Victim-Defendant Settlements under Asymmetric Bargaining Positions: The Role of Wealth and Connections in Access to Justice

Anastasia Antsygina, Madina Kurmangaliyeva*

November 18, 2019

Abstract

Both civil and criminal legal practices deal with victim-defendant settlements. Understanding how the power imbalance affects the negotiation process is crucial for the optimal design of the justice system and access to it. We develop a theoretical model where the victim, or she (resp. the defendant, or he) must exert costly effort for the case to end up with a conviction (resp. acquittal). Before the game starts, the defendant can settle with the victim by making her a take-it-or-leave-it offer. Using the data on criminal traffic offenses in Russia for 2013–2014, we structurally estimate the model. Our results show that law enforcement officers and government officials are approximately twice stronger than a comparable wealth group – white collar workers. The finding is robust to missing court outcomes and amnesty announced in December 2013. With the observed asymmetries in bargaining positions, the ban of victim-defendant settlements can worsen access to justice by more than 8.5%.

*Anastasia Antsygina, National Research University Higher School of Economics, aantsygina[at]hse.ru; Madina Kurmangaliyeva, Tilburg University, and the Institute for the Rule of Law at the European University at St Petersburg, m.kurmangaliyeva[at]uvt.nl,

1 Introduction

Victim-Defendant settlements are widespread in civil litigation and some criminal justice systems too. Courts and prisons are costly to society, so monetary settlements are a cheaper way to deter crime or increase level of care (Polinsky and Shavell, 1984).[1] Moreover, Victim-Defendant settlements can increase the victim's compensation when civil courts cannot enforce payments.[2] However, as argued by a legal scholar Fiss (1984), settlements are not cost-saving when the disputing parties are unequal in resources. Since poor victims do not have enough resources to litigate, rich defendants may settle with such victims at lower offers. Hence, poor victims still incur implicit costs of litigation through a reduced settlement amount, which goes against the core idea of justice that should prevent such distortions.[3] In this paper, we develop a theoretical model and, using Russian data on criminal traffic offenses, we provide (indirect) empirical evidence that higher-resource defendants are able to settle at lower settlement amounts.

Russia allows judges or prosecutors to stop criminal prosecution for unintentional or non-severe crimes if the defendant monetarily compensates and apologizes to the victim's satisfaction. In other words, criminal justice of Russia uses civil-style Victim-Defendant settlements within the criminal justice framework. This is in contrast, to most other countries that leave no or limited say for the victim in sentencing decisions. Moreover, Kurmangaliyeva (2018) provides causal evidence that the conviction probability in Russian criminal justice depends on the interaction between victim's and defendant's wealth, something that we expect to see in civil litigation but not in criminal cases: it is the prosecutor, not the victim, who exert effort to convict the defendant. We exploit the blurred line between civil and criminal justice in Russia to study the institute of civil-style Victim-Defendant settlements.

[1] For a discussion on whether rich defendants should be allowed to transform their resources into better legal counsel, see Lott Jr (1987) and Garoupa and Gravelle (2003).

[2] See Polinsky (2006) who shows that when the offender can hide his wealth, it is optimal to provide him with the choice between incarceration or fine, so that wealthier offenders pay voluntarily to avoid prison.

[3] Glaeser et al. (2003) show that judicial inequality can be a cause and the result of the subversion of institutions by the wealthy.

Usually, an empirical analysis of civil settlements is complicated by non-random selection of victim-defendant pairs in litigation. Victims are often endogenously related to the potential offender. Moreover, not every victim decides to file a suit, which may well depend on the wealth imbalance between the victim and the offender.

We use Russian data on criminal traffic offenses that involve drivers who hit and severely injure a pedestrian. Victim-Defendant settlements are allowed for criminal traffic offenses, even if the victim dies in the crash (i.e., the victim's family members can settle). Since this is the crime that happens in public spaces and victim is in need of immediate medical attention, police is notified and must investigate the crash. Thus, it is not a decision of the victim whether to start the litigation or not. Moreover, given the unintentional nature of traffic accidents, we can assume that the victim is random, i.e., she is a random draw from the underlying distribution of potential pedestrians at risk at a given location and time. Hence, there is no endogenous relation between the victim and the defendant prior to the crash.[4] To sum up, studying criminal traffic offenses involving pedestrians helps avoiding the two pesky sample selection problems.

We propose a stylized model of Victim-Defendant settlements in criminal justice where the disputing parties can spend their resources on legal efforts. We model criminal conviction as a perfect-information Tullock-type contest between two players – the victim and the defendant – who may differ in their bargaining positions. An improvement in the bargaining position of a player can be caused by a looser budget constraint, or lower costs of converting money to contest effort, or higher valuation of the prize.[5] Before the contest, the defendant can make a take-it-or-leave-it offer to the victim subject to his monetary wealth, and if accepted, the game ends.

Under perfect information, the settlement offer decreases (increases) when the

[4]Glaeser and Sacerdote (2003) use traffic accidents for the same reason.

[5]The contest stage of the model is based on Yamazaki (2008). Szidarovszky and Okuguchi (1997) study an asymmetric Tullock contest and prove the existence of a unique pure strategy Nash equilibrium. Yamazaki (2008) extends their result by adding player-specific budget constraints and focusing on a very general contest success function. Baye et al. (1994) analyze a discrete Tullock rent-seeking model with two homogeneous players and a contest success function that displays increasing returns to scale. In this class of games, the equilibrium cannot be derived from first-order conditions. The authors, however, prove that a symmetric equilibrium in mixed strategies exists and develop an algorithm to construct it. We use the indicated results to analyze contestants' decision to settle among themselves.

defendant's (victim's) bargaining position strengthens. Hence, relaxing the defendants's resource constraint allows him to settle more often for two reasons. First, assuming a continuity of victim's bargaining positions, it makes him able to afford better offers for stronger victims (the *volume effect*). Second, higher wealth allows the player to increase his effort in the contest stage, reducing the victim's equilibrium payoff and driving the optimal settlement offer down (the *price effect*). Whether the richer defendant in expectation pays more (the volume effect dominates) depends on the distribution of potential victims bargaining positions.

The data comes from the police database that includes information on the population of criminal offenses investigated in 2013-2014.[6] Our empirical analysis is complicated by two limitations: we do not observe the price at which the two parties settle and the cases in which the defendant was acquitted. We overcome the first limitation by focusing on policemen as defendants and policemen as victims. Policemen – who are not wealthier than white-collar workers and CEOs – can still be expected to drive down settlement prices through their connections and knowledge of the system. We overcome the second limitation by using the model to predict the direction of the sample selection bias for the reduced form analysis and by structurally estimating the model.

Controlling for time, location, the victim's socio-economic status, and defendants' wealth – their education level and the expected price of their cars – policemen are $22.5pp$ (8.2 sd) more likely to settle with their victims than white-collar offenders and CEOs in the sample without acquittals. Vice versa, when policemen are involved as victims, the share of settlements drop by -13.0pp (7.5 sd). It goes in line with the prediction of the model. Our model suggests that the observed gaps are not driven by the selection bias. Since better bargaining position is associated with both higher rate of settlements and higher rate of winning, a higher proportion of settled cases among the observed cases without acquittals can happen only if there is a higher proportion of settled cases in the full sample that includes acquittals. Hence, assuming same wealth and distribution of victims, the results tell us that policemen are for some reason in a better bargaining

[6]The access to data is provided by the Institute for the Rule of Law at the European University in Saint-Petersburg.

position. The results are robust to missing court outcomes and amnesty announced in December 2013.

Next, we structurally estimate a simplified model by dropping budget constraints. Now, cost of effort would incorporate information about both the player's budget constraint and his/her non-monetary resources. Assuming uniform distribution of victims' bargaining positions, we map the estimates of a reduced-form regression into a closed-form solution for structural parameters using Classical Minimum Distance estimator. We find that policemen as defendants have twice lower cost of effort than white-collar workers. Given that policemen are not wealthier than the white-collar workers, this signifies that their connections and knowledge of the system provides the policemen with better prospects in criminal justice.

We also recover counterfactual estimates of the probability of conviction. In a world without settlements, we would see that policemen as victims are 20% more likely to convict their offender than a white-collar victim, while policemen as defendants are 15% less likely to be convicted than a white-collar defendant. We also show that with the observed asymmetries in bargaining positions, the ban of victim-defendant settlements can worsen access to justice by more than 8.5%.

Overall, this paper argues that wealthier defendants while settling more are not necessarily paying more. Since the defendants' wealth improves their bargaining position in the contest against conviction, victims facing a wealthier defendant may be ready to settle at a lower price. Loosely speaking, from the prospective of the pedestrian, it can be "better" for her to be hit by a major than by a general. While, unfortunately, we do not observe prices directly, we still can capture the price effect through policemen who are settling more as defendants than a comparable wealth group. The finding is raising the question of unequal access to civil and criminal justice. Moreover, the unequal access to justice may be even more pronounced for intentional crimes. For intentional crimes, when the defendant can choose his victim (e.g. rape), the defendant will always choose the weaker type, and pay less than a poorer defendant would.

This paper brings together two strands of the Law and Economics literature: the

research on settlements and the literature dealing with resource imbalances and unequal access to justice. The former field uses game-theoretic models of settlements. Here, settlements are praised as cost-efficient, both in civil litigations and in the criminal justice (the plea bargaining between the prosecutor and the defendant), whereas trials are treated as a failure to achieve an agreement.[7] A considerable body of research has been trying to explore what provokes this inefficiency (Spier, 2007).[8]

Moreover, our work contributes to the body of research that structurally estimate the models of settlements. The most recent studies include Silveira (2017), Merlo and Tang (2016), Watanabe (2006), and Sieg (2000).[9] Our work differs from the aforementioned studies in several respects. First, we concentrate on Victim-Defendant settlements in the criminal justice. Second, imperfect information concerns are left out, and the research focuses on resource asymmetries. Third, we build upon a different model: if no settlement happens, the case outcome depends on the efforts of the conflicting parties. Fourth, we do not observe settlement offers, but our theoretical framework and case-specific controls available allow us to build a parametric estimator and to recover the distributions of players' preferences and effort costs.

The paper proceeds as follows. Section 2 describes how Russian justice system processes criminal traffic offenses. Section 3 introduces the model and states our main theoretical results. Section 4 characterizes the data, reduced form, and the structural setup, reports estimation results and robustness checks. Section 5 shows our estimates of access to justice (as defined in the section) in the counterfactual without settlements.

[7]For the literature that criticises settlements, see the papers that raise questions about the increased coercion of guilty pleas from innocents (Langbein, 1978; Alschuler, 1981) and about the inability to reach socially desirable outcomes (Polinsky and Rubinfeld, 1988; Garoupa and Stephen, 2008)

[8]Among the reasons, the literature cites asymmetric information (Reinganum (1986)), divergent beliefs of the parties (Landes (1971); Priest (1984)) and, for civil disputes, binding budget constraints defendants may face.

[9]Silveira (2017) focuses on Bebchuk (1984)'s model of bargaining under asymmetric information and proposes a non-parametric estimator to recover the distribution of defendants' types (their probabilities to be found guilty). Merlo and Tang (2016) look at civil settlements in medical malpractice disputes and recover beliefs of the conflict participants. As the authors claim, a failure to reach a pre-court agreement may arise from excessive optimism of the parties involved. They find that the plaintiff's perception of winning the trial changes with the harm made and the identity of his opponent (in this case, a doctor). Sieg (2000) and Watanabe (2006) also employ the data on medical malpractice litigations. The former paper shows that the bargaining model with settlements replicates all observed patterns quite well. Watanabe (2006) studies dynamic aspects of the negotiation process and emphasizes the role of learning about the opponent's beliefs in achieving the settlement.

Section 6 concludes.

2 A Legal Process for Criminal Traffic Offenses in Russia

According to the Criminal Code, all traffic offenses are classified into civil and criminal cases. The accident enters the latter group if it resulted in grave bodily injuries, which must be certified by the forensic medical exam results. A grave bodily injury must be "hazardous for human life" or involve the loss of sight, speech, hearing, or any organ or the loss of the organ's functions. Also, the legal definition accounts for a permanent loss of a general ability to work, an interruption of pregnancy, mental derangements, or post-traumatic addictions.

The Criminal Code of Russia categorizes respective traffic offenses based on a number of fatalities (namely, no death, one death, or multiple deaths). Moreover, it distinguishes between sober and driving while intoxicated (DWI) cases. The combination of these two characteristics defines six categories of criminal traffic offenses. For any of them, the judge can imprison the defendant for a period starting from several days and up to several years, where the upper cap varies with the offense category. For example, in the period of study 2013-2014, offenders in the "no death & sober" category could get at most two years of incarceration as the main punishment. At the same time, offenders from "multiple deaths & DWI" group could spend up to nine years in prison. As we study traffic accidents involving one driver causing grave injuries to one pedestrian, only four offense categories are relevant for our study.

When a traffic accident happens, the police station responsible for the accident location sends an investigator who must initiate a criminal investigation if there is death or medical reports of grave bodily injuries. The police investigator collects and analyzes all pieces of evidence: medical certificates, witness testimonies, experts' reports, photographs and video materials etc. In case of a hit-and-run, the investigator's task includes finding the responsible driver. At the investigation stage, both the suspected driver and the injured

may involve lawyers to help with the criminal case. If the pedestrian dies, his/her close relatives are recognized as victims. By the end of the process, the investigator passes all the materials to a prosecutor.

The prosecutor must decide whether there is enough evidence against the driver. One remarkable feature of Russian criminal system is that acquittals in court are very rare: less than 1% out of all cases, including criminal traffic offenses.[10] Hence, de-facto the prosecutor's indictment is the conviction. Whereas the court can be seen as a sentencing stage.

At this stage, the defendant with no criminal history of past traffic offenses and the victim can settle in a civil case fashion and dismiss the criminal charge. In particular, the offender voluntarily compensates all moral damages to the victim. The victim forgives the defendant and officially, in a written form, asks for the criminal prosecution to be stopped, subject to the approval of the investigator (with the permission of the prosecutor) or of the judge.Usually, the prosecutor – to avoid scrutiny and to improve her own internal statistics – forwards the settling parties to court and the bulk of the settlements happen with the permission of the judge. While in fact, the parties were ready to settle before court.

If the settlement is approved, the offender gets no criminal record because his guilt has not been verified in court. Hence, in case of a new criminal traffic offense, he can settle again. Nevertheless, the record that the case has resolved by settlement enters the police database and can be observed by external parties (for example, by potential employers during background checks).

If no settlement agreement has been reached, and the prosecutor indicts (i.e., convicts) the offender, the judge decides on the duration of the prison sentence. The judge may also suspend the prison sentence with a trial period, conditional the offender has no criminal record. For the "no death & sober" offense category – the lightest out of six – the judge may also replace a real incarceration term with a restriction of freedom, which is milder than prison. It allows the offender to live usual life, except for certain restrictions on

[10]This feature of the Russian criminal justice system is well-known and widely discussed by scholars. See for example, **Trochev2014**.

movement at night and outside the municipality. In fact, most of the non-settled cases for the "no death & sober" offense category end up as a restriction of freedom or a suspended prison sentence.

Additionally, judges usually revoke the offender's driver license up to several years. Moreover, the court decides how much the defendant must pay in order to cover all moral damages the victim faced. The compensation of medical expenses and property damages, however, is often a separate civil case, involving insurance companies.

Related to the period of our analysis, in December 2013, Russia announced amnesty for every offender under investigation for "no death & sober" offense category, but not the other categories. Every defendant in the relevant offense category who was still subject to an investigation or court case at the time of the amnesty was automatically amnestied, conditional on accepting guilt. We discuss the amnesty and the robustness of our results to its influence in Section 4.5.

3 The Model

3.1 Model Setup

To characterize the interaction between the victim (V, or she) and the defendant (D, or he), we introduce a simple contest model with two heterogeneous players. Such a setup is commonly used in the conflict literature to represent situations where parties exert costly effort in order to win a battle.[11] In our instance, V fights for D being convicted.

There is a continuum of *potential* victims who differ in their vindictiveness a where $a \sim \text{unif}\,(0,\, \bar{a})$. When an accident happens, V gets randomly matched with a representative defendant (D) who learns her characteristics. First, consider the "in-court" scenario. Let $P_C \in [0,\, 1]$ be a probability that the court convicts D. In this case, D gets punishment $x \geq 0$ and faces the total disutility of $\{-bx\}$ where $b < \bar{a}$.[12] [13] At the same time, V

[11] For example, see Esteban and Ray (2011), Sambanis (2017), Robson and Skaperdas (2008).

[12] To keep the analysis as general as possible, we do not specify how b and a relate to each other (both $a \geq b$ and $a < b$ can realize).

[13] Generally, the punishment x is case-specific and depends on the level of harm made to a victim and the degree of guilt.

gains $\{ax\}$. Clearly, V and D have misaligned preferences. Conviction is desirable for the victim. The defendant, however, would like to be acquitted. It results in a conflict where both V and D are willing to exert costly effort (e_V and e_D, respectively) and change the outcome in their favor. For simplicity, we assume linear effort costs:

$$c_i(e) = c_i e,\ i = V,\ D$$

where c_i captures heterogeneity in fighting abilities and does not depend on preferences for punishment (namely, a and b).[14] One can think of c_i as a composite measure of monetary and non-monetary resources available to player i. For example, victims and defendants with tight budget constraints cannot hire good lawyers or run extra examinations. Hence, they display weaker fighting abilities. Individuals who have good social connections or know the legislative system better hold a clear non-monetary advantage. This translates into lower effort costs. In Appendix A, we illustrate how exactly a contest model with budget constraints maps into the unconstrained version of the game with heterogeneous effort costs.

To specify how the probability of being convicted (P_C) depends on players' effort choices and their identities, we employ a standard Tullock contest success function:

$$P_C(e_V,\ e_D) = \frac{(e_V)^r}{\Sigma_{i=V,\,D}(e_i)^r},\ r = 1$$

Also, we state that if no party exerts positive effort, the case certainly results in conviction, i.e. $P_C(0,\ 0) = 1$. This assumption is non-standard in the literature; however, in case of criminal offenses, it makes perfect sense to break a "0–0" tie in favor of victims.

At the contest stage, V and D choose their effort levels to maximize the expected payoffs:

$$\begin{aligned} V: &\quad \max_{e_V} \pi_V(e_V,\ e_D) \\ &s.t.\ \pi_V(e_V,\ e_D) = axP_C(e_V,\ e_D) - c_V e_V \\ D: &\quad \max_{e_D} \pi_D(e_V,\ e_D) \\ &s.t.\ \pi_D(e_V,\ e_D) = -bxP_C(e_V,\ e_D) - c_D e_D \end{aligned}$$

[14] For simplicity, we assume that all potential victims face the same cost of effort.

Now, we introduce a pre-contest stage where V and D can settle. Assume the defendant makes an offer S to the victim before the active conflict phase starts.[15] [16] For simplicity, if S makes the victim indifferent between settling and fighting, she accepts the offer.[17] Further, we define contestants' bargaining positions.

Definition. Contestant i's bargaining position is a combination of his / her (dis)-utility of punishment and fighting ability c_i.

The game proceeds as follows:

1. D and V get matched at random and learn the preferences and effort costs of each other.
2. D makes an offer S to V. If V accepts the proposal, the game ends. Otherwise, D and V move to the contest stage.
3. D and V simultaneously choose their effort e_D and e_V, respectively.
4. The contest outcome realizes (the court either convicts D or acquits him), and the agents get their payoffs.

We solve the game by backward induction.

3.2 The Contest Stage

When V and D do not manage to settle, they move to the contest stage. Proposition 1 provides a general equilibrium characterization of the contest game:

Proposition 1. *The equilibrium of the contest stage exists and is unique.*

Proof. See Appendix B. □

The existence and uniqueness results are proven by construction. The game we are looking at represents a standard asymmetric Tullock contest with two participants, which is well-studies in the literature. With the given contest success function, the equilibrium is

[15] As lawyers claim, in most of the cases it is indeed the defendant who makes a settlement offer.

[16] In principle, one could model the pre-contest stage as a Nash bargaining game where V and D split the surplus among themselves. However, to identify contestants bargaining power, it is crucial to observe the settlement amount, which is never reported. For this reason, we stick to a simplistic assumption of D making a first move and extracting all the surplus.

[17] The analysis extends to the case when V can randomize between settling and fighting.

always interior and unique. Also, it features pure strategies. Proposition 2 summarizes how contestants' equilibrium effort depends on their fighting abilities and preferences for punishment.

Proposition 2. *Contestant i's equilibrium effort e_i^* always increases in his / her valuation of punishment and decreases in c_i:*

$$\frac{\partial e_V^*}{\partial a} \geq 0,\ \frac{\partial e_D^*}{\partial b} \geq 0\ and\ \frac{\partial e_i^*}{\partial c_i} \leq 0\ for\ i = V,\ D$$

For $\frac{a}{c_V} \geq \frac{b}{c_D}$:

1. *e_V^* increases in b and decreases in c_D;*
2. *e_D^* decreases in a and increases in c_V.*

For $\frac{a}{c_V} < \frac{b}{c_D}$:

1. *e_V^* strictly decreases in b and strictly increases in c_D;*
2. *e_D^* strictly increases in a and strictly decreases in c_D.*

Proof. See Appendix B. □

Some results stated in Proposition 2 are straightforward. The equilibrium effort never decreases in the valuations both players attach to punishment. Higher a and b drive contestants' willingness to win up and make the competition tighter.[18] Also, better fighting abilities, associated with lower values of c_V and c_D, allow the players to exert less effort without affecting their winning probabilities. These two facts are well-documented in the contest literature.

The other effects indicated in Proposition 2 depend on relative bargaining positions. Take the case of $\frac{a}{c_V} \geq \frac{b}{c_D}$ when V displays a stronger relative willingness to compete than her opponent. Here, winning is more desirable for the victim. If D's bargaining position improves (namely, b grows or c_D declines), V tends to increase her effort and fight back to keep the probability of conviction as high as possible. The opposite holds for the defendant. When a increases, the victim who already has an advantage (recall

[18] If a goes up, V extracts more utility from D being punished. Higher values of b translate into bigger costs of conviction for D, and his incentives to avoid the court stage increase.

$\frac{a}{c_V} \geq \frac{b}{c_D}$) gets even stronger incentives to fight. This discourages D, and in equilibrium, he exerts less effort. Similar arguments apply when $\frac{a}{c_V} < \frac{b}{c_D}$ holds and, hence, D has a better bargaining position.

3.3 The Settlement Stage

In this subsection, we move one step back and analyze when V and D settle. Let π_i^*, $i = V, D$ be i's equilibrium payoff at the contest stage. First, characterize the optimal settlement offer S.

Lemma 1. *The optimal settlement offer equals to V's equilibrium payoff at the contest stage, i.e. $S^* = \pi_V^*$.*

Proof. The proof is straightforward. Without loss of generality, suppose D's budget is unlimited, and he can afford any settlement offer. Also, assume D strictly prefers to avoid the contest stage. Formally, fix $\pi_D^* \ll -(\pi_V^* + \tau)$ where $\tau \gg 0$ is sufficiently high.[19] First, take $S = \pi_V^* + \varepsilon$ where $\varepsilon > 0$ is small enough. The victim strictly prefers to accept the offer, and D's payoff becomes $\pi_D^{+\varepsilon} = -(\pi_V^* + \varepsilon)$. Next, consider $S = \pi_V^* - \varepsilon$. Now, the victim does not want to settle, the game proceeds to the contest stage, and $\pi_D^{-\varepsilon} = \pi_D^*$. Finally, check $S = \pi_V^*$. In this case, V accepts the proposal (see the assumptions of Subsection 3.1), and D gets $\pi_D = -\pi_V^*$. $S = \pi_V^*$ strictly dominates all other alternatives:

$$\pi_D = -\pi_V^* > \max\left\{\pi_D^{+\varepsilon},\, \pi_D^{-\varepsilon}\right\}$$

and D prefers this strategy. □

Lemma 1 illustrates a typical first-mover advantage. Since D makes a "take-it-or-leave-it" offer in the absence of private information, he extracts all the surplus. If D prefers to avoid the contest stage ($\pi_D^* < -\pi_V^*$), proposing $S^* = \pi_V^*$ allows him to terminate the game, save on settlement costs and get the highest possible expected payoff.

When the optimal settlement offer is defined, we can check how it depends on the victim–defendant characteristics.

Proposition 3. *The optimal settlement offer S^* always:*

[19] The extreme case would be $\pi_D^* = -\infty$.

- *Decreases in D's willingness to win b and in V's fighting cost c_V and*
- *Increases in V's willingness to win a and in D's fighting cost c_D.*

Proof. See Appendix B. □

These results are quite intuitive. If the defendant gets stronger incentives to compete (either his winning benefit increases or fighting abilities improve), he exerts more effort. Then, the victim can either fight back or give up.[20] Under the former scenario, V faces higher effort cost; in the latter case, her winning probability decreases. Overall, V's equilibrium payoff declines, and it becomes easier to settle for the defendant.

The opposite happens when V's willingness to win grows or her ability to fight rises. In this case, D faces a stronger opponent who exerts significant effort, wins with a high probability and, consequently, obtains a larger equilibrium payoff. To prevent the fight, D must give the competitor a sufficient amount of money. Hence, settling with a mighty victim is more expensive.

Proposition 3 implies that matching with a richer defendant does not result in a better settlement offer. As Appendix A illustrates, more money available to D reduces the effort cost c_D and improves his bargaining position. Thus, the value of S^* must go down, *ceteris paribus*. The victim still accepts the offer made; however, her equilibrium payoff declines. This result goes against a conventional perception developed in the literature on "victim-defendant" settlements. The difference stems from the way we model the interaction between the two conflicting parties. Specifically, we use a contest setting where V and D challenge each other. In this case, D's fighting abilities affect V's equilibrium payoff directly, and vice versa. The previous studies on the topic did not employ this competitive approach and could not discover the pattern we find here.

Next, we analyze when the settlement takes place. To prevent the conflict, the following condition must hold:

$$D \text{ makes an offer if and only if } S^* \leq -\pi_D^*$$

$$\Leftrightarrow (a-b)\, x P_C\left(e_V^*,\, e_D^*\right) \leq c_V e_V^* + c_D e_D^* \tag{1}$$

[20] See Proposition 2 for more details.

where asterisks denote equilibrium values. Condition (1) states that the defendant must find it profitable to terminate the game, i.e. his payoff from entering the contest stage needs to be smaller than the settlement cost. The first observation connects players' preferences over punishment and D's willingness to settle. When the victim is not sufficiently vindictive (i.e. $a \leq b$), condition (1) always holds. In this case, the settlement is efficient. Otherwise, the defendant may prefer to fight for some values of a.

Proposition 4. *There exists* $\tilde{a} \equiv \tilde{a}(b, c_V, c_D) > b$ *such that* V *and* D *settle for any* $a \in [0, \tilde{a}]$.

Proof. See Appendix B. □

To find the threshold value of a, we directly compare S^* and $(-\pi_D^*)$ and show when condition (1) holds. The defendant offers the victim the amount of S^* if and only if she is not too vindictive. Importantly, the $(a \leq b)$ case constitutes only a subset of preference profiles for which settlements can be observed. Intuitively, higher values of a drive the optimal offer S^* up and make the settlement more expensive (see Proposition 3). Suppose V's bargaining position improves, and $a > \tilde{a}$ is reached. If the defendant makes the offer to the victim, it is certainly accepted, and D must pay S^* with probability 1. However, if the game proceeds to the contest stage, the defendant faces the punishment with probability less than 1. When V has a strong bargaining position (namely, $a > \tilde{a}$), D's equilibrium payoff turns to be higher in the latter case. Hence, the defendant decides to fight.

To look at Proposition 4 from a different angle, we rewrite condition (1) in terms of the optimal probability of being convicted $P_C^* \equiv P_C(e_V^*, e_D^*)$:

$$D \text{ makes an offer if and only if } S^* \leq -\pi_D^*$$

$$\Leftrightarrow P_C^* \geq \frac{a-b}{a+b} \cdot \frac{ac_D}{bc_V} \equiv \tilde{P}_C$$

In words, the defendant wants to escape the contest stage if and only if P_C^* is sufficiently high. In this case, fighting most likely results in punishment. To avoid such an outcome, the defendant makes the victim a settlement offer, and the game terminates. For $P_C^* < \tilde{P}_C$, the chances to get punished become relatively low. Then, the defendant prefers to engage

in a contest game with a probabilistic outcome instead of paying the amount of S^*, which must be transferred with certainty.

Finally, we take one step back and consider the settlement offer from an ex-ante prospective. Suppose no victim-defendant match has been realized yet. We trace how D's fighting ability c_D, which relates to his monetary and non-monetary resources, affects the expected value of S^* and the probability to settle P_S. Formally, the objects of interest are:

$$P_S = \frac{\tilde{a}}{\bar{a}} \text{ and the sign of } \frac{\partial P_S}{\partial c_D}$$

$$E\left(S^* \middle| a < \tilde{a}\right) = \frac{1}{\tilde{a}} \int_0^{\tilde{a}} S^* da = \frac{c_D^2 x}{\tilde{a}} \int_0^{\tilde{a}} \frac{a^3}{\left(ac_D + bc_V\right)^2} da \text{ and the sign of } \frac{\partial ES^*}{\partial c_D}$$

Two effects must be emphasized:

1. The settlement threshold $\tilde{a}$ (weakly) decreases in c_D:

$$\frac{\partial \tilde{a}}{\partial c_D} = -\frac{c_V}{c_D^2} \frac{b}{\sqrt{\frac{1}{4} + 2\frac{c_V}{c_D}}} I_{\{\tilde{a} < \bar{a}\}} \leq 0$$

   which implies $\frac{\partial P_S}{\partial c_D} \leq 0$. In words, defendants with better fighting abilities can settle with a bigger mass of potential victims. We call this the "volume effect".

2. The conditional expected settlement offer $E\left(S^* \middle| a < \tilde{a}\right)$ always increases in c_D:

$$\operatorname{sgn}\left(\frac{\partial ES^*}{\partial c_D}\right) = \operatorname{sgn}\left(-\frac{\tilde{c}}{c_D}\left[6\tilde{c}\ln\left(\frac{1+\tilde{c}}{\tilde{c}}\right) + \frac{\tilde{c}^2\left(6+5\tilde{c}\right)}{\left(1+\tilde{c}\right)^2} - 5\tilde{c} - 2\right]\right) I_{\{\tilde{a} < \bar{a}\}} +$$

$$\operatorname{sgn}\left(\frac{\tilde{c}}{c_D^2}\left[6\tilde{c}\ln\left(\frac{\tilde{c}}{1+\tilde{c}}\right) + \frac{6\tilde{c}^2 + 9\tilde{c} + 2}{\left(1+\tilde{c}\right)^2}\right]\right) I_{\{\tilde{a} = \bar{a}\}} = 1 \text{ where } \tilde{c} = \frac{c_V}{c_D}$$

   Thus, defendants with an effort cost advantage pay less (in expected terms) when settle. This defines the "price effect".

Overall, the reduction in c_D benefits the defendant in two ways. First, he avoids the contest stage more often (namely, P_S grows). Second, the defendant needs to pay less in order to convince the prospective victim to settle, i.e. $E\left(S^* \middle| a < \tilde{a}\right)$ declines.

With all the observations made, we briefly discuss the institute of "victim-defendant" settlements from the social welfare prospective. Generally,

defendants who escape the contest stage are richer or/and have good social connections. In other words, they face lower effort costs and tend to display better fighting abilities. If we focus on a particular "victim-defendant" match, the presence of settlements makes no party worse off.[21] However, if the society has preferences that are more than just a sum of V's and D's payoffs, the settlements may be abandoned.[22]

Another argument against victim-defendant settlements in the presence of asymmetric bargaining positions relates to deterrence concerns. If advantaged individuals know that in case of a norm violation their victims are likely to have worse fighting abilities, the settlement becomes cheaper. Consequently, they get stronger incentives to break the law than their less advantaged peers. As a result, the settlements make it more problematic to sustain uniform deterrence across different socioeconomic groups.

Deterrence concerns may be less important in case of accidental crimes, such as traffic offenses. However, they become crucial when one focuses on intentional felonies. Now, offenders can decide which victim to target. Since individuals with lower income or/and weaker connections are easier to settle with, they are more likely to be victimized. Roughly speaking, the possibility of settlements in the criminal law may create a "market" for potential victims. This argument can also convince the policy maker against the given institution.

4 The Empirical Analysis

In this section, we bring the proposed theoretical model to the data on criminal traffic offenses in Russia.

[21] Both V and D obtain their contest equilibrium payoffs at least.

[22] One example comes from incapacitation concerns when the society wants to keep dangerous criminals in prison.

4.1 The Data

4.1.1 Data Sources

We use police data on criminal cases that have been under investigation in 2013-2014. The personnel in the police and prosecutor office must fill in (digital) statistical forms (the template provided by law). The statistical forms help collecting information about different stages of the criminal case, following up the case up to court outcomes. The centralized police database aggregate data from the forms across all 84 Russian federal subjects in a centralized manner.[23]

The first database represents the universe of criminal traffic offenses that have been registered by police stations. Here, a unit of observation is a case. The information available, among other things, includes (1) the police department number, the administrative code OKATO of the location of the police department; (2) the date and time of the accident; (3) The aggregate data on victims such as a number of deaths and injured, their social/employment status of up to two victims; (4) the outcome of the investigation stage, e.g. forwarded to court, unresolved hit-and-run, settled, etc.[24] [25] [26]

Another database incorporates information about the defendant. We observe the data on defendants' demographic attributes such as gender, socio-economic status, history of criminal and administrative records etc. If the case reached court, there is also usually information on the court outcome, including the type of punishment and its duration.[27]

[23]The Institute for the Rule of Law at the European University at Saint Petersburg got the access to the filled-in statistical forms. The Institute processed them into several STATA databases with the support of Russian Science Foundation grant 17-18-01618. We received access to the databases from the Institute for research purposes under a restricted user agreement.

[24]The OKATO stands for the *Russian Classification of Objects of Administrative Division*. The code consists of nine digits. The first two identifies the federal subject (e.g., republics, krai, oblast, etc.), i.e., the largest administrative subdivisions of Russia. There are in total 84 federal subjects coded from '01' to '84.' The third to fifth digits identify the second largest subdivision, which usually corresponds to big cities or municipalities. The next three digits refer to city districts within a city, or towns and villages within a municipality. The last digit is a control number.

[25]The investigators code the socio-economic status of participants according to an official guide-book. The information includes different social groups (retiree, rentier, worker, etc.) combined with the more detailed employment information for workers (up to 100 professions) and the type of the organization they work in (public, private).

[26]The case-level database was compiled by the Institute for the Rule of Law at the European University at Saint Petersburg from the police statistical forms 1 and 2.

[27]The offender-level database was compiled by the Institute for the Rule of Law at the European University at Saint Petersburg from the police statistical forms 2 and 6.

28

The third database provides detailed information for each victim: gender, age, ethnic group, citizenship, local residency status, social/employment status, harm caused by the accident.[29]

The data also include a so-called *fabula* that describes the case shortly. Often, investigators use this document for their own easy reference. The description style and the amount of details it contains show significant variation across police departments. Usually, the *fabula* consists of two parts. First, it provides general information on the situation (time, location, weather conditions etc.) and the participants, starting from the description of the offender's actions. The text usually specifies the types of cars driven by the defendant and the victim (where applicable). It also mentions whether pedestrians were involved. The second part of the *fabula* describes the harm made and clarifies who the victim is: a pedestrian, a passenger, or a driver. We process *fabula* using regular expressions search.

We also use imputed car prices for the cars mentioned in the *fabula*, using prices of same-brand second-hand cars posted on *https://auto.ru/* in October, 2014.[30] The first car mentioned in the *fabula* is attributed to the offender.

We merge the three datasets and fabula, using the case identifier, the code of the police department, and the year when the accident happened.

4.1.2 Descriptive Statistics

We restrict the sample to traffic offenses that involve one driver and one pedestrian who has sustained grave bodily injuries.[31] We keep only cases with the accident date between July 2012 to June 2014, for which the investigation had to happen in 2013-2014 and the

[28] For every registered case, there can be no defendant (an offender has not been caught or did not get an accusation), exactly one defendant, and more than one defendant (the crime was committed by a group).

[29] The victim-level database was compiled from form 5.

[30] *https://auto.ru/* is one of the largest on-line platforms for private car sales in Russia.

[31] Any match with the phrase "hit a pedestrian" and its variations raises the flag for the variable *pedestrian*. Any match with the word "passenger" and its variations raises the flag for the variable *passenger*.We keep the cases that mention pedestrians but remove those that also mention passengers.

outcome known by the end of 2014.[32]

Among 14′629 cases, only 10′805 cases were indicated as prosecuted or settled. Mostly, the gap is due to hit-and-runs (in around 2′000 cases the driver was not found). On top of this, the police did not press charges for some of the identified drivers, and the prosecutors happened to drop cases as well. In total, around 800 drivers were acquitted. However, some investigations took longer to be concluded, so the outcome was still missing by the time of the database collection.[33] Finally, some of these cases might have been prosecuted or settled, but the investigator did not update the records in time. Indeed, there is a substantial geographical heterogeneity in the degree of missing outcomes, both when aggregated at the city/municipality level (see Figure C.1 in Appendix C), or at the level of federal subjects.[34] We assume that the regional heterogeneity is linked to the quality of the bureaucracy, rather than to the relative power of the participants. In fact, there is no statistically significant relation between the incidence of the missing outcome and the socio-economic status of the defendant and the victim.[35]

We restrict our analysis to the following seven socio-economic groups of offenders and victims (as regrouped from initially more detailed group coding): 1) Individuals with no permanent employment (NoEmpl); 2) Blue-collar workers (BC); 3) White-collar workers (WC); 4) Individual entrepreneurs (Entr); 5) Company (co-)owners or chief executive officers (CEO); 6) Law enforcement officers, e.g., policeman, court clerk, judge, etc., or government officials, e.g. deputy, lawmaker (LEOGVT);[36] 7) Retirees (Ret).[37]

As you can see from Table 1, the offenders with no permanent employment drive

[32]There are (a few) accidents dating back to much earlier years, up to 1990s. Such accidents for some reason took longer to be registered or received new investigation to be in the reporting period of 2013-2014. Hence, we eliminate them due to sample selection concerns.

[33]We check our results for robustness in Section **??** by restricting the sample only on earlier observations from July 2012 to December 2013.

[34]For example, 31% out of 200 cases in our working sample for the Republic of Bashkortostan have missing information on the investigation outcome, compared to just 5.3% median level across all other federal subjects with number of cases of 150 cases or more.

[35]See Table **??** for the results of the regression with police department and date of crime fixed effects. It shows that the socio-economic statuses of the two parties are not individually or jointly significant.

[36]The list of professions included in LEOGVT: policeman, lawyer, deputy, judge, prosecutor, employee at the judiciary, the federal penitentiary, fire department, the Investigative Committee of Russia, the Federal Security Service, Customs Office, Tax office, court, prosecution office, the Ministry of Justice, the Federal Drug Control Service, the Ministry of Emergency Situations, the Federal Bailiffs Service.

[37]In other words, we excluded children, students, dependants, the military (as they are under the jurisdiction of military courts), and a small group of workers indicated as *other* by the police.

the cheapest cars and have the lowest share of college graduates among the working-age population.[38] Company owners and CEOs have the best cars and the highest level of educational attainment, followed by white-collar workers. Individual entrepreneurs seem to have more wealth and education than the blue-collar workers, but they lag behind the white-collar workers. The majority of entrepreneurs are probably small retail traders or engaged in providing small services. Retirees tend to drive cheaper cars, which probably reflects that they still drive old cars. However, in terms of educational attainment retirees are close to average. Driving a truck, bus, or a motorbike seem to be correlated with lower socio-economic status. Most importantly, law enforcement officers and government officials drive worse cars than the white-collar workers. LEOGVT are also more educated than the average worker, but less than the white-collar workers and the CEOs.

Table 1: Characteristics of offending drivers by employment/social status

	Car price, th RUR		Truck, bus,motorbike	College degree
	mean	sd	share	share
NoEmpl	0.26	0.25	0.14	0.11
BC worker	0.29	0.24	0.15	0.19
WC worker	0.42	0.27	0.09	0.60
Entrepreneur	0.37	0.30	0.14	0.28
CEO	0.50	0.33	0.07	0.65
LEOGVT	0.36	0.30	0.10	0.50
Military	0.42	0.34	0.21	0.14
Child	0.22	0.20	0.40	0.00
Student	0.27	0.24	0.14	0.05
Dependant	0.27	0.30	0.32	0.15
Retired	0.25	0.23	0.09	0.26
Other	0.31	0.25	0.10	0.28

The statistics is based on all criminal traffic offenses, including car-to-car, car-to-pedestrian, and car-to-obstacle crashes.

Importantly, for around 20% of cases in our sample of interest, we know that the case reached court, but we do not know whether the case resulted in a settlement or conviction. We drop these cases from the sample. In this way, we assume that court outcomes are missing at random. This is probably not true, since the rate of missing outcomes likely reflects the duration the case takes to reach court or sentence in court.

[38]The sample in Table 1 is based on all criminal traffic offenses, including car-vs-car, car-vs-pedestrian, and car-vs-obstacle crashes.

Hence, it could be that among cases that took longer to reach every stage, there would be fewer settlements. Hence, we believe that dropping such cases may result in overestimation of the differences in the degree of settlements among some groups, but it will still serve as a useful benchmark. In fact, when we drop such cases, the share of settlements in our sample coincides with the official court statistics.[39] Later we perform robustness checks by cutting the sample of observations for the first few months at the beginning of our sample (See Section 4.5).

The final sample includes 6′244 observations coming from 1′292 distinct police departments and decisions by 1′099 courts.

Table 2 provides descriptive statistics. Victims died in 30% percent of cases. Nine percent of cases involved a DWI driver. Women represent half of the victims, but only 11% of offending drivers, which is not surprising given that women are in general underrepresented among drivers. Offenders are much younger than victims, with 40% of offenders aged between 18 to 29 years old in comparison to just 17% among victims. Also, offenders on average have higher socio-economic status than their victims, with higher incidence of CEOs and white-collar workers. It is also not surprising, as the population of drivers represent a selected sample of economically active layer of Russian population. Note that retirees represent 30% of victims, while only 6% of offending drivers. Overall, the demographic statistics for victims and offenders seem to correspond to what we expect to see as pedestrians and drivers at risk.

We obtain information about the driver's cars only in 40% of cases. Since the fabula is written by the investigator for the internal use, it seems that some investigators use car brands as a salient feature of the accident while other investigators do not. Indeed, as Figure C.2 shows, the incidence of car information in the fabula of a police department follows a bimodal distribution. Hence, the use of car prices is subject to sample selection concerns.

Outcomes vary with the offense category (See the lower part of Table 2). For the lightest category "no death & sober" almost no offender has been incarcerated. Most of

[39]See statistics for 2013 from the Judicial Department at the Supreme Court of the Russian Federation available at `http://www.cdep.ru/index.php?id=79&item=2362`

Table 2: Summary Statistics

GENERAL INFO:	
n obs	6,244
distinct police deps	1,292
distinct courts	1,099
n federal subjects	79

SHARE OF VICTIMS:	
victim died	0.307
female	0.482
was under influence	0.039
by age:	
16 to 17	0.002
18 to 24	0.080
25 to 29	0.085
30 to 59	0.495
60 or older	0.337
by employment:	
no perm job	0.408
bc worker	0.249
wc worker	0.028
entrepreneur	0.004
ceo	0.004
leogvt	0.007
retiree	0.300

SHARE OF OFFENDERS:	
female:	0.108
was under influence	0.091
by age:	
16 to 17	0.000
18 to 24	0.186
25 to 29	0.206
30 to 59	0.542
60 or older	0.066
by employment:	
no perm job	0.375
bc worker	0.475
wc worker	0.038
entrepreneur	0.033
ceo	0.015
leogvt	0.008
retiree	0.056
by education:	
college	0.200
vocational	0.354
car info present	0.439
out of which:	
bus, truck, motorcycle	0.113
cheap car[1]	0.473
medium-price car[2]	0.404
expensive car[3]	0.010

BY OFFENSE CATEGORY:

	sober no death	dwi no death	sober death	dwi death
n obs	3,948	377	1,624	295
SHARE BY OUTCOME:				
settled	0.25	0.20	0.25	0.06
out of which before court	0.02	0.01	0.03	0.01
amnesty	0.43	0.00	0.01	0.00
limitation of freedom	0.24	0.02	0.01	0.00
suspended prison sentence	0.02	0.44	0.34	0.11
imprisoned	0.02	0.31	0.35	0.79
SENTENCE IN MONTHS (if imprisoned):				
average	20.65	19.89	23.83	38.06
st dev	11.31	9.30	9.59	12.44

The sample includes one-driver-one-pedestrian criminal traffic offenses dating between July 2012 to June 2014 (convictions or settlements only, excluding the cases that do not have court outcome stated). [1] expected resale price <250K RUR; [2] expected resale price $\in (250\text{K}, 500\text{K}]$ RUR; [3] expected resale price >500K RUR.

the cases have resulted in amnesty, limitation of freedom, or settlement. However, for the other three categories involving a death or/and DWI the probability of real incarceration and the prison term length are increasing with the gravity of the offense, while the probability of prison sentence suspension and settlements are decreasing.

4.2 The Value of Connections: The Reduced Form Evidence

To illustrate that the relative strength of the defendant to victim affects the price at which the victim agrees to settle, we focus on law enforcement officers and government officials. Belonging to this socio-economic group has two non-monetary returns. First, law enforcement officers and government officials know the institutional setting better and can defend themselves more efficiently in case of committing a crime / becoming a victim. Second, these people are connected to the networks of lawyers, legislators, and other key individuals. Hence, they may exploit the latter channel to affect the case outcome at the contest stage. For these reasons, we expect the given group to display different patterns in terms of settlements when compared to contestants similar in wealth.

Consider the following regression equation:

$$\begin{aligned} S_i = \alpha &+ \beta_D leogvt_i^D + \beta_V leogvt_i^V + \beta_{DV} leogvt_i^D \times leogvt_i^V \\ &+ \psi_1 highses_i^D + \psi_2 highses_i^V + \psi_3 highses_i^D \times highses_i^V \\ &+ \psi_4 college_i^D + \psi_5 college_i^D \times highses_i^V \\ &+ \psi_6 car_i^D + \psi_7 car_i^D \times highses_i^V \\ &+ \gamma other\_controls_i + u_i \end{aligned} \tag{2}$$

where

- $S = 1$ if settled, $(= 0$, otherwise);
- $leogvt^l = 1$ if $l = \{V, D\}$ is a law enforcement officer or a gvt official $(= 0$, otherwise);
- $highses^l = 1$ if $l = \{V, D\}$ belongs to high socioeconomic status, i.e., white-collar

workers, CEO, or LEOGVT. The baseline group $highses^l == 0$ includes individuals with no permanent employment and blue-collar workers;

- car^D includes three dummies for cheap cars, medium-price cars, and expensive cars (the dummy for the base category of bus, truck, or motorcycle drivers is omitted);
- $college^D$ is a dummy that equals to one if the defendant has a college degree;
- *other_controls* include different controls for offense category, time, and location fixed effects:
    1. *dwi* and *death* are dummies for the driving-while-intoxicated accident and the victim's death after the accident;
    2. *police*, *time*, and *accident_month* are dummies that help removing fixed effects of police departments, hour×day-of-week, and the date of the accident as aggregated in monthly intervals;

Using regression 2, we compare law enforcement officers and government officials with individuals of comparable wealth level. If connections and knowledge of the system does not affect the contest stage, then LEOGVT defendants should not settle more often than white-collar workers and CEOs with comparable level of education driving the same type of car. It means that β_D should be equal to zero, as well as β_V and β_{DV}.[40]

Table E.2 summarizes the estimates of β_D, β_V, and β_{DV} specified in Regression (2).[41] Columns 1 to 4 of Table E.2 differ in the set of controls used for proxying defendant's wealth: with and without including education and car information.

Results show that law enforcers and government officials as defendants have higher rate of settlements with their victims than other defendants with comparable socio-economic status, education, and car type. Controlling for education level, a probability to settle for LEOGVT exceeds its baseline counterpart of white-collar workers and CEOs by 27 percentage points. This is equivalent to doubling the share of settlements. If we restrict

[40] β_D can also be negative, since LEOGVT are likely to be less wealthy than WC workers and CEO combined, given the summary statistics from Table 1.

[41] For the all other coefficients please see Table C.1

Table 3: Settlement Rates for Law Enforcement Officers and GVT Officials

	(1)	(2)	(3)	(4)
β_D	0.225	0.265	0.464	0.471
	(0.082)	(0.088)	(0.129)	(0.128)
p-value for H0: $\beta_D = 0$	0.006	0.003	0.000	0.000
β_V	-0.130	-0.119	-0.188	-0.207
	(0.075)	(0.076)	(0.108)	(0.106)
p-value for H0: $\beta_V = 0$	0.085	0.117	0.083	0.052
β_{VD}	-0.588	-0.626	-0.204	-0.316
	(0.322)	(0.316)	(0.207)	(0.211)
p-value for H0: $\beta_{DV} = 0$	0.068	0.048	0.325	0.134
Wealth controls:				
$highses^{D,V}$, $highses^D \times highses^V$	yes	yes	yes	yes
$college^D$, $college^D \times highses^V$	no	yes	yes	yes
car^D, $car^D \times highses^V$	no	no	no	yes
Sample restricted?	no	no	car info	car info
n obs	4,347	4,265	2,529	2,529
n police deps	1,008	998	685	685

Note:
This table reports the estimates of coefficients β_D, β_V, and β_{DV} for regression (2). Standard errors (in parentheses) are clustered at the police department level. All regressions include police department, hour× day-of-week, date of accident (in months), and offense-category fixed effects. Column 1 reports results only controlling for the socio-economic status of the defendant. Starting from Column 2, regression includes the control for D's education, while in Column 4, it also accounts for the car category. The difference between results in Column 2 and 3 is in the sample, for regression 3 it is restricted to cases that have information on vehicles driven by the defendant at the time of the accident. The data includes one-car-one-pedestrian criminal traffic offenses that happened between July 2012 to June 2014, convictions or settlements only. Cases with missing court outcomes have been excluded. The sample of defendants and victims is restricted to NoEmpl, BC, WC, CEO, and LEOGVT workers only.

the sample to the cases where there is information on the car driven by the defendant, the gap increases to 46 percentage points. However, the difference between β_D in Column 2 and Column 3 is not stastically different (the confidence intervals overlap). Adding controls for the car category (Column 4) does not affect the estimate of β_D, which remains significantly higher than zero.

Notably, β_V is negative across all four columns of Table E.2 with the borderline statistical significance at 10% level. For some reason LEOGVT victims are associated with lower degree of settlements than other white-collar workers or CEOs. Also, the interaction term β_{DV} is negative, although insignificant when using all the controls (which most probably is due to sample size issues).

Overall, law enforcement officers and government officials tend to behave differently in the case resolution. Controlling for wealth proxies does not eliminate the discrepancies observed across groups. Indeed, the direction of the discrepancies is in line with the non-monetary channels explanation. In other words, LEOGVT defendants are able to use their connections and knowledge of the system to be a stronger contestant against their victims, decreasing the price at which their victims are ready to settle and allowing LEOGVT to settle more often than the defendants with comparable or even higher wealth constraints. Similarly, LEOGVT are also stronger as victims, and that is why their settlement price is higher than the the one asked by victims of comparable wealth.

We are cautious about interpreting the extent of the differences in settlement rates across groups. Let us remind that the sample consists only of observed cases where the defendant have been either convicted or settled with the victim. Hence, although we can formally test whether the LEOGVT defendants settle more than to WC workers and CEOs based on the observed sample, to give the true magnitude of the effect needs correction for acquitted cases. We try to overcome this drawback in the next section.

4.3 Estimating Costs of Efforts

In this part, we estimate the parameters of the settlements model from Section 3. The empirical approach aims to account for the acquittals and provides a common metric to

compare different groups.

Using parameters of the theretical model, we can express the expected share of settlements among the cases that we observe as settlements or convictions – $Prob(S|S \vee C)$. We assume that the vengeance utility a is distributed uniformly over an interval $[0, \bar{a}]$. Then, $Prob(S|S \vee G)$ is a combination of the parameter $\bar{a}$ and the costs of efforts of the victim and the defendant, i.e., c_v, c_d:

$$Prob(S|S \vee C; c_v, c_d, \bar{a}) = \frac{Prob(S|c_v, c_d, \bar{a})}{Prob(S|c_v, c_d, \bar{a}) + Prob(C|c_v, c_d, \bar{a})} \tag{3}$$

The formula above comes from the Bayes rule for conditional probabilities, where the unconditional probability of settlement is:

$$Prob(S|c_v, c_d, \bar{a}) = \frac{1}{\bar{a}}\tilde{a}(c_v, c_d, \bar{a}) \tag{4}$$

and the unconditional probability that defendant is convicted can be expressed as:

$$\begin{aligned} Prob(C|c_v, c_d, \bar{a}) &= \frac{1}{\bar{a}} \int_{\tilde{a}(c_v, c_d, \bar{a})}^{\bar{a}} \left(1 + \frac{c_v}{c_d a}\right)^{-1} da \\ &= \frac{1}{\bar{a}} \left(\bar{a} - \tilde{a}(c_v, c_d, \bar{a}) - \frac{c_v}{c_d} ln\left(\bar{a} + \frac{c_v}{c_d}\right) + \frac{c_v}{c_d} ln\left(\tilde{a}(c_v, c_d, \bar{a}) + \frac{c_v}{c_d}\right)\right) \end{aligned} \tag{5}$$

Both unconditional probabilities 4 and 5 depend on the threshold value $\tilde{a}$ which has the following solution:

$$\tilde{a}(c_v, c_d, \bar{a}) = min\left\{\frac{1}{2} + \sqrt{\frac{1}{4} + 2\frac{c_v}{c_d}}, \bar{a}\right\} \tag{6}$$

In equations (3)-(6), we normalize the disutility parameter b for every defendant to one, i.e., $b = 1 \; \forall \; D$. For the estimation, we further normalize the costs of effort c_v and c_d for the baseline group, which consists of individuals who have no permanent work. The normalizations are necessary since we cannot separately identify all the parameters of the model, especially since c_v and c_d enter all the equations as ratios. Hence, any estimate of c_v and c_d captures relative cost of effort to the baseline group, and may confound with changes in b, which we will discuss later.

We modify Regression (2) to accommodate each of seven employment groups separately (for now omitting D's education and car information):

$$S_i = \beta_0 + \sum_{k \in SES} \beta_v^{(k)} \mathbb{1}\{V_i = k\} + \sum_{k \in SES} \beta_d^{(k)} \mathbb{1}\{D_i = k\} + \sum_{k \in SES} \sum_{l \in SES} \beta^{(k,l)} \mathbb{1}\{V_i = k, D_i = l\} + \gamma other\_controls_i + u_i \tag{7}$$

where $SES = \{BC, WC, Entr, CEO, LEOGVT, Ret\}$ as defined in Section 4.1.2; S_i is a dummy which equals to one if the case i settles; $\beta_v^{(k)}$ captures the difference in the shares of settlements with respect to the baseline group of victims with no permanent employment for the victim that belongs to one of the other six socio-economic groups; similarly, $\beta_d^{(k)}$, for the defendant that belongs to a socio-economic group k, and $\beta^{(k,l)}$, for the interaction term between the socio-economic status of the victim, k, and the defendant, l.

Parameters β from Regression 7 provide estimators for the moments of interest. For all $k, l \in SES$, we have:

$$\hat{\beta}_0 \xrightarrow{p} Prob(S|S \vee C; c_v = 1, c_d = 1, \bar{a}) \tag{8}$$

$$\hat{\beta}_v^{(k)} \xrightarrow{p} Prob(S|S \vee C; c_v^{(k)}, c_d = 1, \bar{a}) - Prob(S|S \vee C; c_v = 1, c_d = 1, \bar{a}) \tag{9}$$

$$\hat{\beta}_d^{(k)} \xrightarrow{p} Prob(S|S \vee C; c_v = 1, c_d^{(k)}, \bar{a}) - Prob(S|S \vee C; c_v = 1, c_d = 1, \bar{a}) \tag{10}$$

$$\hat{\beta}_d^{(k,l)} \xrightarrow{p} Prob(S|S \vee C; c_v^{(k)}, c_d^{(l)}, \bar{a}) - Prob(S|S \vee C; c_v^{(k)}, c_d = 1, \bar{a}) - Prob(S|S \vee C; c_v = 1, c_d^{(l)}, \bar{a}) + Prob(S|S \vee C; c_v = 1, c_d = 1, \bar{a}) \tag{11}$$

Hence, we can map empirically estimated moments into theoretical parameters of interest. We combine formulas 8, 9, 10, and 11 with equations 3, 4, 5, and 6 into a system of non-linear equations $\beta = h(\theta)$, which maps a vector of β into the 13×1 vector of θ, where $\theta = \{\bar{a}, c_v^{(BC)}, c_v^{(WC)}, c_v^{(Entr)}, c_v^{(CEO)}, c_v^{(LEOGVT)}, c_v^{(Ret)}, c_d^{(BC)}, c_d^{(WC)}, c_d^{(Entr)}, c_d^{(CEO)}, c_d^{(LEOGVT)}, c_d^{(Ret)}\}'$.

Table 4: Moments for these groups of victims and defendants are used in estimation

		Defendants						
		NoEmpl	BC	WC	Entr	CEO	LEOGVT	Ret
	NoEmpl	✓	✓	✓	✓	✓	✓	✓
	BC	✓	✓	✓	✓	✓	✓	✓
	WC	✓	✓	✓				
Victims	Entr	✓	✓					
	CEO	✓	✓					
	LEOGVT	✓	✓					
	Ret	✓	✓	✓	✓	✓	✓	✓

Using Classical Miminum Distance Estimation we can estimate θ as:

$$\hat{\theta} = argmin_{\theta \in \Theta} \{\hat{\beta} - h(\theta)\}' \widehat{W} \{\hat{\beta} - h(\theta)\} \tag{12}$$

where $\widehat{W}$ is the estimated optimal weighting matrix, which is the inverse of the asymptotic variance of $\hat{\beta}$.

The asymptotic variance of $\hat{\theta}$ is calculated as:

$$\widehat{Avar}(\hat{\theta}) = \left(\hat{H}' W \hat{H}\right)^{-1} \tag{13}$$

where $\hat{H}$ is the 30×13 Jacobian of $h(\theta)$. See the derivations for the Jacobian in Appendix D.

The classical minimum distance estimator has an advantage over maximum likelihood ...[42]

Since we have seven socio-economic groups of offenders and victims, we could have had up to 49 moments to map. However, due to low sample size for some combinations of V and D, not every interaction term is estimable or it is estimated imprecisely. Hence, we focus on $M = 30$ combinations of V and D – mostly populous cells of socio-economic groups – as represented in Table 4. With this, we still have 17 overidentifying restrictions.

Strictly speaking, it is not true that we have overidentifying restrictions. In reality, the cost of effort for any given defendant is likely to vary with the strength of the victim. For example, for the same level of wealth and connections of the defendant, facing a

[42]See more on the Classical Minimum Distance Estimation in Wooldridge, 2010.

richer victim drives up the costs of effort for the defendant. This is sraightforward if we think about a wealth-constrained optimization problem and its mapping back to the uncostrained problem, which we discuss in Appendix B. Hence, ideally, we should have estimated a match specific c_v and c_d for each combination of V and D: $c_v^{(d)}$ and $c_d^{(v)}$. However, in our analysis, we specify just one c_d or c_v per given SES group, and hence our estimates can be seen as a weighted average between different values c_d or c_v when facing different opponents. For example, c_d will mostly reflect the costs of fighting against such victims as $NoEmpl$, BC, and Ret, while c_v will mostly reflect costs against $NoEmpl$ and BC defendants.

4.4 Estimation results

Overall, the estimates of c_v and c_d in Column 1 of Table 5 – our baseline results – capture the expected wealth distribution across different socio-economic groups, mirroring the expected car prices and education levels from Table 1. You can notice that BC workers have lower costs than the NoEmpl group, WC has lower costs than WC, retirees are not that far away from BC workers, etc. It is not surprising, since a tighter wealth constraint at contest and settlement stages is equivalent to a higher cost of effort in our uncostrained model (as discussed in Appendix B). Following this logic, our approach simplify the reality of budget-constrained interactions in criminal justice by solving an equivalent problem with an unconstrained budget but heterogeneous costs of effort.

The higher costs of effort then can be seen as lower access to justice if all groups were equal in wealth. Had CEO defendants and workers without permanent employment have the same high level of wealth, the current criminal justice outcomes for traffic offenses mean that the costs of effort for CEO defendants is just a third of the baseline group. Similarly, among victims, the costs for CEOs is just a fifth the costs of the victims without permanent employment. Note that, since c_v and c_d values are each normalized to their respective baseline groups of NoEmpl workers, we cannot compare directly the value of c_v to c_d.

Most importantly, we see that the cost of effort for LEOGVT defendants is as low (or

Table 5: The estimation results for costs of effort

	(1)	(2)	(3)	(4)
c_d: No permanent empl	1.00	1.00	1.00	1.00
c_d: BC worker	0.71	0.76	0.67	0.48
	(0.06)	(0.07)	(0.09)	(0.13)
c_d: WC worker	0.42	0.56	0.78	1.16
	(0.59)	(0.88)	(0.43)	(1.73)
c_d: Entrepreneur	0.77	0.93	0.78	0.54
	(0.17)	(0.25)	(0.25)	(0.27)
c_d: CEO	0.34	0.43	0.26	0.12
	(0.08)	(0.12)	(0.09)	(0.06)
c_d: LEOGVT	0.26	0.30	0.03	0.04
	(0.07)	(0.09)	(0.02)	(0.02)
c_d: Retiree	0.71	0.84	0.72	0.50
	(0.11)	(0.16)	(0.19)	(0.22)
c_v: No permanent empl	1.00	1.00	1.00	1.00
c_v: BC worker	0.74	0.80	0.59	0.79
	(0.08)	(0.10)	(0.12)	(0.23)
c_v: WC worker	0.70	0.84	0.31	1.06
	(0.15)	(0.21)	(0.17)	(0.51)
c_v: Entrepreneur	1.67	2.81	1.42	0.00
	(0.87)	(1.39)	(1.12)	(0.77)
c_v: CEO	0.23	0.63	0.88	0.67
	(0.16)	(0.43)	(0.43)	(0.68)
c_v: LEOGVT	0.36	0.29	0.55	0.58
	(0.26)	(0.28)	(0.40)	(0.60)
c_v: Retiree	0.79	0.78	0.53	0.24
	(0.08)	(0.11)	(0.12)	(0.14)
$\bar{a}$	6.57	7.17	6.54	12.41
	(0.87)	(1.06)	(1.19)	(3.78)
Additional controls:[1]				
$college^D \times \mathbb{1}\{V_i = k\}$	no	yes	yes	yes
$car^D \times \mathbb{1}\{V_i = k\}$	no	no	no	yes
Sample restricted?	no	no	car info	car info
n obs	6,301	6,195	3,620	3,620
n police deps	1,179	1,169	826	826
χ^2 statistics[2]	25.10	25.53	54.25	54.52

[1] the terms are interacted with all seven SES groups of victims, i.e., in $\mathbb{1}\{V_i = k\}$ we have $k \in \{NoEmpl, BC, WC, Entr, CEO, LEOGVT, Ret\}$. [2] The critical value for the overidentification test using the distribution of χ^2_{17} is 24.77 at 10% siginificance level and 27.59 at 5% significance level.

even lower) than the one for CEO defendants. Given that LEOGVT are not as wealthy, this is likely capturing the effect of non-monetary channels which we discussed in Section 4.2. If we assume that LEOGVT are wealthier than BC workers and entrepreneurs but less wealthy than WC workers and CEOs (based on car prices and education levels), then the wealth-implied cost of effort for LEOGVT defendants should have been somewhere closer to 0.5. This implies that connections and knowledge of the system helps LEOGVT defendants to decrease the costs by half. In other words, the LEOGVT defendants are approximately twice stronger contestants than a comparable wealth group of defendants.

In the second column of Table 5, we report the results for the same regression but adding the control for college education of the defendant interacted with the socio-economic group of victims. The baseline group of defendants now changes to individuals with no permanent employment who have no college degree. Controlling for the defendant's education does not change much the estimates and the relative hierarchy among the SES groups.

Finally, we also look at the sample that has the information on the defendants' cars (see Columns 3 and 4 of Table 5). In this sample, the costs of effort for CEOs is lower than in the overall sample (see Column 3): 0.26 vs 0.43. This could mean that CEOs are especially wealthy in the selected sample of police departments that include information on the cars. However, what is more striking, is that the realtive cost of effort for LEOGVT defendants is especially low, just one tenth of the effort costs of CEOs. This means that non-monetary channels of influence makes LEOGVT defendants at least ten times stronger contestant than comparable wealth group. For some reason, the police departments that post informaiton about the cars seem to be the police departments where the connections matter especially a lot. This hints at some regional heterogeneity in the value of connections, which we will explore in the next section.

When we include the car categories as the control interacted with the victim's SES, the gap in costs of effort between LEOGVT and CEO again narrows down (See Column 4 of Table 5). Note that here the baseline group of defendants chages to individuals with no permanent employment who have no college degree and who drive a medium-price car.

The strikingly low level of effort costs that we estimate for LEOGVT defendants could be alternatively explained by a higher b – the disutlity of conviction – for this group of defendants. We, however, do not believe that this is the case. So far, we normalized b to be the same level across all defendants, and it may potentially confound with our estimates of c_d. Indeed, it may seem reasonable to predict that the disutility of conviction is higher for LEOGVT as opposed to the baseline group of defendants with no permanent work, which can be argued on the basis of a higher opportunity costs linked to reputation and time loss. However, in our analysis we are comparing LEOGVT defendants to white-collar workers, company owners and chief executives. If indeed b is high for LEOGVT, it should be also high for WC workers and CEOs.

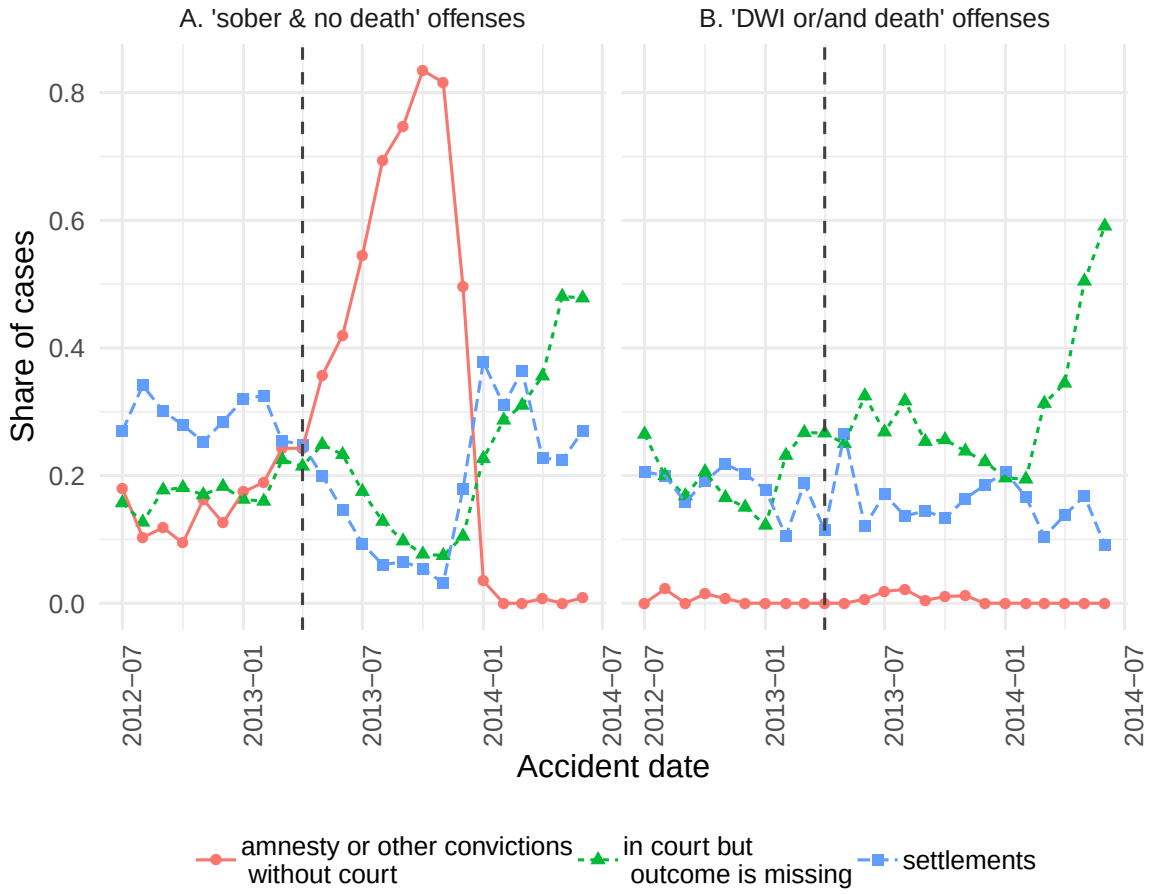
Nevertheless, we still report the χ^2 statistics to test for the overidentifying restrictions. Despite the fact that we know that our restrictions are likely to fail, we still cannot reject the null hypothesis at 5% significance level for the baseline result in Column 1 of Table 5. However, when we use the selected sample with car information, it becomes apparent that our restrictions fail. This means that there is much more variance in the real match-specific costs around the estimated weighted-average costs in the sample with car information. Nevertheless, the estimated averages are still meaningful.

Additionally, in Appendix E we investigate regional heterogeneity and find that the effect of connections and system knowledge is the strongest in smaller cities, rather than in big cities such as Moscow, St Petersburg, and regional capitals. The result probably indicates that the value of connections is the strongest in smaller populations with tighter networks.

4.5 Robustness to missing court outcomes and amnesty

One of the threats to our empirical analysis comes from the fact that we dropped cases with missing court outcome from our analysis. The second concern is the amnesty which was announced in December 2013 and which affected the outcomes of 'sober & no death' category of traffic offenses. Figure 1 shows the evolution of average outcomes for cases grouped by month in which the accident happened. Panel A reports the statistics for

Figure 1: Outcomes by accident date



Note: The vertical dashed line corresponds to April 2013, the date we use as a sample restriction for the robustness check.

'sober & no death' category of traffic offenses in our sample: Panel B, for the cases with DWI or/and death of the victim.

As you can see from both Panel A and Panel B of Figure 1, the share of cases in court with (yet) unkown outcome shoots up for the cases at the end of our sample. It reaches 50-60% of the sample in June 2014 – the last month of accident dates in the sample. It is indicative that the missingness can be (partially) attributed to case duration, i.e., investigators did not have enough time to update records in the database for the latest cases.

Moreover, considerable share of offenders in 'sober & no death' category were amnestied. To remind, almost any driver charged with 'sober & no death' offense category and whose case was still under open investigation or already in court by the time of the amnesty official announcement (mid-December 2013) could plee guilty in

order to be amnestied (and the amnesty applied automatically). Not surprisingly, there was a particular spike in the share of amnestied for the offenses that happened in the period from May to December 2013. The highest numbers (at around 80% of all cases) are for the accidents that happened in September, October, November 2013 – the cases that had not yet been processed by the investigation. Alternatively, it could be explained by the anticipation of the an amnesty, as the Presidential Council for Civil Society and Human Rights – a consultative body to the President of the Russian Federation – first proposed a potential amnesty on September 4, 2013.[43] Notice that with the amnesty the rates of settlements went down. Also the amnesty mechanically brought down the share of missing court outcomes.

In our regressions we use fixed effects for the date (at monthly level) of the traffic offense. However, both the amnesty and missing court outcomes could relate to the duration of the investigation and court hearing, which in turn may relate to the relative strength of the defendant and the victim. Hence, we re-run Regression (2) by restricting the accident dates from July 1, 2012, to April 30, 2013, i.e., the cases prior to the spike in amnesties, prior to any anticipation of the amnesty, or the spike in missing court outcomes.

As you can see from Table 6, the results based on the restricted sample of earlier traffic accidents are broadly in line with the baseline estimates. The sample has lost more than a half of the observations and roughly a third of police departments in comparison to the results in Table E.2. Nevertheless, β_D remains significantly above zero, meaning that the LEOGVT workers manage to settle more often than white-collar workers or CEOs. The results provide confidence that our general conclusions hold.

5 Victim-Defendant Settlements and Access to Justice

In this section, we investigate how the institute of victim-defendant settlements affects access to justice. First, define what access to justice is in our setting.

[43] See the news in Russian language here: `https://web.archive.org/web/20190723090146/https://gulagu.net/news/4421.html`

Definition. Consider the defendant of type D and take two victims of types V_1 and V_2, respectively. V_1 and V_2 have equal access to justice if and only if the probability of conviction does not depend on the victim's type, i.e. $P_C(D, V_1) = P_C(D, V_2)$. Otherwise, V_1 and V_2 have uneqal access to justice.

This notion brings as back to the existing law and economics literature where the identity of a victim does not matter for the case resolution. Indeed, prosecutors must implement all legislative procedures on victims' behalf. Hence, all victims matched against the same defendant should have identical chances of the conviction outcome. Any deviations from this state signal that the criminal justice system does not work well: given the defendant and the severity of crime, some victim types are less likely to end up with the conviction outcome.

To trace the effect of settlements on access to justice, we focus on three defendant types: (1) WC workers, (2) law enforcers and government officials (LEOGVT), and (3) CEOs. As Table 1 shows, WC and LEOGVT workers have comparable wealth. Then, if only monetary channel matters, these two groups matched against the same opponent type must have similar bargaining positions and the probability of conviction. CEO drivers represent the richest group (see Table 1). Nevertheless, as our estimates indicate, their bargaining position is weaker than the one LEOGVT workers hold (see Table 5). This provides evidence in favor of non-monetary fighting abilities.

The analysis builds on three measures. For each victim-defendant match, we restore (1) the settlement probability P_S, (2) the expected unconditional probability of conviction $E(P_C)$ and (3) the expected probability of conviction for non-settled cases $E(P_C|\, a > \tilde{a})$:

$$P_S = \frac{\tilde{a}}{\bar{a}},\ E(P_C) = \frac{1}{\bar{a}} \int_0^{\bar{a}} \frac{ac_v}{ac_v + c_d} da$$

$$E(P_C|\, a > \tilde{a}) = \frac{1}{\bar{a} - \tilde{a}} \int_{\tilde{a}}^{\bar{a}} \frac{ac_v}{ac_v + c_d} da$$

We use the unconditional probability of conviction $E(P_C)$ to measure access to justice in the absence of settlements ("No S" policy). When settlements are available ("S" policy), victims who accept the offer are never worse off compared to the "No S" scenario. Then, we must focus only on victim-defendant matches which move to the contest stage, and

$E(P_C|\, a > \tilde{a})$ covers exactly these cases. To estimate access to justice, we compare $E(P_C)$ (resp. $E(P_C|\, a > \tilde{a})$) over different victim groups under "No S" (resp. "S") policy.

Table 7 reports settlement and conviction probabilities for different victim-defendant matches. Among the three defendant groups, LEOGVT workers have the highest chance to settle with any victim type and the lowest probability of conviction. In this dimension, they over-perform CEOs who are known to be the richest group (see Table 1). As we have already discussed, unobservable non-monetary factors (namely, connections or a better knowledge of the criminal justice system) can explain this pattern. Entrepreneurs represent the weakest victim type: they display the smallest (resp. highest) probability of conviction (resp. settlement).

Table 8 shows how access to justice evolves when settlements are available. According to our definition, access to justice becomes more equal if for a given defendant type, the ratio of conviction probabilities derived for different victim groups gets closer to a unity. In Table 8, we take WC workers as a reference group and compute all the values with respect to it:

$$\text{Access to Justice}^{D}_{V,\,WC} = \frac{E\left(P^{p}_{C}\left(V,\, D\right)\right)}{E\left(P^{p}_{C}\left(WC,\, D\right)\right)} \text{ where } p = \{\text{No S, S}\}$$

First, consider what happens when "No S" policy applies. As Table 8 indicates, BC, WC and RET victims have almost equal chances to reach the conviction outcome. The biggest distortion is observed for entrepreneurs and CEOs. In particular, the latter group has worse access to justice than its WC counterpart. On the contrary, CEO victims are 1.43 times more likely to end up with the conviction outcome.

Second, check how access to justice evolves under "S" policy. If the ratio of conviction probabilities is smaller (resp. bigger) than 1, access to justice improves when the percentage change (namely, $\Delta(\%)$ in Table 8) is positive (resp. negative). Remarkably, the introduction of settlements improves access to justice for all victim types: the ratios become closer to a unity. The biggest gain corresponds to the "ENTR vs WC" ratio in LEO and CEO defendant matches (the 8.1% and 8.5% improvement, respectively). Also, settlements reduce the gap between WC victims and their CEO (resp. LEO) counterparts by more than 6% (resp. 4%). Thus, the positive effect of

settlements on access to justice is more pronounced for groups where the difference in conviction probabilities under "No S" policy is the greatest.

The observed improvement in access to justice can be explained as follows. In non-settled cases, victims are more vindictive (recall the $a > \tilde{a}$ condition) and, hence, exert greater effort, *ceteris paribus*. This drives the equilibrium probability of conviction up. On the contrary, settled cases represent matches where the vindictiveness parameter a is relatively low. Here, victims' incentives to win are not sufficiently strong, and their equilibrium effort declines. If such cases move to the contest stage, the conviction outcome emerges with a relatively small probability. Hence, the expected value of P_C becomes smaller in case of "No S" policy. Moreover, the $\tilde{a}$ threshold increases in c_v implying that victims with a weaker bargaining position settle more often. Hence, if "S" policy applies, the expected value of P_C grows more for disadvantaged groups, and the gap in conviction probabilities gets smaller. Thus, access to justice improves.

6 Conclusion

Most states use Victim-Defendant settlements to solve civil and criminal conflicts. This paper explores how bargaining positions of the parties involved (namely, their preferences, non-monetary fighting abilities and resource constraints) define the case outcome. Also, we discuss the effect Victim-Defendant settlements may have on social welfare. With this approach, the previous work devoted to out-of-court case resolution connects to the literature that focuses on resource imbalances and the inequality before the law.

We construct a stylized theoretical model where two individuals with conflicting interests, the victim and the defendant, must exert effort in order to achieve / avert the court stage. The defendant has an option to settle with the victim before the fight starts, and the optimal offer decreases in his bargaining position. Reaching the agreement is always efficient when the defendant encounters sufficiently high winning benefits. If the victim displays strong preferences for revenge, but the opponent has better fighting abilities, the latter player is willing to enter the contest stage. Hence, even feasible settlements can fail to happen.

To estimate the model, we employ the data on criminal traffic offenses in Russia. Our results show that policemen, when they are involved as defendants, have approximately twice lower cost of effort than a comparable wealth group. We attribute their lower costs to non-monetary resources thanks to their connections or/and the knowledge of the criminal justice system. As a result, policemen manage to settle more often than a comparable wealth group of defendants, which indicates that the policemen's victims were ready to accept lower settlement prices.

The results raise fairness concerns. Since both monetary and non-monetary resources increases the chances of defendant's success in the contest stage, an increase in the defendant's wealth also decrease the settlement price, all other things being equal. The paper contributes to the empirical literature on settlement process.

Although we focused on criminal traffic offenses, the model and the estimation approach proposed in the paper turn to be very general. To push the analysis further, one must specify the objective function of the society and concentrate on the optimal design of the justice system. The criterion may include deterrence and incapacitation concerns, as well as equality considerations. Without this step, it is impossible to give a precise answer when Victim-Defendant settlements must be abandoned, and we leave it for the future.

References

1. Alschuler, A. (1981). "The changing plea bargaining debate." *California Law Review*, 69(3), 652–730.

2. Baye, M., Kovenock, D., de Vries, C.G. (1994). "The solution to the Tullock rent-seeking game when R > 2: mixed-strategy Equilibria and mean dissipation rates." *Public Choice*, 81(3-4), 363-380.

3. Baye, M., Kovenock, D., de Vries, C.G. (2005). "Comparative analysis of litigation systems: An auction-theoretic approach." *Economic Journal*, 115, 583–601.

4. Bebchuk, L.A. (1984) "Litigation and settlement under imperfect information." *RAND Journal of Economics*, 15, 404–415.

5. Esteban, J., Ray, D. (2011). "Linking conflict to inequality and polarization." *American Economic Review*, 101(4), 1345–74.

6. Fiss, O. (1983). "Against settlements." *Yale Lj*, 93, 1073.

7. Garoupa, N., Gravelle, H. (2003). "Efficient deterrence does not require that the wealthy should be able to buy justice." *Journal of Institutional and Theoretical Economics*, 159(3), 545–552.

8. Garoupa, N., Stephen, F. (2008). "Why plea bargaining fails to achieve results in so many criminal justice systems: A new framework for assessment." *Maastricht Journal of European and Comparative Law*, 15(3), 323–358.

9. Glaeser, E., Sacerdote, B. (2003) "Sentencing in homicide cases and the role of vengeance." *Journal of Legal Studies*, 32(2), 363–382.

10. Scheinkman J., Glaeser, E., Shleifer, A. (2003). "The injustice of inequality." *Journal of Monetary Economics*, 50(1), 199–222.

11. Grossman, H., Mendoza, J. (2001). "Butter and guns: Complementarity between economic and military competition." *Economics of Governance*, 2, 25–33.

12. Harland, A. (1982). "Monetary remedies for the victims of crime: Assessing the role of the criminal courts." *UCLA L. Rev.*, 30, 52.

13. Landes, W. (1971). "An economic analysis of the courts." *The Journal of Law and Economics*, 14(1), 61– 107.

14. Langbein, J. (1978). "Torture and plea bargaining." *The University of Chicago Law Review*, 46(1), 3–22.

15. Lott, J. (1987). "Should the wealthy be able to buy justice?" *Journal of Political Economy*, 95(6), 1307– 16.

16. McBride, M., Skaperdas, S. (2014). "Conflict, settlement, and the shadow of the future." *Journal of Economic Behavior and Organization*, 105, 75–89.

17. Merlo, A., Tang, X. (2016). "Bargaining with optimism: A structural analysis of medical malpractice litigation." Working Paper.

18. Okuguchi, K., Szidarovszky, F. (1997). "On the existence and uniqueness of pure Nash equilibrium in rent-seeking games." *Games and Economic Behavior*, 18(1), 135-140.

19. Pérez-Castrillo, D., Verdier, T. (1992). "A general analysis of rent-seeking games." *Public Choice*, 73(3), 335-350.

20. Polinsky, A.M., Rubinfeld, D. (1988). "The deterrent effects of settlements and trials." *International Review of Law and Economics*, 8(1), 109–116.

21. Priest, G., Klein, B. (1984). "The selection of disputes for litigation." *The Journal of Legal Studies*, 13(1), 1–55.

22. Reinganum, J., Wilde, L. (1986). "Settlement, litigation, and the allocation of litigation costs." *RAND Journal of Economics*, 17(4), 557–566.

23. Robson, A., Skaperdas, S. (2008). "Costly enforcement of property rights and the Coase theorem." *Economic Theory*, 36(1), 109–128.

24. Sambanis, N., Skaperdas, S., Wohlforth, W. (2017). "External intervention, identity, and civil war." Working Paper.

25. Sebba, L. "Third Parties: Victims and the Criminal Justice System." The Ohio State University Press, 1996.

26. Shapland, J. (1984). "Victims, the criminal justice system and compensation." *The British Journal of Criminology*, 24(2), 131–149.

27. Sieg, H. (2000). "Estimating a bargaining model with asymmetric information: Evidence from medical malpractice disputes." *Journal of Political Economy*, 108, 1006–21.

28. Silveira, B. (2017). "Bargaining with asymmetric information: An empirical study of plea negotiations." *Econometrica*, 85(2), 419–452.

29. Skaperdas, S. (1996). "Contest success functions." *Economic Theory*, 7, 283–290.

30. Spier, K. "Litigation" in "Handbook of Law and Economics", vol. 1, 259–342. North-Holland, 2007.

31. Strang, H., Sherman, L. (2003). "Repairing the harm: Victims and restorative justice." *Utah L. Rev.*, 1, 15–42.

32. Waldfogel, J. (1995). "The selection hypothesis and the relationship between trial and plaintiff victory." *Journal of Political Economy*, 103, 229–260.

33. Watanabe, Y. (2006). "Learning and bargaining in dispute resolution: Theory and evidence from medical malpractice litigation." Working Paper.

34. Yamazaki, T. (2008). "On the existence and uniqueness of pure-strategy Nash equilibrium in asymmetric rent-seeking contests." *Public Economics Theory*, 10(2), 317-327.

35. Zedner, L. (1994). "Reparation and retribution: Are they reconcilable?" *The Modern Law Review*, 57(2), 228–250.

A Contests with Budget Constraints vs Unconstrained Asymmetric Games

Consider two versions of the contest game. The unconstrained model (UC) was specified in Subsection 3.1. The game with budget constraints (C) has the same primitives, and only two extra assumptions are introduced:

1. $c_i \equiv 1$ for any $i = V,\ D$ and
2. $e_i \leq w_i$ for any $i = V,\ D$ where w_i represents the budget constraint of player i and constitutes common knowledge.

The proof of Proposition 1 finds the unique equilibrium of the UC game:

$$e^*_{V,\,UC} = \frac{a^2 b c_D x}{(a c_D + b c_V)^2},\ e^*_{D,\,UC} = \frac{b^2 a c_V x}{(a c_D + b c_V)^2}$$

Depending on w_V and w_D, the solution of the C model looks differently:

1. Both budget constraints are slack if and only if:

$$\begin{cases} w_V > \frac{a^2 b x}{(a+b)^2} \\ w_D > \frac{b^2 a x}{(a+b)^2} \end{cases}$$

   and the optimal effort levels become:

$$e^*_{V,\,C} = \frac{a^2 b x}{(a+b)^2},\ e^*_{D,\,C} = \frac{b^2 a x}{(a+b)^2}$$

   Imposing $c_i \equiv 1$ for $i = V,\ D$, we get $e^*_{i,\,C} = e^*_{i,\,UC}$.

2. Suppose V's budget constraint binds, but the defendant still has enough resources to sustain the interior solution:

$$\begin{cases} w_V \leq \frac{a^2 b x}{(a+b)^2} \\ w_D > \frac{b^2 a x}{(a+b)^2} \end{cases}$$

   In this case, the optimal effort looks as follows:

$$e^*_{V,\,C} = w_V \text{ and } e^*_{D,\,C} = \sqrt{b x w_V} - w_V \Leftrightarrow w_D > \sqrt{b x w_V} - w_V$$

   Next, we check if there exist c_V and c_D such that $e^*_{i,\,C} = e^*_{i,\,UC}$ for $i = V,\ D$:

$$e^*_{i,\,C} = e^*_{i,\,UC} \Leftrightarrow \begin{cases} w_V = \frac{a^2 b c_D x}{\left(a c_D + b c_V\right)^2} \\ \sqrt{bx w_V} - w_V = \frac{b^2 a c_V x}{\left(a c_D + b c_V\right)^2} \end{cases}$$

$$\Leftrightarrow c_D \equiv 1,\; c_V = \frac{a}{b}\left[\sqrt{\frac{bx}{w_V}} - 1\right] \geq 1 \,\text{for}\, w_V \leq \frac{a^2 bx}{\left(a+b\right)^2}$$

In words, a player whose budget constraint binds faces a higher effort cost and, hence, holds a weaker bargaining position.

3. Assume D's (V's) budget constraint becomes active (slack):

$$\begin{cases} w_V > \frac{a^2 bx}{\left(a+b\right)^2} \\ w_D \leq \frac{b^2 ax}{\left(a+b\right)^2} \end{cases}$$

and the optimal effort is:

$$e^*_{V,\,C} = \sqrt{axw_D} - w_D \Leftrightarrow w_V > \sqrt{axw_D} - w_D \,\text{and}\, e^*_{D,\,C} = w_D$$

To find c_V and c_D such that $e^*_{i,\,C} = e^*_{i,\,UC}$ for $i = V,\, D$, we solve:

$$e^*_{i,\,C} = e^*_{i,\,UC} \Leftrightarrow \begin{cases} \sqrt{axw_D} - w_D = \frac{a^2 b c_D x}{\left(a c_D + b c_V\right)^2} \\ w_D = \frac{b^2 a c_V x}{\left(a c_D + b c_V\right)^2} \end{cases}$$

$$\Leftrightarrow c_D = \frac{b}{a}\left[\sqrt{\frac{ax}{w_D}} - 1\right] \geq 1 \,\text{for}\, w_D \leq \frac{b^2 ax}{\left(a+b\right)^2},\; c_V \equiv 1$$

4. Both budget constraints bind under the following conditions:

$$\begin{cases} w_V \leq \frac{a^2 bx}{\left(a+b\right)^2} \\ w_D \leq \frac{b^2 ax}{\left(a+b\right)^2} \end{cases}$$

and $e^*_{i,\,C} = e^*_{i,\,UC}$ for $i = V,\, D$ requires:

$$e^*_{i,\,C} = e^*_{i,\,UC} \Leftrightarrow \begin{cases} w_V = \frac{a^2 b c_D x}{\left(a c_D + b c_V\right)^2} \\ w_D = \frac{b^2 a c_V x}{\left(a c_D + b c_V\right)^2} \end{cases}$$

$$\Leftrightarrow c_D = \frac{b w_V}{\left(w_V + w_D\right)^2},\; c_V = \frac{a w_D}{\left(w_V + w_D\right)^2}$$

This example shows how one can transform a contest game with budget constraints into an unconstrained model by choosing c_V and c_D appropriately.

B Proofs

Proposition 1. *The equilibrium of the contest stage exists and is unique.*

Proof. Rewrite the optimization programs as follows:

$$\begin{aligned} V: &\quad \max_{e_V} \left\{ \tilde{a}x \left(\tfrac{e_V}{e_V+e_D} \right) - \tfrac{e_V}{c_V} \right\} \\ D: &\quad \max_{e_D} \left\{ \tilde{b}x \left(\tfrac{e_V}{e_V+e_D} \right) - e_D \right\} \end{aligned}$$

where $\tilde{a} = \frac{a}{c_V}$ and $\tilde{b} = \frac{b}{c_D}$. First-order conditions are:

$$\begin{aligned} BR_V(e_D): &\quad \tilde{a}xe_D = (e_D + e_V)^2 \\ BR_D(e_V): &\quad \tilde{b}xe_V = (e_D + e_V)^2 \end{aligned}$$

The second-order derivatives of $\pi_D(\cdot)$ and $\pi_V(\cdot)$ are always negative. Thus, any e_V and e_D that satisfy first-order conditions correspond to an interior maximum. Solving the system of FOCs delivers:

$$\begin{aligned} e_V^* = \tfrac{\tilde{a}^2\tilde{b}x}{(\tilde{a}+\tilde{b})^2},\ e_D^* &= \tfrac{\tilde{a}\tilde{b}^2x}{(\tilde{a}+\tilde{b})^2} \\ \pi_V^* &= \left(\tfrac{\tilde{a}}{\tilde{a}+\tilde{b}} \right) \left(a - c_V \tfrac{\tilde{a}\tilde{b}}{\tilde{a}+\tilde{b}} \right) x \\ \pi_D^* &= -\left(\tfrac{\tilde{a}}{\tilde{a}+\tilde{b}} \right) \left(b + c_D \tfrac{\tilde{b}^2}{\tilde{a}+\tilde{b}} \right) x \end{aligned}$$

where asterisks denote equilibrium effort levels and expected payoffs. Since both $\pi_V(\cdot)$ and $\pi_D(\cdot)$ are strictly concave in e_V and e_D, respectively, the uniqueness follows. □

Proposition 2. *Contestant i's equilibrium effort e_i^* always increases in his / her valuation of punishment and decreases in c_i:*

$$\frac{\partial e_V^*}{\partial a} \geq 0,\ \frac{\partial e_D^*}{\partial b} \geq 0,\ \frac{\partial e_i^*}{\partial c_i} \leq 0 \, for \, i = V,\ D$$

For $\frac{a}{c_V} \geq \frac{b}{c_D}$:

1. e_V^* *increases in* b *and decreases in* c_D*;*

2. e_D^* *decreases in* a *and increases in* c_V*.*

For $\frac{a}{c_V} < \frac{b}{c_D}$*:*

1. e_V^* *strictly decreases in* b *and strictly increases in* c_D*;*

2. e_D^* *strictly increases in* a *and strictly decreases in* c_D*.*

Proof. First, check how e_V^* and e_D^* depend on $\tilde{a}$ and $\tilde{b}$, respectively:

$$\frac{\partial e_V^*}{\partial \tilde{a}} = \frac{2\tilde{a}\tilde{b}^2}{\left(\tilde{a}+\tilde{b}\right)^3} \geq 0$$

$$\frac{\partial e_D^*}{\partial \tilde{b}} = \frac{2\tilde{b}\tilde{a}^2}{\left(\tilde{a}+\tilde{b}\right)^3} \geq 0$$

Since $\tilde{a} = \frac{a}{c_V}$ and $\tilde{b} = \frac{b}{c_D}$, we obtain:

$$\frac{\partial e_V^*}{\partial a} = \frac{\partial e_V^*}{\partial \tilde{a}}\frac{\partial \tilde{a}}{\partial a} = \frac{\partial e_V^*}{\partial \tilde{a}} \geq 0,\ \frac{\partial e_V^*}{\partial c_V} = \frac{\partial e_V^*}{\partial \tilde{a}}\frac{\partial \tilde{a}}{\partial c_V} = -\frac{\partial e_V^*}{\partial \tilde{a}} \leq 0$$

$$\frac{\partial e_D^*}{\partial b} = \frac{\partial e_D^*}{\partial \tilde{b}}\frac{\partial \tilde{b}}{\partial b} = \frac{\partial e_D^*}{\partial \tilde{b}} \geq 0,\ \frac{\partial e_D^*}{\partial c_D} = \frac{\partial e_D^*}{\partial \tilde{b}}\frac{\partial \tilde{b}}{\partial c_D} = -\frac{\partial e_D^*}{\partial \tilde{b}} \leq 0$$

Second, compute the derivatives of e_V^* and e_D^* with respect to $\tilde{b}$ and $\tilde{a}$:

$$\frac{\partial e_V^*}{\partial \tilde{b}} = \frac{\tilde{a}^2\left(\tilde{a}-\tilde{b}\right)}{\left(\tilde{a}+\tilde{b}\right)^3} \geq 0 \Leftrightarrow \tilde{a} \geq \tilde{b}$$

$$\frac{\partial e_D^*}{\partial \tilde{a}} = \frac{\tilde{b}^2\left(\tilde{b}-\tilde{a}\right)}{\left(\tilde{a}+\tilde{b}\right)^3} \leq 0 \Leftrightarrow \tilde{a} \geq \tilde{b}$$

where $\frac{\partial e_V^*}{\partial b} \geq 0$, $\frac{\partial e_V^*}{\partial c_D} \leq 0$ and $\frac{\partial e_D^*}{\partial a} \leq 0$, $\frac{\partial e_D^*}{\partial c_V} \geq 0$ for $\tilde{a} \geq \tilde{b}$ follow. Similarly, the signs of $\frac{\partial e_V^*}{\partial \tilde{b}}$ and $\frac{\partial e_D^*}{\partial \tilde{a}}$ for $\tilde{a} < \tilde{b}$ can be found. □

Proposition 3. *The optimal settlement offer* S^* *always:*

- *Decreases in* D*'s willingness to win* b *and in* V*'s fighting cost* c_V *and*

- *Increases in* V*'s willingness to win* a *and in* D*'s fighting cost* c_D*.*

Proof. The optimal settlement offer S^* equals to V's equilibrium payoff π_V^* (Lemma 1):

$$S^* = \frac{a^3 c_D^2 x}{\left(ac_D + bc_V\right)^2}$$

First, consider how S^* depends on preference parameters a and b:

$$\frac{\partial S^*}{\partial a} = \frac{a^2 c_D^2 x \left(3bc_V + ac_D\right)}{\left(ac_D + bc_V\right)^3} \geq 0$$

$$\frac{\partial S^*}{\partial b} = -\frac{2a^3 c_D^2 x}{\left(ac_D + bc_V\right)^3} \leq 0$$

Second, check the response of S^* to changes in c_V and c_D:

$$\frac{\partial S^*}{\partial c_V} = -\frac{2a^3 c_D^2 x}{\left(ac_D + bc_V\right)^3} \leq 0$$

$$\frac{\partial S^*}{\partial c_D} = \frac{2a^3 b c_V c_D x}{\left(ac_D + bc_V\right)^3} \geq 0$$

□

Proposition 4. *There exists* $\tilde{a} \equiv \tilde{a}\left(b,\, c_V,\, c_D\right) > b$ *such that* V *and* D *settle for any* $a \in [0,\, \tilde{a}]$.

Proof. The optimal settlement offer is $S^* = \pi_V^*$ (Lemma 1). The game does not proceed to the contest stage if and only if:

$$\pi_V^* = S^* \leq -\pi_D^* \Leftrightarrow a^2 c_D \leq b\left(ac_D + 2bc_V\right)$$

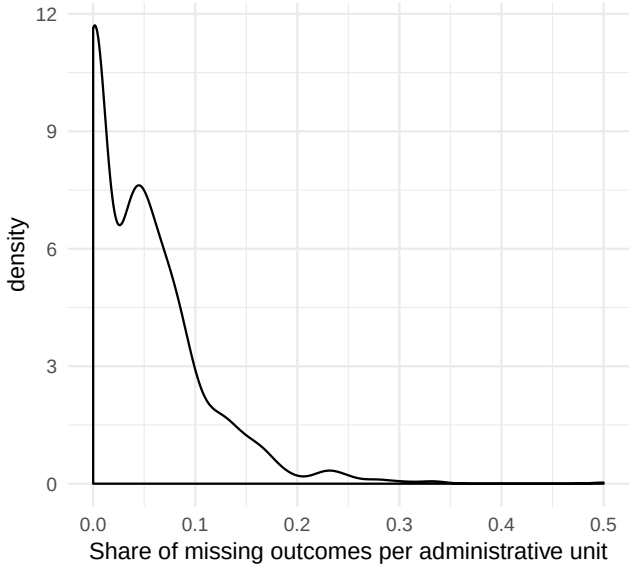
This inequality holds for:

$$a \in \left[b\left(\frac{1}{2} - \sqrt{\frac{1}{4} + 2\frac{c_V}{c_D}}\right),\, b\left(\frac{1}{2} + \sqrt{\frac{1}{4} + 2\frac{c_V}{c_D}}\right) \right]$$

where the lower bound is negative. Defining $\tilde{a} = \min\left\{ b\left(\frac{1}{2} + \sqrt{\frac{1}{4} + 2\frac{c_V}{c_D}}\right),\, \bar{a} \right\} > b$, we obtain the statement of the proposition. □

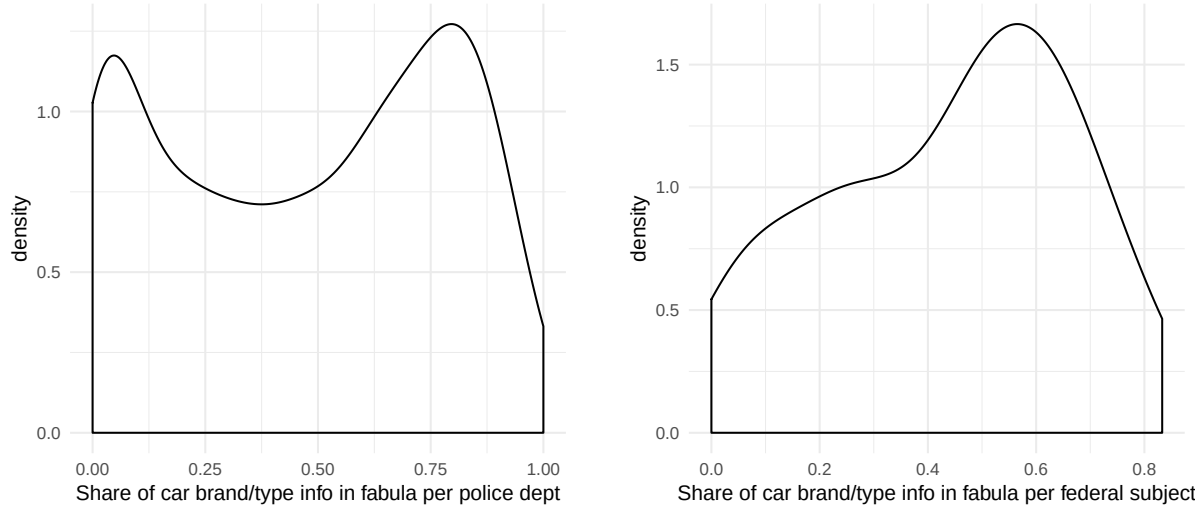
C Tables and Figures

Figure C.1: The incidence of missing case outcome information per administrative unit (only units with >10 obs)



The administrative unit is defined on page X. To remove noise, we kept only the administrative units with more than ten observations. This figure shows that the investigators in most administrative units timely fill in the updates on the outcomes per each case. However, there are some administrative units where the investigators do not update the information in a timely manner (the long right tail of the distribution). This heterogeneity cannot be explained by the difference in the timing of the accidents across the administrative units, as there is no correlation between the average time of the accident per adm unit and the share of missing information per unit. The figure is based on all traffic offenses that happened between July 2012 to June 2014, including car-to-car, car-to-pedestrian, and car-to-obstacle crashes.

Figure C.2: The presence of car type/brand information



(a) per police department (only those with >10 obs)

(b) per federal subject (only those with >30 obs)

On the subfigure (a) you can see that distribution of the presence of car information is bimodal. It means that investigators in some police departments almost always indicate the type of car in the fabula while in the other police departments they usually omit such information. Moreover, the difference in the fabula filling across police departments translates into substantial heterogeneity at the regional level, as seen from subfigure (b). The figures are based on all traffic offenses that happened between July 2012 to June 2014, including car-to-car, car-to-pedestrian, and car-to-obstacle crashes.

D Additional derivations

D.1 The Jacobian of the Observed Share of Settlements

The first order derivatives for $Prob(S|S \vee G; c_v, c_d, \bar{a})$ with respect to any of its parameters have the following formula:

$$\frac{\partial Prob(S|S \vee G; c_v, c_d, \bar{a})}{\partial x} = \frac{Prob(G|c_v, c_d, \bar{a}) \frac{\partial Prob(S|c_v, c_d, \bar{a})}{\partial x} - Prob(S|c_v, c_d, \bar{a}) \frac{\partial Prob(G|c_v, c_d, \bar{a})}{\partial x}}{\left(Prob(S|c_v, c_d, \bar{a}) + Prob(G|c_v, c_d, \bar{a})\right)^2}$$

where x can be c_v, c_d, or $\bar{a}$.

$$\begin{aligned}
\frac{\partial Prob(S|c_v, c_d, \bar{a})}{\partial \bar{a}} &= -\frac{\tilde{a}(c_v, c_d)}{\bar{a}^2} \\
\frac{\partial Prob(S|c_v, c_d, \bar{a})}{\partial c_v} &= \frac{1}{\bar{a}} \frac{\partial \tilde{a}(c_v, c_d)}{\partial c_v} \\
\frac{\partial Prob(S|c_v, c_d, \bar{a})}{\partial c_d} &= \frac{1}{\bar{a}} \frac{\partial \tilde{a}(c_v, c_d)}{\partial c_d}
\end{aligned}$$

Since $\hat{\tilde{a}} << \hat{\bar{a}}$, we know that the partial derivative of $\tilde{a}(c_v, c_d)$ with respect to $\bar{a}$ is zero at the estimated values.

$$\frac{\partial Prob(G|c_v, c_d, \bar{a})}{\partial \bar{a}} = \frac{1}{\bar{a}}\left(1 - \frac{c_v}{c_d}\frac{1}{(\bar{a} + \frac{c_v}{c_d})} - Prob(G|c_v, c_d, \bar{a})\right)$$

$$\frac{\partial Prob(G|c_v, c_d, \bar{a})}{\partial c_v} = \frac{1}{\bar{a}}\left(-\frac{\partial \tilde{a}(c_v, c_d)}{\partial c_v} + \frac{1}{c_d} ln\left(\tilde{a}(c_v, c_d) + \frac{c_v}{c_d}\right) + \frac{c_v}{c_d}\frac{1}{(\tilde{a}(c_v, c_d) + \frac{c_v}{c_d})}\left(\frac{\partial \tilde{a}(c_v, c_d)}{\partial c_v} + \frac{1}{c_d}\right.\right.$$
$$- \frac{1}{\bar{a}}\left(\frac{1}{c_d} ln\left(\bar{a} + \frac{c_v}{c_d}\right) + \frac{c_v}{c_d}\frac{1}{(\bar{a} + \frac{c_v}{c_d})}\frac{1}{c_d}\right)$$

$$\frac{\partial Prob(G|c_v, c_d, \bar{a})}{\partial c_d} = \frac{1}{\bar{a}}\left(-\frac{\partial \tilde{a}(c_v, c_d)}{\partial c_d} - \frac{c_v}{c_d^2} ln\left(\tilde{a}(c_v, c_d) + \frac{c_v}{c_d}\right) + \frac{c_v}{c_d}\frac{1}{(\tilde{a}(c_v, c_d) + \frac{c_v}{c_d})}\left(\frac{\partial \tilde{a}(c_v, c_d)}{\partial c_d} - \frac{c_v}{c_d^2}\right.\right.$$
$$+ \frac{1}{\bar{a}}\left(\frac{c_v}{c_d^2} ln\left(\bar{a} + \frac{c_v}{c_d}\right) + \frac{c_v^2}{c_d^3}\frac{1}{(\bar{a} + \frac{c_v}{c_d})}\right)$$

and finally

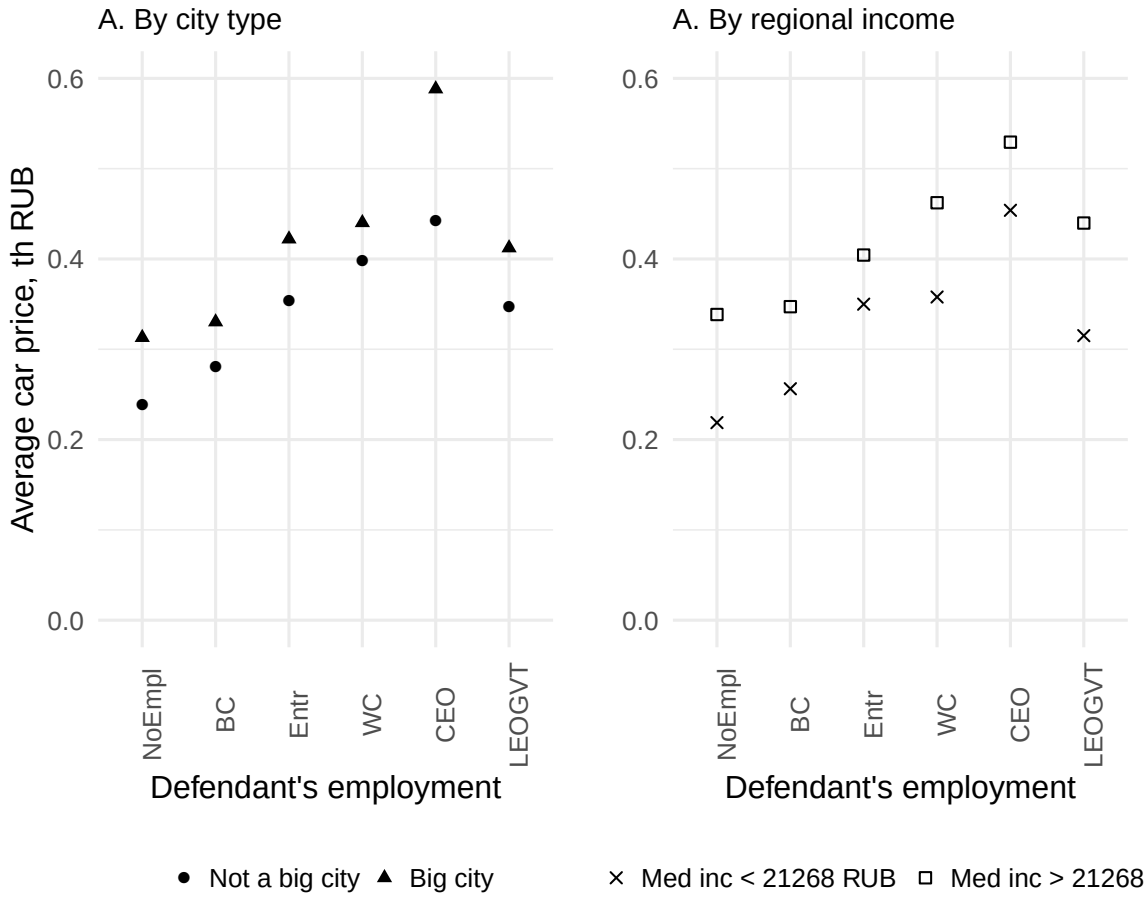
$$\frac{\partial \tilde{a}(c_v, c_d)}{\partial c_v} = \left(\frac{1}{4} + 2\frac{c_v}{c_d}\right)^{-0.5}\frac{1}{c_d}$$
$$\frac{\partial \tilde{a}(c_v, c_d)}{\partial c_d} = -\left(\frac{1}{4} + 2\frac{c_v}{c_d}\right)^{-0.5}\frac{c_v}{c_d^2}$$

E Regional heterogeneity

In this section, we explore whether the additional benefit of connections and knowledge of the system differs across regions. We return back to our reduced form setup in Regression (2) without education or car category controls, i.e.:

$$\begin{aligned} S_i = {} & \alpha + \beta_D leogvt_i^D + \beta_V leogvt_i^V + \beta_{DV} leogvt_i^D \times leogvt_i^V \\ & + \psi_1 highses_i^D + \psi_2 highses_i^V + \psi_3 highses_i^D \times highses_i^V \\ & + \gamma other\_controls_i + u_i \end{aligned} \tag{E.1}$$

Figure E.3: Average car prices by employment group across different samples



First, we estimate Regression (E.1) on two different samples: cases in big cities versus cases in smaller cities and localities. Big cities include Moscow, Saint-Petersburg, and the biggest city in each federal subject as identified by OKATO code.[44] We cannot reject the null hypothesis that LEOGVT are no stronger than white-collar workers or CEO in big cities, although the point estimates still point in the same direction. At the same time, the effect of the non-monetary channel seems to be pronounced for smaller cities or rural areas.

Second, we split the federal subjects into two subsamples based on the median monthly wage income in the region.[45] We do not find any substantial heterogeneity in the results across the regional income groups.

[44]The biggest city in a federal subject is denoted with 401 as the third to fifth digits in an OKATO code. Moscow and Saint-Petersburg are federal subjects in itself.

[45]The statistics on the median wage income across different regions of Russia can be found at the Federal State Statistics Service website: `www.gks.ru/free_doc/new_site/population/bednost/tabl/3-1-5.doc`.

If we look at the average car prices for different defendant groups across the different samples (Figure E.3), we see that the population is wealther in big cities (rather than in smaller cities or villages) and in regions with higher median income (rather than in lower income regions). However, overall, the relative values of car prices across different employment groups remain similar across all subsamples. Notably, the gap between the car prices of CEOs and of LEOGVT workers is the largest in big cities. Perhaps, this creates a larger downward bias – from the uncaptured wealth differential between the two groups – for β_D in the sample of big cities, explaining the failure to reject the null hypothesis.

Ideally, we would love to investigate the heterogeneity for each region. However, given that the sample of LEOGVT workers is not that big as we have only two years of observations, we refrain from investigating the regional heterogeneity any further. However, if we or other researchers get data for more years of observation, this kind of analysis would be certainly possible.

Table 6: Settlement Rates for Law Enforcement Officers and GVT Officials: The restricted sample

	(1)	(2)
β_D	0.321	0.715
	(0.127)	(0.141)
p-value for H0: $\beta_D = 0$	0.012	0.000
β_V	-0.157	-0.155
	(0.123)	(0.195)
p-value for H0: $\beta_V = 0$	0.202	0.427
β_{VD}	-0.300	-0.732
	(0.242)	(0.370)
p-value for H0: $\beta_{DV} = 0$	0.216	0.049
Wealth controls:		
$highses^{D,V}$, $highses^D \times highses^V$	yes	yes
$college^D$, $college^D \times highses^V$	yes	yes
car^D, $car^D \times highses^V$	no	yes
Sample restricted?	no	car info
n obs	1,841	1,053
n police deps	719	467

Note:
This table reports the estimates of coefficients β_D, β_V, and β_{DV} for regression (2). Standard errors (in parentheses) are clustered at the police department level. All regressions include police department, hour× day-of-week, date of accident (in months), date of accident (in months), × 'sober & no death' offense category, and offense-category fixed effects. Column 1 reports results controlling for the socio-economic status of the defendant and the control for D's education, while in Column 2, it also accounts for the car category. The data includes one-car-one-pedestrian criminal traffic offenses that happened between July 2012 to April 2013 for 'sober & no death' offenses and between July 2012 to April 2014 for other offense categories, which were prosecuted or settled. Cases with missing court outcomes have been excluded. The sample of defendants and victims is restricted to NoEmpl, BC, WC, CEO, and LEOGVT workers only.

Table 7: Settlement and Conviction Probabilities for Different Defendant Groups

		P_S			$E(P_C \mid a > \tilde{a}) \equiv E(P_C^{\mathrm{S}})$			$E(P_C) \equiv E(P_C^{\mathrm{No\ S}})$		
		D**'s Type**								
		LEO	WC	CEO	LEO	WC	CEO	LEO	WC	CEO
V**'s Type**	U	0.40	0.34	0.36	0.52	0.67	0.58	0.41	0.56	0.46
		(0.07)	(0.16)	(0.06)	(0.07)	(0.16)	(0.06)	(0.08)	(0.18)	(0.06)
	BC	0.36	0.31	0.32	0.59	0.72	0.64	0.47	0.61	0.53
		(0.07)	(0.14)	(0.05)	(0.08)	(0.16)	(0.06)	(0.08)	(0.18)	(0.07)
	WC	0.35	0.30	0.32	0.60	0.73	0.65	0.49	0.62	0.55
		(0.07)	(0.14)	(0.06)	(0.09)	(0.16)	(0.07)	(0.09)	(0.18)	(0.08)
	ENTR	0.48	0.39	0.43	0.45	0.61	0.50	0.34	0.48	0.39
		(0.14)	(0.19)	(0.11)	(0.14)	(0.18)	(0.14)	(0.14)	(0.20)	(0.14)
	CEO	0.25	0.23	0.23	0.79	0.85	0.83	0.70	0.78	0.75
		(0.06)	(0.10)	(0.05)	(0.11)	(0.14)	(0.09)	(0.13)	(0.17)	(0.11)
	LEO	0.29	0.26	0.26	0.72	0.81	0.76	0.62	0.72	0.67
		(0.08)	(0.13)	(0.06)	(0.13)	(0.16)	(0.11)	(0.15)	(0.19)	(0.14)
	RET	0.37	0.31	0.33	0.57	0.71	0.63	0.46	0.60	0.52
		(0.07)	(0.15)	(0.05)	(0.08)	(0.16)	(0.06)	(0.08)	(0.18)	(0.07)

Note:
Groups correspond to: U – unemployed individuals or those who have no permanent job; BC (WC) – blue (white) collar workers; ENTR – entrepreneurs; LEO – law enforcement officers; RET – retired individuals. The calculation is based on the "Car vs Pedestrian" sample and 10'000 random draws. Standard deviations in parentheses.

Table 8: "Victim-Defendant" Settlements and Access to Justice

		D**'s Type**								
		LEO			WC			CEO		
		No S	S	Δ (%)	No S	S	Δ (%)	No S	S	Δ (%)
V**'s Type**	U vs WC	0.84	0.87	+3.6	0.90	0.92	+1.6	0.84	0.89	+6.7
	BC vs WC	0.96	0.98	+2.5	0.98	0.0.99	+0.2	0.96	0.98	+2.2
	ENTR vs WC	0.69	0.75	+8.1	0.77	0.84	+7.9	0.71	0.77	+8.5
	CEO vs WC	**1.43**	**1.32**	**−7.8**	**1.26**	**1.16**	**−7.4**	**1.36**	**1.28**	**−6.4**
	LEO vs WC	**1.27**	**1.20**	**−5.2**	**1.16**	**1.11**	**−4.5**	**1.22**	**1.17**	**−4.0**
	RET vs WC	0.94	0.95	+1.2	0.97	0.97	+0.5	0.95	0.97	+2.5

Note:
Groups correspond to: U – unemployed individuals or those who have no permanent job; BC (WC) – blue (white) collar workers; ENTR – entrepreneurs; LEO – law enforcement officers; RET – retired individuals. The calculation is based on the "Car vs Pedestrian" sample and 10'000 random draws. Standard deviations in parentheses. Percentage change in access to justice is

$$\Delta = \left(\frac{E\left(P_C^{\mathrm{S}}\left(V,\, D\right)\right)}{E\left(P_C^{\mathrm{S}}\left(WC,\, D\right)\right)} \cdot \left[\frac{E\left(P_C^{\mathrm{No\ S}}\left(V,\, D\right)\right)}{E\left(P_C^{\mathrm{No\ S}}\left(WC,\, D\right)\right)} \right]^{-1} - 1 \right) \cdot 100\%$$

Table C.1: Settlement Rates for Law Enforcement Officers and GVT Officials (all coefficients)

	(1)	(2)	(3)	(4)
$leogvt^D$	0.225	0.265	0.464	0.471
	(0.0817)	(0.0883)	(0.129)	(0.128)
$leogvt^V$	-0.130	-0.119	-0.188	-0.207
	(0.0752)	(0.0759)	(0.108)	(0.106)
$leogvt^D \times leogvt^V$	-0.588	-0.626	-0.204	-0.316
	(0.322)	(0.316)	(0.207)	(0.211)
$highses^D$	0.0748	0.0110	0.00556	0.00977
	(0.0416)	(0.0424)	(0.0581)	(0.0563)
$highses^V$	-0.0350	-0.00941	-0.0158	-0.125
	(0.0333)	(0.0380)	(0.0463)	(0.152)
$highses^D \times highses^V$	0.0717	0.121	-0.112	-0.127
	(0.136)	(0.157)	(0.156)	(0.150)
$college^D$		0.145	0.151	0.150
		(0.0196)	(0.0288)	(0.0283)
$college^D \times highses^V$		-0.129	-0.0982	-0.0659
		(0.0869)	(0.101)	(0.112)
$cheap\_car$				0.0249
				(0.0361)
$mediumprice\_car$				0.0102
				(0.0337)
$expensive\_car$				0.167
				(0.126)
$cheap\_car \times highses^V$				0.0201
				(0.152)
$mediumprice\_car \times highses^V$				0.214
				(0.151)
$expensive\_car \times highses^V$				-0.309
				(0.237)
dwi&no death	-0.0272	-0.0247	-0.0916	-0.0889
	(0.0341)	(0.0345)	(0.0386)	(0.0388)
sober&death	0.0336	0.0422	0.0137	0.0143
	(0.0264)	(0.0266)	(0.0316)	(0.0322)
dwi&death	-0.185	-0.179	-0.143	-0.146
	(0.0347)	(0.0346)	(0.0502)	(0.0495)
_cons	0.304	0.286	0.343	0.320
	(0.0648)	(0.0663)	(0.0868)	(0.0937)
N obs	4347	4265	2529	2529

This table reports also all other coefficients for Regression (2). See notes for Table E.2 for more information.

Table E.2: Settlement Rates for Law Enforcement Officers and GVT Officials

	Big city?[1]		Regional median income[2]	
	yes	no	low	high
	(1)	(2)	(3)	(4)
β_D	0.076	0.421	0.268	0.213
	(0.110)	(0.107)	(0.125)	(0.111)
p-val H0: $\beta_D = 0$	0.487	0.000	0.032	0.055
β_V	-0.039	-0.172	-0.021	-0.262
	(0.105)	(0.098)	(0.124)	(0.093)
p-val H0: $\beta_V = 0$	0.708	0.078	0.863	0.005
β_{VD}	n/a	-0.502	-0.120	-0.931
	n/a	(0.296)	(0.210)	(0.188)
p-val H0: $\beta_{DV} = 0$	n/a	0.091	0.568	0.000
n obs	1,844	2,503	2,278	1,881
n police deps	255	753	670	309

Note:
This table reports the estimates of coefficients β_D, β_V, and β_{DV} for regression (E.1) across different samples. Standard errors (in parentheses) are clustered at the police department level. All regressions include police department, hour× day-of-week, date of accident (in months), and offense-category fixed effects. The data includes one-car-one-pedestrian criminal traffic offenses that happened between July 2012 to June 2014, which were prosecuted or settled. The sample of defendants and victims is restricted to NoEmpl, BC, WC, CEO, and LEOGVT workers only. [1] Big cities include Moscow, St Petersburg, and the largest administrative city in each other Federal subject; [2] The median monthly income per Federal subject in 2013 as reported by Goskomstat (the statistical office of Russian Federation). It is considered to be low if below the median value for the Russian Federation, which was 21,268 Rubles in 2013, or high, otherwise. [3] Petty corruption in the Federal Subject is low if the petty corruption index 2010 is $\in \{A, B, C\}$ and high if the index is $\in \{D, E\}$.