FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

Political connections, corruption and road traffic safety: a Russian case study

Seppe Van Den Berge
Student number: 01708259

Supervisor: Tom Eeckhout

Master's dissertation submitted in order to obtain the degree of Master of Science in Business Engineering - Operations Management

Academic year 2024-2025

GHENT UNIVERSITY

CONFIDENTIALITY AGREEMENT

PERMISSION

I declare that the content of this Master's Dissertation may be consulted and/or reproduced, provided that the source is referenced.

Seppe Van Den Berge

FOREWORD

Who would have thought I'd be writing a foreword section for a thesis? I applied for this subject five years ago, during my first year as a master's student. By the end of that year, I stopped studying as I felt the need to explore life outside of academics.
And yet, here I am. I came back, and this is the final assignment before I can close this chapter. Having now worked as a student employee for three years, this moment has been closing in on me steadily.

Throughout this journey, I often wondered: why am I studying? But then again: why not? There's a lot one could say about it, but one thing is certain: these years have taught me how to cope with unbearable stress and those end-of-the-world feelings. I've learned what it's like to cram an entire semester's worth of knowledge into a single week, to somehow meet impossible deadlines, and to discover deep and surprising parts of myself in the process.

That being said, I should probably thank a few people here!

First of all, I want to thank Frauke Cuelenaere,
who made me see the bridge again towards studying. She made every decision path seem as an evident choice, when I was lost on the crossroads.

Evidently, Tom Eeckhout, for being a 'super'visor.
With weekly meetings of one or two hours, we dove together deep into the next steps to take, towards obtaining this end result.

Then the personal circle:
My parents, for always giving me everything I needed during study period, and beyond.
Always on standby when deadlines loomed.

My three older sisters, who were there if I couldn't see the forest for the trees. I want to especially thank Sanne. I'm not sure I would have finished this course of study without her.

Ellen, for proposing three times (!) to read my thesis, and then also doing it.
Are you crazy?

And of course, the person closest to me during these six months of thesis work: Kalliopi. You gave me the space, peace and energy to get started. You sometimes wondered if you were too much of a distraction, but in reality, you made the whole process so much easier.

Thank you all, and now, let's move on to the next stage of life!

TABLE OF CONTENTS

1 INTRODUCTION 19

2 LITERATURE REVIEW 21

2.1 Theoretical foundations: corruption, nepotism, and informal influence 21

2.2 Why informal enforcement happens: cultural and institutional roots 22

2.3 Why informal enforcement matters: fairness, risk and social cost 23

2.4 Vanity plates as a symbol of informal privilege 25

2.5 Measuring nepotism and its effects 27

3 DATA 29

3.1 Data Sources 29

3.1.1 Traffic Transaction Data (1993–2006) 29

3.1.2 Pension and Employment Data (1999–2004) 31

3.2 Data cleaning and preprocessing 32

3.2.1 Period restriction and identification logic 32

3.2.2 Filtering non-individuals 33

3.2.3 Address standardization and key construction 34

3.2.3.1 Street name standardization 34

3.2.3.2 Address validation and filtering 35

3.2.3.3 Key construction and matching 36

3.2.4 Family and household construction 37

3.2.5 Cleaning pension and employment data 38

3.2.5.1 Standardizing name fields 38

3.2.5.2 Handling missing birthdates 39

3.2.5.3 Ministry employee identification 40

3.2.5.4 Aggregating to the individual level 40

3.2.6 Dataset merging 41

3.2.6.1 Ministry flag assignment via record linkage 41
3.2.6.2 False match risk 41
3.2.6.3 Identifying family members of ministry employees 42
3.2.6.4 Summary statistics 43
4 METHODOLOGY 45
4.1 Overview of empirical strategy 45
4.2 Defining connectedness and informal influence 45
4.2.1 Ministry employment and family ties (Level 0 & Level 1) 46
4.2.2 Informal signals of privilege 46
4.2.3 Stages of enforcement and points of influence 47
4.3 Hypotheses and empirical framing 49
4.3.1 Hypotheses and conceptual framing 49
4.3.2 Analytical setup and measurement framework 51
4.3.3 Model specifications and estimation strategy 53
4.3.3.1 Descriptive analysis approach 53
4.3.3.2 Regression model specification 54
5 RESULTS 57
5.1 Descriptive results by political connection 57
5.2 Conditional discretion and behavioral risk (H3 & H1) 58
5.3 Regression results: preferential treatment in fines 61
5.4 Regression results: additional enforcement stages (H1-H4) 62
5.5 Role of informal signals: luxury & vanity 66
6 DISCUSSION 68
6.1 Enforcement timeline: where privilege bends the rules 68
6.1.1 Detection 68
6.1.2 Sanctioning 69

6.1.3 Judicial Adjudication 69
6.1.4 Accidents 70
6.1.5 Data quality caveats 71
6.1.6 The layered nature of leniency 71
6.1.7 The role of aerial discretion 72
6.2 Interpretation of control variables 72
6.2.1 Gender and detection-stage discretion 72
6.2.2 Age and experience 74
6.3 Implications of Nepotism/Institutional and sociopolitical implications 75
6.4 Limitations of the study 77
6.5 Direction for future research 79
7 CONCLUSION 81
8 REFERENCE LIST 83
9 APPENDIX 87
9.1 ANNEX A. Supplementary text notes 87
9.2 ANNEX B. Supplementary Tables and Figures 88
9.3 ANNEX C. Model Specifications and Output 93
9.3.1 Descriptive comparisons by group 93
9.3.2 Regression results: political connections 95
9.3.2.1 Fine 95
9.3.2.2 Violation 96
9.3.2.3 Fine given violation 97
9.3.2.4 Accident 98
9.3.2.5 License deprivation 99
9.3.3 Regression results: informal privilege signals 100
9.3.3.1 Fine 100
9.3.3.2 Violation 101

9.3.3.3 Fine given violation 102
9.3.3.4 Accident 103
9.3.3.5 License deprivation 104

LIST OF TABLES

Table 1: Summary statistic for traffic sanction variables 30
Table 2: Variable completeness in the raw pension dataset (1999–2004) 31
Table 3: Summary of dataset reduction across cleaning steps 34
Table 4: Completion and diversity of address fields 34
Table 5: Field completion and uniqueness before and after cleaning 39
Table 6: Birthdate completeness before and after imputation 39
Table 7: Unique values for name and birthdate fields in traffic and pension datasets 41
Table 8: Violation, fine, license deprivation, and accident rates by group 43
Table 9: Group-level summary of key driver and vehicle attributes 44
Table 10: Enforcement outcomes by group 57
Table 11: Ten most represented ministries (by unique employees) in the pension dataset 88
Table 12: Distribution of luxury brands in the dataset 89
Table 13: Distribution of vanity plate numbers 91
Table 14: Summary of enforcement outcomes: ministry vs. unconnected 93
Table 15: Summary of enforcement outcomes: family vs. unconnected 93
Table 16: Summary of enforcement outcomes: luxury vs. non-luxury 94
Table 17: Summary of enforcement outcomes: vanity vs. non-vanity 94
Table 18: Average fines per violation per group 95
Table 19: Logit estimates – Fine receipt by ministry and family 95
Table 20: Logit estimates – Traffic violations by ministry and family 96
Table 21: Logit estimates – Fine receipt conditional on violation (ministry and family) 97
Table 22: Logit estimates – Accident involvement by ministry and family 98
Table 23: Logit estimates – License deprivation by ministry and family 99
Table 24: Logit estimates – Fine receipt (ministry, with luxury and vanity) 100
Table 25: Logit estimates – Fine receipt (family, with luxury and vanity) 100
Table 26: Logit estimates – Traffic violations (ministry, with luxury and vanity) 101
Table 27: Logit estimates – Traffic violations (family, with luxury and vanity) 101
Table 28: Logit estimates – Conditional fine receipt (ministry, with luxury and vanity) 102
Table 29: Logit estimates – Conditional fine receipt (family, with luxury and vanity) 102
Table 30: Logit estimates – Accident involvement (ministry, with luxury and vanity) 103
Table 31: Logit estimates – Accident involvement (family, with luxury and vanity) 103
Table 32: Logit estimates – License deprivation (ministry, with luxury and vanity) 104
Table 33: Logit estimates – License deprivation (family, with luxury and vanity) 104

LIST OF FIGURES

Figure 1: Global Absence of Corruption Index (WJP Rule of Law Index, 2024). 22
Figure 2: Conceptual framework of informal privilege: from blat to nepotism 23
Figure 3: Welfare effects of selective enforcement in traffic law 24
Figure 4: Informal market prices for vanity plates in Russia (Gosnomerus.com, 2024). 25
Figure 5: Legal workaround to transfer vanity plates between vehicles (Gosnomerus.com, 2024) 26
Figure 6: Distribution of police interaction events by year (Traff_OPdate) 30
Figure 7: Distribution of vehicle registration years (Entry_date) 30
Figure 8: Distribution of observations and unique car-person pairs (full vs. restricted sample) 32
Figure 9: Unique license plate distribution 33
Figure 10: Flowchart for address key assignment 36
Figure 11: Overlap between address keys 37
Figure 12: Distribution of individuals per surname core 38
Figure 13: Distribution of car–person combinations by group status 43
Figure 14: Enforcement timeline: from behavior to record 47
Figure 15: Enforcement outcomes by political connection group 58
Figure 16: Conditional fine rates with 0.95 CI 59
Figure 17: Average number of fines per violator by group 60
Figure 18: Estimated effects of political connection on fine receipt across model specifications 62
Figure 19: Gender differences (F-M) in (conditional) fine rates across police districts 73
Figure 20:Behavioral model of risk and enforcement under privilege 79
Figure 21: Comparison of vanity plate definitions 91
Figure 22: Age at entry: Ministry vs Family 92
Figure 23: Age at entry: Unconnected 92

1 INTRODUCTION

Traffic laws are designed to regulate behavior on the road and ensure safety and equal treatment for all drivers. However, in many countries, enforcement is not applied uniformly. Individuals with political, social, or economic privilege often benefit from exemptions, undermining public trust in state institutions. In Russia, traffic enforcement has long faced criticism for favoring the well-connected. This has been highlighted both by grassroots activism and empirical research. Movements like the Blue Buckets Society (Evans, 2018) and studies using the distribution of special license plates as a proxy for corruption (Eeckhout et al., 2021) have drawn attention to these disparities.

My interest in this topic was first sparked by the protests of the Blue Buckets Society, a grassroots campaign that mocked the misuse of emergency lights (*migalki*) by elites and officials. Their symbolic acts of placing blue toy buckets on car roofs, highlighted the gap between formal rules and informal privilege. While the movement brought visibility to the issue, it left open deeper institutional questions: What mechanisms allow some individuals to escape accountability? Are these privileges grounded in wealth, informal influence, or deeper nepotistic ties?

Academic work has begun to explore these dynamics. Eeckhout (2017) showed a strong correlation between vanity plates and luxury vehicles, while Eeckhout et al. (2021) used the prevalence of such plates as a regional proxy for institutional corruption. Their findings indicate that in regions where these plates are more common, drivers receive fewer citations and are involved in a higher number of accidents. This pattern suggests that underlying corruption may lead to weaker enforcement and riskier driving behavior. However, vanity plates function as indirect signals of informal privilege. They hint at unequal treatment but do not directly reveal the political or familial ties that may drive it.

Privileges granted to high-ranking officials, especially in state security or political contexts, may be institutionally justified in some sense. But when these privileges extend to family members or close affiliates without formal roles, they raise concerns about nepotism and systemic abuse. In other words, even if certain benefits for politicians can be rationalized, there is little justification for extending those benefits to their relatives by default.

A tragedy in January 2025 brought a more personal dimension to the issue. An acquaintance of my girlfriend was killed in a high-speed crash in Crete. The driver of the luxury Porsche involved in the incident had reportedly been stopped by police earlier that night for suspected drunk driving but was not detained. The resulting accident killed a 22-year-old man (*ROSA.gr*, 2025). This incident, and the public outrage that followed, underscored a larger concern: when individuals escape accountability, is it due to their wealth, or because of personal connections within law enforcement and government?

This thesis adopts a forensic economic approach to investigate that question in the Russian context: ***Do political connections, beyond wealth or informal privilege signals, drive preferential treatment in traffic law enforcement?***

While previous studies have addressed informal signals of privilege, this study takes a more direct approach. It shifts the focus from symbols to substance by examining whether individuals with confirmed institutional or familial ties to government ministries receive systematically different treatment under traffic law.

To achieve this, the study integrates two key sources of administrative data. The Russian Pension Fund dataset (1999–2004), previously used by Szakonyi (2019) to trace political and familial ties for advantages in the labor market, as well as the Traffic transaction dataset (1993-2006) analyzed by Eeckhout et al. (2021). This linkage enables a novel analysis of enforcement outcomes across detection, citation, sanctioning, and accident records, using car–person combinations as the unit of analysis. Both ministry employees and their household-linked family members are examined.

Despite the fragmented and inconsistent nature of the original data, extensive cleaning and matching enabled a robust analysis. The study directly identifies politically connected individuals and evaluates whether those connections influence legal treatment, something prior work using status proxies like vehicle type or vanity plates could only imply.

In doing so, this thesis sheds new light on how informal privilege operates not just in hiring or procurement decisions, but in daily, consequential interactions between citizens and the state. It shows that legal inequality in Russia extends to the streets, where rules are applied unevenly depending on who you are, or who you are connected to.

2 LITERATURE REVIEW

2.1 Theoretical foundations: corruption, nepotism, and informal influence

"The policeman asked his supervisor for a raise and got the reply, 'I cannot give you a raise, but I can give you a stop sign.'"
— Russian Proverb

This popular joke illustrates the embedded nature of corruption in law enforcement. It highlights the corrupt practice of traffic officers using portable stop signs to fine unsuspecting drivers, thereby supplementing their income through informal bribes. This joke reflects a well-known reality. In many countries, like Pakistan, inadequate salaries combined with unchecked discretionary power can encourage such corrupt practices among officers (Malik & Qureshi, 2020). But corruption in law enforcement goes well beyond petty misconduct. It includes systemic patterns like favoritism, cronyism, and nepotism. All of which distort the principle of equal legal treatment.

Corruption is commonly defined as the misuse of public office for private gain (Treisman, 2000). It includes both explicit acts such as bribery and embezzlement, as well as more diffuse practices like patronage and informal favoritism. Svensson (2005) identifies three broad drivers of corruption: economic, structural, and institutional. Economic drivers include poverty and low wages, which heighten the incentive for public officials to seek illicit income. Structural drivers arise from overregulation and administrative complexity, which create more opportunities for rent-seeking behavior. Institutional drivers include weak judicial independence, authoritarian legacies, and a lack of transparency and accountability.

Within this broader landscape, nepotism and *blat* emerge as two distinct yet overlapping channels of informal influence. Nepotism refers to preferential treatment of relatives, particularly in areas such as hiring, promotions, or legal leniency. *Blat,* a Russian term rooted in Soviet-era practice, is more expansive: it refers to the use of personal networks and reciprocal favors to access goods, services, or outcomes that would otherwise be unavailable through formal procedures.

These informal channels differ in character: each reflects their own distinct mechanism of influence and social obligation. Bribery is transactional: a one-time exchange of money for a specific outcome. Nepotism is relational: it operates through family bonds and obligations rather than monetary payment. *Blat* is networked and normalized: a long-term system of reciprocal exchange embedded in social norms. These dynamics are particularly powerful in contexts where frontline enforcement involves high discretion, such as traffic policing. Officers decide whom to stop, cite, or penalize. In such moments, informal influence can quietly override formal rules.

In practice, these influences rarely operate in isolation. Bribes may coexist with favoritism, and personal networks may shield individuals even in the absence of any financial exchange. Such entanglements raise serious concerns about institutional integrity and legal equality.

The World Justice Project's (WJP) Rule of Law Index (2024) provides a global benchmark that can be used to measure how legal systems perform across four core dimensions: accountability, just laws, open government, and accessible and impartial justice. Russia scores poorly across all four pillars, particularly in the sub-indicator for "absence of corruption in the executive branch", placing it among the lowest-performing countries in the region. This index not only captures more than formal rules, it reflects how people actually experience the law in daily life. As the WJP emphasizes, routine encounters with police, courts, and public officials are where the rule of law is either realized or eroded.

Figure 1 shows Russia's global position on the "absence of corruption" dimension, highlighting the systemic nature of unequal enforcement.[1] This empirical backdrop helps frame the next section, which zooms in on Russia's specific legacy of informal governance, particularly the persistence of *blat* and the moral normalization of nepotism.

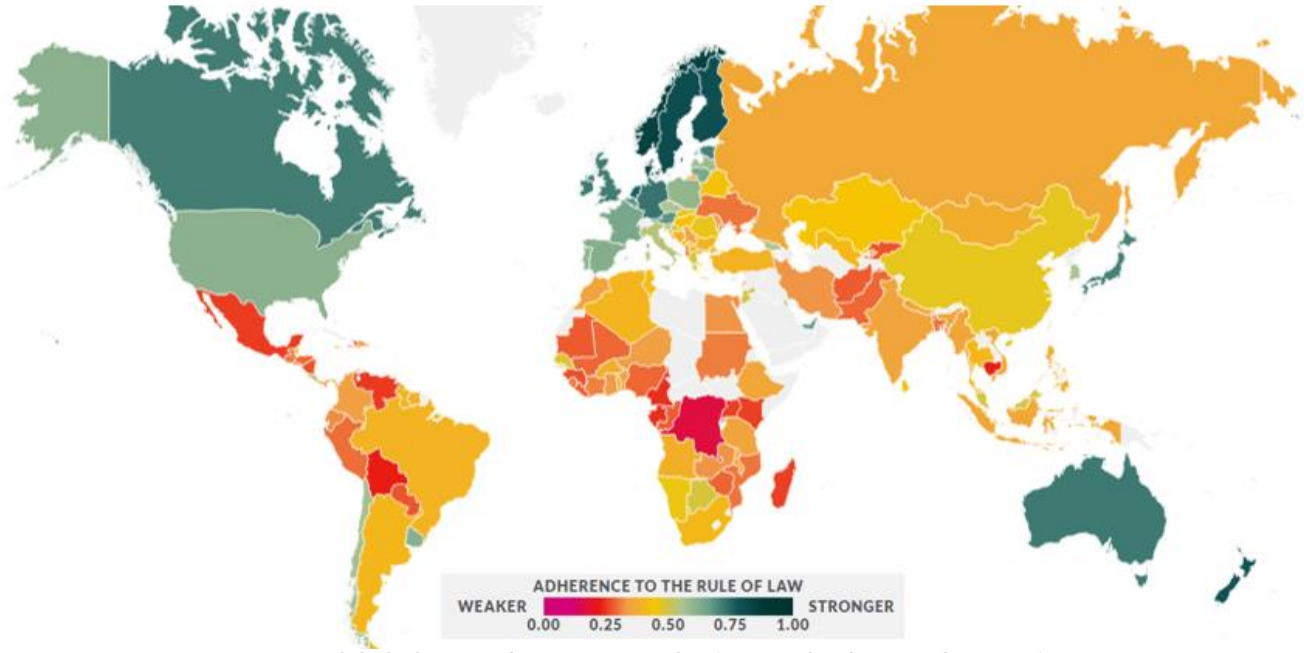


Figure 1: Global Absence of Corruption Index (WJP Rule of Law Index, 2024).

2.2 Why informal enforcement happens: cultural and institutional roots

The period examined in this study (1999-2004) corresponds with the formative years of post-Soviet Russia. Svensson (2005) identifies transition economies, particularly those emerging from socialist regimes, as especially vulnerable to corruption. These countries often inherit weak institutions, centralized political authority, and deeply rooted cultures of informality. Throughout this timeframe, the collapse of formal structures in Russia opened space for informal networks to consolidate influence. In such contexts, formal legal procedures often coexist with personalized enforcement, allowing individual connections to shape who is sanctioned and who is shielded.

To understand how corruption operates in Russian law enforcement, it is essential to examine the legacy of *blat*. Originally a survival strategy under the Soviet command economy, *blat* referred to the use of personal connections to access scarce goods and services outside official distribution channels (Ledeneva, 2009). In a system marked by chronic shortages and bureaucratic rigidity, informal networks emerged as a parallel infrastructure to formal governance.

[1] Russia ranks in the bottom quartile globally, indicating persistent institutional challenges in combating corruption.

Following the Soviet Union's collapse, *blat* did not disappear, it adapted to new political and economic realities. While no longer essential for everyday commodities, it remains instrumental in accessing jobs, licenses, educational placements, and, perhaps most crucially, leniency in legal enforcement (Ledeneva, 2009). Aliyev (2013), studying the South Caucasus, notes that these practices persist in post-Soviet states as mechanisms for accessing both public and private benefits. Though terminology may evolve, the mechanism of reciprocal favors embedded in social networks, remains durable.

Blat has thus evolved from a rationing mechanism to a normalized governance practice. It no longer centers on exchanging food or appliances but on strategic favors involving legal discretion and institutional access. This evolution blurs the boundary between legality and social obligation, making it increasingly difficult to distinguish between corruption and culturally sanctioned behavior.

Figure 2 illustrates how *blat* functions as a foundational mechanism in informal governance. Favoritism refers broadly to any preferential treatment rooted in personal relationships. Cronyism narrows this to favoritism among friends and professional allies. Nepotism, the most specific category, denotes preferential treatment extended to family members.

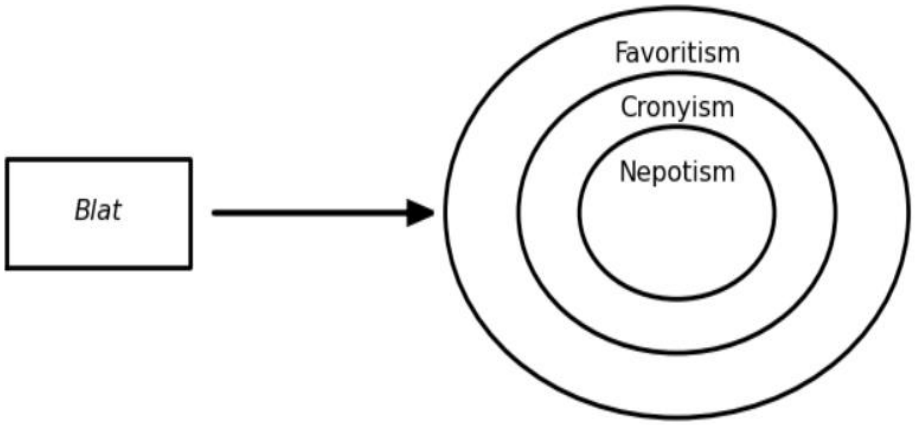


Figure 2: Conceptual framework of informal privilege: from blat to nepotism

In the former Soviet Union, nepotism was often not regarded as unethical or illegal but as a moral duty to one's kin. Kupatadze (2012, as cited in Byulegenova et al., 2024) emphasizes that shielding a family member from legal sanction may not be seen as corruption but as fulfilling an expected obligation. This moral framing helps explain why informal privilege remains entrenched in law enforcement. Officers may feel socially compelled to offer leniency to relatives or the relatives of high-ranking officials. Over time, such behaviors become routinized, no longer perceived as favoritism, but as common sense.

This normalization is particularly consequential in domains with high discretion, such as traffic policing. A police officer might stop and fine a stranger, but refrain from intervening when the offender is the son of a ministry official. The formal structure remains intact, but enforcement is selectively applied and often invisibly.

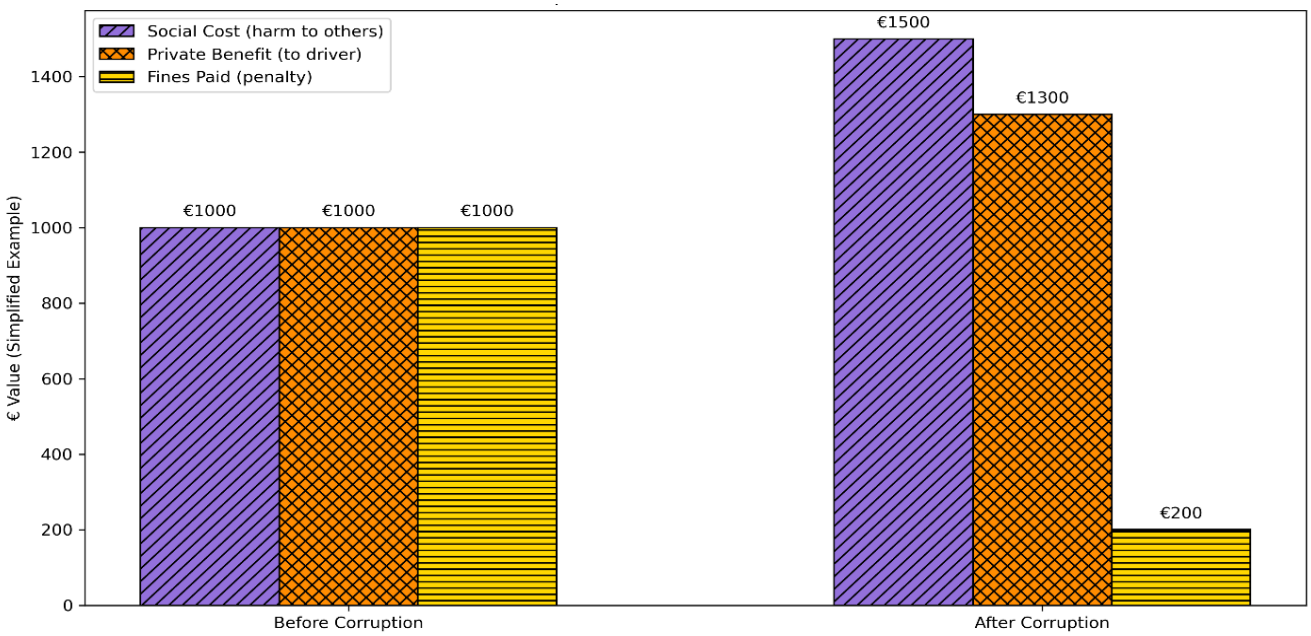
2.3 Why informal enforcement matters: fairness, risk and social cost

Regardless of legality, behaviors like nepotism and favoritism undermine fair competition and economic efficiency. Even in the absence of explicit bribery, the practice of officials acting based on personal connections rather than merit or legality results in unfair enforcement. This is particularly damaging in the public sector, where resources and penalties should be distributed based on law, not loyalty (Khatri, Tsang, & Begley, 2006; Yu et al., 2023).

Although nepotism and cronyism are not always explicitly illegal, they represent forms of institutional inefficiency. As Fisman and Golden (2017, cited in Ragauskas & Valeškaitė, 2020) argue, public trust relies on perceptions of procedural fairness. When the well-connected evade punishment, public confidence in the legal system deteriorates.

The effects of such leniency are not only normative, but also measurable. Traffic laws, for example, are designed to internalize the externalities of risky driving behavior. Speeding increases the likelihood of fatal accidents, and fines are intended to deter such behavior. In economic terms, the optimal fine should equal the expected social cost of the infraction.

Figure 3 presents a simplified model of this mechanism. If, for instance, driving 20 km/h over the speed limit raises the risk of a fatal accident by 0.1%, and the statistical value of a human life is estimated at €1 million, then the optimal fine would be €1,000 (0.1% × €1 million).



Example: Speeding increases chance of fatal accident by 0.1%. If Value of a Statistical Life (VSL) = €1,000,000, the optimal fine = €1,000. After corruption, privileged drivers avoid penalties, take more risks, gain more benefit, and shift cost to society.

Figure 3: Welfare effects of selective enforcement in traffic law

In the absence of corruption, the private benefit a driver gains from speeding (e.g., time saved or convenience) is valued at €1,000. Since the fine for speeding is also €1,000, the driver fully internalizes the external cost of their behavior. From a welfare economics perspective, this alignment means there is no welfare loss: the driver speeds only if their personal benefit matches the societal harm, and the fine ensures their private decision reflects that cost.

In contrast, the scenario that includes corruption illustrates the consequences of uneven enforcement. A driver who expects leniency (due to personal connections or selective enforcement) may take greater risks, such as exceeding the speed limit by more than 20 km/h. Suppose this elevated behavior increases the chance of a fatal accident to 0.15%. Given a statistical value of a life of €1 million, the expected social cost rises to €1,500 (0.15% × €1 million). Yet the driver only pays €200 because of informal protections. The resulting welfare loss is the gap between the social cost and the fine paid:

$$€1{,}500 - €200 = €1{,}300$$

This €1,300 is effectively shifted onto others (pedestrians, other drivers, and society at large) through higher accident risks, stress, and reduced safety. The driver enjoys greater private benefit while externalizing more of the public harm, illustrating how selective enforcement creates both inequality and inefficiency.

This gap constitutes a form of moral hazard[2]: when drivers expect leniency due to personal connections, they are more likely to engage in riskier behavior. The resulting harms are more accidents, more injuries, and higher public costs, which are borne by others. While some early work questioned the empirical link between enforcement and road safety (Schafer & Mastrofski, 2005), other studies find that consistent enforcement reduces accident rates (DeAngelo & Hansen, 2014), and that weakened expectations of punishment can reduce compliance through a moral hazard channel (Levitt & Porter, 2001).

Selective leniency, then, is not a victimless practice. It distorts driver behavior, undermines the effectiveness of enforcement, and erodes the credibility of the legal system. In the chapters that follow, this thesis examines whether such patterns can be empirically observed among politically connected individuals in Russia.

2.4 Vanity plates as a symbol of informal privilege

Vanity plates are among the clearest symbols of informal privilege in Russian traffic culture, but their meaning depends on institutional context. In some countries, vanity plates are a form of conspicuous consumption. In the United Arab Emirates, for example, a plate bearing the number “1” sold for $14.5 million (Emirates Auction, 2021). These plates are purchased for status, not protection. In Russia, however, they convey something different: informal power. Officially, license plates are randomly assigned. Yet certain desirable combinations such as "001," "AMR," "EKX" , are disproportionately found on luxury cars and vehicles linked to state institutions (Eeckhout, 2017).

One notable example is the "EKX" combination, which is widely recognized in Russia as the abbreviation for *Еду Как Хочу* (*Edu Kak Khochu*), meaning "I drive as I want." This unofficial interpretation reflects a perceived sense of impunity among drivers bearing these plates. According to Russia Beyond (2017), "EKX" plates are often associated with the Federal Protective Service (FSO), which provides security for top state officials. Another example is “AMR,” informally linked to the “Administration of Russian Militia”. These combinations are typically reserved for powerful institutions such as the Ministry of Internal Affairs, the Ministry of Defense, or the Federal Security Service (FSB). Although officially non-transferable and randomly assigned, such plates are widely believed to circulate through elite networks by exploiting legal loopholes, including temporary registrations and staged transfers coordinated by intermediaries (see Figure 5). This legal fiction enables vanity plates to trade in a shadow market, often at steep prices (see Figure 4), while preserving a formal appearance of compliance (Manaev, *Russia Beyond*, 2017).

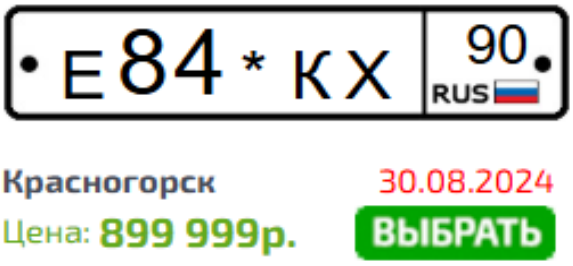


Figure 4: Informal market prices for vanity plates in Russia (Gosnomerus.com, 2024).

[2] The concept of moral hazard originated in 19th-century insurance markets, where it described the tendency of insured individuals to take on greater risks because they were protected from the full cost of those risks. In economics, the concept was later formalized through principal–agent theory, notably by Holmström (1979), who showed how hidden actions and asymmetric information complicate contract design.

Crucially, these plates function as visible signals of informal privilege. Their distinctive combinations and institutional connotations are instantly recognizable, conveying to traffic officers that the driver may be connected to powerful circles, and therefore should not be interfered with. In this sense, vanity plates operate less as direct causes of impunity and more as badges of untouchability within a broader system of selective enforcement.

In regions where these plates are more frequently observed, traffic outcomes differ. Eeckhout et al. (2021) use the prevalence of vanity plates as a proxy for regional corruption, finding that areas with higher concentrations of such plates tend to have fewer citations and more accidents. Their study interprets this as evidence that institutional corruption reduces enforcement quality and increases behavioral risk. In this view, vanity plates are not the cause of these outcomes but a symptom of broader informal systems. They serve as markers of regions where the rule of law is selectively applied.

Still, the plates themselves are highly visible, which makes it easy to misinterpret them as direct drivers of impunity. Why would a traffic officer refrain from stopping a luxury vehicle with a distinctive plate? If the plate merely signaled wealth, it might actually attract bribe-seeking behavior. Yet in regions where such plates are prevalent, these vehicles are consistently stopped less often. This pattern suggests that their presence correlates with a broader institutional climate. One in which enforcement norms are weaker, and corruption is more entrenched. In this context, vanity plates are not powerful because officers consciously recognize and avoid them, but because they circulate in environments where privilege is systematically tolerated. The plate signals informal power, but it cannot tell us whether it is the plate, the car, or the person behind the wheel that truly matters.

This study moves beyond symbolic signals to investigate the actual presence of political connections. Unlike prior work that relies on status markers as proxies for power, the empirical analysis in this thesis directly identifies politically connected individuals and their relatives, and tests whether those connections, rather than visible signals alone, drive disparities in enforcement. In doing so, it shifts the focus from where corruption is likely to happen (regions with many vanity plates) to who benefits from informal protections in practice.

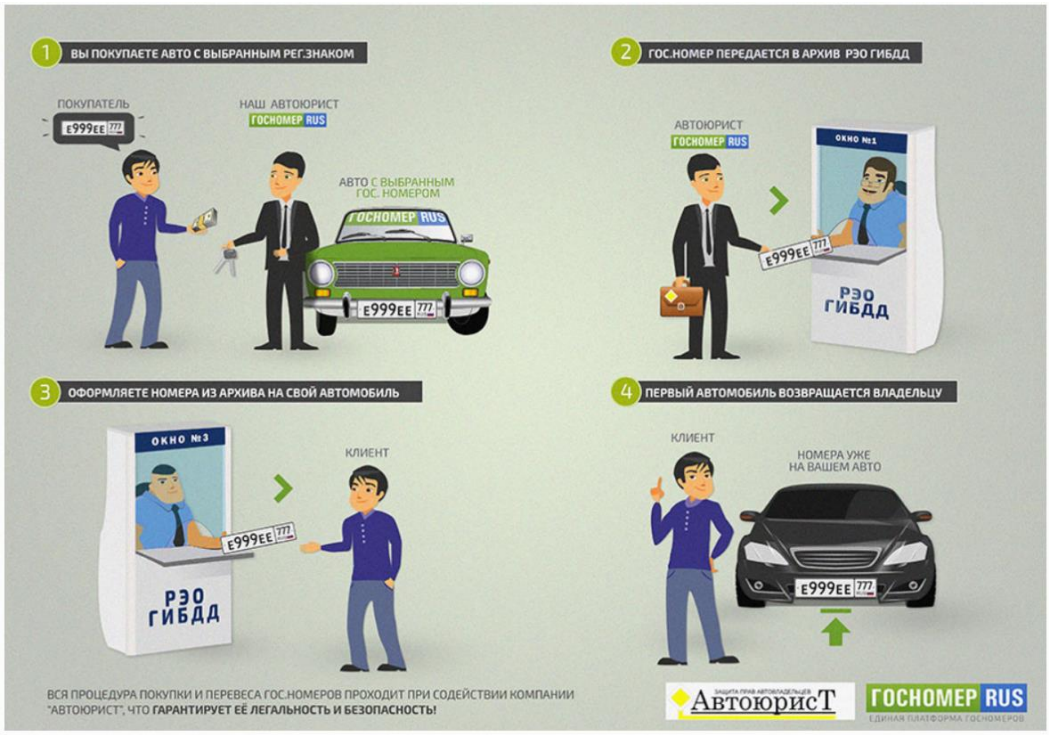


Figure 5: Legal workaround to transfer vanity plates between vehicles (Gosnomerus.com, 2024)

2.5 Measuring nepotism and its effects

Nepotism is notoriously difficult to measure because it operates through informal, undocumented relationships, making the phenomenon hard to quantify at both national and individual levels. At the macro level, studies have attempted to estimate the prevalence of nepotism indirectly. For example, Byulegenova et al. (2024) use socio-demographic data and the Corruption Perceptions Index (CPI) to construct an aggregate indicator of nepotism in post-Soviet countries. Their study suggests that a society's tolerance for corruption, along with traditional kinship structures and a high proportion of older individuals, correlates with the penetration of nepotism across public life. This type of research provides insight into the extent and institutionalization of nepotism within certain regimes, but it does not allow for direct assessment of the behaviors or advantages of specific individuals.

In contrast, micro-level studies have begun to focus on the mechanisms and consequences of nepotism at the level of individual households and actors. One of the most rigorous contributions in this regard is Szakonyi (2019), who uses administrative data from the Russian Pension Fund to link federal officials to their relatives and, in a next step, tracks their labor market outcomes.

His difference-in-differences strategy reveals that relatives of public officials experience increases in employment and wages, as well in the public as in the private sector. These benefits arise not only from public hiring but also from firms strategically employing relatives of officials to gain access to procurement opportunities and favorable regulation. This is an informal corporate strategy known as the "princeling effect". To measure nepotism at this level, one must identify networks of relatives, which is a complex task as it involves accessing sensitive and confidential information (Ragauskas & Valeškaitė, 2020).

While existing literature has extensively examined nepotism's impact on employment outcomes and the accumulation of elite privileges, its implications for law enforcement remain underexplored. Investigative work by Navalny's Anti-Corruption Foundation has exposed high-profile examples of impunity and unjust enrichment among Russia's political elite, particularly in the realm of public procurement and state contracting. Although these findings have had tangible financial repercussions (Healy et al., 2012; Enikolopov et al., 2018), they remain largely anecdotal and focused on top-level corruption. The question thus remains: does nepotism also distort legal accountability in everyday life, such as traffic enforcement?

This thesis seeks to fill that gap by shifting the focus from labor markets to law enforcement. Building on Szakonyi's approach, it uses the same pension data to identify political and familial connections which are then merged with the traffic enforcement dataset from Eeckhout et al. (2021). This enables to empirically test whether connected individuals, identified via household links and employment in ministries, receive preferential treatment in traffic enforcement. Unlike prior studies that relied on symbolic proxies (e.g., luxury cars or vanity plates), this study uses direct relational data to examine enforcement outcomes across multiple stages: detection, citation, sanctioning, and accident involvement.

While a difference-in-differences design might be feasible in principle, it was not pursued here due to the fragmented and sparse nature of the available data. The administrative sources used in this study, particularly the traffic and pension registries, pose significant challenges for constructing reliable longitudinal comparisons or household-level panels. These challenges are especially pronounced when compared to administrative sources like the Larix-10 database used by Szakonyi (2019), which contains more complete and standardized address information, allowing for more reliable household-level linkage. Nevertheless, the empirical logic remains rooted in treatment comparison: do politically connected individuals face systematically different legal consequences, even after accounting for observable characteristics?

By shifting the focus from signals to structures, from symptoms to mechanisms, this study aims to clarify whether nepotism in Russia functions as a shield not only in career advancement but also in avoiding legal sanction. In doing so, it contributes to the broader literature by expanding the scope of nepotism research into the realm of rule-of-law enforcement, offering new insights into how informal privilege interacts with legal institutions in the daily experience of citizens.

From proxy to mechanism

Prior research has highlighted visible status symbols such as vanity plates as proxies for informal privilege, suggesting but not directly demonstrating selective enforcement. This study advances the literature by identifying actual political and familial ties through administrative records. By linking these connections to enforcement outcomes, it moves beyond symbolic inference to test the mechanism more directly: whether informal power, embedded in institutional relationships, systematically distorts the application of the law.

Research question: Do political connections, beyond wealth or informal privilege signals, drive preferential treatment in traffic law enforcement?

3 DATA

This chapter describes the administrative data used in this study.[3] The analysis draws on two large-scale datasets that were originally not intended for research purposes. One contains information on vehicle ownership and traffic-related events, while the other includes employment and income records from the Russian Pension Fund. These datasets are merged at the car–person level, allowing for the identification of individuals with ministry affiliations and the assessment of their traffic enforcement outcomes. The analysis focuses on the period from 1999 to 2004, during which both sources provide overlapping coverage.

3.1 Data Sources

3.1.1 Traffic Transaction Data (1993–2006)

The traffic related variables are derived from an administrative transaction database of the Main Directorate for Road Safety of the Ministry of Internal Affairs of Russia (abbreviated GIBDD). This dataset was previously processed and analyzed by Eeckhout et al. (2021), and forms the basis for the current study. It includes 20,397,711 observations, each with 138 variables covering license plate details, vehicle characteristics, transaction types, registration events, and limited personal information. As is typical with large-scale administrative records, the raw data contain numerous inconsistencies, such as missing values and misspelled entries. Substantial effort was therefore required to clean and standardize the dataset to ensure it was suitable for subsequent analysis.

Key date variables in the traffic dataset include *Entry_date* and *Traff_OPdate*. While *Entry_date* records the date a vehicle was first officially registered, *Traff_OPdate* marks the date of an interaction with the traffic police. Such interactions include four main categories: violations, fines, license deprivations, and accidents.

Figure 6 displays the number of observations involving a recorded police intervention (i.e., where *Traff_OPdate* is non-missing). The vertical dashed lines at 1999 and 2004 mark the period selected for analysis, corresponding to the years for which employment and income data are available. Figure 7 shows the distribution of vehicle registration years. The left panel covers the full registration window (1993-2006), while the right panel focuses on the period from 1995 onward. A notable spike in 1993 likely reflects vehicles already in circulation before digitization, which were assigned a default registration year. The vertical dashed line at 2004 indicates the cutoff used in the analysis, excluding vehicles first registered thereafter.

[3] See Appendix A, Note 1 for acknowledgment of software tools used.

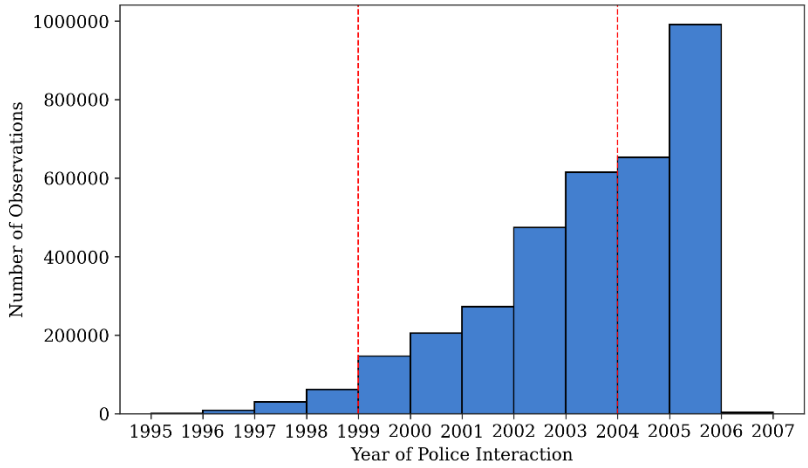


Figure 6: Distribution of police interaction events by year (Traff_OPdate)

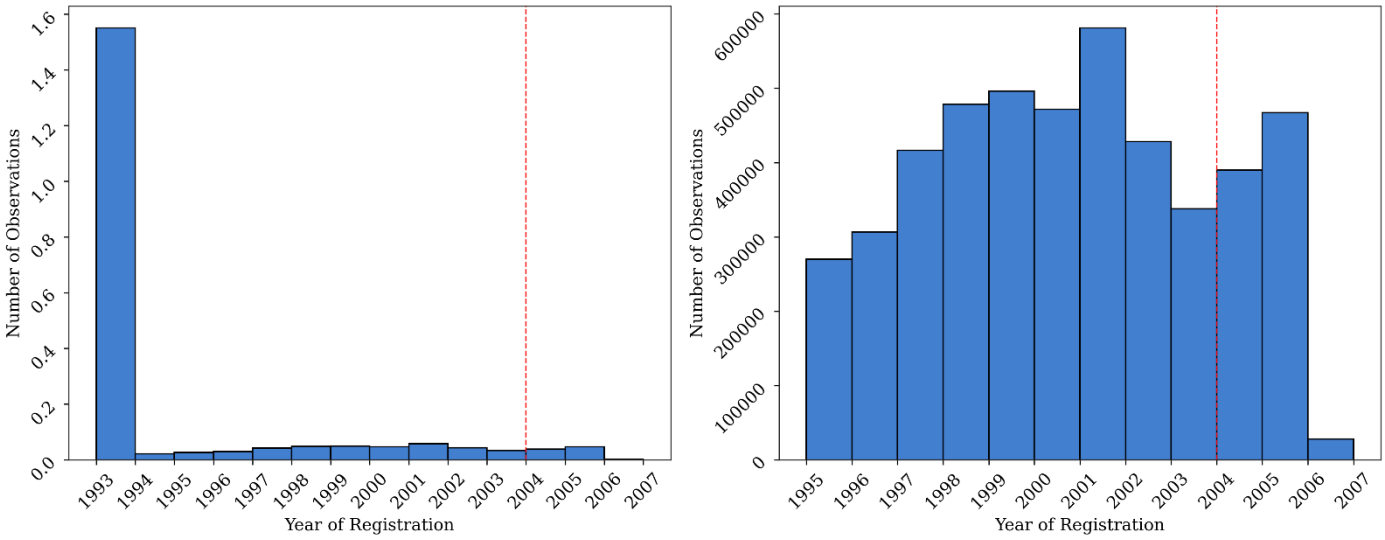


Figure 7: Distribution of vehicle registration years (Entry_date)

To simplify identification of traffic-related events, a binary indicator variable, *Traffic_data*, is used. This variable takes the value 1 when *Traff_OPdate* is non-missing, signifying that the observation involves one of the four main types of traffic incidents. Table 1 summarizes the raw frequency and relative share of each incident type in the dataset.

Variable	# of Events	% of Total	# Non-Events	% Non-Events
Violated	2,850,847	13.98	17,546,864	86.02
Fined	1,928,496	9.45	18,469,215	90.55
LicDeprived	66,467	0.33	20,331,244	99.67
Accid	615,932	3.02	19,781,779	96.98

Table 1: Summary statistic for traffic sanction variables

In addition to traffic-specific fields, the dataset contains personal identifiers such as full name, date of birth and registered address information. Full names consist of *Familya* (surname), *Imya* (first name), and *Otchestvo* (patronymic). These fields are essential for matching records and identifying kinship ties. To support the nepotism analysis, addresses are standardized into structured keys, which are then combined with lemmatized surname stems to construct household-level family identifiers.

Russian surnames often vary by gender or marital status (e.g., Ivanov vs. Ivanova); therefore reducing them to a shared core allows for more accurate linking. These steps are detailed further in Section 3.2.

3.1.2 Pension and Employment Data (1999–2004)

The second dataset used in this study originates from the Russian State Pension Fund and contains administrative records of reported earnings. It spans the years 1999 to 2004 and includes 53,578,600 observations. Income information is submitted to the fund by employers on an annual basis. Each entry contains a range of personal and institutional identifiers: full name (first, middle, and last), birthdate, employer name (*employer*), tax identification number (*employer_inn*), individual tax identifiers (*fl_inn, fl_gni, home_gni*), passport numbers, and geographic information such as *zipcode*.

Table 2 summarizes the structure and completeness of key fields in the raw pension data. Core fields such as *firstname*, *lastname*, *fullname*, *employer* and *employer_inn* are nearly complete. Others, such as *birthdate* (78%), *document_num* (75%), and *zipcode* (69%, show moderate levels of missingness. Personal tax identifiers (*fl_inn*) are available only for a subset of the sample (37%), reflecting underlying administrative inconsistencies. This variation in data completeness motivated a variety of preprocessing steps, including the imputation of missing birthdates (see Section 3.2).

Variable	Filled (count)	Filled (%)
employer	53,578,511	100.00%
employer_inn	53,578,600	100.00%
fl_inn	19,877,755	37.10%
fl_gni	22,144,110	41.33%
home_gni	35,610,510	66.46%
firstname	53,578,600	100.00%
middlename	53,278,136	99.44%
fullname	53,578,600	100.00%
birthdate	41,764,252	77.95%
document_series	34,621,048	64.62%
document_num	40,270,724	75.16%
zipcode	37,141,398	69.32%
income_year	50,866,324	94.94%
year	53,578,600	100.00%
lastname	44,970,081	83.93%
taxable_year	42,259,717	78.87%

Table 2: Variable completeness in the raw pension dataset (1999–2004)

Merging the traffic and pension datasets entails typical challenges of administrative records digitized from paper sources. As noted by Szakonyi (2019), entries often contain typographical errors, inconsistent formatting, or abbreviated names. Individuals are matched deterministically across datasets using full name and date of birth. While this approach does not guarantee perfect identification, the exact-match procedure yields a sufficiently low false-positive rate for the purposes of this study (see Section 3.2.6.2).

Political connections are identified by flagging individuals in the pension dataset who were employed by a federal ministry or presidential institution at any point from 1999 to 2004. This flag is subsequently merged onto the traffic data. Household-level relationships are inferred using detailed address components from the traffic dataset, rather than ZIP codes or external registry data. As a result, family structures are identified less comprehensively than in Szakonyi (2019). He did use residence registry data, which, being the core source in that context, was likely more complete and reliable.

3.2 Data cleaning and preprocessing

As these datasets are prone to errors and noisy variables, substantial effort was invested in cleaning, restructuring, and exploring the data in order to build a robust analytical playground.

3.2.1 Period restriction and identification logic

Since the pension dataset covers only the years 1999 to 2004, the traffic dataset is trimmed accordingly: all traffic interactions occurring outside this period, as well as vehicles first appearing after 2004 are excluded. This restriction ensures that the merged dataset reflects individuals who are simultaneously observable in both sources during the same timeframe.

Enforcement decisions in traffic policing are typically directed at vehicles, not individuals: license plates receive fines, not people. However, the identity of the vehicle's operator or registered owner, particularly if they are politically connected, can influence whether a vehicle is sanctioned at all. To capture this dynamic, this study constructs a combined unit of analysis: the car–person combination. It links individuals (by name and birthdate) to vehicles (by license plate). This enables the assignment of political connection status to specific vehicles and allows retrieval of vehicle attributes like brand.

This approach allows us to answer questions that neither vehicle- nor person-level data could address alone. For instance, we can test whether vehicles associated with connected individuals are fined less often, or whether individuals avoid sanctions across any of the vehicles they use. It also ensures that enforcement outcomes are consistently matched to the correct person–vehicle pairing. This is especially relevant in cases where one individual owns or operates multiple cars, or where cars are shared among family members.

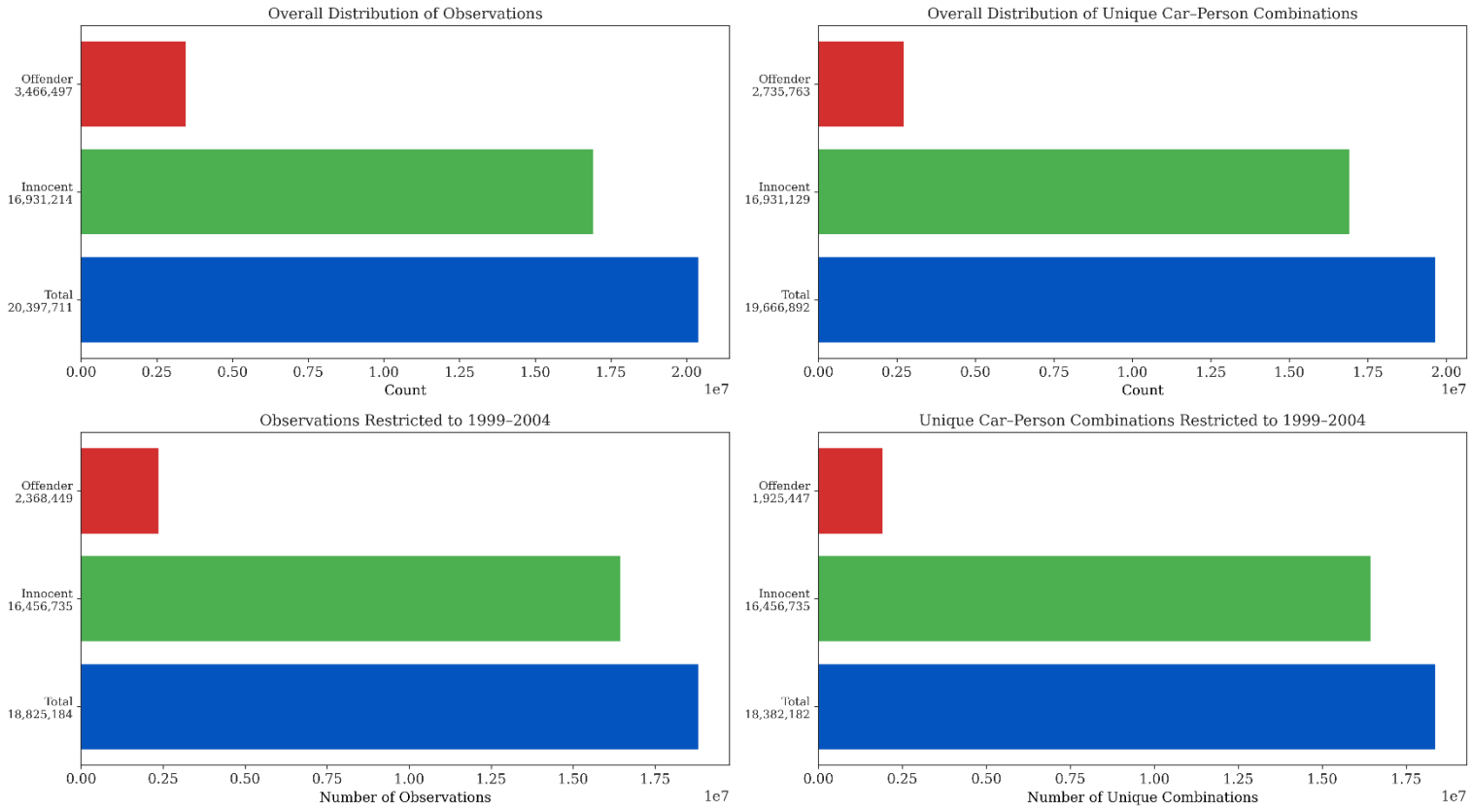


Figure 8: Distribution of observations and unique car-person pairs (full vs. restricted sample)

Figure 8 illustrates how the dataset breaks down across two dimensions: (1) whether an individual had a police-recorded traffic incident (Offender vs. Innocent), and (2) whether the data refers to an observation or a unique car–person combination. The four panels show that the distribution remains broadly consistent, regardless of the time restriction or the aggregation to car–person units. A single car–person combination may be involved in multiple traffic incidents, which accounts for the slight difference in magnitude between offender counts when measured by unique car-vehicle combinations (18,382,182) versus total observations (18,825,184).

Importantly, the Offender and Innocent categories are not mutually exclusive with political connectedness. This separation between connection status and recorded enforcement outcomes enables the analysis to examine whether politically connected individuals are treated more leniently by traffic authorities. As discussed in Section 2.3, unequal enforcement may not only distort sanctioning, but also encourage riskier conduct, leading to measurable externalities such as increased accident rates. Since the data primarily reflect outcomes rather than direct behavior, accident incidence serves as a valuable proxy for underlying risk, allowing limited inference about behavioral differences across groups.

3.2.2 Filtering non-individuals

During the initial data exploration, it became clear that not all entries in the traffic dataset correspond to private individuals. Many records included organizations or company names in place of personal identifiers. These typically contained patterns such as: ОАО, АООТ, ЗАО, ООО, ГУП, and АОЗТ. These stand for *Open Joint Stock Company*, *Joint-Stock Company of Labor*, *Closed Joint Stock Company*, *Limited Liability Company*, *State Unitary Enterprise*, and *Joint Stock Company with Foreign Participation*, respectively.

To systematically exclude these entities, I relied on the *Gender* variable. This field is almost always filled in for real individuals but is missing for organizational entries. Using this heuristic, I excluded 4,617,028 observations with missing gender values, as well as an additional 6,363 records where the gender was filled but the name still matched a corporate pattern. These steps reduced the sample to 14,201,823 observations, moving the analysis toward a more realistic foundation for studying individual-level traffic behavior.

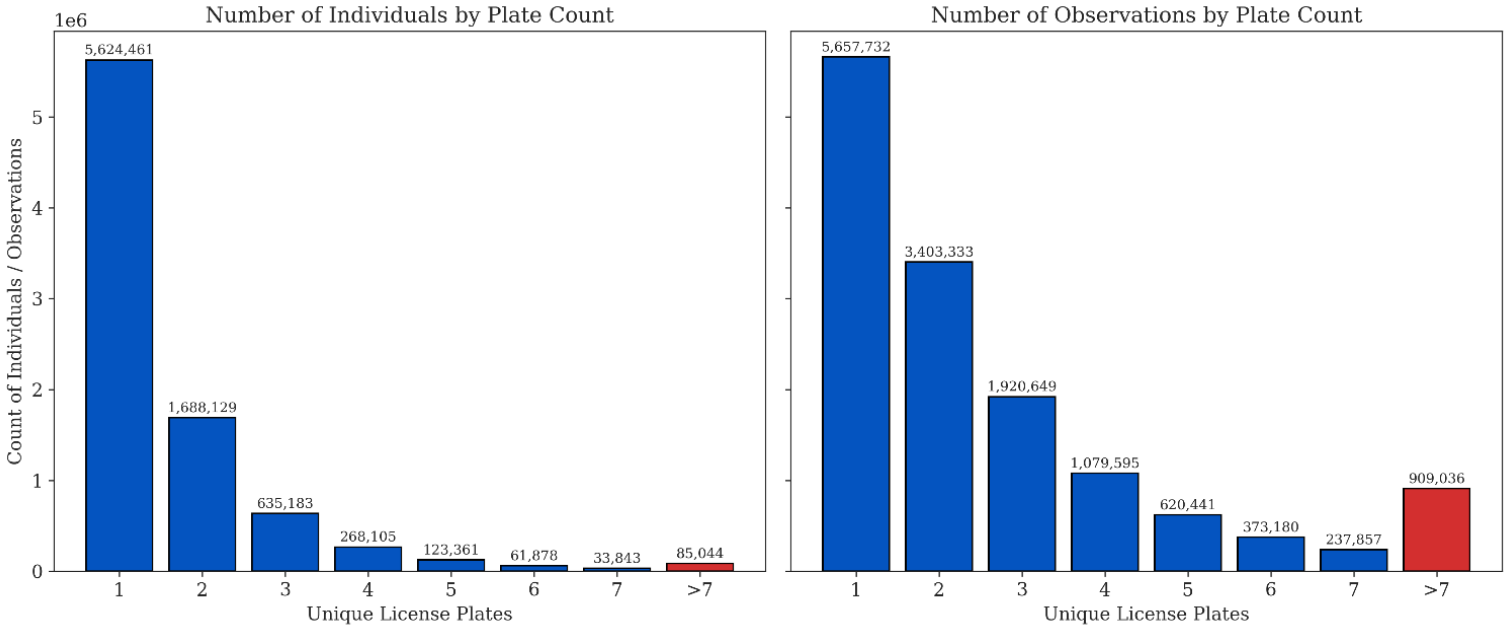


Figure 9: Unique license plate distribution

Another irregularity that emerged early on during the cleaning process involved filtering individuals with an implausibly high number of associated vehicles. Some records appeared with over 1,000 unique license plates. Anomalies of this nature likely reflect company fleets, car dealerships or erroneous entries. Figure 9 presents a two-panel visualization: the left panel displays the number of individuals by their count of distinct license plates, while the right panel shows the corresponding number of observations. Over half of all individuals (5,657,732 or 66.5%) are linked to exactly one plate, around 20% have two, and about 7.5% have three. While the majority of individuals cluster at the lower end of the spectrum, the distribution has a long tail, with a small number of entities associated with dozens or even hundreds of plates.

To reduce the influence of such outliers, a 99th-percentile cutoff was applied, excluding individuals with more than 7 unique license plates. In Figure 9, this threshold is visually highlighted by the red bar, which marks the group of individuals filtered out due to excessive vehicle ownership. This cutoff filtered out approximately 85,044 individuals (1% of the sample), removing a total of 909,036 observations. This is around 6.4% of the dataset following the gender-based filter. The resulting dataset retained 13,292,787 observations and 8,419,374 individuals. This step preserves realistic variation in car ownership while minimizing distortive effects from fleet records or clerical errors.

Post-filter diagnostics indicate that the dataset still captures meaningful behavioral diversity: individuals now appear with up to 7 distinct vehicles, and car-person combinations register up to 41 police interaction records. Table 3 summarizes each data cleaning step and its impact on sample size.

Step	Observations	Individuals	Car–Person Combos	Obs Removed
Raw dataset	20,397,711 (100.0%)	10,924,426 (100.0%)	19,666,892 (100.0%)	–
Period restriction	18,825,184 (92.3%)	10,322,973 (94.5%)	18,382,100 (93.5%)	1,572,527
Gender/Corporate filter	14,201,823 (69.6%)	8,504,418 (77.8%)	14,111,100 (71.8%)	4,623,361
99th percentile LP_c filter	13,292,787 (65.2%)	8,434,960 (77.2%)	13,203,662 (67.1%)	909,036

Table 3: Summary of dataset reduction across cleaning steps

3.2.3 Address standardization and key construction

To prepare for address-based household linking, I followed the methodology outlined in Szakonyi (2019), standardizing registration addresses into a structured Russian postal format. This format consists of four hierarchical elements: street name, *dom* (house number), *korpus* (building or block), and *kvartira* (apartment number).

3.2.3.1 *Street name standardization*

Table 4 reports the fill rates of these address components. Coverage is highest for street names (filled in 94.9% of records), followed by house numbers (*dom*, 81.1%) and apartment numbers (*kvartira*, 72.9%). The *korpus* field, used primarily in larger residential complexes, is the sparsest, with only 22.3% of entries filled.

Field	Filled (n, %)	Missing (n, %)	Unique Values
Street	12,608,845 (94.9%)	683,942 (5.1%)	1,280,473
Dom (House)	10,782,221 (81.1%)	2,510,566 (18.9%)	20,994
Korpus (Building)	2,966,041 (22.3%)	10,326,746 (77.7%)	1,640
Kvartira (Apt)	9,690,908 (72.9%)	3,601,879 (27.1%)	7,894

Table 4: Completion and diversity of address fields

While the street field is present in most entries, it contains over 1.28 million unique values across 12.6 million filled records. This amounts to approximately one unique street name per ten individuals, highlighting widespread inconsistencies in address recording, likely stemming from spelling variations, abbreviations, punctuation, and other formatting issues. Such a low number of individuals per street would be expected if the dataset covered a wide and evenly distributed geographic area.

However, the *AdressReg* variable indicates otherwise. While 4.7 million entries lack a regional label, 20 out of the 51 distinct regions account for over 99.99% of all labeled observations. The dataset is highly concentrated in a few areas, most notably Moscow Oblast (2.3M), Samara (1.1M), and St. Petersburg (944K). This early exploration highlights the messy and fragmented nature of the address data and underscores the importance of thorough standardization in subsequent cleaning steps.

As a first step, basic cleaning was performed. Whitespace was trimmed, and all letters were converted to uppercase. Leading prefixes such as "УЛ." or "УЛИЦА" (*ulitsa*, meaning "street") were stripped. Non-alphanumeric characters, such as dots, commas, and hyphens, were removed. Sequences of spaces or hyphens were collapsed into a single underscore. Common street-type abbreviations, such as "ПРОСП" (*prospekt*, avenue), "БУЛ" (*bulvar*, boulevard), and "ПР" (*proyezd*, passage), were also eliminated. After this initial cleaning, the number of unique street names was reduced from 1,280,473 to 1,221,610.

Although preliminary cleaning resolved some inconsistencies, spelling variations still persisted across similar street names. To further consolidate the street field, I implemented a region-aware fuzzy matching procedure. Within each region, the 20 most common cleaned street names were identified. For observations with an "UNKNOWN" regional label, a global list of frequent street names was used instead. Fuzzy matching was then conducted based on string similarity (threshold ≥ 90%), while also ensuring consistency in any numerical components (e.g., distinguishing "7-Я" from "3-Я").

This final fuzzy standardization step further reduced the number of unique street names to 129,038, resulting in a nearly 90% reduction relative to the initial number of unique entries.

During address standardization, street names were cleaned into a simplified and consistent format. While this process effectively reduces minor formatting differences, it also inevitably collapses street types (e.g., *prospekt, bulvar, proyezd*) into a single stem when their base names are identical (e.g., "ЛЕНИНА"). This could theoretically introduce false matches. However, because household matching relies not only on the street name but also on strict matching of house numbers, apartment numbers, regional identifiers (*AdressReg*), and even surname cores for family construction, the risk of mislinking distinct households is minimized. The goal is not to perfectly reconstruct the original address, but rather to achieve reliable consistency for household and family linkage.

3.2.3.2 *Address validation and filtering*

To define a valid address, each record was required to contain at least a non-empty street name and house number (*dom*). While Szakonyi (2019) emphasizes the importance of including apartment numbers (*kvartira*) where available, he also acknowledges that in cases of incomplete addresses, the combination of street and *dom* alone can allow useful linking. Based on this, I opted for the relaxed street-and-*dom* criterion as my minimum filter for a valid address. This more inclusive rule filtered the working sample from 13,292,787 to 10,405,710 address records, corresponding to 6,503,582 individuals included in key construction.

3.2.3.3 *Key construction and matching*

With standardized and validated address data, I constructed three types of address keys to support household-level linking: a full key (including region), a relaxed key (excluding region), and a final key that combines both. It uses fuzzy matches in cases where the region is missing. This process yielded approximately 5.6 million unique *full_address_key*, 5.2 million *relaxed_address_key*, and 5.4 million *final_address_key*. The most common keys corresponded to large residential complexes in cities such as Irkutsk, Moscow, and Saratov. At this stage, the dataset includes 6.5 million individuals, implying an average of just 1.20 individuals per final address key. While this may seem low relative to the average household size of 3.34 reported for Moscow in the 2002 census (Shcherbov et al., 2023), the discrepancy is expected. Unlike census or population registry data, this dataset captures only individuals who were vehicle owners or involved in police-reported traffic activity between 1999 and 2004. Many non-driving household members are thus missing by design. In this context, a ratio of 1.20 reflects the sparse and selective nature of administrative traffic records and remains consistent with the data's partial coverage of households.

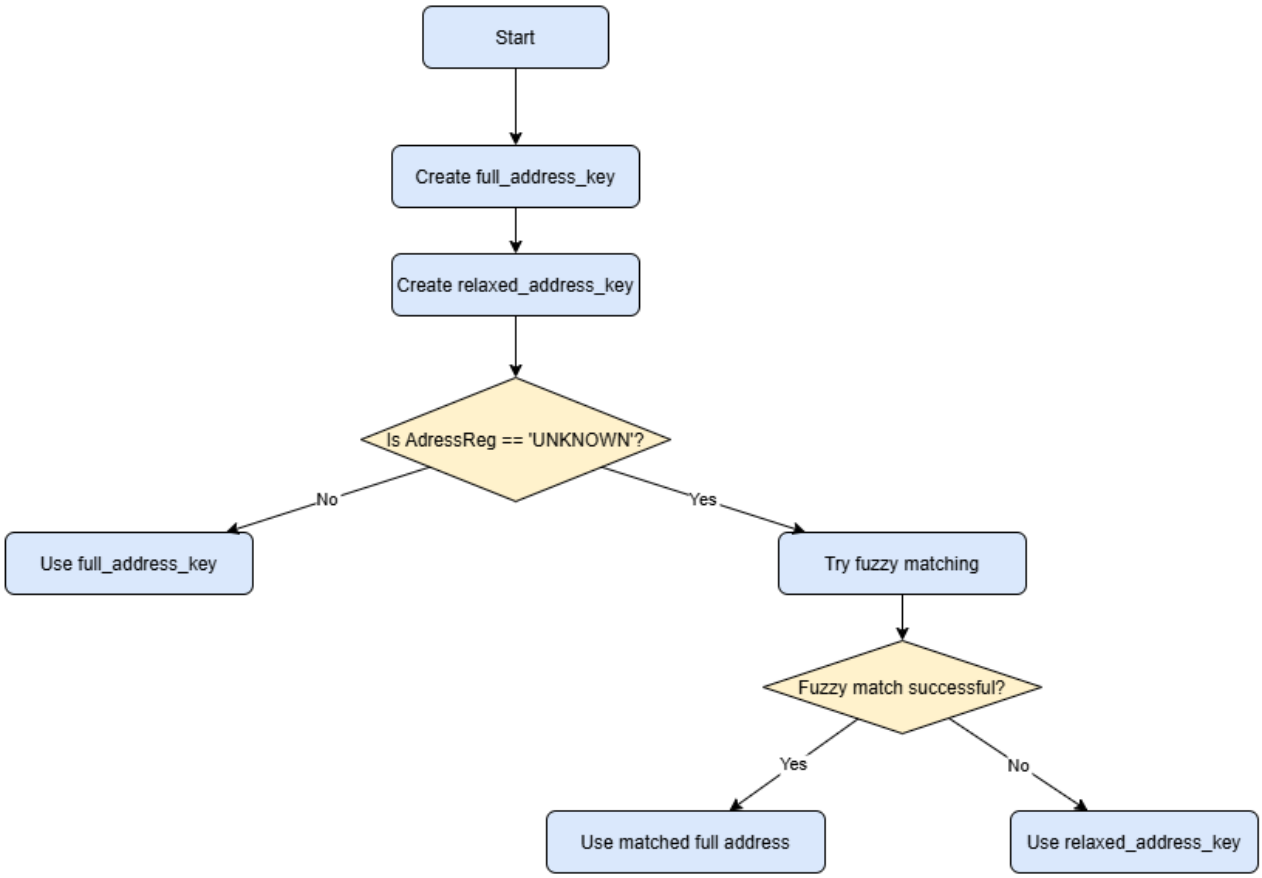


Figure 10: Flowchart for address key assignment

The logic behind the final key assignment is summarized in Figure 10. After applying this method, approximately 58.63% of full address keys appeared unchanged in the final set, while around 41% required adjustment either through fuzzy matching or fallback to the relaxed key. The resulting overlaps between the different key sets are visualized in Figure 11, demonstrating the final composition of address keys after the matching process. For all subsequent analysis steps, the final address key is used to identify household membership and potential family links.

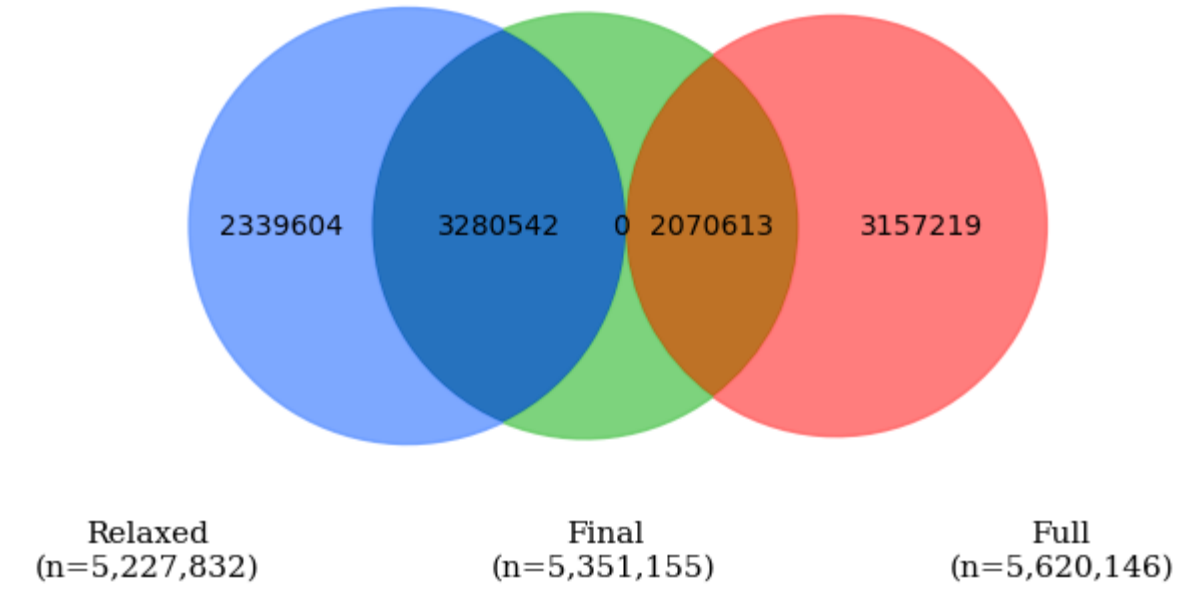


Figure 11: Overlap between address keys

3.2.4 Family and household construction

With standardized address information in place, I proceeded to build proxies for household and family structures. Family-level identifiers were constructed by combining each individual's surname core with their assigned address key. This approach enables the identification of likely family units within a shared residential setting.

Surnames were first cleaned and lemmatized to remove gendered suffixes (e.g., -ова, -ева, -ина, -ская), following Russian linguistic norms. This process created a neutral surname core (e.g., "ИВАНОВ" derived from both "ИВАНОВ" and "ИВАНОВА"), improving the robustness of kinship identification by capturing relatives with gender-differentiated surnames.

I created family identifiers for all three address key types, resulting in roughly 6.8 million full, 6.7 million relaxed, and 6.7 million final family IDs. Around 75% of individuals had a unique family ID (i.e., no matching surname at their address). Again, this outcome is to be expected, given that traffic data does not exhaustively cover the population. The remaining 25% were grouped into multi-person families.

Although Szakonyi (2019) removes individuals with unique surnames, I retain them in my analyses, as I am also interested in how ministry employment may affect individuals without family ties. Rather than excluding these cases, I flag them accordingly.

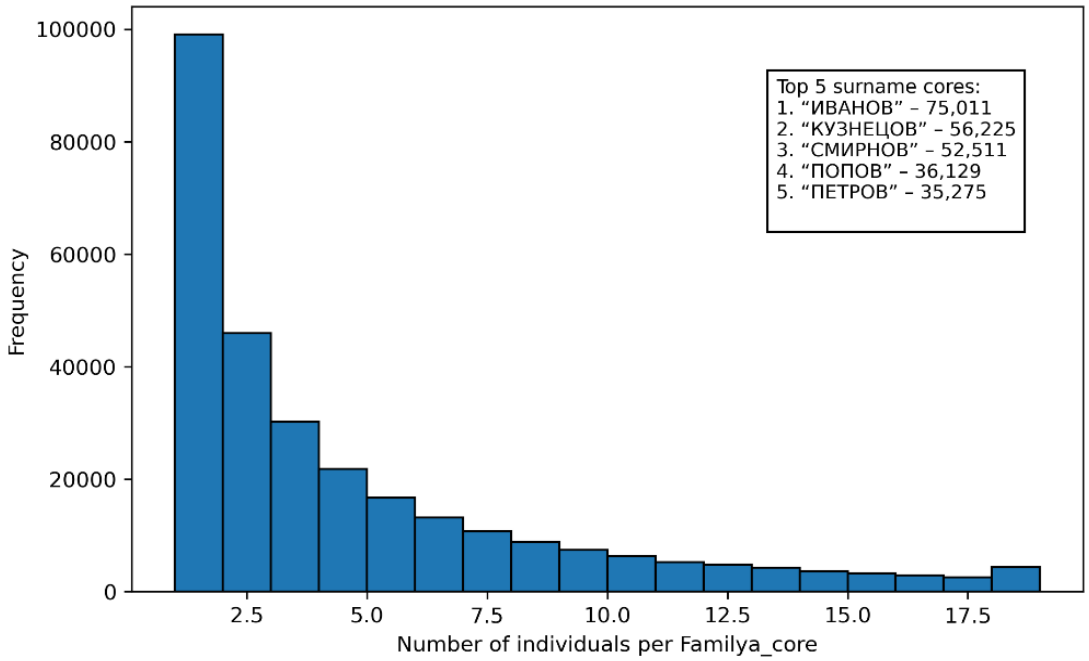


Figure 12: Distribution of individuals per surname core

Figure 12 displays the distribution of individuals per surname core across the dataset. The distribution is highly skewed: a small number of surname cores (such as Иванов, Кузнецов, and Смирнов) account for disproportionately large shares of the population, while most cores are relatively rare. This skewness reflects broader naming patterns in Russia and highlights one of the limitations of surname-based linking: common names increase the risk of false positives when inferring kinship. However, since the method combines surname cores with detailed address keys, the likelihood of incorrect grouping is substantially reduced in practice. A more detailed analysis of false positive risk in family linkage follow in Section 3.2.6.2. This combination offers a pragmatic and transparent strategy for identifying family structures within the constraints of partial administrative data.

3.2.5 Cleaning pension and employment data

The pension dataset is used to identify individuals employed by federal ministries or presidential institutions. These individuals form the core group from which potential nepotism links are constructed in subsequent steps. To prepare the dataset for merging, I cleaned and standardized key personal identifiers (e.g., full names and dates of birth), as well as institutional names and employment records. Given that these records were digitized from legacy administrative systems, the data often contained typographical inconsistencies, incomplete entries, or variations in ministry naming conventions.

3.2.5.1 *Standardizing name fields*

To harmonize the pension data with the traffic records, I first renamed the name-related fields to align with the traffic dataset (*Familya*, *Imya*, and *Otshestvo*). Initial exploration revealed significant inconsistencies and missing values. For instance, the *Familya* field had over 8.6 million missing records out of 53 million total.

Fortunately, these missing entries could be recovered using the *fullname* field, which stored the complete name as a single string. I split this string into its components and filled in missing values where possible. A standard cleaning routine was then applied to all name fields: it removed non-standard characters, standardized text to uppercase, and stripped redundant whitespace.

Table 5 summarizes the number of filled and unique values before and after cleaning. Note that *Otshestvo* values could not further be filled based on the *fullname* field.

Field	Filled (Before)	Unique (Before)	Filled (After)	Unique (After)
Familya	44,970,081	975,492	53,578,600	629,245
Imya	53,578,600	144,582	53,578,600	102,688
Otshestvo	53,278,136	242,826	53,578,136	172,547

Table 5: Field completion and uniqueness before and after cleaning

This cleaning step substantially reduced spelling and formatting variation in the name fields, improving the consistency of identifiers and the reliability of exact matching in the merge process.

3.2.5.2 *Handling missing birthdates*

The *DoB_c* (cleaned date of birth) field was initially missing in over 11.8 million records, representing roughly 22.05% of the dataset. While dropping these records might appear like a straightforward solution, doing so could pose significant limitations. Since the dataset spans five years (1999–2004), many individuals appear in multiple rows across different years. An individual may lack a birthdate in one year but have it recorded in another. More importantly, a person might only be recorded as working for a federal ministry in a year where their birthdate is missing. Excluding these rows would prevent correctly flagging such individuals as politically connected.

To mitigate this, I applied a targeted imputation strategy. Using a candidate-based approach, I constructed a reference index from rows with non-missing birthdates, keyed by (*Familya*, *Imya*, *Otshestvo*) and cross-referenced with auxiliary identifiers such as passport infomation (*document_series*, *document_num*) and tax ID (*fl_inn*). This allowed imputation of missing *DoB_c* values wherever matches were found on any of these fields.

The procedure reduced the number of missing birthdates from 11.8 million to 6.8 million, which is a 42.5% decrease. Table 6 summarizes the change in completeness before and after imputation.

Metric	Before Imputation	After Imputation
Missing birthdates (n)	11,814,348	6,788,675
Missing birthdates (%)	22.05%	12.67%
Count of filled birthdates	41,764,252	46,789,925

Table 6: Birthdate completeness before and after imputation

3.2.5.3 *Ministry employee identification*

Finally, I flagged individuals employed by federal ministries or the presidential administration using a conservative string-matching procedure based on employer names. Specifically, only entries containing standalone instances of "МИНИСТЕРСТВО" (ministry) or "ПРЕЗИДЕНТСКАЯ" (presidential), using a word-boundary regular expression to avoid partial or misleading matches. This approach avoided partial or ambiguous matches and ensured a focus on formal federal-level institutions.

Initially, I explored broader keyword searches and the use of tax identification numbers (INNs) of known ministries as alternative identification strategies. However, these methods often captured entities that were only loosely affiliated with ministries. These included technical schools, regional branches, or research institutes, which do not constitute formal ministry employment. For example, organizations like "Diplomatic academy of the ministry of foreign affairs" or "Institute of microeconomics under the ministry of economic development" appeared regularly, despite not being direct ministry employers. Additionally, a substantial number of INNs in the dataset were mapped to commercial banks or reused across multiple, inconsistently labeled employers, introducing significant noise and undermining confidence in the accuracy of INN-based classification.

Moreover, even when excluding known INNs of banks and filtering for central institutions, INN-based linking remained noisy, with many INNs mapped to multiple, inconsistently labeled employers over time. Given this ambiguity and the lack of comprehensive INN-employer mapping for the early 2000s, I opted for the simpler keyword approach. Although conservative, it provides greater transparency, limits false positives, and ensures that flagged individuals had explicit textual references to formal ministries or presidential bodies.

This yielded 88,883 rows flagged as ministry employment (*is_ministry)*. Aggregating across years using individual identifiers, I identified 19,858 unique federal employees. Of these, 2,673 still lacked a usable birthdate even after targeted imputation, which prevented successful linkage to the traffic dataset. As a result, they do not appear in the final analysis sample due to insufficient identifying information. A breakdown of the top ten ministry employers flagged by this method is provided in Appendix B, Table 11.

3.2.5.4 *Aggregating to the individual level*

Once the cleaned and imputed dataset was prepared, I aggregated all records to the individual-level using a unique person identifier (*fio_dob_index*), combining name fields and date of birth. During aggregation, the *is_ministry* indicator was set to the maximum value across all years, meaning that individuals were flagged as ministry employees if they held such a role at any point between 1999 and 2004. Although variables such as employer names, ZIP code, and inferred gender (based on the patronymic) were also retained during this step, they were not used in the core analysis. These fields were preserved for completeness and potential exploratory use.

3.2.6 Dataset merging

3.2.6.1 *Ministry flag assignment via record linkage*

The concluding step before analysis involved assigning political connection indicators to the traffic database. The *is_ministry* flag was merged onto the traffic dataset using an exact match on four cleaned fields: surname (*Familya*), first name (*Imya*), patronymic (*Otshestvo*), and date of birth (*DoB_c*).

To maximize match accuracy, a final duplicate check was performed on the traffic data, revealing 80 duplicate records which were removed prior to merging. After cleaning, the pension dataset contained 17,238,593 rows, and the traffic dataset contained 10,405,707 records for 6,503,581 individuals. The *DoB_c* field, essential for exact matching, was available for 89.4% of traffic entries and 87.3% of pension entries.

The merge procedure resulted in 4,849,763 matched rows, or approximately 46.6% of the traffic dataset. Among these, 17,696 rows (corresponding to 17,653 car–person combinations) were flagged with *is_ministry* = 1. These correspond to 10,054 unique individuals who were employed by a federal ministry or presidential institution at some point during the observation window.

3.2.6.2 *False match risk*

Table 7 presents the number of unique values for each of the four merge keys . Because exact matches were required on all four fields, the chance of mistakenly linking unrelated individuals is extremely low.

Field	Traffic Dataset	Pension Dataset
Familya	418,196	629,245
Imya	41,104	102,688
Otshestvo	73,473	172,547
DoB_c	29,861	37,722

Table 7: Unique values for name and birthdate fields in traffic and pension datasets

Assuming independence across the four fields, the number of possible combinations is calculated as:

$$418{,}196 \times 41{,}104 \times 73{,}473 \times 29{,}861 \approx 3.77 \times 10^{20}$$

Given the size of both datasets (≈6.5 million and ≈17.2 million individuals), the expected number of false positive matches due to random collisions, referred to as the false positive rate (FPR) is:

$$FPR = \frac{6.5 \times 10^{6} \times 17.2 \times 10^{6}}{3.77 \times 10^{20}} \approx 2.96 \times 10^{-7}$$

In practical terms, this corresponds to fewer than 2 accidental matches across the entire merged dataset.

However, the assumption of full statistical independence does not perfectly hold. Russian names follow strong cultural conventions, leading to high-frequency combinations like *Ivan Ivanovich Ivanov*[4], and certain birthdates (e.g., January 1st, 1911 or 1901) are overrepresented due to placeholder defaults. For instance, 0.63% of traffic individuals (40,769) share the date 01.01.1911, and another 0.19% (12,568) share 01.01.1901. Together, these two dates alone account for over 0.8% of all individuals. The merged dataset contains nearly 30,000 distinct birthdates, but the prevalence of placeholder or default entries limits the effective uniqueness of *DoB_c*.

Even accounting for this, the estimated error remains acceptably low. A tenfold inflation of the baseline FPR would still imply fewer than 30 false matches, which is negligible relative to the 4.8 million merged rows. Moreover, the strict requirement for a four-way match guards against the risk of systematic mislinkage.

3.2.6.3 *Identifying family members of ministry employees*

To study nepotistic links, I identified individuals who were not ministry employees themselves but who shared both a household address and surname core with someone who was. This was done using the *family_id_final* constructed in Section 3.2.4.

Although 10,054 individuals are identified as ministry employees in the merged dataset, only 2,848 non-ministry individuals are flagged as family members through shared addresses and surname cores. At first glance, this may seem surprisingly low. Given an average household size of 3.34 in Moscow in 2002 (Shcherbov et al., 2023), one might expect approximately 23,774 family members across 10,054 employees ($10{,}054 \times (3.34 - 1)$, excluding the employee themselves). However, the dataset is not a population registry. It only includes individuals who: (1) Owned a car, or (2) Were involved in police traffic interactions (e.g., traffic violations or accidents), during the period 1999-2004.

During this period, car ownership in Russia ranged from 120 to 180 passenger vehicles per 1,000 people (Helgi Library; BOFIT Weekly Review 11/2017, March 17, 2017), suggesting that roughly 12–18% of the general population would be captured via vehicle ownership alone. This macro-level figure likely underestimates actual coverage in the dataset, since some individuals may appear due to traffic events even if they do not personally own a car. Nonetheless, applying this 12–18% range to the estimated 23,774 family members yields an expected coverage of roughly 2,853 to 4,279 vehicle-linked relatives.

The observed count of 2,848 family-linked individuals falls just below the lower-bound estimate (≈99.8%) and covers approximately 66.6% of the upper bound. This is consistent with a dataset shaped by administrative vehicle and enforcement records. These 2,848 unique family members account for a total of 4,729 rows in the merged traffic dataset, reflecting multiple registered vehicles or enforcement interactions per person.

In sum, while this approach cannot reconstruct the full household networks of politically connected individuals, it yields a conservative and analytically useful subset of relatives who appear in administrative traffic data. This subset is sufficient to evaluate potential nepotism effects in enforcement outcomes.

[4] See Figure 12 in Section 3.2.4. The distribution of surname cores, as well as first names and patronymics, is heavily skewed, and the variables are not independent, i.e. certain combinations (e.g., *Ivan Ivanovich*) occur disproportionately often.

3.2.6.4 *Summary statistics*

To conclude the data preparation, I present a structural summary of the final merged dataset. Figure 13 displays the distribution of unique car–person combinations by group affiliation and traffic outcome. As in earlier figures (see Section 3.1), I maintain consistent color coding: red bars indicate offenders, green innocents, and blue total entries. The left panel presents these distributions for non-ministry individuals. The right panel disaggregates the same categories for individuals identified as ministry employees and their family members.

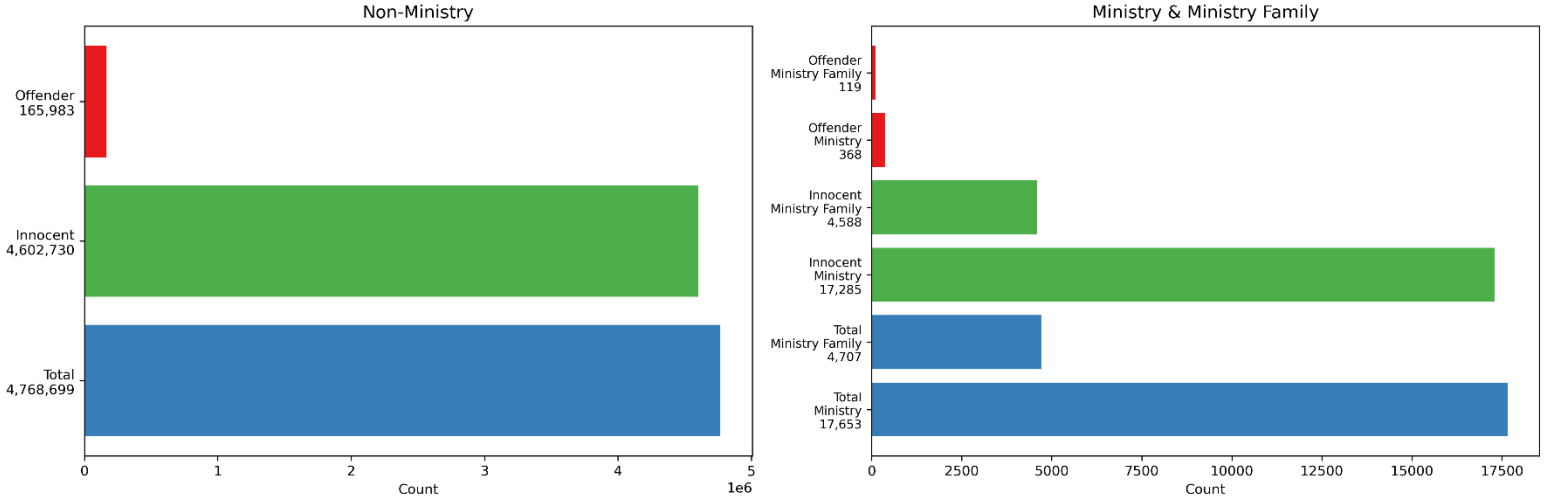


Figure 13: Distribution of car–person combinations by group status

Each bar represents the number of unique car-person combinations per group, emphasizing the sample size available for outcome analysis. The offender category is particularly important, as it serves as the foundation for identifying potential differences in enforcement. It includes 165,983 non-ministry combinations (3.48% of that group), compared to 368 for ministry employees (2.08%) and 119 for their relatives (2.53%). While the absolute number of offenders differs substantially across groups, the relative shares across groups are more comparable. This suggests that outcome-based comparisons are technically feasible, and the lower incidence rates among politically connected individuals already hint toward the possibility of more lenient treatment.

Group	Violated	Fined	LicDeprived	Accid
Ministry	286 (1.62%)	189 (1.07%)	6 (0.03%)	87 (0.49%)
Ministry Family	91 (1.93%)	64 (1.36%)	1 (0.02%)	28 (0.59%)
Unconnected	150,018 (3.15%)	112,501 (2.36%)	3,517 (0.07%)	16,964 (0.36%)

Table 8: Violation, fine, license deprivation, and accident rates by group

Importantly, Table 8 confirms that each group includes at least one observed instance of each enforcement outcome (violation, fine, license deprivation, and accident) across its car–person combinations. While the total number of events per group exceeds the number of unique combinations shown in Figure 13, this reflects the fact that multiple outcomes can occur for the same individual. For example, a single car–person combination may incur a violation, receive a fine, and later be involved in an accident. This completeness is essential as it ensures that outcome rates can be calculated and compared across all groups, even if some event types are relatively rare. That said, license deprivation presents a particular limitation: only 6 events are recorded for ministry employees and just 1 for family members. While Fisher's exact test is technically valid even with such small counts, the extreme sparsity limits the interpretability of the results.

Some early patterns already start to emerge. Across the three enforcement outcomes (violation, fine, and license deprivation) event rates are consistently lower for the connected groups, while accident rates are slightly higher. This combination is particularly noteworthy: if those with connections face fewer sanctions yet are involved in more accidents, it signals a form of impunity. This observation aligns with the theoretical framework outlined in Section 2.3, where selective leniency increases moral hazard, ultimately leading to higher public risk and greater welfare loss. In that model, lower expected fines for privileged individuals distort incentives, encouraging behavior that imposes external costs on society.

To conclude, Table 9 provides a summary of essential demographic and vehicle characteristics by group. These include age, gender, vehicle type, and license plate status: factors relevant for understanding observed enforcement differences.[5] Visualizations of age distributions across groups are presented in Appendix B, Figures 22 and 23.

Group	N	Mean Age	% Male	% Luxury Car	% Vanity Plate
Ministry	17,652	40.3	77.09%	12.73%	5.85%
Ministry Family	4,707	36.5	67.98%	17.04%	6.46%
Unconnected	4,768,700	36.2	77.34%	10.09%	5.19%

Table 9: Group-level summary of key driver and vehicle attributes

[5] On average, connected individuals are older (mean age = 40.1 vs. 36.2) and more likely to be female (24.8% vs. 22.7%) compared to unconnected individuals.

4 METHODOLOGY

4.1 Overview of empirical strategy

This chapter presents the empirical framework that is used to evaluate whether political connections and informal privilege signals affect traffic enforcement outcomes in Russia. The analytical strategy unfolds in three main stages.

First, I conduct descriptive comparisons of key enforcement outcomes (violations, fines, license deprivations, and accidents) across groups with and without political connections. These comparisons reveal preliminary patterns in sanction rates and behavioral risk. Second, I estimate logistic regression models to test whether connected individuals are sanctioned differently when controlling for demographic and contextual variables. This includes district fixed effects and individual-level controls. Third, I extend the analysis to test the role of informal privilege indicators (luxury vehicle ownership and vanity plates) and examine whether they independently predict leniency or if the inclusion of these variables alters the estimated effects of political connectedness.

Subsequent sections define the key variables and connection groups, introduce the enforcement timeline and points of discretion (Section 4.2), present the core hypotheses (Section 4.3.1), describe the construction of explanatory and outcome variables (Section 4.3.2), and outline the statistical models used for estimation (Section 4.3.3). The empirical results are presented and interpreted in Chapter 5.

4.2 Defining connectedness and informal influence

This study conceptualizes connectedness as the ability to access favorable treatment from state institutions, specifically traffic police, due to political, familial, or social capital. This influence may be formal (e.g., employment in a federal ministry), relational (e.g., family ties to such individuals), or perceived (e.g., visible signs of status or privilege). While previous research on nepotism (e.g., Szakonyi 2019) has focused on formal political or bureaucratic ties, I distinguish these from informal signals such as vehicle type or license plate, which may serve as perceived markers of privilege. These symbolic cues are not treated as direct evidence of connectedness but are included to control for visible status attributes that may influence enforcement discretion.

This distinction is particularly important in the context of traffic enforcement, where officers typically lack access to employment registries or administrative records during roadside stops. Instead, their judgments may rely on visible, immediate cues, such as vehicle type or license plate, that signal a driver's status, influence, or perceived capacity to retaliate or bribe. To capture these different channels, I define four separate indicators of potential connectedness or influence:

1. Direct political employment in a federal ministry (Level 0 connection)
2. Family ties to a ministry employee via shared address and surname core (Level 1 connection)
3. Ownership of a luxury vehicle
4. Possession of a vanity license plate

Together, these variables serve as alternative, and potentially complementary, pathways through which individuals may secure favorable treatment from traffic enforcement authorities.

4.2.1 Ministry employment and family ties (Level 0 & Level 1)

The primary explanatory variables in this study capture formal ties to the Russian federal government. I distinguish between two levels of political connectedness.

Level 0 includes individuals directly employed by a federal ministry or presidential institution between 1999 and 2004. This flag (*is_ministry*), derived from administrative earnings records, identifies persons with explicit bureaucratic authority or elite institutional affiliation (Section 3.2.5). It serves as the core measure of formal political connection.

Level 1 captures likely family members of Level 0 individuals. These are identified using standardized addresses and lemmatized surname cores to infer household-based kinship links within the traffic dataset (Section 3.2.4). If two individuals share both a household address and surname core, and one is flagged as Level 0, the other is flagged as Level 1 (*ministry_family_final_flag*). This approach infers relational ties to formal power in the absence of household registry data.

To maximize coverage, I use the "final" version of the family flag, which combines strict and relaxed address matching. While this method captures only co-resident individuals who appear in the traffic dataset, it offers a meaningful proxy for observable familial connections within enforcement contexts. These two flags are used throughout the analysis to test whether proximity to formal political power correlates with more lenient treatment in traffic enforcement.

4.2.2 Informal signals of privilege

In addition to formal political connections, this study incorporates two visible indicators of informal privilege: luxury vehicle ownership and vanity license plates. These variables are used in both descriptive comparisons and regression models to test whether symbolic markers of social status, influence traffic enforcement outcomes. Although not the primary focus, these indicators provide alternative pathways through which social privilege may shape law enforcement discretion.

Such effects are particularly relevant in roadside encounters, where officers are often required to make swift decisions based on limited observable information. In the absence of access to background records or employment data, they may instead rely on visual cues (such as a high-end vehicle or an atypical license plate) as immediate indicators of status, influence, or potential connections. These impressions can shape enforcement behavior on the spot, even when no actual institutional protection exists. This reasoning is further developed in Section 4.2.3, which outlines the key stages of discretionary enforcement.

The luxury indicator (*lux*) is a binary variable coded 1 if the vehicle's make appears in a curated list of premium automotive brands (e.g., BMW, Mercedes-Benz, Lexus, Porsche). This list was developed prior to analysis and is included in the dataset. It reflects brands that were particularly rare, expensive, or symbolically loaded in the early 2000s Russian context. Using this definition, 483,182 car-person combinations (approximately 10.1% of the dataset) containing 488,031 luxury car records are flagged. The full brand list is provided in Appendix B, Table 12.

The vanity plate indicator (*van*) flags license plates that match one out of 56 stylized three-digit sequences (e.g., 111, 777, 505), identified in prior work by Eeckhout et al. (2021) as culturally meaningful or status-signaling. Appendix B, Table 13, provides a breakdown of the number of combinations flagged under each variant of the vanity plate definition. Based on this definition, approximately 248,630 car–person combinations (roughly 5.2% of the dataset) are classified as having vanity plates.[6]

Originally included as controls, both indicators show statistically significant associations with enforcement outcomes. This motivates their inclusion not only as covariates but also as descriptive comparison groups. In later models, I test whether including these indicators alters the estimated effects of political connectedness. This makes it possible to assess whether informal privilege overlaps with or accounts for part of the observed disparity.

Together with the formal connection flags, these informal indicators help isolate how different forms of social privilege (both institutional and symbolic) correlate with the uneven application of traffic law.

4.2.3 Stages of enforcement and points of influence

Traffic enforcement in Russia between 1999 and 2004 followed a multi-stage process, governed primarily by the Code of Administrative Offences (KoAP RSFSR, 1984, amended through 2002; and later, the new KoAP RF, No. 195-FZ, effective July 1, 2002)[7]. These laws stipulated that administrative traffic violations could result in sanctions ranging from warnings and monetary fines to more severe penalties such as license suspensions or administrative detention.

The enforcement process allowed for discretion at multiple levels. For minor violations, such as speeding or running a red light, police officers from the traffic authority (GIBDD) had the authority to issue fines directly. However, more serious sanctions, such as license deprivation or administrative arrest, a judicial decision was required, typically issued by a magistrate's court (Zakopyrin et al., 2021, p. 1191)[8].

This multi-tiered structure creates several junctures where political connections or visible privilege may influence outcomes. As shown in Figure 14, I distinguish four stages along the enforcement chain:

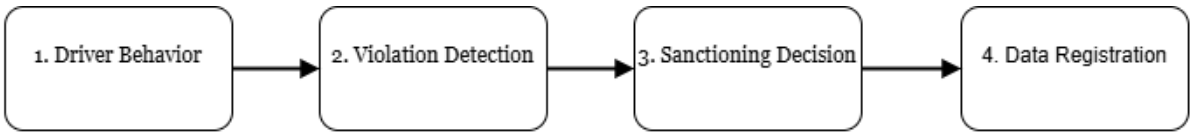


Figure 14: Enforcement timeline: from behavior to record

[6] Alternative vanity definitions were explored using variants that emphasize digit repetition or pattern strength. These yielded subsample sizes ranging from 154,736 to 365,747 flagged cases. I choose to work with the latest vanity definition variable, constructed by Eeckhout et al. (2021). Alternative definitions and their corresponding flagged counts are provided in Appendix B Figure 21

[7] Code of Administrative Offenses of the RSFSR. Adopted June 20, 1984. Amended through 2002. Code of Administrative Offenses of the Russian Federation (KoAP RF), Federal Law No. 195-FZ, adopted December 30, 2001, entered into force July 1, 2002.

[8] See Zakopyrin et al. (2021) for a general discussion of administrative penalties and the distinction between police and judicial authority in administrative law across Russia and other CIS countries.

1. Driver behavior: As discussed in Section 2.3, perceived protection, whether due to institutional ties or visible signals of privilege, may alter how individuals behave on the road. When drivers believe that the consequences of infractions are lower for them, they may take more risks. This behavioral adaptation, or moral hazard, may lead to a paradoxical pattern: fewer sanctions, yet higher accident rates among connected individuals.

2. Violation detection: Police officers have discretion over whether to register a traffic offense. This decision typically occurs during a roadside stop, often based on the officer's direct observation. At this stage, visible signals such as vehicle brand or license plate may serve as cues of status or influence. Officers may opt not to register minor violations committed by individuals who appear privileged, whether due to implicit bias, perceived bribery potential, or fear of retaliation or repercussions.

3. Sanctioning: Once a violation is formally recorded, officers (for minor cases) or courts (for more serious sanctions) determine whether and how to punish the offender. Fewer fines conditional on a registered violation would suggest discretionary leniency, while fewer license deprivations among the connected might indicate preferential judicial treatment.

4. Data registration: Finally, some degree of discretion may occur during data entry. Although this step is largely administrative, it is not immune to informal influence. Omitting or modifying records may occur under pressure, especially in high-profile or politically sensitive cases.

This staged structure is central to the empirical strategy. By examining differences in violation rates, fine issuance rates, license deprivation, and accidents, I am able to isolate the specific stage in the process at which connectedness appears to exert its strongest influence. If connected individuals have lower fine rates conditional on a violation, this suggests discretion at the sanctioning stage. If they are less likely to receive a violation to begin with, then discretion likely occurs at the detection stage. If they have more accidents despite fewer sanctions, this supports the behavioral hypothesis that perceived impunity increases risk-taking. Since license suspension requires judicial action, this can point to preferential treatment at the court level, not just informal leniency at the roadside.

The variables introduced in Sections 4.2.1 and 4.2.2 map onto this framework accordingly. Formal connections may operate across all stages, whereas informal signals like luxury vehicles and vanity plates are particularly relevant during roadside interactions, where enforcement decisions are shaped by what is immediately visible.

4.3 Hypotheses and empirical framing

Building directly on the staged enforcement framework outlined above, this section formalizes the empirical hypotheses to be tested throughout the remainder of the analysis. By mapping political connectedness and informal signals of privilege to different stages of the enforcement process, I aim to distinguish where and how preferential treatment arises. This structured approach allows for disentangling behavioral responses, frontline police discretion, and judicial outcomes, each potentially reflecting different mechanisms of influence.

Section 4.3 is organized in three parts. First, I lay out the main hypotheses in intuitive language terms and link them to relevant enforcement stages (Section 4.3.1). Second, I define the outcome variables and comparison groups used to measure disparities in enforcement (Section 4.3.2). Finally, I outline the statistical models and estimation strategies used to test each hypothesis, including discussion of control variables, fixed effects, and statistical testing procedures (Section 4.3.3).

4.3.1 Hypotheses and conceptual framing

The enforcement process involves several potential points of influence. Each hypothesis is linked to a specific stage and is grounded in the logic of strategic interaction between police officers, courts, and individuals who possess (or signal) privileged status.

- **Hypothesis 0: Preferential fine avoidance (core hypothesis)**
    - *H0*: Politically connected individuals are equally likely to receive fines as others.
    - *H1*: Politically connected individuals are less likely to receive fines.

This is the core hypothesis of the thesis. Monetary fines represent the clearest and most consequential discretionary sanction in the enforcement process. While violations may occur more frequently overall, the issuance of a fine directly reflects a deliberate decision by officers to escalate a detection into a formal penalty. A reduction in fine rates among connected individuals would represent strong evidence of leniency or preferential treatment.

- **Hypothesis 1: Behavioral adaptation: moral hazard**
    - *H0*: Connected individuals are no more likely to cause accidents.
    - *H1*: Connected individuals are more likely to cause accidents.

This hypothesis explores the possibility of moral hazard arising from anticipated leniency. If individuals anticipate protection from enforcement, they may engage in riskier behavior. A higher accident rate among connected individuals would suggest that reduced sanctioning has real behavioral consequences. This effect would be particularly compelling if observed in combination with lower sanction rates. The underlying dynamic is elaborated in Section 2.3 and illustrated in Figure 3, which models how weakened accountability can distort private incentives and increase public risk.

- **Hypothesis 2: Discretion at the violation detection stage**

    - *H0*: Connected individuals are equally likely to be cited for violations.
    - *H1*: Connected individuals are less likely to be cited for violations.

Violation registration occurs at the roadside and represents the first discretionary decision in the enforcement process. Police officers may choose not to record infractions committed by high-status individuals. Although the underlying discretion is not directly observable, systematically lower violation rates among connected individuals may imply that fewer offenses are being documented. This hypothesis thus captures potential discretion at the initial stage of enforcement.

- **Hypothesis 3: Discretion at the sanctioning stage (conditional)**

    - *H0*: Conditional on a violation, connected individuals are equally likely to be fined.
    - *H1*: Connected individuals are less likely to be fined.

This hypothesis isolates decision-making after a violation has been registered. Once a violation has been formally registered, police officers still retain discretion over whether to fine the individual. If connected individuals are systematically less likely to receive fines after being stopped, this would suggest leniency at the sanctioning stage.

- **Hypothesis 4: Preferential judicial treatment**

    - *H0*: Connected individuals are equally likely to receive license deprivation sanctions.
    - *H1*: Connected individuals are less likely to be deprived of their license.

License suspension requires a formal court decision. A lower incidence of license deprivation among connected individuals may point to judicial leniency or the influence of external factors within legal proceedings.

- **Hypothesis 5: Informal signals of privilege**

    - *H0*: Vehicle luxury or vanity status has no effect on enforcement.
    - *H1*: Individuals with luxury vehicles or vanity plates are treated differently (either more leniently or more harshly).

This hypothesis examines the role of visible cues. These may influence officer behavior at the point of detection. On one hand, they may signal influence or institutional backing, which deters enforcement. On the other hand, they may mark individuals as “bribeable” or embedded in informal networks, prompting more stops or sanctions in anticipation of informal payments.

- **Hypothesis 6: Informal privilege as a potential pathway**
    - *H0*: The estimated effect of political connectedness is unaffected by the inclusion of luxury and vanity controls.
    - *H1*: The estimated effect of political connectedness diminishes when luxury and vanity are included.

This hypothesis explores whether visible indicators, such as luxury vehicles or vanity plates, partially account for the observed effects of formal political connections. While no mediation test is conducted, I assess whether the coefficient on political connectedness meaningfully declines when these informal indicators (*lux*, *van*) are added to the model. A substantial reduction would suggest a potential mediating role for informal privilege.

4.3.2 Analytical setup and measurement framework

This section defines the key outcome variables, explanatory groups, and control variables used in both the descriptive comparisons and regression analyses. It also outlines the measurement formulas applied throughout the empirical analysis.

Unit of analysis

The unit of observation is the *car–person combination*, as introduced in Chapter 3. Each record links a unique vehicle to its registered owner. This allows for consistent measurement of whether a person–vehicle pair was ever sanctioned (probability), and how many such events occurred (intensity), without double-counting across events or individuals. This unit is particularly appropriate in a setting where enforcement is vehicle-specific, yet political or social privilege is person-specific.

Outcome variables

Five binary outcome variables are constructed at the car-person level:

- *received_fine*: at least one monetary fine recorded
- *had_violation*: at least one traffic violation recorded
- *license_deprived*: any license suspension
- *was_in_accident*: involvement in at least one traffic accident
- *fine_given_violation*: a fine was issued, conditional on having a violation

Each variable equals 1 if the event occurred at least once for the combination, and 0 otherwise. These outcomes are derived from raw event-level traffic data and aggregated to the combo level.

Group indicators

To compare enforcement outcomes across strata of connectedness, the sample is split in three mutually exclusive groups:

- Level 0 (L0): ministry employees (*is_ministry* = 1)
- Level 1 (L1): family members of ministry employees (*ministry_family_final_flag* = 1 and *is_ministry* = 0)
- Unconnected individuals (UC): neither ministry employees nor family members

Additionally, the binary luxury vehicle and vanity plate indicators (*lux* = 1 and *van* = 1) capture informal or symbolic indicators of privilege. These variables are used descriptively to illustrate how informal status cues relate to enforcement outcomes, and as controls in regression models to test whether the effects of political connectedness hold when accounting for them.

Control variables and fixed effects

Two individual-level characteristics are included as controls in the regression models:

- *AgeAtEntry*: age when the individual enters the dataset (median-imputed for 14.8% missing values, n=711,430)
- *Gender*: a binary indicator for gender (1 = female, 0 = male)

To account for enforcement heterogeneity across geographic areas, the analysis includes fixed effects for traffic police districts (PDs). Since enforcement events can span multiple locations for the same individual, and the unit of analysis is the car–person combination, each combo is assigned a single PD based on the most frequently observed district across their enforcement records. Therefore, each car-person combination is assigned the most frequently observed PD across all their enforcement records. While this is an approximation, it serves as a reasonable proxy for the typical enforcement environment an individual is exposed to.

To avoid oversaturation and instability in models, the raw PD variable (259 categories) is collapsed into a coarser *PD_grouped* variable:

- *PD_grouped*: recoded version where districts with fewer than 10,000 observations are grouped into an "Other" category

This reduces the dimensionality to 77 categories, including "Other" and "Missing".

Measurement strategy

Two key summary statistics are used in the descriptive comparison:

1. Probability of event (extensive margin):

$$P(Event) = \frac{Number\ of\ car-persons\ with\ at\ least\ one\ event}{Total\ number\ of\ car-persons\ in\ group}$$

2. Average number of events per combo (intensive margin):

$$Avg.Events = \frac{1}{N}\sum_{i=1}^{N} Event\ count_i$$

These complementary measures show both how many individuals are affected, and how frequently, helping to distinguish targeted leniency from broader underenforcement.

These measures provide complementary views of enforcement differences. They explain whether a group is more likely to be sanctioned at all, and if so, how often.

4.3.3 Model specifications and estimation strategy

This section outlines the estimation framework used to formally test the hypotheses from Section 4.3.1, using the variables and measurement strategy introduced in 4.3.2.

4.3.3.1 *Descriptive analysis approach*

To explore baseline differences in enforcement outcomes, initial comparisons are conducted without control variables. For each outcome variable, both extensive and intensive margin statistics are computed and compared across Level 0, Level 1, and unconnected individuals. Similar comparisons are also made for individuals with luxury vehicles and vanity plates.

To assess the statistical significance of observed differences, I employ two complementary hypothesis testing strategies depending on the frequency of the outcome variable.

Two-sample t-tests are used for comparing means or proportions across groups when outcomes are frequent enough for large-sample approximations to hold. In practice, this applies to most enforcement outcomes, such as *received_fine*, *had_violation*, and *was_in_accident*, where the number of events is sufficiently high. These tests assess whether the difference in average rates (e.g., fine rates) between groups like ministry-connected and unconnected individuals is statistically distinguishable from zero.

Fisher's exact test is used for rare outcomes where standard approximations may be invalid. This is especially relevant for *license_deprived* (6 cases among 17,659 ministry employees; 1 case among 4,714 family members) and *fine_given_violation* (conditional analysis within a subset of violators). In such cases, the t-test is inappropriate due to sparse data or conditioning. Fisher's test computes an exact p-value based on the 2×2 contingency table, offering a valid and conservative test of group differences.

Using this tailored approach ensures appropriate inference across the full range of enforcement outcomes, from frequent administrative sanctions to rare judicial penalties.

4.3.3.2 *Regression model specification*

A set of logistic regression models is specified to estimate how formal connection status and informal privilege indicators affect the likelihood of binary enforcement outcomes, controlling for relevant covariates.

General logit model

The logit model estimates how the explanatory variables affect the log-odds of the event occurring. By applying the inverse logit (logistic) function, these log-odds can be transformed into predicted probabilities of the outcome.

The generic form of the model is:

$$logit(Pr(Y_i = 1)) = \beta_0 + \beta_1 is_ministry_i + \beta_2 ministry_family_final_flag_i + \beta_3 lux_i + \beta_4 van_i + \boldsymbol{X_i}\gamma + \delta j + \varepsilon_i$$

Where:

- Y_i be the binary outcome of interest for individual i
- $is_ministry_i$: indicator for direct political employment
- $ministry_family_final_flag_i$: indicator for family ties to ministry employee
- lux_i: indicator for luxury vehicle
- van_i: indicator for vanity plate
- $\boldsymbol{X}_i$: vector of individual-level controls (age and gender)
- δj: police district fixed effects (capturing local enforcement differences)
- ε_i: error term
- $Predicted\ probability$: $logit(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki}$

$$\Rightarrow p_i = 1 + \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki})}}$$

Each hypothesis corresponds to a specific implementation of this model:

Ho (Preferential fine avoidance)

- Ho: $\beta_1 = 0, \beta_2 = 0$
- $H1$: $\beta_1 < 0, \beta_2 < 0$

$$logit(Pr(received_fine = 1)) = \beta_0 + \beta_1 is_ministry_i + \beta_2 ministry_family_final_flag_i + \boldsymbol{X_i}\gamma + \delta j + \varepsilon_i$$

H1 (Behavioral adaptation)

- $H0: \beta_1 = 0, \beta_2 = 0$
- $H1: \beta_1 > 0, \beta_2 > 0$

$$logit(Pr(was_in_accident = 1)) = \beta_0 + \beta_1 is_ministry_i + \beta_2 ministry_family_final_flag_i + \boldsymbol{X_i}\gamma + \delta j + \varepsilon_i$$

H2 (Violation detection discretion)

- $H0: \beta_1 = 0, \beta_2 = 0$
- $H1: \beta_1 < 0, \beta_2 < 0$

$$logit(Pr(had_violation = 1)) = \beta_0 + \beta_1 is_ministry_i + \beta_2 ministry_family_final_flag_i + \boldsymbol{X_i}\gamma + \delta j + \varepsilon_i$$

H3 (Sanctioning discretion, conditional)

- $H0: \beta_1 = 0, \beta_2 = 0$
- $H1: \beta_1 < 0, \beta_2 < 0$

$$logit(Pr(fine_given_violation = 1)) = \beta_0 + \beta_1 is_ministry_i + \beta_2 ministry_family_final_flag_i + \boldsymbol{X_i}\gamma + \delta j + \varepsilon_i$$

H4(Judicial leniency)

- $H0: \beta_1 = 0, \beta_2 = 0$
- $H1: \beta_1 < 0, \beta_2 < 0$

$$logit(Pr(license_deprived = 1)) = \beta_0 + \beta_1 is_ministry_i + \beta_2 ministry_family_final_flag_i + \boldsymbol{X_i}\gamma + \delta j + \varepsilon_i$$

H5 (Informal signals of privilege)

- *H0*: No difference in event rates based on luxury or vanity indicators
- *H1*: Individuals with luxury vehicles or vanity plates differ in event rates

Assessed using descriptive statistics only. No regression model is estimated.

H6 (Attenuation by informal signals)

Compare nested versions of the main model:

Model A: $logit(Pr(Y_i = 1)) = \beta_0 + \beta_1 is_ministry_i + \varepsilon_i$
Model B: $logit(Pr(Y_i = 1)) = \beta_0 + \beta_1' is_ministry_i + \beta_3 lux_i + \beta_4 van_i + \varepsilon_i$

Model C: $logit(Pr(Y_i = 1)) = \beta_0 + \beta_2 ministry_family_final_flag_i + \varepsilon_i$
Model D: $logit(Pr(Y_i = 1)) = \beta_0 + \beta_2' ministry_family_final_flag_i + \beta_3 lux_i + \beta_4 van_i + \varepsilon_i$

And evaluate whether $|\beta_1'| < |\beta_1|$ & $|\beta_2'| < |\beta_2|$, which would indicate that part of the effect attributed to political connection overlaps with informal signals such as luxury and vanity plates. This is not a mediation test but a comparison of nested models to assess attenuation.

When PD fixed effects are included, the *PD_grouped* variable enters as a categorical covariate. A lower grouping threshold (e.g., 1,000 observations per PD) led to convergence issues in some models. The threshold was therefore increased to 10,000 observations, which ensured stable estimation and consistent inclusion of district fixed effects across all models.

All models cluster standard errors at the individual level to adjust for repeated observations across car-person combinations. Standard errors are clustered at the individual level to account for the fact that the same person may appear in multiple car–person combinations, introducing dependence across observations.

This approach allows for a stage-by-stage evaluation of the enforcement process, linking formal political connections to enforcement outcomes, and examining the role of informal privilege indicators, while controlling for confounders and contextual heterogeneity.

5 RESULTS

In this chapter, I test the empirical hypotheses developed in Chapter 4 to assess whether individuals with political or bureaucratic connections are treated differently in Russian traffic enforcement. Using the administrative data described earlier, I compare enforcement outcomes for connected individuals (L0 and L1) to those of unconnected (UC) drivers. In addition, I examine whether informal status signals such as luxury vehicles and vanity license plates confer similar advantages.

The chapter proceeds in stages. Section 5.1 provides a descriptive overview of enforcement and accident outcomes across key groups (H0, H1, H2, H4), while Section 5.2 groups H1 and H3 under the theme of conditional discretion and behavioral risk. Sections 5.3 and 5.4 present formal regression results. Section 5.3 evaluates the core hypothesis by testing whether political connections reduce the overall probability of receiving a fine (H0), while Section 5.4 tests H1 through H4 using fully specified logit models. Section 5.5 turns to informal privilege indicators, luxury vehicles and vanity plates, to evaluate H5 and H6.

	# Fines	P(Fine)	# Violations	P(Violation)	# Fines | Violation	P(Fine | Violation)	# License Depr.	P(License Depr.)	# Accidents	P(Accident)
Ministry	189	1.0707%	286	1.6202%	189	66.0839%	6	0.0340%	87	0.4929%
Family	64	1.3597%	91	1.9333%	64	70.3297%	1	0.0212%	28	0.5949%
Unconnected	112501	2.3592%	150018	3.1459%	112501	74.9917%	3517	0.0738%	16964	0.3557%
Luxury	9063	1.8757%	12632	2.6143%	9063	71.7464%	161	0.0333%	2147	0.4443%
Non-Luxury	103563	2.4108%	137609	3.2034%	103563	75.2589%	3361	0.0782%	14903	0.3469%
Vanity	5116	2.0577%	6908	2.7784%	5116	74.0591%	122	0.0491%	843	0.3391%
Non-Vanity	107638	2.3696%	143487	3.1588%	107638	75.0159%	3402	0.0749%	16236	0.3574%

Table 10: Enforcement outcomes by group

Before turning to the formal hypotheses, I begin with a general comparison of enforcement outcomes across groups. Table 10 summarizes the share of car-person combinations in each category who received a fine, committed a violation, lost their license, or were involved in a traffic accident. These differences already suggest at a pattern of unequal treatment. For instance, ministry officials are far less likely to be fined or cited than unconnected drivers, yet they are more likely to be involved in traffic accidents. These patterns motivate the more detailed analysis that follows.

5.1 Descriptive results by political connection

The data reveal stark differences in enforcement outcomes between politically connected drivers (both officials and their family members) and unconnected drivers. As shown in Appendix C, Tables 14 and 15, only 1.07% of ministry officials and 1.36% of officials' family members received a traffic fine during the observation period, compared to 2.36% of unconnected drivers. These correspond to relative reductions of approximately 54.7% and 42.4%, respectively. Recorded violations exhibit similarly notable disparities: ministry officials and their families were 48.5% and 38.5% less likely to have a violation on record than unconnected drivers. While both fine and violation rates are substantially lower for the connected, the fact that the relative reduction in fines is consistently larger than that in violations suggests an additional layer of leniency. This implies that even when a violation is recorded, connected individuals may be less likely to face a resulting fine. This discrepancy anticipates the conditional enforcement patterns explored in Section 5.2.

Similarly, license revocations are extremely rare among connected drivers: only 0.034% of ministry officials and 0.021% of their family members had their licenses suspended or revoked, compared to 0.074% among unconnected drivers. Although the absolute rates are low, connected individuals appear less likely to face license deprivation ($p < 0.01$ for both). These are observed relative differences of approximately 54.1% and 71.6% for ministry officials and their family members, respectively. By contrast, connected groups show higher involvement in accidents: 0.493% of officials and 0.595% of family had an accident, versus 0.356% of unconnected drivers. These are relative differences of 38.5% and 67.2% in the positive direction this time. All reported differences come back highly statistically significant.

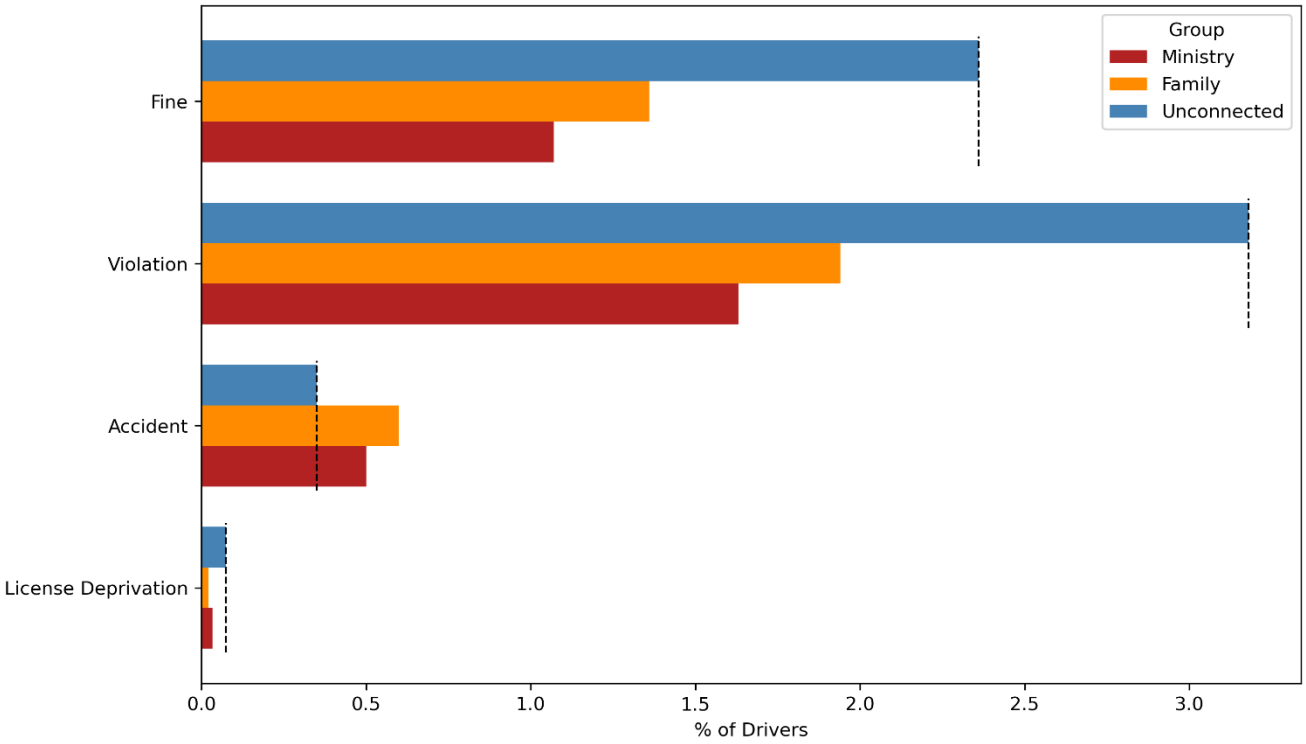


Figure 15: Enforcement outcomes by political connection group

The trends are visually summarized in Figure 15. For each enforcement outcome, the bars representing ministry officials and their family members consistently fall below the unconnected baseline (dashed line), confirming their lower likelihood of receiving fines, accumulating violations, or losing their license. The one notable exception is accident involvement, where both connected groups exhibit higher rates than unconnected drivers.

5.2 Conditional discretion and behavioral risk (H3 & H1)

In this section, I explore whether the enforcement gaps identified earlier extend to more specific mechanisms of discretion. H3 investigates whether connected individuals are fined less likely to receive a fine after a violation is recorded, while H1 examines whether reduced sanction risk is associated with increased accident involvement. Together, these hypotheses assess whether preferential treatment not only affects enforcement outcomes but also shapes driver behavior through perceived immunity (as discussed in Section 2.3).

H3: fine given violation

While connected drivers are less likely to be fined or cited overall, they also appear to benefit from leniency even when a violation is formally recorded. Among car-person pairs with at least one documented violation, only 66.1% of ministry officials received a fine, compared to 75.0% of unconnected violators. Family members fall in between, with 70.3% of such pairs receiving a fine. However, this difference is not statistically significant at conventional thresholds. This is largely due to the small sample size: only 91 family members committed at least one violation, of whom 64 were fined. With such limited data, even a seemingly meaningful difference in proportions yields wide confidence intervals that overlap substantially with the unconnected baseline.

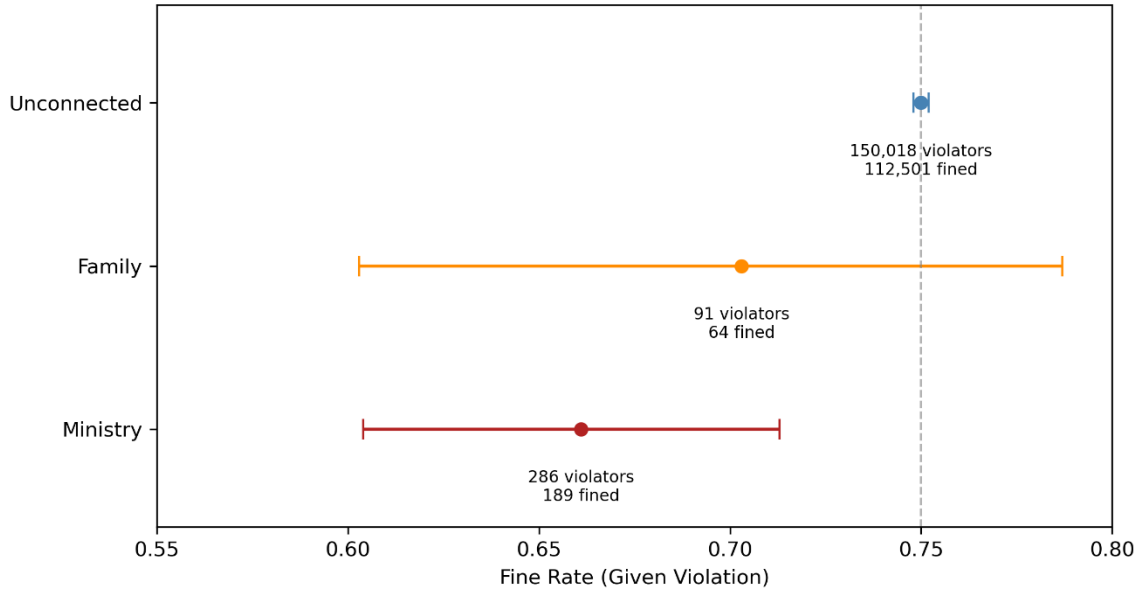


Figure 16: Conditional fine rates with 0.95 CI

Figure 16 visualizes these conditional fine rates along with their 95% confidence intervals. The wider intervals for ministry and family members illustrate the imprecision resulting from smaller sample sizes. The estimate for ministry officials is clearly below the unconnected baseline and statistically significant at the 1% level ($p \approx 0.0008$), offering strong support for H3. The family group also shows a lower share of violators who received at least one fine, but the effect is statistically inconclusive due to limited power. A power analysis suggests that detecting a gap of this magnitude with 80% confidence would require more than 1410 family violators, nearly eight times the number observed in the data (see Appendix A, Note 2).

To complement the binary indicator of whether a fine was issued, I also examine the average number of fines per violator as an alternative measure of sanctioning intensity. *Avg fines per violator* is based on all car–person combinations who had at least one recorded violation, and reflects the average number of fines received across those individuals. Unlike the binary *P(fine given violation)* rate, which measures whether at least one fine occurred, this continuous indicator captures how many fines violators typically received. It reflects repeated enforcement actions for individuals who were flagged for violations, and provides a more sensitive indicator of differential treatment. As shown in Figure 17, ministry officials receive an average of 0.741 fines, and their family members 0.824, compared to 0.995 among unconnected violators. These differences are both substantial and statistically significant: ministry officials receive 25.5% fewer fines per violator ($p < 0.001$), and family members 17.2% fewer ($p \approx 0.021$). The gap persists despite holding the denominator

(violators) constant, which strengthens the case that connected individuals benefit from sanctioning leniency even when a violation is documented. This pattern supports H3 more robustly than the binary fine indicator alone and helps compensate for the limited power of the conditional fine rate test. It strongly suggests the fine-receiving-gap is not just about fewer violations being recorded, it's also about what happens after.[9]

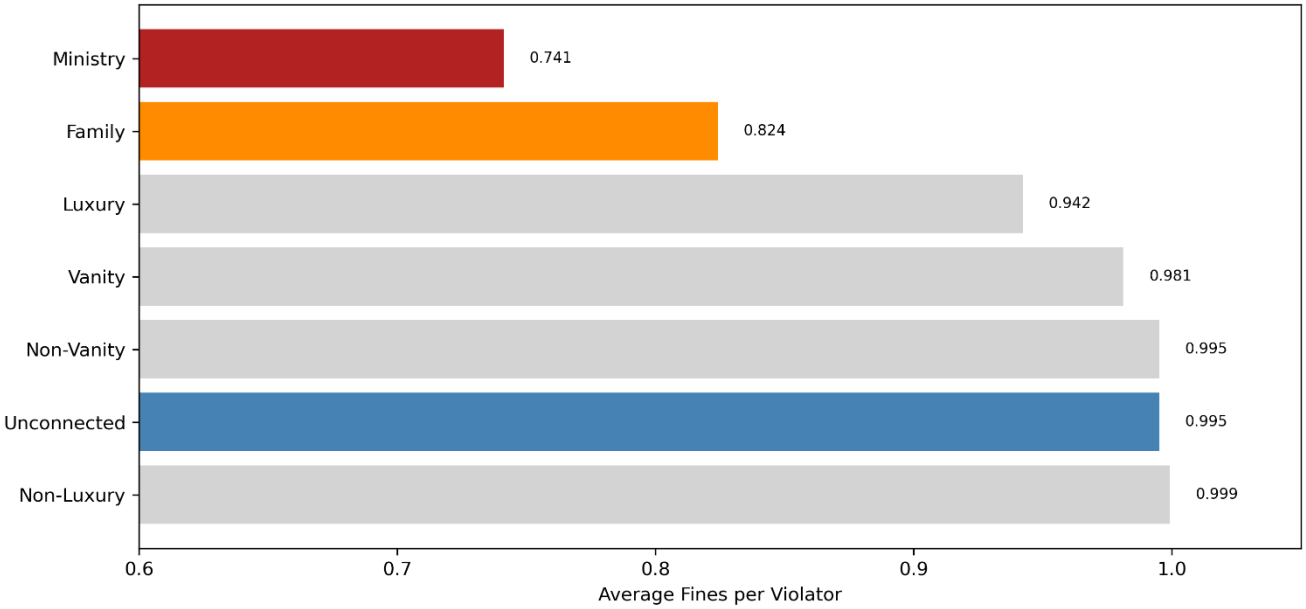


Figure 17: Average number of fines per violator by group

H1: Accident risk and behavioral adaptation

Despite being sanctioned less frequently, connected individuals are involved in traffic accidents at significantly higher rates. As reported in Section 5.1, 0.493% of ministry officials and 0.595% of their family members experienced an accident, compared to 0.356% of unconnected drivers. The relatively higher differences of 38.5% and 67.2% suggest a reversal in the enforcement pattern: connected individuals benefit from more lenient sanctioning, yet simultaneously they are more likely to experience adverse outcomes, such as traffic accidents.

This reversal is consistent with the moral hazard interpretation discussed in Section 2.3: when individuals expect to face fewer consequences for risky behavior, they may respond by engaging in riskier actions. The higher accident rates among connected groups, despite their lower rates of recorded violations and sanctions, suggest that enforcement gaps may have behavioral consequences. In this context, the accident data reinforce the idea that unequal enforcement is not merely symbolic, it may actively shape how connected individuals behave behind the wheel. This inverse pattern implies that many risky acts by connected drivers go unrecorded until they culminate in accidents. This discrepancy underscores how enforcement gaps can translate into real safety costs.

[9] As a robustness check, I also computed the group-level average number of fines per violation by dividing the total number of fines by the total number of violations within each group. This alternative summary aggregates across individuals, rather than averaging individual fine rates, and captures sanctioning intensity at the population level. Results confirm the main pattern: ministry officials received significantly fewer fines per violation than unconnected drivers (0.654 vs. 0.721, $p = 0.006$), while the difference for family members (0.670 vs. 0.721) was not statistically significant ($p = 0.264$). This aligns with the main findings, where family effects are directionally consistent but imprecise due to smaller sample sizes. P-values were estimated via bootstrapping to accommodate the uncertainty of ratio comparisons across unequal group sizes. Detailed summary statistics are presented in Appendix C, Table 18.

5.3 Regression results: preferential treatment in fines

This section tests Hypothesis H0, which posits that politically connected drivers (L0 & L1) are less likely to be fined for traffic violations than unconnected individuals. Descriptive evidence already suggested a large disparity: ministry officials are approximately half as likely to be fined as unconnected drivers. However, raw differences could arise from other factors, such as underlying demographic or behavioral differences. Connected individuals are a small (<1% of sample) and potentially distinct subset of drivers. Many are indeed older on average (see Table 9, Chapter 3), and may drive in different areas or face distinct enforcement contexts.

To isolate the effect of connection status on the likelihood of receiving a fine, we estimate logistic regressions of fine receipt on connection indicators, incrementally adding control variables. Appendix Table 19 reports the full results. The models are specified as follows:

- Model 1&2: bivariate model with only the connection indicator (either L0 or L1)
- Model 3&4: adds control variables (driver's gender and age).
- Model 5&6: adds both controls and PD fixed effects (capturing unobserved differences across regional police departments)

In the baseline model, ministry officials have a coefficient of −0.803 (SE = 0.074), corresponding to roughly 55.2% lower odds of receiving a fine compared to unconnected drivers. Family members show a smaller but still significant effect of 43% lower odds (−0.559, SE = 0.125). Adding gender and age (Models 3 & 4) modestly reduces these effects: officials drop to −0.712, family to −0.484. This reduction partly reflects demographic differences: in our data, female drivers are fined about 74% less often, and each additional year of age reduces fine odds by 1.6%. Connected individuals differ from the average driver on both dimensions: they are on average older (40.1 vs. 36.4) and more likely to be female (24.8% vs. 22.4%). These demographic differences help explain part of the raw fine gap. Including age and gender as controls narrows the disparity, but a significant effect remains.

Including PD fixed effects has a substantial impact. The ministry coefficient drops to −0.333 (SE = 0.075), while the family coefficient becomes small and statistically insignificant (−0.042, SE = 0.128). This suggests that location plays a large role in observed differences.

However, interpretation of the family result requires caution. The PD fixed effects absorb variation across 77 districts, each represented by a dummy variable. This approach uses substantial degrees of freedom and requires large, balanced sample sizes across areas. Yet family-flagged individuals constitute only about one-third of the ministry group, and are fewer than 0.1% of the full driver dataset. It's difficult to determine whether the family effect disappears due to substantive similarity within districts, or due to limited statistical power once fixed effects are included. Nonetheless, adding PD fixed effects increases the model's pseudo R^2 from 0.0247 to 0.1358, indicating a substantial improvement in explanatory power. Future research should therefore account for regional variation in enforcement practices. Failing to control for location may conflate institutional discretion with geographic clustering of connected individuals.

Even so, the ministry effect remains significant and sizable even within districts, suggesting that preferential treatment in fine issuance exists not only across areas, but within them. Officials receive fewer fines than comparable unconnected individuals even when demographics and local enforcement environment are held constant.

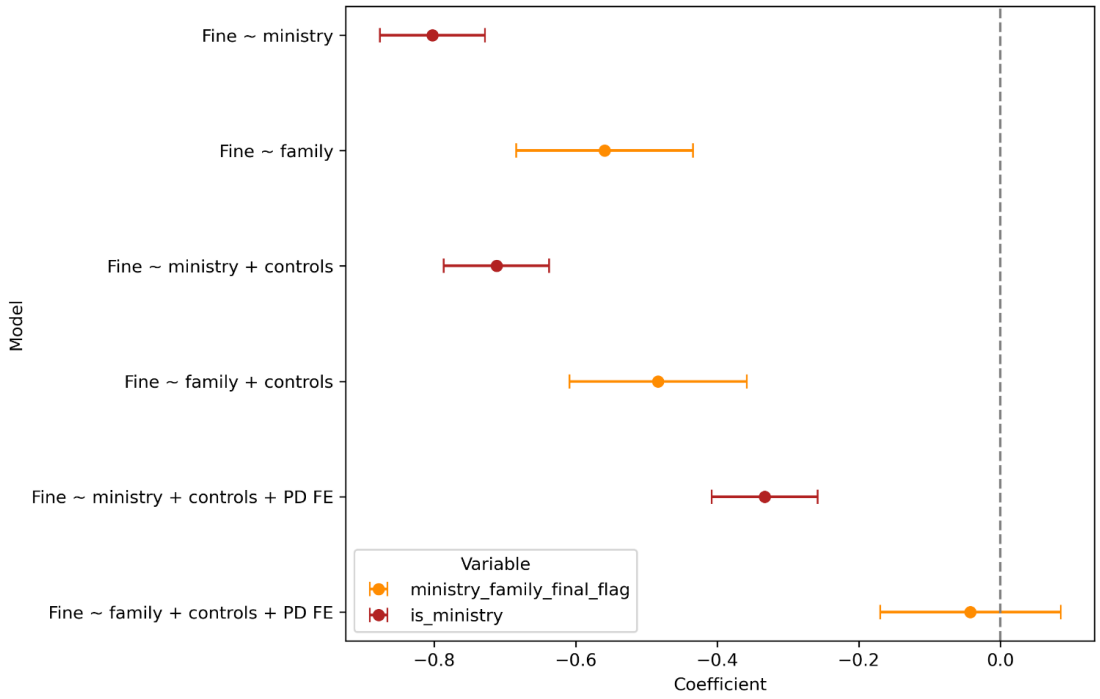


Figure 18: Estimated effects of political connection on fine receipt across model specifications

Figure 18 shows the coefficients from the logistic regressions predicting fine probability. Points show estimated log-odds coefficients and error bars represent 95% confidence intervals. Ministry officials retain a significant negative coefficient across all models. Wider intervals for family members reflect a smaller group size and higher statistical uncertainty.

In sum, H0 is strongly supported: connected drivers, especially ministry officials, are sanctioned at significantly lower rates. The estimated effect remains robust to controls and fixed effects, though somewhat attenuated. This naturally prompts a rhetorical follow-up question: does this pattern reflect safer driving by connected individuals? As shown in Section 5.2, connected drivers are actually more likely to be involved in traffic accidents. This casts doubt on the safer-driving interpretation and instead supports the idea of discretionary leniency in enforcement.

5.4 Regression results: additional enforcement stages (H1-H4)

The main hypothesis is supported, and hypotheses H1 through H4 also find strong descriptive backing. Taken together, the patterns are internally consistent and suggestive of enforcement asymmetries: connected individuals are cited less often for traffic violations, face fewer fines, and are less likely to have their license revoked. At the same time, they are more often involved in traffic accidents.

In the next sections, I test these associations more rigorously using logistic regression models analogous to the analysis of fine issuance. For each hypothesis, the same key independent variables (L0 & L1) are included, while incrementally adding demographic controls and police area fixed effects to assess robustness.

H1: Accident risk, evidence of behavioral moral hazard
Appendix C, Table 22 presents the full regression results. Since accidents are rare, the models feature large negative intercepts. In the baseline specification, ministry officials show significantly higher odds of accident involvement (+0.327, SE = 0.110), as do their family members (+0.515, SE = 0.189). These coefficients translate to an approximate 39% and 67% increase in odds, respectively, confirming the descriptive pattern of increased accident risk among connected individuals. However, there is no statistically significant difference between the two connected groups.

Adding control variables for gender and age has minimal impact on the estimated accident effects. In fact, the ministry coefficient slightly increases to +0.386 and the family coefficient remains nearly unchanged at +0.521. This indicates that demographic differences do not account for the elevated accident rates among connected individuals. In the data, male drivers and younger drivers are significantly more likely to be involved in accidents: being male increases the log-odds of accident involvement by approximately 0.14, while each additional year of age reduces it by about 0.015. Since ministry officials and their family members tend to be older on average and have a higher proportion of female drivers, adjusting for age and gender actually amplifies the relative risk associated with political connection. This stands in contrast to the fine regressions in Section 5.3, where adding demographic controls attenuated the connection effects. This reinforces the interpretation that it is the behavioral factor (risk-taking under perceived protection) driving the accident difference, not demographic characteristics.

Including PD fixed effects reduces both effects: ministry officials drop to +0.173 ($p \approx 0.117$), while the family effect becomes marginally significant at +0.313 ($p \approx 0.098$). This attenuation suggests some clustering of connected individuals and accidents across police areas. Nevertheless, the point estimates remain positive, and the inclusion of fixed effects notably improves model fit (pseudo R^2 increases from 0.0019 to 0.0376), indicating that area-level context matters, but does not fully explain the observed disparities.

The persistence of effects after controlling for location suggests that individual-level mechanisms, such as behavioral risk-taking or implicit enforcement leniency, remains at play. Notably, accidents are the only outcome for which the family flag remains statistically significant (albeit marginally) after including PD fixed effects.

H2: Violation detection, discretion at the point of contact
Hypothesis H2 examines whether connected individuals are less likely to have violations formally recorded, due to informal immunity or discretionary leniency in the detection stage of enforcement. Appendix C, Table 20 reports the full regression results, using *had_violation* as the outcome. The results closely resemble those from H0. In the bivariate models, ministry officials are 49.3% less likely to have a recorded violation (−0.679, SE = 0.061), while family members show a 39.2% reduction (−0.498, SE = 0.106).

Adding age and gender controls narrows these effects slightly, consistent with the demographic profile of connected individuals, namely, they are older and more often female. Both factors are associated with fewer recorded violations. In the full model with police district fixed effects, the ministry coefficient remains negative and significant (−0.292, SE = 0.061, $p < 0.01$), but the family effect drops to −0.074 and loses significance. As with the fine receipt models, this attenuation likely reflects a combination of true district-level clustering of connected individuals and limited power to detect within-district variation for small subgroups. The improved model fit (pseudo R^2 increases) confirms the importance of accounting for location, but even within the same districts, officials remain significantly less

likely to have violations on record.

H3: Sanctioning discretion, fines conditional on violation

Hypothesis H3 tests whether connected drivers are less likely to receive a fine after a recorded violation. This isolates the sanctioning stage of enforcement, examining whether leniency persists even once misconduct is formally documented. The sample is limited to drivers who committed at least one violation (N = 150,395), which reduces statistical power, especially for connected subgroups.

Regression results (see Appendix C, Table 21) provide partial support. In baseline models, ministry officials are significantly less likely to be fined (log-odds = −0.431, SE = 0.125), and the effect remains after adding gender and age controls (−0.389, SE = 0.125), suggesting roughly 35–40% lower odds of sanctioning. However, once police district fixed effects are included, the effect becomes statistically insignificant (−0.103, SE = 0.128), indicating that much of the observed variation is explained by location. The model's pseudo R^2 increases from 0.0012 to 0.0691, reflecting this improvement in explanatory power.

For family-connected drivers, no significant effect is detected in any specification. The initial fine rate gap (70.3% vs. 75.0%) corresponds to a small and unstable coefficient (−0.234, SE = 0.227), which fluctuates and even turns positive in the fixed effects model (+0.197, SE = 0.244). While this reversal is likely not substantive, it illustrates how sparse subgroup data can lead to unstable estimates. More importantly, it exemplifies a broader interpretive tension that recurs across hypotheses: once police district fixed effects are introduced, many effects weaken or disappear, but this does not necessarily imply that substantive disparities are absent. Rather, the fixed effects absorb a large portion of the variation, especially when groups are unevenly distributed across districts, making it difficult to disentangle genuine within-district equality from a simple loss of power. In that sense, this reversal serves as a useful caution when interpreting effect attenuation after strong controls are applied. Further research using larger and more regionally diverse samples of connected individuals would be needed to resolve this ambiguity.

Notably, the gender effect also weakens in this model. While earlier results showed that women were substantially less likely to be fined, the gender coefficient here drops to −0.046 and is only marginally significant ($p < 0.1$) after including district controls. This suggests that some gender disparities reflect regional enforcement norms rather than uniform individual-level treatment. Unlike political connectedness, gender is widely distributed across regions, so this attenuation likely reflects real contextual variation rather than statistical noise. This will be further explored in Chapter 6.

In sum, H3 receives mixed support. Sanctioning leniency appears to apply to ministry officials in simpler models, but not once regional variation is accounted for. For family members, the evidence remains weak throughout. Whether this reflects a genuine absence of discretionary leniency at the sanctioning stage, or a lack of statistical power due to a smaller and more clustered sample, remains an open question.

H4: License deprivation, judicial leniency or external pressure

Hypothesis H4 concerns the most severe enforcement outcome in the data: license deprivation (i.e., suspension or revocation of one's driving license). This typically results from repeated or egregious violations and is often imposed at a later adjudication stage, such as through judicial or administrative review. As such, it provides a test of preferential treatment further downstream, beyond initial enforcement or roadside discretion.

H4 posits that connected individuals are less likely to lose their driver's license, even when committing sanction-worthy offenses. This outcome may result from various forms of institutional leniency, including informal influence, favorable treatment in court, or administrative "pulling of strings" that shields high-status individuals from formal penalties.

However, license deprivation is an extremely rare event in the data: only 0.074% of unconnected drivers, 0.034% of ministry officials, and 0.021% of family members experienced it. This rarity imposes clear limits on statistical inference.

In the descriptive analysis, Fisher's exact tests confirmed significant differences in deprivation rates between connected and unconnected drivers ($p < 0.01$ for both L0 and L1), consistent with H4. But Fisher's test is specifically designed for sparse binary data and tests raw contingency tables. Logistic regression, by contrast, relies on large-sample assumptions, and the standard errors become unreliable when event frequencies are extremely low.

Still, logistic regressions were estimated (Appendix C, Table 23), mirroring the models used for prior hypotheses. In all specifications, the direction of the coefficients is negative, meaning connected drivers are less likely to lose their license. But most estimates are not statistically significant.

For ministry officials, the bivariate model yields a negative and marginally significant effect (−0.774, SE = 0.409, $p < 0.1$). This effect shrinks and loses significance with the addition of controls (−0.640) and becomes very small under district fixed effects (−0.201, SE = 0.410). For family members, the coefficients are large and negative throughout (e.g., −1.243), but consistently non-significant, and fluctuate across models.

As with previous outcomes, the loss of significance for family members (and the diminishing estimate for officials) likely reflects a combination of low statistical power, rare events, and the burden of high-dimensional fixed effects. The inclusion of 77 district dummies absorbs considerable variance, and the pseudo R^2 jumps from 0.0024 to 0.1110, confirming that location plays a substantial role.

In sum, the results for H4 are directionally consistent with preferential treatment, but the logistic models lack the precision needed for definitive conclusions[10]. The most suggestive evidence comes from the raw event rates, which do indicate a meaningful disparity. Whether this reflects systematic bias in adjudication or simply the difficulty of detecting rare effects in small samples remains an open question.

[10] Firth's Penalized Logistic Regression is often recommended for rare event data, especially when one group has few or no occurrences. It corrects for small-sample bias and reduces the risk of separation, yielding more reliable estimates than standard maximum likelihood methods (Heinze & Schemper, 2002). Given the extremely low incidence of license deprivation among ministry employees and their family members, future work could benefit from applying this approach to obtain valid inference without inflated or unstable coefficients.

5.5 Role of informal signals: luxury & vanity

Beyond formal political connections, some drivers may display visible cues of status or influence that could shape how traffic enforcement unfolds. This section explores the role of such informal indicators, namely luxury vehicles and vanity license plates. Luxury cars, such as high-end German sedans or sports models, are widely recognized as signals of wealth and prestige. In Russia, they are often associated with powerful individuals or business elites, and may prompt officers to exercise restraint, whether due to fear of reprisal or implicit social pressure to defer to elites. The luxury vehicle indicator flags cars from premium brands like Mercedes, Lexus, and Porsche (see Appendix B, Table 12).

Vanity plates can also reflect privilege, as they are costly and impossible to obtain through official channels. These plates, often featuring stylized number sequences like 777 or 505, are flagged using the definition from Eeckhout et al. (2021) and detailed in Appendix B, Table 13. While these plates may indicate social status, they can also signal a willingness to engage in informal transactions, including bribery, given that their acquisition often relies on personal connections or under-the-table payments. As such, they may influence enforcement either positively or negatively, depending on the officer's interpretation.

Hypothesis 5 (H5) tests whether these informal markers are associated with differences in enforcement outcomes. Hypothesis 6 (H6) investigates whether the estimated effect of political connection attenuates after adding luxury and vanity indicators to the model. While no formal mediation test is conducted, a substantial drop in the coefficient would suggest that these informal signals account for part of the observed disparity.

H5: Informal signals and enforcement outcomes

Appendix C, Table 16 and 17 present descriptive comparisons of enforcement outcomes across individuals with and without luxury vehicles or vanity plates. The results show notable and statistically significant differences. For example, luxury vehicle owners are fined less often (–22.2%), cited less frequently (–18.4%), and have fewer license deprivations (–57.4%) than non-luxury drivers, all with $p < 0.001$. Vanity plate holders display similar patterns, albeit with smaller effect sizes and weaker significance. However, both groups exhibit higher accident rates, again echoing the pattern observed for connected drivers. This repeating pattern motivates the development of a behavioral model to explain the observed adaptation in driver behavior (see Section 6.5, future research).

These differences suggest that informal signals may influence enforcement behavior to some extent. However, their practical impact is less pronounced than that of formal political connections, and it remains unclear whether the effects reflect independent privilege or overlap with broader status cues, such as institutional position or influence. While H5 is supported descriptively, the observed disparities are more modest in size and interpretation.

H6: Attenuation by informal signals

To assess whether luxury and vanity indicators explain part of the advantage enjoyed by politically connected individuals, nested logit models were estimated. The key question is whether the coefficient for political connection becomes smaller when these informal status indicators are added.

As shown in Appendix C, Tables 24-33, including luxury and vanity indicators results in only marginal changes to the estimated connection effects. For instance, in the fine regression, the coefficient for ministry officials shifts from –0.8027 to –0.7959, while for family members, it changes from –0.2545 to –0.2502. Comparable minor reductions appear across other enforcement outcomes. These shifts are minimal and nowhere near enough to suggest

substantial attenuation, let alone mediation. No formal test was performed, as the coefficient changes were too small to warrant it. This also implies that the indicators operate largely independently, rather than capturing the same underlying source of privilege.

This pattern indicates that visible signals of privilege (luxury cars and vanity plates) account for only a very small part of the enforcement disparity. The core finding persists: political connection itself, rather than mere markers of status or wealth, is the dominant factor driving differential enforcement. Thus, while H6 is directionally supported, the attenuation effect is negligible.

Each indicator (political connection, vehicle luxury, and vanity plates) captures a distinct dimension of privilege. Together, they reveal how both political connection and informal signals can influence enforcement. But taken individually and without controls, these indicators explain very little of the observed variation. Across all specifications, pseudo R^2 values remain low, ranging from just 0.0005 to 0.0029.

And that's probably a good thing. If just three visible markers of power were enough to predict who gets sanctioned and who doesn't, we would be facing a very different and far more troubling story. Fortunately, most of the variance in enforcement outcomes lies elsewhere. This limited explanatory power invites a broader reflection. In the next chapter, I turn to a synthesis of the findings and their implications: what do these patterns tell us about the functioning of traffic enforcement in Russia, and how might they relate to the broader theme of institutional inequality?

6 DISCUSSION

This chapter moves beyond statistical results to explore what they mean: how political connections shape traffic enforcement in Russia. Not as isolated advantages, but as a systemic pattern of informal immunity. To return to the central research question: ***yes political connections, beyond wealth, drive immunity in traffic enforcement in Russia.*** Luxury vehicles and vanity plates, examined under H5 and H6 as proxies for wealth or informal status, were also found to influence enforcement outcomes independently. While they exhibit similar patterns of leniency, these effects are largely distinct from those associated with political connections, with only minimal overlap. For this reason, the focus of this chapter centers on H0 through H4.

6.1 Enforcement timeline: where privilege bends the rules

While Chapter 5 presented these results in statistical detail, the goal here is to interpret them thematically (stage by stage) to understand how privilege operates not just as an outcome, but as a structured process.

As outlined in Chapter 4.2.3, enforcement in Russia follows a multi-stage sequence, from risky behavior to judicial sanctioning. Each stage (detection, sanctioning, adjudication) contains opportunities for discretion, and this discussion revisits those stages through the lens of the empirical findings. Rather than beginning with behavior, this section starts where discretion first leaves a data trace: at the moment of detection. Driving behavior is a logical starting point, since riskier or rule-breaking behavior is what ultimately leads to infractions and penalties.

However, this framing is also incomplete: in a world where enforcement is selective, behavior can itself be shaped by expectations of leniency. If drivers are aware that they are less likely to be stopped or sanctioned by police officers, they may adjust their behavior accordingly, taking more risks behind the wheel. In that sense, behavior can also be seen as a downstream outcome of enforcement practices. This circular dynamic, where behavior drives outcomes, but outcomes also feed back into behavior, makes causal direction tricky to pin down. For this reason, the discussion here begins not with driver intent but with detection, the first point at which discretion becomes visible in the administrative record. From there, we move through sanctioning, adjudication, and finally to accidents, which offer a revealing glimpse into real-world consequences.

6.1.1 Detection

Are violations recorded?

The data indicate that politically connected drivers are significantly less likely to be cited for traffic infractions. They appear roughly half as likely to be officially written up for violations compared to unconnected drivers. This gap highlights the role of early-stage discretion: officers may choose not to stop or report an influential driver, or may issue informal warnings instead of formal citations.

Yet this gap is especially striking when paired with what we know about accidents. Accidents are difficult to conceal: they result in physical damage, insurance reports, or hospital records. In contrast, violations only appear in the data if an officer chooses to record them. So when connected drivers appear more frequently in accident records, but less frequently in violation data, it suggests that infractions may be occurring but are simply not being officially logged. In that sense, the data reveal a detection or registration gap: risky behavior

goes unrecorded, reinforcing the idea of discretion not just in enforcement, but in data entry itself.

The fact that the reduction in fines exceeds the reduction in violations (a 54.7% drop in fine rates vs a 48.5% drop in violations for officials) hints at an extra layer of leniency beyond just driving behavior alone. In short, many infractions by the well-connected likely go unrecorded, insulating them right at the point of contact.

6.1.2 Sanctioning

How are fines distributed between different groups?
In terms of receiving fines, connected drivers enjoy a substantial advantage. Overall fine rates are 40–55% lower for officials and their family members, compared to similar unconnected drivers. Alongside the violations, this represents one of the most robust effects extracted from the dataset. Even after controlling for age, gender, and location, a substantial and significant reduction in fine likelihood remained. For family members, the effect weakens after controlling for local enforcement patterns.

Most of this observed fine disparity stems from the earlier stage of enforcement: connected drivers are cited for violations less frequently. However, as explored in H3, a second layer of discration appears to exist after a violation has already been recorded.

Are penalties imposed after a violation?
Even when a violation is officially documented, connected drivers often benefit from leniency in the consequences. Among those cited for a violation, a smaller share of ministry officials actually received fines compared to regular violators (about 66% versus 75%, respectively). This gap suggests that discretion is exercised not only in whether to document a violation, but also in whether to impose a penalty once misconduct is recorded. Whether through officer judgment, administrative leniency, or informal influence, the sanctioning stage appears to offer yet another layer of insulation for the politically connected.

For family members, the conditional effect was smaller and not statistically significant. This likely reflects both the small number of connected offenders in the sample and potential noise in family identification. Moreover, even the effect for officials lost significance when police district fixed effects are applied, highlighting how statistical significance can be sensitive to geographic clustering and sample size.

Among all stages, this was one of the more difficult effects to detect statistically. Still, the pattern fits the broader picture: if a connected individual is unlucky enough to be cited, they are still less likely to feel the full sting of enforcement. That, too, is a meaningful form of leniency. Just one that is harder to detect in the available dataset.

6.1.3 Judicial Adjudication

Do the harshest legal consequences apply equally?
Although instances are rare, the data suggest judicial leniency for high-status individuals when it comes to the most severe outcomes. Only 0.034% of officials and 0.021% of family members lost their licenses, compared to 0.074% of unconnected drivers. While these absolute rates are low, the relative reductions for connected individuals range from 54% to 71%. These gaps are statistically significant in descriptive terms, though interpretation of the regression results is limited due to low event counts.

Nevertheless, the practical significance remains clear: the observed disparity suggests that elites may face a fundamentally different system of accountability. This interpretation aligns with widely publicized incidents in Russia, where officials' relatives received remarkably light punishment for fatal or reckless traffic violations. One case involved the son of a high-ranking official who avoided serous sanctions after a deadly crash (Kravtsova, 2013). This reinforces perceptions that the law is selectively enforced. Similarly, another case described how an official's family member involved in a fatal collision was shielded from prosecution, highlighting the broader dysfunction of a system where status can outweigh culpability (Franchetti, 2008).

Such examples underscore the broader concern that informal privilege extends even to life-altering sanctions, undermining both deterrence and public trust in legal institutions.

6.1.4 Accidents

Do connected drivers behave more safely, or are they simply caught less?
At this point in the analysis, the question almost answers itself.

If leniency in enforcement were truly reflecting safer behavior, we would expect connected drivers to have fewer accidents than others or at the very least, similar rates. After all, risky driving behavior is closely linked to accident involvement. But the data tell a different story: connected drivers are more likely to be involved in accidents than unconnected drivers. While absolute rates are low (under 1%), the relative differences are striking. Ministry officials are about 38% more likely to appear in an accident record. For their family members, that number rises to over 60%.

Accidents are the only outcome where the family flag remains marginally significant, even after controlling for geography. This raises the possibility (albeit tentatively) that family members may engage in particularly risky driving. Perhaps because they feel less constrained than officials due to lower visibility or weaker institutional oversight. Still, due to sample size limitations and overlapping confidence intervals, I refrain from making strong claims about within-group behavioral differences.

These elevated accident rates are especially revealing when contrasted with earlier findings: the same individuals who are less likely to be fined are more likely to crash. This contradiction casts serious doubt on the idea that connected drivers are simply safer. Instead, it reveals a gap between actual risk and recorded enforcement.

One way to interpret this is in proportional terms. If connected drivers are 60% more likely to be in accidents, and if accidents correlate with rule-breaking, then under proportional enforcement, they should also be 60% more likely to be fined. But instead, we find that they are less likely to be flagged or fined. This mismatch suggests that many of their risky or rule-breaking behaviors go unrecorded, either because they aren't stopped, or because the event isn't logged as a formal infraction.

This leads to an important distinction: while connected drivers appear to break fewer rules on paper, their elevated accident rate implies the opposite. However, we cannot directly observe behavior in this dataset, only enforcement and outcomes. So while it may be tempting to say that these individuals "break more rules," the data only allow us to say that they are involved in more accidents while being penalized less, a pattern consistent with selective enforcement and potential moral hazard. We observe more real-world harm but less official punishment.

This discrepancy also underscores the societal cost of discretionary enforcement. These types of corruption I have been analyzing may seem relatively harmless, so long as no one gets hurt. But once people start ending up in accidents because of it, it's no longer a small deal. Why have traffic laws at that point? If elites are free to drive however they want, while others should stay off the streets to survive, the system resembles the Wild West more closely than it would a supposed legal system. While this framing may seem exaggerated, the sentiment captures the core issue: when enforcement becomes optional for some, traffic laws lose their legitimacy for all.

The risks posed by elite impunity on the roads are not merely theoretical. In 2016, the Kremlin publicly criticized the dangerous driving habits of well-connected youths, after viral videos showed luxury cars street racing through Moscow. These incidents, involving the children of high-ranking officials, sparked widespread backlash. Kremlin spokesman Dmitry Peskov acknowledged the growing public anger, calling the behavior unacceptable and warning that such displays damage the image of the state (Tétrault-Farber, 2016). However, there were very few meaningful sanctions. This reinforces the perception that elite drivers can act with near-total immunity. This political sensitivity highlights the broader implication of the findings: traffic enforcement is not just a technical issue, it's a visible arena where social inequality becomes materially dangerous. When laws are applied differently based on who you are, the risk is not just more crashes, but also public disillusionment with the rule of law itself, potentially leading to vicous circles fueling even worse behavior.

6.1.5 Data quality caveats

Do events get logged correctly?
Even the final, seemingly administrative step of enforcement (i.e. recording infractions and penalties) may involve informal discretion. If violations are selectively omitted or modified, this discretion would be invisible in the data. Since the analysis relies entirely on officially logged events, any such filtering would bias the results in unknowable ways.

This is particularly important for rare outcomes or small subgroups, where a few missing records can distort conclusions. Because administrative data only capture what is recorded, not what actually occurred, registration itself marks a final but unobservable frontier of enforcement discretion. Future research should treat these boundaries of observability with caution.

6.1.6 The layered nature of leniency

Taken together, these stages reveal a layered system of discretion: at each point in the enforcement process (detection, sanctioning, and adjudication) there are opportunities for those enforcing the law to cut slack to the well-connected. A politically connected driver benefits from multiple "safety valves." They might avoid being written up at all when pulled over. If written up, they might receive a warning or a lower fine. And if they would ordinarily lose their license, that ultimate penalty may be quietly averted. Each layer by itself may only tilt outcomes a bit, but cumulatively they provide a form of immunity. This layered leniency is vividly illustrated by the end result: these drivers rarely face serious repercussions, yet the toll shows up in higher accident rates, implying that the normal feedback loop of deterrence is broken. When consequences are repeatedly defanged, bad driving habits come home to roost in real-world harm.

6.1.7 The role of aerial discretion

One final layer in the enforcement puzzle is geography. Enforcement practices differ sharply across Moscow's 77 grouped police districts, which is why district-level fixed effects were included. These help isolate individual privilege from broader patterns of local leniency.

The results reveal an important nuance: ministry officials still receive significantly fewer fines than others in the same district, but the advantage for their relatives often fades once these controls are applied. In plain terms: officials get fewer fines than their neighbors, relatives may not. This suggests that favoritism toward officials operates not just across districts, but within them. For family members, apparent advantages may reflect regional clustering or limited data, rather than robust within-district privilege.

Controlling for police districts thus shows that both location and status matter: where you drive shapes your exposure to enforcement, but who you are still makes a difference, even on the same streets.

6.2 Interpretation of control variables

Beyond political ties, the analysis also accounted for basic driver characteristics like gender and age, both to control for their influence on driving outcomes and to explore how they relate to enforcement disparities. These traits proved relevant and raise an imprtant question: do observed differences (for example, between male and female drivers) reflect genuine behavioral differences, or unequal treatment under the law?

6.2.1 Gender and detection-stage discretion

Consistent with prior research and administrative statistics, men are significantly overrepresented among serious traffic violators in Russia. One recent study analyzing ~160,000 court decisions (2010–2022) found that over 90% of convicted traffic offenders were male (Kazun, 2024). On average, men drive more and engage in riskier behaviors, such as speeding and aggressive maneuvers, than women. Adjusted for exposure, Russian men were found to be approximately 3.3 times more likely to be involved in accidents. This aligns with other findings: Bruckheim & Patel (2022) report that men receive over 70% of all traffic violations.

The data in this thesis echoed these patterns.[11] Being male was consistently associated with a higher likelihood of accidents, while age (used as a proxy for experience) was negatively associated with risk. A substantial part of the observed gender gap in traffic outcomes thus appears to be behavioral: male drivers, on average, engage in riskier driving, which leads to more infractions and accidents.

However, behavior may not fully explain these differences. Enforcement bias remains a possibility. For instance, a Kansas State University study (2012) found that women were 23% less likely to receive tickets during traffic stops. It is sometimes suggested that officers may exercise more leniency toward female drivers in discretionary encounters, while male drivers may be subject to stricter judgment.

[11] In this study, being female was associated with substantially lower odds of receiving fines (−76.6%), being cited for violations (−75.3%), or losing a license (−89.9%). The effects were smaller for conditional outcomes like fine given violation (−16.5%) and accident involvement (−4.9%), but still directionally consistent.

The fact that women in this study consistently exhibit lower violation, sanction, and accident rates cannot be linked solely to officers going easier on them or women showing greater law-abiding behavior. It is likely the result of a combination of factors, including driving more cautiously as well as instances of differential treatment contributing to the gap.

An especially noteworthy result emerges when looking at whether a fine is issued after a violation is already on record. For most models in our dataset, gender was a strong and significant predictor: being female consistently reduced the likelihood of being fined or cited. However, for the outcome *fine given violation*, this effect weakens considerably. In fact, it is the only model where the gender coefficient becomes only marginally significant (ministry p = 0.0524, family p = 0.0514, compared to 0.000 for all others). This occurs despite being based on still a large sample of over 150,000 violators.

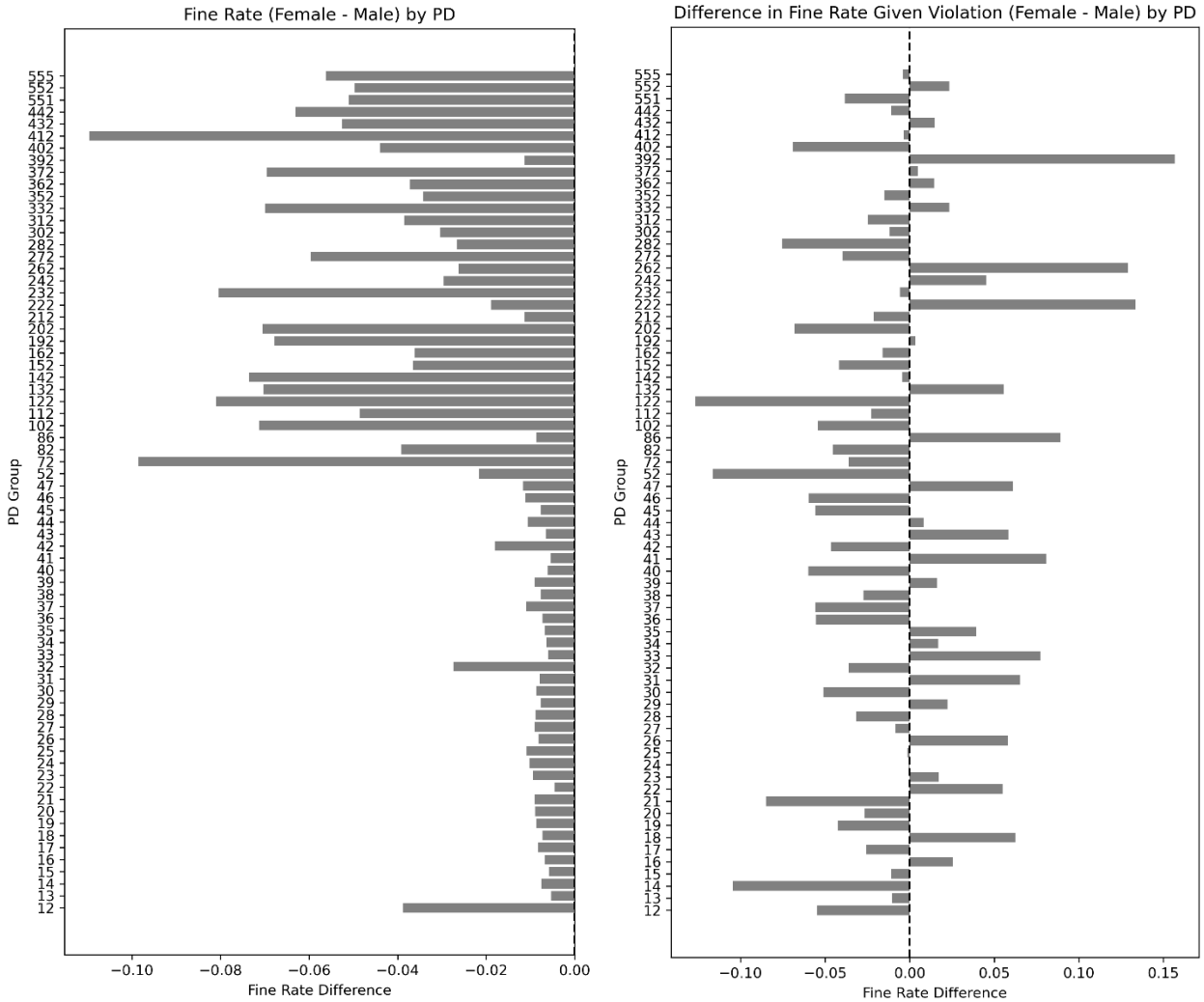


Figure 19: Gender differences (F-M) in (conditional) fine rates across police districts

While the *fine given violation* sample contains fewer observations and is more male-dominated than the full dataset, the reduction in the gender effect is not merely due to limited statistical power. The coefficient for being female shrinks substantially in magnitude (from −1.46 to −0.19), even though the standard error remains low (0.0227). This indicates that the relationship between gender and fines fundamentally differs once violations are held constant. This can be visualised in Figure 19, which shows strong and consistent male-overrepresentation in overall fines, but heterogeneous and much weaker gender patterns in fines conditional on violations.

In short, men are more likely to be cited, but once a citation occurs, gender plays a much smaller role in whether a fine is imposed. This points to detection-stage discretion as the key mechanism behind observed gender disparities.

6.2.2 Age and experience

Age, like gender, plays an intuitive and consistent role in traffic outcomes. Younger drivers had higher rates of both violations and accidents, while older drivers tended to be more cautious. In the regression models, each additional year of age slightly reduced the likelihood of both being cited and being involved in a crash. This pattern is consistent with the expectation that experience and risk aversion increase with age.

Politically connected drivers, especially ministry officials, tended to be older than the general driving population. Many are mid-career or senior professionals. Their family members were younger on average but still older than the unconnected group. Appendix Figures 22 and 23 show these distributions, including a spike around age 38 caused by imputing missing birth years with the median (n = 711,430).

Interpreting the gaps

As summarized in Table 9 (Chapter 3), connected groups (L0 and L1) tend to combine characteristics commonly associated with more favorable enforcement outcomes. Ministry family members, in particular, appear the most "advantaged": they are more likely to be female, drive luxury vehicles (17.04%), and use vanity plates (6.46%) than both ministry officials and unconnected drivers. Ministry officials also exceed the unconnected group on nearly every metric, including average age (40.3 vs. 36.2), luxury car ownership (12.73% vs. 10.09%), and vanity plate use (5.85% vs. 5.19%). In contrast, unconnected drivers consistently show the least favorable profile across these traits.

When age and gender were included as control variables in our regressions, the estimated enforcement gaps for violations, fines, and license deprivation narrowed slightly, suggesting that some of the apparent leniency was simply due to demographic composition i.e., the fact that connected drivers are older and more likely to be female. However, these traits account for only a fraction of the disparity: the advantage remained large and statistically significant even after adjustment. In other words, a slice of the “privilege effect” was just an age/gender effect, but most of it was not.

Interestingly, for accidents, the opposite occurred. Since older drivers generally have fewer accidents, adjusting for age and gender actually made the disparity larger: the accident risk associated with political connection increased after controlling for demographics. This reinforces the idea that the elevated accident rates among connected individuals are not due to who they are (older, more experienced or female), but rather to how they behave on the road: possibly emboldened by perceived impunity.

In simpler terms: controlling for age and gender narrows the leniency gap in penalties, but widens the danger signal in accidents. This asymmetry adds further weight to the behavioral interpretation: connected individuals may take more risks not despite, but because of, the enforcement shield around them.

6.3 Implications of Nepotism/Institutional and sociopolitical implications

This thesis began with a broad concern: whether politically connected individuals in Russia enjoy unequal protection under traffic law, distinct from the effects of wealth alone. The study moved from this normative question through detailed data cleaning, sample construction, and empirical testing. Now, we return to reflect on what the findings reveal about nepotism as a broader social and institutional phenomenon.

The results suggest a clear pattern. Ministry officials are significantly less likely to be fined or have violations recorded, yet they appear more frequently in traffic accident records. Unlike violations or fines, which depend heavily on officer discretion, accidents typically involve physical damage or third-party reports, making them less susceptible to underreporting. The contrast between low citation rates and elevated accident involvement points to reduced accountability and a form of moral hazard: shielded from sanctions, these individuals may adopt riskier driving behavior. While similar advantages appear for family members, the evidence is weaker and more sensitive to context and statistical power. This underlines both the limits of the data and the more diffuse nature of familial favoritism.

By directly measuring unequal legal outcomes for politically connected individuals, this study expands the conversation around nepotism. It moves the focus beyond public sector hiring and into the realm of traffic law enforcement, showing how institutional and familial ties can shape who faces consequences (and who does not) on the road.

Embedded discretion in institutions

One striking aspect of the findings is how deeply preferential treatment appears to be embedded in the enforcement process. It is not a matter of a single officer informally letting someone off. The data suggest a systematic pattern of leniency for connected individuals, pointing to a broader institutional norm. Such patterns likely persist only if tolerated (or quietly encouraged) by higher-ups. In this context, it may be understood within enforcement bodies that “important people” should be handled more carefully, whether due to formal directives, fear of repercussions, or longstanding habit.

This kind of behavior is difficult to trace directly in the data. We cannot observe whether bribes were exchanged or if leniency stems purely from institutional deference to status. But the consistency of favorable treatment suggests that informal influence plays a structural role, regardless of its exact mechanism.

Discretion is inherent to traffic enforcement, officers have leeway in whom to stop, cite, or penalize. This flexibility, while necessary in many contexts, also creates space for unequal treatment to become normalized. Over time, such bias can embed itself so deeply that it is no longer even perceived as corruption, but simply as common sense or adherence to social hierarchy. This quiet erosion of legal equality poses a serious challenge to institutional integrity.

Public safety and moral hazard
The fact that connected drivers have more accidents is a real-world consequence that should concern policymakers. It implies that unequal enforcement is not just an ethical or legal issue, it has human costs. Traffic accidents can result in injuries or fatalities to pedestrians and other drivers, not just the offending party. So when a certain group is allowed to skirt the rules, it endangers the broader public.

This research suggests a classic moral hazard scenario: shielded from normal repercussions, these drivers take extra liberties on the road, leading to more crashes. Figure 3 (Chapter 2) illustrates this dynamic in simplified welfare-economic terms: privileged drivers internalize less of the social cost of their behavior, leading to greater private gain at the expense of public safety. Figure 20, shown in section 6.5, proposes a behavioral model to formalize how such enforcement asymmetries can shift driver behavior toward a new, riskier equilibrium.

This represents a lose-lose situation: privileged drivers face greater harm (despite temporary protection, they end up in more wrecks), and innocent bystanders are put at greater risk. From a policy perspective, this undercuts any claim that selective leniency has no victims. On the contrary, it likely increases the social costs (medical, insurance, even economic losses from accidents). It underscores that consistent enforcement is not merely a matter of fairness, it's essential for public safety.

More broadly, these findings emphasize that traffic enforcement is not only about road safety, it also reflects how power operates in everyday law. When accountability is selective, harm is not only unequally distributed, but also amplified.

Institutional trust and rule of law
At its core, the rule of law is meant to ensure governance through impartial and predictable legal standards. Yet our findings point to a departure from this ideal, what Angelis & Harrison (2003) describe as a collapse of "rule by law" into "rule by men". Politically connected individuals in Russia appear shielded from the legal consequences that apply to others. When social status overrides objective law, the principle of equality before the law collapses into selective privilege.

This matters beyond traffic violations. As emphasized in the *Rule of Law Index* by the World Justice Project (2024), the rule of law rests on four universal principles: accountability, just laws, open government, and accessible and impartial justice. These ideals are often experienced most directly through routine encounters with law enforcement. As Angelis and Harrison (2003) put it, "For most people, their primary interactions with the legal system are through the police, courts, and government agencies that enforce the law. These day-to-day encounters shape the public's understanding of what the rule of law means in practice." If drivers observe that elites receive fewer fines, more lenient treatment, or avoid accountability after accidents, public trust erodes. The system ceases to function as a neutral enforcer of rules and instead reinforces existing hierarchies.

As previously discussed, this aligns with the concept of *blat*: the informal exchange of favors embedded in post-Soviet institutional culture. While wealth can enable privilege through bribery, the patterns in this study persist even after controlling for vehicle type and vanity plates, which serve as proxies for economic status. Although the administrative data do not include income or capture bribes directly, the consistent advantage observed for ministry officials suggests a broader, institutionalized form of protection. Rather than stemming from financial leverage alone, legal leniency appears to flow from personal and bureaucratic connections. Law still applies, but not uniformly.

This pattern reflects more than individual anomalies, it signals a systemic issue with rule of law. Russia continues to rank low on international indicators of legal equality and corruption control (see Figure 1 in Chapter 2), with nepotism and elite impunity remaining persistent challenges. If traffic enforcement, arguably one of the most routine forms of legal application, cannot escape these influences, it raises broader concerns about institutional integrity across other domains.

This erosion of equal treatment under the law carries multiple consequences. It undermines impartial enforcement: when a minister's son and a taxi driver commit the same offense, chances are that only the latter reliably faces sanction. It imposes social and moral costs by enabling riskier behavior among the privileged, who as a result contribute disproportionately to accident rates. It also diminishes public trust in legal institutions, weakening voluntary compliance. And it raises the bar for reform, since addressing entrenched privilege requires more than legal measures. It demands deeper institutional and cultural change, as well as technical mechanisms that reduce discretion.

6.4 Limitations of the study

While this study provides novel insights into discretionary enforcement and political privilege in Russian traffic policing, several important limitations must be acknowledged. Some of these were already mentioned, particularly under data quality caveats (Section 6.1), but are consolidated here for clarity.

1. Data provenance and observability constraints

The dataset used in this analysis is not officially sanctioned and was obtained through non-transparent channels. As a result, it is impossible to validate the integrity of data collection or ensure that it comprehensively captures all registered drivers and enforcement actions. There is no external benchmark to assess registration completeness, accuracy, or possible selection biases in who is recorded. This limitation pervades all inferences: we can only study what is visible, and the structure of what remains unrecorded is unknown.

2. Cross-sectional design and lack of temporal dynamics

The analysis is cross-sectional in nature, providing a snapshot of enforcement patterns over a defined period. This limits the ability to capture dynamic effects or behavioral adaptation over time. For example, we cannot observe whether enforcement changes after a driver gains or loses political connection, or whether risk-taking escalates following repeated leniency. A panel dataset would allow for much richer longitudinal analysis, including before/after comparisons and individual-level fixed effects.

3. Limitations in identifying family ties and political rank

Family members are identified based on shared surnames, address co-residence, and vehicle linkages. This methodology captures active drivers linked to officials but it fails to identify non-driving family members and more distant kin. Furthermore, there is also the incompleteness of address records, which weakens co-residence matching. The sample size of identified family-connected drivers is modest, which reduces statistical power in some tests and prevents detailed subgroup analysis (spouses vs. children).

The identification of politically connected individuals is inherently constrained. Ministry officials are grouped together as a single category, without distinction between ranks, agencies, or positions of influence. This aggregation may mask important heterogeneity as some ministries may command more deference than others.

4. Measurement proxies and spatial imprecision

Location is proxied using police district (PD) codes, which are primarely based on vehicle registration rather than actual driving exposure. To approximate actual exposure, I assigned each car–person combination to its most common observed PD (the mode). While this helps, it still introduces potential mismatch: registration location may not fully reflect real driving behavior or exposure to local enforcement.

Moreover, if politically connected individuals tend to register their vehicles in districts with laxer enforcement, either strategically or due to institutional privilege, then controlling for PD risks absorbing some of the very favoritism under investigation. In other words, part of the connection effect may be "controlled away" by fixed effects if geographic privilege overlaps with political privilege. PD fixed effects thus reduce certain spatial confounds but remain a blunt instrument. They help correct for regional variation, but at the same time may also obscure how institutional leniency is linked to location.

5. Ambiguity in outcome interpretation

Some key outcomes, such as *fine given violation*, are difficult to interpret precisely. This variable may reflect differences in how officers respond once a violation is recorded (e.g., warnings vs. fines), or differences in what gets logged in the system post-facto (e.g., successful appeals, informal cancellations). The data do not allow us to distinguish between being let off at the scene and having a fine administratively waived at a later stage in the administrative process. This ambiguity affects how strongly we can interpret sanctioning gaps as active leniency.

6. Cleaning and aggregation tradeoffs

The cleaning and merging of administrative datasets involved a range of assumptions and imputations, particularly for missing or inconsistent address fields, birthdates, and INN identifiers. While efforts were made to transparently document and cross-check key steps, the final dataset inevitably reflects these decisions and assumptions. For example, co-residence was inferred from standardized address strings, many of which were incomplete or variably formatted. Conservative heuristics were applied to reduce false positives in family linkage, with an emphasis on ensuring reliability over completeness. As a result, some genuine family ties were likely missed due to strict matching criteria. Similarly, inconsistent INN codes across records required simplified matching rules to identify ministry employees, again prioritizing precision over recall. Birthdates were imputed using median values within vehicle-linked profiles, a method that preserved data volume but may introduce bias if the missingness was non-random. These tradeoffs reflect the tension between retaining valuable observations and minimizing noise, and they underscore the caution needed in interpreting fine-grained patterns.

7. Age of the dataset

Finally, the dataset is over two decades old. While many of the structural patterns in Russian governance and enforcement may persist, the age of the data poses questions about current relevance. Enforcement technologies (e.g., automated systems), public attitudes, and institutional norms may have evolved since the time of observation. This study should therefore be read as a historical window into systemic privilege rather than a diagnosis of present-day enforcement.

6.5 Direction for future research

This study opens up several promising avenues for future research, particularly on the behavioral mechanisms that may underpin observed enforcement disparities. While the analysis provides strong statistical evidence of differential treatment for connected individuals, it cannot fully capture the decision-making processes or enforcement dynamics that lead to such outcomes.

One direction for future research, is to develop a formal behavioral model of driver decision-making under conditions of unequal enforcement risk. From a micro-economic perspective, each driver weighs the benefits of risky driving (e.g., time saved, utility of arriving earlier) against the expected cost of being sanctioned, calculated as the probability of detection multiplied by the severity of the fine. For politically connected individuals, both components of this expected cost appear to be systematically lower: they are less likely to be detected, and even when detected, less likely to be fined. This effectively reduces the marginal cost of rule-breaking.

Under such conditions, rational drivers may adjust their behavior toward a new equilibrium, increasing their risk-taking until the perceived cost again reaches a tolerable level. This compensatory response could help explain why connected individuals appear disproportionately in accident records, despite receiving fewer official sanctions. The basic behavioral logic is visualized in simplified form in Figure 20.[12]

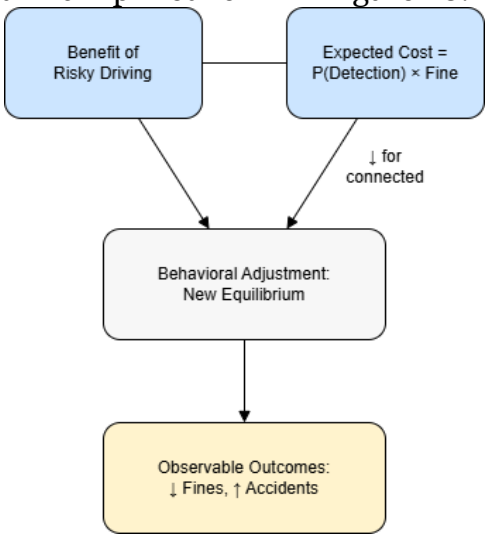


Figure 20:Behavioral model of risk and enforcement under privilege

[12] This schematic illustrates how drivers adjust their behavior in response to changes in the expected cost of rule-breaking. The model assumes that each driver faces a trade-off between the benefit of risky driving (e.g., arriving sooner) and the expected cost of being caught, calculated as the probability of detection multiplied by the severity of the fine. For politically connected individuals, both of these cost components are systematically

A model of this kind would allow researchers to simulate counterfactual enforcement regimes (e.g., what would accident rates look like if connected drivers faced equal sanction probabilities? How large is the behavioral offset between reduced fines and increased accidents?) To answer these questions credibly, more granular data would be needed: ideally, panel data tracking individual drivers over time, or natural experiments involving changes in enforcement policy or visibility.

Future work could also explore how connected individuals weigh external costs (like fines) versus private costs (like crash damage or injury). If fines are more easily dismissed or avoided by the well-connected, but private costs are not, then they may not adjust their behavior fully, resulting in the partial behavioral compensation observed in this study: a drop in fines, but only a moderate increase in accidents. This mismatch suggests incomplete risk internalization and points toward moral hazard under unequal legal deterrence. Richer administrative data (linking driving behavior to job position, wealth, or officer discretion) could help unpack the mechanisms behind this selective accountability.

Another avenue, already discussed in the limitations section, involves disaggregating political connectedness further. This study pooled all ministry officials into a single category for analytical tractability. However, future work could investigate whether certain ministries receive more lenient treatment than others (for instance, distinguishing between interior, defence, or finance ministries). This would reveal whether privilege is evenly distributed across the state apparatus or concentrated in certain institutions.

Similarly, expanding the scope of family connections could provide deeper insights. The current approach was limited to drivers identifiable in the vehicle registry. With more complete household or kinship data, future research could distinguish between spousal, parental, and extended family ties, and explore whether different types of familial proximity influence enforcement outcomes differently.

Finally, with longitudinal data, researchers could examine career-linked enforcement patterns. This could detect whether an individual's treatment changes after a relative enters or exits public office, or whether enforcement norms shift in response to broader institutional changes.

Taken together, these directions point to a broader research agenda that moves beyond surface-level disparities and aims to understand the mechanisms, networks, and incentives that sustain unequal enforcement in settings marked by institutional privilege.

lower. This reduction in enforcement risk shifts the individual's cost–benefit calculus, prompting a behavioral adjustment. To maintain a preferred equilibrium between risk and reward, connected drivers may compensate by engaging in riskier driving, thereby raising their true probability of being fined back toward a tolerable threshold. This behavioral offset helps explain why connected individuals are found to be involved in more accidents, even though they face fewer recorded sanctions. The bottom panel in the figure reflects the observable outcome of this adjustment: a drop in fines, paired with a rise in accident involvement, consistent with moral hazard under uneven enforcement.

7 CONCLUSION

This thesis set out to answer a clear and pressing question: ***do politically connected individuals in Russia receive preferential treatment under traffic law, distinct from the effects of wealth?*** Using administrative data on traffic outcomes, residency, and state employment, I constructed a cleaned sample capable of isolating enforcement patterns for individuals linked to government ministries and their families.

Despite the complexity and messiness of the original datasets, extensive cleaning, matching, and variable construction allowed for robust statistical testing. The results consistently pointed in the hypothesized directions: politically connected individuals, particularly ministry officials, were significantly less likely to receive fines or recorded violations, yet appeared more frequently in accident records. These findings suggest a consistent pattern of reduced accountability and a form of moral hazard, whereby leniency encourages riskier behavior.

Crucially, these effects persisted even after controlling for visible wealth indicators, such as luxury vehicles and vanity plates. This reinforces the central claim that familial and institutional ties to the state, not just economic status, shape who gets punished and who gets protected on Russian roads.

Summary of hypotheses

- Core hypothesis (H0): Politically connected individuals were less likely to receive fines, a result strongly supported by the data.

- H1: These individuals were more often involved in accidents, indicating behavioral adaptation consistent with moral hazard.

- H2: They were cited for fewer violations, suggesting discretion in detection.

- H3: Conditional on a violation, connected individuals were slightly less likely to be fined, but this effect weakened under fixed effects.

- H4: Descriptive statistics suggested lower rates of license deprivation among connected individuals. However, in regression models with fixed effects, these differences were no longer statistically significant. While this hints at judicial leniency, the evidence remains inconclusive, and future work could apply more suitable models (e.g., Firth logit) to better address rare event bias.

- H5: Informal signals of privilege (vanity plates, luxury cars) mirrored these patterns, receiving fewer sanctions and more accidents.

- H6: While informal signals of status correlate with enforcement gaps, they do not explain away the advantages conferred by political connectedness, pointing to separate forms of privilege.

Taken together, these findings demonstrate that ***political nepotism, not merely wealth, contributes to unequal treatment in traffic law enforcement.***

Contribution
This thesis advances the study of nepotism by shifting the focus from labor markets to law enforcement as an underexplored yet critical domain. While previous work like Szakonyi (2019) documented the economic returns to familial ties among public officials, this study reveals how those same ties distort legal accountability.

Szakonyi's approach provided a rigorous blueprint for identifying nepotistic advantages using administrative data, while Eeckhout et al. (2021) demonstrated that vanity plate prevalence correlates with fewer citations and more accidents, interpreting this as a symptom of regional enforcement bias and informal institutional norms.

This thesis extends that logic to the domain of political connectedness, where privilege manifests not through formal job appointments but through discretionary enforcement outcomes. While no formal difference-in-differences strategy was used, the identification approach remains rooted in isolating treatment effects between connected and non-connected individuals, controlling for location, demographics, and informal wealth indicators like luxury vehicle ownership and plate prestige.

By doing so, it shows that legal inequality is not limited to elite hiring practices, procurement decisions, nor to the institutional bias associated with high-status indicators such as vanity plates and luxury vehicles. It filters into the everyday state-citizen interface, where a police officer decides whom to stop, fine, or overlook. This reinforces the need to see nepotism not merely as a human resource issue, but as a broader governance challenge with direct implications for institutional trust and public safety.

8 REFERENCE LIST

Aliyev, H. (2013). Post-communist informal networking: Blat in the South Caucasus. *Demokratizatsiya: The Journal of Post-Soviet Democratization, 21*(1), 89–114. https://demokratizatsiya.pub/archives/21_1_RW41W8336734714X.pdf

Angelis, T. J., & Harrison, J. H. (2003). *History and importance of the rule of law* [White paper]. World Justice Project. https://worldjusticeproject.org/our-work/publications/working-papers/history-and-importance-rule-law

Bank of Finland Institute for Economies in Transition. (2017, March 17). Russia's car ownership remains well below EU average. *BOFIT Weekly Review*, *2017*(11). https://www.bofit.fi/en/monitoring/weekly/2017/vw201711_2/

Bruckheim, M., & Patel, S. (2022, October 13). *Male vs. female driving statistics*. Bruckheim & Patel. https://www.brucklaw.com/male-vs-female-driving-statistics/

Byulegenova, B., Prasolov, V., Sheryazdanova, G., Bratanovsky, S., & Sabirova, L. (2024). Nepotism in post-Soviet states: The quantitative assessment on socio-demographic factors and the Corruption Perceptions Index. *Journal of Ethnic and Cultural Studies, 11*(1), 96–118. https://doi.org/10.29333/ejecs/1814

DeAngelo, G., & Hansen, B. (2014).
Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy, 6*(2), 231–257. https://doi.org/10.1257/pol.6.2.231

Eeckhout, T. (2017). *Do vanity plates pay off? A study of privilege and impunity in Russian traffic enforcement* (Master's thesis). Ghent University. https://libstore.ugent.be/fulltxt/RUG01/002/351/271/RUG01-002351271_2017_0001_AC.pdf

Eeckhout, T., Natkhov, T., Polishchuk, L., Schoors, K. J. L., & Hoefman, K. (2021). *Statistical footprints of corruption: 'Vanity Fair' of automobile license plates in Russia*. SSRN. https://doi.org/10.2139/ssrn.3973625

Emirates Auction. (2008, February 16). *License plate "1" sold for 14.5 million US $*. https://www.emiratesauction.com/media-center/news/148/license-plate-"1"-sold-for-14.5-million-us-$

Enikolopov, R., Petrova, M., & Sonin, K. (2018). Social media and corruption. *American Economic Journal: Applied Economics, 10*(1), 150–174. https://doi.org/10.1257/app.20160089

Evans, A. B. (2018). *Protests in Russia: The Example of the Blue Buckets Society*. *Demokratizatsiya: The Journal of Post-Soviet Democratization, 26*(1), 3-24. Retrieved fromhttps://www.researchgate.net/publication/324692659_Protests_in_Russia_The_example_of_the_blue_buckets_society

Franchetti, M. (2008, June 29). *Traffic incident gives insight into Russia's corrupt legal system*. The Times. https://www.thetimes.com/comment/register/article/traffic-incident-

gives-insight-into-russias-corrupt-legal-system-mtmk3t2xqfz

Gosnomerus.com. (2024). *License plate listing: E84 KX 90*. Retrieved February 26, 2025, from https://gosnomerus.com/plate/search?number=ekx&region=&region_id=&city_id=&price_min=850000&price_max=950000&category_list%5B0%5D=3

Gosnomerus.com. (n.d.). *Buying a license plate in Russia*. Retrieved February 26, 2025, from https://gosnomerus.com/page/buying.html

Healy, P. M., Ramanna, K., & Shaffer, M. (2012). *Rospil.info* (Harvard Business School Case No. 112-033). Harvard Business School Publishing. https://hbsp.harvard.edu/product/112033-PDF-ENG

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine, 21*(16), 2409–2419. https://doi.org/10.1002/sim.1047

Holmström, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics*, 10(1), 74–91. https://doi.org/10.2307/3003320

Kansas State University. (2012, June 21). *Racial and gender profiling can affect outcome of traffic stops*. ScienceDaily. https://www.sciencedaily.com/releases/2012/06/120621130716.htm

Kazun, A. (2024, May 20). *Men behind the wheel: Three times more violations and accidents than women*. HSE University. https://www.hse.ru/news/science/956617071.html

Khatri, N., Tsang, E. W. K., & Begley, T. M. (2006). Cronyism: A cross-cultural analysis. *Journal of International Business Studies, 37*(1), 61–75. https://doi.org/10.1057/palgrave.jibs.8400171

KoAP RF. (2001). *Kodeks Rossiyskoy Federatsii ob administrativnykh pravonarusheniyakh* [Code of Administrative Offences of the Russian Federation], Federal Law No. 195-FZ, adopted December 30, 2001, effective July 1, 2002. Retrieved from https://www.wipo.int/wipolex/en/legislation/details/12772

KoAP RSFSR. (1984). *Kodeks ob administrativnykh pravonarusheniyakh RSFSR* [Code of Administrative Offences of the RSFSR], as amended to June 15, 1998. Retrieved from https://natlex.ilo.org/dyn/natlex2/r/natlex/fe/details?p3_isn=51265

Kravtsova, Y. (2013, March 31). *Traffic law enforcement disparity destabilizes society*. The Moscow Times. https://www.themoscowtimes.com/2013/03/31/traffic-law-enforcement-disparity-destabilizes-society-a22814

Ledeneva, A. V. (2009). From Russia with blat: Can informal networks help modernize Russia? *Social Research: An International Quarterly, 76*(1), 257–288. https://doi.org/10.1353/sor.2009.0052

Levitt, S. D., & Porter, J. (2001). How dangerous are drinking drivers? *Journal of Political Economy, 109*(6), 1198–1237.

https://doi.org/10.1086/323281

Malik, N. A., & Qureshi, T. A. (2020). A study of economic, cultural, and political causes of police corruption in Pakistan. *Policing: A Journal of Policy and Practice*, 15(2), 1446–1462. https://doi.org/10.1093/police/paaa016

Ragauskas, R., & Valeškaitė, I. (2020). Nepotism, political competition and overemployment. *Political Research Exchange, 2*(1), 1781542. https://doi.org/10.1080/2474736X.2020.1781542

ROSA.gr. (2025, January 13). Χανιά: Οργή για το χαμό του 22χρονου στο δυστύχημα με την Porsche. [Chania: Outrage over the death of the 22-year-old in the Porsche accident]. https://www.rosa.gr/koinonia/chania-orgi-gia-to-chamo-tou-22chronou-sto-dystychima-me-tin-porsche/

Russia Beyond. (2018, March 22). *Hidden messages of Russian license plates*. https://www.rbth.com/lifestyle/327032-hidden-messages-of-russian-license

Schafer, J. A., & Mastrofski, S. D. (2005). Police leniency in traffic enforcement encounters: Exploratory findings from observations and interviews. *Journal of Criminal Justice, 33*(3), 225–238. https://doi.org/10.1016/j.jcrimjus.2005.02.003

Shcherbov, S., Botev, N., & Zaletova, N. (2023). Household size and composition in Russia: Regional and urban-rural differences and changes between 2002 and 2021. *Population and Economics*, *7*(1), 1–22. https://doi.org/10.3897/popecon.7.e107735

Svensson, J. (2005). Eight questions about corruption. *Journal of Economic Perspectives, 19*(3), 19–42. https://doi.org/10.1257/089533005774357860

Szakonyi, D. (2019). Princelings in the private sector: The value of nepotism. *Quarterly Journal of Political Science, 14*(3), 349–381. https://doi.org/10.1561/100.00018087

Tétrault-Farber, G. (2016, May 30). *Kremlin targets Russian elite's street racer kids in pre-vote crackdown*. Yahoo News. https://sg.style.yahoo.com/kremlin-targets-russian-elites-street-racer-kids-pre-150117052.html

Treisman, D. (2000). The causes of corruption: A cross-national study. *Journal of Public Economics*, 76(3), 399–457. https://doi.org/10.1016/S0047-2727(99)00092-4

Wikipedia contributors. (n.d.). *Russian jokes*. In *Wikipedia, The Free Encyclopedia*. Retrieved January 29, 2025, from https://en.wikipedia.org/wiki/Russian_jokes

World Justice Project. (2024). *Rule of Law Index: Absence of Corruption*. Retrieved from https://worldjusticeproject.org/rule-of-law-index/global/2024/Absence%20of%20Corruption

World Justice Project. (2024). *Washington, D.C.: World Justice Project. Retrieved from* https://worldjusticeproject.org/rule-of-law-index/

Yu, J., De Klerk, S., & Hess, M. (2023). The influence of cronyism on entrepreneurial resource acquisition. *Asia Pacific Journal of Management, 40*(1), 121–150.

https://doi.org/10.1007/s10490-021-09769-7

Zakopyrin, V. N., Fateev, A. V., Bolshakova, N. V., & Gubina, E. O. (2021). Administrative offences legislation in Russia and CIS countries: Problems and development trends. In E. G. Popkova (Ed.), *European Proceedings of Social and Behavioural Sciences* (Vol. 113, pp. 1190–1195). European Publisher. https://doi.org/10.15405/epsbs.2021.02.02.148

9 APPENDIX

9.1 ANNEX A. Supplementary text notes

All data cleaning, preprocessing, and analysis were conducted using *Python* in *Jupyter Notebooks*, hosted on UGent's remote server environment. This platform provides regularly updated versions of standard scientific libraries. The data pipeline relied on *pandas* for data manipulation, *NumPy* for numerical operations, *difflib* and *re* for string processing and matching, and *os* for file handling and image export. Visualizations were produced using *matplotlib*, *seaborn*, and *matplotlib_venn*, while *python-docx* was used to format and export regression output tables.

Schematic diagrams, specifically Figures 10, 14, and 20, were manually designed in *draw.io* to complement the programmatically generated outputs and improve visual clarity.

This workflow enabled efficient processing of large administrative datasets, ensured transparency and reproducibility, and supported the production of high-quality analytical and visual outputs.

Note 1: Software and workflow

The observed fine rate among unconnected violators is 75.0%, compared to 70.3% for family members of ministry officials, this is a difference of 4.7 percentage points. A power analysis using a two-proportion z-test (two-sided, $\alpha = 0.05$, power = 0.80) indicates that detecting a difference of this size would require a sample of approximately 1,410 family violators. Only 91 such cases are observed in the dataset. This explains the wide confidence interval and lack of statistical significance, despite a notable gap in point estimates.

Note 2: Power calculation for conditional fine rate comparison

9.2 ANNEX B. Supplementary Tables and Figures

Ministry name (Russian)	Ministry name (English)	Unique employees
МИНИСТЕРСТВО ФИНАНСОВ РОССИЙСКОЙ ФЕДЕРАЦИИ	Ministry of Finance of the Russian Federation	3349
МИНИСТЕРСТВО ПУТЕЙ СООБЩЕНИЯ РФ	Ministry of Railways of the Russian Federation	3194
МИНИСТЕРСТВО ЭКОНОМИКИ РОССИЙСКОЙ ФЕДЕРАЦИИ	Ministry of Economic Development of the Russian Federation	2425
МИНИСТЕРСТВО ЮСТИЦИИ РФ	Ministry of Justice of the Russian Federation	2187
МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ	Ministry of Transport of the Russian Federation	2152
МИНИСТЕРСТВО ТОРГОВЛИ РОССИЙСКОЙ ФЕДЕРАЦИИ	Ministry of Trade of the Russian Federation	1925
МИНИСТЕРСТВО ТРУДА И СОЦИАЛЬНОГО РАЗВИТИЯ РФ	Ministry of Labor and Social Development	1587
МИНИСТЕРСТВО ОБРАЗОВАНИЯ МОСКОВСКОЙ обл.	Ministry of Education of the Moscow Region	1551
МИНИСТЕРСТВО ПРИРОДНЫХ РЕСУРСОВ РОССИЙСКОЙ ФЕДЕРАЦИИ	Ministry of Natural Resources of the Russian Federation	1325
МИНИСТЕРСТВО РФ ПО НАЛОГАМ И СБОРАМ	Ministry for Taxes and Levies of the Russian Federation	1284

Table 11: Ten most represented ministries (by unique employees) in the pension dataset

Car Brand	Count	% of Luxury Cars
Audi	132,374	27.1%
Mercedes	110,711	22.7%
Nissan	83,283	17.1%
Volvo	42,571	8.7%
Subaru	18,846	3.9%
Chevrolet	18,738	3.8%
Jeep	14,848	3.0%
Saab	12,415	2.5%
Lexus	11,082	2.3%
Chrysler	7,623	1.6%
Rover	6,770	1.4%
Dodge	6,500	1.3%
Land Rover	5,375	1.1%
Daimler	2,768	0.6%
Ssang Yong	2,727	0.6%
Lincoln	2,132	0.4%
Jaguar	1,611	0.3%
Cadillac	1,410	0.3%
Infiniti	1,187	0.2%
Porsche	1,087	0.2%
Mercury	1,026	0.2%
Gmc	853	0.2%
Oldsmobile	541	0.1%
Acura	403	0.1%
Hummer	353	0.1%
Range Rover	353	0.1%
Mini	175	<0.1%
Bentley	144	<0.1%
Ferrari	72	<0.1%
Rolls Royce	39	<0.1%
Eagle	8	<0.1%
Triumph	3	<0.1%
Saturn	3	<0.1%
Total	**488,031**	**100.0%**

Table 12: Distribution of luxury brands in the dataset

Note: The table reports 488,031 car records flagged as luxury vehicles. These correspond to 483,182 unique car–person combinations, as some vehicles are associated with multiple individuals in the administrative data.

Vanity number	Count	% of Vanity cars
808.0	4,843	1.9%
777.0	4,814	1.9%
800.0	4,803	1.9%
909.0	4,782	1.9%
555.0	4,771	1.9%
888.0	4,762	1.9%
400.0	4,761	1.9%
999.0	4,752	1.9%
600.0	4,739	1.9%
444.0	4,735	1.9%
111.0	4,726	1.9%
787.0	4,705	1.9%
101.0	4,702	1.9%
63.0	4,688	1.9%
222.0	4,661	1.9%
707.0	4,657	1.9%
88.0	4,649	1.9%
64.0	4,611	1.9%
11.0	4,609	1.9%
202.0	4,608	1.9%
30.0	4,574	1.8%
80.0	4,573	1.8%
900.0	4,572	1.8%
77.0	4,561	1.8%
700.0	4,555	1.8%
20.0	4,548	1.8%
70.0	4,536	1.8%
97.0	4,530	1.8%
10.0	4,521	1.8%
55.0	4,518	1.8%
90.0	4,503	1.8%
40.0	4,495	1.8%
666.0	4,472	1.8%
200.0	4,466	1.8%
333.0	4,439	1.8%
100.0	4,438	1.8%
99.0	4,419	1.8%
21.0	4,391	1.8%
44.0	4,337	1.7%
717.0	4,328	1.7%
303.0	4,300	1.7%
300.0	4,294	1.7%
22.0	4,286	1.7%
8.0	4,264	1.7%
23.0	4,260	1.7%
505.0	4,249	1.7%
9.0	4,246	1.7%
50.0	4,241	1.7%
7.0	4,226	1.7%

6.0	4,196	1.7%
5.0	4,182	1.7%
3.0	4,138	1.7%
4.0	4,102	1.6%
2.0	4,054	1.6%
500.0	4,040	1.6%
1.0	4,035	1.6%
Total	**251,267**	**100.0%**

Table 13: Distribution of vanity plate numbers

Note: Total count exceeds number of vanity-flagged car–person combinations (248,630) because some numbers appear across mµltiple entries.

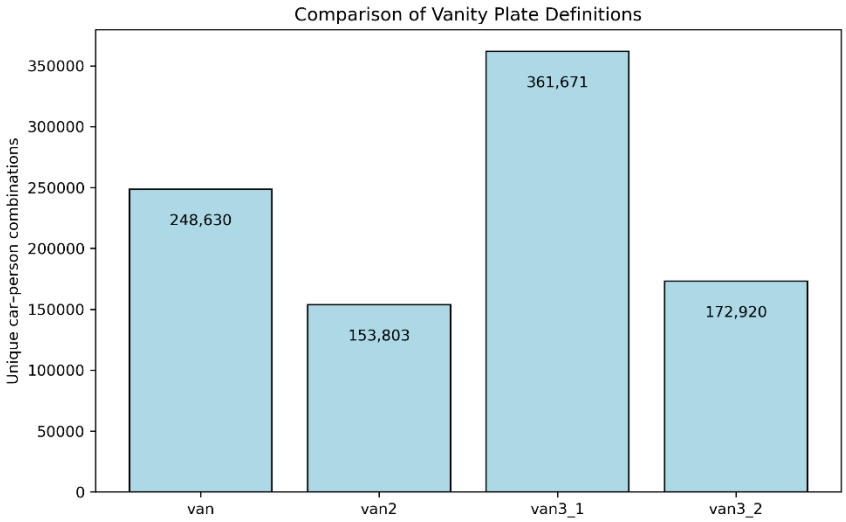


Figure 21: Comparison of vanity plate definitions

Number of unique car–person combinations captured under different vanity plate definitions. The final indicator (*van*) was used in the analysis for its conceptual clarity and alignment with prior literature (Eeckhout et al., 2021). Other definitions are shown for comparison.

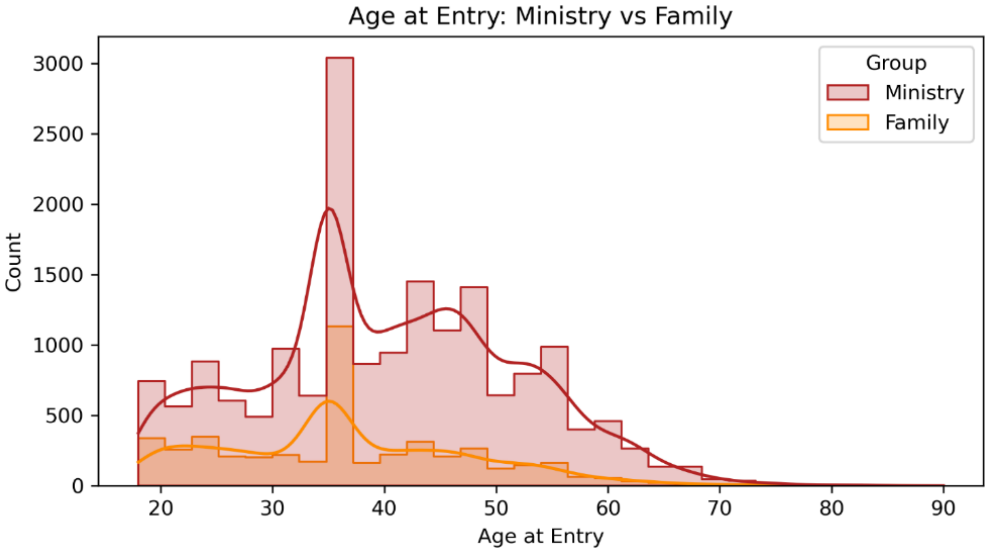


Figure 22: Age at entry: Ministry vs Family

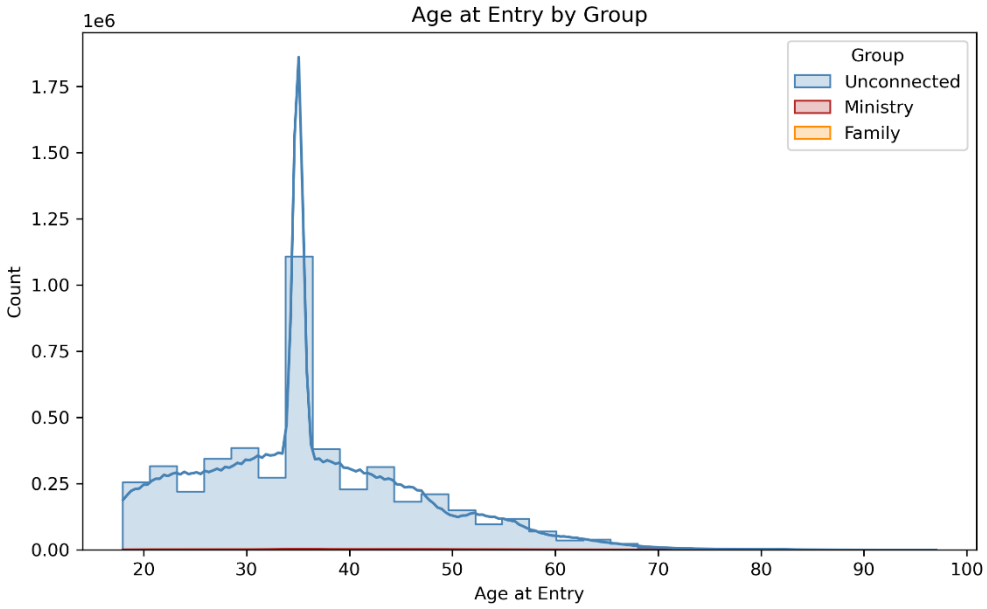


Figure 23: Age at entry: Unconnected

9.3 ANNEX C. Model Specifications and Output

9.3.1 Descriptive comparisons by group

Outcome	Ministry	Unconnected	% Difference	p-value
P(Violation)	1.6202%	3.1459%	-48.5%	0.00000
P(Fine)	1.0707%	2.3592%	-54.6%	0.00000
P(Fine given Violation)	66.0839%	74.9917%	-11.9%	0.00078
P(License Deprivation)	0.0340%	0.0738%	-53.9%	0.00190
P(Accident)	0.4929%	0.3557%	+38.5%	0.00939
Avg Fines	0.0120	0.0313	-61.6%	0.00000
Avg Fines per Violator	0.7413	0.9951	-25.5%	0.00000
Avg Violations	0.0184	0.0434	-57.7%	0.00000
Avg License Deprivations	0.0003	0.0008	-56.0%	0.00190
Avg Accidents	0.0050	0.0037	+34.7%	0.01674

Table 14: Summary of enforcement outcomes: ministry vs. unconnected

This table compares enforcement outcomes between ministry-connected and unconnected drivers. Binary outcomes (e.g., P(Fine), P(Violation)) are based on the proportion of car–person combinations with at least one event. Continuous outcomes (e.g., average fines per combo) reflect the mean number of events per combo. Two-sided t-tests are used for continuous outcomes, and Fisher's exact test is used for rare binary outcomes (license deprivation) and conditional outcomes (fines given violation). Significance tests are two-tailed.

Outcome	Family	Unconnected	% Difference	p-value
P(Violation)	1.9333%	3.1459%	-38.5%	0.00000
P(Fine)	1.3597%	2.3592%	-42.4%	0.00000
P(Fine given Violation)	70.3297%	74.9917%	-6.2%	0.33224
P(License Deprivation)	0.0212%	0.0738%	-71.2%	0.00849
P(Accident)	0.5949%	0.3557%	+67.2%	0.03301
Avg Fines	0.0159	0.0313	-49.1%	0.00000
Avg Fines per Violator	0.8242	0.9951	-17.2%	0.02083
Avg Violations	0.0238	0.0434	-45.2%	0.00000
Avg License Deprivations	0.0002	0.0008	-72.5%	0.00849
Avg Accidents	0.0062	0.0037	+66.5%	0.03710

Table 15: Summary of enforcement outcomes: family vs. unconnected

Definitions, measurement strategy, and statistical tests are consistent with Table 14

Outcome	Luxury	Non-Luxury	% Difference	p-value
P(Violation)	2.6143%	3.2034%	-18.4%	0.00000
P(Fine)	1.8757%	2.4108%	-22.2%	0.00000
P(Fine given Violation)	71.7464%	75.2589%	-4.7%	0.00000
P(License Deprivation)	0.0333%	0.0782%	-57.4%	0.00000
P(Accident)	0.4443%	0.3469%	+28.1%	0.00000
Avg Fines	0.0246	0.0320	-23.0%	0.00000
Avg Fines per Violator	0.9423	0.9992	-5.7%	0.00000
Avg Violations	0.0357	0.0442	-19.2%	0.00000
Avg License Deprivations	0.0004	0.0008	-56.3%	0.00000
Avg Accidents	0.0046	0.0036	+28.2%	0.00000

Table 16: Summary of enforcement outcomes: luxury vs. non-luxury

Definitions, measurement strategy, and statistical tests are consistent with Table 14

Outcome	Vanity	Non-Vanity	% Difference	p-value
P(Violation)	2.7784%	3.1588%	-12.0%	0.00000
P(Fine)	2.0577%	2.3696%	-13.2%	0.00000
P(Fine given Violation)	74.0591%	75.0159%	-1.3%	0.07323
P(License Deprivation)	0.0491%	0.0749%	-34.5%	0.00000
P(Accident)	0.3391%	0.3574%	-5.1%	0.12544
Avg Fines	0.0273	0.0314	-13.3%	0.00000
Avg Fines per Violator	0.9812	0.9951	-1.4%	0.22274
Avg Violations	0.0382	0.0436	-12.4%	0.00000
Avg License Deprivations	0.0005	0.0008	-34.4%	0.00000
Avg Accidents	0.0035	0.0037	-6.3%	0.06359

Table 17: Summary of enforcement outcomes: vanity vs. non-vanity

Definitions, measurement strategy, and statistical tests are consistent with Table 14

Group	Total Fines	Total Violations	Fines per Violation	% Difference (vs. UC)	Bootstrap p-value
Ministry Officials (L0)	189	286	0.654	-9.3	0.00075
Ministry Family (L1)	64	91	0.67	-7.1	0.264
Unconnected (UC)	112501	150018	0.721	Reference	—

Table 18: Average fines per violation per group

This table summarizes the total number of fines and violations for each group, and reports the group-level average number of fines per recorded violation. The "% Difference" column compares each group to the unconnected baseline. Bootstrap p-values are based on 1,000 resamples and reflect two-tailed significance for the observed differences. Bootstrapping is used to obtain more robust inference when comparing ratios across groups, particularly under skewed distributions and unequal or small group sizes.

9.3.2 Regression results: political connections

All regression models are estimated using logistic regression. Tables report marginal effects. Robust standard errors are shown in parentheses. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1. Models use car–person combinations as the unit of analysis. “PD FE” indicates the inclusion of police district fixed effects.

9.3.2.1 *Fine*

Variable	Fine ~ ministry	Fine ~ family	Fine ~ ministry + controls	Fine ~ family + controls	Fine ~ ministry + controls + PD FE	Fine ~ family + controls + PD FE
Intercept	-3.7234 (0.0031)*** p = 0.0000	-3.7251 (0.0031)*** p = 0.0000	-2.8154 (0.0109)*** p = 0.0000	-2.8141 (0.0109)*** p = 0.0000	-1.6770 (0.0267)*** p = 0.0000	-1.6765 (0.0267)*** p = 0.0000
is_ministry	-0.8027 (0.0744)*** p = 0.0000		-0.7120 (0.0746)*** p = 0.0000		-0.3331 (0.0748)*** p = 0.0000	
ministry_family_final_flag		-0.5592 (0.1250)*** p = 0.0000		-0.4836 (0.1252)*** p = 0.0001		-0.0422 (0.1276) p = 0.7408
C(Gender)[T.1.0]			-1.4586 (0.0122)*** p = 0.0000	-1.4587 (0.0122)*** p = 0.0000	-1.3346 (0.0122)*** p = 0.0000	-1.3348 (0.0122)*** p = 0.0000
AgeAtEntry			-0.0202 (0.0003)*** p = 0.0000	-0.0202 (0.0003)*** p = 0.0000	-0.0159 (0.0003)*** p = 0.0000	-0.0159 (0.0003)*** p = 0.0000
PD Fixed Effects	No	No	No	No	Yes	Yes
Observations	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0000	0.0247	0.0246	0.1358	0.1357

Table 19: Logit estimates – Fine receipt by ministry and family

9.3.2.2 *Violation*

Variable	Violation ~ ministry	Violation ~ family	Violation ~ ministry + controls	Violation ~ family + controls	Violation ~ ministry + controls + PD FE	Violation ~ family + controls + PD FE
Intercept	- 3.4275 (0.0027)*** p = 0.0000	-3.4290 (0.0027)*** p = 0.0000	-2.5753 (0.0096)*** p = 0.0000	-2.5742 (0.0096)*** p = 0.0000	-1.4114 (0.0235)*** p = 0.0000	-1.4109 (0.0235)*** p = 0.0000
is_ministry	-0.6788 (0.0606)*** p = 0.0000		-0.5953 (0.0608)*** p = 0.0000		-0.2916 (0.0611)*** p = 0.0000	
ministry_family_final_flag		-0.4975 (0.1060)*** p = 0.0000		-0.4216 (0.1061)*** p = 0.0001		-0.0740 (0.1073) p = 0.4901
C(Gender)[T.1.0]			-1.4177 (0.0104)*** p = 0.0000	-1.4178 (0.0104)*** p = 0.0000	-1.3214 (0.0104)*** p = 0.0000	-1.3215 (0.0104)*** p = 0.0000
AgeAtEntry			-0.0186 (0.0003)*** p = 0.0000	-0.0186 (0.0003)*** p = 0.0000	-0.0150 (0.0003)*** p = 0.0000	-0.0150 (0.0003)*** p = 0.0000
PD Fixed Effects	No	No	No	No	Yes	Yes
Observations	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0000	0.0247	0.0247	0.1159	0.1158

Table 20: Logit estimates – Traffic violations by ministry and family

9.3.2.3 *Fine given violation*

Variable	Fine given violation ~ ministry	Fine given violation ~ family	Fine given violation ~ ministry + controls	Fine given violation ~ family + controls	Fine given violation ~ ministry + controls + PD FE	Fine given violation ~ family + controls + PD FE
Intercept	1.0980 (0.0060)*** p = 0.0000	1.0973 (0.0060)*** p = 0.0000	1.3534 (0.0220)*** p = 0.0000	1.3541 (0.0220)*** p = 0.0000	1.2654 (0.0532)*** p = 0.0000	1.2654 (0.0532)*** p = 0.0000
is_ministry	-0.4310 (0.1246)*** p = 0.0005		-0.3892 (0.1251)*** p = 0.0019		-0.1029 (0.1283) p = 0.4226	
ministry_family_final_flag		-0.2342 (0.2270) p = 0.3022		-0.2343 (0.2279) p = 0.3040		0.1974 (0.2440) p = 0.4185
C(Gender)[T.1.0]			-0.1885 (0.0227)*** p = 0.0000	-0.1892 (0.0227)*** p = 0.0000	-0.0462 (0.0238)* p = 0.0524	-0.0463 (0.0238)* p = 0.0514
AgeAtEntry			-0.0070 (0.0006)*** p = 0.0000	-0.0070 (0.0006)*** p = 0.0000	-0.0020 (0.0006)*** p = 0.0011	-0.0020 (0.0006)*** p = 0.0011
PD Fixed Effects	No	No	No	No	Yes	Yes
Observations	150,395	150,395	150,395	150,395	150,395	150,395
Pseudo R-squared	0.0001	0.0000	0.0012	0.0012	0.0691	0.0691

Table 21: Logit estimates – Fine receipt conditional on violation (ministry and family)

9.3.2.4 *Accident*

Variable	Accident ~ ministry	Accident ~ family	Accident ~ ministry + controls	Accident ~ family + controls	Accident ~ ministry + controls + PD FE	Accident ~ family + controls + PD FE
Intercept	-5.6345 (0.0077)*** p = 0.0000	-5.6337 (0.0077)*** p = 0.0000	-5.0863 (0.0285)*** p = 0.0000	-5.0874 (0.0285)*** p = 0.0000	-5.9728 (0.1877)*** p = 0.0000	-5.9738 (0.1877)*** p = 0.0000
is_ministry	0.3268 (0.1100)*** p = 0.0030		0.3859 (0.1100)*** p = 0.0005		0.1732 (0.1105) p = 0.1170	
ministry_family_final_flag		0.5151 (0.1891)*** p = 0.0065		0.5210 (0.1891)*** p = 0.0059		0.3127 (0.1893)* p = 0.0985
C(Gender)[T.1.0]			-0.0494 (0.0186)*** p = 0.0078	-0.0495 (0.0186)*** p = 0.0076	-0.1440 (0.0185)*** p = 0.0000	-0.1442 (0.0185)*** p = 0.0000
AgeAtEntry			-0.0152 (0.0008)*** p = 0.0000	-0.0151 (0.0008)*** p = 0.0000	-0.0184 (0.0007)*** p = 0.0000	-0.0183 (0.0007)*** p = 0.0000
PD Fixed Effects	No	No	No	No	Yes	Yes
Observations	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0000	0.0019	0.0018	0.0376	0.0376

Table 22: Logit estimates – Accident involvement by ministry and family

9.3.2.5 *License deprivation*

Variable	License deprived ~ ministry	License deprived ~ family	License deprived ~ ministry + controls	License deprived ~ family + controls	License deprived ~ ministry + controls + PD FE	License deprived ~ family + controls + PD FE
Intercept	-7.2122 (0.0169)*** p = 0.0000	-7.2135 (0.0169)*** p = 0.0000	-5.9712 (0.0573)*** p = 0.0000	-5.9697 (0.0573)*** p = 0.0000	-4.7426 (0.1343)*** p = 0.0000	-4.7421 (0.1343)*** p = 0.0000
is_ministry	-0.7743 (0.4087)* p = 0.0581		-0.6404 (0.4088) p = 0.1172		-0.2015 (0.4101) p = 0.6232	
ministry_family_final_flag		-1.2431 (1.0000) p = 0.2138		-1.1612 (0.9999) p = 0.2455		-0.6593 (1.0019) p = 0.5105
C(Gender)[T.1.0]			-2.2173 (0.0956)*** p = 0.0000	-2.2173 (0.0956)*** p = 0.0000	-2.0396 (0.0956)*** p = 0.0000	-2.0397 (0.0956)*** p = 0.0000
AgeAtEntry			-0.0291 (0.0016)*** p = 0.0000	-0.0291 (0.0016)*** p = 0.0000	-0.0253 (0.0016)*** p = 0.0000	-0.0253 (0.0016)*** p = 0.0000
PD Fixed Effects	No	No	No	No	Yes	Yes
Observations	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0000	0.0241	0.0240	0.1110	0.1110

Table 23: Logit estimates – License deprivation by ministry and family

9.3.3 Regression results: informal privilege signals

This section presents logistic regression results analyzing the probability of enforcement outcomes while controlling for informal privilege indicators: luxury vehicle ownership (*lux*) and vanity plate ownership (*van*).

For each outcome variable, results are reported in two separate tables:

- One for the sample of ministry officials (*is_ministry*)
- One for the sample of their connected family members (*ministry_family_final_flag)* All models report marginal effects. Robust standard errors are shown in parentheses. Significance levels: *** p < 0.01, ** p < 0.05, * p < 0.1. The unit of analysis is the car–person combination.

9.3.3.1 *Fine*

Variable	Fine ~ ministry	Fine ~ ministry + lux	Fine ~ ministry + van	Fine ~ ministry + lux + van
Intercept	-3.7234 (0.0031)*** p = 0.0000	-3.7004 (0.0032)*** p = 0.0000	-3.7164 (0.0032)*** p = 0.0000	-3.6945 (0.0033)*** p = 0.0000
is_ministry	-0.8027 (0.0744)*** p = 0.0000	-0.7967 (0.0744)*** p = 0.0000	-0.8018 (0.0744)*** p = 0.0000	-0.7959 (0.0744)*** p = 0.0000
lux		-0.2543 (0.0111)*** p = 0.0000		-0.2499 (0.0112)*** p = 0.0000
van			-0.1441 (0.0146)*** p = 0.0000	-0.1291 (0.0146)*** p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0007	0.0002	0.0008

Table 24: Logit estimates – Fine receipt (ministry, with luxury and vanity)

Variable	Fine ~ family	Fine ~ family + lux	Fine ~ family + van	Fine ~ family + lux + van
Intercept	-3.7251 (0.0031)*** p = 0.0000	-3.7020 (0.0032)*** p = 0.0000	-3.7180 (0.0032)*** p = 0.0000	-3.6961 (0.0033)*** p = 0.0000
ministry_family_final_flag	-0.5592 (0.1250)*** p = 0.0000	-0.5431 (0.1250)*** p = 0.0000	-0.5575 (0.1250)*** p = 0.0000	-0.5418 (0.1250)*** p = 0.0000
lux		-0.2545 (0.0111)*** p = 0.0000		-0.2502 (0.0112)*** p = 0.0000
van			-0.1442 (0.0146)*** p = 0.0000	-0.1292 (0.0146)*** p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0006	0.0001	0.0006

Table 25: Logit estimates – Fine receipt (family, with luxury and vanity)

9.3.3.2 *Violation*

Variable	Violation ~ ministry	Violation ~ ministry + lux	Violation ~ ministry + van	Violation ~ ministry + lux + van
Intercept	-3.4275 (0.0027)*** p = 0.0000	-3.4084 (0.0028)*** p = 0.0000	-3.4210 (0.0028)*** p = 0.0000	-3.4029 (0.0029)*** p = 0.0000
is_ministry	-0.6788 (0.0606)*** p = 0.0000	-0.6738 (0.0606)*** p = 0.0000	-0.6780 (0.0606)*** p = 0.0000	-0.6731 (0.0606)*** p = 0.0000
lux		-0.2070 (0.0095)*** p = 0.0000		-0.2029 (0.0095)*** p = 0.0000
van			-0.1320 (0.0126)*** p = 0.0000	-0.1196 (0.0126)*** p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0005	0.0002	0.0006

Table 26: Logit estimates – Traffic violations (ministry, with luxury and vanity)

Variable	Violation ~ family	Violation ~ family + lux	Violation ~ family + van	Violation ~ family + lux + van
Intercept	-3.4290 (0.0027)*** p = 0.0000	-3.4098 (0.0028)*** p = 0.0000	-3.4225 (0.0028)*** p = 0.0000	-3.4043 (0.0029)*** p = 0.0000
ministry_family_final_flag	-0.4975 (0.1060)*** p = 0.0000	-0.4841 (0.1060)*** p = 0.0000	-0.4959 (0.1060)*** p = 0.0000	-0.4829 (0.1060)*** p = 0.0000
lux		-0.2072 (0.0095)***p = 0.0000		-0.2032 (0.0095)*** p = 0.0000
van			-0.1321 (0.0126)*** p = 0.0000	-0.1197 (0.0126)*** p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0004	0.0001	0.0005

Table 27: Logit estimates – Traffic violations (family, with luxury and vanity)

9.3.3.3 *Fine given violation*

Variable	Fine given violation ~ ministry	Fine given violation + lux	Fine given violation + van	Fine given violation + lux + van
Intercept	-3.7234 (0.0031)*** p = 0.0000	-3.7004 (0.0032)*** p = 0.0000	-3.7164 (0.0032)*** p = 0.0000	-3.6945 (0.0033)*** p = 0.0000
is_ministry	-0.8027 (0.0744)*** p = 0.0000	-0.7967 (0.0744)*** p = 0.0000	-0.8018 (0.0744)*** p = 0.0000	-0.7959 (0.0744)*** p = 0.0000
lux		-0.2543 (0.0111)*** p = 0.0000)		-0.2499 (0.0112)*** p = 0.0000
van			-0.1441 (0.0146)*** p = 0.0000	-0.1291 (0.0146)*** p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0007	0.0002	0.0008

Table 28: Logit estimates – Conditional fine receipt (ministry, with luxury and vanity)

Variable	Fine given violation ~ family	Fine given violation + lux	Fine given violation + van	Fine given violation + lux + van
Intercept	-3.7251 (0.0031)*** p = 0.0000	-3.7020 (0.0032)*** p = 0.0000	-3.7180 (0.0032)*** p = 0.0000	-3.6961 (0.0033)*** p = 0.0000
ministry_family_final_flag	-0.5592 (0.1250)*** p = 0.0000	-0.5431 (0.1250)*** p = 0.0000	-0.5575 (0.1250)*** p = 0.0000	-0.5418 (0.1250)*** p = 0.0000
lux		-0.2545 (0.0111)*** p = 0.0000		-0.2502 (0.0112)*** p = 0.0000
van			-0.1442 (0.0146)*** p = 0.0000	-0.1292 (0.0146)***p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0006	0.0001	0.0006

Table 29: Logit estimates – Conditional fine receipt (family, with luxury and vanity)

9.3.3.4 *Accident*

Variable	Accident ~ ministry	Accident ~ ministry + lux	Accident ~ ministry + van	Accident ~ ministry + lux + van
Intercept	-5.6345 (0.0077)*** p = 0.0000	-5.6626 (0.0083)*** p = 0.0000	-5.6318 (0.0079)*** p = 0.0000	-5.6592 (0.0084)*** p = 0.0000
is_ministry	0.3268 (0.1100)*** p = 0.0030	0.3196 (0.1099)*** p = 0.0037	0.3271 (0.1100)*** p = 0.0029	0.3199 (0.1099)*** p = 0.0036
lux		0.2490 (0.0232)*** p = 0.0000		0.2514 (0.0232)*** p = 0.0000
van			-0.0531 (0.0354) p = 0.1328	-0.0712 (0.0354)** p = 0.0444
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0005	0.0000	0.0005

Table 30: Logit estimates – Accident involvement (ministry, with luxury and vanity)

Variable	Accident ~ family	Accident ~ family + lux	Accident ~ family + van	Accident ~ family + lux + van
Intercept	-5.6337 (0.0077)*** p = 0.0000	-5.6618 (0.0082)*** p = 0.0000	-5.6311 (0.0079)*** p = 0.0000	-5.6585 (0.0084)*** p = 0.0000
ministry_family_final_flag	0.5151 (0.1891)*** p = 0.0065	0.4963 (0.1891)*** p = 0.0087	0.5158 (0.1891)*** p = 0.0064	0.4969 (0.1891)*** p = 0.0086
lux		0.2489 (0.0232)*** p = 0.0000		0.2513 (0.0233)*** p = 0.0000
van			-0.0531 (0.0354) p = 0.1330	-0.0712 (0.0354)** p = 0.0445
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0005	0.0000	0.0005

Table 31: Logit estimates – Accident involvement (family, with luxury and vanity)

9.3.3.5 *License deprivation*

Variable	License Deprivation ~ ministry	License Deprivation ~ ministry + lux	License Deprivation ~ ministry + van	License Deprivation ~ ministry + lux + van
Intercept	-7.2122 (0.0169)*** p = 0.0000	-7.1527 (0.0173)*** p = 0.0000	-7.1941 (0.0172)*** p = 0.0000	-7.1372 (0.0176)*** p = 0.0000
is_ministry	-0.7743 (0.4087)* p = 0.0581	-0.7582 (0.4087)* p = 0.0636	-0.7720 (0.4087)* p = 0.0589	-0.7562 (0.4087)* p = 0.0643
lux		-0.8513 (0.0807)*** p = 0.0000		-0.8397 (0.0807)*** p = 0.0000
van			-0.4228 (0.0922)*** p = 0.0000	-0.3825 (0.0922)***p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0001	0.0026	0.0005	0.0029

Table 32: Logit estimates – License deprivation (ministry, with luxury and vanity)

Variable	License Deprivation ~ family	License Deprivation ~ family + lux	License Deprivation ~ family + van	License Deprivation ~ family + lux + van
Intercept	-7.2135 (0.0169)*** p = 0.0000	-7.1540 (0.0173)*** p = 0.0000	-7.1954 (0.0172)*** p = 0.0000	-7.1385 (0.0176)*** p = 0.0000
ministry_family_final_flag	-1.2431 (1.0000) p = 0.2138	-1.2000 (0.9999) p = 0.2301	-1.2387 (1.0000) p = 0.2155	-1.1963 (0.9999) p = 0.2315
lux		-0.8513 (0.0807)***p = 0.0000		-0.8397 (0.0807)*** p = 0.0000
van			-0.4229 (0.0922)*** p = 0.0000	-0.3826 (0.0922)*** p = 0.0000
Observations	4,791,059	4,791,059	4,791,059	4,791,059
Pseudo R-squared	0.0000	0.0026	0.0005	0.0029

Table 33: Logit estimates – License deprivation (family, with luxury and vanity)