Working Paper Series **637**
(ISSN 1211-3298)

Under Pressure? Performance Evaluation of Police Officers as an Incentive to Cheat: Evidence from Drug Crimes in Russia

Ekaterina Travova

CERGE-EI
Prague, April 2019

ISBN 978-80-7343-444-1 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-493-8 (Národohospodářský ústav AV ČR, v. v. i.)

Under Pressure? Performance Evaluation of Police Officers as an Incentive to Cheat: Evidence from Drug Crimes in Russia*

Ekaterina Travova†(CERGE-EI‡, EUSP§)

April, 2019

Abstract

This paper provides an empirical analysis of possible manipulations of amounts of seized drugs, based on a unique dataset that contains full information on drug crimes in Russia reported during 2013-2014. First, using a standard bunching estimator, I investigate the incentives for police officers to manipulate and find that the motivation most likely arises from the officers' performance evaluation system. Second, applying a novel bunching technique, I determine that police officers are more likely to manipulate the drug amounts seized from repeat offenders. The overall effect of manipulation is an additional year of incarceration, and this is not dependent on a guilty plea.

JEL Classification: H11, H76, K14, K42.

Keywords: Drug Crimes, Police Discretion, Performance Evaluation, Incentives.

---

*I am grateful to Andreas Menzel for generous support at all stages of this project. I also thank Stepan Jurajda, Randall Filer, Nikolas Mittag, Sergei Mikhalishchev, Dmitriy Skougarevskiy, Alexey Knorre, Jan Palguta and seminar participants at CERGE-EI and EUSP for helpful discussions and comments. All errors remaining in this text are the responsibility of the author. This study was supported by Charles University, GAUK project 702218.

†e-mail: ekaterina.travova@cerge-ei.cz

‡CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic.

§European University at Saint-Petersburg, Shpalernaya st. 1, 191187 Saint Petersburg, Russia.

1 Introduction

For decades, global anti-drug policies have been based on general principals of eliminating the production, sale and consumption of any illegal psychoactive compounds, and have involved harsh law enforcement and even militarization. Even though the likely failure of a war on drugs has been established, these policies are still in place in many countries, leading to unequal treatment of different groups of drug offenders (International Drug Policy Consortium 2018). While the harmful consequences of tough anti-drug laws are well-studied, there is little evidence as to how this war creates specific incentives that affect the behavior of law enforcers.

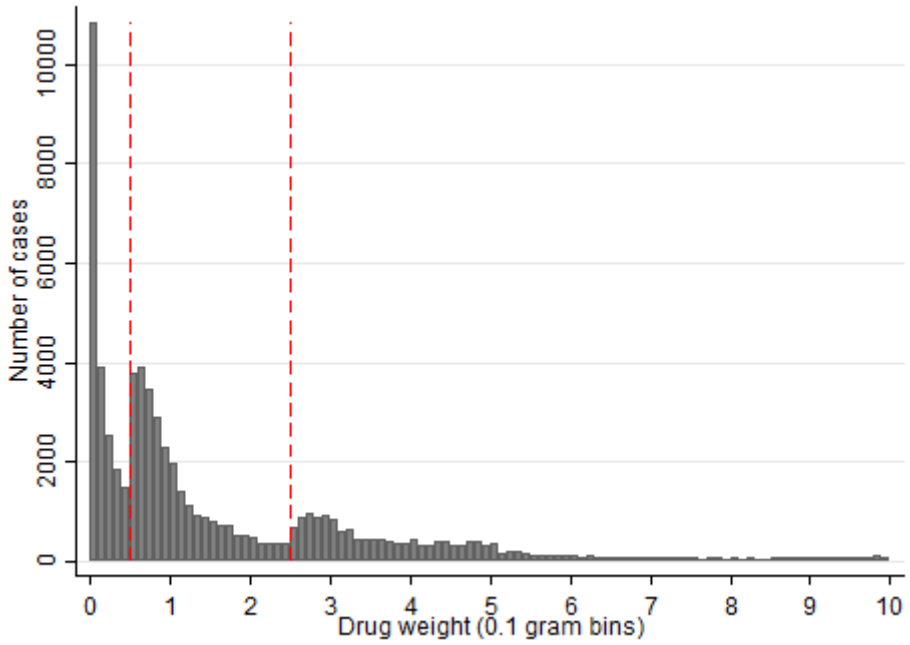
Russia is a particularly notable example of the war on drugs, as according to The Federal Penitentiary Service of Russia (2017), at the end of 2016 almost a quarter of all prisoners were convicted of drug related crimes. Moreover, 70% of these drug offenders were imprisoned for using, not for producing or selling drugs[1]. At the same time, the Russian law for drug possession for personal use carries penalties of up to 15 years of incarceration. This raises questions about the effectiveness and fairness of the criminal justice system.

Figure 1 shows the distribution of drug related cases across quantities of heroin seized in Russia during 2013-2014. The thresholds determine the scales of seizures (less than significant, significant and large drug amounts) that define the severity of crime and punishment[2] (Appendix, Tables B1 and B2). It seems suspicious that, at the moment of arrest, many people apparently possess a drug amount just above a threshold beyond which they will be convicted of a more serious crime. In addition, there is a missing mass of cases just below the defined thresholds. This might be evidence of manipulation of drug quantities seized by the police, which so far has only been alleged by various media reports.

[1] In comparison, worldwide, 1 in 5 prisoners are incarcerated for drug offences. 83% of them are convicted of drug possession for personal use (International Drug Policy Consortium 2018).

[2] There is also a third threshold (at 500 grams for heroin) that is not depicted on the graph. Above it the amount of heroin is especially large.

Figure 1: Distribution of cases across quantities of heroin seized in Russia during 2013-2014



Note: The baseline sample consists of all heroin related cases registered in Russia during 2013-2014. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

This phenomenon might be driven by a rational response of police officers to performance evaluation requirements that, in the context of tough anti-drug laws, establish a strong motivation to present as many prosecutions of serious and most serious crimes as possible. In this paper, I provide evidence of the importance of incentives, exploiting changes in punishment across thresholds and differences in the evaluation approaches of the two law enforcement agencies which were responsible for drug control during 2013-2014. While one of these agencies compared the performance of regional offices cross-sectionally in order to set incentives, the other agency used a comparison within each office over time. To study the differences in magnitudes of manipulation and how they could be related to the incentive structure, I apply the standard bunching estimator (Saez 2010, Chetty et al. 2011, Kleven and

Waseem 2013) focusing on cases related to heroin, one of the most widely used and most potent drugs in Russia at that time. I analyze manipulation only at the second threshold, due to unclear incentives and data limitations around the first and third thresholds (see Section 5 for more detailed discussion).

In addition, adopting the bunching technique recently developed by Diamond and Persson (2016), I identify the mean characteristics of victims of manipulation of seized drug amounts, and the effect of the manipulation on sentence length. The results suggest that while demographics and low socio-economic status of offenders do not have a significant effect on a police officer's decision to manipulate, having a criminal history increases the probability of becoming the victim of such a manipulation. The overall estimated effect of the manipulation on sentence length is around one additional year of incarceration, and the magnitude of this effect is not dependent on a guilty plea.

Using the estimates obtained, and information on the annual budget of a prison authority per convict, I calculate an additional pressure on the overall country budget of around $13.5 million. However, this is a lower bound of the total social cost induced by manipulations, since it does not take into account that longer incarceration strengthens barriers to reintegration after release, increases the probability of recidivism and amplifies the spillover effect.

2 Related Literature

This paper is related to the broad scope of literature on the Economics of Law Enforcement. One strand of this literature focuses on the performance evaluation of law enforcers and incentive schemes at different levels of the criminal justice system. For example, Mas (2006) studies the effect of arbitration decisions on the performance of police officers. He finds that, in the case of favorable outcomes, crime reports rise and arrest rates and average sentence length decline. In contrast,

Prendergast (2001) shows that increased external oversight leads police officers to reduce crime-fighting activities to avoid possible investigation. As was highlighted in the seminal paper by Holmstrom and Milgrom (1991), strong incentives could be inappropriate in government jobs, resulting in a negative effect. A civil servant's job is characterized by multitasking. At the same time, some objectives that the civil servant has to attend to are more easily measured than others. In this context, strong incentives could detract attention away from tasks that are not easily measured, leading to inefficiency. Banerjee et al. (2012) run a sequence of experiments in India and find that reduced autonomy of police station managers reduces police effectiveness. This supports another explanation of the possible negative effect of some incentive systems: the issue may not be with the incentives themselves, but with their implementation (Banerjee et al. 2012).

Another strand of this literature concentrates on disparities in the responses of law enforcers to punishment cliffs, their discretion and its consequences. Anbarci and Lee (2014) and Goncalves and Mello (2017) use US data on speeding tickets and find an excess mass at speeds just below the first threshold, above which the fine increases. They take this bunching as evidence of manipulation by police officers, who may wish to avoid onerous punishment for drivers. The authors also determine that these police officers are less lenient to minority drivers. In addition, racial disparities were spotted in the context of police stops, searches and use of force (Knowles et al. 2001, Persico 2009, Goel et al. 2016, Legewie 2016). At the level of the prosecutor's office, Bjerk (2005), Ulmer et al. (2007) and Rehavi and Starr (2014) find that some prosecutors are more likely to charge offenders who were initially arrested for crimes under a mandatory minimum sentencing law with a lesser crime not covered by this law. Bjerk (2017) focuses on drug crimes in the US and finds that first-time drug offenders are likely to avoid prosecution under a mandatory minimum law. At the court level, some papers determine a strong bias in judges' decisions against blacks (Abrams et al. 2012, Alesina and La Ferrara

2014, Arnold et al. 2017), while others do not find this effect (Anwar and Fang 2015, Mechoulan and Sahuguet 2015). In countries where race and ethnicity play a lesser role in defining status differences, studies show the significant effect of social networks in China (Liu et al. 1998), gender and age in South Korea (Lee et al. 2011), and other personal characteristics of both the offender and the victim in the Netherlands (Johnson et al. 2010).

Despite Russia having one of the highest incarceration rates in the world (Institute for Criminal Policy Research, n.d.), there is little empirical evidence relating to the Russian law enforcement system and its specific incentive structure. Most of the existing literature is descriptive, based on interviews and simple summary statistics. For example, Paneyakh (2014) examines the Russian legal system and describes how the system of performance evaluation of law enforcers could be a critical source of prosecution and conviction bias. More rigorous analysis of the law enforcers' response to punishment thresholds can be found in Skougarevskiy (2017). Applying RDD analysis to data on cannabis and heroin cases, the author finds that the length of unconditional incarceration increases by 0.84 years when drug weight crosses the threshold. Volkov (2016) analyzes all felony cases processed by federal district courts during 2009-2013, and finds a significant bias in judges' decisions against entrepreneurs, and offenders of low socio-economic status. Kurmangaliyeva (2017) determines that the Russian judicial system is more lenient to wealthier defendants. My paper adds to this literature by presenting rigorous evidence on how law enforcement in Russia is shaped by its performance evaluation system in the context of drug control, providing incentives for police officers to manipulate the amounts of drugs seized.

Methodologically, the paper follows the rapidly growing literature on bunching, which analyzes agents' responses to various thresholds. The literature distinguishes between two different bunching designs, one of which exploits behavioral discontinuities at kink points, discrete changes in the choice sets slope, and was

developed by Saez (2010), and Chetty et al. (2011). The other type of design is based on notch points, discrete changes in the choice sets level, and was introduced by Kleven and Waseem (2013). These approaches were initially developed to study the response to tax regulation. However, increasingly they are applied in many other settings, for example education (Diamond and Persson 2016, Brehm et al. 2017), pensions (Brown 2013, Manoli and Weber 2016), social insurance (Einav et al. 2015, Le Barbanchon 2016), car speed regulation (Goncalves and Mello 2017), welfare programs (Camacho and Conover 2011), procurement (Palguta and Pertold 2017) and others. An extended review of the bunching literature can be found in Kleven (2016). My study creates a new application for bunching techniques in the context of drug control.

3 Institutional Context and Data

This section briefly discusses the institutional background, providing information on Russian anti-drug laws and the system of performance evaluation for police officers. Additionally, it describes the dataset used for the empirical analysis.

3.1 Institutional Context

The first Russian Independent Anti-Drug Agency was established in 2002. Since then it has been reorganized multiple times, and in 2004 was renamed the Russian Federal Service for Drug Control (FSKN)[3], also known as the Drug Police. The increased responsibilities of this agency included control of legal and illicit drug trafficking, prevention of drug abuse, drafting of state policy, and legal regulation. The FSKN shared concurrent jurisdiction with the Public Security Service (Police) of the Ministry of Internal Affairs (MVD)[4], but was solely responsible for coordinating

[3]Federal'naya sluzhba Rossiiskoi Federacii po kontrolyu za oborotom narkotikov, FSKN.
[4]Ministerstvo vnutrennih del Rossiiskoi Federacii, MVD.

and pursuing Russian drug investigations abroad (The Ministry of Internal Affairs of the Russian Federation, n.d.). While the main focus of the FSKN had to be on significant cases (drug trafficking, organized crime, large drug amounts), the MVD mostly dealt with routine low-profile cases, such as drug use and small-quantity drug sales. Almost two thirds of all drug related cases registered during 2013-2014 were initiated by the MVD. At the same time, the MVD provided many other public security functions, and drug control was not its only responsibility. This could explain geographical differences in the numbers of cases initiated by the two agencies. According to Knorre and Skougarevskiy (2015), the FSKN more often operated in less populated localities, taking a targeted approach to searching for drug offenders, while the MVD more often seized drugs in densely populated regions.

Anti-drug legislation. 95% of all drug crimes registered in Russia in 2013-2014 were prosecuted under articles 228 and 228.1 of the Criminal Code of the Russian Federation. The severity of a penalty under these articles depends on the amounts of drug seized, which are classified as "significant", "large" or "especially large" (Appendix, Table B1).

Article 228 imposes criminal responsibility for the illegal acquisition, storage, transportation, manufacture or processing of drugs. The punishment for drug possession on a significant scale, with no intention to sell, (the least serious crime) is imprisonment for up to three years. On large and especially large scales (serious and most serious crimes), the punishment is imprisonment for three to ten and ten to fifteen years, respectively (Appendix, Table B2). In the case of voluntary surrender of drugs to a police officer and active assistance during the investigation, an offender is exempted from criminal liability. If the amount of drug seized is less than significant, the person can only be brought to administrative responsibility punished with a fine or administrative arrest for up to 15 days.

Article 228.1 imposes criminal responsibility for the illegal manufacture, sale or dispatch of drugs, which is punishable by imprisonment for four to eight years if

the amount is less than significant, eight to fifteen years for a significant amount, and ten to twenty years for a large amount. Especially large amounts carry a fifteen to twenty years, or life, sentence (Appendix, Table B2). In this case, the crime is serious if the amount of drug seized is less than significant, and most serious if the amount is significant or higher (Knorre 2017).

The practice of plea bargaining was introduced in 2001 and became quite common in Russia. During 2013-2014 around 60% of all cases (30% of drug related offences) were adjudicated in relation to plea bargaining. Pleading guilty significantly simplifies the procedure: a conviction is pronounced without the actual examination of evidence at a court hearing. In addition, a person that accepts a plea bargain waives the right to appeal. In return, by pleading guilty the offender lowers the upper bound of the sentencing range by one third.

The performance evaluation of anti-drug agency personnel. During the 2013-2014 period, when both the FSKN and the MVD were responsible for enforcing drug laws, each had their own evaluation system. The system used by the FSKN was based on parameters that included the number of drug crimes solved, the number of serious and most serious drug crimes solved and the total amount of drugs seized (per 100 officers). For each parameter, the FSKN regional offices received a position in cross-region ratings. The final evaluation was determined by the overall rank of the office in relation to others, which was calculated as the sum of these positions.

On the other hand, the system of performance evaluation of the MVD was based on an overall score for each regional office, which was calculated as the weighted average of values of non-departmental and departmental assessments. Among parameters that characterized a police office's work were the number of crimes solved, the number of serious and most serious crimes solved and the number of serious and most serious drug crimes solved (per 100 officers). These statistics belonged to the highest-weighted group of parameters in the overall score. Crucially, in contrast to the FSKN, the MVD stations compared performance with their own

evaluation in the previous year[5] (Novikova 2014).

Regarding the consequences for police officers, there were disciplinary and financial measures that could be applied to personnel based on their performance. If the officer met or surpassed the targets, he might receive a monetary bonus to his monthly salary or even promotion (for high-profile cases). While there was no guarantee that the officer would be rewarded for good performance, he certainly was reprimanded, warned or even fired in the case of unsatisfactory performance. In addition, the officer could be deprived of monthly bonuses, in addition to the fixed salary, if he did not fulfill the plan.

Thus, the system of performance evaluation presented strong incentives for police officers to show the required level of cases and prosecutions using any tools, including manipulations of drug amounts. At the same time, even though the FSKN's system was more transparent, it was more difficult for the FSKN stations to set the "necessary" amount of manipulations, since it had to take into account the performance of other stations in the current period. In contrast, the MVD officers always knew what numbers they should reach. These institutional features could significantly contribute to the difference in the magnitudes of manipulation by these two agencies, which I investigate in more detail in Section 5.2.

3.2 Data

This paper uses a database provided by the Institute for the Rule of Law at the European University at St.-Petersburg, Russia[6]. It contains information on almost 300,000 drug crimes reported in Russia during 2013-2014. The information is based on five forms that are created at each stage of the investigation of a specific case and include the following data:

[5]Formally, after the reforms in 2011, the MVD offices had to compare their performance across units rather than relative to the previous period. However, locally this did not work due to the complexity of the system.

[6]Initial data was compiled and prepared at the Institute for the Rule of Law at the European University at St. Petersburg with support from the Russian Science Foundation grant 17-18-01618.

- form 1: identified crime and investigation results;
- form 2: socio-economic characteristics of offender;
- form 3: criminal proceedings;
- form 4: reparation for damages and the seizure of crime objects;
- form 6: trial results[7].

Form 1 is completed by an investigator when he or she decides to initiate criminal proceedings that should be approved by a prosecutor. During the investigation, forms 2, 3 and 4 are created. These forms have to be checked by the prosecutor's office before referring the case to the judicial authorities. Form 6 is filled in by a judge. After closing the case, all forms should be converted from written to electronic form and submitted to an information center (Shklyaruk and Skougarevskiy 2015).

Knorre and Skougarevskiy (2015) and Skougarevskiy (2017) extracted and analyzed all information on primary drug types, weights of drugs seized, offenders' characteristics and court decisions from this database. I follow their logic. Both forms 1 and 4 contain information on weights of drugs seized, which coincide only for 92.8% of cases. However, the distributions of cases across drug amounts do not differ significantly. I use data from form 1 to estimate the magnitude of possible manipulation and investigate to what extent it varies by drug type, article and agency. I believe that this better reveals the initial picture of how police offices work, since form 4 is created at a later stage after the prosecutor's approval of case initiation and contains drug amounts measured in the laboratory. In order to identify characteristics of victims of manipulation, I merge data from forms 1 and 2. For the investigation of manipulation consequences, I turn to combined data from forms 1, 2 and 6, merged with drug weights from form 4. I restrict the sample to

[7]Form 5 is not in the database since it should contain information on victims, while drug crimes are victimless.

cases related only to drug use for two reasons. First, separation by article is needed due to the existing specifics of different types of crime. Second, the drug dealers sample from forms 1, 2, 6 and 4 contains an insufficient number of observations.

The initial dataset based on form 1 contained data on 518,979 drug crimes including 89,152 heroin related cases. 14% of cases related to heroin were excluded from consideration because the amount of drug seized was missing[8]. Missing values are likely to be caused either by inaccurate completion of forms by police officers or by mistakes during the conversion of the forms into electronic files. Additionally, under some circumstances, a case can be initiated without drug seizure. See Table B3 in the Appendix for more information on the samples discussed in this paper and missing values. While differences in means are statistically significant for almost all factors, their values themselves are small in most cases. As expected, the documents are more complete for more serious crimes (with longer sentences), when there is conclusive evidence (being arrested under the influence of drug) or in the case of refusing to plead guilty, which leads to a full investigation, compared to the simplified procedure under the plea bargain.

4 Empirical Strategy

To study the differences in magnitudes of manipulation and how they could be related to the incentive structure, I apply the standard bunching estimator (Saez 2010, Chetty et al. 2011, Kleven and Waseem 2013).

To check that the results are insensitive to the choice of estimation parameters, I repeat the procedure described in Appendix C.1, using different polynomial orders k, values of upper bound r_u and starting points after the exclusion of the area around the first threshold. I also vary the upper point for drug weight where I cut the

[8]The form 4 dataset included information on 236,989 drug crimes out of which 50,782 were related to heroin. Due to missing drug weight, 8% of heroin related cases were also excluded from the analysis.

sample since the long tail with few observations does not contain much information. This robustness check also allows me to choose the estimation parameters that are applied in further analysis.

To estimate the mean characteristics of manipulation victims, and the impact of manipulation on sentence length and the probability of pleading guilty, I follow Diamond and Persson (2016). Their technique is based on comparison of the observed and estimated counterfactual distributions of a parameter studied. This allows me to recover the average summary statistics of offenders moved above the threshold and the causual effect of this movement on outcome variables. Appendix C.2 and C.3 provides additional technical details. The standard errors are calculated using a bootstrap.

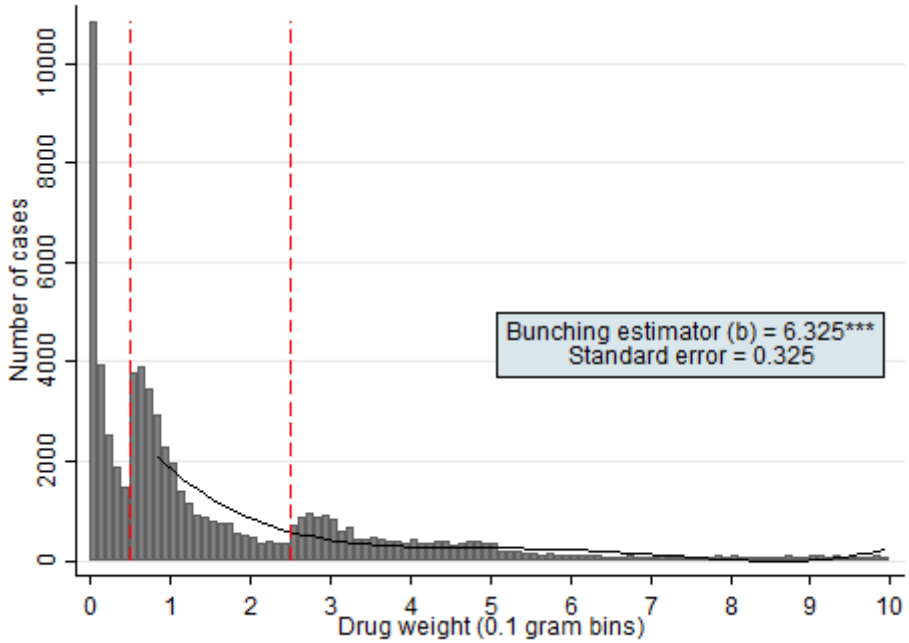
5 Results

5.1 Manipulation of Seized Drug Amounts

The analysis of distributions of cases by type of drug seized finds significant bunching around the second threshold, but only in the case of heroin. Graphs with distributions of some of the other most often seized drugs are in the Appendix, Figure A1. The bunching estimator for all heroin cases from form 1 is 6.325 (Figure 2). This means that the excess mass above the 2.5 grams threshold is almost six times greater than the average number of cases that would be in the manipulation window above the threshold in the counterfactual world without manipulation. The effect is slightly stronger in merged samples from forms 1 and 2 and forms 1, 2, 6, and 4 (Appendix, Figure A2), supporting the result observed in the initial dataset from form 1.

A sensitivity analysis shows that the value of the bunching estimator varies insignificantly (Appendix, Table B4). Therefore, I can conclude that the estimates

Figure 2: Distribution of cases across quantities of heroin seized



Note: The baseline sample consists of all heroin related cases from form 1 registered in Russia during 2013-2014. The series shown in bars is a histogram of the observed distribution of cases. The solid line is a fourth-degree polynomial fitted to the empirical distribution. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

are robust to the change of parameters. For all estimations of the mean victim characteristics and the effect of manipulation I set the upper bound of the manipulation window equal to 3.3 and the polynomial degree equal to 4. In addition to good fit, these parameters give the smallest possible estimate of bunching that does not allow overstatement of the effect of manipulation.

A simple explanation for the bunching only in the case of heroin could be that it is easier for a police officer to manipulate this drug by the small amounts needed to cross the threshold. In addition, heroin was one of the most popular and potent drugs in Russia at that time. A large share of heroin users were from the lowest socio-economic class, which made the manipulation even easier for police officers.

Another point that should be clarified is the choice of the counterfactual

distribution's form. Thinking about an offender as a rational agent would lead to a counterfactual distribution with humps just below the thresholds. Since these humps could not be estimated, I make the assumption of the counterfactual distribution with a smoothly decreasing shape. In this case, the bunching estimator yields a lower bound. Indirect evidence in support of a smoothly decreasing shape comes from the distributions of cases related to the other types of drugs (Appendix, Figure A1), which do not have bunching (at least around the second threshold).

The visual inspection of distributions determines the significant difference in manipulation magnitudes across Russian regions. However, rigorous estimation is unfeasible in this case, since splitting the sample into 83 subsamples (as many as regions) significantly reduces statistical power when estimating the region-specific extent of manipulation. Therefore, I divide all regions into only two groups, on or away from the main drug-trafficking routes[9]. Figure A3 in the Appendix shows that the magnitude of manipulation in regions along the routes is more than twice as high as that in regions away from the routes. This could be explained by the following factors. First, in regions which are on the drug-trafficking routes, the share of population that could potentially be manipulated (drug users, drug dealers) is greater. Second, police officers in these regions might be more experienced in dealing with drug related crimes. Additionally, there might be differences in incentive structures in the two drug control agencies, which is investigated in more detail in the following subsection.

5.2 Incentives for Manipulation of Seized Drug Amounts

The significant bunching determined above raises the question of what causes the police officers' response to the threshold. According to Paneyakh (2014), the main driving force for dishonest behavior is the system of performance evaluation of police

[9]Information on drug-trafficking routes is taken from the website of Russia's international news agency https://ria.ru/20100603/242406939.html. Accessed on December 1, 2018

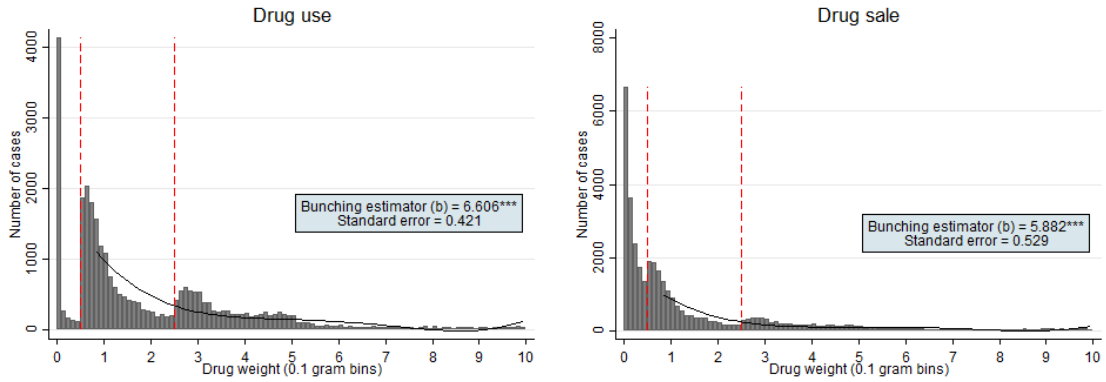
officers. To investigate this, I exploit changes in punishment across thresholds, and differences in the evaluation approaches of the two drug control agencies.

Combining the information on sanctions for drug related crimes and the systems of performance evaluation of police officers discussed in Section 3.1. suggests the following incentives for moving offenders from below a threshold to above it. In the case of drug use (article 228), crossing the first threshold changes the status of offence from administrative to criminal that improves the performance statistics, since in the MVD's system the weight for the number of (drug) crimes solved is much higher than for the number of administrative offences registered. In the FSKN system, administrative offences are not taken into account at all. In the case of drug sale (article 228.1), crossing the first threshold increases the severity of crime from serious to most serious, but does not contribute to the overall evaluation because the number of serious and most serious (drug) crimes are calculated together. I do not analyze possible manipulations around the first threshold for two reasons. First, data on offences below the first threshold could be incomplete due to police officers' reluctance to deal with cases that do not affect their performance evaluation significantly. In addition, some officers might show leniency towards minor offences and not register them. Second, the number of weight bins that could be defined below the threshold is insufficient for estimating the counterfactual distribution.

These issues do not arise in the analysis of police officers' responses to the second threshold. In the case of drug use, crossing the second threshold increases the severity of crime from least serious to serious that, in turn, positively affects the evaluation. The incentive for moving offenders from below to above the threshold in the case of drug sale is ambiguous, since manipulation does not directly contribute to performance statistics. However, it could be explained by police officers' concern about losing "points" if a drug sale case is requalified to a drug use case (for example, storage without the purpose of sale). At the same time, if the drug amount seized is large (above the second threshold), a requalification only decreases the severity of

the crime (from most serious to serious). However, that does not change the number of serious and most serious (drug) crimes solved by the police and, hence, does not worsen the performance statistics. Figure 3 presents a sharper graph and slightly higher bunching estimate for drug users (left) than for drug dealers (right), which could be explained by different incentives at the threshold. In addition, drug users are the significantly larger group of drug offenders, as well as much easier to locate and, hence, manipulate.

Figure 3: Distributions of cases related to drug use (left) and drug sale (right) across quantities of heroin seized



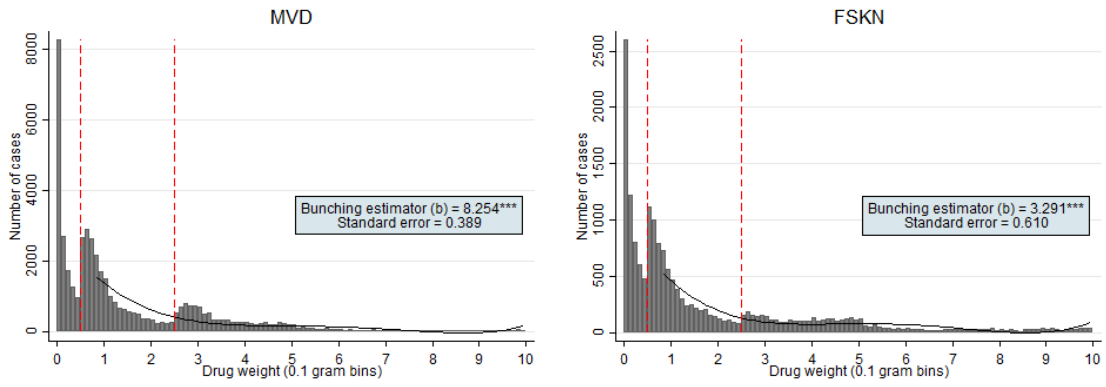
Note: The baseline sample consists of all heroin related cases from form 1 registered in Russia during 2013-2014. The series shown in bars is a histogram of the observed distribution of cases. The solid line is a fourth-degree polynomial fitted to the empirical distribution. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

At the third threshold (500 grams for heroin), which is not presented in graphs, bunching is not observed, probably due to weak incentives and (or) insufficient number of observations. Therefore, I do not explore the police officers' responses to this threshold, and even exclude the long tail from the analysis, since it does not affect the counterfactual distribution and estimates.

Turning to differences in the systems of performance evaluation, I break all heroin related cases into two groups: those initiated by the MVD and those initiated by the FSKN (Figure 4). The estimation determines a difference in the values of the

bunching estimator, significant at the 1% level. The bunching estimate for the MVD cases is 8.254, while for the FSKN cases it is only 3.291. This can be explained by the different responsibilities of two agencies and specific features of two systems of performance evaluation. In the case of the FSKN, final crime statistics are compared with the performance of other police stations and, eventually, other regions. The FSKN officers do not know the exact level that should be reached in order to obtain a satisfactory performance evaluation. Even though they can predict this level to some extent, the incentives to manipulate in the case of the FSKN are, therefore, weaker. In turn, the MVD offices compare results with their own performance in the previous period, which is well known to them. Given that the most recent performance should not be worse than previously, the performance evaluation system may incentivize some police officers to behave dishonestly, manipulating drug amounts seized in order to improve their statistics.

Figure 4: Distributions of cases initiated by the MVD (left) and the FSKN (right) across quantities of heroin seized



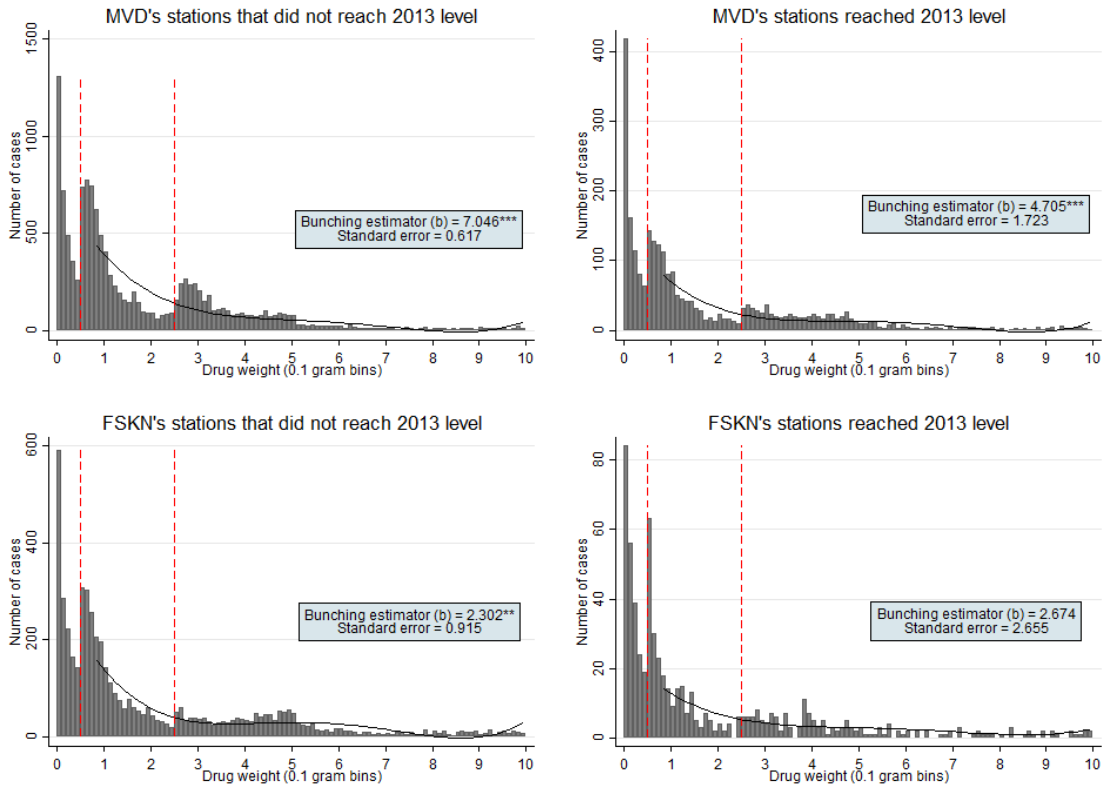
Note: The baseline sample consists of all heroin related cases from form 1 registered in Russia during 2013-2014. The series shown in bars is a histogram of the observed distribution of cases. The solid line is a fourth-degree polynomial fitted to the empirical distribution. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

As mentioned above, moving offenders from below the second threshold to above it increases the number of serious and most serious drug crimes solved, which

improves the performance evaluation. This feature, in combination with the different comparison schemes in the two agencies, can differently affect the behavior of police officers during a year. Thus, the FSKN officers might manipulate the seized drug amounts more before submitting the final report, if the regional office did not reach the number of serious and most serious drug crimes of the best performing in the previous year office (assuming that these previous results are taken into account for forecasting). In contrast, the MVD officers should manipulate more if they did not reach their own previous year's level. For the evaluation, the number of serious and most serious drug crimes is calculated per 100 officers. I do not have information on the size of each regional office, and therefore I cannot exploit the FSKN's cross-region comparison scheme[10]. Instead, assuming that the number of officers at each station is fixed during the 2013-2014 period, I determine the total absolute number of serious and most serious drug crimes solved by each station in 2013. Then, I divide all stations (by agency) into two groups: those which by the end of the third quarter of 2014 did and did not achieve their 2013 levels. Using data from the fourth quarter of 2014, I check whether the bunching varies between these four groups. As expected, the estimation shows that in the case of the MVD, the magnitude of manipulation is higher for those police stations which had not yet met their previous year's performance level, and that the difference is statistically significant (Figure 5). At the same time, the magnitude of manipulation by the FSKN police stations does not significantly depend on reaching, or not reaching the "benchmark". Thus, these results support the hypothesis that the driving force for manipulation of drug amounts is the performance evaluation system.

[10]I conducted an exploratory analysis of the effect of the cross-region comparison scheme, assuming the number of officers to be proportional to the total number of drug crimes solved by each station during each year. First, I identified the FSKN's and MVD's stations with the highest relative number of serious and most serious drug crimes solved in 2013. Second, for each agency, I split all stations into two groups: those which by the end of the third quarter of 2014 reached and did not reach the "best" level of 2013 (established by either the FSKN's or MVD's station). Finally, I calculated the bunching estimate for four groups of stations using data from the fourth quarter of 2014. The difference in magnitudes appeared to be insignificant.

Figure 5: Distributions of cases across quantities of heroin seized during the 4 quarter of 2014 by the MVD's and FSKN's police stations that by the end of the 3 quarter of 2014 did nor reach (left) and reached (right) the total number of serious and most serious drug crimes solved in 2013



Note: The baseline sample consists of all heroin related cases from form 1 registered in Russia during 2013-2014. The series shown in bars is a histogram of the observed distribution of cases. The solid line is a fourth-degree polynomial fitted to the empirical distribution. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

Bribery might be another motive for police officers to manipulate the drug amounts or for moving people in the opposite direction - from above to below the threshold. However, while there is no direct evidence, it seems logical to assume that if an offender decides to pay a bribe he does it to buy himself out of prison, not just to decrease the sentence. This means that those individuals are most likely not in the database at all and bribery cases are undetectable. At the same time, if bribery cases are not distributed uniformly across seized drug amounts, their

omission might affect the manipulation window estimation and, hence, the bunching estimate. However, as the sensitivity analysis shows, the change of parameters does not significantly influence the result.

5.3 Mean Characteristics of Possible Victims of Manipulation

There are a number of criteria that a police officer can use to select offenders to push above the threshold. My analysis begins by calculating summary statistics for the whole population of heroin offenders and for those who fall into the manipulation region. Table B5 in the Appendix shows that means are similar across these two samples, suggesting the absence of self-selection into the area around the second threshold. To determine the mean characteristics of victims of manipulation, I use the technique described in Appendix C.2 and present results in Table 1.

Table 1: Mean characteristics of possible victims of manipulation

	Eligible for manipulation	Manipulated	Difference	s.e.
Male	0.809	0.858	-0.049^{*}	0.025
Russian	0.856	0.869	-0.013	0.023
At least college	0.386	0.400	-0.014	0.029
Unemployed	0.761	0.756	0.005	0.018
Repeat offender	0.670	0.733	-0.063^{**}	0.030
Administrative offence[11]	0.076	0.059	0.017^{*}	0.010
Under the influence of drug	0.518	0.467	0.051	0.033
Under the influence of alcohol	0.013	0.016	-0.003	0.008

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The baseline sample consists of all heroin related cases from forms 1 and 2 registered in Russia during 2013-2014. Column 1 presents the predicted mean characteristic of all drug offenders who possessed an unmanipulated amount of drug which fell into the manipulation window below the threshold. Column 2 presents the predicted mean characteristic among the compliers, i.e., the offenders who were actually moved above the threshold. Column 3 tests the difference. To obtain the estimates, I apply the method described in detail in Appendix C.2.

[11] A wrongful, guilty action (omission) of a natural person or legal entity which is administratively punishable under The Code Of Administrative Offences Of The Russian Federation. This violation of the law is not serious enough to be considered criminal.

First of all, I check whether such demographics as gender and nationality affect a police officer's decision to manipulate the seized drug amounts. I find the difference in the mean shares of men among those who were eligible for manipulation but did not receive the "treatment", and those who were pushed above the threshold. However, this difference is only weakly significant. At the same time, there is no effect related to the offender being Russian.

Turning to the offender's socio-economic status, which is approximated by employment status and education level, I find that the data do not support the hypothesis of the drug amounts of low status individuals being manipulated to a greater extent. This could partly be explained by a possible corruption motive, when a police officer expects either to extract a bribe from the offender or improve his performance statistic. As the analysis shows, there is also no significant effect related to an offender being under the influence of drugs or alcohol at the moment of arrest, which, in general, should make this group of offenders a more vulnerable target. The only factor that appeared to be significant in a police officer's decision making is the offender having a previous criminal history: repeat offenders are more likely to be pushed above the threshold. This could be explained by it being easier to manipulate a person whose socio-economic characteristics are known. All in all, the results support the idea proposed in some media reports that almost any average drug offender could become the victim of manipulation.

5.4 The Effect of Manipulation on Sentence Length

A case by case comparison of heroin weights from forms 1 and 4 shows that they coincide only in 92.8% of the full sample; the weight difference for the rest of the data varies from -3586 to 1604 grams[12]. Significant deviations seem suspicious and might be the consequence of mistakes made when filling in the card or converting it into an electronic form. At the same time, observations with large discrepancies in weights

[12] Average nonzero weight difference is -1.231 gram.

are randomly distributed and, therefore, could be excluded from the analysis. Table 2 presents the results of an estimation conducted for full and restricted samples, which includes observations with absolute weight differences of less than 14 grams; this being the 95th percentile among absolute nonzero deviations.

Table 2: The effect of manipulation on sentence length and probability of pleading guilty

	Absolute difference ≤ 14		Full sample	
	Coefficient	s.e.	Coefficient	s.e.
	Panel A. Sentence length			
First stage	1.071***	0.009	0.366	2.461
ITT	1.089***	0.052	1.091***	0.054
LATE(sentence)	1.016***	0.217	2.978***	0.695
	Panel B. Pleading guilty			
LATE(plea)	−0.167	0.818	−0.484	0.430
LATE(sentence)^{plead}	0.942***	0.060	0.942***	0.065
LATE(sentence)$^{not\ plead}$	0.924***	0.131	−0.780	1.104

* p < 0.10, ** p < 0.05, *** p < 0.01

Note: The baseline sample consists of all heroin use related cases from forms 1, 2, 6 and 4 registered in Russia during 2013-2014. See the text for further details defining the subsample of observations with absolute difference in weights of less than 14 grams. Panel A presents estimates of the impact of drug weights from form 1 on drug weights from form 4 (First stage), as well as ITT effect of manipulation on the sentence length of all individuals in the manipulation region, and LATE of manipulation on the sentence length of compliers only. Panel B presents LATE of manipulation on the probability of pleading guilty and on the sentence length of those drug offenders who did and did not plea guilty. To obtain the estimates, I apply the method described in detail in Appendix C.3

The first stage effect of manipulation of heroin amounts registered by police officers in form 1 on heroin weights recorded in form 4 after the expertise is significant and shows the 1.1 grams increase of drug seized amount for individuals in the manipulation area. This implies that police officers work and manipulate in collaboration with laboratory experts. There is also a significant effect of being in the manipulation window on sentence length (ITT). However, in order to see the impact of manipulation on compliers' years of imprisonment, I divide ITT by the first stage effect and obtain LATE(sentence), which suggests a one year increase in

sentence length for individuals who were pushed above the threshold. This estimate is close to that obtained in the RDD paper by Skougarevskiy (2017). Using data on cannabis and heroin cases from Russia, he finds that the length of unconditional incarceration increases by 0.84 years when the drug weight crosses the threshold. My estimate could be higher because I focus solely on heroin cases, which might be considered to be more serious offences than cannabis related crimes.

Turning to the possible heterogeneity of the effect of manipulation, I estimate LATE for those who pleaded guilty and who did not accept a plea bargain. According to Titaev and Pozdnyakov (2012), in general, pleading guilty in Russia does not reduce the sentence significantly and even worsens the offender's situation in some cases. Nevertheless in 2013-2014, almost 60% of all cases (30% of drug related offences) were processed under a plea agreement. This quite large share could be explained by the legal illiteracy of offenders when they simply do not know how the plea bargain may influence their legal situation. Additionally, police officers could offer the agreement more forcefully if the credibility of evidence collected is in doubt as in the case of, for example, manipulation of drug amounts. In turn, a plea bargain leads to a conviction without the actual examination of evidence at a court hearing. However, my analysis shows that crossing the threshold does not increase the probability of pleading guilty (LATE(plea) in Table 2). At the same time, the difference in the effects of manipulation on the sentence length of those who accepted the plea bargain and those who did not plead guilty (LATE(sentence)^{plead} and LATE(sentence)$^{not\ plead}$) is statistically insignificant. The reason could be the quite common practice of pinning the unsolved crime on a person who is already convicted of something, and then to push this person to plead guilty to both crimes.

The estimates obtained allow me to calculate the approximate welfare loss due to manipulation of seized drug amounts by police officers. For example, in 2013-2014, the average yearly government spending on one prisoner was around \$2705[13].

[13]All amounts are expressed in U.S. dollars using the average 2013-2014 exchange rate

According to the estimation results from Section 5, there were around 3000 offenders who were moved above the threshold as a result of manipulation during the 2013-2014 period, and each of them was sentenced to an additional year in prison than they would otherwise have been. Thus, the additional annual government expenditure is $\$2705 * 1500 * 1$, which is almost $4.1 million. At the same time, the above-noted yearly spending on prisoners is only 30% of the whole budget of The Federal Penitentiary Service of Russia. Therefore, the estimate of welfare loss obtained can be considered as a lower bound. If I use an annual prison authority budget per convict of $9016, on average, for calculation, the additional pressure on the overall country budget is around $13.5 million. However, even this amount is still a minor share of the total social cost induced by manipulations, since it does not take into account that longer incarceration strengthens barriers to reintegration after release, increases the probability of recidivism and amplifies the spillover effect.

6 Conclusion

The likely failure of a war on drugs has been widely acknowledged, yet in many countries anti-drug policies are still based on harsh law enforcement and even militarization. This leads to inefficient budget spending and unequal treatment of different groups of drug offenders, with a strong focus on drug users. Russia is a particularly notable example. A recently published report on drug crimes in Russia (Knorre 2017) illuminates revealing statistics on the distribution of criminal cases across quantities of heroin seized. These statistics suggest the bunching of offenders who were arrested with an amount of drugs just above the threshold sufficient to be convicted for a more serious crime. At the same time, there is a missing mass of cases just below the threshold. This might be evidence of manipulation of drugs quantities seized by the police, which so far has only been alleged by various media

(RUB/USD = 35.158).

reports.

This paper provides an empirical analysis of possible manipulation of amounts of drugs seized using a unique dataset that contains rich information on drug crimes reported in Russia during 2013-2014. Exploiting the specific features of the Russian institutional context, I show the importance of incentives from performance evaluation in driving the manipulations by police officers. Additionally, the results suggest that individuals with a criminal history are more likely to have their drug amounts manipulated by the police. The overall effect of this manipulation on sentence length is an additional year of incarceration, which is not dependent on a guilty plea. Finally, I calculate the additional pressure on the overall country budget of around $13.5 million. However, this is a lower bound of the total social cost induced by manipulations, since it does not take into account that longer incarceration strengthens barriers to reintegration after release, increases the probability of recidivism and amplifies the spillover effect.

This paper clearly shows the inefficiency of existing performance evaluation system and motivation scheme and raises a serious question of optimal incentive structure. Forecasting expected results itself is a common practice in many public organizations; this provides guidance for coming period. However, the way it is implemented could become an issue as in the case of drug control in Russia. The fact that police officers are punished for not achieving their targets without taking into account any reasons for this failure, may trigger their dishonest behavior and manipulation of drug amounts. Thus, one step on the way to efficiency could be to decentralize the performance evaluation system and give regional offices some flexibility in setting the performance requirements.

References

Abrams, D. S., M. Bertrand, and S. Mullainathan (2012). Do Judges Vary in Their Treatment of Race? *Journal of Legal Studies 41*, 347–383.

Alesina, A. and E. La Ferrara (2014). A Test of Racial Bias in Capital Sentencing. *American Economic Review 104*, 3397–3433.

Anbarci, N. and J. Lee (2014). Detecting Racial Bias in Speed Discounting: Evidence from Speeding Tickets in Boston. *International Review of Law and Economics 38*, 11–24.

Anwar, S. and H. Fang (2015). Testing for Racial Prejudice in the Parole Board Release Process: Theory and Evidence. *Journal of Legal Studies 44*, 1–37.

Arnold, D., W. Dobbie, and C. S. Yang (2017). Racial Bias in Bail Decisions. National Bureau of Economic Research Working Paper 23421.

Banerjee, A., R. Chattopadhyay, E. Duflo, D. Keniston, and N. Singh (2012). Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy and Training. National Bureau of Economic Research Working Paper 17912.

Bjerk, D. (2005). Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion Under Mandatory Minimum Sentencing. *Journal of Law and Economics 48*, 591–627.

Bjerk, D. (2017). Mandatory Minimums and the Sentencing of Federal Drug Crimes. IZA Discussion Paper 10544.

Brehm, M., S. A. Imberman, and M. F. Lovenheim (2017). Achievement Effects of Individual Performance Incentives in a Teacher Merit Pay Tournament. *Labour Economics 44*, 133–150.

Brown, K. M. (2013). The Link Between Pensions and Retirement Timing: Lessons from California Teachers. *Journal of Public Economics 98*, 1–14.

Camacho, A. and E. Conover (2011). Manipulation of Social Program Eligibility. *American Economic Journal: Economic Policy 3*, 41–65.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *Quarterly Journal of Economics 126*, 749–804.

Diamond, R. and P. Persson (2016). The Long-Term Consequences of Teacher Discretion in Grading of High-Stakes Tests. National Bureau of Economic Research Working Paper 22207.

Einav, L., A. Finkelstein, and P. Schrimpf (2015). The Response of Drug Expenditure to Nonlinear Contract Design: Evidence from Medicare Part D. *Quarterly Journal of Economics 130*, 841–899.

Goel, S., J. M. Rao, and R. Shroff (2016). Precinct or Prejudice? Understanding Racial Disparities in New York City's Stop-and-Frisk Policy. *Annals of Applied Statistics 10*, 365–394.

Goncalves, F. and S. Mello (2017). A Few Bad Apples? Racial Bias in Policing. Princeton University Working Paper 608.

Holmstrom, B. and P. Milgrom (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization 7*, 24.

Institute for Criminal Policy Research (n.d.). *Highest to Lowest - Prison Population Rate*. `http://www.prisonstudies.org/highest-to-lowest/prison_population_rate?field_region_taxonomy_tid=All`. Accessed January 1, 2019.

International Drug Policy Consortium (2018). *A Civil Society Shadow Report "Taking Stock: A Decade of Drug Policy"*. http://fileserver.idpc.net/library/Shadow_Report_FINAL_ENGLISH.pdf. Accessed January 30, 2019.

Johnson, B. D., S. Van Wingerden, and P. Nieuwbeerta (2010). Sentencing Homicide Offenders in the Netherlands: Offender, Victim, and Situational Influences in Criminal Punishment. *Criminology 48*, 981–1018.

Kleven, H. (2016). Bunching. *Annual Review of Economics 8*, 435–464.

Kleven, H. J. and M. Waseem (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *Quarterly Journal of Economics 128*, 669–723.

Knorre, A. (2017). *Narkoprestupleniya v Rossii: Analiz Sudebnoi i Kriminal'noi Statistiki*. Technical report, The Institute for the Rule of Law at the European University at St. Petersburg. http://enforce.spb.ru/images/Knorre_Drug_crimes_in_Russia.pdf. Accessed October 1, 2017.

Knorre, A. and D. Skougarevskiy (2015). *Kak MVD i FSKN Borjutsja s Narkotikami: Sravnitel'nyj Analiz Rezul'tativnosti Dvuh Vedomstv*. Technical report, The Institute for the Rule of Law at the European University at St. Petersburg. http://enforce.spb.ru/images/analit_zapiski/FSKN_MVD_memo_2015_web.pdf. Accessed October 1, 2017.

Knowles, J., N. Persico, and P. Todd (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy 109*, 203–229.

Kurmangaliyeva, M. (2017). Criminal Justice and Wealth Inequality: How Much Freedom Can Money Buy in Russia? Available at SSRN https://ssrn.com/abstract=2816363. Accessed October 1, 2017.

Le Barbanchon, T. (2016). Optimal Partial Unemployment Insurance: Evidence From Bunching in the US. Bocconi University Working Paper.

Lee, M., J. T. Ulmer, and M. Park (2011). Drug Sentencing in South Korea: The Influence of Case-Processing and Social Status Factors in an Ethnically Homogeneous Context. *Journal of Contemporary Criminal Justice 27*, 378–397.

Legewie, J. (2016). Racial Profiling and Use of Force in Police Stops: How Local Events Trigger Periods of Increased Discrimination. *American Journal of Sociology 122*, 379–424.

Liu, J., D. Zhou, A. E. Liska, S. F. Messner, M. D. Krohn, L. Zhang, and Z. Lu (1998). Status, Power, and Sentencing in China. *Justice Quarterly 15*, 289–300.

Manoli, D. and A. Weber (2016). Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions. *American Economic Journal: Economic Policy 8*, 160–182.

Mas, A. (2006). Pay, Reference Points, and Police Performance. *The Quarterly Journal of Economics 121*(3), 783–821.

Mechoulan, S. and N. Sahuguet (2015). Assessing Racial Disparities in Parole Release. *Journal of Legal Studies 44*, 39–74.

Novikova, A. (2014). Sistema Ocenki Policii. Moscow. Public Verdict Foundation. Paper presented at International conference "Human rights and law enforcement agencies". `http://publicverdict.org/articles_images/12001_60506_reform_police.pdf`. Accessed October 1, 2017.

Palguta, J. and F. Pertold (2017). Manipulation of Procurement Contracts: Evidence from the Introduction of Discretionary Thresholds. *American Economic Journal: Economic Policy 9*, 293–315.

Paneyakh, E. (2014). Faking Performance Together: Systems of Performance Evaluation in Russian Enforcement Agencies and Production of Bias and Privilege. *Post-Soviet Affairs 30*(2-3), 115–136.

Persico, N. (2009). Racial Profiling? Detecting Bias Using Statistical Evidence. *Annual Review of Economics 1*, 229–254.

Prendergast, C. (2001). Selection and Oversight in the Public Sector, With the Los Angeles Police Department as an Example. National Bureau of Economic Research Working Paper 8664.

Rehavi, M. M. and S. B. Starr (2014). Racial Disparity in Federal Criminal Charging and its Sentencing Consequences. *Journal of Political Economy 6*, 1320–1354.

Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy 2*, 180–212.

Shklyaruk, M. and D. Skougarevskiy (2015). *Kriminal'naya Statistika: Mehanizmy Formirovaniya, Prichiny Iskazheniya, Puti Reformirovaniya*. Technical report, The Institute for the Rule of Law at the European University at St. Petersburg. `http://enforce.spb.ru/images/Products/Crimestat_report_2015_IRL_KGI_web.pdf`. Accessed October 1, 2017.

Skougarevskiy, D. (2017). *What Do Graduated Sanctions Tell Us About the Functions of the Law: A case of Drug Crimes*. Ph. D. thesis, Graduate Institute of International and Development Studies.

The Federal Penitentiary Service of Russia (2017). *Characteristics of Prison Population*. `http://fsin.su/structure/inspector/iao/statistika/Xar-ka%20lic%20sodergahixsya%20v%20IK/`. Accessed October 1, 2017.

The Ministry of Internal Affairs of the Russian Federation (n.d.). *The History of*

the Ministry of Internal Affairs of the Russian Federation. https://xn--b1aew.xn--p1ai/history. Accessed October 1, 2017.

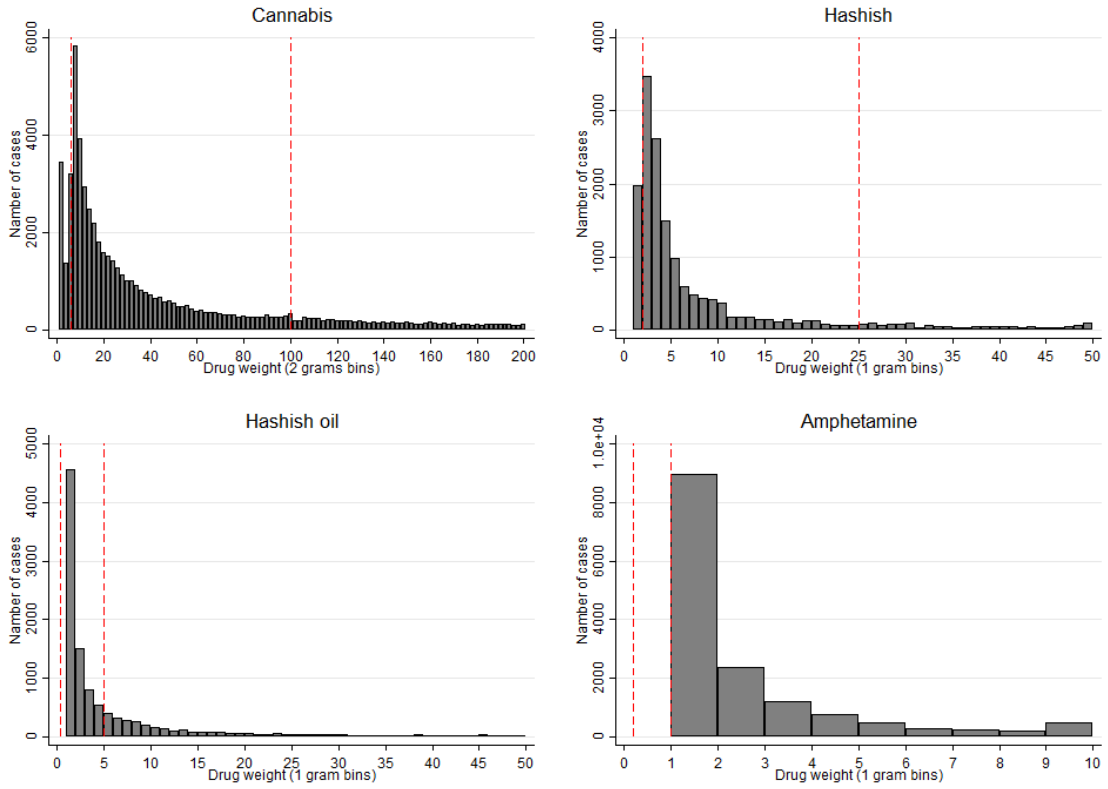
Titaev, K. and M. Pozdnyakov (2012). *Poryadok Osobyi - Prigovor Obychnyi: Praktika Primeneniya Osobogo Poryadka Sudebnogo Razbiratel'stva (gl. 40 UPK RF) v Rossiiskih Sudah*. Technical report, The Institute for the Rule of Law at the European University at St. Petersburg. http://www.enforce.spb.ru/images/analit_zapiski/pm_gl_40_UPK_fin.pdf. Accessed January 30, 2019.

Ulmer, J. T., M. C. Kurlychek, and J. H. Kramer (2007). Prosecutorial Discretion and the Imposition of Mandatory Minimum Sentences. *Journal of Research in Crime and Delinquency 44*, 427–458.

Volkov, V. (2016). Legal and Extralegal Origins of Sentencing Disparities: Evidence from Russia's Criminal Courts. *Journal of Empirical Legal Studies 13*, 637–665.

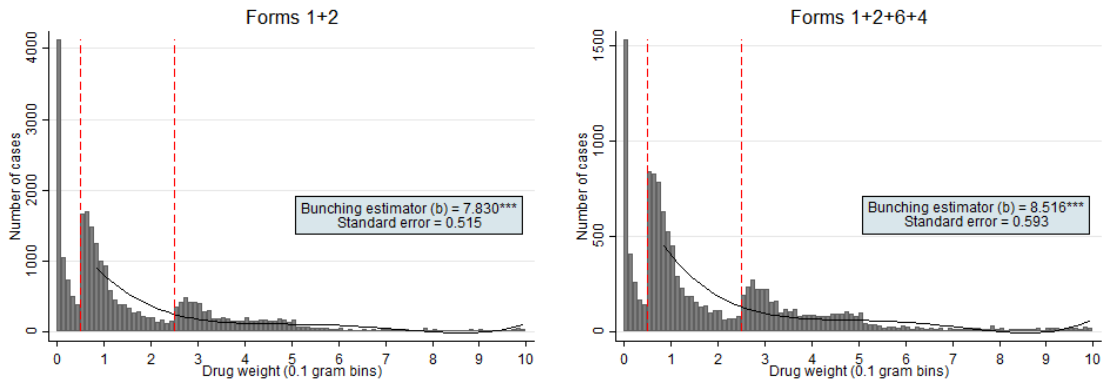
A Supplemental Figures

Figure A1: Distributions of cases across quantities of drugs seized by drug type



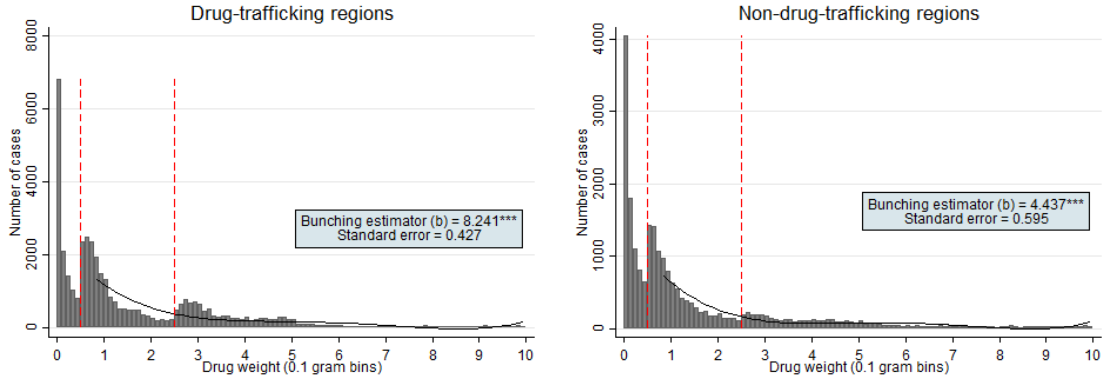
Note: The baseline sample consists of all drug related cases from form 1 registered in Russia during 2013-2014. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

Figure A2: Distributions of cases from forms 1, 2 (left) and forms 1, 2, 6, 4 (right) across quantities of heroin seized in Russia during 2013-2014



Note: The series shown in bars is a histogram of the observed distribution of cases. The solid line is a fourth-degree polynomial fitted to the empirical distribution. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

Figure A3: Distributions of cases across quantities of heroin seized in regions which are along (left) or away from (right) the main drug-trafficking routes



Note: The baseline sample consists of all heroin related cases from form 1 registered in Russia during 2013-2014. The series shown in bars is a histogram of the observed distribution of cases. The solid line is a fourth-degree polynomial fitted to the empirical distribution. The thresholds (dashed lines) determine the scale of seizure (less than significant, significant and large drug amounts), the severity of crime and punishment.

B Supplemental Tables

Table B1: Amounts of drugs (grams above) for purposes of articles 228, 228.1 of the Criminal Code of Russian Federation

	Significant	Large	Especially large
Cannabis	6	100	100000
Heroin	0.5	2.5	500
Amphetamine	0.2	1	200
Papaver	20	500	100000
Desomorphine	0.05	0.25	10
Hashish	2	25	10000
Cocaine	0.5	5	1500

Table B2: The severity of offence and sanctions according to articles 228, 228.1 of the Criminal Code of Russian Federation

Drug amount	Article 228 (use)		Article 228.1 (sale)	
	Severity	Sentence (years)	Severity	Sentence (years)
Less than significant	Administrative offence	Fine/15 days	Serious	4-8
Significant	Least serious	0-3	Most serious	8-15
Large	Serious	3-10	Most serious	10-20
Especially large	Most serious	10-15	Most serious	15-20

Table B3: The comparison of means within the missing values analysis

	Form 1			Form 4			Forms 1+2			Forms 1+2+6+4		
	(1)	(0)	(1) - (0)	(1)	(0)	(1) - (0)	(1)	(0)	(1) - (0)	(1)	(0)	(1) - (0)
Initiated by the MVD	0.682	0.643	0.039***	0.749	0.796	−0.047***	0.753	0.817	−0.064***	0.794	0.942	−0.148***
Initiated by the FSKN	0.314	0.355	−0.040***	0.250	0.204	0.046***	0.245	0.182	0.063***	0.205	0.058	0.147***
Article 228 (use)	0.447	0.364	0.083***	0.530	0.440	0.090***	0.727	0.670	0.057***	0.837	0.793	0.045***
Article 228.1 (sale)	0.550	0.592	−0.042***	0.467	0.537	−0.070***	0.269	0.291	−0.022***	0.160	0.187	−0.027***
Male							0.814	0.837	−0.023***	0.822	0.844	−0.022**
Russian							0.851	0.895	−0.044***	0.876	0.920	−0.045***
At least college							0.381	0.350	0.031***	0.392	0.325	0.068***
Unemployed							0.784	0.810	−0.026***	0.772	0.828	−0.056***
Student							0.0001	0.0005	-0.0004	0.0001	0.0007	−0.0006*
Worker							0.182	0.150	0.032***	0.196	0.133	0.063***
White collar							0.014	0.010	0.004*	0.015	0.014	0.001
Repeat offender							0.672	0.584	0.088***	0.669	0.565	0.103***
Administrative offence							0.067	0.025	0.042***	0.070	0.026	0.045***
Under the influence of drug							0.509	0.354	0.155***	0.536	0.358	0.178***
Under the influence of alcohol							0.014	0.014	0.000	0.016	0.016	0.000
Sentence length										3.009	2.417	0.592***
Pleaded guilty										0.346	0.604	−0.258***
Observations	76,735	12,417		46,593	4,189		30,728	4,268		14,350	1,516	

* p < 0.10, ** p < 0.05, *** p < 0.01

Note: The table compares means in four samples used in the analysis: from form 1, form 4, forms 1 and 2, and forms 1, 2 and 6 merged with weights from form 4. The samples consist of all heroin related cases registered in Russia during 2013-2014. Columns (1) present means in the subsamples without observations with missing drug weight, columns (0) present means in the subsamples of observations with missing drug weight, columns (1) - (0) shows differences in means.

Table B4: Robustness check

Starting point	Polynomial degree k	Manipulation window		Bunching estimator b	s.e.
		Lower bound r_l	Upper bound r_u		
0.7	4	1.5	3.3	7.463***	0.305
0.7	4	1.5	3.4	8.879***	0.377
0.7	4	1.5	3.5	10.828***	0.526
0.7	5	1.1	3.3	7.057***	0.415
0.7	5	1.1	3.4	8.123***	0.509
0.7	5	1.1	3.5	9.515***	0.691
0.8	**4**	**1.6**	**3.3**	**6.325*** **	**0.325**
0.8	4	1.6	3.4	7.256***	0.319
0.8	4	1.6	3.5	8.453***	0.398
0.8	5	1.2	3.3	7.885***	0.579
0.8	5	1.2	3.4	9.165***	0.766
0.8	5	1.2	3.5	10.851***	0.956
0.9	4	1.7	3.3	6.398***	0.246
0.9	4	1.7	3.4	6.787***	0.277
0.9	4	1.7	3.5	7.421***	0.309
0.9	4	1.3	3.3	7.459***	0.473
0.9	4	1.3	3.4	8.466***	0.537
0.9	4	1.3	3.5	9.751***	0.572

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The baseline sample from form 1 consists of all heroin related cases registered in Russia during 2013-2014. To obtain the estimates of bunching, I apply the method described in detail in Appendix C.1.

Table B5: Summary statistics

	Overall	Manipulation region
Male	0.814	0.826
Russian	0.851	0.859
At least college	0.381	0.400
Unemployed	0.784	0.748
Student	0.0001	0.000
Worker	0.182	0.210
White-collar	0.014	0.019
Repeat offender	0.672	0.713
Administrative offence	0.067	0.060
Under the influence of drugs	0.509	0.541
Under the influence of alcohol	0.014	0.017
Observations	30,728	5,026

Note: The baseline sample from forms 1 and 2 consists of all heroin related cases registered in Russia during 2013-2014. See the text for further details defining the subsample around the threshold.

C Estimation Details

C.1 Detecting Manipulation of Seized Drug Amounts

First, I estimate the magnitude of the response of police officers around the crime severity threshold. I quantify this effect, adapting the standard method from the bunching literature (Saez 2010, Chetty et al. 2011, Kleven and Waseem 2013).

To obtain the bunching estimator, I estimate the counterfactual density of seized drug amounts by fitting a high-order polynomial to the observed distribution, excluding the region$[r_l, r_u]$ around the threshold $\overline{D}$:

$$C_j = \sum_{k=0}^{p} \beta_k R_j^k + \sum_{r=r_l}^{r_u} \gamma_r * \mathbb{1}[R_j = r] + \nu_j, \tag{1}$$

where C_j is the number of cases i in bin j, k is the order of the polynomial, R_j is the midpoint of bin j. For heroin related cases, bin size is set to 0.1 gram, which is approximately the smallest dose that can be bought. To obtain the counterfactual distribution I estimate the predicted values from (1), omitting the γ_r shifters for smoothing the density around the threshold:

$$\hat{C}_j = \sum_{k=0}^{p} \hat{\beta}_k R_j^k. \tag{2}$$

Comparing the counterfactual and observed distributions, I can estimate the missing mass to the left of the threshold, and the excess bunching mass to the right of the threshold:

$$\hat{M} = \sum_{j=r_l}^{\overline{D}} (\hat{C}_j - C_j) \text{ and } \hat{B} = \sum_{j=\overline{D}}^{r_u} (C_j - \hat{C}_j). \tag{3}$$

To determine the lower and upper bounds of the excluded interval, I follow Kleven and Waseem (2013). Because the excess bunching above the threshold is quite sharp (compared to the missing mass), the upper bound can be determined visually. With r_u fixed I set the lower bound r_l such that $\hat{B} = \hat{M}$.

Finally, I can obtain a bunching estimate for the magnitude of manipulation, calculating the ratio of excess mass to the average height of the counterfactual density above the threshold:

$$\hat{b} = \frac{\hat{B}}{\sum_{j=\overline{D}}^{r_u} \hat{C}_j / N}, \tag{4}$$

where N is the number of bins in the interval $[\overline{D},\ r_u]$.

Since the paper studies the rational response of the police only around the second threshold, I exclude the area around first threshold from estimation.

C.2 Identifying Victims of Manipulation

In order to recover the characteristics of those who were manipulated by the police, I adopt the technique designed by Diamond and Persson (2016).

First, I estimate the counterfactual expected values of observable characteristic Y at any drug quantity bin R inside the manipulation area, using offenders outside of this area, if there was no manipulation:

$$Y_j = \sum_{k=0}^{p} \beta_k R_j^k + \varepsilon_j, \tag{5}$$

where $R_j < \overline{D} - r_l$ or $R_j > \overline{D} + r_u$. Then I can calculate the observed average values of characteristic Y for offenders inside the manipulation region below ($\overline{Y}^{never}$) and above ($\overline{Y}^{up}$) the threshold $\overline{D}$:

$$\overline{Y}^{never} = \frac{1}{N^{never}} \sum_i Y_i, \text{ where } \overline{D} - r_l \leq r_i < \overline{D}, \tag{6}$$

$$\overline{Y}^{up} = \frac{1}{N^{up}} \sum_i tY_i, \text{ where } \overline{D} \leq r_i \leq \overline{D} - r_u. \tag{7}$$

Here $\overline{Y}^{never}$ is the average characteristic of those offenders who were arrested with the amount of drug just below the threshold and were not selected for manipulation ("never-takers"):

$$\overline{Y}^{never} = \frac{N^{down}}{N^{down} - N^{compliers}} \overline{Y}^{down} - \frac{N^{compliers}}{N^{down} - N^{compliers}} \overline{Y}^{compliers}. \tag{8}$$

Accordingly, $\overline{Y}^{up}$ is the average characteristic of all those offenders who were manipulated ("compliers") and who actually were arrested with a drug amount just above the threshold ("always-takers"):

$$\overline{Y}^{up} = \frac{N^{always}}{N^{always} - N^{compliers}} \overline{Y}^{always} - \frac{N^{compliers}}{N^{always} - N^{compliers}} \overline{Y}^{compliers}. \tag{9}$$

Using estimates from (2) and (5), I can obtain values of $\overline{Y}^{down}$ and $\overline{Y}^{always}$ in the following way:

$$\overline{Y}^{down} = \frac{\int_{\overline{D}-r_l}^{\overline{D}-\sigma} \hat{Y}_j^R \hat{C}_j^R dR}{N^{down}} \tag{10}$$

$$\overline{Y}^{always} = \frac{\int_{\overline{D}}^{\overline{D}+r_u} \hat{Y}_j^R \hat{C}_j^R dR}{N^{always}}. \tag{11}$$

The number of offenders in each part of the manipulation region can be calculated as:

$$N^{never} = N^{down} - N^{compliers}, \text{ where } N^{down} = \int_{\overline{D}-r_l}^{\overline{D}-\sigma} \hat{C}_j{}^R dR, \tag{12}$$

$$N^{up} = N^{always} + N^{compliers}, \text{ where } N^{always} = \int_{\overline{D}}^{\overline{D}+r_u} \hat{C}_j{}^R dR. \tag{13}$$

Plugging these into (8) and (9) and using estimates from (6), (7), (10) and (11), I solve for the compliers' average value of characteristic Y:

$$\begin{aligned}\overline{Y}^{compliers} = 0.5(&\frac{N^{never}}{N^{never} - N^{down}}\overline{Y}^{never} - \frac{N^{down}}{N^{never} - N^{down}}\overline{Y}^{down})+ \\ &+ 0.5(\frac{N^{up}}{N^{up} - N^{always}}\overline{Y}^{up} - \frac{N^{always}}{N^{up} - N^{always}}\overline{Y^{always}}).\end{aligned} \tag{14}$$

Finally, I can compare the mean characteristics of those offenders who were manipulated by the police ("compliers") with the mean characteristics of all offenders who were "eligible" for manipulation but did not receive it ("never-takers"):

$$\Delta Y = \overline{Y}^{never} - \overline{Y}^{compliers}. \tag{15}$$

C.3 Estimating the Effect of Manipulation on Sentence Length

I identify the effect of manipulation of drug quantities on sentence length (and on the probability of pleading guilty) in two steps, following Diamond and Persson (2016).

First, I estimate the relationship between sentence length S and the amount of drug seized from form 1:

$$S_j = \sum_{k=0}^{p} \beta_k R_j^k + \gamma_R * \mathbb{1}[R_j \geq \overline{D}] + \omega_j, \tag{16}$$

where $R_j < \overline{D} - r_l$ or $R_j > \overline{D} + r_u$. Equation (16) gives the expected length of sentence at each drug amount inside the manipulation region in the counterfactual world where no offender is manipulated.

Then, I calculate the counterfactual expected sentence length across the whole set of drug offenders inside the manipulation region:

$$\overline{S} = \int_{\overline{D}-r_l}^{\overline{D}+r_u} \hat{S}_j \frac{\hat{C}_j^{\,R}}{\int_{\overline{D}-r_l}^{\overline{D}+r_u} \hat{C}_j^{\,R}} dR. \tag{17}$$

Comparing observed and estimated counterfactual average sentence lengths, I obtain the "intent-to-treat" effect, which shows a change in the length of imprisonment due to the offender having been caught with the actual amount of drug that falls within the manipulation region:

$$ITT = \frac{\sum_{i \in manip\, region} S_i}{N^{manip}} - \overline{S}, \tag{18}$$

where N^{manip} is the number of offenders in the manipulation area.

The procedure described above can be repeated with drug quantities from form 4 instead of sentence length. This constitutes the effect of being manipulated on the amount of drug seized that is determined officially at the laboratory and then considered by judge at court. The ratio of ITT from equation (18) to this effect, in turn, identifies the local average treatment effect (LATE) of being manipulated on the length of imprisonment.

Abstrakt

Tento článek předkládá empirickou analýzu možných manipulací s množstvím zabavených drog. Analýza je založena na unikátních datech, která obsahují kompletní informace o drogové kriminalitě v Rusku za období 2013-2014. V první části analýzy využívám standardní odhad shlukování (bunching estimator) ke zkoumání motivace policistů k manipulaci. Zjišťuji, že motivace nejpravděpodobněji pramení ze systému hodnocení výkonnosti policistů. V druhé části aplikuji novou techniku odhadu shlukování a zjišťuji, že policisté častěji manipulují s množstvím zabavených drog od pachatelů, kteří se již dříve dopustili drogové kriminality. Celkový vliv manipulace se projevuje jedním dodatečným rokem odnětí svobody a je nezávislý na přiznání pachatele.

Working Paper Series
ISSN 1211-3298
Registration No. (Ministry of Culture): E 19443

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

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

Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: http://www.cerge-ei.cz

Editor: Byeongju Jeong

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

ISBN 978-80-7343-444-1 (Univerzita Karlova, Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-493-8 (Národohospodářský ústav AV ČR, v. v. i.)