LUISS GUIDO CARLI

ESSAYS IN PUBLIC ECONOMICS

A DISSERTATION SUBMITTED TO
THE FACULTY OF LUISS GUIDO CARLI
DEPARTMENT OF ECONOMICS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY

DARIA BOTTAN

ROME, ITALY
DECEMBER 2015

Introduction

This dissertation consists of two chapters. Both chapters use unique data to answer questions in public economics.

In Chapter 1, I study optimal law enforcement in presence of corruption using unique data on bribes given to traffic police in Russia. Corruption in law enforcement is thought to decrease the expected punishment for a violation of law and hence decreases deterrence of crimes. I examine whether it is possible to offset the effect of corruption and restore deterrence by increasing the magnitude of sanctions. Whether an increase in punishment will restore deterrence depends on how sensitive the bribes are to changes in sanctions. I estimate sensitivity of bribes to changes in different type of sanctions: monetary and nonmonetary, using difference-in-differences method. I find that bribes are more sensitive to changes in nonmonetary rather than monetary sanctions. In presence of corruption the socially costly monetary sanction is transformed into a bribe a socially costless monetary transfer. This further increases the attractiveness of nonmonetary sanctions in presence of corruption. I also discuss red tape as a nonmonetary sanction and its potential benefits in comparison to other nonmonetary sanctions.

In Chapter 2, I study economic returns to infrastructure using historic data from the city of Chicago in 1830-1930. The city experienced rapid growth during this time making access to clean water and sewage a pressing issue. Using panel data with fixed effects and an instrumental variables approach, I estimate the causal effect of infrastructure access (water and sewage) on land values. I construct an instrument for access to infrastructure by taking advantage of a rule by which a suburb could join Chicago. One of the main motivations for joining Chicago was the provision of water and sewage. Using both panel data with fixed effects and 2SLS, I find that a that access to water positively affects land values; while access to sewage does not have a significant effect. The estimated coefficients are not statistically different between both approaches. Results suggest that access to clean water is a valuable amenity for both individuals and firms.

Acknowledgments

I am very thankful to my advisors:

- Professor Daniel Bernhardt for all the insightful comments, exciting discussions, trust in my research potential and all the support.
- Professor Liliane Giardino-Carlinger for her patience, very thoughtful and useful comments and constant support and encouragement through the process.
- Professor David Albouy for motivating me; asking me hard and interesting questions and pushing me to do a better job; for making me see the big picture and helping me discover the topics I am really interested in.

I could not possibly imagine having better advisors!

I would also like to thank Professor Mark Borgschulte for our numerous talks and taking time to understand my issues in depth; thoughtful, useful and interesting comments and being approachable. I thank Professor Maria Yudkevich for making me interested in research in the first place. I thank Professor Leonid Polishchuk for showing me by example what a good scholar is, teaching me to think like a researcher and support during the challenging times.

I thank Professors Richard Akresh, Fabiano Schivardi, Evan Starr, Yann Algian, Ben Marx, Anil Bera, Werner Guth, Sergei Kovbasyuk, Andrea Pozzifor for multiple discussions and suggestions.

I am also grateful to Professor Martin Perry and Professor Pierpaolo Benigno for supporting my idea of splitting my time during the Ph.D. between LUISS Guido Crali and University of Illinois at Urbana-Champaign. I thank other Professors and staff of UIUC and LUISS Guido Carli for their comments and providing me with an opportunity to be a Teaching Assistant which provided a financial support for my Ph.D.

I would like to thank the students in LUISS Guido Carli and UIUC for their comments, interest in my research and interesting discussions.

I am very thankful to my friends who helped me to preserve my mental health and not to give up during the process of writing this dissertation. I am also grateful to my parents and sister and the family of my husband for making me feel loved and supported, listening to my troubles, happy for my achieves and having trust in my potential.

This dissertation could not be written without love, patience, advice, support and jokes of my wonderful husband Nicolas Bottan.

Contents

1 Optimal Law Enforcement In Presence Of Corruption: Evidence From Bribes To Traffic Police In Russia.

1.1 Introduction

Economic models of crime suggest that potential offenders respond to the expected value of punishment (Becker, 1968, Polinsky and Shavell, 1984). There is empirical evidence from developed countries supporting this idea: using the data on red light running in U.S. and Israel Bar-Ilan and Sacerdote (2004) show that deterrence increases when fines increase. The empirical evidence from the developing countries is limited.

Corruption is a central issue in law enforcement in developing countries. Corruption in law enforcement can take different forms. The main two forms are organized crime (i.e. structures similar to those described in the God Father movie) and casual corruption (i.e. one time interactions between police and offender), see Bowles and Garoupa (1997) for a more detailed discussion. In this paper I focus on casual corruption. For example, a detective hides convicting evidence in exchange for a bribe; or a traffic police officer does not report running red light in exchange for a payment. Corruption decreases the expected punishment and, hence, deterrence of crimes (Bowles and Garoupa, 1997, Polinsky and Shavell, 2001): potential offenders will discount the punishment if they know that they can avoid it by paying a lower bribe. It might be possible to offset the corruption deterrence diluting effects by increasing the magnitude of the punishment. Whether increasing the punishment will restore deterrence depends on how sensitive bribes are to changes in sanctions. I estimate the sensitivity of bribes to changes in different type of sanctions: monetary and nonmonetary.

I study the effect of a change in sanctions on the average size of the bribe using data on self-reported bribes paid to traffic police in Russia. The data was obtained from a website that encourages drivers to submit information on bribes they paid to avoid an official sanction. To estimate the causal effect of changes in sanction on the bribe, I exploit the variation in changes in sanctions to different types of violations. I use the types of violations for which sanctions did not change as a counterfactual. I find that bribes are more sensitive to changes in nonmonetary rather than monetary sanctions: elasticity of 0.63 vs. elasticity of 0.05. This suggests that if the government wanted to offset the deterrence-diluting effect of corruption, a nonmonetary sanction might be the more effective.

A typical example of nonmonetary sanction is imprisonment. For the case of traffic violations one can think of other examples, such as driver's license revocation. In this paper I emphasize that bureaucratic costs, for example creating long lines to pay a fine or completing numerous forms, can also be considered a nonmonetary sanction. Adding bureaucratic costs to a fine can increase the bargaining power of the enforcement agent and hence result in a higher bribe and higher deterrence. If the expected bribe is higher, smaller number of potential offenders will find it worthwhile to violate the law. In absence of corruption nonmonetary sanctions create social costs for the government (for example cost of maintaining a prison) and for the offenders (private cost of time lost in prison). If all offenders are corrupt, the nonmonetary sanction is threatened but not implemented in practice: socially costly sanction is substituted by a bribe (see Garoupa and Klerman (2004) for a formal proof). This makes nonmonetary sanctions more attractive in presence of corruption in law enforcement.

The empirical investigation of optimal law enforcement is important from a policy implication point of view. Many developing countries suffer from bribery in various areas of life: traffic violations, violations in construction, criminal acts. Bribery decreases deterrence resulting in high death rates from traffic violations, inefficiently and unsafely designed cities and criminals avoiding prosecution. Understanding which tools work best for offsetting deterrence diluting effects of corruption can help to avoid substantial social losses, including human life. To my knowledge this is the first study providing empirical evidence on this issue in the developing countries.

The structure of paper is as follows. Section 1.2 provides information on bribery in Russia and describes the data. Section 1.3 provides the theoretical framework. Section 1.4 discusses the identification strategy. Section 1.5 provides results and section 1.6 studies their robustness. Section 1.7 discuses how bureaucratic costs can help mitigate the deterrence-diluting effect of corruption. Section 1.8 concludes.

1.2 Background and Data

1.2.1 Corruption and Traffic in Russia

Russia is one of the most corrupt countries in the world: In 2014 Transparency international ranked Russia 136 out of 174 countries in the corruption index (where countries with lowest levels of corruption are ranked first). Corruption is Russia is present in various aspects of life. When asked what are the most corrupt institutions respondents name: municipal and federal government, police and traffic police, public medical and educational personnel[1]. Bribery is a socially accepted practice in Russia: 41% of respondents state that they do not disapprove those who give bribes[2]. Bribing traffic police is a particularly common form of corruption: almost every second driver admits to having given a bribe to traffic police at leas once.[3]

Traffic fatalities and traffic congestion are acute issues in Russia. Fatalities in traffic accidents are 18.6 per 100,000 inhabitants, this is higher than in Pakistan (17.4) and

[1] Press release number 2437 by VCIOM, 23 of October 2013

[2] FOMnibus survey "Problems of corruption in Russia" of 1500 respondents in 1000 places, 2011

[3] Data from Russian Public Opinion Research Center (VCIOM). 2010. "The work of Russian traffic police: Evaluation of Russians"

Burundi (13.2).[4] In 2012, Moscow was ranked as the most congested city in Europe.[5]The fact that drivers systematically violate traffic rules contributes to the severe traffic situation. For example, over 720 drivers parked on a tram-line in Moscow during the first half of 2010.[6] The amount of time loss, human life loss and the scope of traffic rule violations suggests that the level of deterrence in Russia is smaller than the optimal. The federal government routinely increases monetary sanctions for traffic violations in an attempt to increase deterrence and mitigate the traffic problems.

The typical process of giving a bribe to traffic police officer is as follows. A driver violates a traffic rule and gets pulled over (traffic police usually does not frame individuals who did not violate any law). The police officer announces the corresponding official sanction to an offender. The driver can either choose to accept the official sanction or to offer a bribe. If a driver accepts a ticket, it has to be paid: the possibility of bribing a government official later is limited. If instead the driver chooses to bribe the police officer, she usually does so by putting the money with the documents requested by the officer. The bribe can either be accepted or rejected by the police officer. Paying a bribe is potentially associated with additional risks for both offenders and traffic police officers. According to Russian laws, giving or receiving a bribe is punishable by a fine and/or imprisonment for a period of up to 8 years. However these laws have rarely been implemented in practice.

If a violator has to pay a monetary fine he often has to bear additional bureaucratic (red tape) costs with it. Drivers can pay a fine by either visiting a bank or using a credit/debit card online.[7] Regardless of the means of payment, it might take a long time for the information about the payment to get processed and in many cases the information gets lost.[8] This creates two additional problems. Firstly, a driver could be summoned to court.

[4]According to data from WHO Global status report on road safety 2013: supporting a decade of action.

[5]according to TomTom European Congestion Index Anual Report 2012

[6]Acoarding to 2010. "Drivers from Moscow will pay 6 mln rubles in fines for parking on the tram line". Driver from St. Petersburg, July 23. http://www.spbvoditel.ru/2010/08/23/011/

[7]Paying online has been available starting from 2012

[8]A driver can bring a copy of his payment receipt to the traffic police office, to make sure the information

The current punishment for failure to pay a fine on time is up to 15 days in prison or paying twice the amount of the fine. Secondly, a person who has unpaid fines can be denied of exiting the country (traveling abroad). The probability of having the payment information lost is not negligible. During 2012, about half a million individuals were summoned to court for an unpaid fine even though they actually paid it on time[9].

1.2.2 Data source

Data on bribes was collected from Skolkodal.ru[10] - the name of the website translates from Russian as "How much did you give". The website provides information to drivers on: new traffic rules, typical locations of traffic police, tips on how to get around police, etc. Among other things, the website encourages drivers to submit information about the bribes they paid.

Drivers submit the size of the bribe, date (day, month, year), city and type of traffic violation (chosen from one of 11 types of traffic violations)[11]. No information on the identity of individuals which submitted bribes is available.

Corruption is illegal by nature so individuals might tend to misreport it out of shame or fear of punishment. For example, one can under-report the size of the bribe to make it look less important. The data in this study is self reported. The evidence from psychological literature suggests that individuals tend to be more willing to report sensitive behaviors when they are not directly observed by a surveyor (for example when the questionnaire is self-administrated, see Turner et al. (1998), Tourangeau and Yan (2007)). In essence it is so because individuals feel less embarrassed and have smaller fear of prosecution when

reaches the traffic police. However it requires a wait and potentially long traveling to a traffic police office

[9] Sichev,Vasiliy.2013."They don't let us sleep".Lenta.ru,May 20. http://lenta.ru/articles/2013/05/20/fines/

[10] www.http://skolkodal.ru

[11] The website only shows the information about the bribes as an average of a bribe for a day per type of traffic violation. I cannot distinguish whether there was one or more bribes submitted per violation per day. I treat the average reported bribe per day as one observation.

they self-report data without a presence of surveyor[12]. An anonymous survey about bribes on the internet provides precisely the kind of setting in which individuals are less likely to misreport.

It is possible that the composition of the offenders and those individuals who submit bribes changes when the sanctions change. The main limitation of the data is that I do not know any individual characteristics of the offenders and hence can not control for them directly. I conciser the selection problem in more detail in the robustness section. However the lack of individual characteristics might decrease the incentives to misreport the bribes.

Individuals may choose to devote time and make information about their bribes available as a way of contributing to a public good. Online-based voluntarily contribution to a public good is common practice, with Wikipedia as the most famous example. There are other websites in Russia that collect information on bribes paid to traffic police, one of them (vashamashina.ru) constructs a confidence interval for a bribe which will be accepted, so that drivers do not overpay. The website used as a data source in this paper states that information on bribes is collected to find the extent of the illegal incomes extracted by traffic police.

I restrict my analysis to Moscow and St. Petersburg since most of the bribes were submitted from these 2 cities. I study the period from January 2011 to August 2013. In September 2013 there were a lot of changes in traffic rules and sanctions which makes finding a valid control group challenging and makes it hard to pin down the exact factors which affect the size bribes. I relay on one website as a data source instead of combining different sources because this website has the largest number of observations and because other sources often do not include comparable types of traffic violations.

[12]See Sequeira (2012) for a discussion of various approaches of measuring corruption in the field.

1.2.3 Descriptive Statistics

Table 1 provides summary statistics of bribes by violation type. Second column shows the official sanction as of January 2012. The bribes and sanctions are presented in Russian currency ruble. One dollar equaled approximately 30 rubles during the period of observation[13]. Most of the types of traffic violations can only result in a monetary sanction (fine) regardless of the number of times the rule is violated by one driver. There is no system of accumulation of fine points in Russia. The sample includes two types of traffic violations which can result in nonmonetary sanction: driver's license revocation. The mean bribe for these types of violations (illegal passing and speeding by more then 60 kilometers per hour) is more then 3 times larger then the average bribes for other types of traffic violations.

Table 1 shows that for most of the types of traffic violations average bribe exceeds official monetary sanction. This observation is not driven by few very high bribes: the median bribe is also larger then the official fine for most of the types of traffic violations.

At first glance it seems unreasonable that individual offer higher payment in bribes than they are required to pay by law. There are several reasons why bribes may be higher than the fine. One potential explanation is that offenders what to avoid having a record of law violation because such record can increase their insurance payment.The number of traffic violations does not affect the price of insurance in Russia, unless a violation resulted in an accident. Hence potential increase in insurance premium is not a reason for the bribe to exceed the official fine. My preferred explanation is that an official fine includes bureaucratic costs in addition to direct monetary costs. Bureaucratic costs include time, money and stress related to paying an official fine; they appear to be quite substantial in Russia as described in background session. When offenders pay bribes they pay to avoid the official sanction and the bureaucratic costs associated with it, that is why they may be

[13]The average salary in 2012 was around 46,000 rubles in Moscow and around 32,000 rubles in St. Petersburg

willing to pay higher bribes than the monetary fines.

The average size of the bribe can depend on many factors: bureaucratic costs, custom, bargaining power of the police officer (it might be easier to extract higher bribes for more dangerous violations). The theoretical framework provides some thoughts regarding it.

1.3 Theoretical Framework

1.3.1 General Framework

The economic literature on optimal law enforcement originated with the seminal article of Becker (1968). A potential offender commits a harmful act only if her private benefits of committing the act outweigh the expected costs. It is typically assumed that there is a distribution of gains from committing a harmful act across potential offenders. With some probability an offender can be caught and sanctioned. In the theoretical literature on optimal law enforcement the social welfare is usually considered to be the sum of all individual expected utilities minus the harm done, minus the costs of law enforcement (see Polinsky and Shavell (2007) for a detailed review of the theoretical literate on the topic). The social planner is maximizing the social welfare by choosing the probability of detection, magnitude of sanctions and their type (monetary or nonmonetary).

The typical consensus in the literature is that monetary sanctions (fines) should be used to the full extend before the nonmonetary sanctions (like imprisonment) are used (Polinsky and Shavell (1984), Shavell (1987)). This is so because monetary sanctions are socially cheaper. A fine is a monetary transfer and hence is typically considered socially costless: the offender paying the fine looses utility but this loss is offset by an increase in utility of the fine recipient. While imprisonment imposes costs both on the enforcement system (price of maintaining and individual in jail) and individuals (private cost of being in jail).

Another example of a nonmonetary sanction is a revocation of driver's license or a restaurant license. However, these are not the only possible forms of nonmonetary sanction. Bureaucratic costs associated with the punishment - for example waiting in line to pay a fine- are also socially costly nonmonetary sanctions.

1.3.2 Corruption In Law Enforcement

The main conclusion of the theoretical literature is that corruption dilutes deterrence. This was first pointed out by Becker and Stigler (1974) and further studied by Polinsky and Shavell (2001) and Bowles and Garoupa (1997). Corruption dilutes deterrence because it decreases the expected punishment of the offender, he can now pay a bribe smaller than the official sanction. These models usually assumes that the bribe is proportional to the official sanction for the offense, for example Bribe=αSanction. The maximum bribe is equal to the sanction itself because otherwise the offenders would be better off by paying the official sanction instead of the bribe. The size of the bribe is determined by alpha, that represents the bargaining power of the enforcement agent versus the bargaining power of the offender. α can vary depending on the type of the sanction. For example the bargaining power of an enforcement agent might be higher when the sanction is nonmonetary (for example imprisonment) rather than for fine.

1.3.3 Offsetting the Deterrence Diluting Effect of Corruption

The typical instruments used by the social planner are adjusting the probability of detection, type and size of sanctions. I will focus on the two later instruments, since I do not have data on the probability of detection. Furthermore, since increasing the probability of detection is costly (for example, more law enforcement agents would need to be hired) while increasing the fine is costless.

If the magnitude of the sanction increases, the bribe is likely to increase since individuals are wiling to pay more to avoid a larger sanction. At the same time higher bribe increases the expected costs of law violation (holding the probability of detection constant) and hence increases deterrence. It remains an empirical question whether this increase will fully offset the deterrence diluting effect of corruption.

The social planner has to choose what type of sanction to use: monetary or nonmonetary. In absence of corruption monetary sanctions should be used to full extend before nonmonetary sanctions are used, sine only monetary sanctions are socially cost less. In presence of corruption the nonmonetary sanction is transformed into a a monetary socially costless transfer (bribe)(Garoupa and Klerman, 2004).

If all offenders pay bribes the marginal social cost of increasing monetary and nonmonetry sanctions are the same. The marginal benefits of an increase in sanctions (increase in deterrence) depend on how sensitive the bribes are to changes in different type of sanctions. This sensitivity in the end determines which type of sanction is more effective in compensating for deterrence diluting effect of corruption. I further estimate those sensitivities empirically for the case of traffic violations in Russia.

Another tool for increasing deterrence is adding bureaucratic costs sanctions to the fine. For example, creating long lines to pay fines. Bureacratic costs are in esemnce nonmonetary sanctions. Adding them to a monetary fine increases the expected punishment and hence deterrence. In absence of corruption bureaucratic costs are socially costly, since time lost in line is a loss to an offender and not a gain to anyone in society. However similarly to the argument of Garoupa and Klerman (2004) in presence of corruption the offender ends up paying a bribe, instead of standing in line. This transforms a socially costly bereucratic cost to a socially costless monetary transfer[14].

[14]This holds if all offenders pay bribes and as a result avoid the official sanction and bureaucratic costs associated with it. Those offender who don't pay bribes will bare the social cost. The more common practice bribery is the less social costs arise with nonmonetary sanctions

1.4 Identification Strategy

In order to combat the traffic problems the government routinely increases the official sanctions for various types of traffic violations in order to increase deterrence. Table 2 presents the details on all the fine changes during the period from January 2011 to August 2013 for all the 11 types of traffic violations available in the data. I focus on changes in fines for three types of traffic violations:

i) Illegal passing (Driving on the wrong side of the road): The nonmonetary sanction decreased. Before the sanction change all the caught violators had their driver's license revoked. After the change first time violators pay a monetary fine of 5 000 rubles and only second time violators have their license revoked [15]

ii) Illegal parking: Monetary sanction (fine) increased 10 times (from 300 to 3,000 rubles)[16]. The website from which I collected the data does not offer a specific category for submitting bribes on parking violations. It only has a broader category (violating the rules connected with signs and road marking) which includes parking violations but also includes violations like not stopping at the stop sign, etc. However only the fines for illegal parking changed during the period of interest.

iii) Running red light: Monetary sanction (fine) increased 30 percent (from 700 to 1,000 rubles). Results for this type of violation are not very robust and do not satisfy identifying assumptions described below. Therefore I will not emphasize these results and present them only to provide the complete picture.

I do not include the period after August 2013 since there were major changes in sanctions for most of the types of traffic violations. In this setting it is hard to study longer term effects of fine change on bribe size since there there are many changes happening simultaneously.

[15]A driver is considered a first time violator if he/she did not commit the same type of violation during the prior year

[16]There was one out of 12 types of parking violation for which the fine have increased to 2,500 rubles

The objective of the study is to identify the causal effect of a change in official sanctions on the size of the bribe. In particular I study how sensitive the size of the bribe is to monetary and nonmonetary (loss of driver's license) changes in sanctions. To estimate the causal effect, I exploit the variation in changes of fines for different types of violations. I use the types of violations for which fines did not change as a counterfactual[17]. This allows to disentangle the effect of a change in sanction from the overall trend in bribes and from the part of the bribe associated with bureaucratic costs.

The main estimating equation examines the effect of different changes in sanctions on the size of the bribe:

$$Log(Bribe_{v,c,t}) = \alpha + \beta_1 PassChange_{v,t} + \beta_2 ParkChange_{v,t} + \beta_3 LightChange_{v,t} + \gamma t + \delta_t + \eta_m + \mu_v + \theta_c + \epsilon_{v,c,t} \quad (1)$$

where v indexes the type of traffic violation, c the city and t the date. The dependent variable is a logarithm of the average bribe for a given violation type at certain date. The variables $PassingChange$, $ParkingChange$ and $LightChange$ are indicator variables, taking value 1 for the affected types of violations after the official sanction changed and zero otherwise. The main specification of the model includes a linear trend t (controlling for any potential long term trend in the data). δ_t and η_m are year and month indicators, capturing seasonality and time heterogeneity. μ_v are violation fixed effects and θ_c is a city dummy (equals one for Moscow and zero for St. Petersburg). The coefficients of interest, βs are the difference-in-differences estimates of the average effect of sanction change on the size of the bribe.

To account for cluster correlation I provide clustered standard errors at the violation

[17]Some types of violations (i.e, having no technical documentation for the car, not giving way to pedestrian, speeding by 10-40 kilometers per hour) had few observations in 2013. I keep those observations in the control group. However, the results are robust to excluding these types of violations from the control group.

level. Given the small number of clusters (11 types of traffic violations) I use wild cluster bootstrap technique. This technique provides more conservative standard errors than the standard clustering procedures when the number of clusters is small.

The underlying assumption for the validity of the identification strategy is that in the absence of fine change the bribes for affected types of violations would have evolved similarly to those of counterfactual violations. In robustness section I provide pre-trend tests suggesting that there were no statistically significant difference between affected and counterfactual violations for illegal passing and parking.

The results for the red light running violation do not satisfy the robustness checks. I still report the main results so that the reader has a complete picture however I do not draw any conclusions based on them.

1.5 Results

1.5.1 Main Results

Table 3 presents the estimated coefficients for changes in sanctions on the logarithm of average bribe. The coefficients are statistically significant regardless of the method used to estimate the standard errors. The coefficients for changes in sanctions have the expected signs: the punishment for illegal passing decreased and so did the bribe, the punishment for illegal parking and red light running increased and so did the bribe. According to base line specification Table 3 column 1 the decrease in the risk of having the driver's license revoked decreased the size of the bribe by 61 percent on average (the 95% confidence interval with wild bootstrap clustered standard errors ranges from -0.78 to -0.45). An increase in the fine for illegal parking increased the bribe by 49 percent on average (95 confidence interval: from 0.23 to 0.75) and increased the bribe for running the red light by 42 percent (the 95% confidence interval 0.21 to 0.63). Since the results for red light running do not hold under the robustness checks, I will not focus on analyzing them further.

The estimated coefficients remain stable when more controls are added: column 2 includes the interaction of violation and city dummies to capture heterogeneity of violations associated with different cities; column 3 includes the intercept of city and time trend to capture long term trends at a city level; column 4 includes the intercept of violation and month dummies to capture seasonality in different type of violations; the last column includes all the controls simultaneously.[18]

The average bribe differs substantially for illegal passing and illegal parking (7500 rubles vs. 2500 rubles). Hence, the effect in levels differ substantially for these violations: 4 180 ruble decrease in bribe for illegal passing vs. 1240 ruble increase in bribe for illegal parking.

1.5.2 Estimating the elasticity of bribe to change in sanctions

The social planner is interested in understanding what tools to use to restore deterrence that is diluted due to corruption. Since monetary and nonmonetary sanctions are different in nature, it is hard to compare their effects directly. In order to compare the effects of changes in different type of sanctions on deterrence I present the results in terms of elasticity. I estimate how sensitive are the bribes to changes in sanctions and use the flowing formula to calculate the elasticity:

$$\epsilon = \frac{\%\Delta B}{\%\Delta S} \tag{2}$$

where B is a bribe and S is an official sanction.

The average bribe for illegal parking increased by 49% when the sanction for illegal parking increased by 900 percent. This gives a bribe-to-sanction elasticity of 0.055 (49/900=0.055). The 95% confidence interval ranges from 0.03 to 0.08[19]. This suggests

[18] The effect of sanction change in St. Petersburg is smaller then in Moscow, however, the difference is not statistically significant.

[19] The 95 percent confidence interval is calculated using robust and wild bootstrap clustered standard

that it might be hard to offset the deterrence-diluting effects of corruption by only increasing fines. Consider a hypothetical example where a fine for a violation is \$100 and the government realizes that the drivers are deterred optimally if the payment for the violation is 100% higher. To make the drivers paying bribes to pay \$200, the government would have to increase the fine all the way up to \$1800 (given the estimated average elasticity of 0.055). Such a big increase in fine might be socially unacceptable some drivers still don't pay bribes; because the government might want to reserve higher fines for more serious violations; potential offenders are risk-averse and wealth differs among individuals[20].

Calculating the bribe-to-sanction elasticity for nonmonetary sanction is less straightforward since it is harder to estimate percentage change in sanction. The sanction for illegal parking decreased: before the change all violators would have their license revoked, now it only happens to second time violators. I need to obtain an estimate of the value of the driver's license in order to to approximate the percentage change in sanction. The upper bound of the value of a driver's license is of particular interest since ultimately I would like to compare the elasticity of a bribe to a change in monetary and nonmonetary sanction. The elasticity of a bribe to a change in monetary sanction is small in magnitude, that is why I am mainly interested in estimating the lower bound for the elasticity of a bribe to a change in nonmonetary sanction. Elasticity is smaller when the change in sanction is bigger (the change in sanction is in the denominator). And for the case of illegal passing the change in sanction is higher when the value of the driver's license is bigger.

I use four alternatives to estimate the monetary value of the driver's license.

1. Legal help. There are companies offering legal services helping drivers with revoked driver's license to get the license back. Prices differ depending on the type of traffic violation which caused the revoke of the license. [21].

errors at the violation level

[20]For a more detailed discussion of why fine is often lower then the maximal see Polinsky and Shavell (1991) and Polinsky and Shavell (1979)

[21]The information comes from this website http://glavnaya-doroga.com/price

2. Obtaining a fake driver's license. There are companies which print fake driver's licenses[22].

3. Bribe payment for other types of traffic violation resulting in driver's license revocation. In the main data set of this paper there are two types of traffic violations which can result in driver's license revocation: illegal passing and speeding more than 60 km per hour above the limit. I use the maximum bribes payed to avoid the official sanction as the proxy for the value of a driver's license.

4. Bribe payed at the driver's license test. I use the data on the bribes individuals payed to obtain a driver's license at the driving tests [23].

Table 6 summarizes how much drivers paid to get a driver's license from the 4 sources listed above. Within each source there is variation in how much different drivers paid. For example the price for a fake driver's license varied from 15 to 30 thousand rubles. The Table includes information on the lowest, average and highest price paid. One thing to notice here is that the upper bound of how much drivers value the license is comparable across different sources. I use the highest of available estimates for the value of the driver's license to estimate the lower bound for the elasticiy of a bribe to a change in nonmonetary sanction.

The average bribe for illegal passing decreased by 61% when the sanction for illegal parking decreased. This gives a lower bound of the elasticity of a bribe of 0.63 (61/97=0.63). The 95% confidence interval for this elasticity is from 0.45 to 0.75[24]. This is a much larger then the elasticity of a bribe to a change in monetary sanction. The point estimates is more then 10 times larger (0.05 vs. 0.63). Even considering the confidence intervals the bribe seems to be at least 5 times more sensitive to changes in nonmonetary rather than

[22]The information comes from this website selling fake driver's licenses from a "donor", a person which actually has a driver's license and has the same name as an offender . The fake driver's license is supposed to work well with the database of the traffic police http://www.2prava.com

[23]The information comes from the website http://www.vashamashina.ru/bill.php

[24]The 95 percent confidence interval is calculated using robust and wild bootstrap clustered standard errors at the violation level

monetary sanctions. This suggests that in presence of corruption, increasing nonmonetary sanctions might be a more efficient way to increase deterrence than increasing monetary sanctions. Potentially this is due to the fact that nonmonetary sanctions increase the bargaining power of the enforcement agents.

1.6 Robustness

1.6.1 Selection

One challenge to the causal interpretation of the results is the composition of people who violate the rules changes with the change in sanction. For example, if the sanction increases, potential offenders with lower gains from crime or lower income can stop violating the rules. As a result the remaining pool of offenders would consist of richer individuals, so the average bribe might increase. This increase could be due to a change in the composition of individuals reporting bribes rather than the change in bribes due to change in sanction. I do not have any information on characteristics of individuals reporting bribes so I cannot test changes in composition directly. However assuming that the share of drivers who report their bribes remains constant, the number of law violation provides information on how potential offenders respond to changes in sanctions.

Unfortunately, the official statistics on the amount of law violations was not made available. However I can check whether the reported number of bribes in my data changed after the changes in fines. I estimate the same equation as equation (1), but use the number of bribe reports made during one month for a specific type of traffic violation in a specific city as dependent variable. Given the distribution of the dependent variable, model (1) is estimated using Poisson regression.

Table 4 provides the results. The coefficients for the change in sanctions for illegal passing and parking are not statistically significant, suggesting that the number of reported bribes didn't change. This provides the suggestive evidence that the results are not driven by the change in the composition of individuals reporting bribes.

The coefficient for red light running is positive and statistically significant. Suggesting that the number of bribe reports increased after the increase in sanctions in comparison to control traffic violations. This suggests that one must be careful when interpreting the results for running the red light.

1.6.2 Testing identifying assumption

The main identifying assumption is that the bribes for both comparison and treatment violations would have developed similarly in the absence of changes in sanctions. Even though it is not possible to observe the evolution of bribes in both states, I can study whether the evolution of bribes for different types of violations was similar before the changes in sanctions were implemented. Since I do not count with a long panel before the changes in sanctions were implemented, I create a placebo variable that takes a value of 1 for the six months prior to the sanction change, and zero otherwise (included in the model as $PrePark$, $PrePassing$ and $PreLight$). This strategy will maximize the power of the estimates to avoid not rejecting the null hypothesis due to having a smaller sample. If bribes evolved similarly between the different treatments and the comparison group of violations, I would expect to find that the estimates for pre-variables to be close to zero.

Table 5 presents the estimates including these pre- variables. The estimates for the pre- change coefficients for illegal passing and parking are close to zero and not statistically significant, suggesting that the bribes for these groups evolved similarly to those in the comparison group prior to the changes in sanctions. However, the pre-change estimate for running a red light does suggest that bribes for this violation in particular were already changing beforehand. It is important to note that the sanction change for this particular type of violation happens close to the beginning of my panel and therefore have few pre-change observations. Additionally, the estimates for the main coefficients of interest ($PassChange$ and $ParkChange$) remain unchanged. Taken together, these results suggest that the main identification assumption for illegal passing and parking hold, while results for running a red light should be interpreted with caution.

1.7 Discussion and additional benefits of bureaucratic costs

This section summarizes different ways in which nonmonetary sanctions, in particular bureaucratic cost, might be useful for optimal law enforcement in a corrupt country context.

Theoretical literature on law enforcement in presence of corruption suggests that corruption dilutes deterrence. To further support the idea that corruption dilutes deterrence empirically, I present evidence across the Russian cities. Figure 2 shows the correlation between the percentage of drivers which report giving bribes at least once during 2007[25] and the number of fatalities in traffic accidents [26]. The cities with higher percentage of drivers giving bribes tend to have higher amount of traffic accidents: a one percentage point increase in number of drivers giving bribes is associated with a 5% increase in fatalities from traffic accidents. This is a suggestive correlation supporting the idea that bigger scope of corruption is associated with less deterrence.

Corruption dilutes deterrence through offenders paying bribes smaller than the official sanction for the violation of law. In my data I do observe that a substantial share of drivers pays bribes below the official sanctions. Figure 3 and 4 provide histograms showing the dencity of bribes for illegal parking and illegal passing with a turn. The red dashed vertical line shows the official fine. The histograms show that some share of offender pay bribes smaller than official fine. However an even larger share of offenders pays a bribe above

[25]The data comes from a survey made by a Russian insurance company Rosgosstrah in 2007. More then 9 000 of drivers were surveyed in 37 Russian cities. The data is provided in a report published by the center of strategical studies. How the drivers deal with troubles on the roads, 2008

[26]The data on traffic fatalities is on the regional (not city level) and comes from official statistics of traffic police in Russia http://www.gibdd.ru/stat/, data on traffic fatalities

the official sanction: the median bribe is above the official fine for most of the types of traffic violations (see Table 1). My preferred explanation of this is that paying an official sanction is associated with bureaucratic costs, hence offenders are willing to pay higher bribes to avoid those bureaucratic costs. As a result bureaucratic costs increase deterrence by increasing expected punishment.

Bureaucratic costs can play a beneficial role in law enforcement in presence of corruption. Adding bureaucratic costs to a fine increases the expected magnitude of the punishment, which in turn helps to restore deterrence. In absence of corruption of corruption, bureaucratic costs are socially costly: they impose costs on offenders without creating gains to anyone (like time wasted waiting in line is a social loss). Corruption eliminates the social costs by turning bureaucratic costs into a monetary transfer. This argument is similar to the argument Garoupa and Klerman (2004) make formally for nonmonetary sanctions. This makes bureaucratic costs and other nonmonetary sanctions an attractive law enforcement tool when corruption is present and deterrence is diluted.

I do not have a source of variation in bureaucratic costs so I can't estimate by how much they dilute deterrence directly. However I do investigate this question for other types of nonmonetary sanctions. I find that bribes are more sensitive to changes in nonmonetary rather then monetary sanctions (elasticity of 0.68 vs. 0.05), suggesting that nonmonetary sanctions have higher potential in increasing deterrence (increasing bribes) than monetary sanctions do.

Bureaucratic costs are often already in place in corrupt countries. Figure 1 provides the cross country correlation between control of corruption (the higher the control the lower corruption) and government effectiveness (the higher the effectiveness the lower bureaucratic cost). This graph is constructed using Worldwide Governance Indicator project ranking of countries for 2013. If we abstract from the usual limitations of these indicators we can observe a strong correlation between red tape and corruption. Countries suffering from corruption also usually suffer from bureaucratic costs. Usually, the red tape decreases the efficiency of the economy. In the particular case of law enforcement it might be beneficial since the presence of bureaucratic costs contributes to restoring the deterrence diluted by corruption.

All the above arguments suggest that in presence of corruption the social planner might want to rely more on nonmonetary sanctions and bureaucratic costs in particular: they have similar costs and higher potential in increasing deterrence.

1.8 Conclusion

This paper contributes to the literature on optimal law enforcement in presence of corruption. I first document that corruption dilute deterrence, confirming the theoretical predictions. I further study if deterrence diluting effect of corruption can be offset by an increase in sanction and which type of sanction in particular. Increasing deterrence is important in practice, it can increase public safety and save lives. In the context of traffic violations in Russia, drivers are not deterred enough and violate rules: hundreds of drivers park on a tram line contributing to making Moscow the most congest city in Europe; a big share of drivers drive drunk[27] making the traffic fatalities rates in Russia comparable to rates in countries with smaller GDP per capita - Pakistan and Burundi. The government recognizes traffic safety as an important issue and increased monetary sanctions for various types of traffic violations in order to increase deterrence. The traffic situation largely remained unchanged.

I study if the deterrence diluting effect can be offset by increasing the sanction using the self-reported data on bribes to traffic police in Russia. I find that when monetary sanctions increase, the average bribe doesn't increase by a lot: estimated elasticity of 0.05. Suggesting

[27]driving after drinking any amount of alcohol is unlawful. Only 50 percent of drivers report that they have never driven while drunk. Source: "how many drivers drive drunk", Steer.ru, March 21, 2012

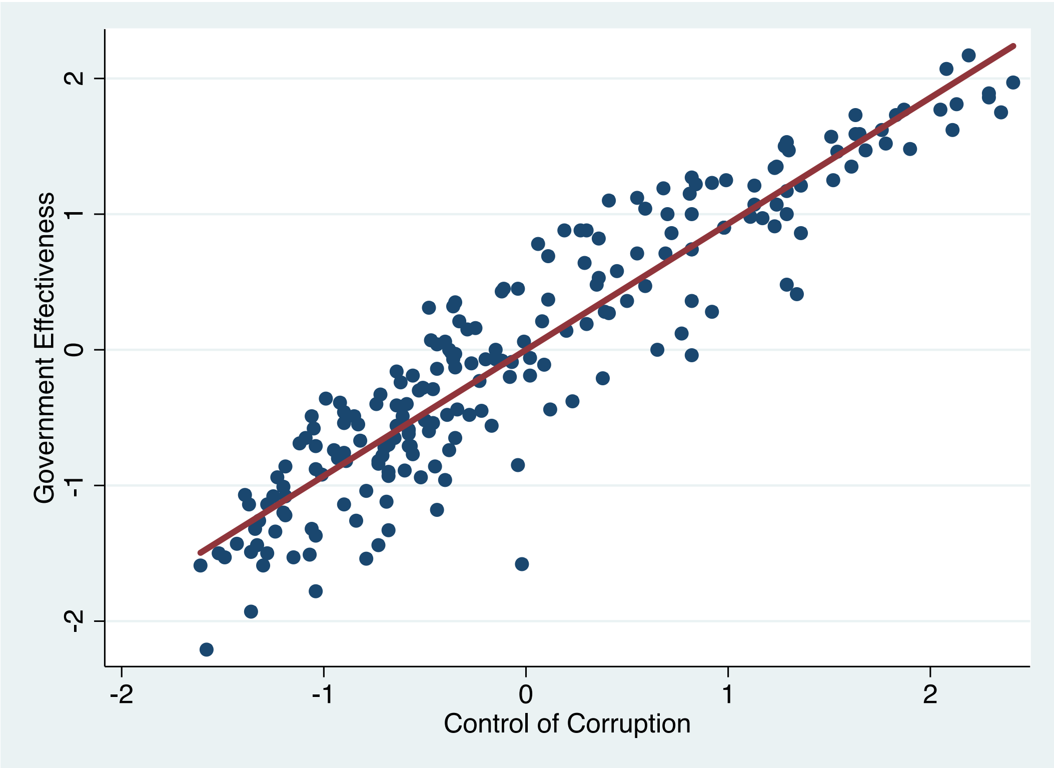
that to offset the deterrence diluting effect the monetary sanction has to be raised to a level which might be socially unacceptable. I find that the average bribe is more sensitive to a change in nonmonetary sanction, estimated elasticity of 0.68. This suggests that it might be possible to increase deterrence by threatening to impose a nonmonetary sanction. In practice however the nonmonetary sanction is not imposed but substituted by a bribe. This eliminates the social costs of nonmonetary sanctions and turns them into a socially costless monetary transfer. Making nonmonetary sanctions more attractive in presence of corruption.

The typical example of nonmonetary costs is imprisonment. I argue that bureaucratic costs can be also considered a nonmonetary sanction. Those include filling up long and confusing forms, long wait in lines. Adding bureaucratic costs (red tape) to monetary sanctions can increase expected punishment and hence is another tool available in restoring the deterrence diluted by corruption.

The framework used in this paper doesn't only apply to corruption in traffic police in Russia. A lot of countries are suffering from bribery in different areas of social life: from traffic violations to sanitary standards violations.The website similar to the one I use in this paper already exists in Kenia and researches can design similar websites in other countries to study corruption in any area of interest. Learning more about bribes could allow to increase deterrence making the society safer, without dramatic increases in social costs.

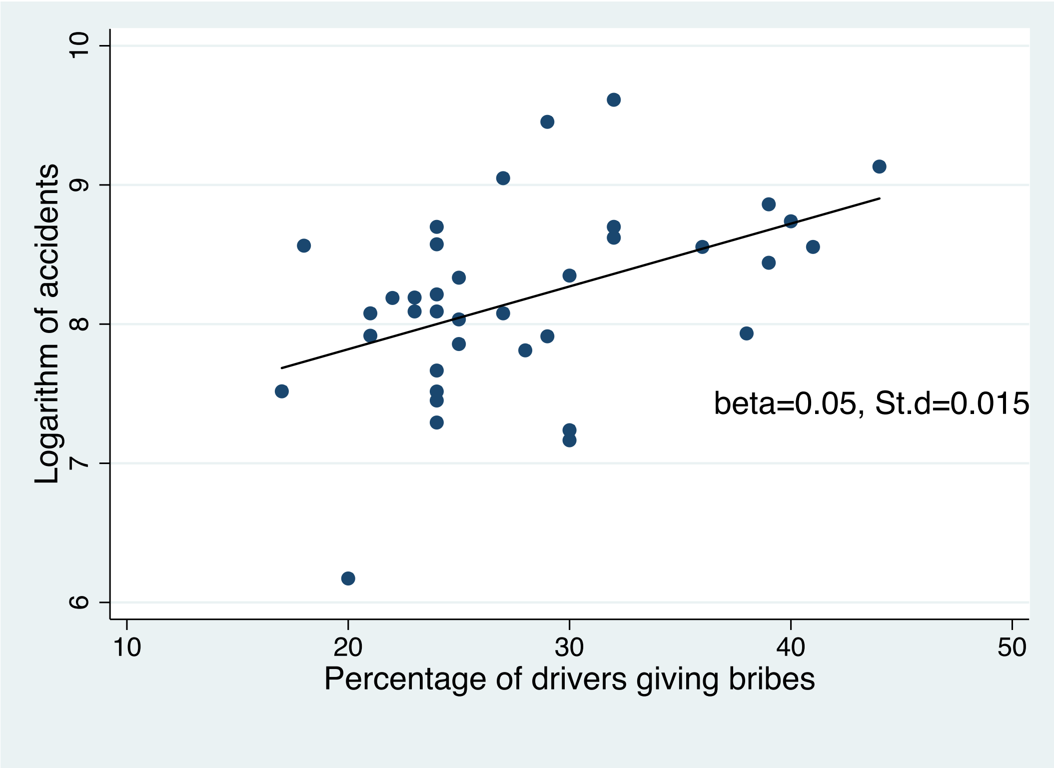
1.9 Figures and Tables

Figure 1: Effectiveness of Government and Corruption Across Countries



Note: Cross-section of countries, data obtained from the Worldwide Governance Indicator project ranking of countries for 2013.

Figure 2: Percentage of individuals giving bribes and number of accidents



Note: The data obtained from a survey conducted by the Russian insurance company Rosgosstrah in 2007. More then 9,000 of drivers were surveyed in 37 Russian cities. Data on traffic fatalities at the regional (not city level) obtained from official statistics of traffic police in Russia (http://www.gibdd.ru/stat/)

Figure 3: Histogram of bribes: illegal parking before the fine change

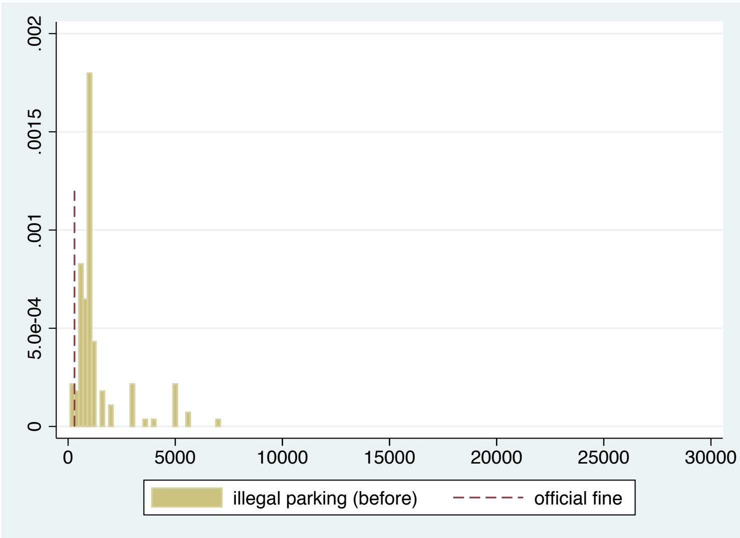


Figure 4: Histogram of bribes: Illegal passing with a turn

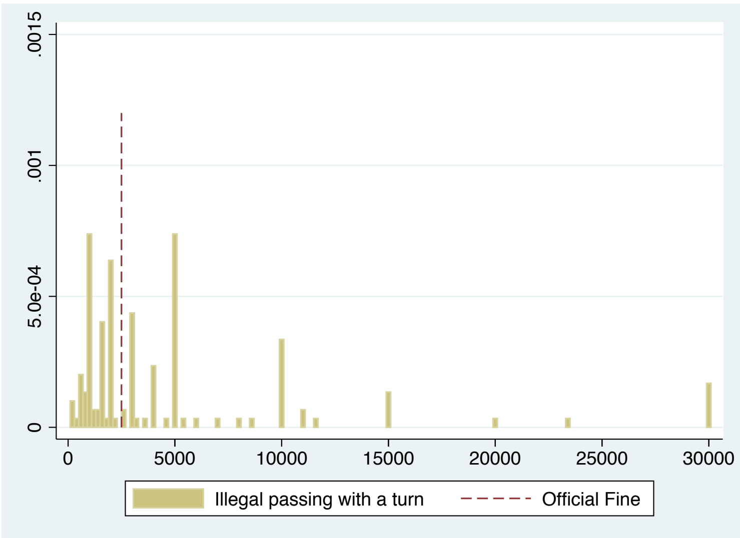


Table 1: Descriptive statistics of bribes

violation type	official fine in January 2012	Obs.	Mean	Median	Std. Dev.
Illegal passing	DLR	269	7581	5000	8107
Violating the rules connected with signs and road marking	300 rubles	217	2494	1000	4514
Driving on the red light	1 000 rubles	485	3606	1200	5743
No car insurance	300-800 rubles	162	1569	467	4370
Illegal passing while turning	500 rubles	149	4757	2500	6140
No seat belt	500 rubles	83	1699	500	5078
No technical documentation for the car	800-1000 rubles	115	791	417	2771
Not yielding to pedestrian	100-300 rubles	12	1667	800	2729
Speeding (10-40km per hour)	100-300 rubles	24	1367	500	2170
Speeding (40-60 km per hour)	1000-1 500 rubles	133	1089	900	1589
Speeding (more then 60km per hour)	2 000- 2 500 rubles or DLR	129	6551	3000	7880

Note: *DLR stands for driver's license revocation

Table 2: Official fine changes

Type of violation	Date of change	Fine before the change	Fine after the change
Illegal passing	1.01.2013	DLW* for a period of 3-6 month (no fine).	If the rule is violated for the first time there is a fine of 5000 . if the rule is violated for the second time (during a year period) DLW for 6-12 month.
Illegal parking	1.06.2012	300 rubles	3 000 rubles
Driving on the red light	1.01.2012	700 rubles	1 000 rubles
No car insurance	1.01.2013	300-800 rubles	300-800 rubles and the licence plate is removed until the insurance is purchased
Illegal passing while turning	No Change	-	-
No seat belt	No Change	-	-
No technical documentation for the car	No Change	-	-
Not yielding to pedastrian	No Change	-	-
Speeding (10-40km per hour)	No Change	-	-
Speeding (40-60km per hour)	No Change	-	-
Speeding (more then 60 km per hour)	No Change	-	-

Note: *DLW stands for driver's license withheld

Table 3: Effect of sanction change on log(bribe size)

	(1)	(2)	(3)	(4)	(5)
Dep. Var.: Log(Bribe)					
PassChange	-0.614***	-0.637***	-0.617***	-0.678***	-0.695***
	(0.210)	(0.211)	(0.209)	(0.232)	(0.234)
	<.01	*<.01*	*<.01*	*<.01*	*<.01*
ParkChane	0.491***	0.488***	0.489***	0.370**	0.371**
	(0.165)	(0.168)	(0.165)	(0.188)	(0.189)
	<.01	*<.01*	*<.01*	*0.026*	*0.018*
LightChange	0.421***	0.416***	0.419***	0.476***	0.476***
	(0.134)	(0.135)	(0.134)	(0.139)	(0.139)
	<.01	*<.01*	*0.012*	*0.018*	*0.012*
Violation*City	-	Y	-	-	Y
City-trend	-	-	Y	-	Y
Violation*Month	-	-	-	Y	Y
R-squared	0.36	0.366	0.36	0.417	0.421
Observations	1,616	1,616	1,616	1,616	1,616

Note: All regressions include violation fixed effects, year, month and city dummies. Heteroskedasticity-robust standard errors in parenthesis, wild-bootstrap clustered p-values in cursive (at the violation level). *** p<0.01, ** p<0.05, * p<0.1

Table 4: Robustness check: number of bribes reported

	(1)	(2)
Dep. Var.: # of bribes per month		
PassChange	-0.117	-0.124
	(0.19)	(0.19)
ParkChane	0.0421	0.0381
	(0.16)	(0.16)
LightChange	0.617***	0.610***
	(0.12)	(0.12)
Violation*City	Y	Y
City-trend	-	Y
Model	Poisson	Poisson
R-squared		
Observations	366	366

Note: All regressions include violation fixed effects, year, month and city dummies. Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Robustness check: parallel trend assumption

	(1)	(2)	(3)	(4)	(5)
Dep. Var.: Log(Bribe)					
PassChange	-0.565***	-0.588***	-0.570***	-0.611***	-0.625***
	(0.218)	(0.220)	(0.219)	(0.243)	(0.245)
	<.01	*<.01*	*<.01*	*<.01*	*<.01*
ParkChane	0.539**	0.537**	0.537**	0.385	0.265*
	(0.169)	(0.174)	(0.169)	(0.196)	(0.198)
	0.034	*0.046*	*0.034*	*0.112*	*0.112*
LightChange	0.930**	0.928**	0.926***	0.846***	0.716***
	(0.146)	(0.147)	(0.146)	(0.164)	(0.164)
	0.012	*<.01*	*<.01*	*<.01*	*<.01*
PrePassing	0.134	0.132	0.128	0.0881	-0.117
	(0.180)	(0.181)	(0.181)	(0.212)	(0.213)
	0.328	*0.352*	*0.334*	*0.580*	*0.584*
PreParking	-0.146	-0.131	-0.146	-0.0854	-0.220**
	(0.216)	(0.219)	(0.216)	(0.219)	(0.224)
	0.504	*0.554*	*0.508*	*0.492*	*0.496*
PreLight	0.940***	0.943***	0.939***	0.757***	0.828***
	(0.149)	(0.150)	(0.149)	(0.221)	(0.222)
	<.01	*<.01*	*<.01*	*<.01*	*<.01*
Violation*City	-	Y	-	-	Y
City-trend	-	-	Y	-	Y
Violation*Month	-	-	-	Y	Y
R-squared	0.36	0.366	0.36	0.417	0.421
Observations	1,616	1,616	1,616	1,616	1,616

Note: All regressions include violation fixed effects, year, month and city dummies. Heteroskedasticity-robust standard errors in parenthesis, wild-bootstrap clustered p-values in cursive (at the violation level). *** p<0.01, ** p<0.05, * p<0.1

Table 6: Estimates of the vaue of driver's licence

Source of information:	Monetary Value of the driver's license		
	lower bound	average	upper bound
Legal help	-	12.5	-
Fake driver's license	15	22.5	30
Bribe for other types of traffic violation resulting in DLR	0.3	18.9	37.5
Bribe at the driver's license exam	1.5	14.9	25

Note: DLR stands for driver's license revocation. Values are in thousands of rubles.

2 The Value of Infrastructure Investment and Urban Growth: Sewage and Water Supply in Chicago 1830-1930.

2.1 Introduction

Public infrastructure spending is a big budget share both in developed and developing countries, however evidence on economic returns to infrastructure is limited.

Starting from the 1850s, Chicago experienced rapid population growth that increased demand for clean water supply and sewer provision. In this paper I estimate the economic returns to sewage and water connections in Chicago during 1850 to 1930 period. The exercise of estimating economic returns to infrastructure access using historic data is of particular interest since many developing countries are currently experiencing an urban growth processes and have a disease environment similar to that of Chicago in late 1800s. Large rural to urban migration increases the population density and creates additional need for infrastructure and water and sewage in particular. To make a decision on the optimal provision of infrastructure it is important to gain better understanding of the returns.

Previous papers on returns to infrastructure in 19th-20th century in the U.S. have focused largely on the effect of water and sewage supply on death rates. For example, Ferrie and Troesken (2008) document a decrease in death rates from various diseases in Chicago due to better water supply. Alsan and Goldin (2015) document a decrease in infant mortality in Massachusetts due to better water and sewerage provision, while Cutler and Miller (2005) find that water filtration and chlorination decreased death rates in 13 U.S. cities. However, an improvement in infrastructure access is likely to have a much broader effect than a decrease in death rates for some groups of the population – it likely has other externalities. Better infrastructure allows more people to live in close proximity to each other, and closer to industrial centers. Introduction of water and sewage increases productivity of firms and provides benefits to individuals, not only through a better disease environment, but also through higher quality of life. In order to capture broader benefits of infrastructure this paper estimates the effect of access to water and sewage on land value. The land value represents economic returns to infrastructure both to firms and individuals: the net present value of all future amenities minus all future taxes.

I use historical data on land values in Chicago between 1830 to 1930. The unit of observation is 1 by 1 square mile land plots. I first explore the effect of increasing water and sewage access on land values using panel data with plot level fixed effects. The data on water and sewer access is available at the block-street level. I aggregate this at the 1 by 1 square mile plot level and calculate the percentage of street blocks within each plot that have access to water and sewage as my measure of infrastructure access. I find a positive and statistically significant relation between water access and land values, that is robust to various specifications. However, sewage does not seem to be a statistically significant factor in explaining land values over time. These effects are robust to classifying plots into three categories depending on the level of access (i.e. no-, low- and high-access), in order not to rely on a linear function between access and land values. Additionally, I perform a placebo test and find that the future values of infrastructure access are uncorrelated with current land values – suggesting that my estimates are not driven by anticipation or by spurious correlation.

However, identification that relies on panel data with fixed effects is making very strong assumptions regarding the independence of the infrastructure access variables and the error term. In practice the decision on infrastructure investment is likely be endogenous to local outcomes or other unobserved characteristics. For example, desirable areas where rich individuals live are more likely to both experience land value growth and better infrastructure provision. An exogenous source of variation in infrastructure access or spending is needed to estimate the causal effects.

I take advantage of the process of urban growth in Chicago to find the source of exogenous variation in access to infrastructure. During 1850 to 1930, Chicago was expending its territory since surrounding suburbs were choosing to join the City. One of the reasons why suburbs chose to join the City was the fact that sanitation was a pressing issue and Chicago promised the suburbs that joined better and cheaper access to waterworks and sewerage. Once the suburb joined, it indeed experienced growth in infrastructure access. Residents of a particular plot may have directly influenced whether their suburb joined. However a suburb was not eligible to join unless it was directly adjacent to Chicago (i.e. had a common boarder with Chicago). I exploit this fact to construct an instrument for water and sewage access and estimate the causal effect of infrastructure on land values. In particular, I use an indicator for whether a neighbor is part of Chicago and a function of the number of years since it joined to instrument for my infrastructure variables. The fact that the neighboring suburb joins makes the suburb of interest eligible to join, and hence potentially can receive better access to infrastructure in future years. The instrumental variables proposed are strongly correlated with the access to infrastructure. It is unlikely that residents of a plot would have significant influence on whether and when a neighboring suburb joins, making my instrument plausibly exogenous.

Using instrumental variables yields similar results as those found by using panel data with plot fixed effects. According to my preferred specification, if access to water increased by 1%, the land value of an acre of land would increase by 0.5 percentage points on average. The effect is also economically significant: increasing infrastructure access from zero to access on all street-blocks is equivalent to moving a plot approximately 4 miles closer to the central business district in 1928.[28] I find that the returns to water supply are larger and more robust than those of sewage. Water has approximately 2 times larger magnitude of an effect and also tends to be statistically significant in most of the specifications. The relatively higher importance of water provision might be explained by the fact that water access is a necessary condition for a safe disease environment and high quality of life. Additionally, in Chicago there were close substitutes for sewage disposal other then being connected to the sewage network (e.g. disposal in vaults), while there were few substitutes for clean lake water.

This paper contributes to several literatures. First, it is related to the literature studying historic growth of American cities. For example, Hornbeck and Keniston (2014) studies barriers to urban growth looking at the trajectory of land and property values after the Boston fire in 1872. Siodla (2014) examines the how the San Francisco earthquake and fire in 1906 affected the location of firms. My paper examines city historic growth in the context of increasing factor productivity through the expansion of infrastructure. Second, it relates to a literature estimating the returns to investments in public infrastructure. For instance, investments in education infrastructure have been found to increase property values in the modern U.S. (Cellini et al., 2010). Finally, this paper contributes to a literature estimating the value of city amenities to firms and individuals (Albouy, 2015, Albouy et al., 2013) and the amenity value of infrastructure (Haughwout, 2002).

Section 2.2 provides historical background of urban growth and sanitation in Chicago. Section 2.3 describes the data. Section 2.4 discusses the identification strategy, Section 2.5 provides panel regression with fixed effects and instrumental viable estimates. The final section concludes and discusses future avenues for research.

[28]The last year for which I have land values available

2.2 Historical Background

2.2.1 A Rapid Growth

The city of Chicago experienced rapid growth in trade, population and territory during 1845 to 1900. The transformation was so fast that Chicago was called "the lightning city"[29]. Until 1848 Chicago was one of many cities located on the Great Lakes with access to a waterway, to the West coast and Europe. In 1848 the Illinois and Michigan canal was completed, and created a direct connection between the Great Lakes and the Mississippi River. This waterway replaced the slow and muddy roads linking Chicago with the hinterland. This allowed for the cheaper delivery of multiple commodities to and from the Great Lakes through Chicago. The decrease in transport costs boosted the trade of lumber and grain in the city. Additionally, the expansion of the railroad system spurred even faster growth. In 1850 only one railroad entered Chicago. By 1856 Chicago was the focus of ten lines. During the 1850s to 1870s, Chicago became one of the world capital for commodity trade, a large center for the distribution of lumber (which was the main construction material at the time), wheat and corn, cattle and hog.[30].

The increase in trade volumes was accompanied by a rapid population growth. The population increased from 30,000 in 1850 to 110,000 in 1860, and continued to increase partly due to a large inflow of migrants from across the world. In 1889 Chicago became the second largest city in the U.S.[31]

2.2.2 Water Supply and Sewage Disposal: Sanitation as a Pressing Issue

Until the 1850s most of the residents in Chicago obtained the water through delivery by means of carts or from wells. The most popular way of sewer disposal were vaults. The waste, however, did not absorb fully because most of the ground in the City was clay and the terrain was flat – at the same level as the lake. As a result the ground was swampy, muddy and unhealthy. As the population, and population density started to grow the quality of water in the wells worsened since what was disposed in the vaults was contaminating the wells. One resident of Chicago at the time wrote to the Sanitary News in 1884: "*In time, the water drawn from the wells began to taste, -a little blackish at first, then saltish, and finally it had a perceptible odor which finally became offensive. A well, at length, had the odorous characteristic of a privy vault*[32]."

The lake was the main source of higher quality water in Chicago. Starting in 1853, the first waterworks

[29] p. 35, Mayer and Wade (1974)
[30] This summary is based on Mayer and Wade (1974)
[31] Based on the discussion in Spinney (2000).
[32] Brown (1894), page 27

were constructed. However, most of the population did not have access and continued to get water delivered by cart from the lake or extracted from wells. Figure 6 shows the map of Chicago and how the water supply grid was expanding. The water grid is colored in blue, while the city limits are in red.

Until the mid 1850s, sewage was mainly disposed in vaults. In 1856 the first sewer lines were constructed, and the city started disposing of the waste into the river. The river was flowing into the lake, contaminating the water supply. However, at that time, the link between diseases (like typhoid or cholera) and contaminated water was not well understood. Additionally, the waste from slaughter houses and distilleries was also largely disposed into the river, contaminating the lake further. The expansion of the sewage grid followed similar patterns to those of water grid shown on Figure 6.

As the population of the city was growing, more waste was dumped into the sewer and as the sewer grid was expending, more of it was ending up in the lake. The quality of the water was particularly bad on days when the river got flooded by rain or when the wind blew from the south pushing more of the polluted river water to the lake. The Governor explained: "*the water as such times was not only horror of all good housewives, but it was justly thought to be very unhealthy... when the wind blew strongly from the south... the water from the river made from the sewage mixed with it into abominably filthy soup, was pumped up and distributed through the pipes alike to the poorest street gamin and to the nabobs of the city.*"[33]

Sanitation was a pressing issue in Chicago and in surrounding suburbs during the second half of the 19th century and the beginning of 20th century. A large effort was made to provide better quality water. There were several stages of the process, including multiple and not fully successful attempts to reverse the flow of the river away from the lake (by deepening existing canals and installing pumps); as well as building 2 and 4 mile tunnels into the lake to draw cleaner the water further from the river. All those measures failed to consistently protect the lake water from the the river and sewer. In 1900 the Sanitary and Ship Canal was completed, it successfully reversed the river flow away from the lake. This provided Chicago with a consistently clean source of water for the first time in many years.

2.2.3 Annexations

During 1850 to 1920, Chicago experienced a rapid growth in territory through annexing surrounding suburbs (i.e. townships and villages). Figure 5 provides a map detailing the territories and dates of the suburb annexations. A suburb could only be annexed to Chicago if it was directly adjacent to the city (i.e. had a common border). Some areas were annexed by an act of the legislature of Chicago. However,

[33] Brown (1894), page 32

the vast majority of suburbs chose to join the city by election. If the suburb voted in favor of annexation, Chicago always allowed it to join. All the suburbs in my sample eventually joined Chicago.

Chicago promised access to water and sewage to suburbs which joined the city. Access to better and cheaper infrastructure was one of the main reasons why the suburbs chose to join. The main alternative to Chicago waterworks drawing the water from the lake was boring for water. The wells neither provided good quality water nor were cheep. Table 7 provides the costs of supplying water in 1885-1886 in Chicago and surrounding suburbs. The costs of providing water were at least 50% higher in the surrounding suburbs. The rapid population growth was potentially inducing suburbs to switch to a more cost efficient technology with smaller marginal costs.

Once a suburb joined the city it experienced growth in access to infrastructure. Figure 6 illustrates that this happens with a lag. Indeed, it takes time to plan and contract the actual infrastructure.

2.3 Data

The main source of data on land values was obtained from Hoyt and Millis (2000), "One Hundred Years of Land Values in Chicago: The Relationship of the Growth of Chicago to the Rise of Its Land Values, 1830-1933". The book contains maps with data on the average land values per acre for each square mile in dollars. The data is based on actual sales and is available for 6 points in time: 1836, 1857, 1873, 1892, 1910, 1928.[34] Some data is available at a more disaggregated level. Those 1 by 1 square mile land plots are the main units of observation.

The data on water and sewage availability is available at a street block level. The data comes from Semi-Annual reports of the Board of Water of Sewerage Commissions. Those reports were digitized and made available by Historical Urban Ecological GIS Data Portal[35]. First, I drop streets containing missing information on the date of water and sewage connection. Next, for a given year, I calculate the percentage of street blocks within each plot that have access to water and sewage.

One of the limitations of the data is that the data on sewage and water is unavailable before a suburb joins Chicago. However once the township joins the city the data becomes available. So the first years of data after the suburb joins can be used to approximate the exposure to sewage before the suburb joined.

[34]All the land values are represented in 1967 dollars. Historic CPI is used to make the adjustments. The Historic CPI is provided by the Federal Reserve Bank of Minneapolis *https://www.minneapolisfed.org/community/teaching-aids/cpi-calculator-information/consumer-price-index-1800*

[35]Link to the Historical Urban Ecological GIS Data Portal *http://hue.uadata.org/gis/*

The data on dates of annexations and boarders of townships is biased on a map by Chicago Department of Public Works, Bureau of Maps and Plats, 1930. See Figure 5 for the map. The map is divided into squares which are in most cases equivalent to my unit of observation: 1 by 1 mile plots.

Table 8 provides descriptive statistics of the data for the years for which land value data is available. In 1836 Chicago used to be very small by area (few plots were a part of it) and had virtually no sewage and waterworks. By 1928 most of the townships and land plots joined Chicago and gain much higher access to infrastructure.

2.4 Identification strategy

The main goal of this paper is to estimate the effect of infrastructure on land values. The main unit of observation is the 1 by 1 mile plots of land over time.

I use a panel data approach and explore the time variation in water and sewage access on the plot level in order to estimate the effect of water and sewage on land values. I estimate the following equation, where the dependent variable is the logarithm of land value:

$$Log(LandValue)_{i,t} = \beta Infrastructre_{i,t} + \theta X_{i,t} + \eta_i + \psi_t + \varepsilon_{i,t} \quad (3)$$

where $X_{i,t}$ is a matrix including distance to the central business district interacted with the year, as well as distances to the river and lake also interacted with year. It is important to include these variables since they capture potential differences in land value trends depending on the relative position of the plots. For example, plots that are further away from downtown could have originally been sparsely populated, but with the rapid growth of the city these areas rapidly urbanized. By including these distance-time interactions, I am able to capture these effects. The variable of interest, $Infrastructure$ is either the percentage of streets with water connection, or sewage. It is not possible to include both at the same time since they are highly correlated (though not perfectly – correlation of 0.89). β, my coefficient of interest, estimates the average effect of increasing infrastructure access in given plot. η_i are the plot-fixed effects, capturing any unobserved plot-level characteristic that remains fixed over time, ψ_t are year effects, and $\varepsilon_{i,t}$ is the error term assumed to be uncorrelated with the variable of interest.

The causal interpretation of β is likely to be problematic in this context. First potential issue is reverse causality. It is possible that those plots which were already growing in land value also decided to get larger access to infrastructure. Than the β coefficient would not represent the effect of infrastructure access on land values. Secondly, there may be omitted variables or other unobserved characteristics mediating the relationship between infrastructure access and land values. For example, richer individuals may live in a location with more desirable characteristics (that are not directly observed) and may also value access to infrastructure more. Including distance-time interactions does control for some unobserved characteristics related to the desirability of plot location, but it may not fully capture other dimensions.

To address the problems listed above I use an instrumental variable approach. I can exploit the variation in sewage and water expansion that can partially be explained by the annexations of the different suburbs to Chicago. As described in the previous section, one of the reasons to join Chicago was gaining access to cheaper and better water and sewage systems. Once the suburb joined Chicago, it experienced growth in water and sewage provision. Consider a particular plot that highly values access to water and at the same time contains rich residents. The decision of whether their suburb (including this plot) joins Chicago can be potentially endogenous since residents of this plot can directly affect the decision during elections. However, a suburb can not join Chicago unless it is directly adjacent to it (i.e. has a common border). The decision made by a neighboring suburb on whether to join Chicago should not be directly affected by this particular plot. Joining Chicago eventually increases water and sewage provision. All the suburbs in my sample eventually join Chicago so the only variation comes from the timing of joining the city.

I take advantage of these facts in constructing instruments for infrastructure access. In particular, I use a dummy variable indicating whether the neighboring suburb joined Chicago by period t ($NeighborJoin_{i,t}$) and a function of the number of years since the neighboring suburb joined Chicago ($f(YSNJ)_{i,t}$). It is important to use the function of years because there is a significant lag between joining Chicago and the eventual expansion of access to infrastructure. The functional form of this lag is described in greater detail in the results section. Using these instruments, I estimate the effect of infrastructure access using two stage least squares. In the first stage, I estimate the following equation:

$$Infrastructure_{i,t} = \pi_1 NeighborJoin_{i,t} + \pi_2 f(YSNJ)_{i,t} + \kappa X_{i,t} + \eta_i + \psi_t + \epsilon_{i,t} \tag{4}$$

In the second stage, I use the predicted infrastructure ($Infrastructure^{*}$) and estimate the same equation as in equation 3:

$$Log(LandValue)_{i,t} = \beta Infrastructure^{*}_{i,t} + \theta X_{i,t} + \eta_i + \psi_t + \varepsilon_{i,t} \tag{5}$$

The causal interpretation of β depends on the instruments being uncorrelated with the error term and the instruments not having a direct impact on land values other than through infrastructure access. The first assumption is discussed above and the exclusion restriction is discussed in detail with the results. In all the regressions presented below I cluster standard errors at the plot level.

2.5 Results

2.5.1 Main Results

In order to estimate the effects of increasing access to water and sewage on land values I estimate equation 3 by OLS, where I first use the percentage of water connections as a main explanatory variable and later the percentage of sewage connections instead. Columns 1 and 4 of Table 9 present the estimated coefficients for % Water and % Sewage, respectively. There is a positive and statistically significant correlation between water expansion and land values. This result remains robust in column 2, where I use a more flexible function for controlling for differential distance trends. Specifically, I interact each distance variable (to central business district, lake and river) with dummy variables for each year. The estimated coefficient is not statistically different from that estimated in the first column, though the estimate appears to be less precise.

One potential shortfall of the estimates in columns 1 and 2 is that the suburbs that joined Chicago were experiencing different trends in growth of land values, which may have led them to join Chicago (and later gain access to water) at different times. To account for this possibility, in column 3 I add suburb-group specific time trends; where a suburb-group is defined as the group of suburb joining Chicago during the same time period. The magnitude of the coefficient of access to water increases under this specification and remains statistically significant. However, the coefficient is not statistically distinguishable from my initial estimate. The estimate in column 3 suggests that if access to water were to increase by 1%, land value would increase in turn by 0.52 percentage points, on average.

Considering the coefficients for sewage in columns 4 to 6, I find a positive though not statistically significant correlation between sewage access and land values. Only in column 6, after accounting for potential differences in trends according to when a suburb joins Chicago, the estimated coefficient is significant at the 10% level and about half the magnitude as that estimated for access to water. In these estimates, I am assuming that the relation between water or sewage access is linear. However, this may not be true since there may be a non-linear relation between infrastructure access and land values. For example, moving from zero access to 5% access may not be the same as moving from 80% to 85% – that is what is implied by the functional form used.

As robustness, I allow for a more flexible function of access to infrastructure, where I now classify each plot into three groups. The omitted category is zero infrastructure. The second group is *low* infrastructure, defined as a plot with less than 50% coverage. The final group is *high* infrastructure, defined as a plot with more than 50% coverage. Instead of the linear percentage access, I include two dummy variables for low or high access to infrastructure in the regression. The estimates are presented in Table 10. The estimates for water access are consistent with the results found including just the percentage of access. There is a monotonically increasing effect of moving from no water, to low, and moving from low to high. The coefficient for high access to water is almost twice the magnitude as that for low and the coefficients are highly statistically significant.

The results for sewage are less clear, where overall there seems to be a positive correlation between having sewage compared to not having it, but the effect is not monotonically increasing. Having some sewage has a larger correlation with land values than having high access to sewage. This non-monotonic relation in increasing access to sewage may explain why the correlation estimated in the previous table was not very strong.

Results suggest that there is a positive and significant correlation between access to water and land values, while the relation between sewage and land values is less strong and measured less precisely. One of the potential reasons why the effect of sewage is less robust is that there might be complementary between water and sewage. Alsan and Goldin (2015) find that both water and sewage are needed to provide healthy disease environment. Unless there is sewage the fecal mater can spread diseases through sources other than water. However estimating the complementary of water and sewage with the data I have is difficult since water and sewer access are highly correlated. It may be possible to study those effects using more disaggregated level data on land values.

For additional robustness, I perform a placebo test, where I include the future value of access to infrastructure (e.g. the lead of % water) instead of the current value. If the future value of infrastructure were positively associated with current land values it would suggest that either individuals were anticipating the expansion of infrastructure or that the estimated effects of infrastructure may be spurious. The results are presented in Table 11. Reassuringly, I find that the estimated coefficients are significantly smaller than those previously estimated and not statistically different from zero. These results provide some suggestive evidence that my previous results may not be driven by anticipation or are spurious.

2.5.2 Instrumental Variables Approach

The causal interpretation of my previous results rely on the assumption that, after accounting for different factors such as plot fixed effects and distance-time trends, access to infrastructure is not correlated with the error term. Given that this assumption likely does not hold, I use an instrumental variables approach. In this setting, I instrument access to infrastructure exploiting the timing when neighboring suburbs joined Chicago. Arguably, the time the neighboring town joins Chicago is independent of the own town's decision, but it does affect whether the own town joins since it becomes eligible to do so (since it would now have a common border with Chicago). As explained previously, there are no records of townships forcing or influencing neighboring townships to join Chicago. The only exception is that Chicago "forcibly" annexed a territory by an act of legislature, but this should not be a problem as long as that was not influenced by other townships.[36]

Given that there is a lag in the time that the neighboring township joins and the time it takes the water and sewage grid to expand, I will use a dummy variable indicating whether the neighboring town joined Chicago, and a polynomial function of the years since they joined.[37] Figure 7 illustrates this for access to water, where the number of years since the neighboring township joined is plotted on the horizontal axis, and access to water (% of street blocks with water) is on the vertical access. Notice that the relation is not linear and some polynomial function of time would fit the data better.

In Table 13 of the Appendix, I present the estimates for the instrumental variables of the first stage described by equation 4 above. Throughout all specifications, the F-Statistic for the joint significance of my instruments is highly significant. This suggests that the correlation between my instruments and the instrumented variable is strong. From the previous figure, it seemed that a cubic function of time fit the data best. This is confirmed by the fact that the F-statistic is higher when using a cubic function rather than a quartic function (F statistics of 116 vs. 94 for water access). Nevertheless, I regress the different functions and different model specifications for additional robustness.

The 2SLS estimates of the effects of infrastructure access on land values are presented in Table 12.

[36]There is also potential for creating additional instrumental variables for exposure to infrastructure. The grid of the sewage was expanding away from the river and the grid of the water works was expanding from the lake. In the end a particular location would most likely get the water or sewer access only after the locations on it's way to river or lake would get the access to infrastructure. This provides a potential for another source of exogenous variation in infrastructure access which is potentially uncorrelated with other unobserved characteristics affecting the land value. Instrumental variables estimation based on this approach gives results similar to the instrumental variables coefficients estimated in this section.

[37]Another way of thinking about this is the number of years the current township has been eligible to join Chicago.

Considering the effect of access to water, the coefficients again are positive and statistically significant. The coefficients are larger than before, but it is not possible to conclude that the coefficients are statistically different with respect to my previous panel regression with fixed effects estimates. Additionally, the coefficients for sewage are closer to the estimate in column 6 of Table 9, but not statistically different from zero.

The causal interpretation of these estimates rely on: (i) the neighbor joining Chicago is independent of any unobserved characteristics related to the own township or plot, and (ii) the fact that the neighbor joins Chicago does not directly affect own land values (exclusion restriction). The plausibility of the first condition was described above. As for the second condition, unfortunately, it is not possible to directly test whether it is satisfied or not (Angrist and Pischke, 2008). One could argue that the fact that a neighboring suburb joined Chicago could have a direct impact on own land value thanks to the expectation of joining Chicago in the future. In case that there were to be a significant persistence affecting the trajectory of land values given by a neighboring suburb joining Chicago, then my IV estimates should be considered as an upper bound. Then the estimates would be a combination of the effects of both infrastructure access and the neighboring suburb joining the City. However, this effect would accrue immediately (or even before the neighbor actually joined because it is expected to join) and should not have a persistent effect on the trajectories of land values over time. Given that my instrument relies mainly on a function of the time since the neighboring suburb joined Chicago, the direct effect of the instrument on land values should not be a large concern.

2.6 Conclusion

This paper estimates the economic returns to access to water and sewage in Chicago in 1850-1930. I use an instrumental variables approach to estimate the causal effect between access to infrastructure and land values. I find that access to water increases land values while the effect of sewage is smaller in magnitude and imprecisely estimated. The effect of increasing water access from 0% to 100% would be equivalent to moving that plot approximately four miles closer to the central business district in 1928.

This paper contributes to research on economic returns to infrastructure. The context of the 19th-20th century is of a particular interest since it is in many ways similar to what is happening in the developing countries: a lot of people are migrating from rural to urban areas in a pursuit of better life. This creates additional challenges for creating a good disease environment and increases demand for infrastructure. To better plan the extend of infrastructure access it is important to understand what potential returns of it are.

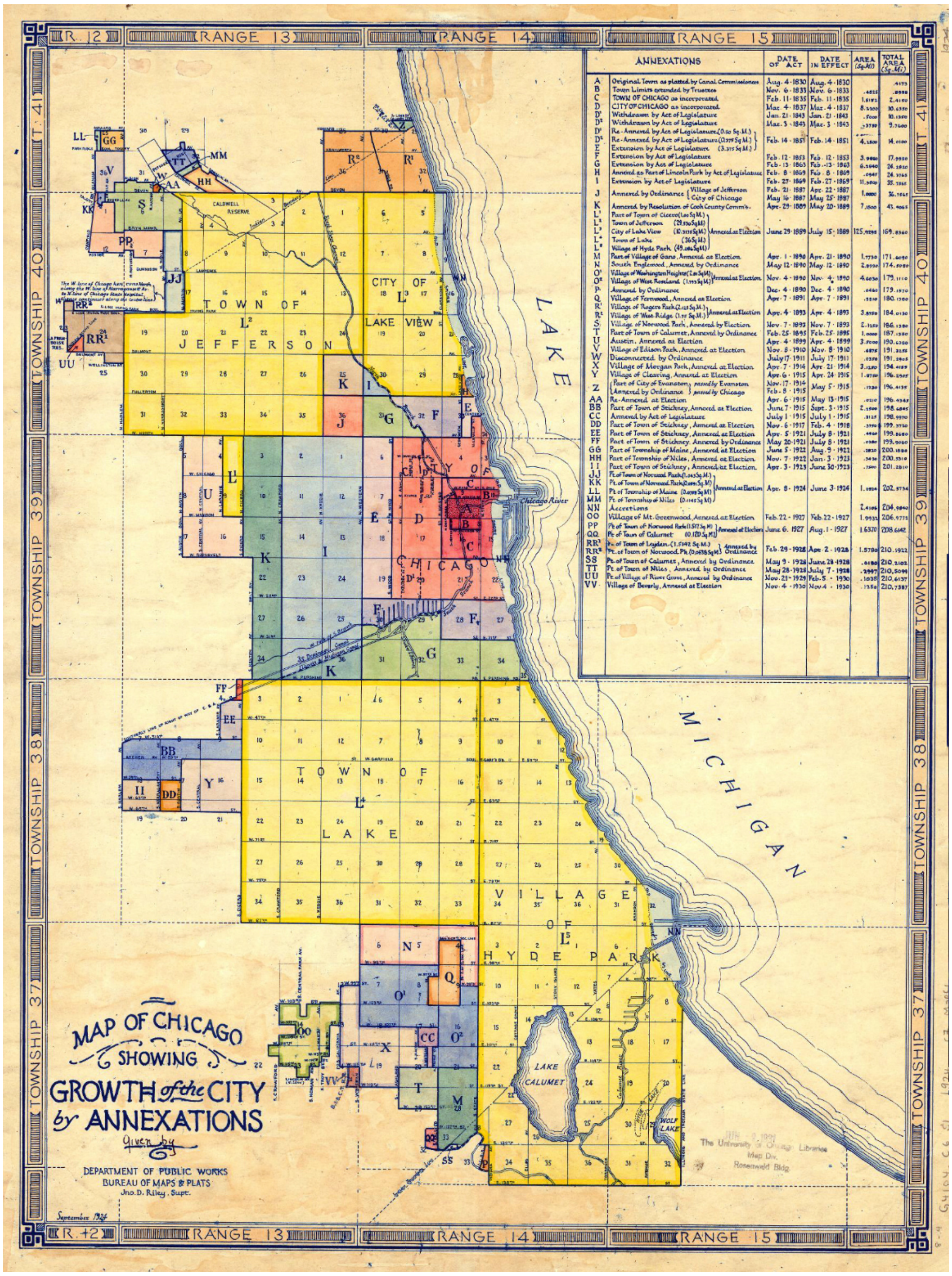
Present cities are shaped by the cities of the past. For example the size of the city and the density gradient is likely to be affected by how much the city was able to expand in the past. Some cities have small core cities and a large ring of suburbs around them (e.g. Boston and Chicago). While other cities have a large core city are and smaller ring of suburbs around (e.g. Austin and Cleveland). The residents that live in the suburbs around the city are likely to use some of the city's amenities while not paying taxes to the central city. Because of it the spatial structure of the city and its suburbs creates significant implications for local public finance. The historic urban growth and annexations are one of the factors that can potentially affect the modern spatial structure of the city. The desire to get access to better and cheaper infrastructure is one of the potential factors affecting the decisions of suburbs to join the city or not, and hence affecting the size of the core city. Modeling how infrastructure and historic urban growth shaped the modern spatial structure of cities is a potential avenue for future research.

The future versions of this paper could include richer set of control variables. For example controls for rail road and cart lines, controls for land use, exposure to Chicago Fire in 1871. Additionally obtaining information on the costs of infrastructure construction would allow to perform a cost benefit analysis. Future work can also make use of a land value data available at a more dissaggregated level (it is available starting from 1913), exploiting a finer variation in access to infrastructure.[38] This finer data will allow to measure the effects more precisely and will allow to draw better conclusion on relative importance of water and sewage.

[38]Currently I do not have access to this data

2.7 Figures and Tables

Figure 5: Map of Annexations

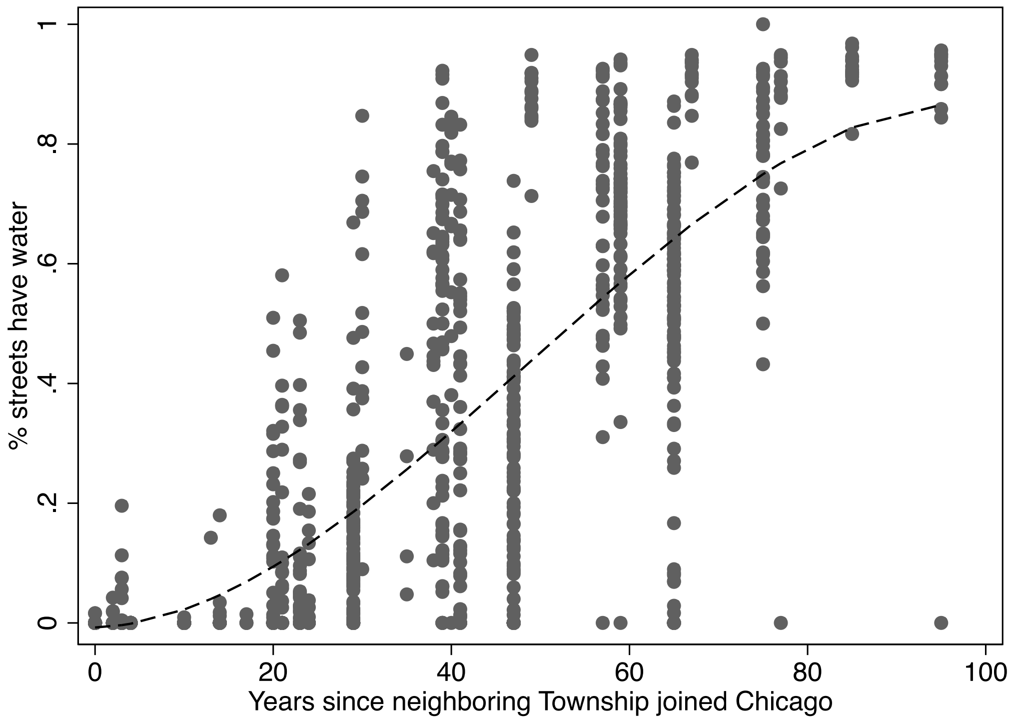


Note: Map obtained from the Chicago Department of Public Works, Bureau of Maps and Plats, 1930. See *https : //chicagology.com/population/*

Figure 6: Water Grid Expansion and City limits

Note: Grey squares represent 1 by 1 mile plots. Red lines denote suburbs that joined Chicago by the year indicated. Blue lines denote water grids at the street-block level by the year indicated. Map complied by author. Plot and township shapefile made by author based on maps from Hoyt and Millis (2000). Water grid shapefile obtained from the Historical Urban Ecological GIS Data Portal.

Figure 7: Water



Note: Grey dots denote pooled observations in sample, dashed line is the best fitting fractional polynomial.

Table 7: Cost of Water Provision in Chicago and Surrounding Townships in 1885-1886

Location	cost per mill gallons	average daily consumption
Chicago	$6	91 mill
Lake View	$12	1.9 mill
Town of Lake	$9	7.2 mill
Hyde Park	$9	3.4 mill

Note: Page 369, Brown (1894).

Table 8: Descriptive Statistics

Year	Mean land value	% Part of Chicago	% Water	% Sewage
1836	924	0.03	0	0
	(6,461)	(0.17)	(0)	(0)
1857	7,248	0.12	0.02	0.00
	(33,894)	(0.33)	(0.09)	(0.04)
1873	21,660	0.20	0.08	0.06
	(69,534)	(0.40)	(0.21)	(0.20)
1892	81,780	0.84	0.21	0.21
	(359,475)	(0.37)	(0.30)	(0.32)
1910	115,944	0.91	0.37	0.38
	(532,874)	(0.29)	(0.31)	(0.32)
1928	131,094	0.98	0.60	0.45
	(431,326)	(0.13)	(0.24)	(0.32)
Nr Townships		46		
Nr Plots		268		

Note: Descriptive statistics for each year of data. Land values represent the average land value per acre for the sample of 1 by 1 square mile plots. All values are expressed in 1967 dollars. % part of Chicago describes the proportion of plots that belong to Chicago. % Water and % Sewage are the proportion of street-blocks in a plot with connection to water or sewage. Standard deviations are presented in parenthesis.

Table 9: Main Results – Panel data with fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.: Log(Land Value)						
% Water	0.433***	0.300*	0.516***			
	(0.138)	(0.166)	(0.142)			
% Sewage				0.0313	0.0389	0.267*
				(0.123)	(0.132)	(0.136)
Observations	1,462	1,462	1,457	1,462	1,462	1,457
R-squared	0.945	0.958	0.951	0.945	0.958	0.951
Number of plots	268	268	267	268	268	267
Plot FE	Yes	Yes	Yes	Yes	Yes	Yes
Dist*Year	Yes	No	Yes	Yes	No	Yes
Dist*I(Year)	No	Yes	No	No	Yes	No
Join*Year	No	No	Yes	No	No	Yes

Note: Each column represents a separate OLS regression including plot-level fixed effects and year dummies. Dist*Year denotes including distance variables (to central business district, lake and river) all interacted with the year variable. Dist*I(Year) denotes including the interaction between distance variables and year dummy variables. Join*Year denotes suburb-group specific time trends; where a suburb-group is defined as the group of suburb joining Chicago in the same time period. Standard errors in parenthesis are clustered at the plot level. Significance levels: *** <1%; ** <5%; * <10%.

Table 10: Robusteness to functional form

	(1)	(2)	(3)	(4)
Dep. Var.: Log(Land Value)				
Water Low	0.327***	0.307***		
	(0.0777)	(0.0744)		
Water High	0.529***	0.494***		
	(0.0964)	(0.111)		
Sewage Low			0.303***	0.410***
			(0.0814)	(0.0846)
Sewage High			0.170*	0.272***
			(0.0912)	(0.102)
Observations	1,462	1,462	1,462	1,462
R-squared	0.947	0.959	0.946	0.959
Number of plots	268	268	268	268
Plot FE	Yes	Yes	Yes	Yes
Dist*Year	Yes	No	Yes	No
Dist*I(Year)	No	Yes	No	Yes

Note: Each column represents a separate OLS regression including plot-level fixed effects and year dummies. Dist*Year denotes including distance variables (to central business district, lake and river) all interacted with the year variable. Dist*I(Year) denotes including the interaction between distance variables and year dummy variables. Low/High variables are dummy variables that equal 1 if infrastructure coverage is larger than 0% and less than 50% for low, more than 50% for high, and the omitted category is none (0%). Standard errors in parenthesis are clustered at the plot level. Significance levels: *** <1%; ** <5%; * <10%.

Table 11: Placebo – Future values of infrastructure

	(1)	(2)	(3)	(4)
Dep. Var.: Log(Land Value)				
Lead % Water	0.154	0.137		
	(0.133)	(0.149)		
Lead % Sewage			-0.152	0.0245
			(0.130)	(0.136)
Observations	1,231	1,227	1,231	1,227
R-squared	0.946	0.963	0.946	0.963
Number of plots	267	266	267	266
Plot FE	Yes	Yes	Yes	Yes
Dist*Year	Yes	No	Yes	No
Dist*I(Year)	No	Yes	No	Yes

Note: Each column represents a separate OLS regression including plot-level fixed effects and year dummies. Dist*Year denotes including distance variables (to central business district, lake and river) all interacted with the year variable. Dist*I(Year) denotes including the interaction between distance variables and year dummy variables. Standard errors in parenthesis are clustered at the plot level. Significance levels: *** <1%; ** <5%; * <10%.

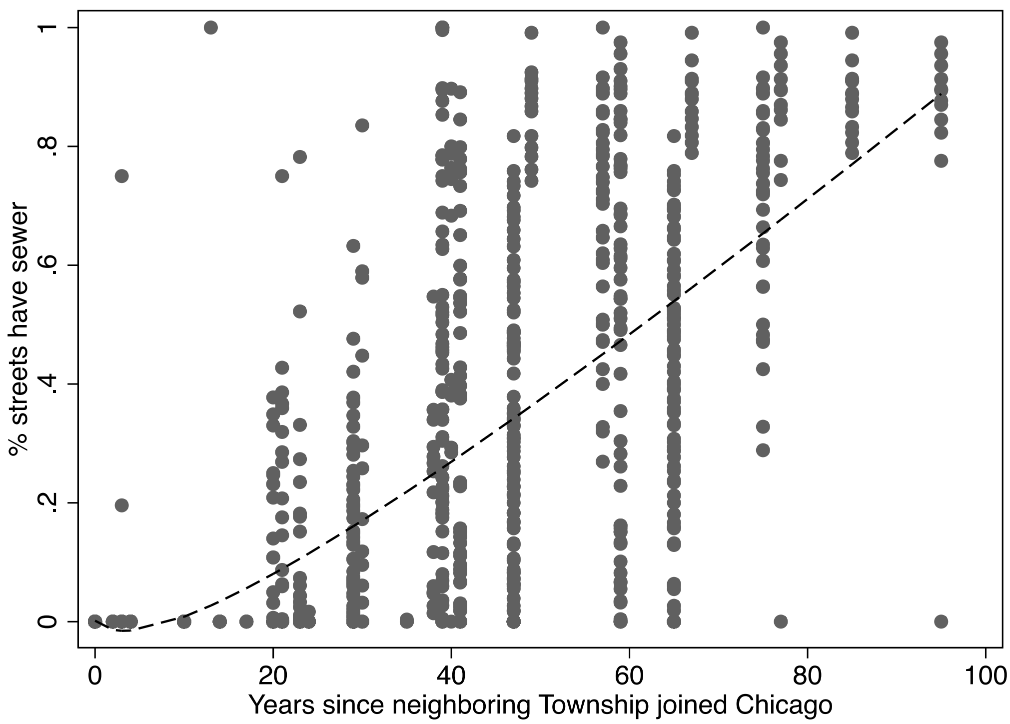
Table 12: Main Results – Instrumental Variables

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.: Log(Land Value)						
% Water	0.800***	0.690***	0.844**			
	(0.274)	(0.266)	(0.377)			
% Sewage				0.370	0.294	0.674
				(0.344)	(0.325)	(0.446)
Observations	1,401	1,401	1,401	1,401	1,401	1,401
R-squared	0.947	0.947	0.960	0.947	0.947	0.959
Number of id	253	253	253	253	253	253
Instrument f(YSNJ)	Cubic	Quartic	Quartic	Cubic	Quartic	Quartic
F-Stat. first stage	116.9	94.64	37.82	61	44.82	18.30
P-Val. first stage	0	0	0	0	0	0
Plot FE	Yes	Yes	Yes	Yes	Yes	Yes
Dist*Year	Yes	Yes	No	Yes	Yes	No
Dist*I(Year)	No	No	Yes	No	No	Yes

Note: Each column represents a separate 2SLS regression including plot-level fixed effects and year dummies. Dist*Year denotes including distance variables (to central business district, lake and river) all interacted with the year variable. Dist*I(Year) denotes including the interaction between distance variables and year dummy variables. Infrastructure variables were instrumented using a dummy variable for whether neighboring suburb joined Chicago and a function of the number of years since the neighboring suburb joined Chicago (YSNJ). This function is specified either as a cubic or quartic function. Standard errors in parenthesis are clustered at the plot level. Significance levels: *** <1%; ** <5%; * <10%.

2.8 Appendix

Figure 8: Sewage



Note: Grey dots denote pooled observations in sample, dashed line is the best fitting fractional polynomial.

Table 13: Instrumental Variables – First Stage

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:		% Water			% Sewage	
Neighbor Joined	-0.093***	-0.040***	-0.031**	-0.087***	-0.043***	-0.025*
	(0.015)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)
Yrs since Neighbor Joined	-0.006***	-0.017***	-0.017***	-0.009***	-0.018***	-0.017***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)
$YSNJ^2$	0.000348***	0.000953***	0.000844***	0.000311***	0.000820***	0.000674***
	(3.85e-05)	(0.000117)	(0.000117)	(3.94e-05)	(0.000137)	(0.000142)
$YSNJ^3$	-3.44e-06***	-1.39e-05***	-1.23e-05***	-2.85e-06***	-1.16e-05***	-9.42e-06***
	(2.75e-07)	(1.93e-06)	(1.92e-06)	(2.82e-07)	(2.23e-06)	(2.30e-06)
$YSNJ^4$		5.69e-08***	5.34e-08***		4.79e-08***	3.95e-08***
		(1.02e-08)	(1.02e-08)		(1.15e-08)	(1.18e-08)
Observations	1,596	1,596	1,596	1,596	1,596	1,596
R-squared	0.815	0.818	0.834	0.723	0.725	0.741
Number of plots	266	266	266	266	266	266
F-Stat. first stage	116.9	94.64	37.82	61	44.82	18.30
P-Val. first stage	0	0	0	0	0	0
Plot FE	Yes	Yes	Yes	Yes	Yes	Yes
Dist*Year	Yes	Yes	No	Yes	Yes	No
Dist*I(Year)	No	No	Yes	No	No	Yes

Note: Each column represents a separate OLS regression including plot-level fixed effects and year dummies. Dist*Year denotes including distance variables (to central business district, lake and river) all interacted with the year variable. Dist*I(Year) denotes including the interaction between distance variables and year dummy variables. YSNJ denotes the number of years since neighboring suburb joined Chicago. Standard errors in parenthesis are clustered at the plot level. Significance levels: *** <1%; ** <5%; * <10%.

References

Albouy, D. (2015). What are cities worth? land rents, local productivity, and the total value of amenities. *Review of Economics and Statistics, Forthcoming.*

Albouy, D., Leibovici, F., and Warman, C. (2013). Quality of life, firm productivity, and the value of amenities across canadian cities. *Canadian Journal of Economics/Revue canadienne d'économique*, 46(2):379–411.

Alsan, M. and Goldin, C. (2015). Watersheds in infant mortality: The role of effective water and sewerage infrastructure, 1880 to 1915. Technical report, National Bureau of Economic Research.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Bar-Ilan, A. and Sacerdote, B. (2004). The response of criminals and noncriminals to fines*. *Journal of Law and Economics*, 47(1):1–17.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy*, 76(2):169–217.

Becker, G. S. and Stigler, G. J. (1974). Law enforcement, malfeasance, and compensation of enforces. *J. Legal Stud.*, 3:1.

Bowles, R. and Garoupa, N. (1997). Casual police corruption and the economics of crime. *International Review of Law and Economics*, 17(1):75–87.

Brown, G. (1894). Drainage channel and waterway. *Chicago: RR Donnelley and Sons Co.*

Cellini, S. R., Ferreira, F., and Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, 125(1):215–261.

Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century united states. *Demography*, 42(1):1–22.

Ferrie, J. P. and Troesken, W. (2008). Water and chicago's mortality transition, 1850-1925. *Explorations in Economic History*, 45(1):1–16.

Garoupa, N. and Klerman, D. (2004). Corruption and the optimal use of nonmonetary sanctions. *International Review of Law and Economics*, 24(2):219–225.

Haughwout, A. F. (2002). Public infrastructure investments, productivity and welfare in fixed geographic areas. *Journal of public economics*, 83(3):405–428.

Hornbeck, R. and Keniston, D. (2014). Creative destruction: Barriers to urban growth and the great boston fire of 1872. Technical report, National Bureau of Economic Research.

Hoyt, H. and Millis, H. A. (2000). *One Hundred Years of Land Values in Chicago: The Relationship of the Growth of Chicago to the Rise of Its Land Values, 1830-1933.* Beard Books.

Mayer, H. M. and Wade, R. (1974). *Chicago: Growth of a Metropolis.* University of Chicago Press.

Polinsky, A. M. and Shavell, S. (1979). The optimal tradeoff between the probability and magnitude of fines. *The American Economic Review*, 69(5):880–891.

Polinsky, A. M. and Shavell, S. (1984). The optimal use of fines and imprisonment. *Journal of Public Economics*, 24(1):89–99.

Polinsky, A. M. and Shavell, S. (1991). A note on optimal fines when wealth varies among individuals. *The American Economic Review*, pages 618–621.

Polinsky, A. M. and Shavell, S. (2001). Corruption and optimal law enforcement. *Journal of Public Economics*, 81(1):1–24.

Polinsky, A. M. and Shavell, S. (2007). The theory of public enforcement of law. *Handbook of law and economics*, 1:403–454.

Sequeira, S. (2012). *Advances in measuring corruption in the field*, volume 15. Emerald Group Publishing Limited.

Shavell, S. (1987). The optimal use of nonmonetary sanctions as a deterrent. *The American Economic Review*, pages 584–592.

Siodla, J. (2014). Making the move: The impact of the 1906 san francisco disaster on firm relocations. Technical report, Colby College. Mimeo.

Spinney, R. G. (2000). *City of Big Shoulders: A History of Chicago.* Northern Illinois University Press; 1st edition.

Tourangeau, R. and Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, 133(5):859.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., and Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science*, 280(5365):867–873.