Trust the Police? Self-Selection of Motivated Agents into the German Police Force

Guido Friebel, Michael Kosfeld, and Gerd Thielmann*

September 4, 2018

Abstract

We conduct experimental games with police applicants in Germany to investigate whether intrinsically motivated agents self-select into this type of public service. Our focus is on trustworthiness and the willingness to enforce norms of cooperation as key dimensions of intrinsic motivation in the police context. We find that police applicants are more trustworthy than non-applicants, i.e., they return higher shares as second-movers in a trust game. Furthermore, they invest more in rewards and punishment when they can enforce cooperation as a third party. Our results provide clear evidence for self-selection of motivated agents into the German police force, documenting an important mechanism that influences the match between jobs and agents in public service.

JEL classification: C9, D64, D73, J45

Keywords: Self-selection, intrinsic motivation, public service, trustworthiness, norm enforcement

*Friebel: Goethe University Frankfurt, CEPR, and IZA, gfriebel@wiwi.uni-frankfurt.de; Kosfeld: Goethe University Frankfurt, CEPR, CESifo, and IZA, kosfeld@econ.uni-frankfurt.de; Thielmann: Deutsche Hochschule der Polizei. We would like to thank Anna Lea Gründl, Wiebke Homann, Markus Kröll, Bernard Richter, and Sascha Wilhelm for excellent research assistance as well as the police academies of Hesse and Rhineland-Palatinate for valuable support in conducting this study. We are grateful to two anonymous referees as well as seminar participants at Goethe University Frankfurt, IZA, IAST-TSE, Columbia University, University of Vienna, QUT Brisbane, University of Adelaide, University of Toronto, University of Oslo, and briq for helpful comments and suggestions. The study was financially supported by the ANR and DFG through the project "Understanding Organisations – The Complex Interplay of Incentives and Identity". Kosfeld acknowledges support from the Alfons und Gertrud-Kassel Stiftung.

1 Introduction

When governments engage in the provision of public goods and services, they require agents to implement these policies. Many of these agents, tax inspectors, social workers, and police agents alike, are endowed with substantial authority over citizens. All too often, unfortunately, it cannot be assumed that public servants' incentives are fully aligned with the interests of the government or the public. The media remind us frequently about the abuse of delegated authority, be it in the context of police violence, red tape, or corruption.

Max Weber (1922) was aware of these threats to the legitimacy of what he called "Herrschaft", i.e., institutionalized authority, pointing to two requirements for legitimate public bureaucracy: intensive control mechanisms should be in place, and public servants should have a high degree of loyalty. Ex-post control mechanisms, however, are often expensive and involve the typical bureaucratic inefficiencies, famously described and analyzed by Crozier (1964), Wilson (1989), and Holmström and Milgrom (1991), amongst others. A natural complement are, therefore, *mechanisms that improve the ex-ante match* between jobs and public servants, i.e., bureaucracies and bureaucrats.[1]

In this paper, we study one such ex-ante mechanism focusing on a particular and important public bureaucracy, the police. Our main interest is in *self-selection*, by which citizens with a specific set of characteristics, including intrinsic motivation, are more likely to apply for the job in question. Our data comes from Germany, and we exploit a unique combination of incentivized behavioral experiments, survey data, and access to an exclusive pool of police applicants right at the time of submission of their application together with a natural group of comparison. The results provide clear evidence for the self-selection of intrinsically motivated agents into the German police force. The selection is advantageous, as the revealed motivation of police applicants is well in line with what is desired from a public interest perspective.

[1]More than 2000 years ago, in imperial China, candidates for the public service had to go through excruciating examinations that tested both applicants' skills and their willingness to provide high levels of effort (Miyazaki, 1979). By examining applicants' knowledge of Confucianism, the government sought to attract public servants who held high the values of imperial China, a crucial requirement given that public servants had authority over most of public life and the economy.

Exploring self-selection into public service in the police context is important and informative for economic research for two reasons. First, police misconduct and abuse of authority represents a critical problem in many countries in the world, being particularly high on the agenda, for example, in the U.S. and in Latin America. The police context, thus, provides a both relevant and interesting case for the main research question at hand. Secondly, it is relatively straightforward what to expect from a good police agent: he or she should be trustworthy and motivated to enforce the law and norms of cooperation (Goldstein 1977, Thielmann and Weibler 2014, Dharmapala et al. 2016). Both qualities are important not only to protect citizens against law violations and to safeguard citizens' cooperation with each other, but also to ensure and uphold citizens' willingness to trust and cooperate with the police, a condition that is key for effective crime detection and prevention.

Identifying self-selection in the field is not easy however. Studies based on employees (even beginners) suffer from the problem that participants have already passed the organization's screening process, have taken part in training programs, and have interacted with others in the same occupation, thereby making it difficult to disentangle self-selection from explicit sorting carried out by the organization, training effects, or social influence and peer effects. Often it is also unclear to whom employees from a given occupation should be compared, i.e., what is the relevant comparison group. In this respect, student samples seem advantageous. However, in many cases the evidence is limited to hypothetical job applications or job aspirations expressed in a survey.

We address these problems by collaborating with two state police agencies in Germany (Hesse and Rhineland-Palatinate) to contact all applicants who submitted their application to the respective police academies in a given recruitment period. These applicants have clearly documented their interest in becoming a police officer but have not yet passed any interview, been tested by the recruitment unit or been subject to any training. We additionally exploit the institutional feature that high school graduation (Abitur) forms a necessary requirement for becoming a police officer in these states, and create a natural comparison group composed of high school graduates from the same re-

gion and age cohort who have not applied for the police.[2] Finally, we employ incentivized experimental games to obtain reliable measures of our key variables of interest, trustworthiness and norm enforcement, in combination with a large survey on socio-economic and psychological covariates. The games we use are a trust game (Berg et al. 1995) and a reward-and-punishment game (Fehr and Fischbacher 2004). In the latter game, subjects play in the role of a third party who can reward or punish two other players who interact in a trust game.

Our results, which are based on about 1,400 individual observations, show that applicants to the police force are significantly more trustworthy than participants from the comparison group: they return on average higher shares in the role of the second mover in the trust game. The result is robust to the inclusion of important covariates. Differences in first-mover transfers, on the other hand, which are a measure of trust, are insignificant once we control for other covariates, in particular risk aversion. With regard to norm enforcement, we find that police applicants spend significantly more resources on both rewarding and punishing others as a third party. The result is, again, robust to the inclusion of controls. When controlling for participants' own trustworthiness, the police coefficient decreases in size, suggesting that own trustworthiness (which is higher among police applicants) plays an important role in norm enforcement. In sum, our results document a clear self-selection of "motivated agents" into the police force in Germany.

The paper complements and contributes to a variety of different strands in the literature. Several theoretical papers have analyzed the role of intrinsically motivated agents in organizations and public bureaucracies emphasizing the importance of finding the right match between public service occupations on the one hand, and motivated agents on the other (e.g., Francois 2000, Besley and Ghatak 2005, Delfgaauw and Dur 2008, Buurman and Dur 2012). Prendergast (2007) and Auriol and Brilon (2014) point to the problem that extreme types can also sort themselves into these organizations, providing an explanation for recurrent scandals involving, e.g., police violence or child abuse in aid organizations. McAdams et al. (2015) and Dharmapala et al. (2016) offer a discussion of

[2]Details on police requirements and the German school system are provided below.

related arguments from a law perspective. We contribute to this literature by providing robust empirical evidence for self-selection of intrinsically motivated agents in an important public service case. While we cannot rule out the existence of extreme types in our sample, results show that, on average, selection into the police is positive. In particular, the majority of police applicants seem to balance punishments and rewards.

Empirical papers studying self-selection have highlighted the effects of different incentive schemes both in the lab (Niederle and Vesterlund 2007, Eriksson et al. 2009, Dohmen and Falk 2011) and in the field (Dohmen and Falk 2010, Buser et al. 2014, Dal Bó et al. 2013, Ashraf et al. 2016). Carpenter and Myers (2010) and Hanna and Wang (2014) as well as Barfort et al. (2016) analyze the role of altruism and (dis-)honesty in selection into public service, using samples of U.S. volunteer firefighters and university students in India and Denmark, respectively. While these studies use experimental games like we do to measure motivation, the results are based on active volunteers or hypothetical job preferences only, thus making it difficult, for reasons explained above, to pin down actual self-selection. Banerjee et al. (2015) use a research design that is closer to ours comparing public and private sector aspirants in India.[3] However, their focus is on corruption in a framed lab experiment (cf. Alatas et al. 2009) rather than measures of trustworthiness or the willingness to enforce norms. Furthermore, participants are general aspirants for government administrative services, whereas our sample includes only candidates who explicitly apply for a job with the police. Serra et al. (2011) provide evidence for pro-social selection into the non-profit health sector in Ethiopia using both survey and behavioral measures. Next, our paper is related to Banerjee et al. (2012), who also study the police as an important case of public bureaucracy. Their analysis, however, does not consider self-selection, but rather the effects of different institutional reforms on police performance in Rajasthan, India. Dickinson et al. (2015) conduct an experimental lab study with trained police commissioners in France analyzing the use and efficacy of different norm enforcement institutions. With respect to the role of organizational factors in the performance of public bureaucracies, see also Rasul and Rogger (2016), who provide

[3]See also Banuri and Keefer (2016) for a related analysis with public sector aspirants in Indonesia using charitable donations as a proxy for prosocial motivation.

evidence on the importance of management practices based on data from public services delivery in Nigeria.

Finally, our paper adds to the growing literature showing that experimental games from the lab provide a valuable method to measure economic preferences in the field. See, e.g., Karlan (2005), Rustagi et al. (2010), Fehr and Leibbrandt (2011), Franzen and Pointner (2012), Kosfeld and Rustagi (2015) for social preferences, Chabris et al. (2008), Meier and Sprenger (2010, 2012), Sutter et al. (2013), Backes-Gellner et al. (2018) for time preferences, and Zhang (2012), Buser et al. (2014), Berge et al. (2015) for preferences for competition.

2 Police in Germany

According to a popular saying, in heaven, the mechanics are German, the chefs French, and the police British, while in hell, the mechanics are French, the chefs British, and the police German. In stark contrast, the police in modern Germany actually enjoys steadily high levels of trust from the German population, both inter-personal trust between citizens and police agents, and trust in the police as an institution.[4] More than 80 percent of the German population say they trust police agents (GfK 2016), a score that is 11 percentage points higher than in other European countries. According to a study by Forsa (2015), 84 percent rank the police as a trustworthy institution; the police takes the top rank, followed by universities, the own employer, courts and doctors. These data are in stark contrast with the U.S., for example, in which according to the Gallup (2015) poll, only one out of two citizens trust the police. It seems that modern Germany has succeeded in devising mechanisms that mitigate the problems associated with the delegation of power to the police. It thus appears an ideal setting to study the role of self-selection, in particular of trustworthy agents, as a potential ex-ante mechanism for achieving these results.

According to the German Constitution, police affairs are in the domain of the states, and only a few police tasks are allocated to the federal level (e.g., border control, railway

[4]This is also reflected in the popular saying "The police – your friend and aide", which almost every child in Germany grows up with.

police and international crime and terrorism). The sixteen state police agencies are responsible for all types of policing, i.e., patrol duty, traffic safety, crime prevention, crime control, and public security in daily life. These state agencies range from 2,800 to 42,000 employees, all of them tenured public servants. Our data come from two neighboring state police forces that cover the Rhine-Main area: Hesse and Rhineland-Palatinate.

Police in Germany are supposed to serve the citizen, not the state. This is reflected in the leadership philosophy, which highlights the importance of values and norms as well as trust and trustworthiness, performance, development, and motivation (Thielmann and Weibler, 2014). Entry barriers into the police are high. The two state police agencies we collaborated with only employ high school graduates with a university entrance qualification (Abitur) who, after a series of exams – measuring physical, psychological and cognitive fitness –, enter into a three-year education that is organized by the governmental universities of applied sciences, and graduate with a full-fledged bachelor degree in public administration/police management. Throughout their education, police agents are taught important psychological, sociological and legal foundations of police work, in addition to traditional self-defense and weapons training. Furthermore, there exists a highly selective masters program in the German Police University that the best career officers can apply to after a number of years in the police force, and which gives access to the top echelons of the hierarchy. Police agents in Germany are tenured public servants who are expected to work until the age of 62 to 64 after which they receive a generous pension and continuing health benefits.

Given the role and tasks of police in a democratic society like Germany, it is relatively straightforward what to expect from a police agent. In particular, police agents should be trustworthy, because otherwise the trust of citizens can neither be expected nor sustained. In addition, they should be "social peace keepers", i.e., cope with conflicts and mediate between conflicting interests of citizens. Finally, police agents should be motivated to punish wrongdoers, i.e., engage in the enforcement of law and norms of cooperation. This is important not only to protect citizens against crime and law violation but also to ensure cooperative and law-abiding behavior from citizens themselves. In the following

we analyze to what extent these characteristics are influenced by self-selection. In our attempt to analyze the presence of self-selection in terms of these characteristics, we abstract from direct criminal law enforcement and the safeguarding of public order and security. Rather, we focus on the enforcement of norms of cooperation. While there is clearly a distinction between legal and social norms, the two overlap and interact to a large extent in democratic societies.[5] Moreover, in many circumstances police officers do have considerable discretion in "applying the law", either because the precise legal circumstances can only be verified ex post or it is left to an officer's own judgement whether and how the police is called to act and intervene (Perez and Barkhurst, 2012).[6] Tyler and Huo (2002) show that citizens' personal experience with the police has an important effect on their willingness to accept legal processes and institutions (cf. also Tyler, 2006). Finally, there exists an established experimental economic literature on the enforcement of norms of cooperation (e.g., Fehr and Gächter 2000, Fehr and Fischbacher 2004, Henrich et al. 2006) enabling us to analyze the potential self-selection along "punitive preferences" properly (cf. Dharmapala et al. 2016).[7]

3 Experimental Set-up

Our research strategy for identifying self-selection of trustworthy and norm-enforcing individuals into the police force exploits the fact that high school graduation (Abitur) forms a necessary requirement for becoming a police officer in Hesse and Rhineland-Palatinate. In parallel to the recruitment of applicants to the police, we therefore recruited high school students from the same region and same main graduation cohort as a natural group of

[5] See, e.g., the literature on social norms and the law, in particular the special issue of the *Journal of Legal Studies*, vol. 27 (1998) or the work by Ellickson (1991), Posner (2000), and more recently Hadfield and Weingast (2012, 2014).

[6] Consider, for example, – as it actually happened to one of the authors (GT) – that a police officer is called to a house at midnight, where he is told by a man that he has been robbed and that he has followed the robber to this house where he saw a light going on on the second floor after the person had entered. On the one hand, the man might tell the truth and the robbery might be significant such that the police officer should enter the house to identify the robber and seize the stolen good. On the other hand, it is also possible that the man is mistaken, or even tells a blunt lie, in which case entering the house and disturbing the innocent inhabitant at night would be unlawful.

[7] Baldassari and Grossman (2011) and Faillo et al. (2013) analyze the role of legitimacy of punishment institutions in the lab showing positive effects on cooperation.

comparison.[8] In the following, we first describe the two experimental games that were played by both police applicants and high school students. We then provide detailed information about the procedures we used for recruiting the two groups of participants and for conducting the experiment.

The experiment consisted of two parts that were followed by a survey. In the first part, participants played a trust game as illustrated in Figure 1. In the second part, participants played a reward-and-punishment game. We explain each game successively.[9]

3.1 Trust Game

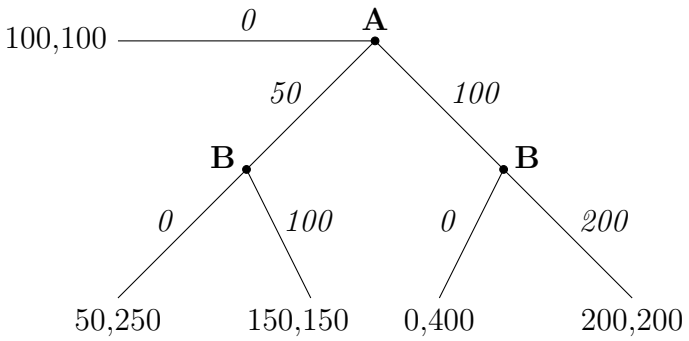

Figure 1: Trust game

In the trust game, there are two players (A and B), each endowed with 100 €. Player A decides first whether to transfer 0, 50, or 100 € to player B. The transfer is tripled, i.e., depending on player A's choice, player B receives either 0, 150, or 300 €. Player B then decides whether to keep the entire transfer or share the returns equally with player A. In the case where player A transfers 50, B can either return 0 (leading to a payoff of 50 € for player A and 250 € for player B) or 100 (leading to a payoff of 150 € for both). If player A transfers 100, B can either return 0 (leading to a payoff of 0 € for A and 400 € for B) or 200 (leading to 200 € for both). In the case where player A transfers 0, player B makes no choice and both players earn their initial endowment of 100 €. Because B has no monetary incentive to share the returns with A in this game (players interact only

[8] Broadly speaking, the secondary school system in Germany comprises three levels: Hauptschule, Realschule, and Gymnasium, the latter offering the possibility to obtain general qualification for university entrance (Abitur). All participants in our study posses this qualification.

[9] See `https://nextcloud.wiwi.uni-frankfurt.de/nextcloud/index.php/s/wQsWsbcLjWmsadq` for experimental instructions. Homann (2012) and Richter (2013) provide additional details about the set-up.

once and anonymously in the experiment), positive transfers from A are interpreted as a measure of the trust player A places in B; similarly, positive backtransfers from B are a measure of the trustworthiness B reveals towards A (cf. Berg et al. 1995, Cox 2004).

All participants played the trust game in both player roles with different anonymous partners, i.e., in particular police applicants were not matched with each other.[10] Therefore, behavior in the trust game is a measure of generalized trust and trustworthiness, which is exactly what we are interested in. We used the strategy method to elicit the choices of player B, i.e., participants in this role decided about their backtransfers *conditional* on player A transferring either 50 or 100 €. This allowed us to elicit the behavior of B for every participant without having to wait for another participant to first make his decision as A. In addition, we obtain a comparable measure of trustworthiness for all participants, which would not have been the case if participants had decided for endogenous, i.e., different, transfer levels of A. All participants correctly answered a set of control questions before they made their decisions in the trust game.

3.2 Reward-and-Punishment Game

The reward-and-punishment game that was played in the second part of the experiment is based on the trust game. The new and distinct feature is that a third player (C) is added, who can reward or punish players A and B conditional on the two players' decisions. Both rewarding and punishing is costly to player C and yields no material benefit. Specifically, player C is endowed with 160 €; after player A and B have made their decisions, player C decides whether to allocate so-called reward or punishment points to any of the two players. Each point that is allocated to a particular player increases or decreases that particular player's payoff by 2 € and at the same time decreases player C's payoff by 1 €. The minimum to which a player's payoff can be decreased is zero, i.e., players A and B cannot make losses. Since we are interested in police applicants' willingness to enforce norms of cooperation, police applicants and high school students were always in the role of player C in this game. The decisions of players A and B were made by students from

[10] Payment rules are explained in Section 3.3.

the FLEX subject pool (see below). We again used the strategy method to collect player C's reward and punishment decisions conditional on all five possible outcomes in the trust game. All participants also correctly answered a set of control questions before they made their decisions.

For two reasons the reward-and-punishment game was always played after the trust game: First, instructions in the reward-and-punishment game are much easier to understand when participants have played the trust game before; thereby, the design reduces any noise that might otherwise be caused by confusion, which is particularly important as the experiment was conducted online (see procedures below). Second, we explicitly wanted participants to go through the strategic situation of the trust game, in order to allow them to make an informed decision as player C in the reward-and-punishment game. For this reason, we also engaged participants in both player roles in the trust game in order to avoid biased decision-making by player C towards either of the two players. Further, we reduced the strategy space of both players in the trust game, as this simplified the decision making of player C in the reward-and-punishment game. Importantly, participants did not learn any outcome of the trust game before they played the reward-and-punishment game.

3.3 Procedures

The experiment was conducted online between fall 2010 and fall 2011 via a secure online server at the Frankfurt Laboratory for Experimental Economic Research (FLEX) at Goethe University. Police applicants were contacted via police academies in Hesse and Rhineland-Palatinate after they had submitted their application to the respective academy. Together with the letter of acknowledgment from the academy that their application had been received, each applicant was sent an invitation from our research team to take part in a study on decision-making and attitudes of job applicants. The invitation was framed neutrally with no emphasis on the police as a particular employer. No information was given about the games to be played in the experiment. We carefully explained in the invitation that the study was an independent research project of Goethe University

and that there was no connection whatsoever to the academy besides the latter's support in sending out the invitation. Further, we emphasized that it was impossible for the research team to link any data from the experiment to any personnel records (which we did not have).[11] Finally, the web interface in which applicants made their decisions was also framed in a neutral manner and did not include any information or links to the police or to the application process. We are confident that these measures minimized any potential demand effects.

Applicants were informed that at the end of the experiment, fifteen participants (ten in the trust game, five in the reward-and-punishment game) would be randomly selected for payment. These participants were randomly assigned to player roles and pairs and were paid out their individual earnings depending on the particular decisions in the game.[12] To participate in the experiment, each applicant received an individual code that had to be entered on the FLEX website and that could be used only once. Due to different timing in the recruitment of police academies, applicants participated in two waves: Applicants in Hesse participated from October 2010 to January 2011, applicants in Rhineland-Palatinate participated between July and October 2011. In total, 617 police applicants completed the experiment. 158 of them did not have or expect a high-school degree and thus did not meet the necessary qualification requirement. They would therefore not be selected by the police, so we exclude them. This leaves us with a sample of 459 police applicants.

For the comparison group of high school students, we contacted 75 public high schools (Gymnasien) in Hesse. These schools were randomly selected out of the full sample of all 224 public high schools in Hesse using a geographic stratification procedure based on zip codes. 42 of the schools we contacted agreed to participate. In each of these schools, students received an invitation to take part in the study that was distributed via their main teacher. As for police applicants, students were informed about the general purpose of the study (but not about the games to be played) and the possibility to earn money

[11] Although it would have been tempting to follow applicants through the screening and training process, we explicitly decided against this possibility in order to make sure that participants are full-informed that their decisions are anonymous and cannot be matched with personal records. For a complementary study on trained police commissioners see Dickinson et al. (2015).

[12] The average payout was 150 €. In addition, three iPod Nanos were raffled among all participants.

in the experiment as well as to participate in a raffle of three iPod Nanos. Again, fifteen participants were randomly selected for payment. Each student received an individual code that had to be entered on the FLEX website and that could be used only once. 959 high school students completed the experiment.[13]

Finally, students from the FLEX subject pool at Goethe University filled up the remaining player roles A and B in the reward-and-punishment game to determine payment of police applicants and high school students in this game.

4 Results

Table 1 displays descriptives of participants' behavior in the two experimental games.[14] As can be seen, police applicants make, on average, about nine percent higher transfers as player A in the trust game (57.52 compared to 52.76 €, Mann-Whitney test: $p < 0.01$), and also return on average around eight percent higher backtransfers as player B compared to the group of high school students (81.70 compared to 74.56 € in the case where A transfers 50, Mann-Whitney test: $p < 0.01$; 160.35 compared to 149.74 € in the case where A transfers 100, Mann-Whitney test: $p < 0.05$).[15] Thus, based on raw data, police applicants appear both more trusting and more trustworthy. A similar picture emerges in the reward-and-punishment game. On average, compared to high school students police applicants invest about twelve percent more resources as player C on rewarding and punishing players A and B (40.56 compared to 36.26 €, Mann-Whitney test: $p < 0.01$). This shows that police applicants are not only more cooperative in the trust game, they are also more willing to enforce norms of cooperation as a third party. In the Appendix we disaggregate player C's decisions in more detail and show that police applicants and high school students exhibit similar reward and punishment preferences, suggesting that

[13] Unfortunately, it is difficult to calculate exact response rates as we do not have precise information how many of the invitations we provided were actually distributed. Based on our records, we estimate the number of invitations that were distributed to police applicants and high school students as 5,312 and 4,522, respectively, yielding a response rate of about 12 and 21 percent. One plausible reason why the response rate at high schools is likely to be higher is that invitations were distributed via teachers during classes, which might have influenced the motivation to participate positively.

[14] See Table 7 in the Online Appendix for distributions of trust and trustworthy behavior.

[15] All tests reported in the paper are two-tailed.

differences in player C's investments are primarily a level effect. We therefore concentrate on participants' average investment in rewards and punishment in our analysis.

Table 1: Behavior in the Experimental Games

	High school students (SD)	Police applicants (SD)
Trust A	52.76	57.52
	(32.47)	(34.23)
Trustworthiness B if A sends 50	74.56	81.70
	(43.58)	(38.71)
... if A sends 100	149.74	160.35
	(86.80)	(79.82)
Average reward and punishment C	36.26	40.56
	(25.06)	(26.57)
N	959	459

Note: Averages with standard deviations in parentheses. Average reward and punishment C is the average of reward and punishment points (in €) allocated by player C to players A or B in the five possible outcomes in the trust game.

We next analyze whether there also exist other differences in observables between police applicants and high school students. It seems very likely that police applicants and high school students differ from each other in additional dimensions. Besides our main criterion of comparison, school-leaving qualification, which we control, there is no reason to believe that both groups should be completely identical. Most likely, selection into the police is multi-dimensional. Based on the survey which we conducted at the end of the experiment, we are able to identify relevant differences between the two groups. The survey comprises key socio-economic variables such as age, gender, income, parents' education, migration background, risk preference, as well as important psychological personality measures (proactivity, sensation seeking). As Table 2 shows, police applicants differ quite a bit from the group of high school students along a number of these dimensions.[16] In particular, police applicants are, on average, more likely to be male, they are both slightly older and taller, more risk tolerant, and they score higher in both proactivity

[16]See also Table 8 in the Online Appendix. The number of observations in these tables (and also in Table 3 and 4 below) is smaller than in Table 1 because of non-responses to parts of the survey.

and sensation seeking. In the following, we will use linear regressions to analyze the role of these variables in the observed differences in behavior in the two experimental games.

Table 2: Differences in Other Covariates

	High school students (SD)	Police applicants (SD)
Female share	0.65	0.40
	(0.48)	(0.49)
Age	19.55	20.48
	(0.89)	(2.60)
Body height	172.87	176.14
	(9.05)	(8.38)
Risk preference	5.04	5.47
	(1.74)	(1.64)
Proactivity	33.75	35.98
	(6.38)	(5.58)
Sensation seeking	49.82	52.53
	(8.94)	(7.10)
N	877	459

Note: Body height in cm; risk preference is the general risk question from the German Socio-Economic Panel (SOEP); proactivity is based on Frese et al. (1997), sensation seeking on NISS (Roth and Hammelstein 2012).

Tables 3 and 4 summarize our main results. We start with the simplest specification and only include a dummy for being a police applicant. We then add control variables, first gender and age, then the additional covariates from Table 8 in the Online Appendix.[17] Column (1) and (2) in Table 3 corroborate our finding above that police applicants are significantly more trusting. Once we control for additional covariates — in particular, risk preferences —, however, this difference becomes insignificant (column (3)). At the same time, a participant's risk preference is found to be significantly associated with trusting behavior, which is intuitive as trust obviously requires the willingness to take risk (of being exploited by the other player). With regard to trustworthiness, columns (4) to (9) in Table 3 show that police applicants are significantly more trustworthy than high

[17]See Table 9 and 10 in the Online Appendix for coefficients of the full set of covariates. Results are qualitatively similar if we use (ordered) probit or a nearest neighbor matching model instead (available upon request).

school students, both without and with additional controls. Risk preferences play no role in this decision, which makes sense as player B's backtransfer decision does not include any risk but is purely distributional. Together these result document our first important dimension – trustworthiness – along which self-selection into the German police can be identified.

The next question is whether police applicants also invest significantly more resources on punishing and rewarding players as a third party. Table 4 shows that this is indeed the case.[18] Column (1) reveals a positive and significant association of being a police applicant with total investment as player C without inclusion of any controls. The association remains significant if we add gender and age (column (2)) as well as additional covariates (column (3)). In columns (4) to (7) we include participants' own behavior in the trust game, either as player A (trust) or as player B (trustworthiness). The rationale is that these variables reflect individual differences in perceived norms of cooperation which are likely to serve as a role model for taking decisions in the reward-and-punishment game. Participants who trust more and/or are trustworthy themselves may be more inclined to reward cooperation as well as punish non-cooperation, relative to participants who act as pure money maximizers in the trust game.

As column (4) shows, adding own trust does not change the size of the police dummy very much. The coefficient becomes slightly smaller but remains significant at the five percent level. Adding own trustworthiness (column (5)) reduces the coefficient further, such that the association is now significant at the ten percent level, while own trustworthiness is highly significant. In this specification we measure trustworthiness separately for the two cases where player A sends either 50 or 100. Results are the same if we combine the two cases (not shown). A similar result is obtained if we include trust and trustworthiness together (column (6)). Finally, in column (7), we combine trust and trustworthy behavior into two behavioral types: "NoTrust_NotTrustworthy" individuals, who do not trust and return zero back transfers, and "Trust_Trustworthy" individuals, who trust (i.e., send either 50 or 100) and return equal shares. As can be seen,

[18]Since the share of censored observations (at 0 or 160) is small, we present results based on OLS. Tobit regressions produce similar results and are available upon request.

Table 3: Trust and Trustworthiness

	Trust A			Trustworthiness B if A sends 50			Trustworthiness B if A sends 100		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Police applicant	0.095**	0.070*	0.061	0.071***	0.074***	0.068**	0.053**	0.079***	0.086***
	(0.038)	(0.040)	(0.043)	(0.024)	(0.025)	(0.027)	(0.024)	(0.025)	(0.027)
Gender		−0.060	0.056		0.021	0.037		0.082***	0.117***
		(0.037)	(0.056)		(0.024)	(0.035)		(0.024)	(0.035)
Age		0.010	0.010		0.003	−0.000		−0.005	−0.006
		(0.011)	(0.012)		(0.007)	(0.007)		(0.007)	(0.007)
Risk preference			0.024**			−0.005			−0.008
			(0.012)			(0.007)			(0.007)
Constant	2.055***	1.890***	0.451	1.746***	1.678***	1.480***	1.749***	1.791***	1.306***
	(0.021)	(0.216)	(0.604)	(0.014)	(0.137)	(0.383)	(0.014)	(0.138)	(0.384)
Additional Controls	NO	NO	YES	NO	NO	YES	NO	NO	YES
N	1,418	1,418	1,331	1,418	1,418	1,331	1,418	1,418	1,331

Note: OLS with standard errors in parentheses. The dependent variables are the transfer sent as player A (trust A) and the backtransfers sent as player B (trustworthiness B) for the two possible cases. Additional controls include log(income), type of city grown up in, migration dummy, education of father and mother, sensation seeking (NISS), proactivity, willingness to take part in lotteries, risk preference. *** p<0.01, ** p<0.05, * p<0.1.

"NoTrust_NotTrustworthy" types invest significantly less in reward and punishment as a third party, while "Trust_Trustworthy" types invest significantly more. In this specification, the police applicant dummy becomes marginally significant at the five percent level ($p = 0.051$).

In sum, the above results provide robust evidence for the self-selection of motivated agents into the German police force. This self-selection can be observed along two key dimensions: First, police applicants are significantly more trustworthy, i.e., they return higher shares as a second mover in a trust game; second, police applicants invest more resources in rewarding and punishing other players as a third party. Given that third-party investments are shown to be significantly associated with own trustworthiness, the results further suggest that police applicants' higher trustworthiness plays an important role also in norm enforcement.

Can these results be interpreted as evidence for why citizens' trust in the police in Germany is particularly high? Homann (2012) analyzes this question by running a lab experiment in which students play the above trust game under two treatment conditions. In both treatments, students are in the role of the trustor (player A) and the game is played one shot. In the baseline treatment, students are told that they are matched with an anonymous stranger. In the other treatment, they are informed that they are matched with a police applicant. If citizens' trust in the police is indeed a reflection of police trustworthiness, this should be observed in the data.[19] The results corroborate the hypotheses. If students are matched with a police applicant, trust is higher, both in terms of higher average transfers (52.53 vs. 46.96 €) and in terms of higher average expected returns (if A transfers 50: 63.13 vs. 50.93 €, if A transfers 100: 108.76 vs. 83.18 €). Thus, police forces in Germany do not only benefit from a positive self-selection of trustworthy individuals, citizens also anticipate this, as is reflected by students' behavior in Homann's study. In other words, Germans trust their police not because trust levels are generally

[19]Note that this hypothesis is based on the assumption that students' expectations about police applicants' trustworthiness are representative for their expectations towards actual police agents. Given that police education in Germany is likely to have a positive influence on police agents' trustworthiness (cf. Section 2), the results thus provide a lower bound for the relationship between citizens' trust and police trustworthiness.

Table 4: Reward and Punishment

	Average Investment C						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Police applicant	4.300***	4.192***	4.215**	3.673**	2.876*	2.813*	3.173*
	(1.451)	(1.536)	(1.682)	(1.640)	(1.638)	(1.619)	(1.626)
Gender		−0.314	−3.107	−3.602*	−4.543**	−4.562**	−4.643**
		(1.429)	(2.159)	(2.104)	(2.104)	(2.080)	(2.089)
Age		0.029	−0.113	−0.199	−0.052	−0.125	−0.020
		(0.417)	(0.450)	(0.439)	(0.437)	(0.432)	(0.434)
Trust				8.851***		6.170***	
				(1.043)		(1.104)	
Trustworthiness50					6.975***	5.140***	
					(1.984)	(1.989)	
Trustworthiness100					10.027***	7.825***	
					(1.981)	(1.998)	
NoTrust_NotTrustworthy							−13.694***
							(2.690)
Trust_Trustworthy							9.386***
							(1.534)
Constant	36.260***	35.905***	68.525***	64.537***	45.108**	47.920**	73.039***
	(0.825)	(8.350)	(23.487)	(22.883)	(22.927)	(22.674)	(22.639)
Additional Controls	NO	NO	YES	YES	YES	YES	YES
N	1,418	1,418	1,331	1,331	1,331	1,331	1,331

Note: OLS with standard errors in parentheses. The dependent variable is the average of reward and punishment points allocated by player C to players A or B in the five possible outcomes in the trust game. Additional controls include log(income), type of city grown up in, migration dummy, education of father and mother, sensation seeking (NISS), proactivity, participation in lotteries, risk preference. *** p<0.01, ** p<0.05, * p<0.1

high compared to other countries, but because those who want to become a police officer in Germany are particularly trustworthy and citizens expect this.

5 Conclusion

Police forces in democratic societies have to strike a subtle balance between protecting the citizen and enforcing the law. Being close to the citizen is a crucial determinant of police efficiency (Blanes i Vidal and Kirchmaier 2016). In order for citizens to be willing to cooperate, the police must be trustworthy, rather than exploiting people's trust. Trust between the police and the citizens they serve is a cornerstone of civilized societies.

Our study has looked at police applicants in two states of Germany who have not even passed any test or interview but have simply revealed their interest in a job with the police by submitting their CV. We found that these applicants feature substantially higher levels of trustworthiness and are willing to invest more of their resources into rewarding trustworthy behavior and punishing the abuse of trust. Our results suggest that the police agencies investigated here manage to attract people with a good-citizen type of motivation which is appreciated in many organizations, but particularly important for the police force. They also suggest that given that ex-post control of public service employees is quite cumbersome, reform of public bureaucracies may rely to a substantial degree on improving the matching process between open positions and candidates.

Of course, the extent to which self-selection is expected to be advantageous elsewhere and also more generally depends on a number of critical factors. Most importantly, institutional factors have to be taken into account such as, for example, the precise role and tasks public service agents have in society (and citizens' expectations thereof) as well as the combination of rewards and benefits together with career and training possibilities (Ashraf et al. 2016). Police agencies in Germany seem to have managed to govern this matching mechanism successfully, as is reflected by our study.

More precisely, we would conjecture that the effects we document for Germany would be more likely to be found in democratic societies in which the police is in general not perceived as an oppressive force of the government against the citizens but rather as

an enforcement organ for law and order and a social peace keeper (cf. Section 2). In democratic societies both trustworthiness and willingness to enforce norms is conducive to the functioning of police. Democracy, however, may only be a necessary but not a sufficient condition for positive self-selection to occur. Police forces in the U.S., for example, tend to enjoy less trust by citizens, and the police uses much more violence in executing law than in Germany (and other countries such as France, Sweden, Britain or Japan). This indicates that there may be multiple equilibria in the interaction between citizens and the police in which some police forces may be in a high-trust equilibrium whereas others may be in an equilibrium with only low trust.[20]

Indeed, the German police force went through a number of reforms in the post-World War II period. The police functions had been heavily centralized in the Nazi era, while after the war, a separation of police functions was carried out (for instance, day-to-day policing versus protection against politically motivated crimes), and the police forces were decentralized to the state level (Kaiser et al., 2012). With the advent of the 1968 generation and, arguably, changed preferences in the German society, the police force was further reformed.[21] An important element in this reform strategy was an investment in education, in particular in terms of legal knowledge, social and leadership skills with a view to transform the police to a citizens' police (Thielmann and Weibler, 2014). This process culminated in the fact that in the two states we investigate (but also in many others), to apply for the police, a high school diploma is required, and every police agent must accomplish a BA degree. Obviously, the three-year education provided by these German police forces, constitutes a contrast to many U.S. state police which offer much shorter training spells that, except for career officers, are more geared towards self-defense and the use of weapons. It is possible that better education may make people more trustworthy and more motivated in enforcing norms of cooperation, but our experiment would only capture the willingness of people to be educated not the effect of education. We are confident though that education has helped in selection but it is not the entire

[20] While it is tempting to draw conclusions for the U.S., there are at least two notable differences in Germany: race issues are much weaker in Germany, and citizens usually have no right to carry fire arms.

[21] A similar observation can be made today when the police in Germans is trying to increase the share of officers with migration background as a response to the higher number of immigrants in the population.

story. Rather we would point to education being part of the cultural change and not tantamount to it.

In general, to what extent institutional and cultural constraints are important in explaining the magnitude to which our effects are present in other countries is a fascinating question that would need a similar design in a cross-country perspective, a challenge we would like to tackle in future research.

References

Alatas, Vivi, Lisa Cameron, Ananish Chaudhuri, Nisvan Erkal, and Lata Gangadharan. 2009. "Subject pool effects in a corruption experiment: A comparison of Indonesian public servants and Indonesian students." *Experimental Economics* 12 (1): 113–132.

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee. 2016. "Do-gooders and go-getters: Selection and performance in public service delivery." Working Paper.

Auriol, Emmanuelle and Stefanie Brilon. 2014. "Anti-social behavior in profit and nonprofit organizations." *Journal of Public Economics* 117: 149–161.

Backes-Gellner, Uschi, Holger Herz, Michael Kosfeld, and Yvonne Oswald. 2018. "Do preferences and biases predict life outcomes? Evidence from education and labor market entry decisions." Working Paper.

Baldassarri, Delia and Guy Grossman. 2011 "Centralized sanctioning and legitimate authority promote cooperation in humans." *Proceedings of the National Academy of Sciences* 108 (27): 11023–11027.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh. 2012. "Improving police performance in Rajasthan, India: Experimental evidence on incentives, managerial autonomy and training." NBER Working Paper 17912.

Banerjee, Ritwik, Tushi Baul, and Tanya Rosenblat. 2015 "On self selection of the corrupt into the public sector." *Economics Letters* 127: 43–46.

Banuri, Sheheryar and Philip Keefer. 2016. "Pro-social motivation, effort and the call to public service." *European Economic Review* 83: 139–164.

Barfort, Sebastian, Nikolaj Harmon, Frederik Hjorth, and Asmus Leth Olsen. 2016. "Sustaining honesty in public service: The role of selection." Working Paper.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, reciprocity, and social history." *Games and Economic Behavior* 10 (1): 122–142.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Armando J. Garcia Pires, and Bertil Tungodden. 2015. "Competitive in the lab, successful in the field?" *Journal of Economic Behavior & Organization* 118: 303–317.

Besley, Timothy J. and Maitreesh Ghatak. 2005. "Competition and incentives with motivated agents." *American Economic Review* 95 (3): 616–636.

Blanes i Vidal, Jordi and Tom Kirchmaier. 2016. "The effect of police response time on crime clearance rates." Working Paper.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, competitiveness, and career choices." *Quarterly Journal of Economics* 129 (3): 1409–1447.

Buurman, Margaretha and Robert Dur. 2012. "Incentives and the sorting of altruistic agents into street-level bureaucracies." Scandinavian Journal of Economics 114 (4): 1318–1345.

Carpenter, Jeffrey and Caitlin K. Myers. 2010. "Why volunteer? Evidence on the role of altruism, image, and incentives." *Journal of Public Economics* 94 (11-12): 911–920.

Chabris, Christopher F., David I. Laibson, Carrie L. Morris, Jonathon P. Schuldt, and Dmitry Taubinsky. 2008. "Individual laboratory-measured discount rates predict field behavior." *Journal of Risk and Uncertainty* 37 (2-3): 37–269.

Cox, James C.. 2004. "How to identify trust and reciprocity." *Games and Economic Behavior* 46 (2): 260–281.

Crozier, Michel. 1964. *The bureaucratic phenomenon*. Chicago: Chicago University Press.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening state capabilities: The role of financial incentives in the call to public service." *Quarterly Journal of Economics* 128 (3): 1169–1218.

Dharmapala, Dhammika, Nuno Garoupa, and Richard H. McAdams. 2016. "Punitive police? Agency costs, law enforcement, and criminal procedure." *Journal of Legal Studies* 45: 105-141.

Delfgaauw, Josse and Robert Dur. 2008. "Incentives and workers' motivation in the public sector." *Economic Journal* 118: 171–191.

Devlin, Matthew. 2010. "Seizing the Reform Moment: Building Georgia's Police, 2004-2006." Innovations for Successful Societies Policy Note 126, Princeton University.

Dickinson, David L., David Masclet, and Marie Claire Villeval. 2015. "Norm enforcement in social dilemmas: An experiment with police commissioners." *Journal of Public Economics* 126: 74–85.

Dohmen, Thomas and Armin Falk. 2010. "You get what you pay for: Incentives and selection in the education system." *Economic Journal* 120: F256-F271.

Dohmen, Thomas and Armin Falk. 2011. "Performance pay and multidimensional sorting: Productivity, preferences, and gender." *American Economic Review* 101 (2): 556–590.

Ellickson, Robert C.. 1991. *Order without law: How neighbors settle disputes.* Harvard University Press.

Eriksson, Tor, Sabrina Teyssier, and Marie-Claire Villeval. 2009. "Self-selection and the efficiency of tournaments." *Economic Inquiry* 47 (3): 530–548.

Faillo, Marco, Daniela Grieco, and Luca Zarri. 2013. "Legitimate punishment, feedback, and the enforcement of cooperation." *Games and Economic Behavior* 77 (1): 271–283.

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. 2000. "Cooperation and punishment in public goods experiments." *American Economic Review* 90 (4): 980–994.

Fehr, Ernst and Andreas Leibbrandt. 2011. "A field study on cooperativeness and impatience in the Tragedy of the Commons." *Journal of Public Economics* 95 (9-10): 1144–1155.

Forsa. 2015. "Polizisten geniessen Vertrauen — Manager nicht-" http://www.stern.de/politik/deutschland/stern-umfrage–wem-vertrauen-die-deutschen-3479178.html

Franzen, Axel and Sonja Pointner. 2013. "The external validity of giving in the dictator game." *Experimental Economics* 16 (2): 155–169.

Francois, Patrick. 2000. "'Public service motivation' as an argument for government provision." *Journal of Public Economics* 78 (3): 275–299.

Frese, Michael, Doris Fay, Tanja Hilburger, Karena Leng, and Almut Tag. 1997. "The concept of personal initiative: Operationalization, reliability and validity in two German samples." *Journal of Organizational and Occupational Psychology* 70: 139–161.

Gallup. 2015. "In U.S., confidence in police lowest in 22 years." http://www.gallup.com/poll/183704/confidence-police-lowest-years.aspx

GfK. 2016. "Worldwide ranking: trust in professions." http://www.gfk-verein.org/en/compact/focustopics/worldwide-ranking-trust-professions

Goldstein, Herman. 1977. *Policing a free society.* Cambridge: Ballinger Publishing Company.

Hadfield, Gillian K. and Barry R. Weingast. 2012. "What is law? A coordination model of the characteristics of legal order." *Journal of Legal Analysis* 4 (2): 471–514.

Hadfield, Gillian K. and Barry R. Weingast. 2014. "Microfoundations of the rule of law." *Annual Review of Political Science* 17 (1): 21–42.

Hanna, Rema and Shing-Yi Wang. 2014. "Dishonesty and selection into public service: Evidence from India." Working Paper.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, et al.. 2006. "Costly punishment across human societies." *Science* 312: 1767–1770.

Herrmann, Benedikt, Christian Thöni, and Simon Gächter. 2008. "Antisocial punishment across societies." *Science* 319: 1362-1367.

Holmström, Bengt and Paul Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *Journal of Law, Economics, & Organization* 7: 24–52.

Homann, Wiebke. 2012. *Essays in behavioural economics: Behaviour in the job and in unemployment*. PhD thesis. Faculty of Economics and Business Administration, Goethe University Frankfurt.

Kaiser, Wolf, Thomas Köhler, and Elke Gryglewski. 2012. *"Nicht durch formale Schranken gehemmt" – Die deutsche Polizei im Nationalsozialismus*. Bonn: Bundeszentrale für politische Bildung und Deutsche Hochschule der Polizei.

Karlan, Dean S.. 2005. "Using experimental economics to measure social capital and predict financial decisions." *American Economic Review* 95 (5): 1688–1699.

Kosfeld, Michael and Devesh Rustagi. 2015. "Leader punishment and cooperation in groups: Experimental field evidence from commons management in Ethiopia." *American Economic Review* 105 (2): 747-783.

McAdams, Richard H., Dhammika Dharmapala, and Nuno Garoupa. 2015. "The law of police." *University of Chicago Law Review* 82: 135–158.

Meier, Stephan and Charles D. Sprenger. 2010. "Present-biased preferences and credit card borrowing." *American Economic Journal: Applied Economics* 2 (1): 193–210.

Meier, Stephan and Charles D. Sprenger. 2010. "Time discounting predicts creditworthiness." *Psychological Science* 23 (1): 56–58.

Miyazaki, Ichisada. 1979. *China's examination hell: The civil service examinations of imperial China*. New Haven: Yale University Press.

Niederle, Muriel and Lise Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *Quarterly Journal of Economics* 122 (3): 1067–1101.

Perez, Douglas W. and Barkhurst, Michael. 2012. *Paradoxes of leadership in police management*. Delmar.

Prendergast, Canice. 2007. "The motivation and bias of bureaucrats." *American Economic Review* 97 (1): 180–196.

Posner, Eric A.. 2000. *Law and social norms*. Harvard University Press.

Rasul, Imran and Daniel Rogger. Forthcoming. "Management of bureaucrats and public service delivery: Evidence from the Nigerian civil service." *Economic Journal*.

Richter, Bernard. 2013. *Essays in behavioral economics — evidence on self-selection into jobs, social networks and leniency*. PhD thesis. Faculty of Economics and Business Administration, Goethe University Frankfurt.

Roth, Marcus and Hammelstein, Philipp. 2012. "The need inventory of sensation seeking (NISS)." *European Journal of Psychological Assessment* 28: 11–18.

Rustagi, Devesh, Stefanie Engel, and Michael Kosfeld. 2010. "Conditional cooperation and costly monitoring explain success in forest commons management." *Science* 330 (6006): 961–965.

Serra, Danila, Pieter Serneels, and Abigail Barr. 2011. "Intrinsic motivation and the non-profit health sector: Evidence from Ethiopia." Personality and Individual Difference 51: 309–314.

Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, Stefan T. Trautmann. 2013. "Impatience and uncertainty: Experimental decisions predict adolescents' field behavior." *American Economic Review* 103 (1): 510–531.

Thielmann, Gerd and Jürgen Weibler. 2014. *Polizeiliche Führungslehre*. Hilden: Verlag Deutsche Polizeiliteratur.

Tyler, Tom R.. 2006. *Why people obey the law*. Princeton University Press.

Tyler, Tom R. and Yuen J. Huo. 2002. *Trust in the law*. Russel Sage Foundation.

Weber, Max. 1922, 1978. *Economy and society: An outline of interpretive sociology*. University of California Press.

Wilson, James Q.. 1989. *Bureaucracy: What government agencies do and why they do it*. Basic Books.

Zhang, Y Jane. 2012. "Can experimental economics explain competitive behavior outside the lab?" Working Paper.

Appendix: Norm Enforcement Preferences

Based on the strategy method we used for player C in the reward-and-punishment game, we can analyze whether police applicants reveal different *preferences* for norm enforcement than high school students. For example, it might be that police applicants have a stronger preference for punishment than for rewarding. Alternatively, police applicants might be less likely to not punish or reward at all. In the following, we focus on player C's reward and punishment decisions targeting player B, because this player's behavior is readily interpretable in terms of cooperation or non-cooperation: if player B shares the returns equally, he cooperates; if he keeps the money, he does not cooperate. Player A's trust decision, on the contrary, is less easily interpretable as it is also influenced by the belief about B's trustworthiness. For example, if player A does not trust, this may be due to pessimism and not because player A is non-cooperative.[22]

Table 5: Norm Enforcement Types: Definition

	A0	A50B0	A50B100	A100B0	A100B200
Never punish nor reward	0	0	0	0	0
Punish non-cooperation	≥ 0	< 0	≥ 0	< 0	≥ 0
Reward cooperation	≤ 0	≤ 0	> 0	≤ 0	> 0
Reward and punish	0	< 0	> 0	< 0	> 0
Punish anti-socially	any value	any value	< 0	any value	< 0

Note: Types are based on player C's reward and punishment decisions targeting player B. AxBy is defined as the outcome in the trust game in which player A sends x and player B returns y.

Table 5 classifies player C's reward and punishment decisions into different norm enforcement types. We say that player C *never punishes nor rewards* if he does not allocate any reward or punishment points in any of the five possible outcomes of the trust game, i.e., total investment in norm enforcement is zero. Next, he *punishes non-cooperation* if he allocates punishment points in case player A trusts and player B keeps the money (A50B0, A100B0) but does not punish otherwise. Similarly, we say he *rewards cooperation* if he allocates reward points in case player A trusts and player B shares equally (A50B100, A100B200) and does not reward otherwise. *Reward and punish* is the combination of

[22]Results, however, do not depend on this. Type shares are also similar and not significantly different, if we include player C's decisions targeting player A.

both types. Finally, we also consider a so-called *anti-social type*, who punishes cooperation, i.e., allocates punishment points in case A trusts and B shares equally (A50B100, A100B200). The latter type has been documented to play a significant role in the success (or better, failure) of stabilizing norms of cooperation in groups and societies (Herrmann et al. 2008, Kosfeld and Rustagi 2015).

Table 6: Norm Enforcement Types: Shares

	Total	High school students	Police applicants	$p-$value (Fisher's exact test)
Never punish nor reward	0.122	0.131	0.102	0.140
Punish non-cooperation	0.623	0.607	0.656	0.079
Reward cooperation	0.263	0.260	0.270	0.699
Reward and punish	0.226	0.224	0.231	0.786
Punish anti-socially	0.009	0.010	0.007	0.566
N	1,418	959	459	

Table 6 shows the shares of norm enforcement types in the two groups of participants. Except for the type who punishes non-cooperation shares are not significantly different. On average, about 12 percent of the participants never punish nor reward, i.e., these participants decide in line with pure money-maximizing preferences. All remaining participants assign reward and/or punishment points in at least one condition although this comes at a personal cost to them. For example, about 62 percent punish non-cooperation, while 26 percent reward cooperation. Interestingly, most of the rewarding types punish as well, as the share of types who both reward and punish is not much smaller (22 percent on average). Finally, about one percent punish anti-socially, i.e., punish B although B cooperates.

The fact that the share of punishing types among police applicants is higher compared to high school students (66 vs. 61 percent, $p = 0.079$) but the share of rewarding types is not (27 vs. 26 percent, $p = 0.699$) seems remarkable and supports the interpretation that police agents' role in society is not to motivate citizens to make others better off but to prevent citizens from making others worse off.

Overall, the data suggest that police applicants' higher investment in norm enforcement is primarily a level effect, i.e., a generally higher motivation to reward or punish any

given behavior in the trust game. We therefore focus on average investment in our main analysis.

Online Appendix: Additional Tables

Table 7: Distribution of Trust and Trustworthy Behavior (in Percentage)

		High school students	Police applicants
Trust A	Send 0	18.46	16.99
	Send 50	57.56	50.98
	Send 100	23.98	32.03
Trustworthiness B if A sends 50	Return 0	25.44	18.30
	Return 100	74.56	81.70
... if A sends 100	Return 0	25.13	19.83
	Return 200	74.87	80.17
	N	959	459

Table 8: Covariates

	Total	High school students	Police applicants	$p-$value
Female	0.569	0.655	0.405	0.000
Age	19.870	19.555	20.473	0.000
ln(Income)	4.957	4.713	5.425	0.000
City type				0.000
large	0.104	0.078	0.155	
medium	0.207	0.199	0.223	
small	0.319	0.352	0.254	
rural	0.370	0.371	0.368	
Migration (y/n)	0.068	0.070	0.066	0.820
Education father				0.000
no	0.011	0.009	0.018	
Volks-/Hauptschule	0.126	0.097	0.179	
Mittlere Reife	0.393	0.363	0.451	
FH-Reife	0.070	0.068	0.074	
Abitur	0.364	0.418	0.260	
other	0.013	0.015	0.009	
don't know	0.023	0.030	0.009	
Education mother				0.000
no	0.010	0.006	0.017	
Volks-/Hauptschule	0.184	0.155	0.241	
Mittlere Reife	0.273	0.251	0.315	
FH-Reife	0.113	0.111	0.116	
Abitur	0.366	0.419	0.265	
other	0.019	0.021	0.015	
don't know	0.035	0.037	0.031	
NISS	50.721	49.794	52.492	0.000
Proactivity	34.505	33.737	35.974	0.000
Participation lotteries				0.000
never	0.619	0.577	0.697	
1-2	0.279	0.293	0.252	
3-10	0.086	0.109	0.044	
11-25	0.011	0.014	0.007	
more often	0.005	0.007	0	
Body height	173.987	172.872	176.118	0.000
Risk preference	5.188	5.041	5.468	0.000
N	1,331	874	457	

Note: Income = monthly income; city type = type of city grow up in until age of 15; participation in lotteries considers last 12 months; risk preference = general risk question from SOEP. Statistical significance is based on t-tests or alternatively, Fisher's exact test (female, migration) and χ^2-test (city type, education father/mother, participation lotteries).

Table 9: Trust and Trustworthiness

	Trust A	Trustworthiness B if A sends 50	Trustworthiness B if A sends 100
	(1)	(2)	(3)
Police applicant	0.061	0.068**	0.086***
	(0.043)	(0.027)	(0.028)
Female	0.056	0.038	0.117***
	(0.056)	(0.035)	(0.035)
Age	0.010	−0.000	−0.006
	(0.012)	(0.007)	(0.007)
ln(Income)	0.012	0.007	−0.008
	(0.020)	(0.012)	(0.012)
Migration	−0.105	0.051	0.075
	(0.073)	(0.047)	(0.047)
Education father	0.071*	−0.004	0.001
	(0.043)	(0.027)	(0.028)
Education mother	−0.068	0.009	0.021
	(0.043)	(0.027)	(0.027)
City type	0.034	0.020	−0.010
	(0.039)	(0.024)	(0.025)
NISS	−0.003	0.001	−0.001
	(0.002)	(0.002)	(0.002)
Proactivity	0.002	0.000	0.002
	(0.003)	(0.002)	(0.002)
Lottery	0.027	−0.004	−0.004
	(0.038)	(0.024)	(0.024)
Risk preference	0.024**	−0.005	−0.008
	(0.012)	(0.007)	(0.007)
Body height	0.007**	0.001	0.003
	(0.003)	(0.002)	(0.002)
Constant	0.451	1.480***	1.306***
	(0.604)	(0.383)	(0.384)
N	1,331	1,331	1,331
R^2	0.021	0.009	0.019

Note: The table reports the full set of covariates for specifications (3), (6), and (9) in Table 3 of the main paper. OLS with standard errors in parentheses. Education father and mother are coded as dummies (1 = Abitur, 0 otherwise), the same for city type (1 = non-rural, 0 = rural) and lottery (1 = participated at least once, 0 = never). *** p<0.01, ** p<0.05, * p<0.1

Table 10: Reward and Punishment

	Average investment C				
	(1)	(2)	(3)	(4)	(5)
Police applicant	4.215**	3.673**	2.876*	2.813*	3.173*
	(1.682)	(1.640)	(1.638)	(1.619)	(1.626)
Trust		8.851***		6.170***	
		(1.043)		(1.104)	
Trustworthiness50			6.975***	5.140**	
			(1.985)	(1.989)	
Trustworthiness100			10.027***	7.825***	
			(1.981)	(1.998)	
NoTrust_NotTrustworthy					−13.694***
					(2.690)
Trust_Trustworthy					9.387***
					(1.534)
Female	−3.107	−3.602*	−4.543**	−4.562**	−4.643**
	(2.159)	(2.104)	(2.104)	(2.080)	(2.089)
Age	−0.113	−0.199	−0.0522	−0.126	−0.020
	(0.450)	(0.439)	(0.437)	(0.432)	(0.434)
ln(Income)	0.103	−0.004	0.136	0.056	0.019
	(0.758)	(0.738)	(0.735)	(0.727)	(0.731)
Migration	4.055	4.983*	2.948	3.853	3.412
	(2.853)	(2.781)	(2.770)	(2.743)	(2.750)
Education father	2.764	2.133	2.797*	2.348	2.779*
	(1.681)	(1.639)	(1.631)	(1.614)	(1.620)
Education mother	1.636	2.237	1.362	1.844	1.565
	(1.668)	(1.627)	(1.619)	(1.603)	(1.609)
City type	−0.781	−1.085	−0.814	−1.013	−0.968
	(1.498)	(1.460)	(1.454)	(1.438)	(1.444)
NISS	0.003	0.031	0.004	0.024	0.019
	(0.093)	(0.090)	(0.090)	(0.089)	(0.089)
Proactivity	0.119	0.099	0.103	0.093	0.114
	(0.118)	(0.115)	(0.115)	(0.113)	(0.114)
Lottery	−1.044	−1.286	−1.053	−1.221	−1.270
	(1.470)	(1.433)	(1.426)	(1.410)	(1.417)
Risk preference	0.817*	0.607	0.928**	0.756*	0.806*
	(0.449)	(0.438)	(0.436)	(0.432)	(0.433)
Body height	−0.219*	−0.284**	−0.258**	−0.294***	−0.276**
	(0.116)	(0.113)	(0.113)	(0.112)	(0.112)
Constant	68.525***	64.537***	45.108**	47.920**	73.039**
	(23.487)	(22.883)	(22.927)	(22.674)	(22.639)
N	1,331	1,331	1,331	1,331	1,331
R^2	0.020	0.070	0.079	0.100	0.091

Note: The table reports the full set of covariates for specifications (3) to (7) in Table 4 of the main paper. OLS with standard errors in parentheses. Education father and mother are coded as dummies (1 = Abitur, 0 otherwise), the same for city type (1 = non-rural, 0 = rural) and lottery (1 = participated at least once, 0 = never). *** p<0.01, ** p<0.05, * p<0.1