Title
Essays on Firms and Institutions

Author
Korovkin, Vasily

Publication Date
2018

University of California
Los Angeles

Essays on Firms and Institutions

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Management

by

Vasily Korovkin

2018

Abstract of the Dissertation

Essays on Firms and Institutions

by

Vasily Korovkin

Doctor of Philosophy in Management

University of California, Los Angeles, 2018

Professor Romain T. Wacziarg, Co-Chair

Professor Nico Voigtländer, Co-Chair

My dissertation contributes towards our understanding of firm behavior in weakly institutionalized environments. It consists of three chapters. The first, "Detecting Auctioneer Corruption: Evidence from Russian Procurement Auctions", develops a novel method for detecting auctioneer corruption in first-price sealed-bid auctions. I study the leakage of bid information by the auctioneer to a preferred bidder. I construct a formal test for the presence of bid-leakage corruption and apply it to a novel data set of 4.3 million procurement auctions in Russia that occurred between 2011 and 2016. With bid leakage, the preferred bidder gathers information on other bids and waits until the end of the auction to place a bid. Such behavior creates an abnormal correlation between winning and being (chronologically) the last bidder. Informed by this fact, I build several measures of corruption. I document that more than 10% of the auctions were affected by bid leakage. My results imply that the value of the contracts assigned through these auctions was $1.2 billion over the six-year study period. I build a model of bidding behavior to show that corruption exerts two effects on the expected prices of the contracts. The direct effect inflates the price of the contract. The indirect effect reduces the expected price since honest bidders are trying to undercut corrupt bidders. I find both effects in the data, with the direct effect being more

pronounced.

My second chapter, "Collusion in Auctions: Evidence from the Timing of Bids", documents collusion between firms using a unique feature of the same Russian procurement data: the timestamps of all bids. Timestamp data allows developing a new method of collusion detection based on the excessive share of simultaneous bids. My method shows that 8–23% of winner-runner-up pairs bid together, which provides a bound on the share of collusive auctions. Next, I document that simultaneous bidding is correlated with higher procurement prices and smaller bid margins in the auctions. We include a battery of controls to state that collusion leads to 8–9% increase in the final price of the contracts and makes joint bids up to 50% closer to each other. The chapter is the first to show how one can enhance methods of collusion detection by using the data on the timing of bids.

In the third and last chapter, I study the effects of armed conflict on trade transactions between firms. The chapter examines trade in the aftermath of the Russian-Ukrainian conflict (2014). The geographic concentration of fighting in a few regions allows me to study the indirect effects of conflict on trade, as opposed to the direct effects of violence or trade embargoes. I employ a highly granular transaction-level dataset for the universe of import and export transactions in Ukraine and find that firms from more ethnolinguistically Ukrainian counties experienced a deeper drop in trade with Russia relative to the firms in more Russian counties. The richness of panel data allows looking beyond explanations unrelated to ethnicities, such as increased transportation costs and bans on certain products. Instead, I focus on two ethnic-specific explanations: a rise in animosity and a decrease in trust. In a stylized model of trade with asymmetric information, I show that one can distinguish these two mechanisms based on whether the effect is more pronounced for homogeneous or non-homogeneous goods, the latter pointing to the trust mechanism. The intuition is that trust mitigates the uncertainty behind goods' quality. Empirically, I show that in contrast to homogeneous goods, the trade of relation-specific goods has not changed differentially across ethnic lines. Hence, I find little evidence in support of a shock to trust. I then use survey data to show that inter-ethnic animosity has indeed escalated in the aftermath of the conflict.

The dissertation of Vasily Korovkin is approved.

John William Asker

Christian Dippel

Nico Voigtländer, Committee Co-Chair

Romain T. Wacziarg, Committee Co-Chair

University of California, Los Angeles

2018

To My Family

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am especially indebted to my Committee Chairs Nico Voigtländer and Romain Wacziarg, for their invaluable support and guidance throughout my PhD studies. I am extremely thankful for the support and advice from my other committee members John Asker and Christian Dippel. I am also grateful to Leonardo Bursztyn, Lorenzo Casaburi, Ernesto Dal Bo, Georgy Egorov, Stefano Fiorin, Gaston Illanes, Seema Jayachandran, Kei Kawai, Lidia Kosenkova, Aprajit Mahajan, Ricardo Perez-Truglia, Nicola Persico, Robert Porter, Nancy Qian, Melanie Wasserman, Noam Yuchtman, and the participants of the Brown Bag at the Anderson School of Management, DEVPEC2016, GEM-BPP workshop, RSSIA2016, NEUDC2016, Northwestern University IO lunch, IIOC2017, MWIEDC2017, and TADC2017, and seminar participants at the University of Wisconsin-Madison, New Economic School, and CERGE-EI Prague for helpful comments and suggestions. I gratefully acknowledge financial support from the UCLA Anderson Center for Global Management. All remaining errors are mine. The first chapter is a version of a working paper with the same title co-authored with Pasha Andreyanov and Alec Davidson (both from UCLA). The third chapter is a version of a working paper with the same title co-authored with Alexey Makarin from Northwestern University; in preparation for publication. For both co-authored chapters, Vasily and other co-authors were co-PIs.

Vita

Education

2012	MA, New Economic School, Moscow
2010	Bachelor, Economics, Moscow State University

Awards

2016-2017	UCLA Dissertation Year Fellowship
2016	Russell Sage Foundation Award
2013-2016	Center for Global Management Grants
2012-2016	UCLA Anderson Fellowship
2010-2012	New Economic School Fellowship

CHAPTER 1

Detecting Auctioneer Corruption: Evidence from Russian Procurement Auctions

1.1 Introduction

An extensive empirical and theoretical literature is dedicated to detecting *collusion* in auctions.[1] In contrast, the literature on detecting *corruption* in auctions—unsavory agreements between bidders and an auctioneer—is virtually nonexistent. Extensive evidence shows that in the context of public procurement—particularly in emerging economies—corruption can exert a strong effect on the inefficient use of public funds (Kenny, 2007; Olken and Pande, 2012).[2] Corruption in procurement can take many forms. So far there has been no systematic evidence on the exact channel by which corruption leads to inefficient allocation of public funds in procurement. Since procurement auctions are a common tool by which public projects are assigned to firms, they are also a natural candidate for such channel.

This chapter examines an endemic type of corruption in procurement—"bid

[1] See Marshall and Marx (2012), Whinston (2008), and Porter (2005) for cases of collusion and methods of detection.

[2] The 2013 OECD "Corruption in Procurement" brochure (http://www.oecd.org/gov/ethics/Corruption-in-Public-Procurement-Brochure.pdf) states that in OECD countries public procurement accounts for 12% of GDP and 29% of government expenditure, totaling 4.2 trillion euros. In Russia (http://www.interfax.ru/business/499881), it was 25% of the 2016 GDP. As the OECD Foreign Bribery Report (http://www.oecd.org/corruption/oecd-foreign-bribery-report-9789264226616-en.htm) shows, more than half of foreign bribery cases concerned obtaining a government contract. Furthermore, Transparency International claims that project costs can, as a result, increase as much as 50%.

leakage" in first-price sealed-bid auctions. Under bid leakage, an auctioneer forms an agreement with one of the bidders to reveal information on all other bids. The preferred bidder can then decide how to place her bid. In the public procurement setting, the auctioneer is a government official who represents a public body, and the bidders are representatives of firms bidding for the contract. We start by developing a method to detect and measure the extent of bid leakage. At the core of our approach is the information on the timing of bids. We apply our method to the universe of first-price sealed-bid procurement auctions held by Russian public bodies from 2011 to 2016. We collected this data set from Russia's centralized official procurement website.[3] The data cover small procurement contracts all over the country, including detailed records of the bids by each contractor, together with the time at which each bid was placed. The time stamps are a unique feature of the data that allows us to detect and measure corruption.[4]

Our first approach exploits the fact that the preferred bidder wants to collect all information on the other participants' bids and thus bids last. We calculate the probability of winning conditional on being the last to bid. In the absence of corruption, bidding and timing decisions are independent, and this probability would be equal to the inverse of the number of participants (that is, three bidders should each have a 1 in 3 chance of winning). In the data, for auctions with three bidders, the probability of winning conditional on bidding last is 0.397, while the probability of winning conditional on not bidding last is 0.301. Comparing the two numbers, we test for the presence of corruption. In addition, the difference between the probabilities allows us to infer the share of corrupt auctions, that is, to measure corruption.

Our baseline approach is intuitive, but it uses only the information on the

[3] http://zakupki.gov.ru.

[4] To our knowledge, time-stamp data are unavailable for the existing sealed-bid auction data sets (as opposed to open auctions). In addition, our new auction data set is one of the largest, with 4.3 million auctions and more than 10 million bids.

probability of winning for the last bidder and for all other bidders. A complementary approach is to use the distribution of timing for the winners and the other bidders to measure the abnormal mass of winning bids placed near the deadline. The share of corrupt auctions can then be estimated from a normalized difference of the cumulative distribution functions of timing for the winner and the runner-up. This approach yields similar quantitative results.

This measure of corruption shows that 10.8% of auctions are affected by bid leakage. The absolute number of contracts affected is 25,000—a number that balloons to over 380,000 if we scale up the estimate to the six-year study period. The value of affected contracts was as high as $1.2 billion.[5]

Once we measure corruption for the whole sample, we can repeat the exercise for a subsample of frequent auctioneers and bidders. That is, we can calculate the difference in conditional probabilities separately for each auctioneer and each bidder. This helps to determine which auctioneers and bidders are more corrupt. In this sense, we follow the literature that detects collusive groups of bidders (Kawai and Nakabayashi, 2014; Porter and Zona, 1999).

In the second part of our analysis, we examine the consequences of corruption for the expected prices that public bodies paid for the contracts. We build a model of bidding behavior. The model predicts that while a direct effect of corruption increases the expected price of the contract, corruption can also affect prices indirectly by changing the behavior of honest bidders. Intuitively, honest bidders behave more aggressively by reducing their bids. As a result, this equilibrium response can *lower* average prices of procurement contracts as opposed to inflating them, as would be natural with corruption. Nevertheless, the net effect on prices

[5]As a benchmark, we can compare these numbers to a large-scale estimate of collusion from Kawai and Nakabayashi (2014). To our knowledge, they have provided the only estimate of collusion measured across a nation, even though they cover only one industry. Kawai and Nakabayashi (2014) show that collusive bidders were awarded 7,600 construction projects worth $8.6 billion in Japan, equal to 19% of national government construction projects. Their scaled-up estimates suggest that 0.85% of the GDP of Japan was affected.

for standard distributions of costs is still positive. That is, the direct effect of corruption is larger in magnitude than the indirect effect.

We test the predictions of the model starting with a reduced-form estimation. We regress the final prices of the contracts and the bids on the measures of corruption of the auctioneers and the bidders, as well as the interaction between auctioneer and bidder corruption. For bidders with low measure of corruption, higher auctioneer corruption is associated with lower bids, while for more corrupt bidders, auctioneer corruption increases the bids. If we interpret our estimates as causal, reducing the share of corrupt auctions from mean value of 0.1 to zero reduces prices by 3%.[6] This estimate hides important heterogeneity. If we concentrate on the subsample of frequent auctioneers and bidders, corruption has a nonlinear effect on prices. For this subsample, there is a nonzero optimal level of corruption.

In the third part of our analysis, we estimate our model structurally to confirm that bidding behavior changes in line with the model. We restrict our analysis to the bidding behavior of bidders with low estimated share of corrupt auctions—"honest bidders." We compare how honest bidders bid when they participate in the auctions with corrupt and with noncorrupt auctioneers. We follow the existing literature on the estimation of the valuations (costs) from the bidding data using the first-order conditions in Guerre, Perrigne, and Vuong (2000), and extend their approach to incorporate bid leakage and to estimate bid functions and the underlying distribution of costs. This approach allows us to estimate bid functions of honest bidders separately for each auctioneer. For more corrupt

[6] If we compare these effects with the existing studies of corruption in similar frameworks, our results are somewhat smaller in magnitude. Di Tella and Schargrodsky (2003) reported that a crackdown on corruption in Argentinean hospitals reduced procurement prices by 10%. Bandiera, Prat, and Valletti (2009) showed that active waste in Italian public procurement increased prices by 11%. Other papers: Ferraz and Finan (2011); Olken (2006, 2007); Reinikka and Svensson (2005) in various settings show estimates of loss from 9% to 24%. Although these papers study procurement and corruption in distribution of pubic funds, they do not study the bidding data in procurement auctions.

auctioneers, bid functions of the honest bidders shift downward. That is, for the same value of the underlying costs, they bid more aggressively, which is in line with our reduced-form results.

Relative to existing literature, we make several contributions. First, we detect corruption (an agreement between an auctioneer and one of the bidders) as opposed to collusion (an agreement between two or more bidders) noted in previous literature.[7] Our paper is the first to build a method for corruption detection in first-price sealed-bid auctions and one of the first papers to empirically study corruption in auctions.[8] At the same time, we draw insights on the bid-leakage corruption in auctions from a vast theoretical literature.[9] Our model of strategic behavior of honest bidders extends Arozamena and Weinschelbaum (2009) to allow the level of corruption to be uncertain and vary across auctioneers. None of the existing theoretical papers provide any empirical evidence on this type of corruption, nor on offsetting effects of corruption on prices in auctions.[10]

We also contribute to the literature on inefficiencies in procurement.[11] This literature perceives the mechanism of allocating a contract as a black box, while

[7]See Porter (2005), and Asker (2010a) for a review of the literature on collusion in auctions, as well as Hendricks and Porter (1988); Porter and Zona (1993); Baldwin, Marshall, and Richard (1997); Porter and Zona (1999); Pesendorfer (2000); Bajari and Ye (2003); Ishii (2009); Asker (2010b); Athey, Levin, and Seira (2011); Conley and Decarolis (2011); Haile, Hendricks, Porter, and Onuma (2012); Kawai and Nakabayashi (2014); Schurter (2016).

[8]With a notable exception of Cai, Henderson, and Zhang (2013), who examine corruption in auction design in land markets in China.

[9]See for instance, Compte, Lambert-Mogiliansky, and Verdier (2005); Menezes and Monteiro (2006); Lengwiler and Wolfstetter (2006); Burguet and Perry (2007); Arozamena and Weinschelbaum (2009); Lengwiler and Wolfstetter (2010).

[10]We are also the first to exploit the timing as a strategic variable in sealed-bid auctions as opposed to open auctions (such as eBay). We show that timing can matter even in sealed-bid auctions, without any considerations that are usually relevant to open auctions (see, for example, Bajari and Hortacsu, 2003; Ockenfels and Roth, 2002, 2006; Hendricks, Onur, and Wiseman, 2012; Hopenhayn and Saeedi, 2015).

[11]Existing papers attribute the inflation in contract prices to different inefficiencies, such as political connections (Schoenherr, 2014; Callen and Long, 2015; Baltrunaite, 2016; Coviello and Gagliarducci, 2016; Guerakar and Meyersson, 2016), red tape (Bandiera, Prat, and Valletti, 2009; Lewis-Faupel, Neggers, Olken, and Pande, 2014; Best, Hjort, and Szakonyi, 2017), or corruption (Shleifer and Vishny, 1993; Di Tella and Schargrodsky, 2003; Ferraz and Finan, 2008; Mironov and Zhuravskaya, 2014).

we stress how corruption changes the outcomes of auctions and subsequently distorts the contract prices. Moreover, little is known on how equilibrium response of honest agents can change the implications of corruption. Although percentage estimates of loss are of a similar order of magnitude to those in the existing literature, we point out that corruption can have severe implications for the behavior of the noncorrupt bidders through the indirect effect.

From the policy side, our paper contributes to the literature on electronic procurement (Elbahnasawy, 2014; Lewis-Faupel, Neggers, Olken, and Pande, 2014; Best, Hjort, and Szakonyi, 2017) and e-governance in general (Banerjee, Duflo, Imbert, Mathew, and Pande, 2014; Muralidharan, Niehaus, and Sukhtankar, 2014). One policy lesson from our paper is that recording the timing of bids is extremely helpful for detecting various types of illegal behavior and for measuring their magnitude.

The rest of our paper is organized as follows. Section 2 describes the institutional background and the details of the procurement procedure that we analyze. Section 3 describes the data and illustrates the ideas behind our methods using simple graphical analysis. Section 4 sets up tests and measures for corruption and shows the results of the estimation. Section 5 presents a stylized model of bidding behavior under bid leakage and examines the expected prices of the contracts with and without corruption. Section 6 explores how corruption varies across auctioneers and bidders and provides reduced-form estimates of the effect on prices in a reduced-form way. Section 7 presents the structural estimation of bid functions of different groups of auctioneers. Section 8 concludes.

1.2 Institutional Background

In Russia public procurement accounted for 25% of 2016 GDP.[12] Most contracts are allocated through auctions. The three most widely used auction types are *requests for quotation*, *open auction*, and *open tender*.[13]

We focus on requests for quotations, or, as we call them interchangeably throughout the paper, first-price sealed-bid auctions, which is the format used for small contracts. The reserve price of such a contract cannot exceed 500,000 rubles (≈\$8,800),[14] and no more than 100 million rubles (≈1.8 million) per year can be assigned this way.[15] Typical examples of these contracts are purchases of office supplies for a municipality, books for a public school, or medical supplies for a public hospital. Small repairs, street cleaning, and other types of services can be purchased through this format as well. Requests for quotations require less paperwork for the public body, but they are also less transparent for the controlling agencies to oversee.

A request for quotations proceeds as follows: a public body posts an announcement of the auction at the designated website. The format of the public notice is standardized; the notice contains exhaustive information about the contract, including the deadline to submit a bid and the requirements to qualify as a bidder. For larger contracts, the public body has to upload an announcement to the website not less than seven business days before the deadline of the auction.

When the auction is live, any eligible firm can submit an application, which has

[12]In absolute numbers, that is ≈\$530 billion. Click here to update the exchange rate. The information is from the Ministry of Economic Development for 2016 (in Russian): http://www.interfax.ru/business/499881.

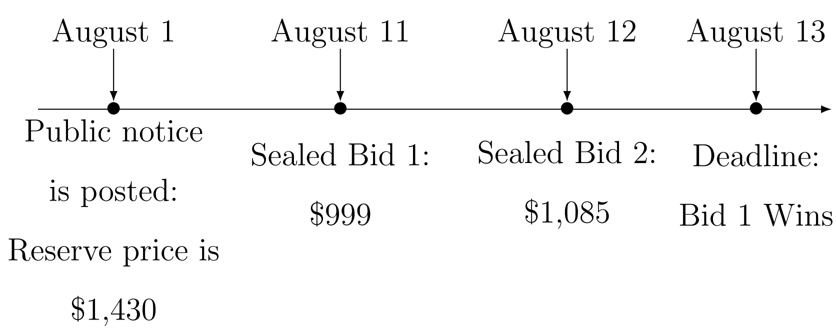
[13]A request for quotation is a first-price sealed-bid reverse auction. An open auction is an online version of a reverse English auction—the participants observe each other's behavior and compete online. The third type, the open tender, resembles a first-price sealed-bid reverse auction, but it takes into account the quality that a bidder can provide. In other words, it is a *scoring auction*.

[14]Click here to update the exchange rate.

[15]Click here to update the exchange rate.

to include a contract price (we call it a "bid"), together with the documentation that confirms the eligibility of the firm. The majority of applications are submitted in sealed envelopes that are usually hand-delivered by a representative of the applicant. In the first part of the study period applications could have been submitted via email with an electronic signature, however recently the number of such applications declined. We discuss the reasons for that and the importance of the difference between paper and electronic submissions later in this section.

After the auction has ended, the applications are opened by the auctioneer. Applications can be rejected if the bidders do not meet the posted requirements. If there was only one bidder and the bidder is eligible, the contract is signed with this bidder. The winner is determined by the lowest bid. If the bids have equal price offers, the earlier bid wins. The public body posts the results of the auctions on the public website.[16] We illustrate the timeline of a typical auction in the graph below, using actual data.



1.2.1 Bid-Leakage Corruption

Corruption in procurement happens in several ways. We concentrate on one generic type of corruption that sealed-bid auctions are especially prone to—*bid-leakage corruption.* Bid leakage is an agreement between an auctioneer and one of the bidders in an auction. The auctioneer shares with one of the bidders—whom

[16]The bidders or their representatives have a right to be present for the opening procedure, can request that any information from the bidding envelopes be disclosed, and can make a video recording of the procedure.

we call the *preferred* or *corrupt* bidder—details of the bids of other participants. Once the corrupt bidder is confident that all other bids have been placed, and if she finds it profitable to bid, she places her bid. If she is not overly concerned with being caught by a higher-level authority, she can undercut the current best bid by as little as possible.[17]

To illustrate our detection approach, we start with several examples from a major online forum that discusses issues of public procurement in Russia.[18] In the first example, Firm A participated in five auctions organized by the same public body. Each auction had four bidders and a reserve price in rubles equal to $6,000. In each of the five auctions, Firm A came in second, by margins of $0.50 to $12. The same company won all five auctions. In another case, Firm B faced a similar situation, and the winner was also always the last firm to bid. In one extreme case, Firm C complained that it was always second in 60 auctions in a row. It always lost by less than 0.5% of the reserve price, always by around 0.5% of its own bid, and always to the same firm, which placed its bids last. This behavior suggests that the winner knew the bid of the complaining firm and waited until the end of the auction to place its sealed bid.[19] These three examples highlight something we will focus on—late bidding by the winner.

Bid leakage is not unique to Russia. Examples abound from many other countries, both developed and developing. For instance, the government of Singapore banned Siemens for five years from participating in any public procurement auctions after the company bribed an official to learn the bids of their competitors. Similar abuses occurred in Berlin over the contract for building a new airport (Lengwiler and Wolfstetter, 2006). In several Italian auctions, corrupt bureaucrats used a laparoscope to get inside the sealed envelopes without damaging

[17]Appendix A discusses a method for detecting bid-leakage based solely on the differences in bids.

[18]http://forum.gov-zakupki.ru/.

[19]Here are the links to these and other cases from this forum (in Russian): (1), (2), and (3).

them.[20] Ingraham (2005) documented cases of corruption in school construction auctions in New York City.

1.2.2 Bidder Response

To what extent do honest bidders react to bid-leakage corruption? We draw our motivation from the same online forum. The bidders discuss two forms of their reaction to bid-leakage corruption: response in timing and response in bids.[21] These bidders argue that response in timing would not be very effective: the preferred bidder will still submit an envelope at the last moment.[22] One of the participants suggested that reducing bids could be effective, since the preferred bidders would find it not profitable to win the contract.[23] Thus, anecdotal evidence suggests that honest bidders should engage in more aggressive bidding. We model the bidding response in Section 5 and examine it empirically in our estimation in Sections 6 and 7.

If bids were submitted via email with an electronic signature, the timing response could be more effective for honest bidders. In the first part of the study period (2011 to 2013), as many as 37% of the applications were submitted electronically. Since we do not know whether those electronic submissions were encrypted or not, we have to be careful with this subsample due to potential timing response. The electronic submission of applications was effectively eliminated in the second part of the study period (2014 to 2016)—only 1.6% after 2013 were submitted electronically. We apply our methods to the latter period, but we also directly test for strategic timing response.

---

[20] The link that discusses the case (in Italian): http://www.ilfattoquotidiano.it/2016/06/27/.

[21] Examples here https://web.archive.org/web/20180123052630/ and here: https://web.archive.org/web/20180123055013.

[22] Potentially, she can submit the envelope even after the deadline: https://web.archive.org/web/20180123053745/http://forum.gov-zakupki.ru/topic40009-10.html

[23] https://web.archive.org/save/http://forum.gov-zakupki.ru/topic40009-30.html#p4720

1.2.3 Change of Regulations

The Russian government has been declaring a fight against corruption for many years.[24] Part of this struggle centers on transparency in public procurement.

One of the government's measures to promote transparency and reduce corruption in procurement was a new piece of federal regulation—Federal law #44. The law mandates that the public be given access to all of the procurement information. It also standardizes the format for all procurement documents, which must be published online, and specifies personal responsibility of public bodies in case of violations, among other steps.[25] Another important measure is an introduction of online standardized bidding system that needs to be used for any type of procurement procedure. In practice, the actual implementation of the platform was postponed, first to 2017, and then to 2018. As a result, submitting bids in any electronic form became illegal, and up to discretion of a public body.[26]

In the first part of the paper we study the period after the introduction of FZ#44, since it allows us to not worry about the strategic response in timing. Once we measure corruption for this subsample we use our auctioneer-level estimates to test the model for the whole sample, including the period of 2011 to 2013.

[24] Putin declared the fight against corruption in 2000 (http:/kremlin.ru/events/president/transcripts/24138) and kept declaring it every year since then. The 2015 quote: https://web.archive.org/save/https://www.1tv.ru/news/2015-03-04/20595-v_putin_borba_s_korruptsiey_vo_vlastnyh_strukturah_budet_vestis_posledovatelno_i_zhestko.

[25] This link (in Russian): http://www.garant.ru/actual/contracts/472245/ describes other measures, including minor changes in the assignment mechanisms, introduction of long-run planning, anti-dumping measures, amended for voiding or changing the contract, and introduction of organized audits of the results.

[26] Discussion between the participants: https://web.archive.org/web/20180123061015/.

1.3 Data

We downloaded a digital archive of all procurement auctions from January 2011 through the end of 2016.[27] The archive contains announcements and protocols, all in .xml format. A typical announcement includes information on the terms of the contract, the reserve price, the auction deadline, and the deadline by which the work specified in the contract should be delivered. Protocols include the bid for each application, the time-stamp, and also whether the bid was accepted or rejected.

We matched the announcements to the protocols and extracted all of the necessary information by parsing each of the .xml files to get the information in each tag. The resulting data set has more than 4.3 million requests for quotations, and more than 10 million bids. We dropped auctions with any quote being rejected and auctions with fewer than three quotes, leaving us with 841,552 request for quotations.[28]

Table 1 of Appendix C shows summary statistics for the whole period covered and for the new law, FZ#44 (2014–2016). The reserve price in requests for quotations is a maximum price that the winning bid cannot exceed in order for the auction to be considered valid. For requests for quotation this reserve price has lower than 500,000 rubles, or $8,700, as of the current exchange rate.[29] In our sample, the mean reserve price is 213,507 rubles ($3,718).[30] The mean winning bid is 157,850 rubles ($2,749),[31] while the mean ratio of the winning bid to the reserve price is 73.3%.[32]

[27] Available in FTP or HTML formats from the official website http://zakupki.gov.ru.

[28] The auctions without quotes do not provide any insights for our analysis. Neither do the rejected bids. We do not use the auctions with one or two bidders in our main analysis since we cannot employ our methods without observing a third bid.

[29] Click here to update the exchange rate.

[30] Click here to update the exchange rate.

[31] Click here to update the exchange rate.

[32] For more details on the reserve price distribution and the distribution of the winning bid

Figure 1 shows that the number of bids placed in the last hours before the deadline is large for all bidders; however, the winner tends to bid later asymmetrically more than the other bidders. Most of the extra density for the winner comes from the last *minutes* of the auction. The densities for the second, third, and fourth bidders are practically indistinguishable from each other. This is an exact pattern that we expect to see in the presence of bid leakage: even with a substantial share of bids placed in the last day, the winners tend to bid asymmetrically closer to the deadline, with a gap more pronounced in the last hours of the auction. In the next section we quantify this asymmetry between the winners and all other bidders, in order to detect and measure corruption.

---

one can resort to Appendix C. Panel A of Figure C1 shows that the distribution of reserve prices has almost a full support from 0 to 500,000 rubles. The empirical density of the reserve price is monotonically decreasing apart from the spikes at round numbers and especially on the maximum level of 500,000 rubles, and except for the neighborhood of 0. Panel B of Figure C1 shows that the winning bids have a similar distribution, but without a spike at 500,000 rubles.

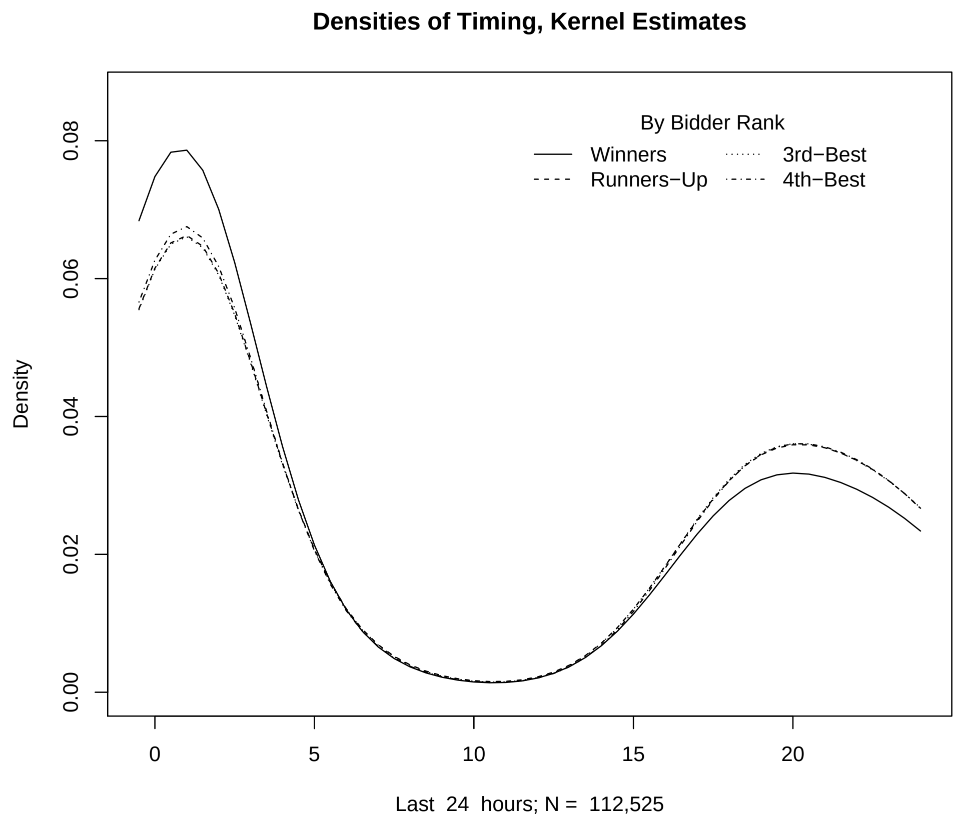


Figure 1.1: Hours to the Deadline for the Top Four Bidders

Notes: Kernel density estimates for the timing of bids. Distance in hours to the deadline over the last twenty four hours. The densities are drawn separately for the winning bids, the runner-up bids, the third-best bids, and the fourth-best bids. Auctions with at least three bidders. The subsample is auctions with no bids placed within 15 minutes corridor from each other to control for artificial correlation coming from correlated bidding. FZ#44. Gaussian kernel with a normal optimal smoothing parameter is used.

1.4 Detecting Corruption

We start this section by showing that corruption in the form of bid leakage is present in the data. We introduce an important assumption–*an independence assumption.* In the next subsection, we discuss why independence is likely to hold in our setting, and provide several tests for it. The independence allows us to measure corruption both for the whole sample and auctioneer-by-auctioneer.

As Figure 1 shows, bids and timing of bids are correlated in a specific way, with winners placing their bids later in the auction. We argue that this is a result of corruption. An alternative explanation could be that there is correlation between timing and bidding decisions even absent corruption. For instance, more efficient bidders could place their bids later in the auction, while estimating the costs of the project, but this is implausible for the most of the contracts in our data. Yet another alternative explanation would be that bidders procrastinate in a specific way—e.g. with more efficient bidders procrastinating more. To rule out all of these alternative explanations, in the next subsection, we provide evidence against correlation of bids and timing absent corruption. For now, however, we just assume that bids and timing are independent, in a spirit of identification assumptions in collusion literature.[33]

Every bidder i in auction j is characterized by a pair (b_{ij}, t_{ij}), where b_{ij} a bid and t_{ij}, the timing of the bid. We omit the indices for the rest of the section. The independence assumption is,

$$b \perp t \quad (I)$$

In this case, rejecting independence in the data is equivalent to accepting corruption. Next, we show that there is corruption in the form of bid leakage, but so far our tests are uninformative about the independece per se. We deal with this issue in the next subsection.

In order to test for corruption, we exploit two specific violations of independence of bids and timing in the data that should be pronounced with bid leakage. The first is that without corruption, bidding last should not predict winning. We are testing the equality of probability of winning conditional on being the last one to bid and a probability of winning conditional on not being the last one to bid

[33]See, for instance Porter and Zona (1993); Bajari and Ye (2003) for an example of how these restrictions are derived.

against the alternative of inequality.

$$H_0 : \mathbb{P}[\text{win}|\text{last}] = \mathbb{P}[\text{win}|\text{not last}] \quad \text{Test 1}$$

We use the classical χ^2 test for no association in contingency tables. Our case is simple, since it is a 2×2 contingency table consisting of two indicator variables– one for winning, and one for being the last in time. In practice, we implement this test by running an OLS regression of the indicator of winning on the indicator of being the last in time, and using Wald test from this OLS regression.[34]

[34] Anatolyev and Kosenok (2009) show that the classical χ^2 test is asymptotically equivalent to the Wald test from OLS in this particular case.

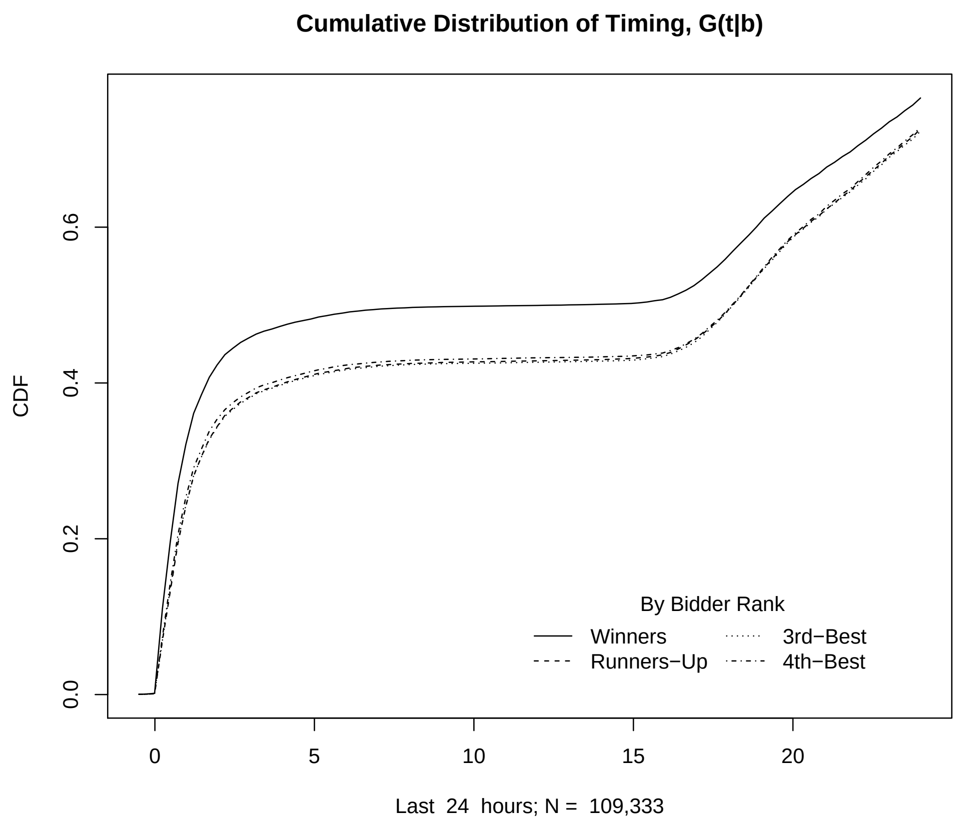


Figure 1.2: Hours to the Deadline for the Top Four Bidders

Notes: Empirical CDFs for the timing of bids. Distance in hours to the deadline over the last twenty four hours. The CDFs are drawn separately for the winning bids, the runner-up bids, the third-best bids, and the fourth-best bids. Auctions with at least three bidders. The subsample is auctions with no bids placed within 15 minutes corridor from each other to control for artificial correlation coming from correlated bidding. FZ#44.

Apart from changing the conditional probability of winning, bid leakage shifts the timing of the winner closer to the deadline. This fact motivates our second test. Figure 2 illustrates the idea behind it. It shows the cumulative distribution functions of timing for winners, runners-up and other bidders. There is a large difference for winner and runner-up, while there is not much of a difference for the rest of the CDFs. If we denote the CDF of timing conditional on having kth rank in bids by $G(t|b = b_{(k)})$, we want to test an equality of the distribution of timings for the winner and any of the other bidders (note that from independence they all should be the same). The H_0 in this case is formulated as

$$H_0 : G(t|b = b_{(1)}) = G(t|b \neq b_{(1)}) \quad \text{Test 2}$$

We test this hypothesis by a two-sample Kolmogorov-Smirnov test.

Table 2 of Appendix C shows the results for the two tests above. Panel A runs Test 1 using Wald test from OLS of winning on being last. The difference in conditional probabilities of winning for the last bidder and for the non-last bidders is 0.096, and the p-value for the Wald test for equality of probabilities is less than 0.001. The fact that the last bidder is much more likely to win, strongly suggests the presence of corruption. We cannot accept the hypothesis of no corruption from Test I.

Panel B of Table 2 implements Test 2 with a Kolmogorov-Smirnov test. We report the test statistics for Kolmogorov-Smirnov test for the equality of CDFs of timing for the winner, the runner-up, and the third bidder. All of the differences are significant at the 0.1% level as well.[35]

1.4.1 Independence and Alternative Explanations

We fail to accept independence of bids and timing, and we interpret it as evidence of corruption. An alternative explanation could be that the independence fails even absent corruption, as in the examples with bidders collecting information, or bidders procrastinating. To test the independence without corruption, we need to construct a noncorrupt subsample. It is hard to build, since the noncorrupt subsample is fundamentally unobserved. However, with bid leakage we know the form of corruption, and hence we can derive such a subsample.

[35]To avoid noise on the right tail of the timing distribution, we keep only the auctions with $t \leq 72$ hours. If we consider the whole sample, the differences will still be significant at the 0.1% level.

We assume that the corrupt bidder waits until the last moment of the auction (or until all potential bidders have placed their bids) and either undercuts the most competitive honest bid or places a high bid.

Assumption 1. A corrupt bidder either wins the auction at the last moment or places a bid higher than the most competitive honest bid.

For the first placebo test, we construct an artificial sample of auctions (A), discarding information on non-winning and non-last bids. We use Wald test from OLS as before. Note that now it is running a regression of being a runner-up on being the second-last, for the subsample of no winners and last bidders. The null hypothesis for this test is formulated as

$$H_0 : \mathbb{P}[\text{bidder second}|\text{second-last}\cap A] = \mathbb{P}[\text{bidder second}|\text{not second-last}\cap A] \quad \text{Test } 1'$$

against a two-sided alternative.

For the second placebo test, we do not consider distribution functions of timings for the winners:

$$H_0 : \; G(t|b = b_{(j)}) = G(t|b = b_{(k)}), \; \forall j, k > 1 \quad \text{Test } 2'$$

Table 3 of Appendix C presents the results of our placebo tests. Panel A shows the results from the Wald test. We cannot reject no association, with a p-value still being less than 0.001, however the difference in conditional probabilities is 7.4 times smaller in magnitude. In addition, the difference is negative, which implies that more efficient bidders do not bid closer to the deadline.

In a similar spirit, Panel B repeats the Kolmogorov-Smirnov test for the equality of the second and the third CDFs (Row 3), the second and the fourth CDFs (Row 4), and the third and the fourth CDFs (Row 5). Two out of three differences are insignificant, one with p-value larger than 0.001, and another one with a p-value of 0.451.

One implication of bid leakage and of Assumption 1 is that most of the results should be driven by the subsample of late bids. If the corrupt bidders are always among the winners and are the last bidders, the asymmetry in conditional probabilities should be more pronounced in the subsample of late bids. In contrast, there should not be much difference in the subsample where only early bids are placed. We define the "All Late" subsample as those bids that were placed within 60 minutes of the deadline. In addition, we define the "All Early" subsample, where all of the bids were placed at least five hours before the deadline. As an additional check, we concoct the "No Early" subsample, comprising auctions where at least one of the bids was placed less than five hours to the deadline. All of the subsample results are reported in Table 4 of Appendix C.

When we run Test 1 for the "All Late" subsample, we find that the differences for this subsample are larger than for the whole sample; the test statistic is 0.153 (Row 1 of Table 4). This confirms the form of corruption. Next, we switch to the "All Early" subsample. The difference for this subsample is small 0.023 (Row 2 of Table 4).

Next, we take the "No Early" subsample. If what we observe is indeed bid leakage, most of the difference should come from last-minute bidding, by the same logic that applied to the "All Late" subsample. As Row (3) of Table 4 shows, it is true that the difference the "No Early" subsample is large in magnitude (0.125). The results for auctions with more than three bidders are qualitatively similar.

1.4.2 Measures of Corruption

Since we cannot reject independence given the data1, Tests 1 and 2 can be modified to derive measures of corruption.[36]

[36]In addition, Appendix A discusses an alternative measure of corruption that is based solely on the information on the differences of bids between the winner and the runner-up, the runner-

One implication of the independence is that without corruption the probability of winning does not depend on the timing. It is equal for all of the bidders. In an auction with K bidders, it is symmetrical and equal to $1/K$. In the presence of bid leakage, the last bidder is always the winner, so this conditional probability is equal to one.

Denoting the probability of corruption in an auction by α, the conditional probability of winning can be written as

$$\mathbb{P}[\text{win}|\text{last}] = \alpha + \frac{1}{K}(1-\alpha).$$

Solving for α, one gets

$$\alpha = \frac{\mathbb{P}[\text{win}|\text{last}] - 1/K}{1 - 1/K}.$$

This is our first measure of corruption, α_I.

The second implication of the independence is that we can use the cumulative distribution function of timings for non-winning bids as counterfactual for the CDF of timings without corruption.

Namely, we can write

$$G(t|b = b_{(1)}) = \alpha \cdot G(t|b = b_{(1)}, \text{corruption}) + (1-\alpha) \cdot G(t|b = b_{(1)}, \text{no corruption}).$$

The independence implies that we can plug in for $k > 1$, $G(t|b = b_{(1)}, \text{no corruption}) = G(t|b = b_{(k)}, \text{no corruption})$ and then plug in $G(t|b = b_{(k)}, \text{no corruption}) = G(t|b = b_{(k)})$.

We also need to choose a parameterization for $G(t|b = b_{(1)}, \text{corruption})$, keeping in mind that with corruption the winner will bid close to the deadline. For now, we use a CDF that is discrete at 0, but we can also use a CDF of a uniform

up and the third bidder, and so on. This method allows to measure corruption for auction data sets, where data on timing are not available.

distribution on $[0, \epsilon]$, where ϵ is relatively small compared to an average distance to the deadline. The results will be similar if we use uniform distribution instead of discrete. The interpretation is that the bid of the winner is by necessity placed at the deadline or in the neighborhood of the deadline.

In this case, we can solve for α as follows,

$$\alpha = \frac{G(t|b = b_{(1)}) - G(t|b = b_{(k)})}{1_{t=0} - G(t|b = b_{(k)})}.$$

A sample analog of it depends on t.

In practice, we implement the method as follows: we cut the sample to the last three days before the deadline, in other words, 72 hours. We are doing it to avoid finite sample bias arising from the skewness of the distribution of timings toward the deadline (see for example Figure 2). We also drop auctions where two bids were submitted together (within 15 minutes of each other).[37] These two constraints leave us with around two-thirds of the initial sample. Next, we get the estimates of $\hat{\alpha}(t)$ for $t = 0.5, 1, 2$, and $t = 5$. We use an average of $G_{(2)}$ and $G_{(3)}$ for $G(t|b = b_{(k)})$ to improve efficiency—we call this measure α_{II}.

We estimate a share of corrupt auctions from the first measure α_I in Table 5 of Appendix C. Columns are split by subsamples by the number of bidders. The largest estimated share of corrupt auctions (9.6%) is for the three-bidder auctions (Column 1).

Table 6 presents the results for the second measure α_{II} with different comparison CDFs: either only the runner-up CDF or the average of the runner-up and the third, and with varying cutoff levels. The estimates vary from 8% to 12% and are all statistically significant at the 5% level.[38] Panel B of Table 6 shows the

[37] In is a part of the sample, firms coordinate their bidding behavior by submitting their bids together or within a 15 to 20 minute interval. This can potentially bias our estimates in an unpredictable way. However, if we cut these simultaneous bids, we can prevent this from happening.

[38] They are significant at the 1% level as well.

results of the placebo tests, where the CDF of the winner is replaced by the CDF of either the runner-up or the third bidder. The placebo estimates are sometimes significant, but many of them are negative and insignificant. Overall, the estimates of the shares of corrupt auctions closely resembles the results from the first measure α_I.

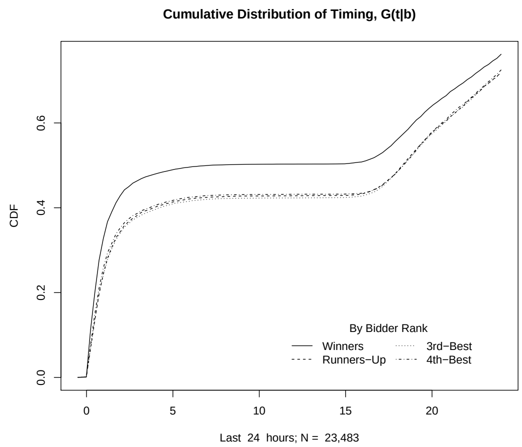
These estimates allow us to provide some back-of-the-envelope calculations. To avoid weighting by the number of bidders we take $\hat{\alpha}_{II}$. Picking $t = 1$ as a medium cutoff we get $\alpha_{II} = 0.108$, or 10.8% of auctions are corrupt. If we multiply these numbers by the total number of contracts with three bidders and from FZ#44 we get that around 25,000 contracts are affected, with a total price of these contracts of $61 million. If we account for the auctions with two bidders and auctions from the first part of the study period (2011 to 2013) these numbers change to 383,000 contracts with a total value of $ 1.2 billion in six years. If we scale up these estimates to the total size of the government purchases—25% of GDP—2.7% of GDP is affected, although this figure is not very reliable, since corruption can take other forms in other types of government purchases.

1.4.3 Timing Response

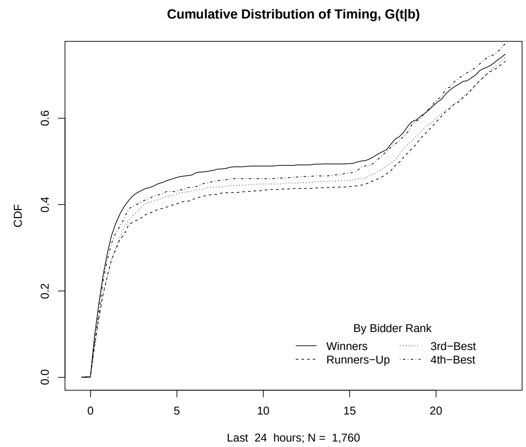
Another concern that we might have is that bidders react to corruption by adjusting their timing. Note that, to violate independence, and thus to threaten our tests, strategic response in timing needs to be correlated with the bid strength, or, in other words, with the efficiency of a bidder.

We formally ruled out that this is the case in Table 3. In addition, as we discussed in Section 2 strategic timing response is highly unlikely with paper applications. However, it is more plausible with electronic auctions. Figure 3 shows the CDFs of timing for the auctions where *only* paper submission was allowed (Panel A) and *only* electronic submission was allowed under FZ#44 (Panel

B). The difference between CDFs for the non winners is more pronounced for the electronic submissions. The electronic submissions is a small part of the sample for FZ#44[39]. Thus, we are not concerned with the differential response in timing.



Panel A: only paper submission



Panel B: only electronic submission

Figure 1.3: Cumulative Distribution of Timing for the Top Four Bidders. Paper and electronic submission

Notes: The subsample is auctions with no bids placed within 15 minutes corridor from each other to control for artificial correlation coming from correlated bidding. Cumulative distribution functions for the timing of bids. Distance in hours to the deadline over the last twenty four hours. The CDFs are drawn separately for the winning bids, the runner-up bids, the third-best bids, and the fourth-best bids. Auctions with at least three bidders. Only the FZ#44.

While we ruled out the differential response in timing, we still might be concerned that honest bidders adjust their timing independently of their bids. This does not affect our tests, but can bias our measures from the previous subsection. To address this issue, we compare the timing behavior of nonwinners in auctions, where our measures of corruption show high levels of corruption, and in those auctions, where measures of corruption, judging from α_{II} is zero. Figure 4 depicts the CDFs of timing for winners, nonwinners for corrupt and noncorrupt auctions.

[39]Less than 2% of bids had to be submitted only this way.

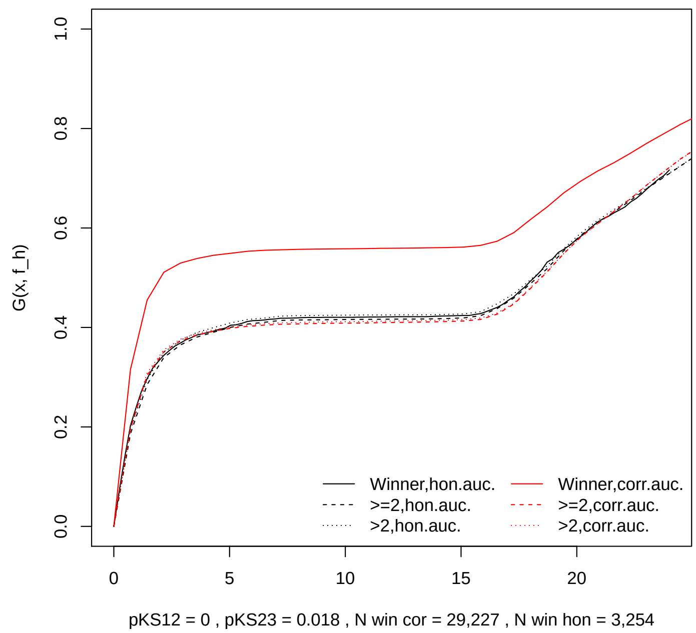


Figure 1.4: Cumulative Distribution of Timing for Corrupt and Noncorrupt Auctioneers, Winners and Nonwinners

Notes: The subsample is auctions with no bids placed within 15 minutes corridor from each other to control for artificial correlation coming from correlated bidding. Cumulative distribution functions for the timing of bids. Distance in hours to the deadline over the last twenty four hours. The densities are drawn separately for the winning bids, and other bids. Auctions with at least three bidders. Corrupt auctioneers are defined by $\hat{\alpha} \geq 0.08$, honest auctioneers are defined by $\hat{\alpha} \in [-0.01; 0.02]$

As one can see from the graphs, while the difference between winning CDFs is large by construction, there is practically no different for the nonwinning CDFs of timing. The Kolmogorov-Smirnov test that we run show that the difference is not significant. Thus, we can rely on our measures α_I and α_{II}.

1.5 The Model

How does corruption affect prices? In Section 4, we used the tests to detect the patterns in the data that are indicative of corruption. Then we applied the measures to estimate the share of corrupt auctions. Naturally, one wonders how bid-leakage corruption affects the *final prices* of contracts.

The final prices of contracts serve as a convenient measure of government effectiveness in the procurement setting (see for example Bandiera, Prat, and Valletti 2009; Lewis-Faupel, Neggers, Olken, and Pande 2014; Best, Hjort, and Szakonyi 2017). To examine the channels through which corruption affects prices, we model the bidding behavior of an honest bidder.

In our model, corruption effectively reduces the level of competition in an auction. Specifically, the direct effect of corruption is proportional to a change in expected prices, when competition in the auction is reduced by one bidder.[40] As we see from the model, an additional channel exists through which corruption affects the equilibrium price—the response of honest firms. Under certain conditions on the distribution of costs, honest firms can place lower bids than without corruption. In other words, with corruption they bid more aggressively. This *reduces* the expected price of the contract, in contrast to the direct effect. We show in a stylized example that this indirect effect can be strong enough for corruption to reduce the overall expected prices.

Note that we use the results from Section 4 and directly assume that the independence holds. This allows us to concentrate only on the bidding decision, avoiding modeling the decision on the timing of bids and the response of honest bidders with adjusted timing.

[40] The change in expected price in our context is the same as the change in expected revenue for value auctions. Corruption can also lead to an inefficient outcome. The outcome of the auction is inefficient, when the corrupt firm has higher costs, than the most efficient honest firm, but it still wins the auction.

We model the bidding decision as a standard first-price sealed-bid procurement auction with independent private costs. We want to point out several facts: (1) there is an equilibrium in monotone bidding strategies; (2) the curvature of the distribution of costs determines whether honest bidders bid more aggressively then they would without corruption; (3) the expected price implications of this behavior depend on the shape of the costs distribution, and therefore, the price effect of corruption is an empirical question.

The public body is buying a good or a service from K buyers, where K is a fixed number of bidders.[41] The cost of bidder is an independent draw from a CDF F on the interval $[0, \bar{c}]$. All the players are risk-neutral. Corruption takes the following form: an auctioneer runs a corrupt auction with a probability α. If the auction is corrupt, one of the bidders learns the bids of other bidders and can place her bid guided by this information. We assume that if upon learning other bids the corrupt bidder does not find it profitable to win the auction, she bids above the most competitive honest bid. Another assumption is that corrupt bidders are drawn from the same distribution than honest bidders.[42] Assumption 3 summarizes all of the conditions for our analysis.

Assumption 2: A. Independence holds.

B. The costs c are i.i.d draws from $F(\cdot)$, same for both types of bidders.

C. A corrupt bidder bids above the most competitive honest bidder, when it is not profitable for her to win the auction.

An honest bidder in this case maximizes the following function:

$$\nu(b, c, \alpha) = (b - c)(1 - G(b))^{K-2}((1 - \alpha)(1 - G(b)) + \alpha(1 - F(b))),$$

[41] We relax this assumption in the estimation procedure in Section 7.

[42] We acknowledge that allowing asymmetries between honest and corrupt bidders, and testing for them is also possible in this setting, but it is not our goal. More important, if we allow for different cost distribution for corrupt and honest bidders, the model below is not identified nonparametrically and requires strong parametric assumptions for identification.

where $G(b) = F(\phi_\alpha(b))$ is the CDF of bids. $\phi_\alpha(b)$ is an inverse of a bidding strategy. Note the α index. It stressed that bidding behavior depends on the extent of corruption measured by the probability of an auction being corrupt.

The equilibrium is given by the first-order condition,

$$c = b - \chi(\alpha, b) :=$$

$$= b - \Big[\frac{(K-2)g(b)}{1-G(b)} + \frac{(1-\alpha)g(b) + \alpha f(b)}{(1-\alpha)(1-G(b)) + \alpha(1-F(b))}\Big]^{-1}. \quad (2)$$

Note that for $\alpha = 0$, this boils down to regular FOC of a first-price sealed-bid auction.

Our setting resembles the one studied in Arozamena and Weinschelbaum (2009), with two differences. First, we study the procurement setting. As a result, the conditions on the underlying distributions of costs for existence of equilibrium, aggressive behavior of honest bidders, and for the revenue implications are slightly different. Second, is that we allow for $\alpha \in (0, 1]$, while their analysis is only for the case of $\alpha = 1$. That is, auctions are always corrupt in their setting. We provide all of the technical results that differ from proofs of Arozamena and Weinschelbaum (2009) in Appendix B.

Definition 1: Log-concavity of survival function. Survival function is defined as $1 - F(x)$. We assume that it is log-concave. Note that it means that $h(x) = \frac{f(x)}{1-F(x)}$ (the hazard function) is increasing.[43]

We argue in Appendix B that log-concavity is sufficient for there to exist a symmetric equilibrium characterized by the solution to the differential equation

[43] A broad class of distributions satisfy log-concavity of survival function, e.g., uniform, normal, logistic, extreme values, exponential, Laplace, Pareto, chi-squared, and chi. Power function distribution satisfies it for a value of the main parameter larger than 1. For the value of the parameter less than for power distribution and for Pareto distribution, the survival functions are not log-concave. However, one can compute the bid functions directly in both of these cases. For a detailed treatment of log-concavity and its applications, see Bagnoli and Bergstrom (2005).

in the FOC (2), and that the differential equation (2) has a strictly increasing solution.

Proposition 1. If the survival function is log-concave, then (2) defines an equilibrium with corruption.

If corruption makes the bidders more aggressive, for the same value of the costs the bid will be lower with corruption. In other words, by monotonicity, the costs have to be higher to achieve the same bid level. Formulating it in terms of the FOC (2), more aggression occurs when $\chi(\alpha, b) < \chi(0, b)$. Proposition 2 establishes whether there is more, or less, aggression with corruption.

Proposition 2. If $w(x) = h^{-1}(x) = \frac{1-F(x)}{f(x)}$ is strictly convex, there is more aggression, i.e., $\chi(\alpha, b) < \chi(0, b)$. If $w(x)$ is strictly concave, there is less aggression, i.e., $\chi(\alpha, b) > \chi(0, b)$. If $w(x)$ is linear then corruption does not affect bidding behavior.[44] [45]

We illustrate aggression in Example 1.

Example 1: Power distribution and aggression. Assume that costs are distributed according to power law, that is $F(x) = x^\theta, \theta > 0, x \in [0, 1]$. For simplicity we consider a case with only two bidders $K = 2$, where the auctioneer is completely corrupt $\alpha = 1$ or completely honest $\alpha = 0$. Solving the differential equations from the FOCs, we get a bidding strategy in an honest auction:

$$\beta_0(c) = \frac{\theta}{1+\theta} \frac{1-c^{\theta+1}}{1-c^\theta}.$$

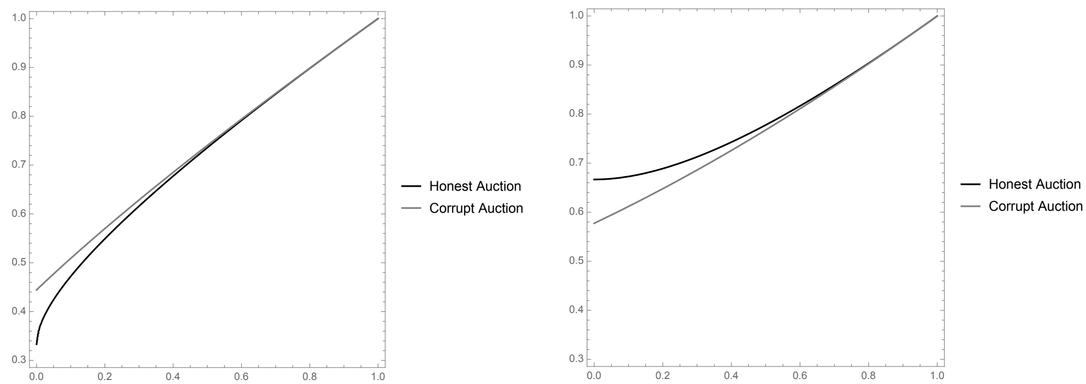
In a corrupt auction, a bid function is defined implicitly through $\phi^c(b)$:

$$\phi_{\alpha=1}(b) = \frac{\theta+1}{\theta} b - \frac{1}{\theta} b^{1-\theta}.$$

[44] Note that it is true for Pareto distribution, $F(c) = 1 - c^{-\gamma}, \gamma > 0$, which does not have a log-concave survival function. However, the bid functions can be derived directly, and α does not affect them.

[45] In Appendix B, we show that Proposition 2 holds for $\alpha \in (0, 1)$.

Figure 5 shows these bid functions for two cases: $\theta = 0.5$ and $\theta = 2$ (Panels A and B, respectively). One bid function being below another means more aggression in this case. The critical value for θ is 1 (uniform distribution). For $\theta = 1$, the distribution is uniform and corruption does not change the bidding behavior of the participants. If $\theta > 1$, corruption makes honest bidders more aggressive, and vice versa, if $\theta < 1$, corruption makes honest bidders less aggressive.

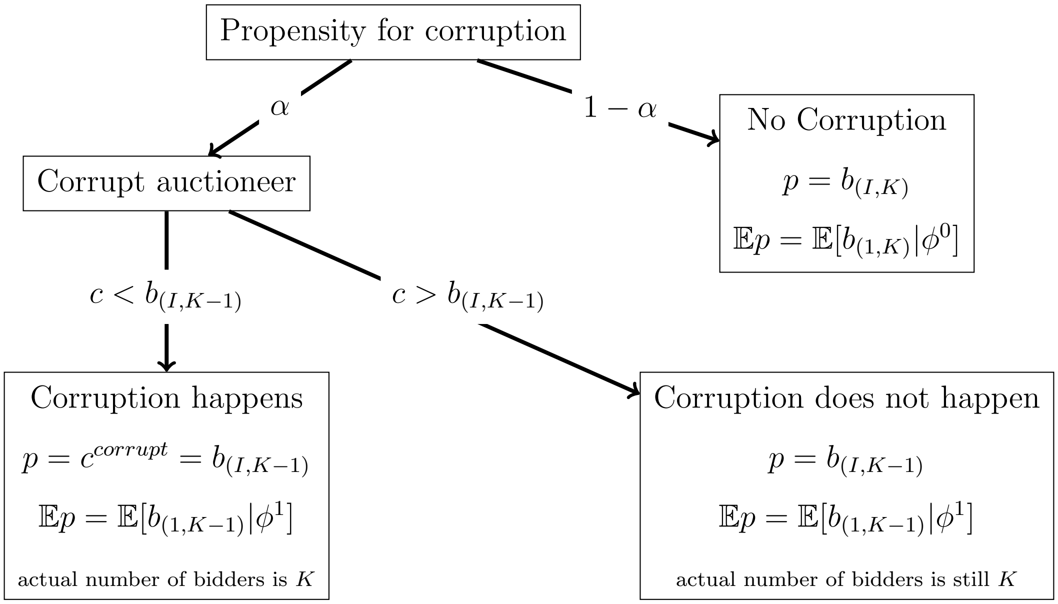


Panel A Panel B

Figure 1.5: Bidding Behavior with and without Corruption

Notes: Panel A: Less Aggression with Corruption, $\theta = 0.5$. Panel B: More Aggression with Corruption, $\theta = 2$

Proposition 2 shows that under fairly general assumptions the expected prices paid to the honest procurers are lower than under corruption. This equilibrium response weights with the direct effect of corruption, so the price implications are potentially ambiguous. Aggression on its own is not enough to make a statement about the expected price. To illustrate this point, we show how the expected price changes with corruption.



The graph above illustrates that even if an auction is corrupt, the corrupt bidder can abstain. This will effectively bring the corrupt auction with no corruption to an auction with $K-1$ bidders.[46] The expected price is given by the following expression:

$$\mathbb{E}[price] = \mathbb{E}[b_{(1)}] = \alpha \cdot \mathbb{E}[b^{(\alpha)}_{(1,K-1)}] + (1-\alpha) \cdot \mathbb{E}[b^{(0)}_{(1,K)}].$$

This expression can be also presented as follows:

$$\mathbb{E}[price] = \alpha \cdot \int (K-1) \cdot b \cdot g_\alpha(b)(1-G_\alpha(b))^{K-2} db + (1-\alpha) \cdot \int K \cdot b \cdot g_0(b)(1-G_0(b))^{K-1} db.$$

Rewriting it once again in terms of the direct and indirect effects, we get,

$\mathbb{E}[price] = \int K \cdot b \cdot g_0(b)(1-G_0(b))^{K-1} db +$

$$\underbrace{\alpha \cdot \Big[\int (K-1) \cdot b \cdot g_\alpha(b)(1-G_\alpha(b))^{K-2} db - \int K \cdot b \cdot g_\alpha(b)(1-G_\alpha(b))^{K-1} db +}_{\text{inflating, direct effect} > 0}$$

[46]The auction can still be a K-bidder auction, but with one high bid being unimportant. Moreover, if we follow the equilibrium choice at the beginning of the section, this bid will still be from the same distribution as the other honest bids.

$$\underbrace{\int K \cdot b \cdot g_\alpha(b)(1-G_\alpha(b))^{K-1}db - \int K \cdot b \cdot g_0(b)(1-G_0(b))^{K-1}db\Big]}_{\text{deflating, indirect effect} > 0}.$$

The integrals in this expression can be calculated explicitly for some distributions or evaluated numerically. We continue with Example 1 to illustrate the changes in expected prices with corruption.

Example 1: Continuation. Without corruption, we need to find

$$\mathbb{E}[price|\alpha=0] = \mathbb{E}[b^0_{(1,2)}] = \int_0^1 \beta_0(c) f(c_{(1,2)}dc = \int_0^1 \frac{2\theta^2}{1+\theta}(c^{\theta-1} - c^{2\theta})dc =$$

$$= \frac{2\theta^2}{1+\theta}\Big(\frac{c^\theta}{\theta} - \frac{c^{2\theta+1}}{2\theta+1}\Big)\Big|_0^1 = \frac{2\theta}{1+2\theta}.$$

With corruption, we need to find

$\mathbb{E}[price|\alpha=1] = \mathbb{E}[b^1_{(1,1)}] = \int_0^1 \theta c^{\theta-1}\beta_1(c)dc =$

$$\{\beta_1(c) = z, dc = \Big(\frac{\theta+1}{\theta} - \frac{1-\theta}{\theta}z^{1-\theta}\Big)dz, z(0) = (1+\theta)^{-1/\theta}\} =$$

$$\theta\Big[\int_{z(0)}^1 (\frac{\theta+1}{\theta}z - \frac{1}{\theta}z^{1-\theta}\Big)^\theta dz + \int_{z(0)}^1 z^{1-\theta}\Big(\frac{\theta+1}{\theta}z - \frac{1}{\theta}z^{1-\theta}\Big)^{\theta-1}dz\Big].$$

This expression needs to be evaluated numerically. Table 7 shows how both expected prices vary with θ. They are approximately equal for $\theta = 2$, and the expected price is lower for $\theta > 2$. This example shows that corruption can reduce expected prices. Note that there is no one-to-one correspondence between aggression and the expected price. The aggression is not high enough for $1 < \theta < 2$ to have a negative indirect effect on prices. In the empirical sections that follow, we examine whether there is more aggression with corruption and we estimate the magnitude of the direct and indirect effects.

1.6 Reduced-Form Estimates of Price Differences

In this section, we bring the results of the model to the data. Our first exercise is to examine the correlation of our measures of corruption with the prices of the contracts. To do so we need to first measure corruption by auctioneers and bidders. Our second exercise is to decompose the effect of corruption into direct and indirect components, which allows us to test our model.

We start by normalizing prices and including regional fixed effects, goods and services categories fixed effects, year fixed effects, and month fixed effects; in line with both the literature on inefficiencies in procurement (Bandiera, Prat, and Valletti, 2009; Best, Hjort, and Szakonyi, 2017) and the literature on observed heterogeneity in auctions (Haile, Hong, and Shum, 2003).

We can still be concerned that omitted variables associated with corruption could directly affect prices. Hence, we also control for the type of procedure (paper or electronic), for the type of public body, for the number of bidders, and for quintiles of the reserve price.[47] We regress the log of prices and the log of bids on the controls and on the measures of corruption.

Finally, we can be concerned that bidders, who participate in auctions with corrupt and noncorrupt auctioneers are fundamentally different from each other. One main difference is selection based on in their costs. We address this selection problem by including bidders fixed effects. Bidders fixed effects allow us to control for any unobserved heterogeneity in prices coming from the type of the bidder. Since we fix the identity of the bidder, the estimates compare the behavior of the same bidder with corrupt and noncorrupt auctioneers.[48]

[47] We use 86 regions in our sample, and 88 first-level categories of goods and services.

[48] We also controlled for the number of bidders in the auction, and it did not change the

A potential remaining concern is that an unobservable characteristic of auctioneers, which causes higher prices, and which is positively correlated with corruption, can bias our estimates upward. It is very unlikely that such a variable exists, because we control for the the choice of the good, we control for differential entry of bidders, and for the reserve price, conditional on the choice of good.

In the second exercise, we want to understand whether honest bidders indeed bid more aggressively, when they expect corruption to happen, and we want to estimate the magnitude of this effect. In order to do so, we include measured corruption of the auctioneer, measured corruption of the bidder, and the interaction between the two in the price regression. We expect that for honest bidders, corruption of auctioneers will have a price-reducing effect, while for corrupt bidders it will inflate the price, and the direct effect will dominate. We can also include bidders fixed effects to control for bidder heterogeneity and estimate the interaction between the fixed effects and corruption as an additional check.

1.6.1 Measuring Corruption for Auctioneers and Bidders

In this Section, we allow the level of corruption—α—to vary between auctioneers and bidders. Thus, we can estimate α for different auctioneers by using either α_I or α_{II}. Likewise, we can estimate α for a bidder.[49] Once we have measures of corruption for the auctioneer and for the bidder, we can correlate those with prices of the contracts.

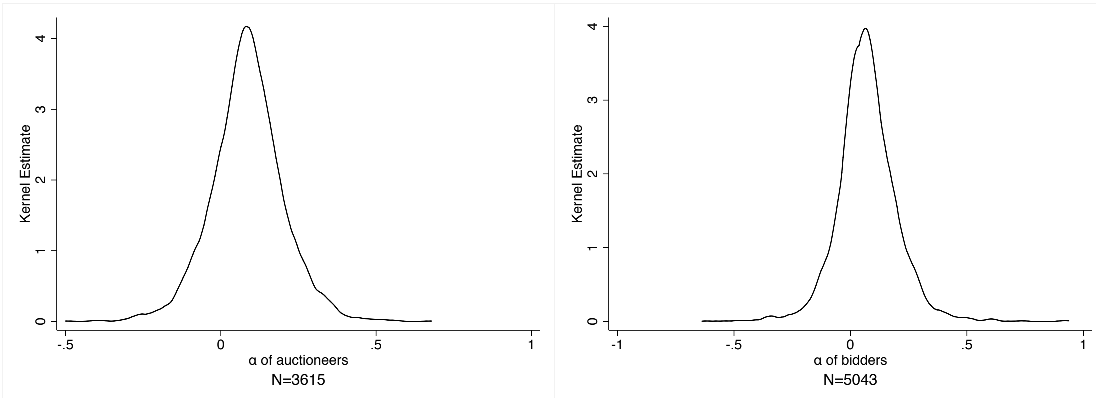
To implement this, we need to have enough power to estimate α for a participant. That is, an auctioneer or a bidder should have participated in enough auctions to have enough power to estimate α.[50] When we keep the frequent par-

results.

[49]Note that corruption of the bidder is not a structural parameter of the model. Still, it will be informative of the asymmetry of winning probabilities for a given bidder, and hence informative on how often she was involved in corruption.

[50]We cut the sample based on the following constraint: a bidder should have bids placed in at

ticipants, we are left with 3,615 auctioneers and 5,043 bidders. Figure 6 shows the distribution of α. Most of the mass is positive, in line with the positive aggregate estimates of α in Section 4.[51]



Panel A: Auctioneers Panel B: Bidders

Figure 1.6: Distribution of $\hat{\alpha}$ for Auctioneers and Bidders

1.6.2 Price Differences

Once we have the measures of corruption for auctioneers and bidders, we want to study how prices of contracts are changing in auctions with bid leakage. First, we run the following regression:

$$\log(price_{aig}) = \theta_0 + \theta_1 \cdot \alpha_a + \theta_2 \cdot X_{ai} + \mu_g + \nu_{aig}, \quad (3)$$

where a indexes auctioneers and i is an auction index. X_{ai} is a set of control variables, μ_g are goods fixed effects. Note that we do not observe corruption directly; that is, we cannot use α_a, so we have to use our estimates of corruption $\hat{\alpha}_{Ia}$.

least 200 auctions; at least 30 of which had three or more bidders. The auctioneer should have at least 200 auctions, at least 30 of which had three or more bidders.

[51]Moreover, the median and mean values of the estimates for the subsamples of frequent participants are 0.074 and 0.082 respectively, which is in line with the estimates for the whole sample in Section 4.

The coefficient of interest in (3) is θ_1. The interpretation of it is partial correlation of corruption and prices. We start by examining raw and partial correlations in Table 8, Columns (1) and (2), respectively. Both estimates are positive and significant, indicating that an increase in corruption is correlated with an increase in prices. Even though we included goods fixed effects, and the set of controls also includes regional fixed effects, years and months fixed effects, a dummy for change in the regulations, and a dummy for a number of bidders in the auction, we are still cautious about interpreting it as a causal effect. Instead, we add bidders fixed effects for each bidder ξ_b, and then rerun (3) as

$$\log(price_{aigb}) = \tilde{\theta}_0 + \tilde{\theta}_1 \cdot \alpha_a + \tilde{\theta}_2 \cdot X_{ai} + \tilde{\mu}_g + \xi_b + \tilde{\nu}_{aigb}, \quad (4)$$

The resulting coefficients of interest are in Column (3) of Table 8. Including bidders fixed effects only increases the coefficient, which is now equal to 0.275.

We get a back-of-the-envelope calculation for price changes from it. Public bodies with different levels of bid leakage pay different prices for the contracts. A public body with $\hat{\alpha} = 0.25$, which corresponds to the 97th percentile, ends up paying 7.1% higher prices than a public body with no bid leakage. If we shift all of the auctioneers with positive estimates of corruption to no corruption, that is, from $\hat{\alpha} = 0.103$ to zero, the savings are still sizable: 2.9%.

1.6.3 Testing the Model

The estimates in Table 8 pool the direct and indirect effects of corruption from the model. We want to test the extent to which the indirect effect of corruption contributes to the price estimates. To do so, we note that if we fix an identity of a bidder to be an honest bidder, she should reduce her bid when she faces a corrupt auctioneer. At the same time, if it is a corrupt bidder facing a corrupt auctioneer, the prices and the bids should both go up. We test it by running following regressions:

$$\log(bid_{aigb}) = \beta_0 + \beta_1 \cdot \alpha_a + \beta_2 \cdot \alpha_b + \beta_3 \cdot \alpha_a \cdot \alpha_b + \beta_4 \cdot X_{ai} + \eta_g + \epsilon_{aigb}, \quad (5)$$

$$\log(price_{aigb}) = \tilde{\beta}_0 + \tilde{\beta}_1 \cdot \alpha_a + \tilde{\beta}_2 \cdot \alpha_b + \tilde{\beta}_3 \cdot \alpha_a \cdot \alpha_b + \tilde{\beta}_4 \cdot X_{ai} + \tilde{\eta}_g + \tilde{\epsilon}_{aigb}, \quad (6)$$

where b indexes bidders, and the rest of the notations are as before.

The coefficients of interest in the regressions are β_1, β_2, and β_3 for the first regression, and $\tilde{\beta}_1, \tilde{\beta}_2$, and $\tilde{\beta}_3$ for the second one. If our model is correct, absent bidder corruption, auctioneer corruption will lead to more aggression—$\beta_1 < 0$. If a corrupt bidder and a corrupt auctioneer face each other, the bid should go up on average due to a direct effect on prices—$\beta_3 > 0$. The direct effect of corruption will be a linear combination of β_1, β_2, and β_3.

We run the equations above for two subsamples. The first one involves auctioneers and bidders with higher than median participation.[52] The second subsample is complementary to the first one. The predictions should hold only if players bid according to the equilibrium described in our model. We expect that they behave as in the model in the subsample of frequent participants. We start by estimating the equations (5) and (6) for the subsample of frequent participants, and we report the results in Table 9.

Table 9 shows the results for both bid-level (5) and price-level (6) regressions for frequent participants. Columns (1) and (2) show a raw and partial correlation of the logarithm of bids and the corruption of the auctioneer. Both of them are positive, statistically significant, and of the same order of magnitude as they are for the prices for the whole sample. The coefficients for prices in Columns (6) and (7) are very similar, though the partial correlation is insignificant.

Column (3) shows our baseline specification for the bids. Both of the baseline

[52] Specifically, these are auctioneers that had at least 685 bids placed in their auctions and participants that put their bids in at least 361 auctions.

coefficients for corruption of auctioneers and bidders are negative and significant $\hat{\beta}_1, \hat{\beta}_2 < 0$, while the interaction term is positive and significant $\hat{\beta}_3 > 0$, as predicted by the model. One way to interpret the magnitude is to see how moving from no corruption $\alpha_a = 0$ to $\alpha_a = 0.25$ changes the bids. Such a change increases the bids by 10.5%.

If we fix the level of corruption for the bidder and the auctioneer as equal to each other ($\alpha_a = \alpha_b$), there is a nonzero level of corruption that minimizes the price. This level is 0.087, close to the median in the sample, but still below the mean for the auctioneers that have a positive level of corruption, 0.103. If instead of moving from $\alpha = 0.25$ to $\alpha = 0$, we move from $\alpha = 0.103$ to $\alpha = 0.087$, the bids will drop by 0.12% of the initial level.

While we treat Column (3) as our main specification for (5), an unobservable characteristic of the contract can exist that is not captured by geography, type of good or service, time, type of public body, law, or competition measured by the number of bidders. One characteristic that may capture this unobservable is the reserve (maximum) price of the contract. The main caveat here is that we cannot directly control for the log-reserve price, because it can be affected by corruption directly and thus is a *bad control*. Corrupt public bodies can inflate the reserve price, and so including it in the regression will be over controlling (see Angrist and Pischke, 2008; Maccini and Yang, 2009). Moreover, the correlation between the bid and the reserve price is 0.95, with an OLS coefficient close to 1, so including it on its own will drain all of the variation. However, we control for the quintiles of the reserve price to address unobserved variables related to the size of the contract and potentially other unobserved characteristics of the contract. Column (4) of Table 9 shows the results controlling for the quintiles of the reserve price. The coefficients get smaller in magnitude but preserve similar patterns as in Column (3). Moreover, the coefficient on bidder corruption shrinks by a factor of two and is only significant at the 10% level.

The richness of our data allows us to pin down all of the unobserved variation coming from the bidder side by directly including bidders fixed effects, instead of including the measure of corruption α_b. The magnitude of the effects in Column (5) shrink, but they remain both economically and statistically significant.[53]

We tested our model and measured the effects on the average bid, rather than the effective final prices of the contract. Columns (6—10) of Table 9 repeat the analysis for final prices (as in equation (6)). The sizes of the coefficients in Columns (8) and (9)—with and without reserve price controls—are larger in magnitude than in corresponding Columns (3) and (4) and are in line with the model. Bidders fixed effects drain a lot of the variation from the sample and make the coefficients imprecisely estimated, keeping the signs of the coefficients in line with the model (Column 10).

Table (9) provides a basis for another important back-of-the-envelope calculation. We use the estimates in Column (8) to derive them. Moving from no corruption to $\alpha = 0.25$ for both bidders and auctioneers increases prices by 15.0%. At the same time, prices are minimized for the level of corruption $\alpha = 0.053$. Moving from the average of auctioneers with positive corruption 0.103 to this minimum level reduces prices by 0.4%.

The results are entirely different for the subsample of infrequent participants, as reported in Table 10. The coefficient on auctioneer corruption is smaller in magnitude compared to Table 9 and positive. All of the interaction terms are insignificant, whether we take the bids (Columns 1—5) or the prices (Columns 6—9). The exception is Column (10), with a negative and significant coefficient on the interaction term. If we do the back-of-the-envelope calculations for this

[53] Note that the coefficient on $\hat{\alpha}_b$ is not identified with bidders fixed effects, however the coefficient on the interaction term $\hat{\alpha}_b \cdot \hat{\alpha}_a$ is identified and it is our quantity of interest in Columns (5) and (10). One can be potentially concerned that it hides a heterogeneity in bidder effect. In addition, in Appendix C, Figure C2 we show the distribution of the interaction terms of bidder fixed effects and corruption of the auctioneers. As can be seen from the graph, those coefficients are predominantly positive.

subsample, switching from zero corruption to $\alpha = 0.25$ increases prices by 9.8%, while moving from $\alpha = 0.103$ to zero corruption reduces prices by 5.3%.

1.6.4 Comparison to Existing Estimates of Corruption

Before proceeding to structural estimation in the next section, we compare our baseline estimates with the estimates of procurement corruption in the existing literature. Our baseline effect on prices is 7.1% if we switch from high corruption to zero corruption, and it is 2.9% if we reduce corruption for all of the auctioneers who have a positive measure of corruption. Both of the numbers are lower than the results from similar contexts in Di Tella and Schargrodsky (2003) and Bandiera, Prat, and Valletti (2009), but not drastically so. The first paper estimates the effect of the crackdown on corruption on procurement prices in Argentinean hospitals and finds and effect of 10%, while the second paper estimates active waste in Italian procurement, with a corresponding price difference of 11%. Other papers (Ferraz and Finan, 2011; Olken, 2006, 2007; Reinikka and Svensson, 2005) study diverted funds in Brazilian local expenditures, loss in subsidized rice and missing road construction expenditures in Indonesia, and diverted education funds in Uganda—show loss estimates ranging from 9% to 24%. Our estimates are lower, but again not dramatically so.

None of the existing studies decompose the effect of corruption into direct and indirect effects. For subsample of frequent bidders, the indirect effect shifts the optimal corruption level from zero to a positive number. In this specific subsample, going from the mean level to the optimal level will reduce prices by 0.4%, but completely eliminating corruption will raise prices.

1.7 Identification and Estimation of the Structural Model

Reduced-form results show that corruption increases prices for the whole sample, with an important heterogeneity in the results. For the subsample of frequent auctioneers and bidders an indirect price-reducing effect of corruption is present. We want to document that these reduced-form changes come from the model in Section 5. Our main objects of interest are the equilibrium bid functions for honest bidders. We want to establish that bid functions of honest bidders, when they face corrupt auctioneers shift downward compared to bid functions of honest bidders, when they face honest auctioneers. That is, for the same level of costs, honest firms in corrupt auctions reduce their bids.

To provide preliminary evidence on the extent of the equilibrium response of honest bidders, we use our model of bidding behavior with corruption and incorporate the identification results of Guerre, Perrigne, and Vuong (2000). While our main empirical contribution from the previous sections to do not rely on the first-order condition, we need to use the FOC from Section 5 to estimate the bid functions.

The first-order condition with corruption is

$$c = b - \chi(\alpha, b) =$$

$$b - \Big[\frac{(K-2)g_\alpha(b)}{1-G_\alpha(b)} + \frac{(1-\alpha)g_\alpha(b) + \alpha f(b)}{(1-\alpha)(1-G_\alpha(b)) + \alpha(1-F(b))}\Big]^{-1}. \quad (2)$$

As before, f and F are the density and the cumulative distribution function of the costs. We modify notation to stress that the distribution of bids depends on the level of corruption; we use a notation G_α and g_α with α index. In this case, the distribution of bids without corruption is G_0 (with the probability density function g_0).

The FOC implicitly defines the central quantity of interest—the bid function

$\beta_\alpha(c)$, which depends on the level of corruption α. To estimate this bid function, we need to observe and estimate the distributions F and G_α, as well as the level of corruption α. We proceed in four steps.

(1) First, we measure corruption for all auctioneers and bidders who have enough data. This gives us the measures $\hat{\alpha}_a$ and $\hat{\alpha}_b$ as in Section 6.

(2) Second, we choose the subsample of honest bidders $\hat{\alpha}_b \approx 0$ and honest auctioneers $\hat{\alpha}_a \approx 0$. Note that in this noncorrupt subsample, the FOC boils down to the standard first-price sealed-bid auction FOC, with

$$\chi(\alpha, b) = \Big[\frac{(K-1)g_0(b)}{1-G_0(b)}\Big]^{-1}.$$

Hence, all of the identification and estimation results from Guerre, Perrigne, and Vuong (2000) hold. Thus, we can estimate the distributions G_0, g_0. Then, we can generate pseudo-costs $\hat{c}$, and estimate the distribution of F, and density f using the generated pseudo-costs.

(3) Third, we vary α for auctioneers, while keeping $\hat{\alpha}_b \approx 0$. That is, since the FOC holds only for honest bidders, we want to examine their behavior. At the same time, we want to study what happens when we move the level of corruption for the auctioneer. This allows us to estimate G_α for a fixed level of corruption α simply from the empirical CDF. Likewise, we get consistent estimates for the density of bids g_α.

(4) Fourth, we already have all of the estimates $\hat{F}$, $\hat{f}$, $\hat{g}_\alpha$, $\hat{G}_\alpha$, and $\hat{\alpha}_a$, which we can plug into FOC with corruption to get $\hat{\chi} = \tilde{\chi}(\hat{F}, \hat{f}, \hat{G}_\alpha, \hat{g}_\alpha, \hat{\alpha}_a, \hat{p}; b)$. Once we do it, we immediately have the estimate for the bid function. Theorem 1 of Guerre, Perrigne, and Vuong (2000), with $b - \chi(\alpha, \cdot)$ instead of $\phi_0(b)$, provides us with the identification result. Plug-in estimators will be consistent in this case.

(4') We also employ an alternative approach: since we assumed that the distribution of costs is the same for any α; we can solve $G_\alpha(b) = F(\phi_\alpha(b)$, where

ϕ is an inverse bid function. This equation can be solved by inverting $\hat{F}$. The resulting estimate for the inverse bid function will be $\hat{\phi}_\alpha(b) = \hat{F}^{-1}(\hat{G}_\alpha(b)$.

1.7.1 Practical Considerations

In the previous subsection, we held the number of bidders fixed at K. Ideally, we want to control for the number of bidders, which varies from auction to auction. We do assume that bidder entry happens exogenously and the distribution of the number of bidders is geometric with parameter p on $\{3, ...\}$. Note that p is not a function of α; that is, corruption on its own does not discourage bidders from participating. Although the assumption might seem strong, the average number of bidders does not differ for auctions with different measured α of auctioneer.

In the case of random entry, expected profit changes to

$$\tilde{\Pi}(b) = (b - c)((1 - \alpha)(1 - G(b)) + \alpha(1 - F(b)))\mathbb{E}_K[(1 - G(b))^{K-2}].$$

For geometric distribution, the FOC becomes[54]

$c = b - \tilde{\chi}(\alpha, b, p) =$

$$b - \Big[\frac{(1-\alpha)g_\alpha(b) + \alpha f(b)}{(1-\alpha)(1 - G_\alpha(b)) + \alpha(1 - F(b)} + \frac{g_\alpha(b)}{1 - G_\alpha(b)} + \frac{(1-p)g_\alpha(b)}{1 - (1-p)(1 - G_\alpha(b))}\Big]^{-1} \quad (2')$$

The last component of the modified FOC is the entry component. We estimate $\hat{p} = 0.49$ and use it throughout.

We assumed that auctions are homogeneous when we outlined the structural model. In the estimation, we relax this assumption and follow a first-stage-regression approach common in the literature (Haile, Hong, and Shum, 2003; Bajari, Houghton, and Tadelis, 2014; Asker, 2010a). At the first stage, we project

[54]The last multiplier in this expression transforms to $\mathbb{E}_K[(1 - G(b))^{K-2}] = \frac{p(1-G(b))}{1-(1-p)(1-G(b))}$. We take logarithms of the expression and we maximize it with respect to b.

the logarithms of bids on the covariates. Note, that it is only possible if the observed heterogeneity enters the valuations in a multiplicatively separable way. Once we do the projection, we run the rest of the analysis with the exponentiated residuals. For consistency, we use the same covariates as in Section 6; that is, dummy variables for regions, for types of goods and services, for years and months, a dummy for FZ#44, and a dummy for whether the public body is a medical or educational institution.

1.7.2 Results

We implement the procedure from the previous subsection. In order to estimate the fundamentals of the model, we need to pick cutoffs. First, we pick a subsample of auctioneers and bidders who took part in a sufficient number of auctions, as in Section 6. We concentrate only on the bidders with low α_b (below the median ≤ 0.06).

Next, for estimating the pseudo-costs in noncorrupt auctions, we need to pick the auctioneers with no corruption. We take 30% of auctioneers with the lowest α_a to estimate the pseudo-costs distribution F. Once, we have the estimate $\hat{F}$, we estimate G_α auctioneer-by-auctioneer.

To illustrate our results, we choose several auctioneers with a sufficient number of auctions and with different values of $\hat{\alpha}_{II} \approx 0$. The table below illustrates our choice.

Number of Auctions	4,175	2,266	2,333	1,980	568	572
$\hat{\alpha}_a$	0.05	0.13	0.21	0.08	0.00	-0.00

Figure 7 shows the resulting estimates for the bid functions. The first four panels show the results for the auctioneers with a positive level of corruption. In three of four cases, corruption moves the bid functions toward more aggression: in one case, it is inconclusive what happens to the bid functions. Still, the evidence points to more aggression. For auctioneers with a low level of corruption, there is no evidence of more aggression, which is also in line with the mechanisms from the model.

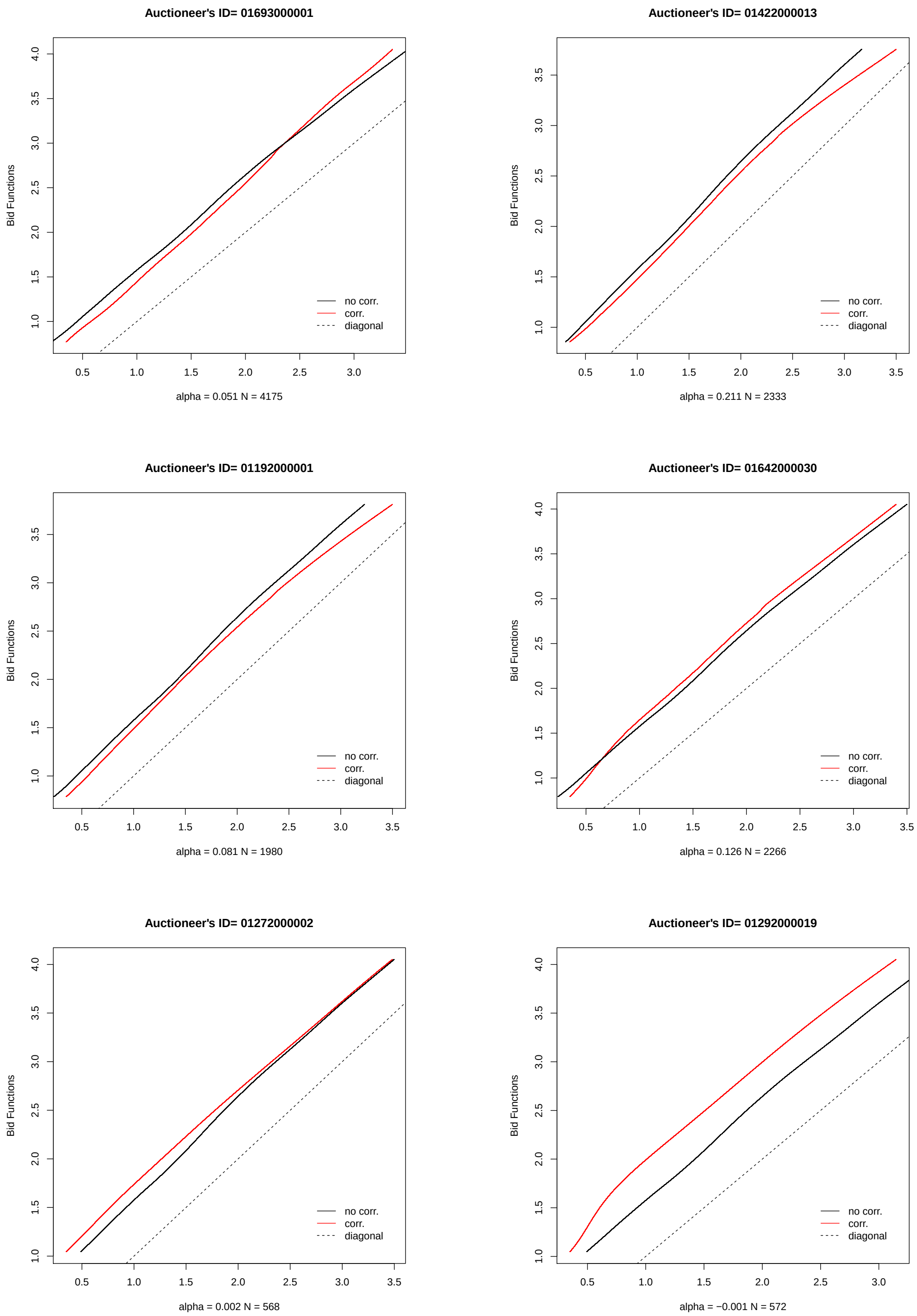


Figure 1.7: Bid Functions for Different Auctioneers, FOC Method

Estimating the model auctioneer-by-auctioneer suggests that we do observe more aggression, but the results are not robust across auctioneers. Finally, we pool auctioneers whose level of α_{II} is in a neighborhood of 0.02 of the median value, $\alpha = 0.09$, and rerun both of the methods. We plug in $\alpha = 0.09$ directly in the FOC. Panel A of Figure 8 shows the results. While parts of the bid function lie below the bid function for no corruption, some parts are above the no-corruption bid function.

We also check what happens when we invert the CDFs at step 4, instead of plugging the estimates into the FOC with positive corruption. The resulting estimates from the CDF method are shown in Figure 8, Panel B. The bid function is below the no-corruption case, and in line with more aggression.

To sum up, we developed a way to test a model of first-price sealed-bid auctions with bid-leakage corruption, and we show that the estimates of bid functions suggest that the indirect effect of corruption is at play.

Panel A: FOC Panel B: CDF method

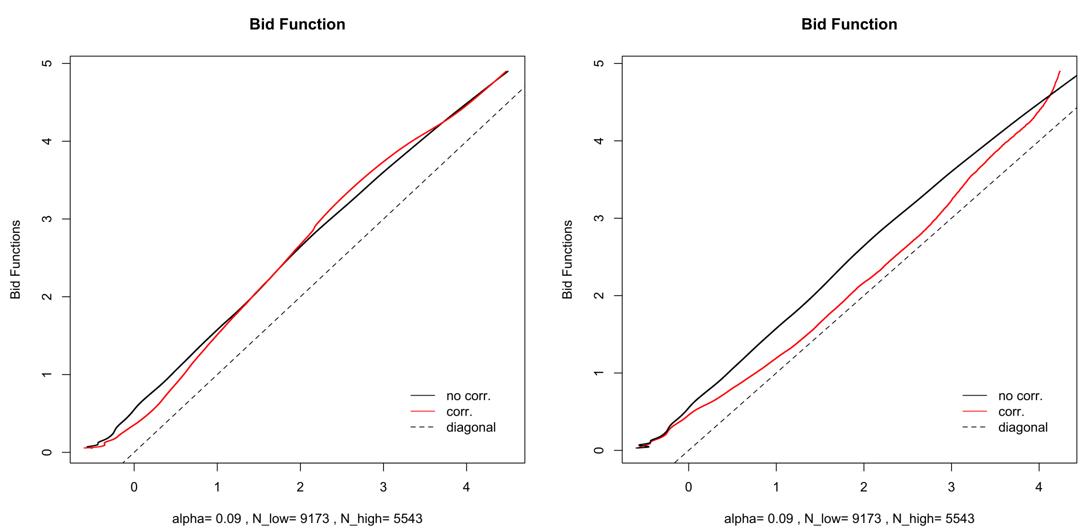


Figure 1.8: Bid Functions for No Corruption and for Pooled Auctioneers

1.8 Conclusion

In this paper, we develop methods to measure corruption focusing on timing of bids. These methods allow us to document large-scale bid-leakage corruption prevalent in Russian procurement auctions. From 2011 to 2016, the cumulative value of the affected contracts was as high as $1.2 billion. The regression of prices on measures of corruption suggests that eliminating bid leakage could have lowered prices by 2.9%. For our subsample of frequent auctioneers and bidders, another effect of corruption—the indirect effect—lowered the prices paid to the honest bidders, who adjusted their bidding behavior to attenuate corruption.

Our paper is the first to provide estimates of this type of corruption empirically, and to estimate the effects of such behavior on prices. We also stress that the change in equilibrium behavior of honest bidders is important, when the monitoring authorities eliminate bid leakage.

We still can discuss several policy implications. The first involves changing the way that contracts are distributed. Running an open auction or a second-price sealed-bid (Vickrey) auction could help remedy bid-leakage corruption. The equilibrium outcomes of these auctions should make participants bid their true costs. Therefore, the preferred bidder and the honest bidders will behave similarly, which would, at first glance, solve the problem of bid leakage. However, open auctions and Vickrey auctions both have problems of their own; both can facilitate collusion between bidders (Robinson, 1985). In the environment of Russian procurement, collusion can ensure even larger losses. In addition, Vickrey auctions come with privacy concerns: auctioneers can invite shill bidders to reduce the margin between the lowest bids. Moreover, once a bidder reveals her true costs, the auctioneer may alter his behavior in the future.[55]

As opposed to changing the auction mechanism, employing technological ad-

[55] For a discussion of why Vickrey auctions are rarely used, see Rothkopf et al. (1990).

vancements is likely the best way forward. First, introducing an encrypted electronic system could reduce bid leakage. Second, tracing the timing of bids to monitor auctioneers could prove useful in targeting this type of corruption.

Appendix B: Proofs

Consistency and Normality of Method II Share Estimator:

$$\hat{\alpha} = \frac{1}{|T|}\sum_{t\in T}\frac{\hat{G}_{b_{(1)}}(t) - \hat{G}_{b_{(2)}}(t)}{H(t) - \hat{G}_{b_{(2)}}(t)} = \frac{1}{|T|}\sum_{t\in T}\frac{\frac{1}{N}\sum 1\{t_{b_{(1)}} < t\} - 1\{t_{b_{(2)}} < t\}}{\frac{1}{N}\sum NH(t) - 1\{t_{b_{(2)}} < t\}},$$

First let's analyze

$$\hat{a}(t) = \frac{\frac{1}{N}\sum 1\{t_{b_{(1)}} < t\} - 1\{t_{b_{(2)}} < t\}}{\frac{1}{N}\sum NH(t) - 1\{t_{b_{(2)}} < t\}}.$$

By uniform consistency of the empirical CDF and Slutsky theorem,

$$\operatorname*{plim}_{N\to\infty} \hat{a}(t) = \frac{G_{b_{(1)}}(t) - G_{b_{(2)}}(t)}{H(t) - G_{b_{(2)}}(t)} = \alpha.$$

Since $\hat{\alpha}$ is an average of consistent estimators, it is itself consistent.[56] Now let's consider

$$\sqrt{N}(\hat{a}(t) - \alpha) = \frac{\frac{1}{\sqrt{N}}\sum 1\{t_{b_{(1)}} < t\} - 1\{t_{b_{(2)}} < t\}}{\frac{1}{N}\sum NH(t) - 1\{t_{b_{(2)}} < t\}} \quad (C)$$

$1\{t_{b_{(1)}} < t\} - 1\{t_{b_{(2)}} < t\}$ is not correlated across auction. Moreover, since we assumed $b \perp t$,

$$Var[1\{t_{b_{(1)}} < t\} - 1\{t_{b_{(2)}} < t\}] = G_{b_{(1)}}(t)(1 - G_{b_{(1)}}(t)) + G_{b_{(2)}}(t)(1 - G_{b_{(2)}}(t)).$$

[56] For two estimators α_n and β_n such that $\operatorname{plim}\alpha_n = \alpha$ and $\operatorname{plim}\beta_n = \beta$, one can write a triangular inequality $|(\alpha_n + \beta_n) - (\alpha + \beta)| \le |\alpha_n + \beta_n| + |\alpha + \beta|$ to derive the result.

The numerator of (C) converges to $H(t) - G_{b_{(2)}}(t)$. By C.L.T.,

$$\sqrt{N}(\hat{a}(t) - \alpha) \xrightarrow{d} \mathbb{N}(0, \mathbb{V}_\alpha)$$

$$\mathbb{V}_\alpha = \frac{G_{b_{(1)}}(t)(1 - G_{b_{(1)}}(t)) + G_{b_{(2)}}(t)(1 - G_{b_{(2)}}(t))}{(H(t) - G_{b_{(2)}}(t))^2}$$

A confidence interval for a fixed t is

$$CI_{0.05}(t) = \hat{a}(t) \pm \frac{\mathbb{V}_\alpha(t) \cdot z_{N(0,1)}}{\sqrt{N}}.$$

Hence we built a confidence interval for $\hat{a}(t)$, for a fixed t. The two functions that we are using for estimation are $H(t) = t/\epsilon$ and $H(t) = 1$.

Now we need to find the asymptotic variance of $\hat{\alpha}$. Consider $\sqrt{N}(\frac{1}{|T|}\sum_{t\in T}\hat{a}(t) - \alpha)$. This can be rewritten as a sum of asymptotically normal estimators. Now the question is whether they are asymptotically independent.

$$\lim_{N->\infty} cov(\hat{a}(t), \hat{a}(s)) = \lim_{N->\infty} \mathbb{E}[\hat{a}(t) \cdot \hat{a}(s)] - \alpha^2 =$$

$$\lim_{N->\infty} \mathbb{E}\Big[\frac{\mathbb{E}[(\hat{G}_{b_{(1)}}(t) - \hat{G}_{b_{(2)}}(t))(\hat{G}_{b_{(1)}}(s) - \hat{G}_{b_{(2)}}(s))|t_{b_{(1)},1}, ..., t_{b_{(1)},N}]}{(1 - \hat{G}_{b_{(2)}}(t))(1 - \hat{G}_{b_{(2)}}(s))}\Big] =$$

$$\lim_{N->\infty} \mathbb{E}\Big[\frac{(\hat{G}_{b_{(1)}}(t) - \hat{G}_{b_{(2)}}(t))(\hat{G}_{b_{(1)}}(s) - \hat{G}_{b_{(2)}}(s))}{(1 - \hat{G}_{b_{(2)}}(t))(1 - \hat{G}_{b_{(2)}}(s))}\Big],$$

where the last equality follows from independence of t's. Applying Slutsky theorem several times,

$$= \frac{(G_{b_{(1)}}(t) - G_{b_{(2)}}(t))(G_{b_{(1)}}(s) - G_{b_{(2)}}(s))}{(1 - G_{b_{(2)}}(t))(1 - G_{b_{(2)}}(s))} = \alpha^2.$$

Hence asymptotic variance of $\hat{\alpha}$ is the weighted average of each of the variances as a function of t.

Results from the Model:

Definition 1, Log-concavity of survival function: $1-F(x)$ is log-concave. $\Leftrightarrow \frac{1-F(x)}{f(x)}$ is decreasing.[57][58]

Proposition 1: If the survival function is log-concave, the FOC (2) defines an equilibrium with corruption.

$c = b - \chi(\alpha, b) :=$

$$= b - \Big[\frac{(K-2)g(b)}{1-G(b)} + \frac{(1-\alpha)g(b) + \alpha f(b)}{(1-\alpha)(1-G(b)) + \alpha(1-F(b))}\Big]^{-1}. \quad (2)$$

Proof: The first step is to show that it holds for $\alpha = 1$, in other words, to restate the Arozamena and Weinschelbaum (2009) result for the case of procurement auctions.

For the case of two bidders $K = 2$, and $\alpha = 1$ complete corruption, (1) takes a following form,

$$c = b - \frac{1-F(b)}{f(b)}.$$

Log-concavity of the survival function assures that the right-hand side of this equation is increasing; it does not depend on the strategies of other player, and thus it gives a unique strictly increasing bid function by dominance.

For $K > 2$ but $\alpha = 1$, we follow similar steps as in Arozamena and Weinschelbaum (2009) and Li and Tan (2000). The FOC can be rewritten as

$$\phi'(b) = \frac{(1-F(\phi))\Big[1 - \frac{f(b)}{1-F(b)}(b-\phi)\Big]}{(K-2)f(\phi)(b-\phi)}.$$

[57] In the weaker condition of regularity that is common in the auction literature (Myerson 1981): the virtual valuation $x - \frac{1-F(x)}{f(x)}$ is weakly increasing. Regularity is implied by log-concavity of survival function, but not the other way.

[58] For power function $F(x) = x^{\theta}$, both regularity and log-concavity holds for $\theta > 1$.

Note that we can apply a following transformation of variables: $v = p_{max} - c$ and $b^v = p_{max} - b$ for some large number p_{max}. For instance $p_{max} = \bar{c}$. Note that

$$F(c) = Pr(\tilde{<}c) = Pr(p_{max} - \tilde{v} < p_{max} - c) = 1 - Pr(\tilde{v} < v) = 1 - F(p_{max} - v) = 1 - F^v(v).$$

This also implies $f(c) = -f^v(v)$ and log-concavity of $1 - F(c)$ is equivalent to log-concavity of $F^v(v)$. Hence the FOC can be rewritten as

$$\phi'^v(b) = \frac{F^v(\phi^v)\left[1 - \frac{f^v(b)}{F^v(b)}(\phi^v - b^v)\right]}{(K-2)f(\phi^v)(\phi^v - b^v)}.$$

The rest of the proof follows Appendix A of Arozamena and Weinschelbaum (2009). The case of $0 < \alpha < 1$ is somewhat different. Note that the FOC can be rewritten as

$$\beta'(c) = \frac{(1-\alpha)(1-F(c)) + \alpha(1-F(b))}{(1-\alpha)(1-F(c)) + \alpha(1-F(b)) - \alpha f(b)(b-c)} \frac{f(c)(b-c)}{A(c,b)},$$

where $A(c,b) = \frac{(K-2)f(c)}{1-F(c)} + \frac{(1-\alpha)f(c)}{(1-\alpha)(1-F(c)+\alpha(1-F(b))}$.

The denominator is never zero (as opposed to the previous case $\alpha = 1$), so we can apply a regular existence theorem for differential equations (see Filippov 1971).

Proposition 2: If $w(x) = \frac{1-F(x)}{f(x)}$ is strictly convex, there is more aggression.

Proof: Note that if we proof Lemma 1 from Appendix B of Arozamena and Weinschelbaum (2009) for our case, the rest of the proofs follow by the change of variables.

Note that the first order condition can be rewritten as

$$b-\phi(b)=\frac{(1-\alpha)(1-G(b))+\alpha(1-F(b))}{(1-\alpha)(K-1)g(b)+\alpha(f(b)+(K-2)g(b)\frac{1-F(b)}{1-G(b)})}.$$

Note that for positive a, b, c, d, $ad < bc$, and $\alpha \in (0,1)$, the following is true:

$$\frac{a}{b}<\frac{\alpha a+(1-\alpha)c}{\alpha b+(1-\alpha)d}<\frac{c}{d} \quad (*).$$

Hence we need to show

$$\frac{1-G(b)}{(K-1)g(b)}>\frac{1-F(b)}{f(b)+(K-2)g(b)\frac{1-F(b)}{1-G(b)}},$$

which is equivalent to

$$w(b)<w(\phi)/\phi'.$$

Hence $b-\phi<\frac{w(\phi)}{\phi'(K-1)}$ is equivalent to $w(b)<w(\phi)/\phi'$.

Likewise,

$$b-\phi>\frac{1-F(b)}{f+(K-2)g(b)\frac{1-F(b)}{1-G(b)}}>\frac{w(b)}{K-1}.$$

Combining the two together,

$$\frac{w(b)}{K-1}<b-\phi<\frac{w(\phi)}{\phi'(K-1)}\Leftrightarrow w(b)<\frac{w(\phi)}{\phi'}.$$

In addition, note that (*) implies for $\alpha \in (0,1)$ and $\omega \in (0,1)$,

$$b-\phi^{\alpha}(b)=\omega(b-\phi^{0}(b))+(1-\omega)(b-\phi^{1}(b))=\omega\frac{a}{b}+(1-\omega)\frac{c}{d},$$

and

$$\partial\omega/\partial\alpha > 0.$$

Let $0 < \alpha' < \alpha'' < 1$, then $\exists \omega' < \omega''$ such that

$$b - \phi^{\alpha'}(b) = \omega'(b - \phi^0(b)) + (1 - \omega')(b - \phi^1(b))$$

and

$$b - \phi^{\alpha''}(b) = \omega''(b - \phi^0(b)) + (1 - \omega'')(b - \phi^1(b)).$$

And hence $\phi^{\alpha''}(b) > \phi^{\alpha'}(b)$, and aggression is monotonic in α.

Table 1.1: Chapter 1 Tables

Appendix C: Figures and Tables

Table 1.2: Summary Statistics

Panel A: Whole Sample

	Variable	Mean	Median	Std. Dev.
Number of bidders	K	4.0	3	1.6
Reserve price, rubles	R	213,507	180,000	153,604
Winner's bid, rubles	$b_{(1)}$	157,850	125,412	124,287
Winner's time to the deadline, hours	$t_{b_{(1)}}$	27.8	4.4	47.1
Winning bid to reserve price	$r_{(1)} = \frac{b_{(1)}}{R}\%$	73.3%	76.5%	18.0%
Winning to second-best bid distance	$db_{(12)} = \frac{b_{(2)}-b_{(1)}}{b_{(2)}}\%$	5.6%	2.3%	8.1%

Notes: The sample includes only auctions with three or more bidders and no rejected bids. $N = 841,552$ auctions

Panel B: Only Law #44

	Variable	Mean	Median	Std. Dev.
Number of bidders	K	3.9	3	1.5
Reserve price, rubles	R	193,233	149,919	150,017
Winner's bid, rubles	$b_{(1)}$	134,112	95,460	116,138
Winner's time to the deadline, hours	$t_{b_{(1)}}$	32.3	17.9	48.2
Winning bid to reserve price	$r_{(1)} = \frac{b_{(1)}}{R}\%$	68.8%	72.0%	19.5%
Winning to second-best bid distance	$db_{(12)} = \frac{b_{(2)}-b_{(1)}}{b_{(2)}}\%$	7.9%	4.6%	9.3%

Notes: The sample includes only auctions with three or more bidders and no rejected bids. $N = 171,539$ auctions

Table 2: Tests for Presence of Corruption

Panel A: Wald Test

	N	P[win|last]	P[win|not last]	Difference	P-value
(1)	286,251	0.397	0.301	0.096	<0.001

Notes: Panel A shows the results for the Wald test from OLS of indicator for winning on indicator for being last. The results are for auctions with three bidders, $K = 3$; however, they hold for $K > 3$. Only the FZ#44.

Panel B: Kolmogorov-Smirnov Test

H_0		N	D-statistic	P-value
$G_1 = G_2$	(2)	222,738	0.081	<0.001
$G_1 = G_3$	(3)	221,858	0.081	<0.001
$G_1 = G_4$	(4)	159,751	0.071	<0.001

Notes: Panel B presents the results of Kolmogorov-Smirnov test. $G_1 = G(t|\text{winner})$, $G_2 = G(t|\text{runner-up})$, and similar for other CDFs. D-statistic is reported from a two-sample Kolmogorov-Smirnov test. Only the last 72 hours are taken for the K-S test to avoid noise on the tails. Only the FZ#44.

Table 3: Placebo Tests

Panel A: Wald Test

		N	P[win|last]	P[win|not last]	Difference	P-value
Artificial Auctions	(1)	286,251	0.482	0.494	-0.013	<0.001

Notes: Panel A shows the results for the Wald test from OLS of indicator for running up on indicator for being second-last for the subsample with last and winning bids dropped. The results are only for auctions with three bidders, $K = 3$; however, they hold for $K > 3$. Only the FZ#44.

Panel B: Kolmogorov-Smirnov Test

H_0		N	D-statistic	P-value
$G_2 = G_3$	(2)	219,546	0.004	0.451
$G_2 = G_4$	(3)	157,439	0.012	<0.001
$G_3 = G_4$	(4)	156,559	0.011	0.001

Notes: Panel B shows the results of a Kolmogorov-Smirnov test. $G_1 = G(t|\text{winner})$, $G_2 = G(t|\text{runner-up})$, and similar for other CDFs. *D*-statistic is reported from a two-sample Kolmogorov-Smirnov test. Only the last 72 hours are taken for the K-S test to avoid noise on the tails. Only the FZ#44.

Table 4: Test I—Subsample Analysis

Subsample		N	P[win|last]	P[win|not last]	Difference	P-value
All Late, Law #44	(1)	14,591	0.435	0.282	0.153	<0.001
All Early, Law #44	(2)	82,369	0.349	0.326	0.023	<0.001
Not Early, Law #44	(3)	203,882	0.417	0.292	0.125	<0.001

Notes: p-values Panel A shows the results for the Wald test from OLS for subsamples indicated below. Row (1) shows the results for the main test, keeping only the auctions, where all of the bids were placed within the last 60 minutes of the deadline. Row (2) is for auctions where all bids were placed more than five hours from the deadline. Row (3) is for the subsample, where there were bids placed in the last five hours. The results are only for auctions with three bidders; $K = 3$, however, they hold for $K > 3$. Only the FZ#44.

Table 5: Measure I—Shares of Corrupt Auctions

	(1)	(2)	(3)	(4)
Number of bidders, K	$K = 3$	$K = 4$	$K = 5$	$K > 5$
P[win|last]	0.397	0.315	0.252	0.196
Share of corrupt auctions	0.096	0.086	0.065	0.069
P-value	<0.001	<0.001	<0.001	<0.001
N	286,251	156,748	90,444	144,517

Notes: All shares are significant at the 1% level. Measures are reported by the number of bidders. Only the FZ#44.

Table 6: Measure II—Shares of Corrupt Auctions

Panel A: Main Measure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t	0.5	1	2	5	0.5	1	2	5
α_{12}	0.083	0.109	0.121	0.123				
CI	[0.079, 0.087]**	[0.104, 0.114]**	[0.115, 0.127]**	[0.116, 0.130]**				
α_{123}					0.082	0.108	0.121	0.124
CI					[0.054, 0.078]**	[0.084, 0.103]**	[0.094, 0.115]**	[0.093, 0.117]**
N	222,738	222,738	222,738	222,738	332,071	332,071	332,071	332,071

Notes: ** - indicates significance at the 5% level. All auctions with three and more bidders, no rejected bids. α_{12} uses $G(t|\text{not winner}) = G(t|\text{runner-up})$, α_{123} uses $G(t|\text{not winner}) = (G(t|\text{runner-up})+G(t|\text{third}))/2$. The cutoff level for the sample is 72 hours. The cutoff level for the window is $t = 0.5, 1, 2, 5$. Only the FZ#44.

Table 6: Measure II—Shares of Corrupt Auctions
Panel B: Placebo Measure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t	0.5	1	2	5	0.5	1	2	5
α_{23}	-0.002	-0.002	0.002	0.003				
CI	[-0.006, 0.001]	[-0.007, 0.003]	[-0.005, 0.008]	[-0.004, 0.009]				
α_{34}					-0.009	-0.011	-0.014	-0.010
CI					[-0.014, -0.004]	[-0.018, -0.005]**	[-0.022, -0.006]**	[-0.020, -0.001]**
N	219,546	219,546	219,546	219,546	156,559	156,559	156,559	156,559

Notes: ** - indicates significance at the 5% level. All auctions with three and more bidders, no rejected bids. α_{23} uses $G(t|\text{runner-up})=G(t|\text{third})$, α_{34} uses $G(t|\text{third})=G(t|\text{fourth})$. The cutoff level for the sample is 72 hours. The cutoff level for the window is t = 0.5, 1, 2 ,5. Only the FZ#44.

Table 7: Expected Prices with and without Corruption

θ	0.5	1	2	4	10
$\mathbb{E}[price\vert\alpha = 0]$	0.5	0.67	0.8	0.89	0.95
$\mathbb{E}[price\vert\alpha = 1]$	0.63	0.73	0.81	0.83	0.78

Table 8: Corruption and Prices—Whole Sample

VARIABLES	(1) No Controls	(2) With Controls	(3) Bidder FE
α_a	0.454***	0.257***	0.276***
	(0.047)	(0.0395)	(0.0402)
Constant	11.40***	12.07***	11.95***
	(0.00642)	(0.0407)	(0.0473)
Only Winning Bids	YES	YES	YES
Controls	NO	YES	YES
Bidder FE	NO	NO	YES
# of Clusters	221,199	218,773	218,773
# of Observations	411,450	407,307	407,307
R-squared	0.001	0.142	0.51

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the auctioneer-bidder pair level. Controls include fixed effects for regions, products/services, years, months, dummy for law (#44 or #94), dummy for whether a public body is a medical or educational institution, a set of dummies for the number of bidders in an auction. α_a is a level of corruption measured for an auctioneer. Column (1) is a raw correlation between auctioneers' corruption and log of bids. Column (2) adds a set of controls including a dummy for the winning bid. Column (3) adds bidders fixed effects.

Table 9. Corruption, Bids, and Prices—Frequent Participants

VARIABLES	(1) No Controls	(2) With Controls	(3) Interaction	(4) Interaction & Reserve Price	(5) Bidder FE	(6) No Controls	(7) With Controls	(8) Interaction	(9) Interaction & Reserve Price	(10) Bidder FE
α_a	0.662***	0.183**	-0.407*	-0.279***	-0.154*	0.650***	0.204	-0.606*	-0.335**	-0.176
	(0.1)	(0.0792)	(0.235)	(0.105)	(0.0861)	(0.173)	(0.142)	(0.32)	(0.147)	(0.13)
α_b			-0.534**	-0.214**				-0.301	-0.314**	
			(0.271)	(0.106)				(0.405)	(0.147)	
$\alpha_b \cdot \alpha_a$			5.361**	1.922**	1.275*			5.863*	2.244*	1.813*
			(2.249)	(0.887)	(0.704)			(3.202)	(1.181)	(1.004)
1_{winner}		-0.206***	-0.193***	-0.162***	-0.157***					
		(0.00644)	(0.0114)	(0.00453)	(0.00401)					
Constant	11.52***	12.15***	11.92***	10.14***	10.19***	11.33***	11.92***	11.85***	9.979***	10.05***
	(0.0117)	(0.0548)	(0.143)	(0.043)	(0.0757)	(0.0195)	(0.109)	(0.194)	(0.0508)	(0.0855)
Only Winning Bids	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES
Controls	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES
Bidder FE	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES
Control for Reserve Price	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
# of Clusters	194,519	192,909	21,010	21,010	21,010	61,521	61,046	9,788	9,788	9,788
# of Observations	562,616	558,713	186,429	186,429	186,429	137,993	137,079	47,514	47,514	47,514
R-squared	0.001	0.185	0.19	0.808	0.818	0.001	0.171	0.174	0.783	0.799

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the auctioneer-bidder pair level. Controls include fixed effects for regions, products/services, years, months, dummy for law (#44 or #94), dummy for whether a public body is a medical or educational institution, a set of dummies for the number of bidders in an auction. The subsample is only the bidders that participated in *at least* 361 auctions, and auctioneers who had *at least* 685 bids placed in their auctions in total (in the approximately 200 auctions that they ran). That is, we take bidders and auctioneers who are *above* the median value for participation. α_a is a level of corruption measured for an auctioneer, α_b is a level of corruption measured for a bidder, 1_{winner} is a dummy for the winning bid. Column (1) is a raw correlation between auctioneers' corruption and log of bids. Column (2) includes a set of controls including a dummy for the winning bid. Column (3) studies the interaction between auctioneers' and bidders' corruption. Column (4) adds controls for the quintiles of the reserve price. Column (5) adds bidders fixed effects, keeping the interaction between bidders and auctioneers corruption. Columns (6) to (10) repeat all the analysis for the subsample of winning bids; in other words, the final prices of the contract.

Table 10: Corruption, Bids, and Prices—Infrequent Participants

VARIABLES	(1) No Controls	(2) With Controls	(3) Interaction	(4) Interaction & Reserve Price	(5) Bidder FE	(6) No Controls	(7) With Controls	(8) Interaction	(9) Interaction & Reserve Price	(10) Bidder FE
α_a	0.476***	0.284***	0.386***	0.0772**	0.0875***	0.411***	0.273***	0.315***	0.077	0.131***
	(0.0277)	(0.0231)	(0.0645)	(0.0302)	(0.0167)	(0.0451)	(0.0396)	(0.119)	(0.0563)	(0.0259)
α_b			0.0695	-0.000806				0.00427	-0.0235	
			(0.0667)	(0.0302)				(0.101)	(0.049)	
$\alpha_b \cdot \alpha_a$			-0.771	-0.229	-0.202			-0.0854	-0.29	-0.486***
			(0.509)	(0.234)	(0.126)			(0.866)	(0.424)	(0.184)
1_{winner}		-0.178***	-0.179***	-0.165***	-0.153***					
		-0.00345	-0.00562	-0.00252	-0.0021					
Constant	11.59***	12.18***	12.10***	10.10***	10.12***	11.43***	12.04***	11.85***	9.906***	9.948***
	(0.00404)	(0.0216)	(0.0533)	(0.0294)	(0.0136)	(0.00611)	(0.0338)	(0.0832)	(0.044)	(0.026)
Only Winning Bids	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES
Controls	NO	YES	YES	YES	YES	NO	YES	YES	YES	YES
Bidder FE	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES
Control for Reserve Price	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
# of Clusters	540,800	534,309	170,041	170,041	170,041	159,678	157,727	58,910	58,910	58,910
# of Observations	1,085,258	1,072,398	525,978	525,978	525,978	273,457	270,228	130,640	130,640	130,640
R-squared	0.002	0.145	0.152	0.824	0.84	0.002	0.131	0.133	0.802	0.831

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered on the auctioneer-bidder pair level. Controls includes fixed effects for regions, products/services, years, months, dummy for law (#44 or #94), dummy for whether a public body is a medical or educational institution, a set of dummies for the number of bidders in an auction. The subsample is complementary to the subsample in Table 9. That is, we *exclude* such auctions where an auctioneer participated in *more* than 685 bids in total and simultaneously the bidder had *more* than 361 bids. That is, we take bidders and auctioneers who are *below* the median value for participation. α_a is a level of corruption measured for an auctioneer, α_b is a level of corruption measured for a bidder, 1_{winner} is a dummy for the winning bid. Column (1) is a raw correlation between auctioneers' corruption and log of bids. Column (2) includes a set of controls including a dummy for the winning bid. Column (3) studies the interaction between auctioneers' and bidders' corruption. Column (4) adds controls for the quintiles of the reserve price. Column (5) adds bidders fixed effects, keeping the interaction between bidders and auctioneers corruption. Columns (6) to (10) repeat all the analysis for the subsample of winning bids; in other words, the final prices of the contract.

Table 11: Panel A—Distance in Log-bids

K	ε	0.50%	1.00%	2.00%	5.00%	N
3		0.197	0.310	0.439	0.613	1,371,041
		0.078	0.139	0.203	0.249	
4		0.178	0.287	0.413	0.597	785,818
		0.108	0.169	0.222	0.245	
5		0.172	0.278	0.402	0.588	467,893
		0.101	0.154	0.194	0.187	
6		0.165	0.267	0.390	0.581	282,006
		0.092	0.138	0.168	0.148	
Pooled		0.180	0.289	0.415	0.597	3,385,982
		0.090	0.145	0.194	0.212	

Table 11: Panel B—Distance in Bids/Reserve Price

K	ε	0.50%	1.00%	2.00%	5.00%	N
3		0.229	0.349	0.478	0.657	1,371,041
		0.107	0.175	0.239	0.288	
4		0.223	0.341	0.471	0.664	785,818
		0.144	0.209	0.259	0.282	
5		0.227	0.343	0.475	0.676	467,893
		0.141	0.196	0.231	0.214	
6		0.229	0.347	0.477	0.686	282,006
		0.135	0.185	0.203	0.163	
Pooled		0.229	0.347	0.478	0.671	3,385,982
		0.125	0.186	0.230	0.246	

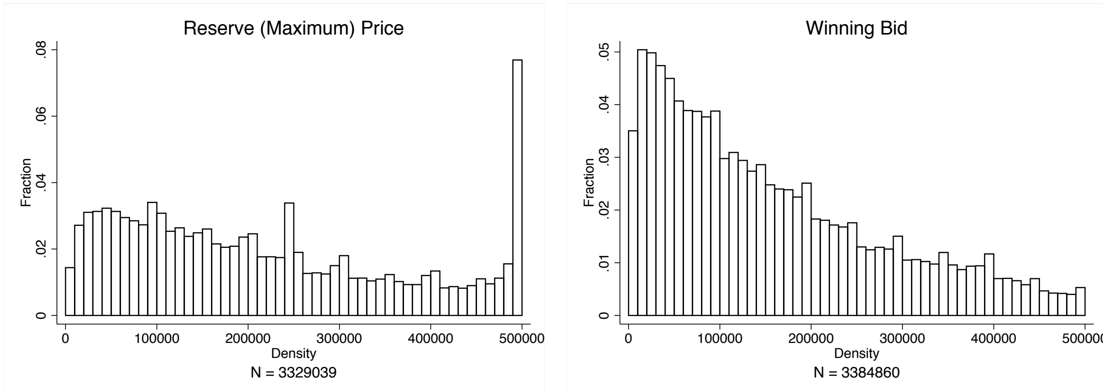
Table 12: Panel A—Baseline Bids Method, Log-Bids

K	3	4	5	6	Pooled
α					
Baseline:	0.31***	0.287***	0.278***	0.267***	0.289***
Upper Bound					
99% Range	(0.003;0.003)	(0.003;0.003)	(0.003;0.003)	(0.003;0.004)	(0.003;0.003)
Baseline:	0.139***	0.169***	0.154***	0.138***	0.145***
Lower Bound					
99% Range	(-0.001;0)	(0;0)	(0;0)	(-0.001;0)	(0;0)
N	1,371,032	785,818	467,893	282,006	3,385,973

Table 12: Panel B—Baseline Bids Method, Log-Bids, Placebo

K	4	5	6	Pooled
α				
Baseline:	0.142***	0.146***	0.15***	0.147***
Upper Bound				
99% Range	(0.003;0.003)	(0.003;0.003)	(0.003;0.004)	(0.003;0.003)
Baseline:	-0.038***	-0.002***	-0.015***	-0.024***
Lower Bound				
99% Range	(0;0)	(-0.001;0)	(-0.001;0)	(0;0)
N	785,818	467,893	282,006	2,014,941

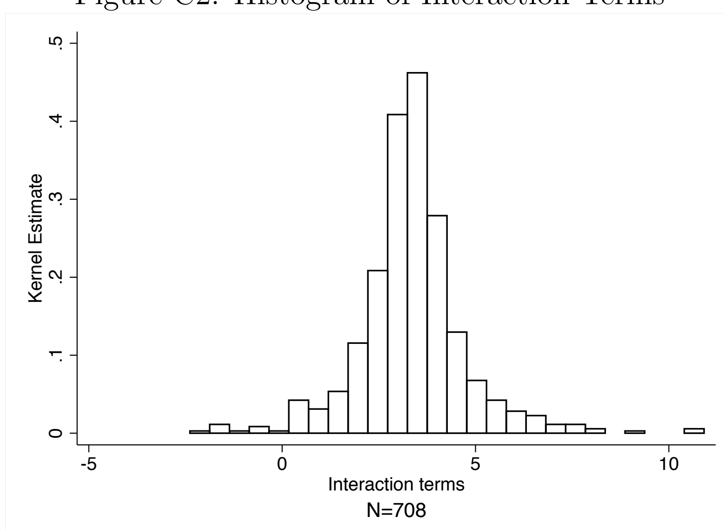
Figure C1: Distribution of Reserve Prices (A) and Winning Bids (B)



Panel A | Panel B

Notes: Histograms for Reserve Price and Winning Bid distributions, fraction of all auctions on the y–axis. All prices and bids are in rubles.

Figure C2: Histogram of Interaction Terms



CHAPTER 2

Collusion in Auctions: Evidence from Timing of Bids

2.1 Introduction

Governments around the world devote considerable resources to detecting and restricting any behavior curtailing competition.[1] One area where this behavior has been observed frequently is auctions: there are many historical cases of collusion, such as bidding rings.[2] Traditional antitrust literature restricts its attention to the developed countries.[3] Until recently micro-level data from less developed countries suitable for measuring collusion was rare and practically non-existent in the auctions setting.

In this paper, we develop a method to detect collusion in first-price sealed-bid auctions. Our method is based on the timing of bids. Bidders in procurement auctions can coordinate to keep the prices higher and thus be involved in the bid rotating schemes. In our context bidders can directly monitor each other by jointly placing their bids.[4] The patterns of such *simultaneous bidding* is something that we use to detect and measure collusion.

---

[1] A notable example of anti-trust and pro-competitive law is the Sherman Act of 1890 in the US. The most relevant regulation in our context is the Federal Law #135 FZ in Russia "On Protection of Competition".

[2] See examples from Marshall and Marx (2012) or Whinston (2008)

[3] For example in Porter (2005) 23 out of 29 empirical papers use the data from the USA, 4 use the data from the UK and 1, from Denmark and Canada each.

[4] For instance, two bidders can submit the envelopes together in case of the paper bid submission or place the bids from the same laptop if the submission of bids is electronic.

Specifically, we collect a massive data set on Russian procurement auctions and observe that there is a large number of bids placed within 30 minutes of each other, for an average length of an auction of two weeks. To control for non-collusive bidding patterns, we compare the probability of bidding at the same time of the day for the same and different days. We test whether the difference in probabilities for the same and different days is zero, which is a baseline test for collusion. It also provides an estimate of the extent of such behavior with up to 28% of auctions with two bidders affected by collusion. The share of collusive bids in all auctions is less and varies between 7% and 23%.

At the next step, we show that simultaneous bidding is correlated with an increase of bids by $4-9\%$ for the auctions with two bidders, with a corresponding increase in prices of $8-9\%$. We show that the findings are robust to different definitions of the simultaneous bidding, and to the inclusion of a battery of control variables. Next, we show that the joint distribution of bids for simultaneous bidders is different than for all other pairs of bids: the joint bids are much closer to each other. We discuss this finding in line with the literature on cartel bidding.

Our paper is the first to use the timing of bids as a tool for detecting collusion. More broadly we contribute to the literature on detecting collusion in auctions, where the authors do not observe, which firms are colluding (Porter and Zona 1999; Baldwin et al. 1997; Bajari and Ye 2003; Ishii 2009; Athey et al. 2011; Conley and Decarolis 2011; Haile et al. 2012; Kawai and Nakabayashi 2014), as opposed to the papers, where the authors have the court cases and such (Porter and Zona 1993; Pesendorfer 2000; Asker 2010b).

We contribute to a scarce literature on the timing decisions in auctions. Previous papers that employ timing (Bajari and Hortacsu 2003; Ockenfels and Roth 2002, 2006; Hopenhayn and Saeedi 2015) describe the strategic behavior of bidders in online auctions such as eBay or Amazon. Those auctions are open-bid as opposed to our sealed-bid auctions.

The rest of the paper is organized as follows. Section 2 describes the institutional background. Section 3 presents the data and summary statistics. Section 4 discusses our detection approach and provides the estimates of the scope of collusion. Section 5 shows the estimates of damages. Section 6 concludes.

2.2 Institutional Background & Data

The three most popular procurement auctions types are *tenders*, *open auctions* and *requests for quotations*.[5] The first type is a reverse English auction, the second type is an auction that includes a quality score weighted with the price, and the third type is a first-price sealed-bid auction. All of the auctions have a reserve price. *Tenders* and *open auctions* are used for larger purchases, and they are more transparent and well-regulated. The *requests for quotations* are used for small purchases. They require less preparation and the also are less transparent. There are other variations of these auction procedures, but they cover an insignificant share of the contracts.

Similar to the previous chapter we focus on the requests for quotations, which are used for small contracts. The reserve price of such a contract can be at most 500,000 rubles ($7,500), and at most 10% of annual expenditures of a public body can be assigned this way, but not more than 100 million rubles ($1.5 million) per year. Typical examples of these contracts are a purchase of office supplies for a municipality, books for a school, or medical supplies for a hospital. Small repairs, street cleaning and all kinds of services can be purchased through this procedure as well.

The public body first posts an announcement of the auction on a public website. The announcement is standardized, and it has exhaustive information about

[5] Sources for this Section: "RusTenders" Website, Federal Law #44, Hramkin and Balsevich et al. (in Russian).

the contract. It also includes the deadline to post the bid and the requirements to qualify as a bidder. The announcement has to be posted no less than seven business days before the deadline.

During the bidding period, anyone can submit an application, which is a price bid paired with some documents showing that this person is eligible for the contract. Applications are accepted in sealed envelopes, by email or online through the website. Sealed envelopes usually have to be submitted on business days from 9 AM to 1 PM or from 2 PM to 5 PM, except for state holidays.

After the auction has ended, the applications are opened and examined by a local committee. Applications can be rejected if the bidders did not meet the posted requirements. The lowest bid wins. If the bids are equal, the earliest application submitted wins. The committee then writes a protocol with the results of the auction, which is also stored on the public website.[6]

Typically all of the potential participants monitor the Internet for the announcements from a given public body and after they submit a bid they receive a letter that notifies when and how the bid was submitted. So, it is relatively easy for them to trace whether the information was entered correctly.

The incentives behind collusion are simple. Once two participants form an agreement, they unambiguously improve their expected profit by not competing against each other. The illegality of such agreements means that they are often difficult to enforce: absent any legal repercussions the urge to cheat on the bidding agreement may be quite strong. However, in the request for quotations, the bidders have a monitoring device: they can place the bids together and observe each other's actions while writing and filing an application. As a result, the bidders will be able to maintain the integrity of the ring, but the bids will be located next to

[6]The bidders or their representatives have a right to participate in the opening procedure, can request to disclose any information from the bidding envelopes and can make a recording of the procedure.

each other in time. We call this *joint bidding*.[7] Balsevich et al, (2012): "It seems that it was a cooperative strategy - keeping the highest price by bid rotation. Even in cases when there was a drop in price... the bidders were applying at the same time."

2.3 Data

We use the same source of data as in the previous chapter. The data contains announcements and protocols. Auction characteristics like reserve price, contract terms, deadlines for submitting applications are stored in the announcements. Protocols have bids of each application, timing of bids and also information about the bid being accepted or rejected. After matching announcements to the protocols, and extracting all the necessary information we obtain a data set of more than 4.3 million requests for quotations. We drop the auctions with any bid being rejected and the auctions with no bids.[8]

Table 2.1 shows the summary statistics. The reserve price in our auctions is a maximum price that needs to be hit by the winning bid for an auction to be considered valid. For the request for quotations, it has to be below or equal to 500,000 rubles. In our sample mean reserve price is 194,693 rubles, which is less than a half of a maximum allowed reserve price. Mean winning bid is 160,855, while the mean ratio of the winning bid and the reserve price is 81.5%.

Other variables that we use are winning bid and time of the winner. We introduce two additional variables as measures of the auction competitiveness: percentage distance from the reserve price to the winning bid and percentage distance between the winning and the runner-up bid.

[7] In some cases for other types of auctions the anti-monopoly committee discovered that the bidding was done from the same IP. [*add links*]

[8] The results for the subsample with rejected bids are similar.

2.4 Detection of Collusion

2.4.1 Patterns in the Data

We start by documenting the following fact: bidders tend to submit their bids simultaneously. Specifically, bids tend to arrive within a 30 minutes interval from each other substantially more often than in during other parts of auction bidding period. Note, that in case of bidding ring, a designated winner wants to monitor the bids of other participants of the ring, to prevent these other participants from deviation. The most direct way of such monitoring is by submitting the bids together: either electronically or using paper envelopes, depending on the rules of the given procurement auction. Such type of monitoring, in turn, leads to simultaneous bidding.

Figure 2.1 shows the timing of bids for winners and runners-up. For visual clarity, we concentrate on the auctions with the deadline on Friday at 9 AM.[9] Note that the estimates of the scope of collusion from the next subsection are presented for auctions with all the deadlines.[10] We depict the hours to the deadline for the bidders with the lowest bids. The runner-up time is on the x-axis and the winner time on the y-axis. The hours are normalized such that $t = 0$ corresponds to the submission of a bid at the deadline.

The most striking pattern of the scatter plot is a very pronounced concentration of pairs of timing at a diagonal. The diagonal corresponds to the simultaneous submission of bids.

To further illustrate and to quantify this pattern, we compare bid pairs that arrived on the same day, to the bid pairs separated by one day. The deadline

[9]It is the deadline that has the maximum number of observations, but the choice of the deadline is not essential for the results.

[10]For this graph, we drop the bids that were submitted less than three hours or more than one hundred hours before the deadline. One can find the scatter plots of the timing for all of the deadlines, and all of the bids, and also for timing normalized by the deadline time in the Supplementary Materials. The patterns in the data are the same.

that we chose was Friday. We also pick such bid pairs that both bids submitted on Thursday and the pairs of bids, when one bid was submitted on Thursday, while another one on Wednesday. Figure 2.2 shows the scatter scatter plots for these pairs. If we pick a corridor around the diagonal that represents simultaneous bidding, we will get a crude estimate of collusion. In practice, we have to choose a bandwidth. On Figure 2.2 we pick the bandwidth of 15 minutes. The share of auctions that have bids within a 30 minutes corridor is 33.2% (from the graph: 3637/10945) for pairs submitted on the same day.

Since our bandwidth is not zero, one can argue that simultaneous bidding can be driven by a convenient time of the day (such as before lunch). In order to address this issue we can normalize the share of simultaneous bidding by the share of auctions that fall within the bandwidth at the same hours of different days. In the example of Figure 2.2, this share is equal to 12.8% (297/2316). One potential estimate of collusion is to take the difference in shares of same hour bidding of different and same days and normalize it by the density of bidding for different day pairs. Specifically, one can use

$$\hat{\gamma} = \frac{Pr\Big(\text{Bandwidth}|\text{Same Day}\Big) - Pr\Big(\text{Bandwidth}|\text{Different Days})}{1 - Pr\Big(\text{Bandwidth}|\text{Different Days}\Big)}$$

$$= \frac{0.332 - 0.128}{1 - 0.128} \approx 0.234.$$

That is, for the subsample under analysis the share of auctions with abnormal simultaneous bidding is 23.4%.

We formalize this intuition, build a more general estimator, and argue that testing for collusion is testing for the abnormal share of simultaneous auctions being zero in the next subsection.

2.4.2 Detection Method

We observe N_a auctions, where each auction is indexed by a. Each auction has K bid, indexed by k. Each bid is characterized by a pair of price for which the firm is willing to deliver the good or service that we call bid magnitude or simply bid throughout the rest of the text, and bid timing (b_k^a, t_k^a). We abstract from bid magnitudes for now, and only study t_k^a. Next, we make an innocuous assumption for our setting and assume that auctions are i.i.d, and we drop the a index.

A pair of timings is drawn from a joint distribution $F(t_k, t_j)$. We do not assume the timing is independent for bidders within auctions. However, we assume that the joint distribution of hours and minutes of bidding should be the same, no matter what the days are. Specifically,

$$t_k = 24 \cdot d_k + h_k, \quad d_k = \{1, ..., 13\}, \; h_k \in [0, 24),$$

$$t_j = 24 \cdot d_j + h_j, \quad d_j = \{1, ..., 13\}, \; h_j \in [0, 24),$$

$$h_k \perp \; h_j | (d_k, d_j), \tag{2.1}$$

where d and h are the days and hours, and the former variable is discrete, while the latter variable is continuous.

One specific corollary of Assumption (1) is that the distribution of minutes is the same for the pair of bids placed on different days:

$$F(h_k, h_j | d_k = d_j) = F(h_k, h_j | d_k \neq d_j) \tag{2.2}$$

Testing for collusion in this framework is testing the equality in (2) against the inequality.

We utilize the specific violation of independence coming from simultaneous bidding. With collusion, the observed joint distribution of hours within the same day is a mixture of two distributions: one is a joint distribution of hours without collusion, and another one is a collusive joint distribution. We can observe the distribution without collusion by using (2) and plugging in the same day distribution of hours.

Denote a probability of collusion happening by γ. Later on, we will also interpret it as a share of auctions with collusive bids. Denote by $G_c(h_k, h_j)$ the distribution of bids with collusion. The formula for mixture will thus take the following form:

$$F(h_k, h_j | d_k = d_j) = (1 - \gamma) \cdot F(h_k, h_j | d_k \neq d_j) + \gamma \cdot G_c(h_k, h_j) \tag{2.3}$$

We are ultimately interested in estimating γ and testing its equality to zero. Since, the latter provides the test for collusion, while the latter provides the estimate of the scope of collusion. We provide the results of a parametric approach first, under which we assume a functional form for $G_c(h_k, h_j)$ similar to the example from the previous subsection. In the next subsection, we discuss a non-parametric procedure of estimation of a two-component mixture model with one known and another unknown symmetric component.

2.4.3 Parametric Approach

For parametric procedure we choose $G_c(|h_k - h_j|) = 1\{|h_k - h_j| \leq \epsilon\}$, similarly to Section 4.1.

The motivation for this assumption comes from Figure 2.3, where we depict the kernel density plots for difference of hours within the day for winners and runners-up for cases when bids arrive at the same and different days. Panel B repeats the exercise for not only the winner – runner-up pairs.

As one can see the spike at zero hourly difference is much more pronounced for same day bidding. A small spike in mass around zero for different day bidding corresponds to the fact that bidders might find it convenient to submit bids at a specific time of the day (e.g., right before lunch).

Based on the parametric assumption, the derivation of estimator from Section 4.1 is straightforward. It follows from (3) that,

$$Pr(|h_k-h_j| \leq \epsilon | d_k = d_j) = (1-\gamma)\cdot Pr(|h_k-h_j| \leq \epsilon | d_k \neq d_j) + \gamma \cdot 1(|h_k-h_j| \leq \epsilon),$$

or
$$\gamma = \frac{Pr(|h_k-h_j| \leq \epsilon | d_k = d_j) - Pr(|h_k-h_j| \leq \epsilon | d_k \neq d_j)}{1 - Pr(|h_k-h_j| \leq \epsilon | d_k \neq d_j)},$$

which for $\epsilon = 1/4$ and two chosen days from Section 4.1 is exactly equal to 0.234.

As a matter of fact, for a fixed choice of ϵ the estimator can be implemented by running an OLS of $1\{|h_k - h_j| \leq \epsilon\}$ on $1\{d_k = d_j\}$. Specifically, we show in the Appendix that

$$\hat{\gamma} = \frac{\hat{\beta}_{1OLS}}{1 - \hat{\beta}_{0OLS}}$$

To explore this approach, we report the estimates derived from OLS in Table 2.2. We use three bands: 10, 30, and 60 minutes corridor. Note that in (3) we restrict days to be not last day: $d_k, d_j > 0$. We need to pick up the cutoff for this choice as well. We use two cutoff levels: 3 and 12 hours. In the next subsection we discuss a non-parametric estimator that does not rely on the choice of ϵ.[11]

Table 2.2 uses only auctions with two bidders. Table 2.3 generalizes the results to all auctions, and to the three pairs of bidders that place the lowest bid (winner, runner-up, and third-best bidder). Depending on the specification and on the choice of cutoffs, the share of collusive shares varies from 7% to 28%. Note that the choice of the last-day cutoff matters much less than the choice of ϵ.

2.4.4 Nonparametric Approach

One issue with the estimator in the previous subsection is that they vary substantially with the choice of cutoff. Ideally, we want to abstract from parametrizing G_c

[11] One can incorporate covariates in this setting. We add Industry, and Public Body fixed effects in the regression in Table 1 of the Supplementary Materials. The results do not change.

and from choosing the cutoff and derive a consistent estimator of γ. Note that, the decomposition of observed joint distribution (3) is a mixture of two distributions: one is unknown, and another one is observed from different days bidding.

For the rest of this section, we only deal with one-dimensional distribution of the distance in hours, instead of the joint density of hours. That is, we write down a corollary of (3) for the distribution of differences, $w_{kj} = h_k - h_j$. We also use densities instead of CDFs

$$f(w_{kj}|d_k = d_j) = (1-\gamma) \cdot f(w_{kj}|d_k \neq d_j) + \gamma g_c(f(w_{kj})$$

We can build on the statistical literature on mixture models[12] and assume symmetry of a collusive component. In our setting, it will mean,

$$g_c(-x) = g_c(x)$$

Imposing some technical identification conditions from **?**, one can use either the method of moments or a minimum contrast estimator to estimate the mixture parameter, i.e., the share of collusive auctions. We illustrate the method of moment procedure in Appendix, and we are currently working on its implementation. For the next Section, we concentrate on the choice of $\epsilon = 1/4$ and the cutoff for the last day of three hours.

2.5 Collusion and Prices

In this section, we want to understand what are the damages from simultaneous bidding, by documenting correlation of simultaneous bidding with bids, contract prices, and bid margins (the difference between bids). We start by looking at the

[12]For example, see Bordes et al. (2006); **?**); Bordes and Vandekerkhove (2010); Hohmann and Holzmann (2013).

scatter plots of bids on Figure 2.4, and Figure 2.5.[13]

First, we show the absolute bids in Figure 2.4. It is apparent from the scatter plot on Figure 2.4 that bids shrink toward each other when they are placed simultaneously. We discuss why this could be happening in the presence of collusion, and we measure the amount of bid margin shrinkage in the next subsection.

At the same time, not only bid margins shrink, but also bids and prices go up in the presence of collusion. To document it, we first normalize bids by the reserve price of the auction (maximum cost estimate of the contract) as in Kawai and Nakabayashi (2014). We show the scatter plot of those normalized bids on Figure 2.5. The bids tend to be closer to the reserve price, and, thus, higher, in auctions with collusion. It is evident from a more pronounced mass for the same day bidding near $(1, 1)$ on Figure 2.5.

2.5.1 Regression Estimates

To quantify the patterns on the scatter plots of Figures 2.4 and 2.5 we run several regressions for prices. We restrict our analysis to normalized bids, although the results are qualitatively similar if we study absolute bids or logarithm of bids (see Supplementary Materials).

Specifically, we run linear regressions of the following form:

$$y_{akj} = \delta_0 + \delta_1 \cdot 1\{d_k = d_j\} + \delta_2 \cdot 1\{|h_k - h_j| \leq \epsilon\} + \delta_3 \cdot 1\{|h_k - h_j| \leq \epsilon\} \cdot 1\{d_k = d_j\} + \xi_{akj},$$

where we use a minimum, a maximum, or an average bid of kj pair as an outcome y_{akj}. In another specification, we use bid margin as an outcome.[14]

We are interested in estimating the following object: $D = (\delta_1 + \delta_2 + \delta_3)/\delta_0$,

[13] We show the scatter plots for Friday 9 AM deadline, the patterns the same in the whole data, and we report them in Supplementary Materials.

[14] That is, $y_{akj} = bid_{aj} - bid_{ak}$.

which corresponds to the increase in prices (decrease of bid margins) with simultaneous bidding. We show the results for prices for the 30 minutes interval, for two and K-bidder auctions.

We start with two-bidder auctions in Table 2.4. The sum of the coefficients relative to the baseline normalized prices, $\hat{D}$ varies from 4.1% to 6.8%. The estimates for auctions with K bidders in Table 2.5 are even higher and vary from4.5% to 8.8%. All of these estimates are statistically different from zero at 1% significance level.

Since auctions are heterogeneous, we might want to take into account observed heterogeneity. We do it by including Public Body fixed effects, and Industry and Region fixed effects in Table 2.6. The results do not differ much. Collusion is correlated with an increase in prices of $8.4 - 9.8\%$.

Similar regression specification allows us to document that simultaneous bidding is associated with smaller bid margins. Bid margins shrink by up to 50% of the initial value, depending on specifications and inclusion of controls. Bid margins are at least 40% lower in auctions with collusive bidding patterns.

This finding is similar to patterns documented in Marshall and Marx (2007) for the case of less than all-inclusive cartels with an enforcing mechanism. The similarity is somewhat surprising and requires further exploration since in our case, collusive bidders monitor each other to avoid the necessity for using any collusive mechanisms.

2.6 Conclusion

This paper builds a new method of detecting and measuring the scope of collusion. Our novel method relies on the information on the timing of bids in first-price sealed-bid auctions. We apply this method to a data set of first-price sealed-bid procurement auctions in Russia and show that collusion contaminates a substantial part of procurement. Next, we provide some crude estimates of losses coming from the price inflation due to collusion, and we document that bid margins shrink with simultaneous bidding.

We argue that recording time-stamps even in the environments where timing is not a strategic variable, can be crucial for antitrust authorities.

As a next step, we will implement a non-parametric estimator of collusion, that will not rely on the ad hoc choices of cutoffs. Studying the variation of collusion across industries and markets is also left for future research.

Appendix

We use the auction index a again, and for simplicity we assume that there are two bidders in the auction, even though the result is generalizable for more bidders. Define variables $w^a = |h_k^a - h_j^a|$ and $z^a = |d_k^a - d_j^a|$, in words, differences in hours and in days.

The least squares estimate of interest is derived from regression of $1\{w^a \leq \epsilon\}$ on $1\{d^a = 0\}$. The slope estimate is equal to

$$\hat{\beta}_{1OLS} = \frac{\frac{1}{N_a}\sum_a 1\{w^a \leq \epsilon\} \cdot (1\{d^a = 0\} - \overline{1\{d^a = 0\}})}{\frac{1}{N_a}\sum_a 1\{d^a = 0\} \cdot (1\{d^a = 0\} - \overline{1\{d^a = 0\}})}$$

The quantities numerator and denominator are sample equivalents of

$$Pr\Big(w^a \leq \epsilon \cap d^a = 0\Big) - Pr\Big(w^a \leq \epsilon\Big) \cdot Pr\Big(d^a = 0\Big)$$

and

$$Pr\big(d^a = 0\big) \cdot \Big(1 - Pr\big(d^a = 0\big)\Big)$$

respectively.

By Slutsky's theorem, and after some straightforward rearrangements the ratio converges to

$$\frac{Pr\Big(w^a \leq \epsilon | d^a = 0\Big) - Pr\Big(w^a \leq \epsilon\Big)}{1 - Pr\big(d^a = 0\big)} = Pr\Big(w^a \leq \epsilon | d^a = 0\Big) - Pr\Big(w^a \leq \epsilon | d^a \neq 0\Big),$$

where the latter equality follows from Bayes rule.

Now, at the same time the estimate of the constant is defined as

$$\hat{\beta}_{0OLS} = \frac{1}{N_a}\sum_a 1\{w^a \leq \epsilon\} - \hat{\beta}_{1OLS} \cdot \frac{1}{N_a}\sum_a 1\{d^a = 0\},$$

which converges to its sample analog of

$$Pr\Big(w^a \leq \epsilon\Big) - \Big(Pr\Big(w^a \leq \epsilon|d^a = 0\Big) - Pr\Big(w^a \leq \epsilon|d^a \neq 0\Big)\Big) \cdot Pr\Big(d^a = 0\Big).$$

Using Bayes rule again, one gets

$$\hat{\beta}_{0OLS} \xrightarrow{p} Pr\Big(w^a \leq \epsilon|d^a \neq 0\Big).$$

Applying Slutsky's theorem again we get that

$$\gamma = \frac{\beta_{1OLS}}{1 - \beta_{0OLS}}.$$

The variance is easily derived via delta-method.

Methods of Moments.

Denote $f_1 = f(x|\text{same day})$, $f_0 = f(x|\text{different days})$, $g = g_c$. Define also, $\tilde{f}(\cdot) = f_1(\cdot + \mu_0)$, where μ_0 is the first central moment of f_0.

In addition, define $\tilde{f}_0(\cdot) = f_0(\cdot + \mu_0)$ and $\tilde{\mu} = \mu - \mu_0$, where μ is the first central moment of g.

In this case the mixture formula becomes

$$\tilde{f}_1(x) = (1 - \gamma) \cdot \tilde{f}_0(x) + \gamma \cdot g(x - \tilde{\mu}).$$

Denote by $\tilde{m}_r$, $r = 1, 2, 3$ the first three central moments of $\tilde{f}_1$, by θ the second central moments of g, by $\tilde{\theta}_0$ and $\tilde{\eta}_0$ the second and the third central moments of

$\tilde{f}_0$. Thus, we get

$$\tilde{m}_1 = \gamma \cdot \tilde{\mu},$$

$$\tilde{m}_2 = \gamma \cdot \tilde{(}\tilde{\mu}^2 + \theta) + (1 - \gamma) \cdot \tilde{\theta}_0,$$

$$\tilde{m}_3 = \gamma \cdot \tilde{(}\tilde{\mu}^3 + 3\theta\tilde{\mu}) + (1 - \gamma) \cdot \tilde{\eta}_0,$$

We are interested in first solving for $\tilde{\mu}$ and then for γ. If all of the moments are non-zero, it boils down to solving a cubic equation,

$$\tilde{\mu}^3 + \frac{3(\tilde{\theta}_0 - \tilde{m}_2)}{2\tilde{m}_1}\tilde{\mu}^2 + \frac{\tilde{m}_3 - \tilde{\eta}_0 - 3\tilde{m}_1\tilde{\theta}_0}{2\tilde{m}_1}\tilde{\mu} + \frac{\tilde{\eta}_0}{2} = 0. \quad (A)$$

All the coefficients of this equation can be estimated by plugging sample analogs for the moments. It will at most have three solutions, and we can choose one that minimizes the discrepancy measure from **?**.

FIGURES

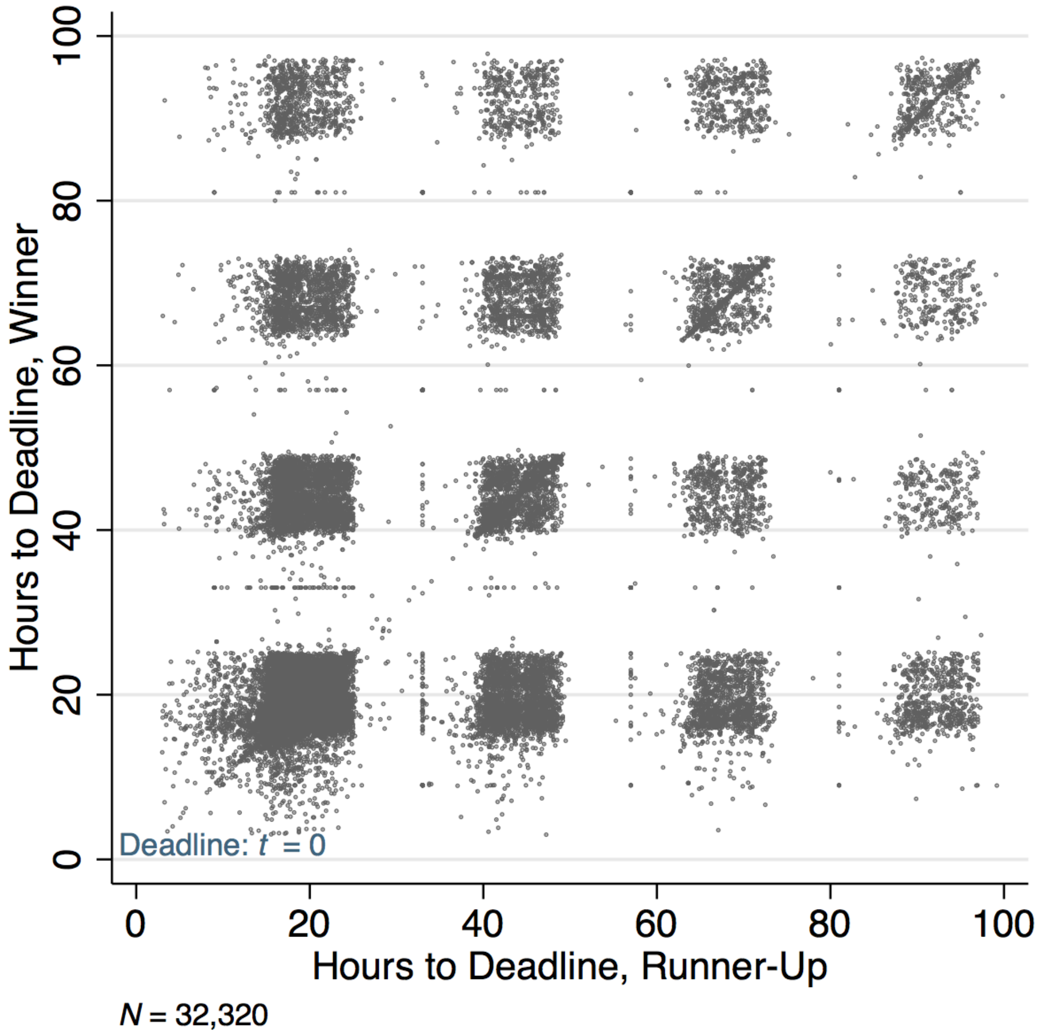


Figure 2.1: Time of Bidding for Winners and Runners-Up: Friday, 9AM Deadline
Note: The figure illustrates the distribution of timing of bids during the bidding period of auctions. Each point corresponds to winner – runner-up pair of timings, expressed in hours to the deadline. The origin is the case of both bids place at the deadline.

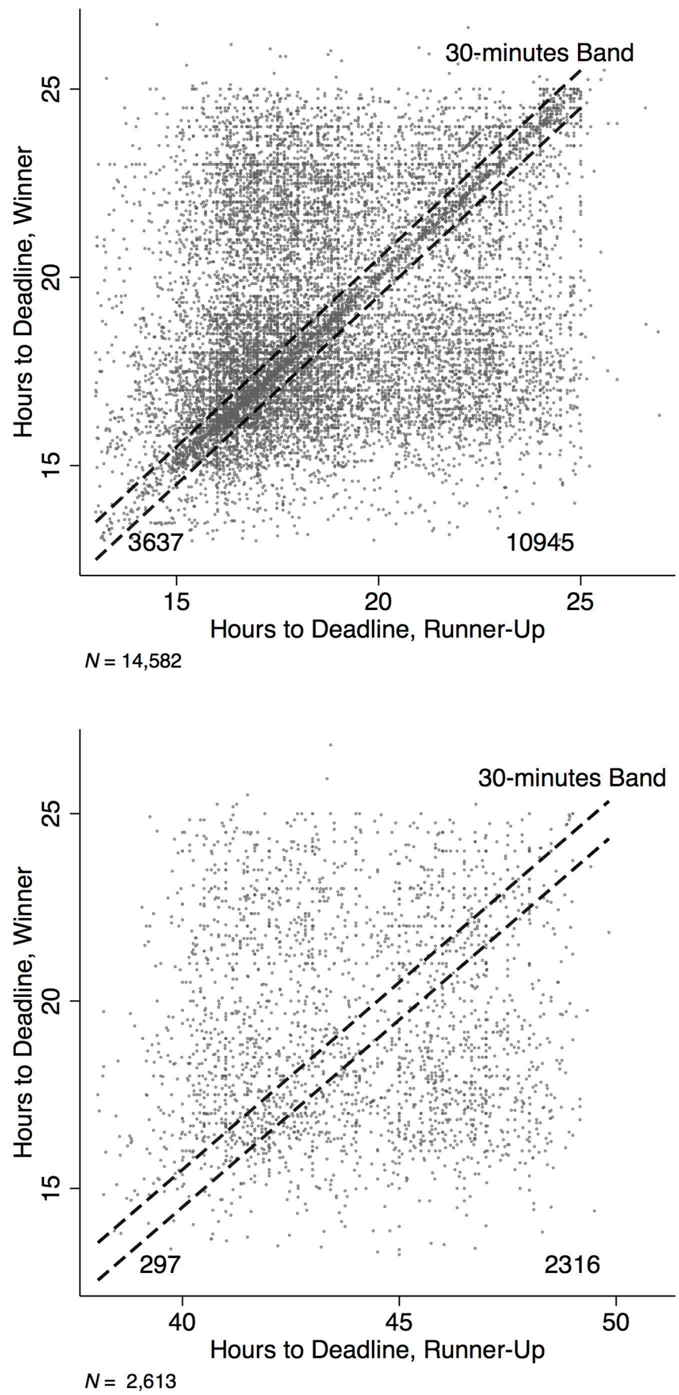
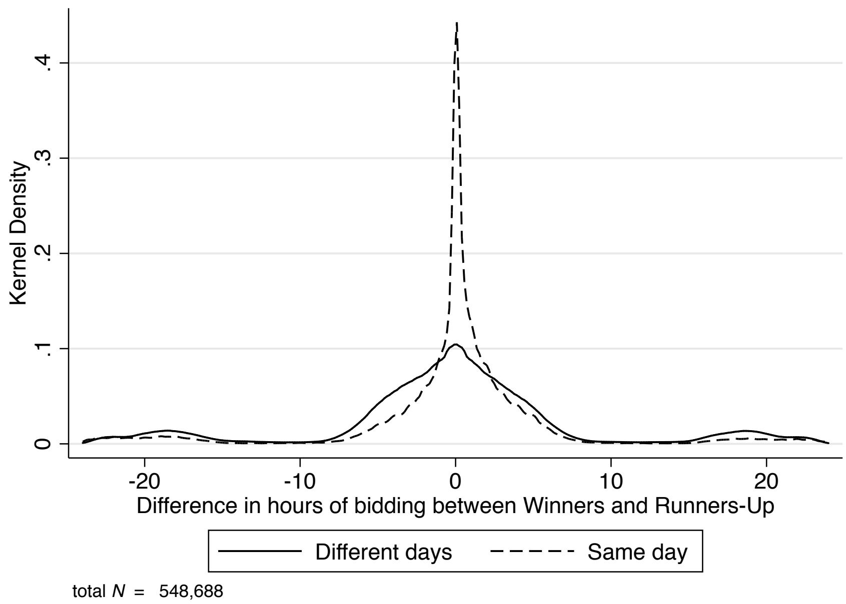
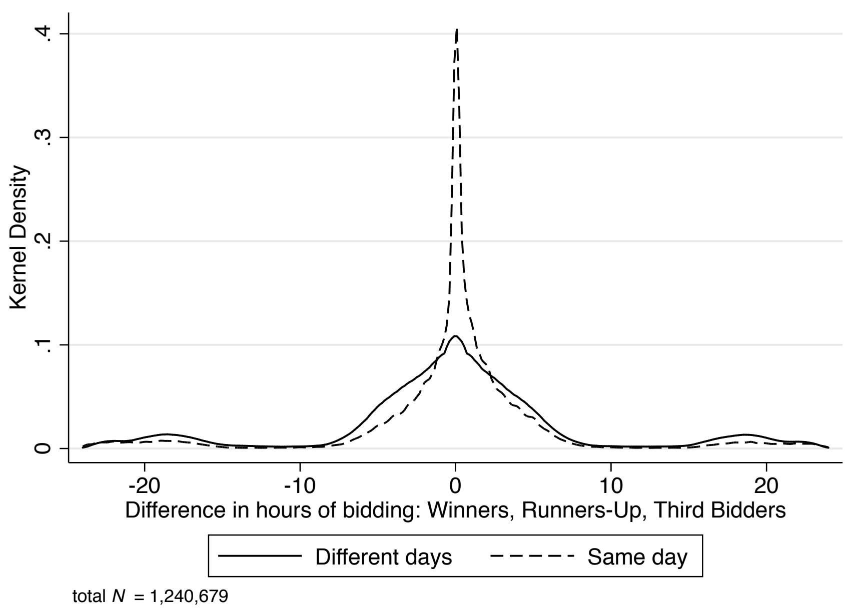


Figure 2.2: Time of Bidding for Winners and Runners-Up: Same to Different Days Comparison

Note: The figure illustrates the distribution of timing of bids during the bidding period of auctions. Each point corresponds to winner – runner-up pair of timings, expressed in hours to the deadline. The first panel shows the distribution of timing for pairs, when both bids were submitted on Thursday; the second panel shows the distribution for Thursday-Wednesday pairs.



Panel A. Only Winners and Runners-Up



Panel B. Three Bidder Pairs

Figure 2.3: Difference in Time of Bidding for Winners and Runners-Up

Note: The figure illustrates the difference in hours of bidding for the bids submitted at the same and at different days. Panel A only shows winners and runners-up, while Panel B also shows third bidders.

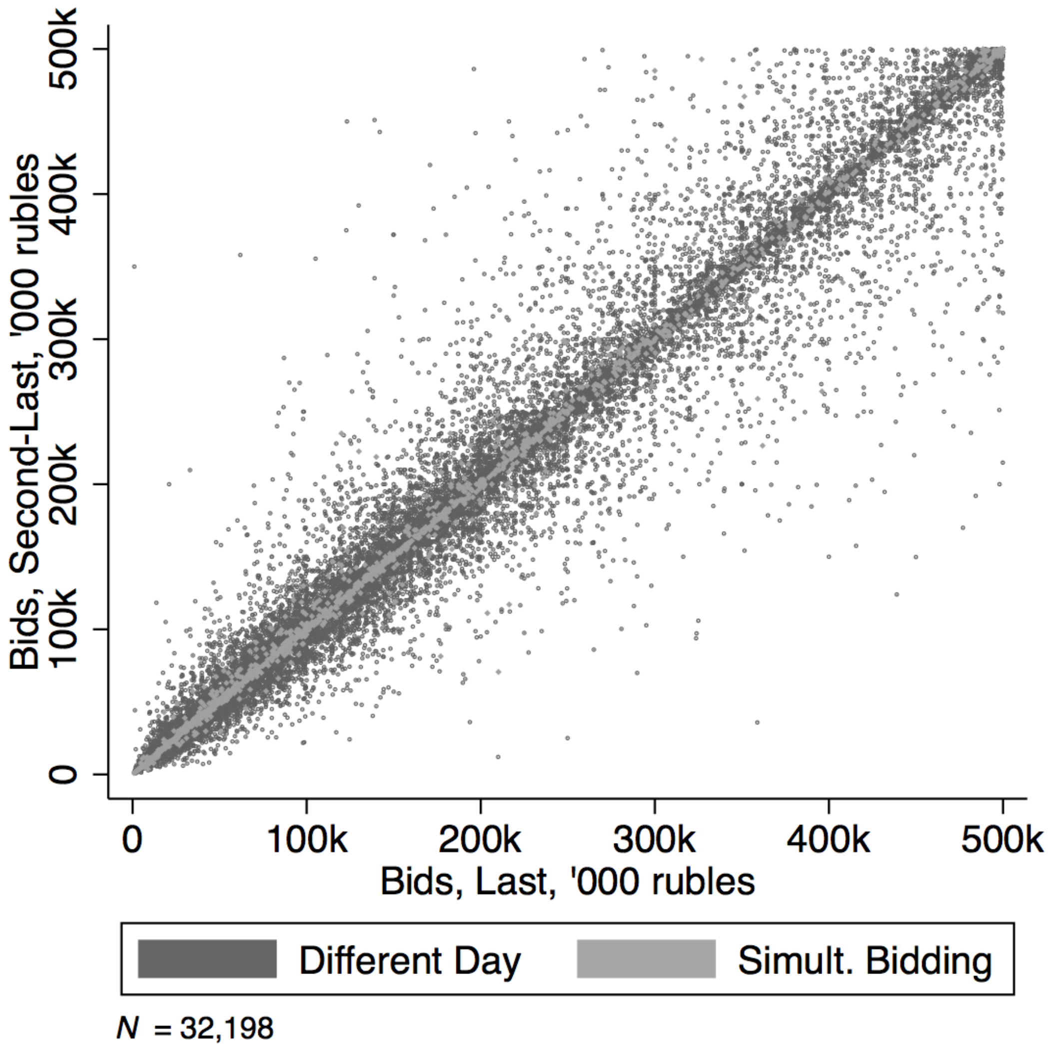


Figure 2.4: Bids of Last and Second-Last Bidders, Friday 9AM Deadline
Note: The figure shows the scatter plot of absolute bids for last and second-last bidder.

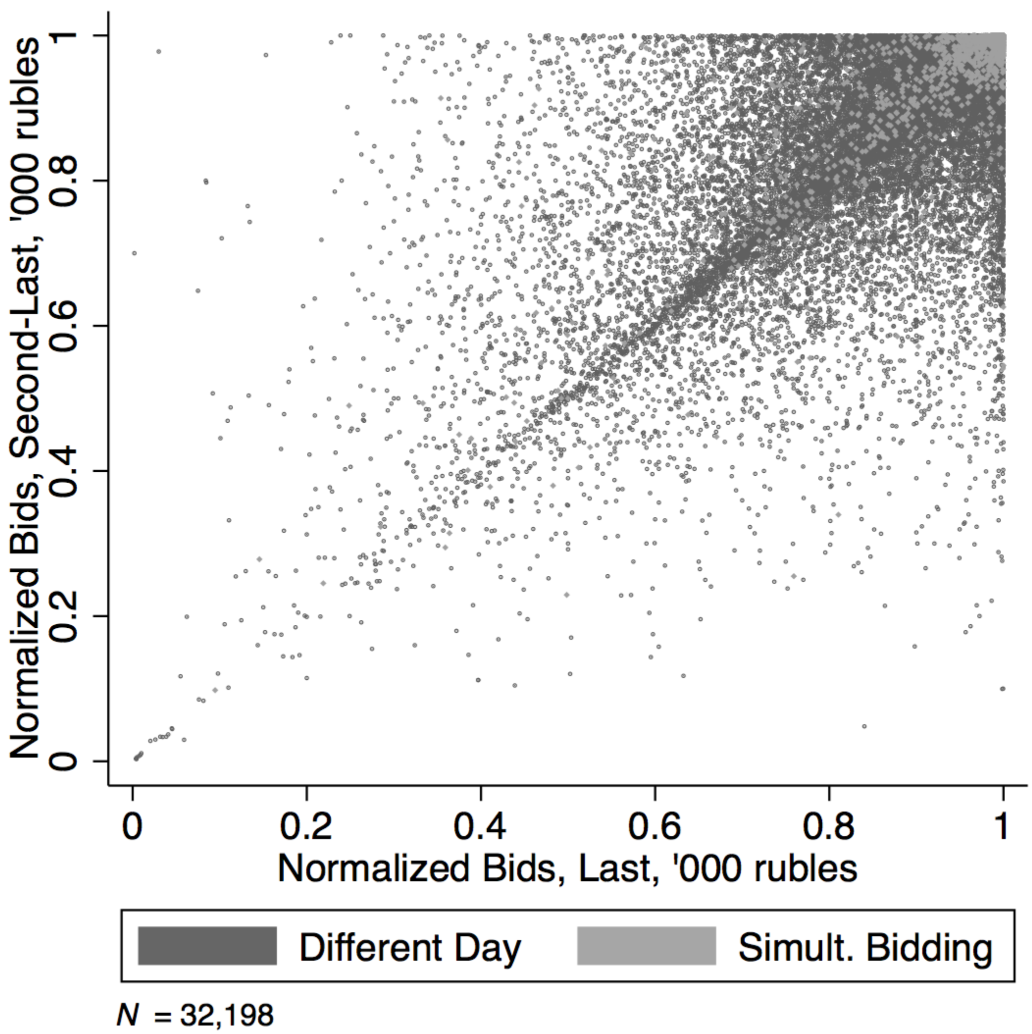


Figure 2.5: Normalized Bids of Last and Second-Last Bidders, Friday 9AM Deadline

Note: The figure shows the scatter plot of normalized bids for last and second-last bidder. Normalized bids are defined by dividing the actual bid by the reserve price of the auction (cost estimate).

TABLES

Table 2.1: Summary Statistics

	Mean	Median	St. Dev.
Number of Bidders	2.79	2	1.34
Winning Bid	160,855	122,208	134,892
Reserve Price	194,693	155,000	151,674
Time of the Winner to Deadline	25.8	5.67	37.9
Winning Bid to Reserve Price %	81.5	85.9	17.7
% Diff.: Winning and Runner-Up Bid	5.23	1.87	8.17
Observations	1,702,585		

Table 2.2: Baseline Measure: Two Bidders, One Pair

Dependent variable: Simultaneous Bidding

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*5-minutes band*		*15-minutes band*		*30-minutes band*	
Same Day Bidding	0.093***	0.090***	0.200***	0.193***	0.247***	0.234***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.038***	0.040***	0.068***	0.072***	0.115***	0.121***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Share $\hat{\gamma}$	0.10	0.09	0.21	0.21	0.28	0.27
R-squared	0.03	0.02	0.07	0.06	0.08	0.07
Observations	450,698	329,993	450,698	329,993	450,698	329,993

Notes: The table shows that same day bidding predicts simultaneous bidding. It provides an estimate of a share of collusive auctions for auctions with two bidders. Regressions are run at auction level; robust standard errors are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01. Odd columns use a cutoff level of 3 hours before the deadline. Even columns use a cutoff level of 12 hours before the deadline.

Table 2.3: Baseline Measure: K Bidders, Three Pairs

Dependent variable: Simultaneous Bidding

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*5-minutes band*		*15-minutes band*		*30-minutes band*	
Same Day Bidding	0.074***	0.080***	0.157***	0.167***	0.208***	0.219***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.017***	0.017***	0.037***	0.037***	0.065***	0.065***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share $\hat{\gamma}$	0.07	0.08	0.16	0.17	0.22	0.23
R-squared	0.03	0.03	0.07	0.07	0.08	0.09
Observations	1,589,287	1,120,282	1,589,287	1,120,282	1,589,287	1,120,282

Notes: The table shows that same day bidding predicts simultaneous bidding. It provides an estimate of a share of collusive pairs for auctions with $K \geq 2$ bidders. Regressions are run at a pair level; robust standard errors are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01. Odd columns use a cutoff level of 3 hours before the deadline. Even columns use a cutoff level of 12 hours before the deadline.

Table 2.4: Prices: Two Bidders

Dependent variable: Normalized Prices

	(1)	(2)	(3)	(4)
Dependent variable:	Min Bid	Max Bid	Average	Absolute diff.
		15-minutes Band		
Same Day Bidding	0.019***	0.012***	0.015***	-0.007***
	(0.000)	(0.000)	(0.000)	(0.000)
15-minutes band	-0.005***	-0.003***	-0.004***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Same Day Bidding × 15-minutes band	0.045***	0.030***	0.038***	-0.015***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.876***	0.928***	0.902***	0.052***
	(0.000)	(0.000)	(0.000)	(0.000)
% Increase in Prices	6.76%	4.14%	5.41%	-39.85%
R^2	0.021	0.016	0.021	0.008
Observations	448,932	448,932	448,932	448,932

Notes: The table reports the correlation of bids (prices) and simultaneous bidding for auctions with two bidders. Regressions are run at auction level; robust standard errors are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01. A cutoff level of 3 hours before the deadline is used. Column (1) shows the results for the minimum bid of a pair (price of the contract), column (2) and (3) show the estimates for maximum and average bid of the pair. Finally column (4) shows the estimate for bid margins (bid difference) of a pair.

Table 2.5: Prices: K Bidders, Winners

Dependent variable: Normalized Prices				
	(1)	(2)	(3)	(4)
Dependent variable:	Min. Bid	Max. Bid	Average	Absolute diff.
		15-minutes Band		
Same Day Bidding	0.016***	0.006***	0.011***	-0.010***
	(0.000)	(0.000)	(0.000)	(0.000)
15-minutes band	0.000	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Same Day Bidding × 15-minutes band	0.057***	0.035***	0.046***	-0.022***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.837***	0.910***	0.873***	0.073***
	(0.000)	(0.000)	(0.000)	(0.000)
% Increase in Prices	8.81%	4.53%	6.58%	-44.49%
R^2	0.020	0.011	0.018	0.012
Observations	677,754	677,754	677,754	677,754

Notes: The table reports the correlation of bids (prices) and simultaneous bidding for auctions with two bidders and only the winner – runner-up pair. Regressions are run at auction level; robust standard errors are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01. A cutoff level of 3 hours before the deadline is used. Column (1) shows the results for the minimum bid of a pair (price of the contract), column (2) and (3) show the estimates for maximum and average bid of the pair. Finally column (4) shows the estimate for bid margins (bid difference) of a pair.

Table 2.6: Prices: K Bidders, Winners

Dependent variable: Normalized Prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:		Min. Bid				Diff. in Bids		
				15-minutes Band				
Same Day Bidding	0.016***	0.017***	0.018***	0.017***	-0.010***	-0.010***	-0.011***	-0.010***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
15-minutes band	0.000	-0.001	-0.001	-0.000	-0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Same Day Bidding × 15-minutes band	0.057***	0.056***	0.064***	0.059***	-0.022***	-0.022***	-0.027***	-0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.837***	0.854***	0.835***	0.821***	0.073***	0.084***	0.074***	0.106***
	(0.000)	(0.026)	(0.000)	(0.024)	(0.000)	(0.020)	(0.000)	(0.015)
Industry and Region FEs		✓		✓		✓		✓
Public Body FEs			✓	✓			✓	✓
% Increase in Prices	8.81%	8.41%	9.78%	9.15%	-44.49%	-38.38%	-51.28%	-33.51%
R^2	0.020	0.123	0.211	0.266	0.012	0.066	0.151	0.178
Observations	677,754	677,754	677,754	677,754	677,754	677,754	677,754	677,754

Notes: The table reports the correlation of bids (prices) and simultaneous bidding for auctions with two bidders and only the winner – runner-up pair. Regressions are run at auction level; robust standard errors are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01. A cutoff level of 3 hours before the deadline is used.

‡ Controls include (2) industry and region fixed effects, (3) public body fixed effects, (4) both industry, region, and public body fixed effects.

CHAPTER 3

Conflict and Trade: Evidence from Russian-Ukrainian Trade Transactions

3.1 Introduction

The consequences of armed conflict is a central topic in political economy and development economics. A large empirical literature has provided strong evidence that conflict, besides its tragic humanitarian effects, can adversely impact economic outcomes such as GDP and stock market indices.[1] However, evidence on exactly what causes these negative economic effects is limited. In particular, little attention has been paid to the way firms respond to conflict. This is partly due to data limitations, and partly due to the focus of existing studies on severe armed conflict, where economic activity nearly ceases.[2]

In this study, we attempt to fill this gap in the literature by studying firm response in an environment of moderate conflict—in other words, in a localized struggle between two nations with no trade embargoes and no violence in the regions where these businesses operate. This type of conflict has become increasingly common in the post-WWII era.[3]

[1]E.g., see Abadie and Gardeazabal (2003), Glick and Taylor (2010), and Miguel and Roland (2011).

[2]One exception is an insightful paper by Amodio and Di Maio (2017) who show, in the context of Palestine during the Second Intifada, that conflict induces firms to switch from domestically produced inputs to the imported ones.

[3]According to UCDP data, the number of internationalized internal armed conflicts, to which the war in Eastern Ukraine has been assigned, has increased from 22 in 1945-1965 to 134 in 1995-2015 while the number of traditional interstate conflicts have gradually decreased over

Conflict can impact firm performance through multiple direct channels, such as destruction of property and diversion of the workforce from productive activities (Ksoll et al., 2014). However, much less is known about the potential indirect effects on firms not located in combat areas. For instance, what happens to firms business relationships when their business partners are associated with a now-hostile country? Conflict may result in external or internal pressure on firms to discontinue such relationships. Moreover, it can damage the mutual trust between the business partners. At the same time, if there are no trade barriers and the infrastructure is not affected, it is not clear whether firms would change their trading behavior by a significant margin. Firms are typically seen as rational and, historically, even during severe armed conflicts, 'trading with the enemy' laws had to be put in place to prevent trade from happening.[4]

This paper studies the decision of firms to "trade with the enemy" in the context of the Russian-Ukrainian conflict (2014–). Specifically, the presence of a large Russian minority in Ukraine allows us to identify the impact of conflict on trade transactions between firms associated with the same or the opposite conflict countries.[5] In a difference-in-differences set-up, we leverage a unique dataset on the universe of export and import transactions of Ukrainian firms in 2013-2016 and compare firm response to conflict in more versus less ethnolinguistically Russian areas of Ukraine. The concentration of combat in a relatively small geographic area of Ukraine gives us the opportunity to analyze the indirect effects of conflict

time. Interstate conflicts have also become increasingly 'moderate' in that they are typically localized and rarely disrupt trade relations (see, e.g., Indian-Pakistan conflict of 2016).

[4]However, sometimes even laws are not enough. Famously, despite an explicit embargo, traders in the US North continued to purchase cotton from the US South during the Civil War (http://nyti.ms/2E0mj4A) while the GM and Ford have been accused of doing business with Hitler's Germany even after Pearl Harbor (http://wapo.st/1HsdxUA).

[5]Conflict triggered a massive but very asymmetric shock to the attitudes of Ukrainians to Russia. Survey data shows that the fraction of ethnic Ukrainians favorable to Russia decreased from 86% to 53% two months after the start of the conflict, while it did not change by much (from 99% to 88%) for ethnic Russians living in Ukraine. The fraction of ethnic Ukrainians in favor of closing the border with Russia went up from 2% to 12% eight months into the conflict with only a trivial change for ethnic Russians (from 0% to 2%). For more details, see Section 3.2.3.

on trade, isolated from the direct effects of violence. The two countries were not directly at war, and, as a result, there has been no disruption of trade at the border.[6] In fact, Russia and Ukraine remained major trade partners years after the start of the conflict.

We show that the negative effect of conflict on trade was more substantial for the firms located in more ethnolinguistically Ukrainian areas of Ukraine relative to the firms from more Russian areas.[7] This effect is present for both the number of trade transactions, as well as total weight and total value traded in a given month. Moreover, the result is observed separately for export and import transactions. Our estimates suggest that moving a firm from a county ('raion') with an average share of ethnic Russians (15%) to a county with the maximum percentage of ethnic Russians among the counties in our sample (50%) would have mitigated the adverse effect of conflict on firm's trade by 45%.

We propose several mechanisms that could drive firms' decisions not to trade with the firm from a hostile country. The first mechanism we examine is animosity which includes both external pressure from activists and consumers, as well as internal bias against trading with firms from hostile regimes. Second, we consider erosion of trust between ethnic groups, which becomes even more important as the importance of formal institutions decreases.[8] We explore these mechanisms in a

[6] Russia and Ukraine have been part of the Commonwealth of Independent States Free Trade Area (CISFTA) since the fall of the Soviet Union in 1991. All import tariffs were set to zero, except for white sugar. While some export tariffs were still in place for a small number of goods, they were unchanged until January 2016 and, as such, would not affect our results.

[7] Overall, the conflict had a massive adverse effect on trade between Russia and Ukraine. The percentage of Ukrainian exports that go to Russia plummeted with the start of the conflict from 25.7% in 2012 to 9.9% in 2016. Likewise, the share of Russian goods among all Ukrainian imports fell from 32.4% in 2012 to 13.1% in 2016. Nevertheless, even after such a severe decline, Russia remained Ukraine's largest trading partner. Similarly, the role of Ukraine in Russian international trade declined but remained significant. For instance, in 2011, Russia imported 5.8% of all goods from Ukraine, making Ukraine the fifth largest importer to Russia. This share dropped to 2.3% in 2015, i.e., after the start of the conflict (eleventh largest).

[8] The fact that bias may impact trade has been extensively studied by the literature on buyer animosity (Edwards et al., 2007; Heilmann, 2016). Furthermore, several studies have documented that conflict affects inter-ethnic trust and that, separately, trust affects trade. Most

stylized model of trade with asymmetric information. In our model, a continuum of sellers possess a good of privately known quality. Furthermore, a buyer may be biased against the seller's ethnic group or identity. With some probability, a buyer meets a seller she can trust, i.e., whose quality she observes perfectly. If business is mutually beneficial, they trade and split the surplus. If they choose not to trade, or if they did not get a chance to meet, they go to the general market. The heterogeneity of good quality is so high that the general market becomes a market of lemons. In this model, we show that both a shift in the distribution of bias (i.e., increased animosity) and a decrease in the probability of meeting a trustworthy partner (i.e., lower trust) would lead to a lower overall probability of trade. Moreover, we derive a testable prediction that an increase in bias caused by conflict should have a bigger impact on homogeneous goods, while a decline in trust should instead have a strong effect on more heterogeneous products.

Informed by the model's predictions, we then attempt to empirically distinguish between an increase in animosity and a decrease in trust. To this end, following Nunn (2007), we split all goods into homogeneous (i.e., traded on exchanges or reference priced) and relation-specific (all others). Nunn (2007) shows that sustaining trade of relation-specific goods requires better institutions, stricter contract enforcement, as well as a higher degree of mutual trust. Therefore, if the main mechanism behind the fall of trade is based on a decrease in trust, one would expect relation-specific goods to be affected more seriously. On the other hand, if one believes that an increase in animosity is the primary mechanism, one would expect our results to be driven by homogeneous goods. We find evidence in support of the latter explanation, as all of our effect comes from trade in homogeneous goods.

If our results are indeed due to animosity, we would expect consumer goods

notably, these two effects were put together in a formal model by Rohner et al. (2013b) in which conflict leads to a decline in trade through erosion of inter-ethnic trust.

to be affected by a larger margin, as consumer goods trade is better observed by customers and activists who could potentially influence firms decisions. We investigate this empirically by estimating our difference-in-differences specification separately for firms trading consumer and intermediate goods. Although our main results hold for both consumer and intermediate goods traders, the estimates are substantially larger for the firms that trade consumer products only. Overall, this pattern supports the animosity explanation, while at the same time suggesting that consumer animosity cannot be the only driving force behind our results.[9]

We are open to alternative explanations of our results beyond animosity and trust. However, there are some obvious ones that we can rule out with the data. For instance, one possible explanation of our results could be that ethnic Ukrainian areas took a bigger overall economic hit as a result of the conflict.[10] Contrary to this argument, we find that sales, profits, and TFP of an average firm (not only the ones trading with Russia) declined by a larger margin in more Russian areas of Ukraine. Hence, if anything, one would expect the firms from more Russian counties to reduce their trade by a larger margin because of a more severe overall economic decline in their home areas – in other words, the reverse of our results. We rule out product-specific consumer boycotts, sanctions, and other product-specific shocks by including differential trends for different four-digit product codes, which do not change our results. Our results are not due to a differential increase in transportation costs since they are robust to flexible controls for the effective distance to Russian border. Finally, it is highly unlikely that firms from ethnically Ukrainian regions are being discriminated against at the Russian border, since their trade with Kazakhstan, which has to go through the Russian border, was not affected.

[9] We are not concerned that consumer bias can be transmitted to trade of intermediate goods, since our results hold for firms that have never traded consumer goods in 2013-2016.

[10] It could be that these areas produced more soldiers which in turn hurt the firms' overall performance.

Our work relates to a vast literature in political economy and development economics on the consequences of armed conflict.[11] Broadly speaking, this literature consists of three sets of studies: (i) studies focusing on the economic effects of conflict, (ii) research studying its psychological impact, and (iii) the work that deals with the political consequences of conflict. As this paper studies the impact of conflict on firm trade decisions, it naturally falls into the first category. At the level of countries and regions, several studies have estimated a large negative macroeconomic impact of international and civil wars (Abadie and Gardeazabal, 2003; Martin et al., 2008; Glick and Taylor, 2010; Miguel and Roland, 2011). When looking at the individual level economic effects of conflict, research has documented a strong negative impact of armed conflict on human capital accumulation and labor market outcomes (Blattman and Annan, 2010; Shemyakina, 2011; Chamarbagwala and Morán, 2011; Leon, 2012). At the intersection of groups (i) and (ii), several papers have documented the effect of armed conflict on fundamental economic preferences, such as the change of risk preferences towards a greater certainty premium (Callen et al., 2014) and a greater present-bias in discounting (Imas et al., 2015). Surprisingly, although many studies find that armed conflict leads to erosion of trust (Rohner et al., 2013a; Cassar et al., 2013; Besley and Reynal-Querol, 2014), a robust finding has recently emerged that war in fact increases pro-social behavior (Bauer et al., 2016).[12]

The closest papers to ours are the ones studying the impact of intra-country violence on firms' decision-making and performance. Guidolin and La Ferrara (2007) provides the time-series evidence that a break-out of civil war in Angola decreased the stock market value of firms operating in the country. Ksoll et al. (2014) analyze the direct effect of violence on the exporters in Kenya that manifested itself mainly through a sharp increase in workers' absence. Amodio and

[11] See Blattman and Miguel (2010) for a survey of the literature on civil conflicts.

[12] Notably, Dell and Querubin (2017) find that U.S. bombing in Vietnam in fact led to a more active insurgency and reduced civic engagement.

Di Maio (2017) show that Palestinian firms from high-conflict areas substitute domestically produced materials for the imported ones as a response to Second Intifada. In contrast to the above studies, this paper provides the first estimates for the effects of conflict on transactions between firms which are not directly affected by violence. It is also the first study to empirically examine the potential of conflict to undermine business relationships through increased animosity and decreased trust.

We also relate to the literature on the role of trust in trade. Rohner et al. (2013b) build a model of inter-ethnic conflict and trade in which conflict sends a negative signal about trustworthiness of the aggressive ethnic group and, as a result, inter-ethnic trade declines. In this paper, we provide the first empirical test of this theory. Guiso et al. (2009) were the first to document a strong relationship between mutual trust and trade between European countries. Nunn (2007) has shown the importance of a hold-up problem in trade by demonstrating that countries with better contract enforcement specialize in production of goods with more relation-specific inputs.[13] In this paper, we deal with a large shock to Russian-Ukrainian relations which enables a test of whether a decrease in trust can have an immediate causal impact on trade.[14]

The paper is organized as follows. section 3.2 gives the historical background on ethnic divisions in the Ukraine and on Russian-Ukrainian trade. section 3.3 de-

[13]The importance of good institutions for trade has also been shown by Levchenko (2007).

[14]Our paper is also connected to the literature on political disputes and boycotts (Heilmann, 2016; Luo and Zhou, 2016; Tanaka et al., 2017; Pandya and Venkatesan, 2016; Fouka and Voth, 2016; Edwards et al., 2007). However, it is distinct from this strand of research in several respects. First of all, consumer boycotts cannot be the only explanation of our results, as we document a significant effect of conflict on firms that have only traded intermediate goods in 2013-2016. In fact, this paper is the first to show the importance of animosity in a business-to-business environment using objective data on trade transactions. Second, in contrast to the peaceful political disputes normally studied by this literature, Russian-Ukrainian conflict was an armed conflict resulting in violence and thousands of deaths. Finally, in contrast to the short-lived impact observed in the existing studies, we show that the Russian-Ukrainian conflict left a long-lasting mark on firms' trade relationships, continuing two years after the annexation of Crimea.

scribes the empirical strategies. section 3.4 discusses the data used in the analysis and provides descriptive statistics. section 3.5 displays our baseline difference-in-differences results, rules out some of the alternative explanations and provides additional robustness checks. section 3.6 outlines our conceptual framework for thinking about indirect effects of conflict on trade. section 3.7 takes the predictions of our model to the data and, thus, attempts to disentangle between potential mechanisms. section 3.8 concludes.

3.2 Background

3.2.1 Ethnic and Regional Divisions within Ukraine

Historically, many regions of Ukraine had a large Russian ethnic minority. The ethnic divide in the country is still pronounced. Figure A1 (in the Appendix) shows the geographical variation in the share of Russians. In the West, the population is exclusively Ukrainian with the share of Russians being less than 1% on average. The central part has a sizable Russian minority varying from 1% to 20% of the population. Finally, the East of the country and Crimea has a Russian majority in some areas.[15]

This ethnic divide manifests itself in a political divide between the Ukrainian West and the 'Russian' East. The Western part of the country usually supported pro-European and nationalistic candidates, while the Eastern part of the country supported pro-Russian candidates. Figures A4 and A5 illustrate this political divide by showing voting patterns in 2012 parliamentary elections and 2004

[15] These data comes from 2001 census. Similar patterns can be seen if one looks at the linguistic heterogeneity measured by 'first language' and 'language used' (Figures A2 and A3, also based on 2001 census), as well as language of social network accounts (Figure A6, where `vk.com` is a Russian equivalent of Facebook). According to census data and independent surveys, 29.6% considered Russian as their mother tongue and approximately 60% used it at home on a daily basis.

presidential elections.[16] This political divide was one of the reasons for the political cycle. Pro-European Victor Yushchenko was the president from 2005 to 2010 following by a pro-Russian Victor Yanukovych, who also lost power to a pro-European Petro Poroshenko.

3.2.2 Russian-Ukrainian Conflict (2014-)

The last transition of power was a result of the Ukrainian revolution of 2014. In November 2013, the President of Ukraine Victor Yanukovych walked back on his promise to enter a political and economic association with the European Union. This step led to massive protests in Kiev and to its violent suppression by Yanukovychs police forces on November 29, 2013. During the next several months, the protests rose tremendously and spread across the country. As a result, on February 22, 2014, Victor Yanukovych fled to Russia. At this moment, the Russian government decided to leverage the political situation, annex Crimea, and promote separatist movements in the Eastern Ukraine. Annexation of Crimea in February 2014 went without a direct military conflict. However, the rebellion in Donetsk and Luhansk regions was not an immediate success for the pro-Russian militants. On the contrary, it led to a long-lasting civil conflict, with up to 10,000 casualties in total.

Figure A7 shows the areas affected by the conflict. These include the annexed Crimea (in white), two quasi-independent states of Donetsk and Luhansk Peoples Republics (in red), and other raions of Donetsk and Luhansk regions which are not part of the separatist territory (in orange). Since all of these areas have been directly affected by conflict, we focus our analysis on the rest of the country (in blue). While the conflict was intense in the affected provinces, the rest of the country enjoyed peace and functioned in a business-as-usual mode.

[16] These electoral maps are the intellectual property of Serhij Vasylchenko.

Despite increasing tensions, trade between Russia and Ukraine is still important for both countries. Ukraine was the fifth largest exporter to Russia in 2011 with 5.8% of all goods imported to Russia coming from Ukraine. This share dropped to 2.3% after the start of the conflict, with Ukraine being the eleventh largest exporter to Russia in 2014. Up to this date, Russia imports a wide variety of products from Ukraine: machines and engines, chemicals, paper, agriculture, processed food, iron, and steel. Russia was and still is the primary trading partner for Ukraine. Ukrainian imports from Russia include oil, gas, and other natural resources.

Another important feature of the environment is that Russia and Ukraine have been part of the Commonwealth of Independent States Free Trade Area (CISFTA) since the fall of the Soviet Union in 1991. Under CISFTA, all import tariffs were set to zero, with a lone exception for white sugar. The tariffs between the two countries started to go up only after Ukraine left CISFTA in January 2016, i.e., almost two years after the start of the conflict. Taken together, the Russian-Ukrainian conflict presents a perfect laboratory for the analysis of the indirect effects of conflict.

3.2.3 Changes in Attitudes after the Conflict

Generally, conflicts present large negative shocks to the relationships between the clashing groups. In this subsection, we examine the attitudes Russians and Ukrainians had towards each other and show that the conflict indeed left a lasting mark on the bond between the two nations.

First, we use the poll data to track the shares of Ukrainians favorable to Russia and of Russians favorable to Ukraine. Figure 1a and 1b display these data plotted over time. As one can see, before the conflict, Russians and Ukrainians shared an overwhelmingly friendly attitude towards each other's countries. The share of

Ukrainians favorable towards Russia before the conflict fluctuated between 80% and 90%, while the share of Russians favorable towards Ukraine has been close to 70-75%.[17] Such high levels of camaraderie reflect a long history of the two nations being part of the same country, which ended with the fall of the Soviet Union in 1991. However, in the immediate aftermath of the conflict, these numbers took a deep dive and almost halved in a matter of months to just above 50% and 35%, respectively. Moreover, they continued to fall until the end of 2014 and stayed low ever since. As shown by the red lines, this change is not due to respondents turning indifferent and is instead driven by more antagonistic attitudes towards the opposing state.

Although Russia and Ukraine are relatively ethnically homogeneous,[18] it could be that Figure 1 is not indicative of a rise in inter-ethnic animosity but instead highlights a worsened relationship between the two nations. If the latter was indeed the case, one would expect ethnic Russians and ethnic Ukrainians living in Ukraine to have a similar reaction to the start of the conflict. However, Figure 2 shows that this is not the case and that the change in attitudes displayed on Figure 1a is driven almost fully by ethnic Ukrainians. Specifically, the share of ethnic Ukrainians favorable towards Russia fell dramatically from 86% to 53% two months after the start of the conflict. Surprisingly, the share of ethnic Russians living in Ukraine and favorable to Russia fell only by 11%, from a near consensus (99%) to an overwhelming majority (88%). These results show that increased antagonism towards Russia among Ukrainians is almost fully driven by ethnic Ukrainians and is generally not shared by ethnic Russians, naturally leading to a rise in inter-ethnic tensions.

[17] For the purposes of brevity, we only present the numbers starting in Sep 2012. However, data before Sep 2012 show that these favorable attitudes persisted over time.

[18] According to the 2001 Ukrainian Census, 77% of Ukrainian population are ethnic Ukrainians. Similarly, per 2010 census, 81% of the population of Russian Federation is ethnically Russian.

3.3 Empirical Strategy

The goal of our empirical exercise is twofold. First, we provide the first time-series evidence on the change in trade between Russian and Ukrainian firms after the start of the conflict. More importantly, the paper identifies the effect of conflict on trade transactions between firms that are associated with different conflict parties.

To provide reduced-form evidence on the overall impact of conflict, we first compare firms trade intensity before and after the conflict has started. Specifically, we estimate the following model:

$$Y_{it} = \alpha_i + \gamma \times Post_t + \epsilon_{it} \tag{3.1}$$

where the outcome variable Y_{it} is the trade activity of firm i at time t; $Post_t$ is an indicator for whether a given time period falls before or after the start of the conflict; α_i are the firm-level fixed effects, and ϵ_{it} are the unobserved firm-time specific shocks. The validity of this specification relies on several assumptions. First, no other simultaneous events should have affected trade activity between the two countries unless they were caused by the conflict itself. For instance, a rapid decline in economic growth after February 2014 would not be a concern because it was one of the consequences of waging the conflict. Second, the model should match the data generating process and should not be misspecified. That is, firm's trade before and after conflict should behave as a firm-specific constant with noise.[19] If these assumptions hold, regression (1) will identify the effect of conflict on firms trade activity.

Next, we study the heterogeneous response to conflict across ethnic lines by utilizing a difference-in-differences strategy. That is, we compare trade intensity

[19]This assumption can potentially be restrictive as it implies that firm's trade cannot exhibit any time trends. However, graphic evidence presented in Section 3.5.1.1 suggests that it may hold in this context.

before and after the start of the conflict for firms located in more versus less ethnolinguistically Russian counties within Ukraine. If we find that, absent any trade restrictions or other mechanical changes, similar firms react to the conflict differently if they are located in more Russian versus less Russian areas of Ukraine, we would attribute this finding to an increase in inter-ethnic tensions.[20] Specifically, we run a regression of the following form:

$$Y_{ijt} = \alpha_i + \theta_t + \gamma_j \times Post_t + \beta \times Rus_i \times Post_t + \epsilon_{ijt} \tag{3.2}$$

where the outcome variable Y_{ijt} is the trade activity of firm i, for good j, at time t; α_i, θ_t and γ_j are, respectively, the firm, time, and product fixed effects; Rus_i is the share of Russian population in the county of firm i in 2001, or any other measure of ethnic alignment with Russia; and $Post_t$ is an indicator for a post-conflict period. We use two main versions of this specification: (1) monthly firm-level trade with firm and year-month fixed effects; (2) monthly firm-product-level trade with firm, year-month, and product-post fixed effects. If trade patterns for more and less Russian areas follow the same time trend absent the conflict, the coefficient on the interaction term identify the differential impact of conflict on firms with various degree of alignment to Russia.

3.4 Data

3.4.1 Data Sources

The empirical analysis in the paper combines transaction-level dataset on Ukrainian trade with demographic census data and firm-level accounting information. In addition, we examine a nationally-representative survey for tracking the changes in popular opinion before and after the start of the conflict.

[20] We will address all of the possible alternative explanations in Section 3.5.2.

Our unique dataset on the international trade of the Ukrainian firms (both export and import) includes dates, weights, values (in Ukrainian hryvnas), and product codes of each transaction. The data are for the period of 2013-2016 and include all trading partners, not only Russian firms.

An important feature of our trade dataset is that it includes addresses of the Ukrainian firms which then allow us to merge it with the census data at the county ('raion') level. We use the latest Ukrainian Census, which was collected in 2001. The most important census information for our analysis is the percentage of Russian-speakers and the share of ethnic Russians within a county.[21]

We then merge firm-level trade transactions to the ORBIS/AMADEUS Database. These dataset contains the accounting information on total sales, profits, and inputs of individual firms. In addition, the dataset contains the names of the managers which we then use to calculate a proxy for the prevailing ethnicity of the firms' key decision-makers.

Based on a ten-digit HS product code available for every trade transaction, we are able to categorize transactions depending on the type of the traded good. For instance, based on the correspondence tables between the HS and BEC codes,[22] we classify each entry as an intermediate good or a consumer good transaction.[23] Similarly, we use the methodology in Rauch (1999) to categorize each transaction into the ones involving differentiated goods and the ones with homogeneous goods.[24]

[21]The census question that determines one's language is 'What is your mother tongue?' Thus, these data may potentially be different from the share of people speaking Russian on a daily basis. However, it may better reflect one's national identity than ethnicity alone.

[22]We use the official conversion table between HS 2012 and BEC 4 product codes which can be found here.

[23]We use the official COMTRADE classification of BEC codes into capital, intermediate, and consumption goods (see details here). We then combine intermediate and capital goods into a single category under the name 'intermediate goods'.

[24]First, we use the official conversion table between the HS 2012 and SITC 2 product codes which can be found here. We then use data from Rauch (1999), available here, to categorize SITC 2 product codes into differentiated, reference priced, or homogeneous goods. For the rest of the

Finally, in order to trace the changes in inter-ethnic attitudes and beliefs, we use nationally-representative surveys of Ukrainians from the Kyiv International Institute of Sociology (KIIS) and of Russians from Levada Center. The surveys were designed to track the opinions and views of Ukrainian and Russian people and were conducted four to five times per year. Although these survey data are available for a longer period of time, we will only use it for the period from September 2012 to September 2016. The sample for each wave of the KIIS survey includes two thousand adults in 110 localities across all Ukrainian regions.

3.4.2 Descriptive Statistics

Before turning to the main analysis, we present the summary statistics of our data. In addition, we provide the descriptive analysis of the overall decline in trading activity between Ukrainian and Russian firms after the start of the conflict.

Table 1 presents the basic summary statistics. In this study, we analyze data from 12,842 Ukrainian firms located in 426 Ukrainian raions over the period of 48 months, from Jan 2013 to Dec 2016.[25] An average firm in our sample traded with Russia every fifth month and, overall, engaged in roughly three trade transactions per month. As for the quantity of trade, an average firm traded 230 tons and 2.3 mln UAH worth of product per month.[26] Notably, the distributions of the total net weight and the total value traded have long right tails which motivates the use of logarithm transformations in our analysis. An average firm traded intermediate goods in 77% of its transactions, stressing the prevalence of the

paper, we combine reference priced products and the goods traded on an organized exchange into a single category we call 'homogeneous goods'. We use the more conservative classification in our analysis, although our results are robust to using the less conservative ('liberal') classification.

[25] Note that, unfortunately, we do not have data for export transactions during Feb-Jun 2013. Thus, for the firms which engage in export activity only, we observe their behavior over the period of 43 instead of 48 months.

[26] 230 tons is equivalent to 11-12 fully loaded trucks. As of Jan 2018, 2.3 mln UAH is equivalent to $80,000 worth of product.

B2B sector transactions in our dataset. Similarly, only 22% of an average firms' transactions involved homogeneous goods.

Although the average shares of ethnic Russians and Russian speakers in Ukrainian raions are not high, they mask large heterogeneity in ethnolinguistic composition across the country. Even after excluding the conflict areas, which were historically more Russian, our sample still contains raions with 53% of ethnic Russians and 75% of native Russian-speakers. Similarly, when we classify managers into Russians and Ukrainians, we find that depending on the classification method either 10 or 30% of the managers in Ukrainian trading firms have Russian last names. The fact that the representation of Russian last names in the Ukrainian firms that trade with Russia is higher than the average share of Russians across the country would be consistent with the classic finding in the trade literature that international commerce often occurs within ethnic networks (Rauch and Trindade, 2002).[27]

Finally, we present the description of the accounting data for all Ukrainian firms in the ORBIS/AMADEUS database. Although the match between the two datasets is not perfect,[28] the numbers reported in Table 1 can still shed some light on whether and in which way the firms that trade with Russia are different from the overall pool of Ukrainian firms. Consistent with the theories of firm productivity and trade (Melitz, 2003), we find that the firms that trade with Russia have, on average, higher sales, profits, and productivity relative to the overall pool of firms.

Table 1 presents a static picture of our data. To display the data in a more dynamic fashion, we examine the overall decline of trade activity in the aftermath

[27] Alternatively, it could be that the method of assigning ethnicity based on last names is over-estimating the share of Russians. For example, it could be the case that some of the managers with Russian last names do not consider themselves ethnic Russians even though they may have had Russian roots. Note that such measurement error would not affect our main results.

[28] Accounting data is available for 8,206 out of 12,842 firms in our main sample.

of the conflict. Figure 3 traces the change in the total number of Ukrainian firms trading with Russia over time. As one can see, the number of firms trading with Russia stayed relatively stable at around 3,500 per month. However, after the start of the conflict, this number declined to about 2,500 firms per month. Note that the number of firms trading with Russia in January is consistently lower than in other months. January is a short business month in Russia because of the New Year and Christmas holidays. After explicitly controlling for the 'January' indicator in a regression form, the effect of conflict on the number of firms is estimated as a loss of 1,000 firms trading with Russia per month.

The overall impact of conflict on trade presented in Figure 3 is sizable, especially given that Russia and Ukraine were major trading partners before the conflict. In Section 3.5, we will identify the ethnolinguistic component of this effect in a difference-in-differences framework. However, to preview our results, Figure 4 offers a visual representation of the trade patterns across raions with the share of Russians above and below the median. To construct this graph, we first regress the log of total weight traded with Russia by a given firm in a given month on firm-level fixed effects.[29] We then calculate the median residuals for two subsets of firms, depending on whether they are located in a county with more or fewer ethnic Russians. As one can see, in 2013, i.e., before the conflict, these two groups of firms behaved very similarly.[30] However, after the start of the conflict, firms from the counties with fewer Russians decreased their trade by a bigger margin relative to the firms from more Russian areas of Ukraine. Moreover, the gap between the two subsets of firms is always of the same sign and is increasing over time.

[29]This procedure allows for a better comparison of firms across industries. Since an average product in some industries (e.g., electronics) weighs significantly less than an average product in others (e.g., manufacturing), simple aggregates may not capture the importance of conflict for the industries with small-weight products. Firm-level fixed effects correct this problem by averaging out the log weight of the products normally traded by firms.

[30]As before, January is a systematic outlier because of a week-long New Year holidays in Russia. We take this into account in our regression results.

3.5 Results

3.5.1 Main Results

3.5.1.1 Time-series Results

This section presents the regression estimates for the overall decline in trading activity between Ukrainian and Russian firms after the start of the conflict.

First, we estimate the regression equation (3.1) for all firms not located in the conflict regions. Under the assumptions that the conflict was unexpected, that there were no other simultaneous shocks of similar magnitude, and that the fixed effects model describes the data generating process correctly, regression model (3.1) will provide consistent estimates for the overall effect of conflict on trade in non-conflict areas.

The resulting estimates are displayed in Table 2. First, we estimate the change in the probability of trade by a given firm in a given month after the beginning of the conflict. Column (1) shows that, after February 2014, this probability for an average firm declined by 7.2 percentage points, or 0.18 standard deviations. Columns (2) and (3) examine the effect of conflict on the volume of trade, measured by log-total weight and log-total value of the shipped goods.[31] The obtained estimates are highly statistically significant and suggest that an average Ukrainian firm experienced a substantial decline in trade volumes with Russia. The estimates correspond to 16.8% to 17.9% of a standard deviation decline in trade activity.

To assess the intensive margin effect of conflict on trade, we estimate equation (3.1) for a subsample of firms which have been trading with Russia both before and after February 2014. Columns (4)-(6) of Table 2 present the results. Evidently, firms that continued trading have substantially decreased their trade

[31] To be precise, we use a $\log(1 + x)$ transformation to allow for zero trade flows.

intensity. They reduced the frequency of their shipments—the probability that an average remaining firm trades with Russia in a given month falls by 17.5 percentage points (44.4% of a standard deviation). Moreover, they decreased the volume of their monthly shipments by 42.2 to 45.1% of a standard deviation. Thus, our findings in Columns (1)-(3) are not exclusively driven by firms exiting trade with Russia.

3.5.1.2 Difference-in-Differences Results

In the previous sections, we have established that the Russian-Ukrainian conflict led to a dramatic decrease in trade between the two countries and to a big shift in attitudes toward Russians. In this part of the paper, we attempt to see whether this two empirical facts are connected. Specifically, we will assess whether reduction in trade was smaller for Ukrainian firms that are closer associated with Russian ethnicity.

Table 3 presents regression estimates of the difference-in-differences equation (3.2), building on the intuition offered by Figure 4. Similar to the time-series results in Table 2, we estimate the effect of ethnic and linguistic divisions on trade using three different outcome variables: (i) a dummy for any export or import activity with Russia by a given firm in a given month, (ii) a logarithm of the total net weight traded by a given firm in a given month, and (iii) a logarithm of the total value traded.

We start with the percentage of ethnic Russians in Ukrainian counties ('raions') as our main measure of ethnic heterogeneity. Columns (1)-(3) of Table 3 show the results for the three outcomes described above. The interaction coefficient of interest for the probability of trade with Russia (column 1) is 0.091 (or 3% of a standard deviation). Together with the time-series results in Table 2, this result suggests that moving a firm from a Ukrainian county with an average share of ethnic Russians (15%) to a county with the maximum share of Russians among the counties in our sample (53%) would mitigate the negative effect of conflict on the probability of trade in a given month by 46%. When we use the log-volumes of trade as an outcome in Columns (2) and (3), we obtain very similar results.

One may wonder whether one would observe analogous patterns with a different measure of Russian-Ukrainian divisions. Columns (4)-(6) of Table 3 look at the shares of native Russian speakers. The results are strikingly similar to Columns (1)-(3), both in terms of statistical significance and in terms of magnitude. As before, all else held equal, if one moved an average firm from a county with an average share of Russian speakers (26%) to a 75% Russian-speaking county (maximum in our sample), this would mitigate the negative effect of conflict on the probability of trade by 31%.

To allow for a graphic exploration of our results, we present our estimates in an event-study form. That is, instead of having one indicator for all months after the start of the conflict, we interact counties' ethnicity with a full set of monthly dummy variables. Figure 5 displays the results. First, we find no evidence of pre-trends, as ethnic divides consistently did not matter for trade before the conflict. Thus, we find support for the main assumption of our difference-in-differences set-up. Second, the Russian-Ukrainian conflict continued to affect inter-ethnic trade long after its start, lasting up until the end of our data in Dec 2016. Such enduring effect radically differs from the short-lived response observed in the literature on political conflicts and consumer boycotts, suggesting that a more severe armed conflict can have a much deeper influence on trade between nations.

It is also of interest to see whether firms decreased trade with Russia due to financial constraints, i.e., if they experienced a simultaneous drop in their overall economic performance. We account for the firms' overall financial health by normalizing the outcome variables for each firm by its overall sales by including the logarithm of the firm's turnover as a covariate. Since the sales data are available only on a yearly basis, we estimate equation (3.2) at a firm-year level.[32] Table 4 presents the results. Note that the normalization of trade intensity by firms' total sales makes these results even stronger. These estimates support the hypothesis that firms in less ethnolinguistically Russian areas of Ukraine were not financially forced to decrease their trade with Russia by a larger margin.

3.5.2 Robustness

The results in the previous section suggest that conflict may have a negative impact on trade that operates through the worsening of inter-ethnic relations. Thus, this paper provides the first empirical evidence on the indirect effects of conflict on trade not related to violence or destruction of property, which previously existed only as theoretical hypotheses (Rohner et al., 2013b). However, before we proceed to exploring the mechanisms, we rule out some mechanical explanations of our findings and test their overall robustness.

First, one can ask whether geographical distance to Russia drives the results in Table 3, and not ethnicity per se. As can be observed on Figure A1, the areas with the fewest percentage of ethnic Russians are located closer to the west of Ukraine, further away from the Russian-Ukrainian border. Therefore, if the conflict substantially increased transportation costs, it could mechanically had a bigger impact on firms located away from the border in less Russian areas. To account for this alternative explanation of our results, we calculate the shortest path to Russia from each of the firms' geographic location and include its interaction with the 'post' dummy as a covariate in our regressions.[33] Table 5 shows that, after accounting for the distance to the border, the results are almost identical

[32]It is important to point out that when we switch from monthly to yearly data, the baseline estimates of equation (3.2) remain very similar to the ones in Table 3. These results are available upon request.

[33]We account for the change in the border after the start of the conflict by re-calculating the shortest path without taking into account the border between Russia and Donetsk and Luhansk regions.

to the ones in Table 2.[34] Thus, it is unlikely that ethnicity acts as a proxy for distance to Russia in our regressions.

Can the effect on the firms be solely driven by the overall economic performance in the firms' home counties? If the conflict had a smaller economic impact on the more Russian areas of Ukraine,[35] it could mechanically lead to the results we find. We address this possibility by tracking the difference between the economic performance of all firms, not only exporters or importers, in the counties with different shares of Russians over 2011-2015. Table 6 shows that firms from more Russian areas tend to perform substantially worse after the conflict, both in terms of their log-profits, log-sales, and their total factor productivity.[36] Overall, this finding is consistent with a stylized fact that the areas with more Russians were, on average, closer to the conflict and, therefore, took a bigger economic hit in the aftermath of the conflict. Hence, our results cannot be due to the disparities in overall economic performance across different areas of Ukraine.

Can the observed relationship be explained by product-specific changes? The main examples of such changes would be sectoral sanctions or consumer boycotts of specific Ukrainian products.[37] To address this issue, we rerun the regressions from Tables 3 and 5 at the firm-product-month level with additional four-digit product fixed effects and their interactions with the conflict indicator. That is,

[34] One may wonder whether a linear control for distance is enough. In a table available upon request, we show that inclusion of higher order polynomials of distance does not change the results either.

[35] For example, this could happen if more people from Ukrainian counties went to fight the rebels.

[36] While we admit the existence of a slight pre-trend for log-sales and some pre-conflict differences for log-profits and TFP, we would like to point out three things: (i) if one extrapolated the pre-trend for sales from 2012-2013 to 2014-2015, one would obtain the coefficients of much smaller magnitude relative to the truth, (ii) differences in growth rates for log-profits and TFP what would constitute a pre-trend, (iii) any pre-existing differences in growth rates of these outcomes would not affect our main results as Figure 5 refutes the existence of a pre-trend for trade with Russia.

[37] Note that all shocks which applied uniformly to all products would be ruled out by time fixed effects.

we go down to a more granular level of product-firm observations rather than just firm level and account for product-specific trends in trade. The coefficients go down in magnitude but are still pronounced and significant (see Table 7). Hence, it makes it highly unlikely that our results in Table 3 can be explained by product-specific sanctions or consumer boycotts from the side of Russian consumers.

3.6 Conceptual Framework

We use an adverse selection setting to illustrate the main indirect effects of conflict on trade: an increase in animosity and a decrease in trust. The model incorporates uncertainty over the quality of the good, trust, and ethnic bias.[38] Lower trust and higher bias both reduce the probability of trade; however, they produce diverging predictions which we then take to the data. Specifically, we find that, in contrast to a reduction of trust, an increase in bias should affect trade more for the goods with a lower baseline variance of quality (i.e., more homogeneous).

In our model, sellers would like to sell an object of privately known quality q to the buyers. The price of the good is denoted as p and is endogenously determined by demand and supply. The quality of the good is drawn from a uniform distribution $q \sim U[\frac{1-\phi}{2}, \frac{1+\phi}{2}]$ with a parameter ϕ guiding the dispersion of quality. Sellers possesses only one unit of quality q, and, conversely, each unit of quality q has only one seller.

The buyers are risk neutral and value the good of quality q at $v(q) = vq$ where v is the buyers' valuation of the good, same for all buyers. Buyers vary in their

[38]Rohner et al. (2013b) study an effect of conflict on trust in a different set-up. In their model, only 'uncivic' (or non-cooperative) population would ever attack other groups. Therefore, whenever conflict happens, it makes the attacked group update their beliefs regarding preference for cooperation of each individual in the attacking group. Our model shares some aspects with the model by Rohner et al. (2013b) but differs from it in several key aspects. First, our model takes into account possible ethnic animosity aroused by conflict that is unrelated to trust. Second, we take conflict as exogenous which may better reflect behavior of large countries as opposed to small ethnic groups. Finally, we put a heavy emphasis on heterogeneous quality of the good which helps us develop testable predictions and disentangle the mechanisms.

animosity level towards the seller $b \sim U[0, \bar{b}]$, which is observable to the buyer only. Sellers are willing to sell only if the quality of the good they possess is below the market price, $q < p$.

With probability ε, a buyer meets a random seller that he can trust. In this case, the buyer perfectly observes the quality of the good. If trade is mutually beneficial, the buyer and the seller split the total surplus equally. If trade is not profitable, the buyer goes to the general market where the quality of the goods is unknown. With the probability $1 - \varepsilon$, the buyer goes to the general market directly, without meeting a reliable seller.

First, consider the general market with goods of unknown quality. The buyers get a payoff of zero if they do not purchase the good. Thus, a buyer trades only if the level of animosity is such that the expected payoff from trade is positive:

$$U(p, b, \phi) = \mathbb{E}[v(q)|q < p] - p - b = \frac{v}{2}\left[p + \frac{1-\phi}{2}\right] - p - b \geq 0$$

Note that if $v > 2$, then $U(\bar{b}, \phi)$ depends positively on p. This is not an interesting case since a seller could increase the price of the good to infinity and ensure a purchase. To prevent this situation, we assume that $v < 2$. Furthermore, for illustrative purposes, we assume that the heterogeneity of the quality of the good is high enough, $\phi > 1$, so that the general market becomes the market of lemons with no trade.

Second, consider a buyer meeting a random seller that he trusts. If trade occurs, each participant gets half of total surplus, i.e., $[(p - q) + (vq - p - b)]/2$. Trade is only beneficial to the players if the surplus is positive, i.e., if $q \geq b/(v-1)$. Hence, the ex-ante probability of trade for a buyer with bias b is $[(1 + \phi)/2 - b/(v - 1)]/\phi$. Integrating over b and multiplying by ε, one gets the overall ex-ante probability of trade in this model:

$$P = \varepsilon \left[1 - \frac{\bar{b}}{2\phi} \left[\frac{\bar{b}}{v-1} - (1-\phi) \right] \right]$$

Now we derive the comparative statics of interest:

Proposition 1: (a) $\frac{\partial P}{\partial \bar{b}} < 0$, (b) $\frac{\partial P}{\partial \varepsilon} > 0$, (c) $\frac{\partial^2 P}{\partial \bar{b} \partial \phi} > 0$, (d) $\frac{\partial^2 P}{\partial \phi \partial \varepsilon} < 0$.

That is, higher average levels of bias, $\bar{b}$, and lower trust, ε, would both lead to a lower probability of trade. Moreover, from (c) we learn that the lower is the baseline variance of quality, the stronger is the effect of an *increase in bias*. At the same time, (d) states that if the conflict is a pure *decrease in trust*, the decline in trade should be more pronounced for the goods with higher baseline variance of quality. Hence, the two effects lead to two diverging empirical predictions regarding the impact of conflict on trade of homogeneous and specific goods.

3.7 Mechanisms

3.7.1 Increase in Bias Versus Decrease in Trust

In our conceptual framework, we focused on two ethnic-specific mechanisms of how conflict may affect trade. First, it could be that conflict increases inter-ethnic bias which presents extra costs to the inter-ethnic trade. Second, conflict may decrease trust between the two ethnic groups. While these mechanisms would both lead to a decrease in trade between ethnic groups, our model suggests a way to distinguish between the two by considering goods with high and low uncertainty regarding their quality. Specifically, Section 3.6 shows that while a decrease in trust should affect goods with high uncertainty more, an increase in bias has an opposite effect.

To test the implications of the model, we split the sample into the firms that trade homogeneous goods and the ones that do not. We define homogeneous goods as in Nunn (2007) as the goods that are either traded on the exchanges or are referenced priced. Table 8 presents the difference-in-differences estimates for both subsamples of firms. As one can see, the differential effect of conflict across ethnicities is fully driven by the traders of homogeneous goods and is zero for the traders of specific goods. Overall, these results suggest that an increase in bias is a better explanation for the rise in ethnic trade differential than a negative shock to inter-ethnic trust.[39]

If our results are really about bias, we would expect that consumer bias has a bigger influence on firms' decisions. However, if the problem is trust, it is not clear why it would affect consumer goods any differently than intermediate goods. To see whether we can find further support for the bias hypothesis, we split our sample of firms into intermediate goods traders and consumer goods traders (see details in Section 3.4). Under the assumption that the trade of intermediate goods is much less salient for the firms' final consumers, a pure consumer bias explanation would predict a null effect of conflict for intermediate goods and a large effect for consumer goods. Table 9 presents the results. Although the effect of conflict on inter-ethnic trade is indeed larger for consumer products, the coefficients for the intermediate goods are still positive and significant in all specifications. While these results provide an overall support for the bias explanation , they also suggest that consumer pressure are unlikely to be only driving force behind our difference-in-differences findings.

We identified two types of product characteristics that affect our results, and it is worth exploring how these interact. Specifically, is it true that most of the effect comes from homogeneous goods, independent of whether they are final or intermediate? Furthermore, does consumer bias manifest itself even for the trade of more complex products where firms ties are most important? Table 10

[39]One may be concerned that it is easier in general to switch partners when trading homogeneous goods and that it must automatically follow that any decrease in trade after conflict would come from homogeneous goods. As a result, the exercise in Table 8 is not a valid way to disentangle the mechanisms. Note, however, that this explanation is not consistent with the fact that traders of heterogeneous goods decreased trade by a larger margin after the start of the conflict (see Columns 7-9 of Table 8).

presents our difference-in-differences estimates for four types of firms: (i) those that trade only specific intermediate goods, (ii) specific consumer goods traders, (iii) homogeneous intermediate goods traders, and (iv) the firms that trade only homogeneous consumer goods. As we would expect, the trade of homogeneous consumer goods have been most affected by conflict across ethnic lines, while the DiD estimates are precise zeros for heterogeneous intermediate goods. The estimates are positive and significant, however, for the firms from groups (ii) and (iii). Overall, Table 10 suggests that both product heterogeneity and its salience for consumers play a role in how it reacts to conflict, further strengthening the case for bias and not trust as the primary mechanism.

3.7.2 Additional Results

3.7.2.1 Switching Trading Partners

In this subsection, we would like to see whether the conflict led to a permanent decrease in trade by firms from the counties with fewer Russians or whether they were able to find new partners in other countries. For the reasons of computation speed, we concentrate on Ukrainian exporters only. Table 11 presents the difference-in-differences results from Table 2 estimated instead for trade with Belarus, Moldova, Poland, Romania, i.e., the biggest trade partners of Ukraine out of its neighboring countries, Kazakhstan, as Kazakh-Ukrainian trade crosses the Russian-Ukrainian border, and the rest of the European Union. As one can see, there is no switching of trade towards other post-Soviet countries or to the most of the EU countries. However firms from the areas with fewer Russians tend to switch to their culturally close European neighbors, i.e., Romania and Poland. These calculations are consistent with a narrative in which conflict increased inter-ethnic tensions between Ukrainians and Russians and strengthened the Ukrainian ethnic identity.[40]

3.7.2.2 Ethnicity of the Area Versus Key Decision-makers

So far, our main variable of interest has been ethnolinguistic composition of the firm's home county. However, it leaves open a question whether what matters is the ethnolinguistic composition of the county as opposed to that of the firms. This question is especially salient given that these variables may be highly correlated and may serve as proxies for each other. While we cannot measure the total fraction of ethnic Russians in a firm due to data limitations, we can attempt to infer the ethnicity of the firms' top-level management from their surnames. Specifically, we measure the share of managers with the Russian-looking last names from ORBIS data set (calculated based on last names endings), and we use it instead of our measure of ethnic heterogeneity.

Table 12 displays the corresponding estimates. Evidently, ethnicity of the managers, as we measure it, does not produce the same results as the ethnicity of the area they work at. We repeat the exercise by running a horse race of the ethnicity of the managers and the ethnicity of the firm's county. The effect of the managers' ethnicity stays null while the effect of the counties' composition stays significant and strong (Table 14). This suggests that conflict does not affect trade through the managers ethnicity and instead operates through the culture and attitudes in the surrounding area.

3.7.2.3 Heterogeneity Across Regions

Given the geographic heterogeneity in ethnolinguistic composition, it would be interesting to explore heterogeneity of our results across regions. Table 13 replicates the main results from Table 3 for three geographic subsamples. The first one

[40] The strengthening effect of conflict on ethnic identity has been well documented in the literature—see, e.g., Rohner et al. (2013a).

is without the capital city, Kyiv, which hosts the big part of economic activity in Ukraine. Dropping Kyiv does not change our results at all, keeping the coefficients almost exactly the same as in Table 3. The second subsample drops three regions neighboring the conflict area of Donetsk and Luhansk. The results suggest that, if anything, coefficients increase in less ethnolinguistically Russian areas. The third subsample drops Western Ukraine that is predominantly Ukrainian. While we loose some statistical power for the results for any trade activity, other coefficients do not change much.

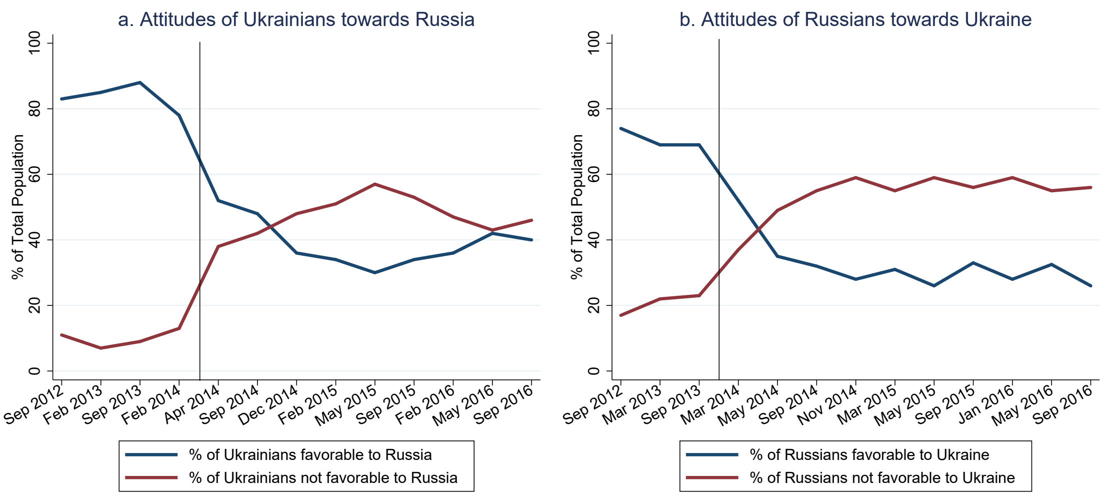
3.8 Conclusion

Conflicts have large and multifaceted effects on the economy. They can impact economic agents directly, through violence and damage to property, or indirectly, for example, by disrupting business relationships. While the existing literature offered some evidence on the former, the indirect effects of conflict are still largely understudied. This paper examines the destructive effect of the Russian-Ukrainian conflict on trade behavior of firms not located in conflict areas. Using uniquely rich transaction-level data on Ukrainian trade, we show that firms located in more ethnolinguistically Ukrainian counties experienced a deeper drop in trade with Russia relative to the firms from more Russian counties. We interpret our findings as arising from increased inter-ethnic tensions after the start of the conflict. Moreover, our evidence suggests that these are less likely to operate through a decrease in inter-ethnic trust and instead supports a hypothesis of an increased ethnic bias.

The Russian-Ukrainian conflict is unique in several respects. In particular, the opposing countries were major trade partners before the start of the conflict, never declared an official war, and did not levy a trade embargo or new tariffs on one another after the conflict broke out. Hence, it presents a rare opportunity to assess the indirect effects of conflict on trade. While this conflict may seem unusual from a historical point of view, we believe that it adequately reflects the implicit nature of modern warfare. As such, we think that our results may help predict the impact of future conflicts between states on trade.

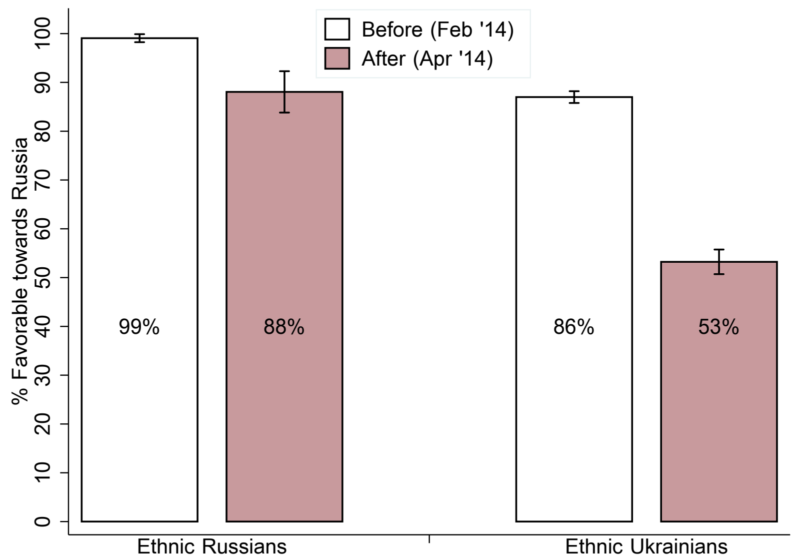
FIGURES AND TABLES

Figure 3.1: Upsurge in mutual antipathy between Russia and Ukraine.



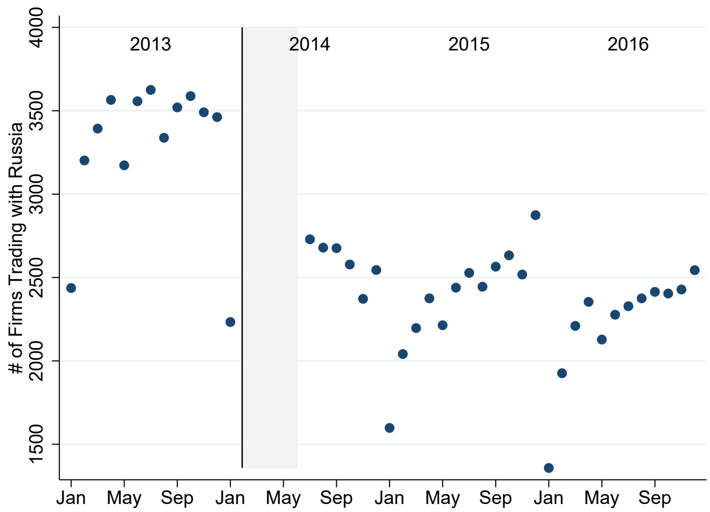
Notes: Data for Panel A come from a regular nationally-representative poll of Ukrainians conducted by Kyiv International Institute of Sociology. The Feb 2014 survey was conducted on 7-17 Feb 2014, i.e. before the start of the conflict. Data for Panel B come from a regular nationally-representative survey of Russians conducted by Levada Center.

Figure 3.2: Disproportional decrease in affinity towards Russia among ethnic Ukrainians.



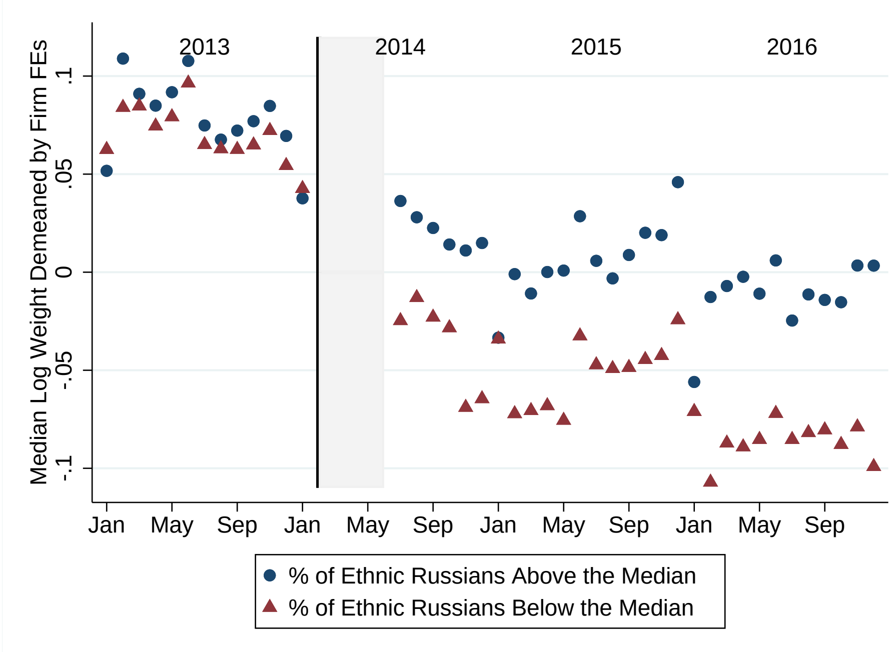
Notes: Data come from a regular nationally-representative poll of Ukrainians conducted by Kyiv International Institute of Sociology. The Feb 2014 survey was conducted on 7-17 Feb 2014, i.e. before the start of the conflict.

Figure 3.3: Change in the number of firms trading with Russia after the start of the conflict.



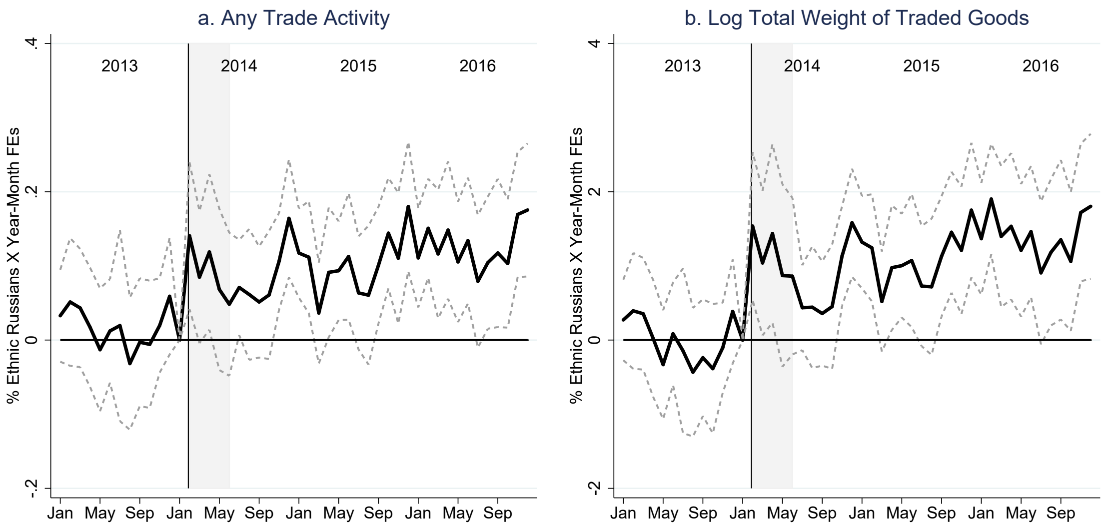
Notes: The number of firms trading with Russia includes both exporters and importers. The graph excludes all the firms located in the areas affected by conflict (see Figure A7). Export data are missing for the period of Feb-June 2014 (colored in gray). Thus, these months are removed for the aggregate comparisons. January is a short business month in Russia because of a long holiday week lasting from Jan 1 to Jan 7. Similarly, Ukraine has two official public holidays in January - New Year's Eve (Jan 1) and Orthodox Christmas (Jan 7). As such, one should view January data as seasonal outliers.

Figure 3.4: Change in firm-level trade. Breakdown by ethnicity of the area.



Notes: The data plotted are the monthly median residuals from a firm-level regression of the logarithm of the total weight traded on firm fixed effects. Data are broken down by the share of ethnic Russian population in the host county of each firm, taken from 2001 Ukrainian Census. Seasonality has been removed by regressing residuals on month fixed effects and on the interaction between January indicator and the share of ethnic Russians in the area. Export data are missing for the period of Feb-June 2014 (colored in gray). Thus, these months are removed for the purpose of this graph until we are able to control for year-month fixed effects. All calculations exclude the firms located in the areas affected by conflict (see Figure A7).

Figure 3.5: Month-by-month Effect of Russian-Ukrainian conflict on Trade



Notes: This graph displays the results of estimating equation (2) with the addition of the interactions between year-month fixed effects and ethnolinguistic composition of the firms' host counties. Note that, aside from this interaction, equation (2) includes firm and year-month fixed effects. Only import data are present for the period of Feb-Jun 2014 (colored in gray). However, year-month fixed effects help account for the mechanical decrease in trade intensity during these months. Removing Feb-Jun 2014 from our analysis does not change the results. The event study graphs for export and import data separately are presented in the Appendix.

Table 3.1: Chapter 3 Tables

Table 1. Summary statistics.

	Observations	Mean	SD	Min	Max
Trade Transaction Data					
Indicator for Any Trade Activity in a Given Month	590,419	0.20	0.40	0	1
Log (1 + Total Net Weight Traded in a Given Month)	590,419	1.97	4.14	0	21.3
Log (1 + Total Value Traded in a Given Month)	590,419	2.73	5.51	0	22.8
Number of Transactions in a Given Month	590,419	3.16	32.2	0	5,420
Total Net Weight Traded in a Given Month, in Tons	590,419	230	6,824	0	1,709,763
Total Value Traded in a Given Month, in UAH '000	590,419	2,361	44,035	0	8,045,764
Types of Goods Traded					
Share of Intermediate Goods Traded by a Firm in 2013-2016	12,842	0.77	0.36	0	1
Share of Consumer Goods Traded by a Firm in 2013-2016	12,842	0.17	0.34	0	1
Share of Homogeneous Goods Traded by a Firm in 2013-2016	12,837	0.22	0.39	0	1
Ethnic Composition of Ukrainian Raions					
Share of Russian Speakers, Census 2001	426	0.07	0.08	0	0.53
Share of Ethnic Russians, Census 2001	426	0.09	0.13	0	0.75
Ethnic Composition of Ukrainian Firms' Management					
Share of Managers w/ Russian Last Names, Method #1	10,302	0.30	0.43	0	1
Share of Managers w/ Russian Last Names, Method #2	10,302	0.10	0.28	0	1
Distance to the Border					
Shortest Path to Russian Border, Pre-Conflict, km	11,756	247.2	163.7	1.5	793.7
Shortest Path to Russian Border, Post-Conflict, km	11,756	253.7	164.9	1.5	793.7
Accounting Data					
IHS Transformation of Profits, All Firms, 2011-2015	1,266,384	10.13	6.89	-19.41	24.68
IHS Transformation of Sales, All Firms, 2011-2015	1,266,532	12.58	4.67	-18.49	26.10
Total Factor Productivity, All Firms, 2011-2015	1,212,987	13.18	1.87	8.94	25.26
IHS Transformation of Profits, Traders Only, 2013-2015	24,616	14.89	4.56	-19.41	22.01
IHS Transformation of Sales, Traders Only, 2013-2015	24,618	16.66	2.97	0	23.56
Total Factor Productivity, Traders Only, 2013-2015	24,403	15.54	1.96	8.94	21.43

Notes: Data on trade include both export and import transactions. Homogeneous goods are defined as in Rauch (1999). Intermediate goods are specified by the standardized BEC classification. Individual is considered a Russian speaker if Russian is his or her mother tongue. Method #1 of calculating the share of managers with Russian last names treats a last name as Russian if it ends in "ov", "ova", "ev", "eva", "in", "ina", "yov", "yova". Method #2 uses a bank of last names that are generally considered ethnic Russian. Shortest path to the Russian border for the periods after conflict excludes parts of the border that are located in conflict regions. IHS stands for inverse hyperbolic sine transformation $L(X) = \log(X + \mathrm{sqrt}(X^2+1))$ as in MacKinnon and Magee (1990). Total factor productivity is derived from a Cobb-Douglas specification regressing turnover on capital and labor (all in logs) with two-digit industry fixed effects.

Table 2. Reduced Form Estimates of the Reduction in Trade

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded
	All Firms			*Firms Trading Before and After*		
Post Feb 2014	-0.072***	-0.736***	-0.911***	-0.175***	-1.857***	-2.290***
	(0.003)	(0.034)	(0.042)	(0.005)	(0.062)	(0.071)
Standardized Beta Coefficient	*-0.182*	*-0.179*	*-0.168*	*-0.444*	*-0.451*	*-0.422*
Year-Month Fixed Effects	NO	NO	NO	NO	NO	NO
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	591,417	591,417	591,417	225,149	225,149	225,149
# of Firms	12,869	12,869	12,869	4,847	4,847	4,847
# of Counties	446	446	446	332	332	332
# of Months	48	48	48	48	48	48
R-squared	0.407	0.482	0.455	0.394	0.504	0.452

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods (export+import) are calculated by transforming the initial variable X with L(X) = log(X + 1). Columns (1) and (4) use an indicator for a firm trading with Russia in a given month. Columns (4-6) restrict the sample to firms that were trading with Russia both before February 2014 and after February 2014.

Table 3. Baseline Results with Firm and Year-Month Fixed Effects

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded
(Post Feb 2014) × (% of Russian Ethnicity)	0.091***	1.153***	1.282***			
	(0.035)	(0.388)	(0.464)			
Standardized Coefficient	*0.030*	*0.037*	*0.031*			
(Post Feb 2014) × (% Russian Language)				0.043***	0.565***	0.608***
				(0.017)	(0.183)	(0.218)
Standardized Coefficient				*0.026*	*0.033*	*0.027*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	590,419	590,419	590,419	590,419	590,419	590,419
# of Firms	12,848	12,848	12,848	12,848	12,848	12,848
# of Counties	426	426	426	426	426	426
# of Months	48	48	48	48	48	48
R-squared	0.412	0.486	0.459	0.412	0.486	0.459

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods (export+import) are calculated by transforming the initial variable X with L(X) = log(X + 1). Data on ethnolinguistic composition is at the county level, and it comes from Ukrainian 2001 Census. Russian language is measured as the % of people who named Russian as their mother tongue ("rodnoi yazik").

Table 4. Difference-in-Differences Results Accounting for Firm Sales, 2013-2015

VARIABLES	(1) Log of Total Weight Traded	(2) Log of Total Value Traded	(3) Log of Total Weight Traded	(4) Log of Total Value Traded
(Post 2014) × (% of Russian Ethnicity)	3.723***	4.327***		
	(0.795)	(0.896)		
Standardized Coefficient	*0.092*	*0.082*		
(Post 2014) × (% Russian Language)			1.761***	2.012***
			(0.361)	(0.429)
Standardized Coefficient			*0.077*	*0.068*
Firm-Level Yearly Revenue	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES
# of Observations	24,531	24,531	24,531	24,531
# of Exporters	8,177	8,177	8,177	8,177
# of Counties	389	389	389	389
R-squared	0.614	0.563	0.614	0.563

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value, net weight of shipped goods, and of sales are calculated by transforming the initial variable X with L(X) = log(X+1). Log-sales are included as a covariate. Due to data availability on sales, the analysis is restricted to 2013-2015. Data on ethnolinguistic composition is at the county level, and it comes from Ukrainian 2001 Census. Russian language is measured as the % of people who named Russian as their mother tongue ("rodnoi yazik"). Instead of splitting data on the monthly level, they are aggregated to the yearly level.

Table 5. Results with Firm and Year-Month Fixed Effects and Controls for Distance to Russian Border

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded
(Post Feb 2014) × (% of Russian Ethnicity)	0.087**	1.243***	1.253**			
	(0.043)	(0.451)	(0.557)			
Standardized Coefficient	*0.029*	*0.040*	*0.030*			
(Post Feb 2014) × (% Russian Language)				0.042**	0.609***	0.600**
				(0.020)	(0.211)	(0.263)
Standardized Coefficient				*0.026*	*0.036*	*0.027*
(Post Feb 2014) × (Shortest Path to Russia, in 1000 km)	-0.006*	-0.034	-0.062	-0.006*	-0.040	-0.069*
	(0.003)	(0.032)	(0.042)	(0.003)	(0.031)	(0.040)
Standardized Coefficient	*-0.015*	*-0.009*	*-0.012*	*-0.015*	*-0.010*	*-0.013*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	542,633	542,633	542,633	542,633	542,633	542,633
# of Firms	11,756	11,756	11,756	11,756	11,756	11,756
# of Counties	421	421	421	421	421	421
# of Months	48	48	48	48	48	48
R-squared	0.414	0.490	0.462	0.414	0.490	0.462

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The results are robust to clustering at the regional level. The logs of total value and net weight of shipped goods (export+import) are calculated by transforming the initial variable X with L(X) = log(X + 1). Data on ethnolinguistic composition is at the county level, and it comes from Ukrainian 2001 Census. Russian language is measured as the % of people who named Russian as their mother tongue ("rodnoi yazik"). All specifications control for the distance to the Russian border, accounting for changes due to conflict in Luhansk and Donetsk regions.

Table 6. Conflict and County-level economic performance

VARIABLES	(1) Log Profit	(2) Log Sales	(3) TFP	(4) Log Profit	(5) Log Sales	(6) TFP
	S = (% with Russian Mother Tongue in a County)			S = (% of Russian Ethnicity in a County)		
(Year = 2012) × (% S)	0.118	-0.003	0.051***	-0.023	-0.070	0.104***
	(0.103)	(0.036)	(0.016)	(0.250)	(0.076)	(0.026)
(Year = 2013) × (% S)	0.253**	-0.077	0.069***	0.343	-0.192*	0.142***
	(0.119)	(0.050)	(0.025)	(0.254)	(0.101)	(0.041)
(Year = 2014) × (% S)	-0.354***	-0.428***	0.008	-0.801***	-0.884***	0.025
	(0.133)	(0.076)	(0.028)	(0.258)	(0.148)	(0.052)
(Year = 2015) × (% S)	-1.005***	-0.803***	-0.060*	-1.966***	-1.653***	-0.107*
	(0.169)	(0.134)	(0.030)	(0.308)	(0.271)	(0.062)
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	839,665	840,005	757,905	839,665	840,005	757,905
# of Firms	167,933	168,001	151,581	167,933	168,001	151,581
R-squared	0.527	0.772	0.940	0.527	0.772	0.940

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. Excludes firms from conflict areas. Data on ethnolinguistic composition is at the county level and comes from Ukrainian 2001 Census. Russian language is measured as the % of people who named Russian as their mother tongue ("rodnoi yazik"). Dependent variables in columns (1), (2), (4), (5) are gross profit and total sales transformed using the inverse hyperbolic sine function L(.), such that $L(X) = \log(X + \mathrm{sqrt}(X^2+1))$ as in MacKinnon and Magee (1990). Total factor productivity in (3) and (6) is derived from a Cobb-Douglas specification regressing turnover on capital and labor (all in logs) with two-digit industry fixed effects.

Table 7. Results with Firm, Year-Month, and Four-Digit Product Code Fixed Effects

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded
Panel A. Product-level Results with Product-Post Fixed Effects						
(Post Feb 2014) × (% of Russian Ethnicity)	0.052**	0.508***	0.676***			
	(0.021)	(0.168)	(0.223)			
Standardized Coefficient	*0.021*	*0.024*	*0.023*			
(Post Feb 2014) × (% Russian Language)				0.027***	0.288***	0.364***
				(0.010)	(0.093)	(0.115)
Standardized Coefficient				*0.020*	*0.025*	*0.022*
Panel B. Product-level Results with Product-Post Fixed Effects and Distance Controls						
(Post Feb 2014) × (% of Russian Ethnicity)	0.043	0.530***	0.609**			
	(0.027)	(0.201)	(0.287)			
Standardized Coefficient	*0.018*	*0.026*	*0.021*			
(Post Feb 2014) × (% Russian Language)				0.023*	0.297***	0.334**
				(0.012)	(0.103)	(0.139)
Standardized Coefficient				*0.017*	*0.026*	*0.020*
4-Digit Product Code-Post Fixed Effects	YES	YES	YES	YES	YES	YES
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	2,148,754	2,148,754	2,148,754	2,148,754	2,148,754	2,148,754
# of Firms	11,722	11,722	11,722	11,722	11,722	11,722
# of 4-Digit Product Codes	1,061	1,061	1,061	1,061	1,061	1,061
# of Counties	421	421	421	421	421	421
# of Months	48	48	48	48	48	48

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The results are robust to clustering at the regional level. The logs of total value and net weight of shipped goods (export+import) are calculated by transforming the initial variable X with L(X) = log(X + 1). Data on ethnolinguistic composition is at the county level, and it comes from Ukrainian 2001 Census. Russian language is measured as the % of people who named Russian as their mother tongue ("rodnoi yazik"). Product code is defined from harmonized system (HS). One observation is a firm-month-four digit product.

Table 8. Homogeneous and Specific Goods

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded	(7) Any Trade Activity	(8) Log of Total Weight Traded	(9) Log of Total Value Traded
	Specific Goods Traders			*Homogeneous Goods Traders*			*Triple Difference*		
(Post Feb 2014) × (% of Russian Ethnicity)	0.024	0.269	0.278	0.389***	5.102***	5.355***	0.025	0.283	0.299
	(0.045)	(0.466)	(0.583)	(0.079)	(1.110)	(1.144)	(0.045)	(0.464)	(0.580)
Standardized Coefficient	*0.008*	*0.009*	*0.007*	*0.128*	*0.162*	*0.129*	*0.008*	*0.009*	*0.007*
(Post Feb 2014) ×							-0.051***	-0.848***	-0.771***
× (Non-Zero Homogeneous Goods Trade)							(0.012)	(0.143)	(0.169)
Standardized Coefficient							*-0.060*	*-0.097*	*-0.067*
(Post Feb 2014) × (% of Russian Ethnicity) ×							0.204***	2.653***	3.014***
× (Non-Zero Homogeneous Goods Trade)							(0.062)	(0.708)	(0.855)
Standardized Coefficient							*0.060*	*0.075*	*0.065*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of Observations	389,288	389,288	389,288	75,362	75,362	75,362	590,371	590,371	590,371
# of Firms	8,537	8,537	8,537	1,634	1,634	1,634	12,848	12,848	12,848
# of Counties	386	386	386	248	248	248	427	427	427
# of Months	48	48	48	48	48	48	48	48	48
R-squared	0.331	0.379	0.361	0.305	0.365	0.335	0.412	0.486	0.459

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods are calculated by transforming the initial variable X with L(X) = log(X + 1). Homogeneous Goods Traders are the firms that have only traded traded homogeneous goods under the classification of Rauch (1999) over the course of 2013-2016. Rauch (1999) defines homogeneous goods as those either traded on the organized exchange or having reference prices. Specific Goods Traders are the firms that have not traded homogeneous goods under the Rauch (1999) classification. Columns (7)-(9) present results of a triple-difference estimation which analyzes patterns of trade before and after the start of the conflict, across the areas with more and less ethnic Russians, for firms which have or have not traded homogeneous goods at some point during our study period.

Table 9. Intermediate and Consumer goods

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded	(7) Any Trade Activity	(8) Log of Total Weight Traded	(9) Log of Total Value Traded
	Intermediate Goods Traders			*Consumer Goods Traders*			*Triple Difference*		
(Post Feb 2014) × (% of Russian Ethnicity)	0.093**	1.206***	1.270**	0.319***	2.976***	3.924***	0.250***	2.677***	3.421***
	(0.040)	(0.439)	(0.505)	(0.084)	(1.003)	(1.209)	(0.073)	(0.846)	(1.026)
Standardized Coefficient	*0.031*	*0.038*	*0.031*	*0.105*	*0.094*	*0.094*	*0.083*	*0.085*	*0.082*
(Post Feb 2014) X (% Intermediate Goods)							0.101***	1.067***	1.386***
							(0.014)	(0.152)	(0.186)
Standardized Coefficient							*0.095*	*0.097*	*0.095*
(Post Feb 2014) X (% of Russian Ethnicity) X X (% Intermediate Goods)							-0.201***	-1.951**	-2.725***
							(0.074)	(0.812)	(1.016)
Standardized Coefficient							*-0.069*	*-0.064*	*-0.068*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of Observations	358,806	358,806	358,806	61,270	61,270	61,270	578,060	578,060	578,060
# of Firms	8,101	8,101	8,101	1,396	1,396	1,396	12,586	12,586	12,586
# of Counties	385	385	385	216	216	216	423	423	423
# of Months	48	48	48	48	48	48	48	48	48
R-squared	0.329	0.399	0.365	0.277	0.314	0.304	0.413	0.491	0.462

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods are calculated by transforming the initial variable X with L(X) = log(X + 1). Intermediate goods traders in columns (1)-(3) are the firms that have only traded intermediate goods under the standardized BEC classification in 2013-2016. Consumer goods traders in columns (4)-(6) are the firms that have only traded consumer goods under the standardized BEC classification in 2013-2016. Columns (7)-(9) use all firms, i.e., not only intermediate or consumer good traders, in a triple difference specification.

Table 10. Homogeneous, Specific, Intermediate, and Consumer Goods

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded
	Specific Intermediate Goods Traders			*Specific Consumer Goods Traders*		
(Post Feb 2014) × (% of Russian Ethnicity)	0.029	0.286	0.348	0.234***	1.841*	2.753**
	(0.044)	(0.427)	(0.541)	(0.082)	(0.951)	(1.133)
Standardized Coefficient	*0.010*	*0.009*	*0.008*	*0.077*	*0.059*	*0.066*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	262,149	262,149	262,149	46,548	46,548	46,548
# of Firms	5,850	5,850	5,850	1,044	1,044	1,044
# of Counties	345	345	345	175	175	175
# of Months	48	48	48	48	48	48
	Homogeneous Intermediate Goods Traders			*Homogeneous Consumer Goods Traders*		
(Post Feb 2014) × (% of Russian Ethnicity)	0.334***	4.693***	4.622***	0.673***	7.411***	8.731***
	(0.084)	(1.248)	(1.188)	(0.149)	(1.600)	(1.978)
Standardized Coefficient	*0.110*	*0.149*	*0.111*	*0.223*	*0.236*	*0.210*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
# of Observations	54,240	54,240	54,240	10,669	10,669	10,669
# of Firms	1,175	1,175	1,175	243	243	243
# of Counties	212	212	212	96	96	96
# of Months	48	48	48	48	48	48

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods are calculated by transforming the initial variable X with L(X) = log(X + 1). Homogeneous Intermediate Goods Traders are the firms that have only traded homogeneous intermediate goods in 2013-2016, where homogeneous goods are defined as in Rauch (1999) and intermediate goods are specified by the standardized BEC classification. Rauch (1999) defines homogeneous goods as those either traded on the organized exchange or having reference prices. Other types of traders are defined accordingly.

Table 11. Trade with other countries

VARIABLES	(1) Belarus Any Export Activity	(2) Belarus Log of Total Weight Shipped	(3) Belarus Log of Total Value Shipped	(4) Kazakhstan Any Export Activity	(5) Kazakhstan Log of Total Weight Shipped	(6) Kazakhstan Log of Total Value Shipped	(7) Moldova Any Export Activity	(8) Moldova Log of Total Weight Shipped	(9) Moldova Log of Total Value Shipped
(Post Feb 2014) X (% of Russian Ethnicity)	-0.050	-0.323	-0.656	-0.007	0.157	-0.037	-0.004	0.097	-0.074
	(0.053)	(0.519)	(0.695)	(0.037)	(0.392)	(0.493)	(0.027)	(0.232)	(0.312)
Standardized Coefficient	*-0.017*	*-0.010*	*-0.017*	*-0.002*	*0.005*	*-0.001*	*-0.001*	*0.003*	*-0.002*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of Observations	194,575	194,575	194,575	97,868	97,868	97,868	239,897	239,897	239,897
# of Exporters	4,525	4,525	4,525	2,276	2,276	2,276	5,579	5,579	5,579
# of Counties	357	357	357	254	254	254	375	375	375
# of Months	43	43	43	43	43	43	43	43	43
R-squared	0.371	0.462	0.417	0.332	0.407	0.369	0.357	0.420	0.397

VARIABLES	(10) Poland Any Export Activity	(11) Poland Log of Total Weight Shipped	(12) Poland Log of Total Value Shipped	(13) Romania Any Export Activity	(14) Romania Log of Total Weight Shipped	(15) Romania Log of Total Value Shipped	(16) Other EU Countries Any Export Activity	(17) Other EU Countries Log of Total Weight Shipped	(18) Other EU Countries Log of Total Value Shipped
(Post Feb 2014) X (% of Russian Ethnicity)	-0.098**	-1.445***	-1.307**	-0.188**	-2.410**	-2.350**	0.020	0.308	0.374
	(0.042)	(0.385)	(0.509)	(0.088)	(0.969)	(1.158)	(0.028)	(0.253)	(0.350)
Standardized Coefficient	*-0.033*	*-0.045*	*-0.033*	*-0.063*	*-0.076*	*-0.059*	*0.007*	*0.010*	*0.009*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of Observations	284,875	284,875	284,875	94,987	94,987	94,987	728,248	728,248	728,248
# of Exporters	6,625	6,625	6,625	2,209	2,209	2,209	16,936	16,936	16,936
# of Counties	399	399	399	324	324	324	500	500	500
# of Months	43	43	43	43	43	43	43	43	43
R-squared	0.328	0.394	0.369	0.433	0.522	0.470	0.347	0.388	0.380

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods are calculated by transforming the initial variable X with L(X) = log(X + 1). Data on ethnolinguistic composition is at the county level, and it comes from Ukrainian 2001 Census.

Table 12. Shares of Russian Managers vs. Russian Ethnicity of a County

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	Any Trade Activity	Log of Total Weight Traded	Log of Total Value Traded	Any Trade Activity	Log of Total Weight Traded	Log of Total Value Traded	Any Trade Activity	Log of Total Weight Traded	Log of Total Value Traded	Any Trade Activity	Log of Total Weight Traded	Log of Total Value Traded
	Difference-in-Differences						*Triple Difference*					
(Post Feb 2014) × (% Managers w/ Russian Last Names)	-0.006	-0.073	-0.098				0.025*	0.238	0.324*			
Calculated using the Surname Endings	(0.008)	(0.083)	(0.112)				(0.014)	(0.150)	(0.194)			
Standardized Coefficient	*-0.007*	*-0.008*	*-0.008*				*0.027*	*0.025*	*0.025*			
(Post Feb 2014) × (% Managers w/ Russian Last Names)				0.004	0.015	0.045				0.018	0.146	0.228
Calculated using a Bank of Surnames				(0.009)	(0.100)	(0.049)				(0.011)	(0.116)	(0.158)
Standardized Coefficient				*0.003*	*0.001*	*0.002*				*0.013*	*0.010*	*0.012*
(Post Feb 2014) × (% of Russian Ethnicity)	0.136***	1.692***	1.883***	0.130***	1.633***	1.798***	0.202***	2.333***	2.755***	0.176***	2.079***	2.422***
	(0.037)	(0.434)	(0.509)	(0.039)	(0.449)	(0.145)	(0.035)	(0.412)	(0.476)	(0.032)	(0.388)	(0.442)
Standardized Coefficient	*0.045*	*0.054*	*0.045*	*0.043*	*0.052*	*0.043*	*0.070*	*0.077*	*0.069*	*0.061*	*0.069*	*0.061*
(Post Feb 2014) × (% of Russian Ethnicity) ×							-0.200**	-1.963**	-2.670**	-0.111**	-1.092**	-1.528**
× (% Managers w/ Russian Last Names)							(0.078)	(0.815)	(1.084)	(0.050)	(0.484)	(0.688)
Standardized Coefficient							*-0.040*	*-0.038*	*-0.039*	*-0.022*	*-0.021*	*-0.022*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of Observations	475,181	475,181	475,181	475,181	475,181	475,181	475,181	475,181	475,181	475,181	475,181	475,181
# of Firms	10,302	10,302	10,302	10,302	10,302	10,302	10,302	10,302	10,302	10,302	10,302	10,302
# of Counties	404	404	404	404	404	404	404	404	404	404	404	404
# of Months	48	48	48	48	48	48	48	48	48	48	48	48
R-squared	0.429	0.508	0.476	0.429	0.508	0.476	0.429	0.508	0.476	0.429	0.508	0.476

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods are calculated by transforming the initial variable X with L(X) = log(X + 1). Last names in (1-3) and (7-9) are treated as Russian if they end in “ov”, “ova”, “ev”, “eva”, “in”, “ina”, “yov”, “yova”. In (4-6) and (10-12) we use a bank of last names that are considered Russian.

Table 13. Heterogeneity Effect across Regions.

VARIABLES	(1) Any Trade Activity	(2) Log of Total Weight Traded	(3) Log of Total Value Traded	(4) Any Trade Activity	(5) Log of Total Weight Traded	(6) Log of Total Value Traded	(7) Any Trade Activity	(8) Log of Total Weight Traded	(9) Log of Total Value Traded
	Without Kiev			*No Regions Close to Conflict Areas*			*No Western Ukraine*		
(Post Feb 2014) X (% of Russian Ethnicity)	0.096***	1.231***	1.335***	0.180***	2.357***	2.400***	0.071*	0.981**	1.060**
	(0.036)	(0.384)	(0.468)	(0.062)	(0.749)	(0.823)	(0.038)	(0.420)	(0.501)
Standardized Beta Coefficient	*0.032*	*0.039*	*0.032*	*0.059*	*0.075*	*0.058*	*0.024*	*0.031*	*0.026*
Year-Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
# of Observations	425,238	425,238	425,238	398,340	398,340	398,340	536,889	536,889	536,889
# of Firms	9,307	9,307	9,307	8,682	8,682	8,682	11,669	11,669	11,669
# of Counties	411	411	411	353	353	353	323	323	323
# of Months	48	48	48	48	48	48	48	48	48
R-squared	0.414	0.487	0.463	0.407	0.478	0.454	0.410	0.485	0.458

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are clustered at the county ("raion") level. The logs of total value and net weight of shipped goods are calculated by transforming the initial variable X with L(X) = log(X + 1). Data on ethnolinguistic composition is at the county level, and it comes from Ukrainian 2001 Census. This table tests whether our results are robust to regions-outliers. In columns (1-3) we restrict our analysis to areas without Kiev – the capital. In columns (4-6) we drop territories that are close enough to conflict territories - Dnipropetrovskaya, Zaporozhskaya, and Kharkovskaya oblasts.

Appendix: Additional Figures

Figure A1.

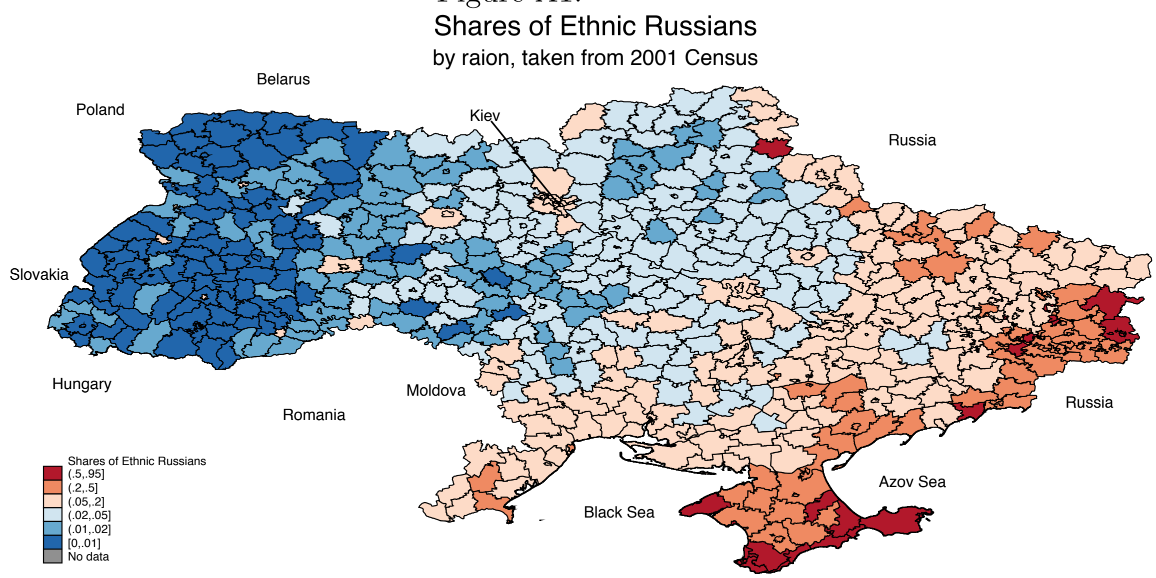


Figure A2.

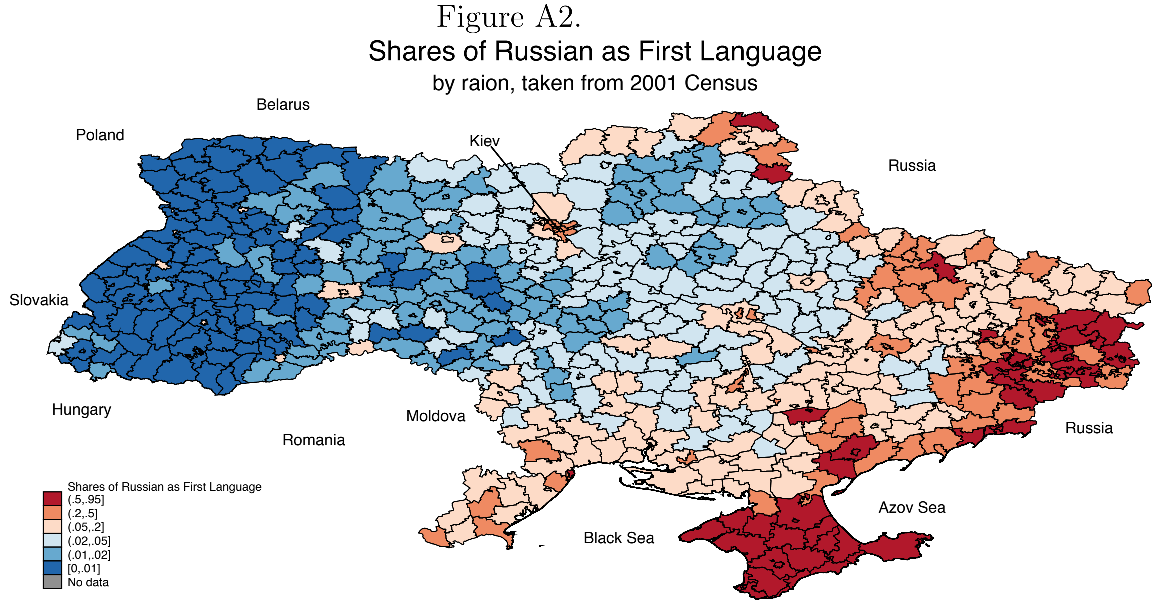


Figure A3.

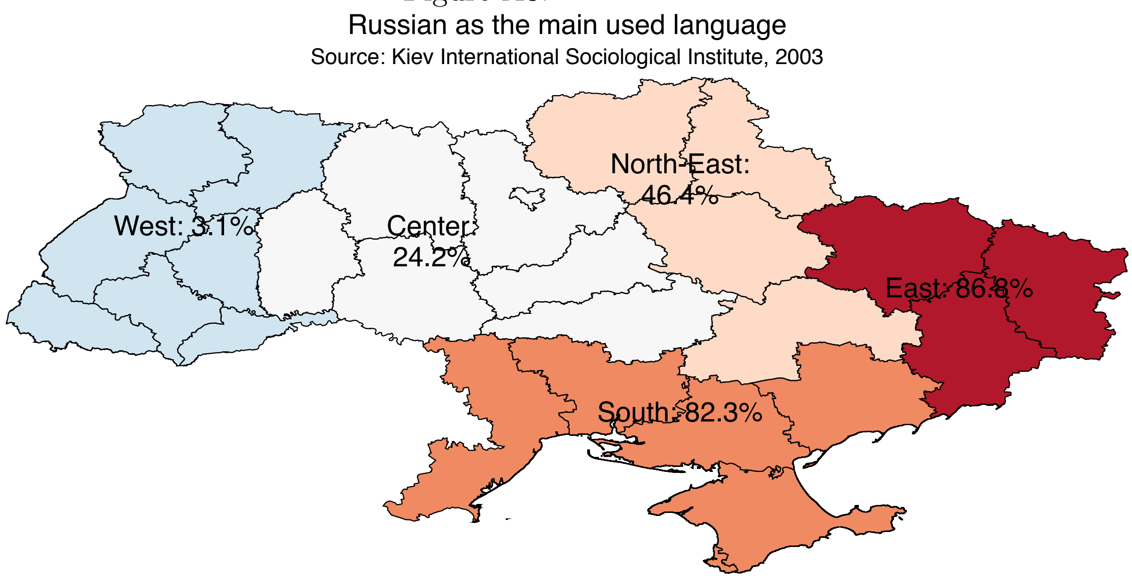


Figure A4. Parties with the Largest Vote Share, October 28, 2012.

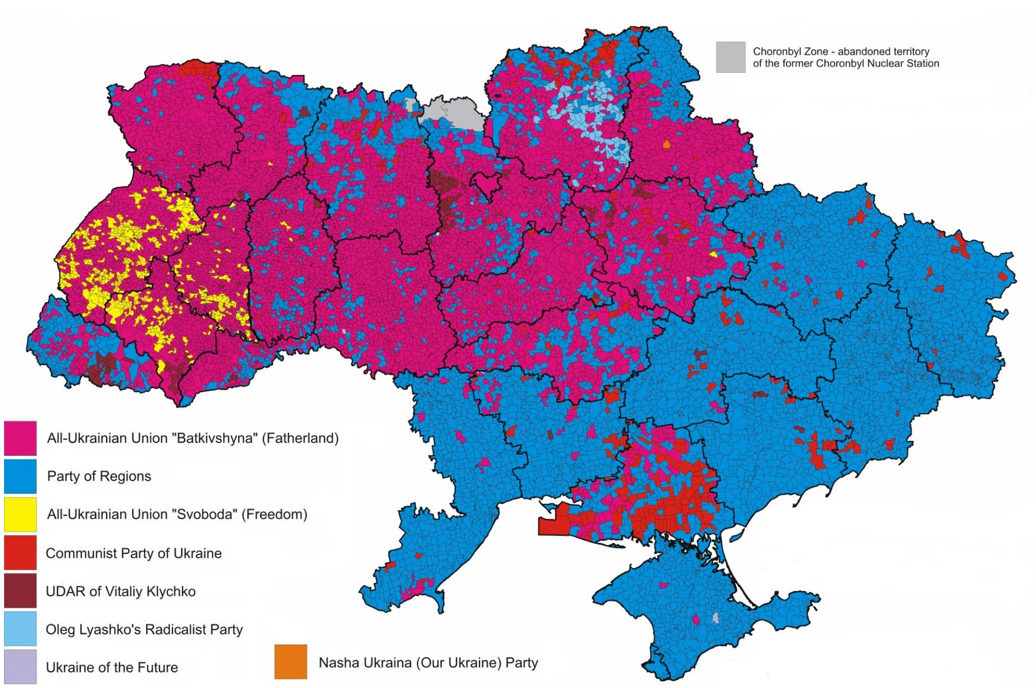


Figure A5. Presidential Candidates with the Largest Vote Share, December 26, 2004.

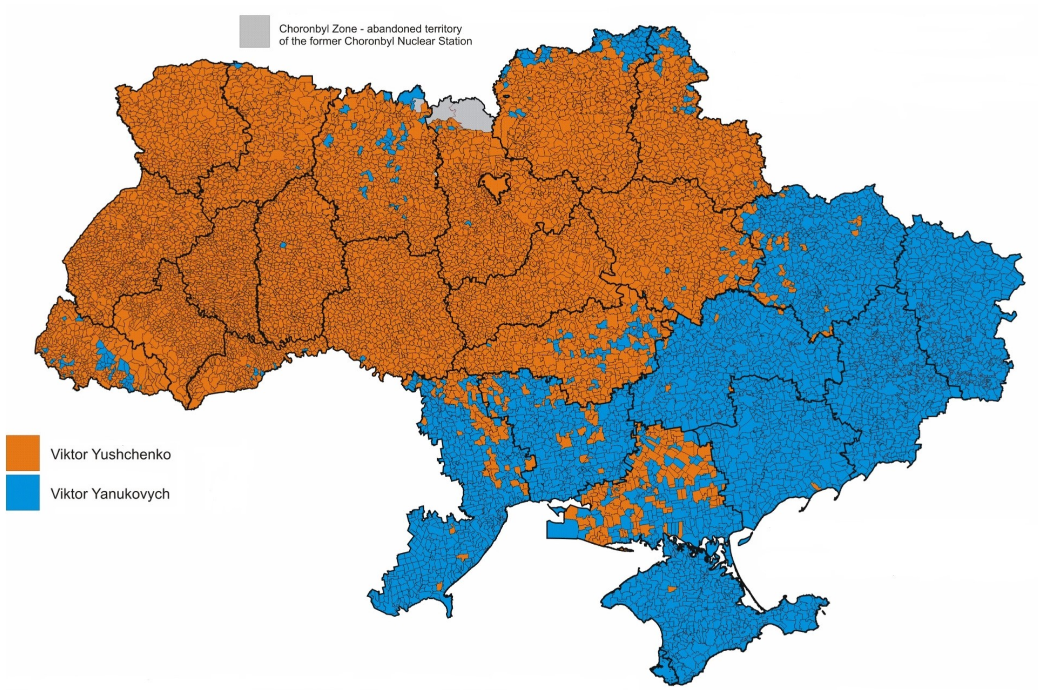


Figure A6. Prevalent language of social media accounts in VK.com.

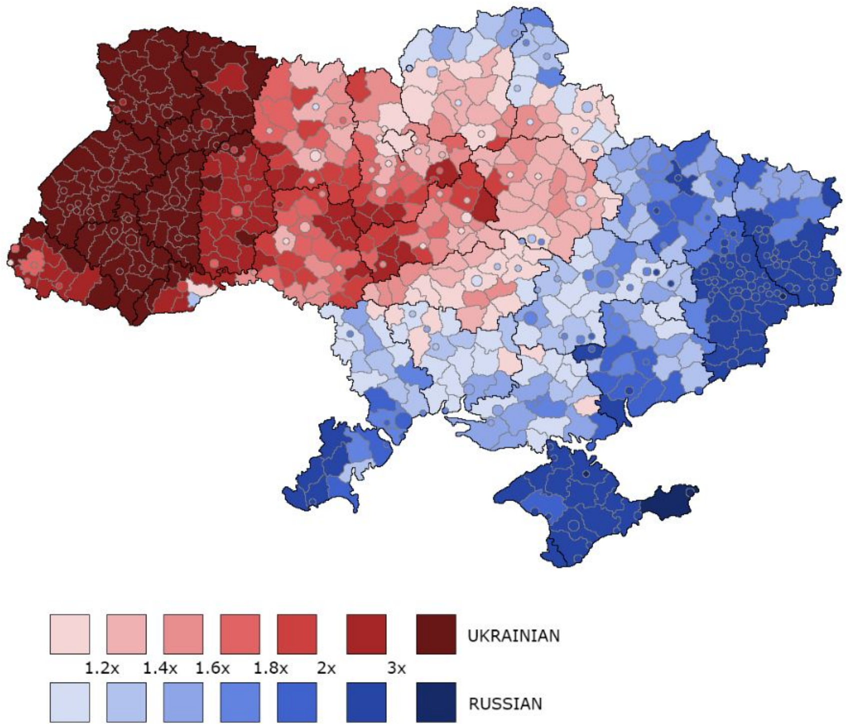
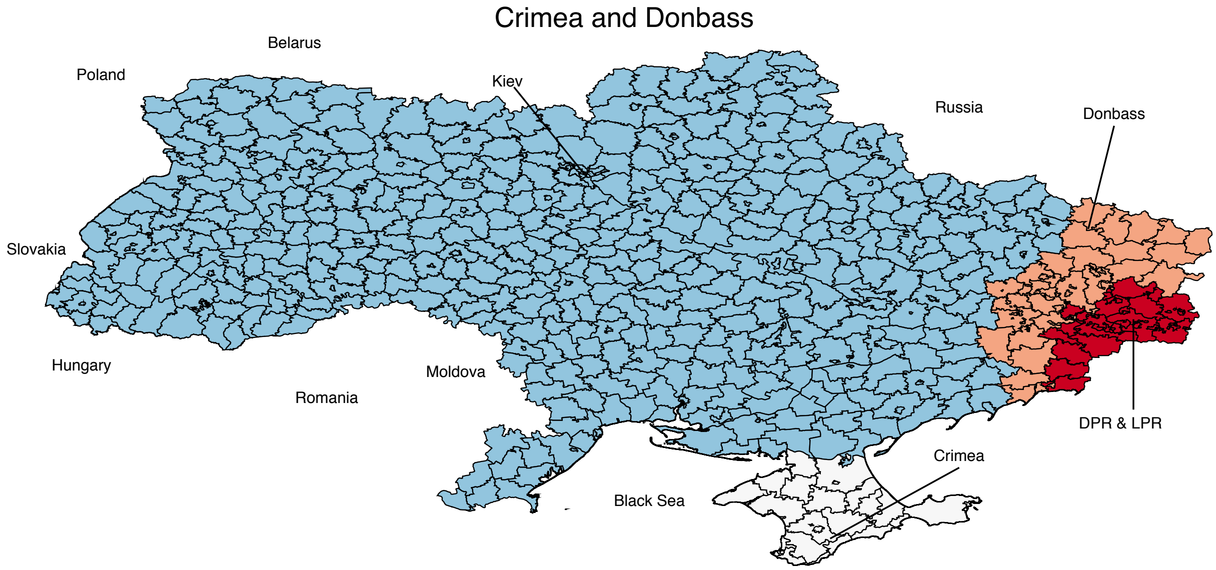

UKRAINIAN
1.2x
1.4x
1.6x
1.8x
2x
3x
RUSSIAN

Figure A7. Conflict areas.



Bibliography

Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the Basque Country. The American Economic Review, 93(1):113–132, 2003.

Francesco Amodio and Michele Di Maio. Making do with what you have: Conflict, input misallocation, and firm performance. The Economic Journal, 2017.

Stanislav Anatolyev and Grigory Kosenok. Tests in contingency tables as regression tests. Economics letters, 105(2):189–192, 2009.

Joshua D Angrist and Jörn-Steffen Pischke. Mostly harmless econometrics: An empiricist's companion. Princeton university press, 2008.

Leandro Arozamena and Federico Weinschelbaum. The effect of corruption on bidding behavior in first-price auctions. European Economic Review, 53(6): 645–657, 2009.

John Asker. Bidding rings. The New Palgrave Dictionary of Economics, 4, 2010a.

John Asker. A study of the internal organization of a bidding cartel. The American Economic Review, 100(3):724–762, 2010b.

Susan Athey, Jonathan Levin, and Enrique Seira. Comparing open and sealed bid auctions: Evidence from timber auctions*. Quarterly Journal of Economics, 126(1), 2011.

Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. Economic Theory, 26(2):445–469, 2005.

Patrick Bajari and Ali Hortacsu. The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. RAND Journal of Economics, pages 329–355, 2003.

Patrick Bajari and Lixin Ye. Deciding between competition and collusion. Review of Economics and Statistics, 85(4):971–989, 2003.

Patrick Bajari, Stephanie Houghton, and Steven Tadelis. Bidding for incomplete contracts: An empirical analysis of adaptation costs. The American Economic Review, 104(4):1288–1319, 2014.

Laura H Baldwin, Robert C Marshall, and Jean-Francois Richard. Bidder collusion at forest service timber sales. Journal of Political Economy, 105(4):657–699, 1997.

Audinga Baltrunaite. Political finance reform and public procurement: Evidence from lithuania. 2016.

Oriana Bandiera, Andrea Prat, and Tommaso Valletti. Active and passive waste in government spending: Evidence from a policy experiment. The American Economic Review, pages 1278–1308, 2009.

Abhijit Banerjee, Esther Duflo, Clement Imbert, Santhosh Mathew, and Rohini Pande. Can e-governance reduce capture of public programs? experimental evidence from a financial reform of indias employment guarantee. economics. mit. edu/files/10557 (accessed July 1, 2015), 2014.

Michal Bauer, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts. Can war foster cooperation? Journal of Economic Perspectives, 30(3):249–74, 2016.

Timothy Besley and Marta Reynal-Querol. The legacy of historical conflict: Evidence from Africa. American Political Science Review, 108(2):319–336, 2014.

Michael Carlos Best, Jonas Hjort, and David Szakonyi. Individuals and organizations as sources of state effectiveness, and consequences for policy design. 2017.

Christopher Blattman and Jeannie Annan. The consequences of child soldiering. The review of economics and statistics, 92(4):882–898, 2010.

Christopher Blattman and Edward Miguel. Civil war. Journal of Economic Literature, pages 3–57, 2010.

Laurent Bordes and Pierre Vandekerkhove. Semiparametric two-component mixture model with a known component: an asymptotically normal estimator. Mathematical Methods of Statistics, 19(1):22–41, 2010.

Laurent Bordes, Stéphane Mottelet, and Pierre Vandekerkhove. Semiparametric estimation of a two-component mixture model. The Annals of Statistics, pages 1204–1232, 2006.

Roberto Burguet and Martin K Perry. Bribery and favoritism by auctioneers in sealed-bid auctions. The BE Journal of Theoretical Economics, 7(1), 2007.

Hongbin Cai, J Vernon Henderson, and Qinghua Zhang. China's land market auctions: evidence of corruption? The Rand Journal of Economics, 44(3): 488–521, 2013.

Michael Callen and James D Long. Institutional corruption and election fraud: Evidence from a field experiment in afghanistan. American Economic Review, 105(1):354–81, 2015.

Michael Callen, Mohammad Isaqzadeh, James D Long, and Charles Sprenger. Violence and risk preference: Experimental evidence from Afghanistan. The American Economic Review, 104(1):123–148, 2014.

Alessandra Cassar, Pauline Grosjean, and Sam Whitt. Legacies of violence: trust and market development. Journal of Economic Growth, 18(3):285–318, 2013.

Rubiana Chamarbagwala and Hilcías E Morán. The human capital consequences of civil war: Evidence from guatemala. Journal of Development Economics, 94 (1):41–61, 2011.

Olivier Compte, Ariane Lambert-Mogiliansky, and Thierry Verdier. Corruption and competition in procurement auctions. Rand Journal of Economics, pages 1–15, 2005.

Timothy G Conley and Francesco Decarolis. Detecting bidders groups in collusive auctions. Technical report, 2011.

Decio Coviello and Stefano Gagliarducci. Tenure in office and public procurement. Available at SSRN 2765159, 2016.

Melissa Dell and Pablo Querubin. Nation building through foreign intervention: Evidence from discontinuities in military strategies. The Quarterly Journal of Economics, 1:64, 2017.

Rafael Di Tella and Ernesto Schargrodsky. The role of wages and auditing during a crackdown on corruption in the city of buenos aires*. Journal of Law and Economics, 46(1):269–292, 2003.

Ron Edwards, Anne-Marie Gut, and Felix Mavondo. Buyer animosity in business to business markets: Evidence from the French nuclear tests. Industrial Marketing Management, 36(4):483–492, 2007.

Nasr G Elbahnasawy. E-government, internet adoption, and corruption: An empirical investigation. World Development, 57:114–126, 2014.

Claudio Ferraz and Frederico Finan. Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. Quarterly Journal of Economics, 123(2), 2008.

Claudio Ferraz and Frederico Finan. Electoral accountability and corruption: Evidence from the audits of local governments. The American Economic Review, 101(4):1274–1311, 2011.

Aleksei Fedorovich Filippov. The existence of solutions of generalized differential equations. Mathematical Notes, 10(3):608–611, 1971.

Vasiliki Fouka and Hans-Joachim Voth. Reprisals remembered: German-Greek conflict and car sales during the Euro crisis. 2016.

Reuven Glick and Alan M Taylor. Collateral damage: Trade disruption and the economic impact of war. The Review of Economics and Statistics, 92(1):102–127, 2010.

Esra Guerakar and Erik Meyersson. State discretion, political connections and public procurement: Evidence from turkey. 2016.

Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. Econometrica, 68(3):525–574, 2000.

Massimo Guidolin and Eliana La Ferrara. Diamonds are forever, wars are not: Is conflict bad for private firms? American Economic Review, 97(5):1978–1993, 2007.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Cultural biases in economic exchange? The Quarterly Journal of Economics, 124(3):1095–1131, 2009.

Philip Haile, Kenneth Hendricks, Robert Porter, and Toshi Onuma. Testing competition in us offshore oil and gas lease bidding. Technical report, Working Paper, 2012.

Philip A Haile, Han Hong, and Matthew Shum. Nonparametric tests for common values at first-price sealed-bid auctions. Technical report, National Bureau of Economic Research, 2003.

Kilian Heilmann. Does political conflict hurt trade? Evidence from consumer boycotts. Journal of International Economics, 99:179–191, 2016.

Kenneth Hendricks and Robert H Porter. An empirical study of an auction with asymmetric information. The American Economic Review, pages 865–883, 1988.

Kenneth Hendricks, Ilke Onur, and Thomas Wiseman. Last-minute bidding in sequential auctions with unobserved, stochastic entry. Review of Industrial Organization, 40(1):1–19, 2012.

Daniel Hohmann and Hajo Holzmann. Semiparametric location mixtures with distinct components. Statistics, 47(2):348–362, 2013.

Hugo Hopenhayn and Maryam Saeedi. Dynamic bidding in second price auction. Technical report, Working paper, The Ohio State University, 2015.

Alex Imas, Michael Kuhn, and Vera Mironova. A history of violence: Field evidence on trauma, discounting and present bias. 2015.

Allan T Ingraham. A test for collusion between a bidder and an auctioneer in sealed-bid auctions. Contributions in Economic Analysis & Policy, 4(1), 2005.

Rieko Ishii. Favor exchange in collusion: Empirical study of repeated procurement auctions in japan. International Journal of Industrial Organization, 27(2):137–144, 2009.

Kei Kawai and Jun Nakabayashi. Detecting large-scale collusion in procurement auctions. Available at SSRN 2467175, 2014.

Charles Kenny. Construction, corruption, and developing countries. World Bank Policy Research Working Paper, (4271), 2007.

Christopher Ksoll, Rocco Macchiavello, and Ameet Morjaria. Guns and roses: Flower exports and electoral violence in Kenya. 2014.

Yvan Lengwiler and Elmar Wolfstetter. Corruption in procurement auctions. Available at SSRN 874705, 2006.

Yvan Lengwiler and Elmar Wolfstetter. Auctions and corruption: An analysis of bid rigging by a corrupt auctioneer. Journal of Economic Dynamics and Control, 34(10):1872–1892, 2010.

Gianmarco Leon. Civil conflict and human capital accumulation the long-term effects of political violence in perú. Journal of Human Resources, 47(4):991–1022, 2012.

Andrei A Levchenko. Institutional quality and international trade. The Review of Economic Studies, 74(3):791–819, 2007.

Sean Lewis-Faupel, Yusuf Neggers, Benjamin A Olken, and Rohini Pande. Can electronic procurement improve infrastructure provision? evidence from public works in india and indonesia. Technical report, National Bureau of Economic Research, 2014.

Huagang Li and Guofu Tan. Hidden reserve prices with risk averse bidders. 2000.

Zijun Luo and Yonghong Zhou. Gainers and losers of political instability: Evidence from the anti-Japanese demonstration in China. 2016.

Sharon Maccini and Dean Yang. Under the weather: Health, schooling, and economic consequences of early-life rainfall. American Economic Review, 99 (3):1006–26, June 2009. doi: 10.1257/aer.99.3.1006. URL `http://www.aeaweb.org/articles?id=10.1257/aer.99.3.1006`.

Robert C Marshall and Leslie M Marx. Bidder collusion. Journal of Economic Theory, 133(1):374–402, 2007.

Robert C Marshall and Leslie M Marx. The Economics of Collusion: Cartels and Bidding Rings. Mit Press, 2012.

Philippe Martin, Thierry Mayer, and Mathias Thoenig. Civil wars and international trade. Journal of the European Economic Association, 6(2-3):541–550, 2008.

Marc J Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71(6):1695–1725, 2003.

Flavio M Menezes and Paulo Klinger Monteiro. Corruption and auctions. Journal of Mathematical Economics, 42(1):97–108, 2006.

Edward Miguel and Gerard Roland. The long-run impact of bombing Vietnam. Journal of development Economics, 96(1):1–15, 2011.

Maxim Mironov and Ekaterina Zhuravskaya. Corruption in procurement: Evidence from financial transactions data. Available at SSRN 1946806, 2014.

Karthik Muralidharan, Paul Niehaus, and Sandip Sukhtankar. Building state capacity: Evidence from biometric smartcards in india. Technical report, National Bureau of Economic Research, 2014.

Roger B Myerson. Optimal auction design. Mathematics of Operations Research, 6(1):58–73, 1981.

Nathan Nunn. Relationship-specificity, incomplete contracts, and the pattern of trade. The Quarterly Journal of Economics, 122(2):569–600, 2007.

Axel Ockenfels and Alvin E Roth. The timing of bids in internet auctions: Market design, bidder behavior, and artificial agents. AI magazine, 23(3):79, 2002.

Axel Ockenfels and Alvin E Roth. Late and multiple bidding in second price internet auctions: Theory and evidence concerning different rules for ending an auction. Games and Economic Behavior, 55(2):297–320, 2006.

Benjamin A Olken. Corruption and the costs of redistribution: Micro evidence from indonesia. Journal of Public Economics, 90(4):853–870, 2006.

Benjamin A Olken. Monitoring corruption: evidence from a field experiment in indonesia. Journal of Political Economy, 115(2):200–249, 2007.

Benjamin A Olken and Rohini Pande. Corruption in developing countries. Annu. Rev. Econ., 4(1):479–509, 2012.

Sonal S Pandya and Rajkumar Venkatesan. French roast: consumer response to international conflictevidence from supermarket scanner data. Review of Economics and Statistics, 98(1):42–56, 2016.

Martin Pesendorfer. A study of collusion in first-price auctions. The Review of Economic Studies, 67(3):381–411, 2000.

Robert H Porter. Detecting collusion. Review of Industrial Organization, 26(2): 147–167, 2005.

Robert H Porter and J Douglas Zona. Detection of bid rigging in procurement auctions. Journal of Political Economy, pages 518–538, 1993.

Robert H Porter and J Douglas Zona. Ohio school milk markets: An analysis of bidding. The Rand Journal of Economics, pages 263–288, 1999.

James E Rauch. Networks versus markets in international trade. Journal of international Economics, 48(1):7–35, 1999.

James E Rauch and Vitor Trindade. Ethnic Chinese networks in international trade. The Review of Economics and Statistics, 84(1):116–130, 2002.

Ritva Reinikka and Jakob Svensson. Fighting corruption to improve schooling: Evidence from a newspaper campaign in uganda. Journal of the European Economic Association, 3(2-3):259–267, 2005.

Marc S Robinson. Collusion and the choice of auction. The RAND Journal of Economics, pages 141–145, 1985.

Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. Seeds of distrust: Conflict in Uganda. Journal of Economic Growth, 18(3):217–252, 2013a.

Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. War signals: A theory of trade, trust, and conflict. Review of Economic Studies, 80(3):1114–1147, 2013b.

Michael H Rothkopf, Thomas J Teisberg, and Edward P Kahn. Why are vickrey auctions rare? Journal of Political Economy, 98(1):94–109, 1990.

David Schoenherr. Political connections and allocative distortions. In 27th Australasian Finance and Banking Conference, 2014.

Karl Schurter. Identification and inference in first-price auctions with collusion, 2016.

Olga Shemyakina. The effect of armed conflict on accumulation of schooling: Results from tajikistan. Journal of Development Economics, 95(2):186–200, 2011.

Andrei Shleifer and Robert W Vishny. Corruption. The quarterly journal of economics, 108(3):599–617, 1993.

Ayumu Tanaka, Banri Ito, and Ryuhei Wakasugi. How do exporters respond to exogenous shocks: Evidence from Japanese firm-level data. 2017.

Michael D Whinston. Lectures on antitrust economics. MIT Press Books, 1, 2008.