls									
Baseline controls	x		x	x	x	x	x	x	x
Design features			x				x		
Sample University (vs. High school)				x			x		
Location fixed effects					x		x		
Education of parents						x	x		
Mean baseline (Wrongdoer SAME, 0 harm)	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.37
Observations	1,685	1,685	1,685	1,685	1,685	1,685	1,685	1,670	1,241
R-squared	0.417	0.408	0.429	0.420	0.440	0.419	0.453	0.417	0.504

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment towards the Wrongdoer (EUR 0-8). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group (Roma minority) than the Punisher. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. Design features include a dummy variable "Scapegoat OTHER" (equal to one if the Scapegoat comes from a different ethnic group than the Punisher) and experimenter fixed effects. Education of parents includes dummy variables for mother/father with a university degree, and dummy variables for unknown education. In Column 8, we exclude all subjects who did not answer all control questions correctly at first or second attempt. In Column 9, we exclude all observations for harm levels, which a decision-maker believed were chosen by none of the twenty potential Wrongdoers. The sample is composed of Punishers from the majority group.

TABLE A8: PUNISHMENT OF THE WRONGDOER, HETEROGENEITY ANALYSIS (MAJORITY SAMPLE)

Dependent variable	Punishment of the Wrongdoer (intensity)					
	Men	Women	High school sample	University sample	Parents with at most secondary education	At least one parent with at university education
Sample	(1)	(2)	(3)	(4)	(5)	(6)
Harm intensity	0.61*** (0.05)	0.51*** (0.05)	0.62*** (0.04)	0.49*** (0.06)	0.57*** (0.04)	0.56*** (0.06)
Wrongdoer OTHER	-0.18 (0.18)	-0.37*** (0.14)	-0.22* (0.13)	-0.40* (0.21)	-0.21 (0.13)	-0.35 (0.21)
Harm intensity*Wrongdoer OTHER	0.14** (0.06)	0.12* (0.07)	0.07 (0.06)	0.23*** (0.07)	0.11* (0.06)	0.18** (0.08)
Controls	baseline	baseline	baseline	baseline	baseline	baseline
Mean baseline (Wrongdoer SAME, 0 harm)	0.30	0.40	0.23	0.51	0.28	0.46
Observations	980	705	960	725	975	685
R-squared	0.477	0.355	0.451	0.404	0.403	0.446

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment towards the Wrongdoer (EUR 0-8). Columns 1 and 2 report results by the gender of the Punisher. Columns 3 and 4 report results for Punishers who are secondary school students and university students, respectively. Columns 5 and 6 report results for Punishers who have both parents with at most secondary school education and those who have at least one parent with a university degree, respectively (self-reported, omitting subjects with missing answers). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group (Roma minority) than the Punisher. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. The sample is composed of Punishers from the majority group.

TABLE A9: BELIEFS REGARDING THE BEHAVIOR OF THE WRONGDOER (MAJORITY SAMPLE)

	Harm	Wrongdoer SAME	Wrongdoer OTHER	Diff (Fisher exact test)
	(1)	(2)	(3)	(4)
Panel A: Beliefs Task 1—Wrongdoer destroys:	0	60.5	63.5	
	2	18.6	14.7	
	4	11.4	11.2	
	6	3.6	2.4	
	8	6.0	8.2	
				0.759
	Harm	Wrongdoer SAME	Wrongdoer OTHER	Diff (ranksum)
	(1)	(2)	(3)	(4)
Panel B: Beliefs Task2—Out of 20 Wrongdoers, how many destroy:	0	8.6	8.5	0.860
	2	3.4	3.0	0.322
	4	3.2	2.9	0.441
	6	2.2	2.4	0.403
	8	2.6	3.1	0.015

Notes: Beliefs of the Punishers from the majority ethnic group regarding which harm levels were chosen by the Wrongdoers. In Task 1 (Panel A), the Punishers were asked to guess which action was chosen by the Wrongdoer and rewarded 1 euro for a correct answer. The shares for majority ethnicity Wrongdoers and Roma ethnicity Wrongdoers are reported in Columns 2 and 3, respectively. The equality of the distributions is tested in Column 4 using a Fisher exact test. In Task 2 (Panel B), the Punishers were asked to guess how many of the twenty potential Wrongdoers chose each level of harm. The estimated number of Wrongdoers is reported in Columns 2 and 3 for the Wrongdoers of majority ethnicity and for the Wrongdoers of Roma ethnicity, respectively (we exclude 20 subjects whose reports do not sum up to twenty). Differences between conditions are tested using the Wilcoxon rank-sum test; p-values are specified in Column 4. The sample is composed of Punishers from the majority group.

TABLE A10: TESTING FOR COLLECTIVE PUNISHMENT

Dependent variable	Punishment of the Scapegoat (intensity)			
	Wrongdoer	Wrongdoer	Scapegoat	Scapegoat
Sample	SAME	OTHER	SAME	OTHER
	(1)	(2)	(3)	(4)
Harm intensity	0.07**	0.09***	0.09***	0.19***
	(0.03)	(0.02)	(0.02)	(0.03)
Scapegoat OTHER	0.06	-0.07		
	(0.14)	(0.08)		
Harm intensity*Scapegoat OTHER	0.12***	0.06		
	(0.04)	(0.04)		
Wrongdoer SAME			0.02	
			(0.13)	
Harm intensity*Wrongdoer SAME			-0.02	
			(0.04)	
Wrongdoer OTHER				-0.15
				(0.09)
Harm intensity*Wrongdoer OTHER				-0.04
				(0.05)
Controls	baseline	baseline	baseline	baseline
Mean baseline	0.19	0.07	0.07	0.14
Observations	835	850	840	845
R-squared	0.101	0.061	0.036	0.084

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment of the Scapegoat (EUR 0-8). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Wrongdoer SAME" and "Wrongdoer OTHER" indicate that the Wrongdoer comes from the same ethnic group or a different ethnic group (Roma minority) than the Punisher, respectively. "Scapegoat SAME" and "Scapegoat OTHER" indicate that the Scapegoat comes from the same ethnic group or a different ethnic group (Roma minority) than the Punisher, respectively. Baseline controls are defined in Table 1. The sample is composed of Punishers from the majority group.

TABLE A11: PERCEIVED SOCIO-ECONOMIC STATUS OF SAME/OTHER WRONGDOERS AND SCAPEGOATS
(MAJORITY SAMPLE)

Dependent variable	Wrongdoer Unemployed (perceived)	Wrongdoer Student (perceived)	Wrongdoer's father low education (perceived)	Scapegoat Unemployed (perceived)	Scapegoat Student (perceived)	Scapegoat's father low education (perceived)
Sample	All	All	All	All	All	All
	(1)	(2)	(3)	(4)	(5)	(6)
Wrongdoer OTHER	0.52*** (0.04)	-0.69*** (0.04)	0.65*** (0.04)			
Scapegoat OTHER				0.40*** (0.04)	-0.61*** (0.04)	0.63*** (0.04)
Constant	0.07*** (0.02)	0.85*** (0.03)	0.02** (0.01)	0.08*** (0.02)	0.78*** (0.03)	0.02* (0.01)
Controls	no	no	no	no	no	no
Observations	337	337	337	337	337	337
R-squared	0.309	0.478	0.466	0.201	0.371	0.443

Notes: OLS, robust standard errors in parentheses. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variables are perceived socio-economic characteristics of the Wrongdoer (Column 1-3) and the Scapegoat (Column 4-6), elicited from the Punishers at the end of the experiment. Punishers are from the majority ethnic group. "Wrongdoer OTHER"/"Scapegoat OTHER" indicates that the Wrongdoer/the Scapegoat come from a different ethnic group (Roma minority) than the Punisher. The sample is composed of Punishers from the majority group.

TABLE A12: PUNISHMENT OF THE SCAPEGOAT, CONTROLLING FOR THE PERCEIVED SOCIO-ECONOMIC ATTRIBUTES

(MAJORITY SAMPLE)

Dependent variable Sample	Punishment of the Scapegoat (intensity)			
	All (1)	All (2)	All (3)	All (4)
Harm intensity	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)
Scapegoat OTHER	-0.00 (0.08)	-0.08 (0.08)	-0.04 (0.10)	-0.08 (0.10)
Harm intensity*Scapegoat OTHER	0.09*** (0.03)	0.07** (0.03)	0.08** (0.04)	0.07* (0.04)
Scapegoat Unemployed (perceived)		0.18** (0.09)		0.18** (0.09)
Harm intensity*Scapegoat Unemployed		0.04 (0.04)		0.04 (0.04)
Scapegoat's father low education (perceived)			0.06 (0.09)	0.01 (0.09)
Harm intensity*Scapegoat's father low education			0.01 (0.04)	0.00 (0.04)
Wrongdoer OTHER	-0.14 (0.13)	-0.10 (0.13)	-0.13 (0.13)	-0.10 (0.13)
Controls	baseline	baseline	baseline	baseline
Mean baseline (Scapegoat SAME, 0 harm)	0.13	0.13	0.13	0.13
Observations	1,685	1,685	1,685	1,685
R-squared	0.072	0.080	0.073	0.080

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment towards the Scapegoat (EUR 0-8). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Scapegoat OTHER" indicates that the Scapegoat comes from a different ethnic group (Roma minority) than the Punisher. "Scapegoat Unemployed (perceived)" indicates that the Punisher believes that the (SAME or OTHER) Scapegoat is unemployed. "Scapegoat's father low education (perceived)" indicates that the Punisher believes that the (SAME or OTHER) Scapegoat has a father with at most primary school education. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group than the Punisher. Perceived socio-economic characteristics of the Scapegoat were elicited from the Punishers at the end of the experiment. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. The sample is composed of Punishers from the majority group.

TABLE A13: PUNISHMENT OF THE WRONGDOER, CONTROLLING FOR THE PERCEIVED SOCIO-ECONOMIC ATTRIBUTES

(MAJORITY SAMPLE)

Dependent variable Sample	Punishment of the Wrongdoer (intensity)			
	All (1)	All (2)	All (3)	All (4)
Harm intensity	0.56*** (0.03)	0.56*** (0.03)	0.56*** (0.03)	0.56*** (0.03)
Wrongdoer OTHER	-0.26** (0.11)	-0.33** (0.16)	-0.23* (0.14)	-0.29* (0.17)
Harm intensity*Wrongdoer OTHER	0.14*** (0.05)	0.11* (0.06)	0.14** (0.06)	0.11 (0.07)
Wrongdoer Unemployed (perceived)		0.13 (0.17)		0.14 (0.18)
Harm intensity*Wrongdoer Unemployed		0.07 (0.06)		0.07 (0.06)
Wrongdoer's father low education (perceived)			-0.04 (0.13)	-0.07 (0.13)
Harm intensity*Wrongdoer's father low education			0.01 (0.06)	-0.01 (0.06)
Controls	baseline	baseline	baseline	baseline
Mean baseline (Wrongdoer SAME, 0 harm)	0.35	0.35	0.35	0.35
Observations	1,685	1,685	1,685	1,685
R-squared	0.417	0.420	0.417	0.420

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable is the extent of punishment towards the Wrongdoer (EUR 0-8). "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the majority ethnic group. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group (Roma minority) than the Punisher. "Wrongdoer Unemployed (perceived)" indicates that the Punisher believes that the (SAME or OTHER) Wrongdoer is unemployed. "Wrongdoer's father low education (perceived)" indicates that the Punisher believes that the (SAME or OTHER) Wrongdoer has a father with at most primary school education. Perceived socio-economic characteristics of the Wrongdoer were elicited from the Punishers at the end of the experiment. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. The sample is composed of Punishers from the majority group.

TABLE A14: PUNISHMENT OF THE SCAPEGOAT (ROMA MINORITY SAMPLE)

Dependent variable	Punishment of the Scapegoat (intensity)	Punishment of the Scapegoat (yes)	Punishment of the Scapegoat (intensity) Punishment of the Scapegoat =	Punishment of the Scapegoat (intensity)	Punishment of the Scapegoat (intensity)
Sample	All (1)	All (2)	yes (3)	All (4)	All CQ correctly (4)
Harm intensity	0.09*** (0.02)	0.01*** (0.00)	0.16*** (0.04)	0.09*** (0.02)	0.10*** (0.02)
Scapegoat OTHER	-0.02 (0.16)	0.03 (0.04)	-0.38 (0.27)	0.01 (0.16)	-0.12 (0.17)
Harm intensity*Scapegoat OTHER	0.01 (0.03)	-0.00 (0.00)	0.03 (0.05)	0.01 (0.03)	0.01 (0.03)
Wrongdoer OTHER	0.04 (0.17)	-0.02 (0.04)	0.27 (0.18)	-0.01 (0.16)	0.04 (0.18)
Controls	baseline	baseline	baseline	full	baseline
Mean baseline (Scapegoat SAME, 0 harm)	0.97	0.26	3.77	0.97	0.80
Observations	2,420	2,420	906	2,420	1,860
R-squared	0.017	0.008	0.064	0.093	0.021

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable in Columns 1,3,4 is the extent of punishment towards the Scapegoat (EUR 0-8). In Column 2, the dependent variable indicates whether the Punisher chose non-zero punishment of the Scapegoat. "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the Roma ethnic group. "Scapegoat OTHER" indicates that the Scapegoat comes from a different ethnic group (majority) than the Punisher. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group than the Punisher. Baseline and full controls are defined in Table 1. In Column 5, we exclude all subjects who did not answer all control questions correctly at the first or second attempt. The sample is composed of Punishers from the Roma ethnic minority group.

TABLE A15: PUNISHMENT OF THE WRONGDOER (ROMA MINORITY SAMPLE)

Dependent variable	Punishment of the Wrongdoer (intensity)	Punishment of the Wrongdoer (yes)	Punishment of the Wrongdoer (intensity) Punishment of the Wrongdoer = yes	Punishment of the Wrongdoer (intensity)	Punishment of the Wrongdoer (intensity)
Sample	All (1)	All (2)	All (3)	All (4)	All CQ correctly (5)
Harm intensity	0.19*** (0.03)	0.02*** (0.00)	0.26*** (0.04)	0.19*** (0.03)	0.21*** (0.03)
Wrongdoer OTHER	-0.09 (0.16)	-0.03 (0.03)	0.06 (0.26)	-0.15 (0.16)	-0.16 (0.16)
Harm intensity*Wrongdoer OTHER	0.06* (0.04)	0.01 (0.01)	0.06 (0.05)	0.06* (0.04)	0.09** (0.04)
Controls	baseline	baseline	baseline	full	baseline
Mean baseline (Wrongdoer SAME, 0 harm)	0.90	0.27	3.41	0.90	0.79
Observations	2,420	2,420	1,141	2,420	1,860
R-squared	0.073	0.032	0.142	0.153	0.094

Notes: OLS, standard errors (in parentheses) are clustered at a Punisher level. *** denotes $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. The dependent variable in Columns 1,3,4 is the extent of punishment towards the Wrongdoer (EUR 0-8). In Column 2, the dependent variable indicates that the Punisher chose non-zero punishment of the Wrongdoer. "Harm intensity" captures the harm caused by the Wrongdoer to the Victim (EUR 0-8). Punishers (and Victims) are from the Roma ethnic group. "Wrongdoer OTHER" indicates that the Wrongdoer comes from a different ethnic group (majority) than the Punisher. Baseline controls include gender and age of the Punisher, and a dummy variable indicating that the Punishing the Wrongdoer Game took place before the Punishing the Scapegoat Game. Full controls include also a dummy variable "Scapegoat OTHER" (equal to one if the Scapegoat comes from a different ethnic group than the Punisher), experimenter fixed effects, location fixed effects, education of parents (dummy variables for mother/father with at least secondary school education, dummy variables for unknown education), and a dummy variable indicating that the subject answered all control questions correctly at the first or second attempt. In Column 5, we exclude all subjects who did not answer all control questions correctly at the first or second attempt. The sample is composed of Punishers from the Roma ethnic minority group.

Working Paper Series

ISSN 1211-3298

Registration No. (Ministry of Culture): E 19443

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

(c) Michal Bauer, Jana Cahlíková, Julie Chytilová, Gérard Roland, Tomáš Želinský, 2021

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by

Charles University, Center for Economic Research and Graduate Education (CERGE)
and

Economics Institute of the CAS, v. v. i. (EI)

CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.

Printed by CERGE-EI, Prague

Subscription: CERGE-EI homepage: <http://www.cerge-ei.cz>

Phone: + 420 224 005 153

Email: office@cerge-ei.cz

Web: <http://www.cerge-ei.cz>

Editor: Byeongju Jeong

The paper is available online at http://www.cerge-ei.cz/publications/working_papers/.

ISBN 978-80-7343-504-2 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)

ISBN 978-80-7344-599-7 (Národohospodářský ústav AV ČR, v. v. i.)