Leniency, Asymmetric Punishment and Corruption
EVIDENCE FROM CHINA

Maria Perrotta Berlin∗ Giancarlo Spagnolo†
August 10, 2016

Abstract

One-sided leniency policies and asymmetric punishment are regarded as potentially powerful anti-corruption tools, also in the light of their success in busting price-fixing cartels. It has been argued, however, that the introduction of these policies in China in 1997 has not helped fighting corruption. Following up on this view, the Central Committee of the Chinese Communist Party passed in November 2015 a reform to the Criminal Code, bringing to its peak the government’s years-long commitment to fight corruption. The reform introduced heavier penalties but also severe restrictions of the leniency offered to bribe-givers. Claims on the effects of the 1997 reform are not backed by data, to our knowledge, while evaluating the effects of a policy on crimes like corruption is difficult. These crimes are typically only observed if detected and convicted by the police, and an increase in observed convictions may as well be due to an increase in the total number of crimes rather than to a positive effect of the policy. We collect data on the investigations of bribery and public official corruption, available for most Chinese provinces for the period 1986-2010, and extend to corruption a method to identify deterrence effects from changes in detected cases, originally developed for cartels. The available evidence so far points to a substantial and stable reduction in the number of major corruption cases around the 1997 reform, a result per se ambiguous but clearly consistent with a positive deterrence effect of the 1997 reform. A case study analysis is under way to corroborate and help the interpretation of these preliminary findings.

1 Introduction

Corruption remains an endemic problem in the developing world and has become a central political issue in emerging countries like India, Brazil and China. This paper focuses on a specific approach to the fight of corruption, being lenient with one party to induce them to denounce the other corrupt parties. As other forms of organized crime, corruption requires cooperation between two or more

∗SITE
†SITE, University of Rome "Tor Vergata", Eief, CEPR. We own a debt of gratitude to Bei Qin, University of Hong Kong, for sharing the Procuratorates’ reports and to Ronald Ren, Associate Senior Counsel at The Linde Group, for substantial feedback in the initial phases of this work. We thank Kaushik Basu for valuable discussions. Erika Gyllström, Martin Rassl and Shangqiu Xu provided excellent research assistance. We also benefited from comments during presentations at the ASWEDE Workshop, the SITE Academic Conference and the FREE Network Retreat. We acknowledge the Jan Wallanders and Tom Hedelius Research Foundations for supporting this work. All remaining errors are our own.
informed parties, so that there is always a witness whose information can be retrieved by suitably structured legal incentives. The story of the Prisoner’s Dilemma is a classic example of leniency - a reduction or cancellation of legal sanctions to a wrongdoer that betrays his partner reporting their illegal behaviour. Formal and informal exchanges of leniency against information and collaboration are a normal feature of law enforcement in most countries. In particular they have been extensively and quite successfully used to fight mafia in the US and Italy, drug dealing and organised crime. Since the US reform in 1993, structured leniency programs, offering immunity to the first party that reports the information (and possibly less generous forms of leniency for other reporting parties) have become Competition authorities’ main instrument to fight cartels and bid rigging.\footnote{See Spagnolo (2008).} The possibility to use leniency to play one part against the other also in the fight against corruption has been at the center of a recent intense policy debate after the popular note, Why, for a Class of Bribes, the Act of Giving a Bribe Should Be Treated as Legal (Basu, 2011). Then chief economist of the Indian government and now of the World Bank, Kaushik Basu advocated asymmetric depenalization, which can be thought of as a form of unconditional, one-sided leniency. More precisely, the note proposed, for one particular type of bribes - harassment bribes (also called extortionary or discharge-of-duty bribes), paid to obtain something one is entitled to - to make bribe-giving legal, while strengthening sanctions against bribe-takers. As for other forms of leniency, the idea is to create a conflict of interests between the partners in crime by tweaking their incentives. One party (in this case the bribe-giver, in the general case the first one who applies for leniency) can now betray and report the illegal act in order to obtain the benefit of the lenient treatment, no sanctions and the restitution of the bribe. In the debate sparked by this note many different arguments have been put forward, both against it and in favor of it. Then a blogpost by a Chinese law scholar, Li (2012), attracted our attention to the case of China, where asymmetric punishment (bribe-giver impunity) has been in place since 1997. She argued, probably reflecting the political debate in the country rather than based on factual evidence, that the system had not been successful. We felt this claim granted a deeper investigation into the details of the Chinese legal reform and the changes it introduced, and of course a careful inspection of the data to back it.

Further motivation for this study comes from the current events in China. Chinese President Xi Jinping since coming into office famously vowed to crack down on both “tigers” and “flies” - powerful leaders and lowly bureaucrats - who engage in corrupt activities. For the past two years, Mr. Xi has carried out a sweeping, highly publicized anticorruption campaign. Even a brand new website (www.ccdi.gov.cn) was launched recently with a handy online feature for reporting corruption, anonymously or not. Most importantly, a reform to the Criminal Law, known as Amendment IX, was proposed in October 2014, then voted in August 2015 and is in force since November 1st, 2015. Among other measures, heavier penalties are envisaged, and severe restrictions of leniency, for those offering bribes. In this study we hope to shed some light onto why the chinese leadership has been so dissatisfied with the anti-corruption legislation as it was since 1997 and on the likely effects of the newly approved changes. In the next Section we review the literature most closely related. In Section 3, we offer a short summary of the evolution of the Chinese anti-corruption legislation. In the rest of the paper, we bring the reform to the data, described in Section 4; the statistical tests in Section 5 suggest a success of the 1997 leniency introduction in reducing corruption cases. However, we acknowledge all the limitations in these data, and propose a way forward. We plan to go beyond aggregate data and look at a sample of cases, as detailed in Section 6, in order to understand which details of the legislation mattered and through which channels. This data collection is of itself another contribution of this study. The paper concludes on a hopeful note.
2 Literature review

The rich theoretical literature sparked by the introduction of leniency programs in antitrust has shown that that these tools can be extremely powerful in deterring collaborative crimes like cartels and corruption\(^2\). For example, Spagnolo (2004) shows that leniency generated an additional deterrence effect operating through “distrust”, and that if fines are severe, the first best of full deterrence can be obtained offering a reward to the first self-reporting party financed by the fines paid by the remaining parties. However, the same literature also showed that these programs can easily be manipulated or misused, becoming counterproductive, so that success depends on the specific details of their design and implementaton (Motta and Polo, 2003; Spagnolo, 2000).

Although Spagnolo (2004) does discuss applications to corruption, specific theoretical analysis of one-sided leniency and corruption starts with Buccirossi and Spagnolo (2006). This paper develops a model of occasional and repeated corrupt transactions emphasizing that corrupt exchanges - not being enforceable by law - expose parties to the risk of hold-up or “double crossing”. A long-term, repeated relationship helps govern these problems, but occasional corrupt deals may be normally unfeasible. It is then shown that the asymmetry in legal punishment linked to one-sided leniency programs could be exploited by wrongdoers to solve the risk of hold-up, making occasional corrupt deals viable. When the incentives generated by these programs are strong enough, e.g. when a reward is paid to the reporting agent, they have a robust deterrence effect on long-term corrupt relationships (and on occasional ones that do not suffer the risk of hold-up). However, when not properly designed, these policies may provide an effective governance mechanism for occasional sequential illegal transactions that would not be feasible in its absence: parties can then structure their corrupt exchange so that the party subject to the risk of hold-up would report to obtain leniency if actually held up, but not if the corrupt transaction is completed. Lambsdorff and Nell (2007) use the static version of the corruption game developed in Buccirossi and Spagnolo (2006) to consider the possibility that different fines are imposed for the acts of paying a bribe, receiving a bribe, giving an illegal advantage (the reason for the bribe), and receiving the illegal advantage. While the complex prescriptions they derive appear to depend on the specific timing assumed and to be relevant to occasional deals only, they confirm the finding in Buccirossi and Spagnolo (2006) that sufficiently strong incentives must be provided for leniency to be effective.

These analyses focus on collusive corruption, where bribes are exchanged against an illegal advantage that produces a distortion, and take into account the risk of hold-up in the corrupt exchange. The 2011 note by Basu, instead, carefully circumscribes the proposal to bribes paid to obtain a service one is legally entitled to, and focuses on situation were the exchange is simultaneous, so that no risk of hold-up is present. The proposal is analyzed in a formal model that maintains Basu’s focus and assumptions in Dufwenberg and Spagnolo (2015). This model also tries to take into account some of the critiques that Basu’s proposal drew during the policy debate, such as that the fear of retaliation could hamper the mechanism, and that moral concerns could arise from the legalization of bribe-paying with possibly counterproductive effect on the frequency of corruption (Dreze, 2011). In Dufwenberg and Spagnolo (2015) it is shown that Basu’s proposal fares poorly in some situations and very well in others, and in particular in those where the bureaucracy and law enforcement institutions are not too inefficient or corrupt, so that reporting costs are low relative to the bribe and there is limited risk of retaliation. This suggests that these mechanisms are better suited to fight serious corruption, where large bribes are at stake. It is also shown that

\(^2\)See e.g. Motta and Polo (2003); Spagnolo (2004); Aubert et al. (2006); Harrington (2008, 2013); Chen and Rey (2013).
modifying the proposal by making immunity conditional on having reported the bribe solves several
of the problems raised in the policy debate, and that the mechanism could be effective even against
collusive corruption, as long as the bribe-givers can be compensated for the loss of the distortive
favor in case they report the corrupt deal. More analysis of the proposal are being advanced. For
example, a recent paper by Oak (2015) considers the possibility that deterring harassment bribes
could lead bureaucrats to increase the amount of distortive corruption, while Basu et al. (2014) focus
on the bargaining game between bribe-giver and bribe-taker and the risk that Basu’s proposal could
increase the size of the bribes while deterring their occurrence.

The literature on leniency in antitrust has taught us that it is very difficult to evaluate empirically
the success of these policies against crimes like corruption and collusion, as only changes in
discovered and convicted cases are typically observed, not in their overall number (Spagnolo,
2008). Indirect methods have been developed to estimate the deterrence effects of these new poli-
cies (Miller, 2009; Harrington and Chang, 2009). By and large, the available evidence supports
the theoretical conclusion that leniency tends to be effective in deterring cartels when accompanied
by strong sanctions, as in the US (Miller, 2009), but much less when sanctions are lower, as in
the EU (Brenner, 2009). Our paper is particularly related to Miller (2009), because we adapt his
identification method developed for long-term price-fixing cartels to the case of more short-term
corrupt exchanges.

Laboratory experiments are particularly valuable to study collusive and white-collar crimes,
as they allow to observe the overall population of infringements, not only the ones discovered,
as in reality, and have by and large confirmed the potential together with the subtlety of these
instruments (Apesteguía et al., 2007; Hinloopen and Soetevent, 2008; Engel et al., 2012). The
experiment by Bigoni et al. (2012) finds, among other things, that deterrence is very strong and
theoretical predictions are approximated only when these schemes allow for a reward to the party
blowing the whistle, as suggested in Spagnolo (2004), a result partly confirmed by the more recent
experiment by Wu and Abbink (2013). When rewards are not allowed, Abbink et al. (2014) find that
the effectiveness of asymmetric punishment depends on the environment, and is strongly dependent
on the (im-)possibility of retaliation by the bribe-taker. Bigoni et al. (2015) focus on the size of
the sanction, and show that one-sided leniency has a tremendous deterrence effect as long as it is
accompanied by sufficiently severe sanctions. Then the probability of independent detection (i.e.
independently from reports by the agents) becomes irrelevant to deterrence, because subjects are
entirely focused on the risk of being betrayed and the large fines this would imply.

3 Anti-bribery legislation in China and the 1997 reform

The major statutes in Chinese anti-bribery legislation are the Criminal Law of the People’s Republic
of China (CL)\(^3\) and the Anti-Unfair Competition Law of the People’s Republic of China (AUCL)\(^4\).
In this paper we will focus on corruption offences investigated and prosecuted under the CL, which

\(^3\)Available in translation in AUCL (1993). Additional sources include: the Interim Provisions on Banning Com-
nercial Bribery, issued by the State Administration for Industry and Commerce of the People’s Republic of China
on November 15th, 1996; the Interim Measures of Hubei Province on Prevention and Administration of Commercial
Bribery in Engineering Construction Fields, issued by People’s Government of Hubei Province on July 11th, 2007; the
Supplementary Provisions of the Standing Committee of the National People’s Congress Concerning the Punishment
of the Crimes of Embezzlement and Bribery, issued by the Standing Committee of the National People’s Congress
on January 21th, 1988 and abolished pursuant to the Criminal Law of the People’s Republic of China promulgated
by the National People’s Congress on March 14, 1997.


covers all public official corruption. The AUCL was introduced in 1993 to address bribery by private sector managers as part of a set of practices that distort competition. If the offence does not violate the CL, the punishment under the AUCL is a fine between 10,000 and 200,000 RMB, plus confiscation of the illegal gain. Art. 22 of the AUCL states explicitly that those guilty of bribery should be investigated and punished in accordance with the CL whenever applicable.

The CL was adopted during the Second Session of the Fifth National People’s Congress on July 1st, 1979 and revised during the Fifth Session of the Eighth National People’s Congress on October 1st, 1997. This revision is a major reform, and constitutes the focus of this study. In the 1979 text, both the crimes of paying and accepting bribes are defined in one single article (Art. 185). Both definitions involve state personnel on at least one end of the corrupt deal. The punishment is slightly more lenient for active bribery: offering bribes could be punished by up to three years imprisonment, while accepting bribes was punishable by up to five years, or more than five in presence of serious losses for the public. Active bribery in the context of elections was also punished to the same extent (Art. 142).

The revised text of the CL promulgated in 1997 is much richer in details than the previous version. The crimes of accepting and paying bribes involving state functionaries, state organs or non-state functionaries are defined and regulated in Chapter VIII. The use of bribery in other contexts is also mentioned in Chapters III, IV and VI regarding the private sector ("Crimes of Disrupting the Order of Administration of Companies and Enterprises"), the electoral context ("Crimes of Infringing upon Citizens’ Rights and Democratic Rights") and the judicial context ("Crimes of Impairing Judicial Administration") respectively.

Between those two versions, the definitions of active and passive bribery and the associated punishments were extensively changed in 1988 by the Standard Committee of the National People’s Congress (the only institution that has the right to revise laws in China), in an official document called Supplementary Provisions of the Standing Committee of the National People’s Congress Concerning the Punishment of the Crimes of Embezzlement and Bribery. Such a document has legal effect, but lower status that the CL. In this text, different levels of punishment are specified in a schedule according to the seriousness of the circumstances, see Tables 1 and 2. Moreover, two important details are added to the discipline. The first one is the introduction of leniency (mitigated punishment or exemption from punishment) for those who confess voluntarily before being investigated. Previously there existed only a generic provision for leniency within the legal system, not specific to the crimes of corruption and bribery. It is noteworthy that there is an asymmetry in the eligibility to leniency: bribe-takers are only eligible if the size of the bribe is below a given threshold, while there is no such limitation for the bribe-giver, see Table 3.

The second one is the introduction of asymmetric punishment. The crime of giving a bribe is now associated with the intent “to secure improper benefits”. This means that a briber either: (1) seeks benefits that are in violation of law, regulations, rules, or state policies; or (2) seeks benefits that are themselves legitimate, but are to be obtained by means of violating laws, regulations, rules, state policies, or industrial norms. Although in practice different judicial authorities have different interpretations on the definition of improper benefit and its importance, and it has never been treated as an absolute prerequisite for a prosecution or conviction on count of bribery (Gintel, 2013; 5).

Note of the Supreme People’s Court and the Supreme People’s Procuratorate on Issuing the Opinions on Issues Concerning the Application of Law in the Handling of Criminal Cases of Commercial Briberies, promulgated by the Supreme People’s Court and the Supreme People’s Procuratorate, 2008, as reported in Gintel (2013).
Table 1: Punishment schedule for bribe-takers

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>p &lt;= 5 or criminal detention plus confiscation of property; p &gt;= 5 if serious losses</td>
<td>p &lt;= 2 or criminal detention; administrative sanctions if not serious</td>
<td>2,000 &lt;= b &lt;= 10,000</td>
<td>5,000 &lt;= b &lt;= 50,000</td>
<td>p &lt;= 3 or criminal detention plus fine; administrative sanctions if not serious</td>
<td>“Relatively large amount”</td>
</tr>
<tr>
<td>1 &lt;= p &lt;= 7; 7 &lt;= p &lt;= 10 if serious</td>
<td>2,000 &lt;= b &lt;= 10,000</td>
<td>5,000 &lt;= b &lt;= 50,000</td>
<td>p &gt;= 10 or life imprisonment plus fine plus confiscation of property; life imprisonment or death, plus fine plus confiscation of property, if serious losses</td>
<td>“Especially huge amount”</td>
<td></td>
</tr>
<tr>
<td>p &gt;= 5 plus confiscation of property; life imprisonment if serious</td>
<td>10,000 &lt;= b &lt;= 50,000</td>
<td>50,000 &lt;= b &lt;= 100,000</td>
<td>p &gt;= 10 or life imprisonment plus fine plus confiscation of property; life imprisonment or death, plus fine plus confiscation of property, if serious losses</td>
<td>“Especially huge amount”</td>
<td></td>
</tr>
</tbody>
</table>

Notes: p = imprisonment, in years; b = size of bribe, in yuan

Table 2: Punishment schedule for bribe-givers

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>p &lt;= 3 or criminal detention</td>
<td>p &lt;= 5 or criminal detention</td>
<td>p &lt;= 5 or criminal detention</td>
<td>p &lt;= 5 or criminal detention plus fine</td>
</tr>
<tr>
<td>Serious circumstances or heavy losses to the public</td>
<td>p &gt;= 5</td>
<td>5 &lt;= p &lt;= 10</td>
<td>5 &lt;= p &lt;= 10 plus fine</td>
<td></td>
</tr>
<tr>
<td>Especially serious circumstances</td>
<td>Life imprisonment plus ev. confiscation of property</td>
<td>p &gt;= 10 or life plus ev. confiscation of property</td>
<td>p &gt;= 10 or life plus ev. confiscation of property</td>
<td></td>
</tr>
</tbody>
</table>

Notes: p = imprisonment, in years
Tanzhuhua, 2011), with a strict literal interpretation, this implies a differentiation in the treatment of extortionary bribes (or harassment bribes), those that do not procure improper benefit but are exchanged for something the giver had right to\(^7\). Under the new legislation, this bribe-giver is not considered guilty. This type of asymmetric punishment has been the object of recent debate, as mentioned in the introduction\(^8\).

The 1997 revision of the CL retains most of the formulations of this 1988 text, although the schedule of punishments is revised in a way that makes punishment less severe (see Table 1). Notice therefore that asymmetric punishment, whether practically relevant or not, was introduced already in 1988 and not in 1997, as claimed in Li (2012). At the same time, the CL has stronger status than the 1988 Supplementary Provisions, which might be interpreted as a reinforcement of this provision.

To sum up, in the 1997 reform two new elements were given strongest legal status: the possibility of leniency and the asymmetric punishment for the case of extortionary bribery. To what extent they have been used in practice remains to be investigated. Concurrently, penalties are by and large increased in 1988 and decreased in 1997, in particular for bribe-takers.

On 23 October 2014, the Central Committee of the Chinese Communist Party passed a Decision concerning Several Major Issues in Comprehensively Advancing Governance According to Law which stressed a commitment to “accelerate State legislation against corruption, perfect systems to punish and prevent corruption, create effective mechanisms so no-one dares to be corrupt, can be corrupt and wants to be corrupt, persist in containing and preventing the phenomenon of corruption. Perfect criminal law systems to punish venality and bribes, broaden the scope of criminal bribery from assets to assets and other property-type interests.”\(^9\) A draft amendment to the Criminal Law (Amendment IX) was also submitted to the NPC Standing Committee in October, and subsequently voted on 29 August, 2015 and in force since 1 November 2015. Heavier penalties are envisaged for those accepting bribes, however the thresholds are made more discretionary, as reported in Table 1\(^10\). Penalties for bribe-givers are kept the same but compounded by fines. What’s more, Amendment IX provides severe restrictions to leniency, as shown in Table 3.

---

\(^7\)The distinction applies to the two situations in which the public official takes the bribe to perform what is her duty (for example, produce a licence the bribe-giver is qualified for) or rather to perform an act in violation of her duty (for example, award a public contract to the bribe-giver); from the point of view of the bribe-giver, in the two situations he would pay for something that is in his right to obtain, or rather something that he has not right to. The first type of bribe is also referred to as extortion.

\(^8\)See Basu (2011) and Li (2012).


\(^10\)Media coverage in English at ChinaDaily (2015) and CCTV (2014). An incomplete translation of the draft can be obtained from the authors upon request.
Table 3: Conditions for leniency

<table>
<thead>
<tr>
<th></th>
<th>Bribe-giver</th>
<th>Bribe-taker</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exemption from punishment</td>
<td>Exemption from punishment (only administrative sanctions)</td>
</tr>
<tr>
<td>1988</td>
<td>Confess prior to investigation</td>
<td>b $\leq$ 5,000 plus confession plus repentance plus restitution</td>
</tr>
<tr>
<td>1997</td>
<td>Confess prior to investigation</td>
<td>b $\leq$ 10,000 plus repentance plus restitution</td>
</tr>
<tr>
<td>2014</td>
<td>Confess prior to investigation plus minor circumstances or critical role of</td>
<td>Confess prior to investigation</td>
</tr>
<tr>
<td></td>
<td>confession or major meritorious service</td>
<td>Large amount plus confession, repentance and restitution before prosecution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Huge amount plus confession, repentance and restitution before prosecution</td>
</tr>
</tbody>
</table>

4 Data

Data on the prevalence of bribery are notoriously hard to come by, because of the secretive nature of the activity. We use several data sources which capture on the one hand actual corruption cases tried in courts and on the other hand surveys of corruption perceptions. Records of actual cases are published by the National Bureau of Statistics China, and report in particular:

- the number of arrests and public prosecutions on suspicion of corruption and bribery\(^\text{11}\) for the period 1998-2010\(^\text{12}\);
- the number of bribery cases accepted by the court, registered and settled in the period 1998-2010\(^\text{13}\);
- and the number of first trials (we disregard the appeals) for corruption and bribery accepted and settled by courts in the period 1999-2010\(^\text{14}\).

For the period prior to 1998, for which the records are not published online, we have accessed the original source in printed version\(^\text{15}\). We collected the corresponding information from the Procuratorates’ Yearly Reports for each of the Chinese provinces since 1986. Reports are available for almost all provinces up to 1995, after which the number of provinces reporting falls sharply. This possibly reflects the switch to electronic reporting. Figure 1 shows the time series of prosecutions

\(^{11}\)Offences of Corruption and Bribery", under “Arrests of Criminal Suspects and Defendants Under Public Prosecution Approved by Procurator’s Offices”.
\(^{12}\)Data missing for 2003.
\(^{13}\)Of cases under direct investigation by Procurator’s offices. Data missing for 2003.
\(^{14}\)Data missing for 2003.
\(^{15}\)We thank Bei Qin, University of Hong Kong, for providing us with access to the reports.
Figure 1: Prosecutions for corruption and bribery

for the counts of corruption and bribery from these two sources, with the number of provinces included in each data point. The red vertical line highlights the date of the reform.

If one expects missing reports across provinces to be random, the sum of observations should be rather stable. As shown in the right panel of Figure 2, this is not the case. Especially for the years surrounding the reform in 1997, the number of observations, i.e. the number of provinces which reported data, is fluctuating\textsuperscript{16}.

We therefore consider the two series as complementary. In particular, the exclusive source of data for the period preceding the reform is the province-level data, which cover a subset of provinces.

\textsuperscript{16}Also, there are only one and three reporting provinces for the years 1986 and 1987, respectively. We therefore drop the data for those two years and base the analysis on the time frame 1988-2010.
We are aware that we are not able to observe the exact same set of provinces for the period after the reform. However we have for these two sources, namely both the provincial and the national data, which can be considered respectively a lower bound (from a somewhat smaller set of provinces) and an upper bound (from the national level series). For robustness checks we also restrict the sample to the subset of 6 provinces that report every year.

The dependent variable for the subsequent analysis is expressed as number of cases per 1 million citizens. Having normalized the data by the population size, the remaining variation across the provinces (right panel of Figure 3) can be due to differences in reporting efforts or other province-specific characteristics (for example, level of economic activity), but could also reflect purely random variation.

![Figure 3: Average number of cases per province without (left) and with (right) controlling for population size](image)

These data present us with an inference problem: they refer only to the reported cases of corruption and bribery, while the pervasiveness of the actual crimes remains unobserved. In the next section, we follow methods and statistical tests applied in other branches of literature (for example Miller (2009) on leniency and cartels) to make inference on trends in (unobserved) actual cases.

Because of the shortcomings of these data, we complement this analysis in two ways. First of all, we look at other available indicators of corruption, widely used indices based on expert assessments and opinion surveys. They are described in Appendix B. More interestingly, we also collect microdata from a sample of case-files, documents and proceedings from trials from before and after the reform. We describe this in details in section 6.

5 Structural break tests

The data on prosecutions mix together corruption and anticorruption activities, as they fail to distinguish occurrence of the criminal activity from detection. A policy that deters crimes but at the same time increases the fraction of those that are successfully prosecuted, will have an ambiguous effect on the number of prosecutions. Often possible deterrence effects of a policy
change are forgotten or disregarded in the public debate, as if the underlying unobserved criminal behavior was assumed to be always constant. Observing fewer detected cases under these implicit assumptions might erroneously lead to the conclusion that law enforcement became less effective under the new law. Taking into account the possibility that the new law might induce deterrence, hence fewer occurrences of the crime in the first place, completely changes the perspective.

Miller (2009) develops a theoretical model that, in the context of collusion, helps to bridge observed and unobserved criminal behavior. The model features a first-order Markov process governing the occurrence of criminal activity (cartel formation, in this case) and derives predictions for how changes in the rate of occurrence and the rate of detection affect the time series of detection. This is then applied to test the effect of the introduction of leniency on cartel formation and detection rates.

Similarly to collusive behavior leading to cartel formation, bribery is also based on trust between the corrupt partners. And leniency similarly may undermine this trust (Bigoni et al., 2015) leading to deterrence effects. We exploit this similarity in the two types of criminal activity to adapt the theoretical results and empirical tests developed by Miller (2009) to the case of the anti-bribery reforms. While cartels are long-term agreements, though, we can think of bribe exchanges as one-shot interactions.

The main results from Miller’s theoretical model are summarized as follows:

**RESULT 1:** The immediate increase in the number of prosecutions after a reform is sufficient to establish a corresponding increase in the detection rate.

**RESULT 2:** The subsequent readjustment of the number of prosecutions below initial levels is sufficient to establish a decrease in the underlying criminal activity (deterrence effect).

Based on this, Miller (2009) expects a peak in discoveries after the reform due to defection of existing cartels’ members, followed by a slump, revealing less cartel formation. Given the instantaneous nature of bribery interactions, we should not expect a separate detection effect of the same kind. On the other hand, the 1997 reform is retroactive, implying that leniency becomes available also for bribe exchanges that have happened before 1997. This could potentially still lead to a peak in discoveries immediately after the reform.

The bar graph in Figure 4, showing the number of cases per 1 million citizens from 1988 until 2010, yields some first insights. The average is relatively high in the first ten years of data and exhibits some time variation, before it experiences a major drop in the year 1998, coinciding with the implementation of the reform in 1997. The average levels off in this low state in the subsequent years.

The box plots in Figure 5 group together observations from before and after the reform, illustrating a difference-in-means test. The left panel uses only the province-level data. Both the median - indicated by the red line - and the variance - indicated by the edges of the box, the 25th and 75th percentile - are considerably lower after the reform. In line with this graphical observation, the two-sample t-test of equal means rejects the null hypothesis of equal means (and equal but unknown variances) at any common significance level.

Since the province-level data has less observations after the reform, we also use the available national data for the post-treatment period as a robustness check. Note that the absolute number of reported cases on the national level is weighed by the national population size to obtain a measure comparable with the province-level data. The box plot in the right panel of Figure 5 yields similar results. The two-sample t-test of equal means again rejects the null hypothesis at any common significance level.
After these first graphical observations and mean tests, we now turn to regression analysis to quantify the effect of the reform. Figure 6 plots the dependent variable, the number of cases per one million citizens (including in red national-level data where available).

In Column (1) of Table 4, we regress the dependent variable only on the reform dummy, which takes the value one for the years after 1997 and zero otherwise. The legal reform resulted in 23.66 less cases per one million citizens, corresponding to a $\frac{23.66}{37.95} \times 100\% = 62.3\%$ decrease.

In order to take into account potential trends over time, the model is augmented in the other columns to include polynomials of different orders in two separate time trends, one for the whole sample and one for the post-reform period. Specifically, the term TIME1 equals one in the first period of the sample (year 1988), two in the second period, and so on. The variable TIME2 equals one in the first period following the reform (year 1998), two in the next period, and so on. The coefficient on the reform dummy, which measures the treatment effect, remains close to 30% and
statistically significant in all but one case: column (4), which includes a second order polynomial in both TIME1 and TIME2. This is probably due to the neat fit of the polynomial, as illustrated in the bottom-left panel of Figure 11 in Appendix A. As one would expect when looking at Figure 6, using the national-level data for the post-reform period yields almost identical results (see Appendix A).

Figure 7 plots the estimated conditional means (i.e. predicted values) for the regression shown in Table 4, column (3), along with the actual number of cases. Similar plots for the other models are reported in Appendix A. The polynomial approximations are very flat in the post-reform period, at a level nowhere near the pre-reform one.

In Appendix A we submit this main result to a battery of robustness checks including Poisson and Negative Binomial models, a subset of data and placebo tests. Our conclusion identifying a drop in prosecutions in 1997 by about 30% stands. This pattern is consistent with a deterrence effect of the 1997 reform, but unfortunately also with alternative interpretations. While a peak-and-slump pattern as in Miller (2009) would have been strong evidence supporting the success of the reform in deterring corruption, we cannot exclude that the drop in prosecutions is due to a general worsening in detection. We deem this, however, unlikely in the light of the general political climate of the time.

We collected supportive qualitative evidence looking at official documents and public speeches, but also press releases and newspapers. In the aftermath of the Tiananmen Square protests, anti-corruption had become a major political objective. Deng Xiaoping’s south talk in the spring of 1992 is regarded as an influential reference point in this respect. After this speech, the China Daily has been reporting news related to anti-corruption on a daily basis for the following three years. While the seeds of the 1997 reform certainly emerged already at this point, we found no evidence of any changes either upwards or downwards in the focus on corruption after the effort of this period. Figure 1 reports the count of words related to corruption or bribery in an important official speech given by the president each year. We can interpret this as the intensity of political commitment and general attention to the problem of corruption. In both plots we see a somewhat heightened

---

The inclusion of other combinations and higher polynomial orders was also tested. The treatment effect remains statistically significant with values in the same range.
Table 4: OLS Regression Results

Dependent variable: Number of cases (Prosecutions for corruption and bribery)

<table>
<thead>
<tr>
<th>Legal Reform Dummy</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFORM</td>
<td>-23.66***</td>
<td>-13.66**</td>
<td>-14.53***</td>
<td>-3.64</td>
<td>-11.87***</td>
</tr>
<tr>
<td></td>
<td>(4.31)</td>
<td>(6.74)</td>
<td>(4.77)</td>
<td>(4.05)</td>
<td>(3.72)</td>
</tr>
</tbody>
</table>

Polynomials in time

<table>
<thead>
<tr>
<th>TIME1</th>
<th>None</th>
<th>2nd order</th>
<th>1st order</th>
<th>2nd order</th>
<th>3rd order</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME2</td>
<td>None</td>
<td>1st order</td>
<td>2nd order</td>
<td>3rd order</td>
<td></td>
</tr>
</tbody>
</table>

| Constant | 37.95*** | 47.37*** | 47.63*** | 29.88** | 5.96 |

| Observations | 23 | 23 | 23 | 23 | 23 |
| DF           | 21 | 19 | 19 | 17 | 15 |
| Adj $R^2$    | 0.63 | 0.64 | 0.66 | 0.71 | 0.74 |
| LL           | -82.20 | -80.71 | -80.30 | -77.16 | -74.33 |
| F-Statistic  | 38.8 | 14.20 | 15.00 | 11.60 | 9.96 |

Notes: Heteroscedasticity and autocorrelation consistent (HAC) estimates, robust standard errors in parentheses.
TIME1 is a time trend for the whole period. TIME2 is a time trend starting after the reform.
*, **, *** Significant at the 10%, 5%, and 1% level, respectively.

Figure 7: Test for structural break – Model (3)
focus on corruption in the central years (1991-2008) of the sample, but no particular peak or dip coinciding with the reform.

Appendix B provides an overview of all the other data sources on corruption we could find for the relevant time period, including corruption perception surveys and composite indices. The picture that emerges from these is ambiguous: some indexes show a worsening around the year of the reform, followed by a recovery, others a stable trend; surveys reveal a lower frequency of harassment bribery compared to other countries in the region and other developing countries, but an increasing acceptance of bribery in general and of the pursuit of improper benefits.

The next section presents our planned work to provide better empirical evidence based on micro-data.

6 Case-file Analysis

As discussed above, the data on prosecutions are subject to several limitations, both theoretical and practical. Starting with the latter, our panel of regional reports is severely unbalanced, while the national ones only cover the post-reform period. Moreover, the data do not distinguish bribery from other corruption offences such as embezzlement, nor giving versus taking of bribes. Although a previous case analysis study (Guo, 2008) found that 82-93% of all corruption cases (between 1978 and 2005) were about bribery, and that only 4-9% of cases were against bribe-givers only, we cannot observe using these data, what part of the drop in cases is simply due to a change in composition. In an extreme scenario, we could imagine that prosecutions against bribe-givers are not carried out anymore for extortionary bribery cases, since they are not considered guilty any longer, and that this would explain all the drop in our aggregate statistic while the number of cases against bribe-takers, and hence the total prevalence of bribery, is unchanged.

Another more general limitation of prosecution data is that they do not disentangle deterrence from detection, or criminal activity from prosecution efforts: a policy that deters crimes but at the same time increases the fraction of those that are successfully prosecuted will have an ambiguous effect on the number of prosecutions.

For these reasons, we analyze here more in depth a stratified random sample of prosecution cases
between 1986 and 2010. Given that we sample a given number of cases, determined by power and budget considerations, in this part of the analysis we cannot gain any insight about the incidence of bribery in general. We can instead observe the impact of the legislative reform on specific details of the corrupt behavior, most notably whether it involves illegitimate benefit or not, the size of the bribe and the favor exchanged, and so on, and the mechanisms through which this behavior occurs or is deterred. In particular we want to distinguish between extortionary (harassment) bribes and bribes for illicit benefit. Moreover we want to shed light on whether and how leniency and asymmetric punishment are applied in practice.

The outcomes that we look at were specified in a pre-analysis plan (Perrotta Berlin and Spagnolo, 2015). Subsequently, in September 2015, we collected a pilot – a small random sample of case files – in order to learn more about what information is available in the case files, while the full data collection is still under way. Here we describe our main outcomes of interest and hypotheses.

- **Relative incidence of report by the bribe-giver.** Ideally we would want to know whether the case was initiated by a spontaneous report. After the introduction of asymmetric punishment and leniency, we should in general expect a higher frequency of cases in which the bribe-givers come forth and report the crime to the authorities in the short run. Although in practice this will depend on the severity of sanctions and the actual likelihood of obtaining leniency, in equilibrium this effect should disappear, due to deterrence. Unfortunately the information on how the case was initiated is not available in the case files. There is however some indication of whether the defendant volunteered information in the cases where leniency was awarded.

- **Relative incidence of harassment bribes in the total.** Looking at the prosecution documents we can first of all form a better idea of whether the presence of improper benefit is actually considered in practice, and so whether harassment bribes are effectively set apart and subject to a different discipline (in particular, the exemption from punishment for the giver). Further, if this is the case, we can observe whether the 1997 reform led to a change in this respect, and hence on the frequency of harassment bribes relative to other types of bribery among the cases detected.

- **Size of the bribe, size of the return, income or status of the bribe-giver (and the bribe-taker)** If the reform, through the introduction of leniency and asymmetric punishment, makes detection easier because of the incentives provided to the bribe-giver, this amounts to an increase in the expected fine for the bribe-taker (even though the actual sanctions decrease). We expect this to lead to an increase of bribe size for distortionary bribes: the threshold to undertake corrupt behavior is now raised, hence individuals will only exchange bribes if the returns from this corrupt agreement are higher. The picture is different though for harassment bribes, when the bribe-giver is not expected to give back the gain. In this case incentives to report increase with the size of the bribe, so the larger bribes will be deterred to a larger extent.

6.1 **Power analysis**

The aim of this exercise is to determine the minimum sample size needed to be able to observe the true impact of the reform. As a reminder, “[…] the power of a statistical test is the probability that it correctly rejects the null hypothesis. [… ] the power of a study depends on sample size,
measurement variance, the number of comparisons being performed, and the size of the effects being studied. (Gelman and Carlin, 2014)

The optimal sample size can be computed using optimal design\textsuperscript{18}, an algorithm that requires as input the standardized effect size, which in turn requires an estimate of the mean and variation of the outcome(s) in the two groups. The main challenge in our case is to specify the expected effect size, given that there is no comparable empirical work done in this area\textsuperscript{19}. We offer here calculations corresponding to ranges of values, going through the outcome variables, discussing what is reasonable to expect in terms of effect on each of them in turns.

Our first outcome is a proportion, the share of cases discovered through a report by the bribe-giver, which makes it easier to specify a reasonable expected value. Starting from a desired significance level of 0.05 and power of 0.8, Figure 9 plots the sample size as a function of the proportion (of reporting bribe-givers) in the control group, i.e. before the reform. The different curves correspond to different sizes of the change in this proportion after the reform. To give an example, if we observe that the proportion of cases initiated by the bribe-giver is 19\% and we expect an increase in it by 60\% after the reform (roughly corresponding to the impact estimated by Miller), we would need about 120 observations in order to correctly reject a false null hypothesis. To identify a (still big but) smaller increase by 40\%, if the original proportion had been 21\%, the required sample jumps already to 483, and a 20\% increase, starting from 25\%, would require more than 1800 observations.

The second outcome is also a proportion, the share of harassment bribes in the total. However, we have very little to go by in terms of expected proportions and effect sizes. There is no distinction on this dimension in previous work, and unfortunately even in our pilot there was no instance of harassment bribes either before or after the reform. Another alternative is represented by a more

\textsuperscript{18}See Raudenbush et al. (2005).

\textsuperscript{19}The alternative would be to specify the minimum size of effect that we will be able to identify given the budget at our disposal. This in turn requires estimates of the factors determining the cost, namely how accessible prosecution documents are, how they look like and consequently how long time it will take to process them and extract the relevant information.
general approach: a shortcut when data on mean and standard deviation of the outcomes are not available is to directly specify the effect size one wishes to detect in multiples of the standard deviation of the outcome. Cohen (1992) proposes that an effect of 0.2 standard deviation is “small”, 0.5 is “medium” and 0.8 is “large”. The sample size required to observe such effects, with the same significance and power as above, is 392, 63 and 25 respectively.

Specifying a plausible range is more complicated for our third outcome of interest, the size of the bribe. The already cited statistical case analysis conducted by Guo (2008) observes that the average amount involved in corruption cases (of which about 85% are bribery) is 17,000 RMB in the period 1978-1990, jumping to around 3 millions RMB in 1991-2004. The first alternative for us would then be to use these values to conjecture an expected standardized effect size. We do not know, however, how representative this sample is for the purposes of our study. The average size of bribe in our pilot was 2.4 million RMB before and 2.3 after the reform, which, taking again the Cohen (1992) shortcut, would suggest a very small change and hence demand a sample size on the larger side.

The ultimate determination of a precise sample size to be collected cannot abstract from cost and budget considerations.

Finally, the sample will be stratified by years. This should allow us to control to some extent for the institutional climate and enforcement effort prevailing at different points in time.

6.2 Secondary outcome variables:

After looking at the pilot, we decided to collect also the following indicators, which can be thought of as mechanisms through which the reform affects our main outcomes. In particular, our main hypotheses are based on the letter of the reform. Since however there is a measure of discretion in the practice of law enforcement, through these secondary outcomes we have the possibility of verifying whether the expected changes in law enforcement do actually happen in practice. We consider this analysis to be exploratory, as it is meant to lend credibility to our main results and clarify the mechanisms. Hence these outcomes are not part of the power analysis determining the size of the sample.

- **Actual likelihood of obtaining leniency.** Leniency is not automatic in the text of the law. Since it is administered at the discretion of the judicial authority, it can be informative to investigate how often it is accorded in practice. This might also highlight differences in attitudes between different courts, which might be exploited at a later stage.

- **Actual sanctions imposed.** Similarly to the previous point, sanctions are specified in the legal text in form of ranges (ex. not less than 2 and up to 5 years imprisonment), and imply discretion for the judicial authority. It is hence interesting to verify how severe are the sanctions administered in practice. Again, this might also highlight differences in attitudes between different courts, which might be exploited at a later stage.

- **Latency.** The analysis of cases allows to some extent to distinguish the time of detection from the time when the corrupt exchange took place. This gives a clearer picture of the time trends in criminal activity as separate from discovery.

---

20This formidable increase is explained by the author with reference to the economic liberalizations occurring during the period, and might have little to do with the legal reform that we are proposing to study. We should however be careful and take into account such secular trends.


• **Bureaucrat characteristics.** Who becomes a bureaucrat after corruption has been made more difficult? Is there selection into this career? We plan to collect as many personal characteristics of the corrupt bureaucrats from the case files as available.

7 Conclusions

This paper provides the first empirical assessment of the effectiveness of leniency and asymmetric punishment as a policy tool against corruption. Leniency has been used before to undermine the internal trust between partners in crime in other law enforcement areas, and this mechanism has been studied theoretically and experimentally in the context of corruption, but never evaluated empirically. Part of the reason lies in the difficulty to obtain good data on corruption. We cannot solve completely the issue of data quality, as we also rely on data from detected corruption cases. However we go a step further by collecting and analyzing microdata from a stratified randomized sample of these cases. Whereas the aggregated data clearly show that something happened to corruption cases in China in connection with the 1997 Criminal Law reform, and that the observed effects are consistent with an increase in deterrence, without the microdata we wouldn't be confident in this interpretation. Through the analysis of the sample we can instead isolate at a higher level of detail the changes in criminal behavior, reporting behavior and prosecution activity and link them to the details of the legal reform, so as to highlight the mechanisms at work. Overall we believe this to be a significant contribution to our understanding of how to best tailor policy to fight corruption.

A further contribution of this paper is the focus on China. This country is home to a sixth of humanity, and currently undergoing a massive crack down on corruption. Whatever we can learn about the effectiveness of anti-corruption policies in this country is likely to have considerable welfare effects.

References


19


Appendices

A Robustness checks

A.1 OLS models

As a first robustness check, we estimate the same models as in Section 5 on the subset of 6 provinces which have a report for all the 23 years of the sample. Figure 10 shows that this makes results even stronger. Figures 11 and 12 report a visualization of the other models estimated in Table 4, which fit different order of polynomials in the time trends.

![Figure 10: Test for structural break in a subset of provinces](image-url)
Figure 11: Test for structural break – other polynomials

Figure 12: Test for structural break – subset of non-missing provinces
A.2 Poisson regression

The linear regression model rests on assumptions that can be at odds with this particular type of data. The dependent variable is assumed to be continuous, normally distributed (hence symmetric around the mean), and linearly related to the independent variables (McClendon, 1994). Crime data rarely adhere to these assumptions. Most crime incidents are distributed as rare event counts. In other words, smaller values are much more common across spatial units than larger values, with zero often being the most commonly observed value. Such a distribution violates the aforementioned assumptions of OLS regression. Although these considerations are attenuated through the aggregation and averaging of the data (remember that the dependent variable is the number of cases per one million citizens), it is worthwhile to compare the OLS results with regression models that are designed to analyze count data, namely the Poisson and negative binomial regression models. The Poisson regression model is often used to model count data and contingency tables. The response variable is assumed to follow a Poisson distribution, and the logarithm of its expected value can be modeled by a linear combination of unknown parameters.

Table 5 reports the results of five Poisson models corresponding to the linear models used in 4. Note that the reported coefficients have to be converted in order to be comparable to the OLS coefficients. The estimated number of cases \( \hat{\lambda} \) in model (1) before the reform can be calculated as \( \hat{\lambda} = \exp(\hat{\beta}_0) = \exp(3.64) = 38.09 \). After the reform, when the treatment dummy takes the value one, the estimated number of cases is \( \hat{\lambda} = \exp(\hat{\beta}_0 + \hat{\beta}_1 \cdot 1) = \exp(3.64 - 0.98) = 14.30 \). This result is very similar to the model (1) in the OLS case, a 62.4% decrease. The other estimates can be calculated in a similar fashion. Take for instance model (2): in the year 1997, the estimate is \( \hat{\lambda} = \exp(\hat{\beta}_0 + \hat{\beta}_2 \cdot 10 + \hat{\beta}_3 \cdot 10^2) = \exp(3.87 - 0.05 \cdot 10 + 0.000929 \cdot 100) = \exp(3.46) = 31.81 \), where \( \hat{\beta}_2 \) and \( \hat{\beta}_3 \) are the coefficients for the second order polynomial in TIME1. In the year 1998, the estimate amounts to \( \hat{\lambda} = \exp(\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 \cdot 10 + \hat{\beta}_3 \cdot 11^2) = \exp(3.87 - 0.64 - 0.05 \cdot 11 + 0.000929 \cdot 121) = \exp(2.79) = 16.28 \). More generally, OLS and Poisson regressions yield very similar results. However, the reform coefficient loses statistical significance in model (4) and (5) and for higher order polynomials (not reported) when using the Poisson model. The similarity in the results is not surprising when considering that the normal distribution is a good approximation to a Poisson distribution for data with a mean above (roughly) 30. The linear model assumes that the values are normally distributed around the expected value and can take any real value. Hence, when the mean is large enough, i.e. negative values are highly unlikely, and the variance is in a similar range, the OLS approximates the Poisson regression estimates quite well.

One drawback of the Poisson regression is the inherent assumption of equal mean and variance. Yet, we saw that the data exhibits different degrees of variation, especially when comparing the dependent variable before and after the reform. To handle overdispersed count variables, the negative binomial distribution is often used, since it allows for variance greater than the mean, making it suitable for count data that do not meet the assumptions of the Poisson distribution. Fitting a negative binomial model to our data delivered identical results\(^{21}\), except slightly inflated standard errors. The pattern of significance is also unchanged, with models (1) - (3) strongly significant but not (4) and (5).

\(^{21}\)Results are not reported but are available upon request to the authors.
Table 5: Poisson and Negative Binomial models

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Legal Reform Dummy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFORM</td>
<td>-0.98***</td>
<td>-0.64***</td>
<td>-0.67***</td>
<td>-0.26</td>
<td>-0.56</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.18)</td>
<td>(0.18)</td>
<td>(0.29)</td>
<td>(0.43)</td>
</tr>
<tr>
<td><strong>Polynomials in time</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIME1</td>
<td>None</td>
<td>2nd order</td>
<td>1st order</td>
<td>2nd order</td>
<td>3rd order</td>
</tr>
<tr>
<td>TIME2</td>
<td>None</td>
<td>None</td>
<td>1st order</td>
<td>2nd order</td>
<td>3rd order</td>
</tr>
<tr>
<td>Constant</td>
<td>3.64***</td>
<td>3.87***</td>
<td>3.88***</td>
<td>3.39***</td>
<td>2.79***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.20)</td>
<td>(0.34)</td>
</tr>
</tbody>
</table>

Observations: 23 23 23 23 23  
DF: 21 19 19 17 15

Notes: Standard errors in parentheses.
TIME1 is a time trend for the whole period. TIME2 is a time trend starting after the reform.
*, **, *** Significant at the 10%, 5%, and 1% level, respectively.

A.3 National data

Table 6 replicates the specification of Table 4, using the national statistics on prosecutions rather than the province-level ones for the period after the reform as a robustness tests. These data provide an upper-bound to the actual number of prosecutions in the group of provinces that we observe for the period before the reform, as they are supposedly including all the 31 provinces. We see that the decrease immediately following the reform is still significant, although smaller in size.

A.4 Placebo Interventions

So far we imposed to the data an exogenous breakpoint at the date of leniency introduction. An alternative approach would be to check whether alternative breakpoints - i.e. a different hypothetical timing of the legal reform - fit the data better. If this were the case, then it would be unlikely that the relationship between the reform introduction and the time series of prosecutions would be causal. If instead the fit is superior when the breakpoint is imposed at the date of the reform, then the data provide support for our hypothesis. Recent literature suggests the Quandt-Likelihood Ratio (QLR) test for detecting structural changes of unknown timing (e.g., Hansen, 2001). The QLR test consists of calculating Chow breakpoint tests at every observation, while ensuring that subsample points are not too near the end points of the sample.

The results are shown in Figure 13. Each point on the graph in panels (a) and (b) represents the maximized log-likelihood of a different regression specification. The x-axis is rescaled with 0 for 1997, the year of the reform. Looking at panel (a), the maximized log-likelihood value is located in 1990, when allowing for breakpoints between 1990-2008 (i.e. a symmetric window, trimming 2 observations from each end of the sample). The corresponding fitted values of the placebo intervention with breakpoint in 1990 are illustrated in panel (c). Looking at panel (c) it becomes evident that 1990 has the highest maximized likelihood because it describes the kink in the dependent variable around 1990. However, looking at the actual data, it is clear that this
Table 6: OLS Regression Results - National-level data for the post-reform period

<table>
<thead>
<tr>
<th>Sample</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal Reform Dummy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.34)</td>
<td>(7.42)</td>
<td>(4.77)</td>
<td>(4.16)</td>
<td>(4.46)</td>
</tr>
<tr>
<td>Polynomials in time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIME1</td>
<td>None</td>
<td>2nd order</td>
<td>1st order</td>
<td>2nd order</td>
<td>3rd order</td>
</tr>
<tr>
<td>TIME2</td>
<td>None</td>
<td>None</td>
<td>1st order</td>
<td>2nd order</td>
<td>3rd order</td>
</tr>
<tr>
<td>Constant</td>
<td>37.95***</td>
<td>47.41***</td>
<td>47.63***</td>
<td>29.88**</td>
<td>5.96</td>
</tr>
<tr>
<td></td>
<td>(4.34)</td>
<td>(14.50)</td>
<td>(10.49)</td>
<td>(13.08)</td>
<td>(17.36)</td>
</tr>
</tbody>
</table>

Observations 22 22 22 22 22
Degrees of Freedom 20 18 18 16 14
Adjusted R^2 0.58 0.59 0.66 0.66 0.70
Log Likelihood -79.02 -77.71 -77.22 -74.22 -71.48
F-Statistic 30.00 10.90 11.70 9.18 7.92

Notes: Heteroscedasticity and autocorrelation consistent (HAC) estimates, robust standard errors in parentheses.
TIME1 is a time trend for the whole period. TIME2 is a time trend starting after the reform.
*, **, *** Significant at the 10%, 5%, and 1% level, respectively.

---

corresponds to the sharp increase in cases around 1990, which plausibly have very little to do with the events in 1997. The test is looking for a global maximum, and hence we should be cautious with the interpretation of the results and the choice of an appropriate window for possible breakpoints.

In fact, the results change when a smaller window is chosen. Allowing for breakpoints between 1992-2006 (again a symmetric window, this time leaving out the first and last 4 observations), the log-likelihood is maximized in the year 1996. In particular, the maximized log-likelihood produced by the placebo intervention in 1996 is the only one greater than the one produced by the actual legal reform in 1997. By visual inspection of the bottom panels of the figure, it is clear that the fitted values approximate the data much better in panel (d) compared to panel (c), both before and after the reform. The year 1990 as a breakpoint with the highest maximum log-likelihood of the placebo interventions might therefore just be an anomaly.
B Soft evidence: other corruption indicators

In the literature on corruption, composite indexes have gained popularity. Well known examples include Transparency International’s widely-cited “Corruption Perceptions Index” and the World Bank Institute (WBI) “Control of Corruption” index Kaufmann et al. (2009). The CPI scores countries based on how corrupt their public sector is perceived to be on a scale of 0-100, where a lower score corresponds to more corruption. Left panel of Figure 14 shows a sharp improvement in the CPI score for China up to 1998, the year after the reform, followed by a few years drop and then a more moderate but still increasing trend.

The Index of Economic Freedom compiled by the Heritage Foundation is another composite index of corruption perceptions, primarily derived from the CPI and complemented with qualitative and quantitative data from other sources\textsuperscript{22}. The component of the index called ‘Freedom from

\textsuperscript{22}From the Heritage Foundation’s webpage: “Unless otherwise noted, the Index relies on the following sources for information on informal market activities, in order of priority: Transparency International, Corruption Perceptions Index, 2011; U.S. Department of Commerce, Country Commercial Guide, 2009-2012; Economist Intelligence Unit, Country Commerce, 2009-2012; Office of the U.S. Trade Representative, 2012 National Trade Estimate Report on Foreign Trade Barriers; and official government publications of each country.” http://www.heritage.org/index/freedom-
of corruption", also varying on a similar scale of 0 to 100, is plotted in the right panel of Figure 14. It shows an abrupt drop in 1997, followed by a steady improvement. These patterns are consistent with an increased emphasis on corruption fighting around the year of the reform, which might reflect in the public perception of how pervasive corruption is.

Within their Worldwide Governance Indicators (WGI) research project, covering 212 countries and territories and measuring six dimensions of governance since 1996, Kaufmann, Kraay and Mastruzzi Kaufmann et al. (2009) report the aggregate “Control of Corruption” index, varying on a scale form -2.5 to 2.5, and the underlying data from all of their sources. Rather than the original format, though, these are reported in the form in which they enter the governance indicators, namely rescaled on a 0-1 scale. In both scales, a lower score indicates a worse outcome, i.e. more corruption. Figure 15 plots the aggregate index and all the sources that are available for China in all the relevant years. Only one data point is available prior to 1997, making it hard to draw any inference on the impact of the reform. In general, the aggregate index gives a rather negative assessment of the trends in corruption, although all the components seem rather stationary.

This illustrates one of the main drawbacks of this type of composite indexes. The sources used in constructing them can change over time. This implies that different values are likely to reflect differing implicit definitions of corruption, depending on what goes into them. The standardization procedure used to place different indicators on a common scale can also impair the ability to track changes meaningfully over time. A final issue with the indexes that use expert sources is their interdependence. If expert assessments display high correlations driven by the fact that they consult each other’s ratings - or that they all base their ratings on the same information sources - this can undermine the main premise of the aggregation methodology that averaging more sources produces more accurate and reliable estimates.

We considered separately also the components of the CPI index. These were ultimately not included here due to either not being publicly available, not covering a sufficient number of years, or not referring specifically to bribery. For some sources, though, we were able to access the unpublished firm-level responses that are behind the publicly released country-level index. Surveys are...
relatively well-suited for evaluating the administrative corruption, as they measure the prevalence of corruption as experienced by users of government services. However, surveys are less effective in assessing the prevalence of corrupt transactions that occur entirely within the state, for example when politicians bribe bureaucrats. The Business Environment and Enterprise Performance Survey (BEEPS) and the World Economic Forum (WEF) “Executive Opinion Survey” are the most research-friendly surveys on corruption-related topics, as they are systematic and comparable across countries and years, have broad coverage and disclose most informations about their definitions and methodology. The BEEPS, funded by the EBRD, are focused on Eastern Europe and Central Asia and not available for China, however, while the WEF includes China as long back as at least 1995\textsuperscript{23}.

In the question of interest, survey respondents were asked how common it is for firms to make undocumented extra payments or bribes connected with imports and exports; public utilities; annual tax payments; awarding of public contracts and licensing; and obtaining favorable judicial decisions. In all of these cases, the assessment is improving with very similar downward trends in the period 2004-2013\textsuperscript{24}.

Another source that similarly elicits the information about what service the bribe was paid for is the World Bank’s Enterprise Surveys, collected since 2002 from 130,000 companies in 135 countries. Unfortunately only one year is available for China. As reported in Table 7 below, according to this source bribery incidence in 2012 is lower in China (11.6% of firms experiencing at least one bribe payment request) both compared to the East Asian and Pacific region (24.2%) and to the whole survey sample (17%). However, bribery associated with illegitimate benefit is more common while extortionary bribery is less common in both comparisons (columns (2) and (3) of Table 7).

Finally, the World Value Survey is another well-known source of data on opinions and attitudes

\textsuperscript{23}Note that the WEF has conducted the Executive Opinion Survey for over 30 years, but due to methodology changes they are unwilling to provide data going further back in time than 2004.

\textsuperscript{24}The results are not reported. The distinction between different circumstances in which a bribe payment might occur can be suggestive about the presence or not of illegitimate benefit, that is relevant for the application of impunity according to the law. However, since there is no difference in the trends, no relevant inference can be drawn. Moreover the time horizon of these data does not cover the implementation of the reform.
Table 7: Unjust-benefit bribes VS harassment bribes in 2012

<table>
<thead>
<tr>
<th></th>
<th>Bribery incidence (percent of firms experiencing at least one bribe payment request)</th>
<th>Percent of firms expected to give gifts to secure government contract</th>
<th>Percent of firms expected to give gifts to public officials “to get things done”</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>11.6</td>
<td>42.2</td>
<td>10.7</td>
</tr>
<tr>
<td>EAP</td>
<td>24.2</td>
<td>31.0</td>
<td>20.4</td>
</tr>
<tr>
<td>All</td>
<td>17.0</td>
<td>26.4</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Source: World Bank Group Enterprise Surveys

around the globe. From the WVS we focus on two variables: the share of respondents supporting the view that it is justifiable to accept a bribe in the exercise of one’s duty; and the share of respondents that think it is justifiable to claim benefits to which one is not entitled. The first one can be thought of as a proxy for how widespread the practice of bribing public officials is in general. The second can instead be related to the practice of using or at least supporting the use of bribery to obtain an illegitimate benefit.

In the left panel of Figure 16, a decreasing trend in the acceptance of passive bribery seems to have been reverted by the legal change. Even though there are too few data points to actually establish a pre-reform decreasing trend, it is clear that the share of respondents supporting this view increased after the reform. This is somewhat reminiscent of the critique moved to asymmetric punishment as a depenalization of a wrong behavior, which would over time undermine the moral sense. The left panel shows instead that the acceptance for the behavior of seeking a benefit one is not entitled to only grew over time since 1990. In this case, the legal change did not impact the secular trend.

Figure 16: Share of respondents who think behavior is justifiable

![Figure 16](image_url)

\[25\] For example in Dreze (2011).