Subordinates in Charge: Does Delegation Improve Bank Supervision?*

Wouter Dessein, Columbia GSB
Di Gong, UIBE
Thomas Lambert, RSM Erasmus
Wolf Wagner, RSM Erasmus

March 1, 2026

Abstract

We develop a model of bias and information loss in supervisory communication and apply it to evaluate a policy reform that delegated supervisory decision-making authority for a subset of bank branches to a lower-level. Affected branches are 57-80% more likely to face supervisory intervention, implying substantial efficiency gains arising from improved detection of banking misconduct and more accurate assessment of its severity. The evidence is inconsistent with alternative explanations, such as higher supervisory stringency or more risk-taking. Our findings highlight decentralization benefits within supervisory hierarchies.

Keywords: Supervision, communication, decentralization, financial architecture

JEL: D23, G21, G28

*We thank our discussants, Margherita Bottero, João Rafael Cunha, Guangqian Pan, Tianyue Ruan, Alessandro Scopelliti, Yang Su, Wenyu Zhu, as well as Andreas Brøgger, Bing Han, Kewei Hou, Max Margolin, José Martín Flores, and Tianyi Wang, for their helpful suggestions. We also thank conference and seminar participants at the AEA Meetings, China Financial Research Conference, EFA Meetings, Sydney Banking and Financial Stability Conference, Central Bank of Ireland, CEPS, CUHK, Dresden University of Technology, Erasmus University, UIBE, University of Limoges, University of St Andrews, and University of the Andes for many helpful comments and discussions. We are grateful to Guangzi Li for sharing the data on bank lending and Yanxin Yu for research assistance. This version substantially extends and supersedes a previous draft that circulated under the title "Local Bank Supervision." Di Gong acknowledges funding from the National Natural Science Foundation of China (Project No. 72273026). Wouter Dessein is affiliated with Columbia Business School and the CEPR. Di Gong is affiliated with the University of International Business and Economics. Thomas Lambert is affiliated with the Rotterdam School of Management. Wolf Wagner is affiliated with the Rotterdam School of Management and the CEPR.

Please send correspondence to Wouter Dessein at wd2179@gsb.columbia.edu, to Di Gong at d.gong@uibe.edu.cn, to Thomas Lambert at t.lambert@rsm.nl, or to Wolf Wagner at wagner@rsm.nl.

1 Introduction

A central question in the design of the financial architecture is how powers should be allocated between local and central authorities. A large part of the discussion has focused on this in the context of local supervisors that are *independent* of the central entity, broadly finding that outcomes are less strict under local decision-making.[1] This is explained by local supervisors being captured by local interests, or failing to internalize spillovers.

This paper shows that within a supervisory *hierarchy*, local supervision may deliver stricter and better outcomes.[2] The reason is that divergent interests of local and central supervisors make communication among them less informative. This results in a loss of local information when decisions are taken at the central level and in interventions not taking place even though they are warranted. As a result, delegation of decision-making to local supervisors can make supervision both more efficient and stricter.

We examine a decentralization reform in China and interpret the findings through the lens of a model of supervisory communication and bias. In 2015, China decentralized supervision for branches belonging to banks classified as "local" but not for branches belonging to "national" banks.[3] Prior to 2015 all the branches were supervised in a hub-and-spoke system in which information was predominantly collected by supervisors at the prefecture-level, but decisions involved the central supervisor, thus requiring communication among both. The reform then fully transferred decision-making for branches

[1] See, for example, Agarwal et al. (2014), Haselmann et al. (2022) and Beck et al. (2023). Note that lenient supervision does not necessarily imply worse supervision. Generally, theory makes divergent predictions about whether centralized or decentralized settings are preferred (e.g., Dell'Ariccia and Marquez (2006); Calzolari et al. (2019); Lóránth et al. (2022)). See Ampudia et al. (2019) for an overview of the policy debate.

[2] Hierarchical, or hub-and-spoke, supervision is common. In the U.S., Fed, FDIC and OCC have headquarters and regional offices. Similar distributed structures exist for instance in Germany (Bundesbank), India (Reserve Bank of India), Canada (the Office of the Superintendent of Financial Institutions) and China (People's Bank of China and National Administration of Financial Regulation).

[3] The Chinese banking sector is the largest in the world with about $40 trillion in assets. It serves more than 800 million individuals through more than 4,000 commercials banks.

of local banks to local supervisors. Importantly, the delegation did not change the overall objective of supervision, as local supervisors remain fully subordinated to the central supervisor and accountable to the latter.

We make use of a large, partly hand-collected, data set covering 5,429 prefecture-level branches over a ten-year window around the 2015 policy reform. We measure supervisory outcomes through formal declarations of banking misconduct ("supervisory interventions"). For this we make use of novel bank branch- and supervisory-office-level data on supervisory interventions.[4] The granularity of the data allows us to exploit variations across branches of the same bank but also within a supervisory office (regarding decisions about branches belonging to different banks).

We find that branches of local banks are more likely to be subjected to a supervisory intervention following the reform. Specifically, the probability for a local branch receiving a supervisory intervention increases by 6.3-8.8 percentage points (pp) relative to a branch of a national bank, which corresponds to 57-80% of the unconditional probability of intervention.

What can explain that decentralization leads to more interventions, even though local supervisors are usually considered to be more lenient? We present a model where local supervisors are better informed about potential misconduct but are biased against interventions. This bias impedes communication (modeled as cheap talk) between local and central supervisor under centralization. Specifically, the local supervisor wants to avoid intervention by the central supervisor when the level of misconduct is moderate. This results in him not relaying such incidences to the central supervisor, deterring (desirable) supervisory intervention. Decentralization restores information in supervisory decision-making and leads to more detection of misconduct even though the local supervisor is biased against interventions.

[4] A supervisory intervention in the Chinese system is a formal step taking when a bank is suspected of not adhering to rules or good practice. It declares the specific violation, corrects the misconduct, and issues a punishment. There are more than 300 local supervisory offices that have received more decision-powers as a result of the reform.

Whereas our model is open about whether decentralization is desirable,[5] it is always desirable in the parameter range in which supervisory interventions increase. The implied welfare improvements are material. In particular, our baseline estimate suggests that the reform eliminates 19% of the efficiency losses under centralized supervision.

Our model has two important implications. First, now also moderate levels of misconduct are detected. Such misconduct warrants lower penalties. We hence expect the average fine issued in a supervisory intervention to decline. Second, the model also predicts that fines becomes more variable. This is because under centralized supervision, the local supervisor strategically withholds information about the severity of misconduct even in the range where interventions ultimately take place, resulting in penalties that are less responsive to the underlying misconduct. We find both predictions to be fulfilled in the data. Our model also identifies two factors that determine the extent to which decentralization improves outcomes: central access to local information and predictability of local information. Using proxies for both, we find the effect of the reform indeed to be more pronounced when (incremental) local information is more valuable.

Our findings regarding the likelihood, severity and dispersion of supervisory intervention as well as the value of local information all speak to an informational interpretation of the reform. They are inconsistent with several alternative explanations. First, there may be a confounding effect around the time of the reform that has increased risk specifically for branches of local banks (though both types of branches are comparable prior to the reform). Whereas this may also lead to more supervisory interventions for local branches, average fines should increase then as well, inconsistent with our findings (the risk explanation is also inconsistent with our finding described below that local banks have become more risk-averse following tighter supervision). Second, the preferences of local supervisors regarding interventions in local banks may have changed around the reform. This may lead to more interventions if for some reason supervisors have

[5] If the importance of local relative to public information is modest, or if the local bias is very large, centralization is in fact preferred.

become stricter towards local banks, but this would again suggest that average fines should increase for them, not decrease. Third, supervisory capacity may have increased following the reform, possibly also explaining better supervisory outcomes (Eisenbach et al., 2022). However, our results are robust to controlling for changes in supervisory capacity as proxied by new hires at the local level.

We conclude our analysis by investigating whether improved supervision curbs risk-taking by banks. Using individual (loan-level) lending decisions, we find that branches of local banks become more conservative in their lending after the policy reform: They require higher compensation for taking on risk, and they reduce the amount they lend, that is, the amount of credit risk they take on. We also find that this has aggregate consequences: prefectures with a higher share of branches from local banks experience lower credit growth in the aftermath of the reform.

Our paper has an important message for policy. A large part of the current discussion on the design of the supervisory architecture (in particular, in the Eurozone) has focused on centralization of policies vis-à-vis independent supervisors (see Ampudia et al. (2019)). Several studies have emphasized stricter outcomes under centralized supervision: Agarwal et al. (2014) show that federal regulators are consistently tougher than state regulators, whereas Haselmann et al. (2022) find that large significant banks included in the SSM are subjected to stricter standards.[6] Our analysis suggests that within a supervisory hierarchy, as present in many countries, outcomes may be stricter (and better) when the local supervisor is in control. This is because the local bias impedes communication, making decisions by the central supervisor less informed about local conditions[7] and ultimately also more biased. In addition, within a hierarchy any

[6] Our model shows that stricter supervision should not be equated with improved efficiency: Delegation can be preferred even when it leads to fewer interventions as the interventions that take place better reflect underlying misconduct.

[7] Several papers have documented the importance of local information in the supervisory process. For example, Lim et al. (2023) and Gopalan et al. (2021) have shown that the closures of local offices of

bias of the subordinated supervisor is probably more easily contained.

Our paper is related to a growing theoretical literature that analyzes optimal financial architectures involving multiple supervisors and regulators. A large part of this literature has studied trade-offs between a single central regulator and several independent local regulators, focusing primarily on inefficiencies arising from externalities across jurisdictions (Dell'Ariccia and Marquez, 2006; Kara, 2016; Foarta, 2018; Calzolari et al., 2019; Bolton and Oehmke, 2019; Colliard, 2020; Lóránth et al., 2022). Closely related, Repullo (2018) and Carletti et al. (2021) study the incentives of local and central supervisors to acquire information when there is hierarchical supervision. Our paper instead focuses on the role of coarse communication (Crawford and Sobel, 1982) for the optimal design of supervision. Such communication has also been shown to shape outcomes within firms, in terms of corporate governance (Harris and Raviv, 2008; Malenko, 2014; Baldenius et al., 2014; Malenko, 2024) and between different layers of management (Harris and Raviv, 2005; Grenadier et al., 2016; Chakraborty and Yilmaz, 2017).

Our model itself builds directly on Dessein (2002) who analyzes the benefits of delegation relative to communication. Dessein (2002) shows that there is a trade off between information and incentives but that decentralization is preferred for a large range of parameter values.[8] Our work extends Dessein (2002) to a setting with both an extensive and an intensive margin (the decision to intervene and the severity of intervention). We show that at the extensive margin there is no trade off with incentives, and that incorporating the latter even eliminates the overall trade off for a large range of parameters.[9] Our model also maps into observable outcomes, allowing us to infer the efficiency implications of an actual decentralization reform.

central supervisors, resulting in higher bank risk, consistent with geographical proximity between banks and supervisors improving supervision

[8] In contrast to Dessein (2002), in Aghion and Tirole (1997) information acquisition is endogenous and is affected by decentralization. Other papers that consider decentralization in a general principal-agent context are Melumad and Reichelstein (1987), Bolton and Dewatripont (1994) and Alonso et al. (2008).

[9] An alternative interpretation of our model is that of a project choice by a local manager involving both a decision on whether to undertake the project, and at what scale. The manager, however, is biased towards inaction and smaller projects.

Empirical studies have emphasized benefits to decentralization in several contexts. Li (2017) shows that experts are more biased but still make better decisions due to better information. Aghion et al. (2021) find that firms that delegated more power to local plant managers prior to Great Recession outperformed centralized counterparts in sectors that were hardest hit by the subsequent crisis. Bandiera et al. (2021) analyze the consequences of autonomy of procurement officers in Pakistan, and find that autonomy reduces procurement prices with no quality reductions. Kala (2024) shows that granting managers in state-owned enterprises more autonomy results in greater value added but reduces investment in activities that bring primarily social benefits.

2 Institutional Background

2.1 An overview of the China's banking sector

The Chinese banking system has experienced substantial growth over the decades and is now the world's largest ($38.98 trillion assets as of the end of our sample period (2020), compared to $27.71 trillion in the United States). There are over 4,000 commercial banks. Eighteen of them operate on a nationwide scale (the six largest state-owned banks and the 12 national joint-stock banks). These national banks collectively held $26.96 trillion assets, accounting for 69.5% of all commercial bank assets. In addition to the national banks, there is a diverse range of regional institutions that we refer to as local banks: 134 prefecture-level commercial banks; approximately 1,600 rural commercial banks; several hundred of rural credit cooperatives; and numerous village banks.

Commercial banks are predominantly organized through branches. The typical organizational structure consists of a headquarter, a branch ("Fenhang" in Chinese) in each prefecture where the bank is active, numerous lower-level offices ("Zhihang"), and even more local branches. Our analysis will focus on the prefecture-level branch, which is responsible for all banking activities in the prefecture, including the offices and the local

branches. A prefecture is an administrative unit similar to a Metropolitan Statistical Area in the United States and has a median population of more than 3.3 million (larger than the median US state).

2.2 Regulatory and supervisory framework

The China Banking Regulatory Commission (CBRC) was created in 2003 as the main authority regulating and supervising the Chinese banking sector. In 2018, the CBRC merged with the China Insurance Regulatory Commission (CIRC), to become the China Banking and Insurance Regulatory Commission (CBIRC). The CBRC has a hub-and-spoke structure akin to the OCC in the United States. Headquartered in Beijing, the CBRC supervises all commercial banks through a network of local supervisory offices. This network comprises provincial offices (CBRC bureaus) in the capitals of the 31 provinces and in five major metropolitan areas (Dalian, Ningbo, Xiamen, Qingdao, and Shenzhen), and municipal offices (CBRC sub-bureaus) in 306 prefectures (see Figure 1, Panel A). Comparable to the field offices of the OCC in the United States, these offices (bureaus and sub-bureaus) serve as local entities overseeing the banks within their respective jurisdictions.

The CBRC has a fully hierarchical management structure. The CBRC's head (or central) office in Beijing establishes rules, guidelines, and policies. It also directly appoints the heads of local offices. There is, in principle, full alignment of the objectives between central and local supervision. This setting is similar to the one in many other countries, see footnote 2.

2.3 The decentralization reform of 2015

In January 2015, the CBRC had its first major reform. In the area of supervision, the primary objective of the reform is to decentralize administrative powers, bringing supervisors closer to financial institutions. Specifically, the reform transfers the supervisory

responsibilities and powers for local banks to local supervisors, without changing the organization of supervision for national banks.

Prior to reform, the central office and local offices of CBRC jointly supervised all banks.[10] However, the ultimate authority was with the central supervisor. The most common case regarding supervisory intervention[11] is that local offices provided recommendations to the central office, which then made the final decision. In other cases, local offices decided on an intervention first and informed the central office afterwards, possibly leading to an overruling of the lower-level decision. In a few cases, investigations by a local office were also joined by staff from the central office, and joint decisions were reached. The reform fully transferred supervision for local banks, and specifically decisions on interventions, to the local offices. The local offices now independently decide on interventions, without reporting obligations to the central office. Figure 1, Panel B, illustrates the change in the supervisory architecture as a result of the 2015 reform. While the reform makes supervision of local banks the sole task of local supervisors, it does not modify who sets the overall objectives of supervision. Those remain in the full realm of the CBRC's head office, to which the local offices are reporting.

3 A Model of Supervisory Communication and Bias

The relation between central and local supervisor in China resembles a principal-agent setting, as the the latter is fully subordinated to the central supervisor who sets the overall objectives of supervision. Prior to the reform, the central supervisor had ultimate authority over decisions (even though supervision was carried out jointly). Information about the banks was primarily in the hand of the local supervisor, who communicated this information to the central supervisor (either directly, or indirectly through an ini-

[10] With the exception of the headquarters of the national banks, which are solely supervised by the central office.

[11] Official supervisory interventions play a vital role in China, incentivizing banks to promptly rectify identified problems.

tial decision). The reform then fully delegated decision-making authority to the local supervisor.

In this section we develop a model that allows analyzing the impact of the reform on supervisory decision-making.[12] The main tension in this model is that the local supervisor has better information about banks, but faces distorted incentives as he is biased against interventions (Section 5.6 provides evidence for such a bias). A central insight of the analysis is that local bias substantially distorts decision-making also under centralized decision-making. The reason is that the local supervisor communicates strategically to the central supervisor, in order to influence the latter's decisions. This results in two costs. First, also decisions by the central supervisor are subject to a bias (that goes in the same direction as the local supervisor's bias). Second, as the local supervisor does not fully transmit information, there is an informational loss, making supervisory outcomes less tailored to underlying conditions.

The model shows that delegation can be both welfare enhancing and reducing, and derives conditions for either case. Based on our empirical estimates this allows to exclude the case where delegation lowers welfare. The model also allows us to derive a range of empirical predictions that can be contrasted against alternative explanations for why supervisory interventions may have changed following the reform.

3.1 Setup

There is a central supervisor, also called the principal (she), and a local supervisor, also called the agent (he). The type of a bank is given by $\theta + \delta$. Only the agent observes θ, where θ is i.i.d. uniformly distributed on $[-1, 1]$. In contrast, δ is observed by both principal and agent, where $\delta \in \left\{-\bar{\delta}, \bar{\delta}\right\}$ with equal probability with $\bar{\delta} \in (0, \frac{1}{2})$ a measure

[12] The model adapts Dessein (2002) to a setting in which the principal makes two decisions that can be observed in the data: the intervention decision (extensive margin) and the size of the fine (intensive margin).

of (the importance of) public information.[13] Bank type can be interpreted as how likely it is that the bank has violated supervisory guidelines (that is, committed misconduct) and/or how severe the violation is. Beyond misconduct, it can also be seen as an indicator of overall risk-taking by the bank.

The central supervisor would like to impose a fine (penalty) on the bank whenever $\theta + \delta > 0$, and her ideal fine is

$$F^*(\theta) = \max\{0, \theta + \delta\}.$$

The local supervisor would like to impose a fine whenever $\theta + \delta > b$ ($b > 0$) and its ideal fine is given by

$$F_A^*(\theta) = \max(0, \theta + \delta - b\}.$$

Preferences over fines are represented by a loss function which is linear in the absolute distance between the actual fine F ($F \geq 0$) and the supervisor's ideal fine. In particular, the central supervisor's preferences over fines are given by

$$U(F, \theta) = -\left|F - F^*(\theta))\right|.$$

Similarly, the local supervisor's preferences over fines are given by

$$U_A(F, \theta) = -\left|F - F_A^*(\theta)\right|.$$

We refer to an outcome with a positive fine ($F > 0$) as a supervisory intervention.

[13] Limiting $\bar{\delta}$ to be smaller than $\frac{1}{2}$ reduces the number of cases to be discussed and, hence, simplifies the exposition.

3.2 Organization design

We consider two institutional settings: (1) Decentralized supervision: Under delegation, the local supervisor decides on the fine. It follows that under delegation, the fine always equals $F = \max\{0, \theta + \delta - b\}$. (2) Centralized supervision: Under centralization, the central supervisor decides on the fine, but she can consult the local supervisor. Any communication is modeled as cheap talk (Crawford and Sobel, 1982; Dessein, 2002). The local supervisor then sends the central supervisor a cheap talk message m and the principal imposes a fine

$$F = \arg\max E(U(F, \theta)|G(\theta, m))$$

where $G(\theta, m)$ is the distribution of θ conditional on the local supervisor sending message m (determined in equilibrium).

3.3 Equilibrium under centralization

We now study the outcome under centralization with cheap talk. The following proposition establishes when informative communication is feasible under centralization, and whether the principal intervenes following communication by the agent.

As is common, we focus on the most informative cheap talk equilibrium (see, e.g. Dessein (2002)) as this equilibrium is preferred by both principal and agent. As is well known, there always exists an equilibrium in which any communication is uninformative (babbling). Define

$$n^* = \left\lfloor 1 + \sqrt{\frac{1+\delta}{2b}} \right\rfloor$$

where $\lfloor X \rfloor$ is the largest integer smaller than X.

Proposition 1. *(i) Whenever* $b > \min\{\frac{1}{2} + \frac{\delta}{2}, \frac{1}{2} - \delta\}$, *then:*

- *If* $\delta < 0$, *there is no informative communication between agent and principal, and*

the principal never issues a fine since $E(\theta+\delta)<0$.

- *If* $\delta>0$, *the principal always issues a strictly positive fine. Whenever* $b>\frac{1}{2}$, *there is no informative communication between agent and principal, but now* $E(\theta+\delta)>0$. *Whenever* $b\in(\frac{1}{2}-\delta,\frac{1}{2})$, *informative communication between agent and principal is possible, but* $E(\theta+\delta|m)>0$ *following any cheap talk message* m.

(ii) Whenever $b<\min\{\frac{1}{2}+\frac{\delta}{2},\frac{1}{2}-\delta\}$, *the most informative cheap talk equilibrium is characterized by a set of* $n^*>1$ *messages* $M=\{m_1,...m_{n^*}\}$,[14] *and a partition* $(a_0,a_1,a_2,...,a_{n^*})$ *of* $[-1,1]$, *with* $a_0=-1, a_{n^*}=1$, *and* $a_{k-1}<a_k$ *for* $k\in\{1,..,n^*\}$ *such that*

- *the agent sends message* m_k *whenever* $\theta\in(a_{k-1},a_k)$ *so that* $G(\theta|m_k)=U[a_{k-1},a_k]$.
- *upon receiving the message* m_1, *the principal issues no fine, that is* $F_1=0$.
- *upon receiving message* m_k *with* $k>1$, *the principal issues a fine*

$$F_k=\frac{a_{k-1}+a_k}{2}+\delta>0.$$

- *The partition* $(a_0,a_1,a_2,...,a_{n^*})$ *is characterized by*

$$\begin{aligned} a_1+\delta &= b\cdot(n^*-1)\left(\frac{2n^*}{2n^*-1}\right)+\frac{1+\delta}{2n^*-1}\\ a_2-a_1 &= 2(a_1+\delta)-4b\\ a_{k+1}-a_k &= a_k-a_{k-1}-4b,\ \text{for } k>1 \end{aligned}$$

- *Whenever* $n^*>2$, *there also exist cheap talk equilibria with* $n\in\{1,..,n^*-1\}$ *messages characterized by the same conditions.*

Proof. See Appendix □

The above proposition shows that that whenever $b<\{\frac{1}{2}+\frac{\delta}{2},\frac{1}{2}-\delta\}$, informative communication between principal and agent is possible and the principal does not issue

[14] Note that $b<\frac{1}{2}+\frac{\delta}{2}$ implies that $n^*>1$

a fine whenever the agent recommends her not to do so (i.e. whenever the agent sends message m_1). For most of the parameter space in which informative communication is feasible, that is whenever $b > \frac{1+\delta}{8}$, the agent can only indicate whether the principal should issue a strictly positive fine or not ($n^* = 2$). Only for b very small, that is $b < \frac{1+\delta}{8}$, the agent may be able to indicate whether a desired fine should be small or large, or even finer gradations ($n^* \geq 3$).

For large biases, that is when $b > \max\{\frac{1}{2} + \frac{\delta}{2}, \frac{1}{2}\}$, no informative communication is feasible and intervention depends on the sign of δ.

Finally, when $\delta > 0$, there is also small parameter range when $b \in (\frac{1}{2} - \delta, \frac{1}{2})$, where informative communication is feasible, but the principal always intervenes.

3.4 Optimality of delegation and likelihood of supervisory interventions

We can now examine under which circumstances delegation is preferred (from the principal's perspective) to centralization, and what this implies for (observable) supervisory outcomes. Supervisory outcomes under delegation are given by $F = \max\{0, \theta + \delta - b\}$. Supervisory outcomes under centralization were described in the previous subsection, and given by $F = \max\{0, E(\theta + \delta | m_i)\}$.

Proposition 2. *(i) Whenever $b > 1 - \sqrt{(1-\bar{\delta})\bar{\delta}}$, centralization is optimal, and there is more frequent intervention under centralization than delegation.*

(ii) Whenever $\frac{1}{2} - \bar{\delta} < b < 1 - \sqrt{(1-\bar{\delta})\bar{\delta}}$, delegation is optimal, but intervention is still more frequent under centralization. There is always intervention under centralization if $\delta > 0$ (which occurs with 50% probability); there is intervention under delegation if and only if $\theta > b - \delta$, which occurs in less than 50% of cases.

(iii) Whenever $b < \frac{1}{2} - \bar{\delta}$, delegation is optimal, and there is strictly less intervention under centralization. Under delegation, there is no intervention whenever $\theta < b - \delta$; under centralization, there is no intervention whenever the agents sends message m_1,

that is when $\theta < a_1$ *where* $a_1 > b - \delta$.

To provide intuitions for this result, it is easy to focus first on the case where $\bar{\delta}$ is small, so the cut-offs for b in the above proposition are approximately 1 and $\frac{1}{2}$. Whenever $b > 1$, the agent then never intervenes under delegation. In contrast, under centralization, the center optimally intervenes whenever $\delta = \bar{\delta} > 0$. It follows that centralization is both strictly preferred and results in more interventions.

In contrast, for $b \in (\frac{1}{2}, 1)$, delegation is strictly preferred. The reason are the informational benefits of delegation. Under delegation, there is then intervention whenever $\theta + \delta > b$. Moreover, whenever the agent intervenes, the principal also strictly prefers intervention (and derives positive utility from it). In contrast, under centralization, no informative communication is feasible and since $\bar{\delta}$ is small, the principal is almost indifferent between intervention and not. It follows that for $b \in (\frac{1}{2}, 1)$, delegation is strictly preferred over centralization. It is still the case, however, that centralization results in more frequent intervention than delegation: the principal intervenes whenever δ is positive (50% of all cases), whereas the agent intervenes only when $\theta + \delta > b$, which occurs in less than 25% of all cases for $b > \frac{1}{2}$.

The above logic in favor of delegation is more subtle when $\bar{\delta}$ is large, as then there is a substantial benefit from intervening under centralization when $\delta = \bar{\delta} > 0$. A trade-off then arises between more, larger but also rather indiscriminate interventions under centralization (always intervene when $\delta > 0$), and less frequent, smaller interventions under delegation that are tailored to the value of θ (intervene when $\theta + \delta > b$). As we show in the Appendix, this trade-off resolves in favor of delegation when b is not too large relative to $\bar{\delta}$, that is when $b < 1 - \sqrt{(1-\bar{\delta})\bar{\delta}}$.

Finally, when $b < \frac{1}{2} - \bar{\delta}$, delegation remains strictly preferred over centralization, but now also results in more intervention than centralization. The latter result may seem counterintuitive given that the local supervisor is more lenient (has a higher threshold for intervention). The agent's bias, however, is now sufficiently small so that the principal

does not intervene whenever the agent sends message m_1 – indicating that the bank is low-risk. The agent, in this sense, has "real authority" over when to intervene or not. Under centralization, however, the agent sends this low-risk message whenever $\theta < a_1$ where one can verify that $a_1 > b - \delta$. For $\theta \in (b - \delta, a_1)$, the local supervisor would like to issue a fine, but she prefers a much smaller fine in those cases compared to what the central supervisor would impose (a big fine). As a result, the local supervisor prefers not to reveal these types and recommends no fine at all, even though he would have intervened himself under delegation.

One can further show that delegation remains strictly preferred over centralization due to the same informational benefits: delegation allows to tailor fines to local information (the realization of θ). Importantly, this is the case both on the extensive margin (whether or not to issue a fine based on θ) as well as the intensive margin (tailoring the size of the fine to θ). On the extensive margin, delegation unambiguously improves efficiency – there is strictly more intervention under delegation and such intervention is always warranted. On the intensive margin (when messages $m_2, m_3, ...$ are sent under centralization), there is a trade-off between fines that are too small under delegation, but also more tailored to the realization of θ. These results mirror those in Dessein (2002), which was the first paper to compare cheap talk with delegation but only considered the intensive margin.

This proposition allows us to infer from our empirical results that we are in the parameter range where delegation is optimal: The proposition shows that whenever delegation is suboptimal, supervisory interventions decrease, which is inconsistent with the evidence. In fact, interventions are increasing only in the third range where the bias is small, there is informative communication, and delegation is optimal.

3.5 Two message equilibrium

In this section we derive several predictions of the model that can be empirically examined. We also consider competing channels that may result in delegation leading to a higher incidence of interventions.

We restrict the analysis to the empirically relevant parameter range. First, we assume that we are in the third range of Proposition 2 ($b < \frac{1}{2} - \bar{\delta}$). Second, the analysis in Section 5.4 shows that we are in the parameter range where $n^* = 2$ ($1 + \sqrt{\frac{1+\delta}{2b}} < 3$). Furthermore we set $\bar{\delta} = 0$ as $\bar{\delta}$ does not affect intervention frequencies in the third range.

Assumption 1. $\frac{1}{8} < b < \frac{1}{2}$ *and* $\bar{\delta} = 0$.

From Proposition 1 we have that the threshold for sending the high message is

$$a_1 = \frac{4}{3}b + \frac{1}{3}.$$

If θ is below this threshold, the agent will thus send the low message and the principal sets a fine of $F_1 = 0$. When the state is above the threshold, the agent sends a high message and the principal will set a fine equal to its expected preferred fine in the high region: $F_2 = \frac{1+a_1}{2}$. Given that $a_1 > b$, it follows that there are hence strictly less interventions than under decentralization (the likelihood of intervention under decentralization is $\frac{1-b}{2}$ whereas under centralization it is $\frac{1-a_1}{2}$)

We can confirm that this is indeed an equilibrium as follows. First, at the threshold a_1, the agent is exactly indifferent to sending a high or low message. Second, we have that $F_1 = 0$ is optimal following m_1 since $a_1 < 1$. Moreover, the assumption $b < \frac{1}{2}$ guarantees that $a_1 < 1$.

3.5.1 Expected fine and fine dispersion

Under delegation, the expected fine in a supervisory intervention ($E(F|F>0)$) is lower (the expected fine under centralization is $\frac{1+a_1}{2}$, whereas the one under decentralization is $\frac{1-b}{2}$). There are two reasons for this. First, the threshold for intervention declines (since $b < a_1$). That is, there are now also interventions in banks with lower θ, which warrant lower fines. Second, under delegation the local supervisor sets the fines, and his preferred fine is lower.

Whereas fines are lower under decentralization, the (normalized) deviation of the fine ($\frac{E(|F-E(F|F>0)|}{E(F|F>0)|}|F>0)$) is higher. Intuitively, this is because under centralization the agent partitions information into intervals, so information about the value of θ within each interval is lost. In the two message equilibrium fine variability is even zero as there is only a single positive fine ($F2$).

3.5.2 Access to local information

Suppose that the principal sometimes also has access to information of the local supervisor. Specifically, assume that with a certain probability she observes θ as well. In such a situation she no longer requires information from the agent under centralization and will simply intervene when $\theta > 0$. Central interventions thus become more frequent (as they otherwise only occur when $\theta > a_1$). It follows that the effect of decentralization on the likelihood of intervention falls (though our estimates show that they are still positive). The intuition is that when the central supervisor is more informed, the agent's bias will play less a role in centralized supervisory decisions. Thus intervention decions under centralization are less distorted, lowering the impact of delegation.

3.5.3 Local uncertainty

The benefit of having access to local information is higher when local conditions are less predictable, that is, when they are more uncertain (Dessein (2002)). To understand the effect of uncertainty in our setting, suppose that with certain probability θ takes the mean value of 0 (instead of being distributed on $[-1, 1]$). In such a situation, the decentralized supervisor would never intervene, whereas the central supervisor would intervene with positive probability (for instance, because she randomizes when indifferent to intervention). Thus, removing uncertainty reverses the effect of decentralization on interventions. Given that the overall effect of decentralization is positive in our data, higher uncertainty hence increases the effect of decentralization on interventions.

3.5.4 Competing channel I: Higher risk of local banks

Around the time of the reform, there may have been a potentially confounding effect that has increased the risks of local banks (but not national banks). This may potentially explain why supervisory interventions have increased following the reform, completely unrelated to supervisory communication.

To examine, assume that the informational frictions of our model are absent. That is, the central supervisor also observes θ (either because she directly observes it, or because the local supervisor relays it truthfully to the central supervisor). Consider now, that for exogenous reasons, the risks of banks increases after the reform. Specifically, assume that the state θ is now uniformly distributed on $[-1 + r, 1 + r]$ ($r \in (0, 1)$) instead of $[-1, 1]$.

Prior to the reform, the fee under centralization is set according to $F = \max\{0, \theta\}$. The resulting likelihood of intervention is $\frac{1}{2}$ and the average fine in an intervention is $\frac{1}{2}$. Following the reform, the fee is set according to $F = \max\{0, \theta - b\}$. The likelihood of intervention is $\frac{1}{2}\frac{1+r-b}{2}$ and the average fine is $\frac{1+r-b}{2}$. We can see that interventions can

indeed increase when the risk effect is sufficiently large ($r > b$). However, in this case, fees should also increase, inconsistent with the evidence. The intuition is simple: higher interventions require local banks to become sufficiently more risky, but in this case, this would also imply higher fees.

3.5.5 Competing channel II: Higher stringency of local supervisors

The reform may also have increased the stringency of the local supervisor, specifically towards local banks. Again, this may cause more interventions into local banks, unrelated to an informational channel.

To investigate, we now assume that the reform lowers the bias of the local supervisor by $\triangle b$ ($\triangle b > 0$). Outcomes prior to the reform, that is, outcomes under centralization, are unchanged and identical to above (the likelihood of intervention and the average fine are both $\frac{1}{2}$). Following the reform, the fee will be set according to $F = \max\{0, \theta - b + \triangle b\}$. The likelihood of intervention is $\frac{1-b+\triangle b}{2}$ and the average fine is $\frac{1-b+\triangle b}{2}$. We see that interventions can indeed increase when $\triangle b$ is sufficiently large ($\triangle b > b$), but in this case the average fee should increase as well, and not decrease.

4 Data and Research Design

4.1 Sample composition and data sources

Supervisory interventions and resulting penalties are disclosed on the CBRC/CBIRC websites. The information disclosed contains the date of the action, the institution or individual to which the action applies, the supervisory authority responsible for the action, the reasons underlying the action, the specific laws or regulations violated, and the penalty (or penalties) imposed.

There are five types of penalties. First, a warning is a formal letter issued by a supervisor stating non-compliance with laws or regulations. Second, a fine entails a

monetary penalty imposed on a bank (like the civil money penalty in the US context), along with the confiscation of any illegal proceeds. Third, a disqualification refers to barring bank managers from holding senior positions in the banking sector (for a specified period or permanently). Fourth, a prohibition refers to barring bank staff from working in the banking sector (for a specified period or permanently). Fifth, a license revocation is the withdrawal of the authorization to operate the branch. Figure 2 provides an example of a penalty as disclosed on the CBRC/CBIRC website.[15]

We construct the sample of supervisory interventions using textual analysis of penalty announcements. Our sample includes 12,078 penalty events during the period between 2010 and 2020. Table 1 shows statistics on penalties. Panel A reports the frequency of penalties by types, highlighting that warnings and fines are the most issued type of penalties. Panel B breaks down penalties based on their underlying reasons, showing that in more than 50% of the cases the main reason is loan related (note that a bank may receive a penalty for several reasons). Panel C breaks down penalties based on their recipient (an individual or a bank), showing that banks are the most frequent recipients.

The information on penalties is manually collected at the disaggregated levels of (prefecture-level) branches and their subordinated offices. Since local supervisors operate at the prefecture level, we aggregate penalties at this level and allocate it to the (single) prefecture-level branch. The final sample covers 5,429 (prefecture) branches of 1,074 banks in 342 prefectures for ten years surrounding the 2015 reform. As shown in Table 2, the 1,074 banks consist of 1,056 local banks (127 prefecture commercial banks and 929 rural commercial banks[16]) and 18 national banks (6 large state-owned banks and 12 joint-stock banks). Overall, our sample represents about 90% of the assets of the Chinese banking sector.

[15] In 2015, the Yichang branch of Bank of Hubei was punished for inaccurately classifying loans based on their risk level and withholding a certain percentage of the loan as a deposit when the loan was granted. Consequently, the branch received a fine of RMB 400,000 from the local supervisory office, Yichang Bureau (Figure 2).

[16] We exclude rural credit cooperatives, village banks, foreign banks, internet banks, and similar institutions.

Table 2 presents several information about local and national banks. One takeaway from this table is that local and national banks are fairly comparable at the prefecture level. Across the sample, branches of local and national banks are of similar size and have profitabilities, and operate in prefectures of similar market concentration, credit availability and distance to Beijing. Branches of national banks have a slightly higher market share (6.6% compared to 5%).

We supplement the penalty data with bank-level, firm-level, prefecture-level, and loan-level data from the Chinese Research Data Services (CNRDS) and the China Stock Market and Accounting Research (CSMAR). Sample summary statistics are presented in Table 3. All variable definitions are given in Table A1 in the Appendix.

4.2 Empirical Specification

We use a difference-in-differences design to examine the impact of the decentralization reform of 2015 on supervisory interventions. We estimate the following specification at the bank-prefecture-year level:

$$Penalty_{ijt} = \alpha_i + \alpha_j + \alpha_t + \beta Localbank_i \times Post_t + \varepsilon_{ijt}, \tag{1}$$

where the subscript i denotes a specific bank i, j the prefecture of the bank branch, and t the year. The dependent variable, $Penalty_{ijt}$, is a one of our measures of penalties imposed on branch belonging to bank i in prefecture j in year t. We use two main variables capturing supervisory activity. The first variable is a dummy taking the value of 1 if a bank branch receives at least one penalty in year i, and 0 otherwise. The second variable is the log of 1 plus the number of penalties received by a bank branch in year i. In further tests, we also create similarly constructed variables for each type and each recipient of penalties. $Localbank_i$ is a dummy taking the value of 1 for branches of local banks i (treated group), and 0 for branches of national banks i (control group). $Post_t$

is a dummy taking the value of 1 from 2015 onwards, and 0 otherwise. The bank fixed effects (α_i) control for differences across banks, while the prefecture fixed effects (α_j) control for differences in local conditions. Since each prefecture has one supervisory office, the prefecture fixed effects also account for any (time-invariant) differences in local supervisory stringencies. We also include year fixed effects (α_t)) to control for any macro movements. In some specifications, we further include prefecture $\times$ year fixed effects, which control for time-varying prefecture-level heterogeneity. For instance, there may be a turnover of supervisory officers, resulting in a change in local supervisory leniency. In other specifications, we also include bank $\times$ prefecture fixed effects (equivalent to branch fixed effects), which control for heterogeneity across branches of the same bank. Following Gormley and Matsa (2014), our main specification does not include endogenous bank-level controls to avoid the "bad control" problem (see also Angrist and Pischke (2009)). ε_{ijt} is the error term. The coefficient of interest is β, which measures the effect of the 2015 decentralization reform on treated branches. We expect β to be positive if local banks are subject to tighter supervision following the reform. Throughout, we report robust standard errors clustered at the level of the prefecture where the branch is located.[17]

5 Decentralization and penalties

5.1 Basic results

We begin our analysis by examining the effect of the 2015 decentralization reform on supervisory interventions. Table 4 presents the results from estimating equation 1. The first four columns focus on the likelihood of receiving a penalty. In column 1, we run the regression without any fixed effects. In column 2, we add separate bank, prefecture,

[17] We experiment with alternative clustering levels—in particular, clustering by bank and double clustering by prefecture and year. Clustering by prefecture used in the reported results produces the most conservative standard errors.

and year fixed effects, implying that only the interaction between $Localbank_i \times Post_t$ is estimated. In column 3, we include prefecture $\times$ year fixed effects to absorb any time-varying local shocks, meaning that we compare branches of national versus local banks in the same prefecture of the same year. In column 4, we include bank $\times$ prefecture fixed effects and year fixed effects, which controls for differences across branches. In columns 1 to 4, the estimated coefficients on the interaction term (statistically significant at the 1% level) show that local banks are 6.3 to 8.8 pp more likely than national banks to receive a penalty following the decentralization reform. Given the unconditional probability of getting a penalty of 11.1%, the effect represents an increased probability ranging from 57 to 80%.[18]

The last four columns focus on the number of interventions as outcome. The results confirm the ones from the previous columns. The estimated coefficients on the interaction term (statistically significant at the 1% level) imply that local banks receive relatively more penalties following the decentralization as compared to national banks. We take the log of 1 plus the number of penalties to retain observations with zero-valued outcomes. However, linear regressions where the dependent variable is the log of 1 plus the count outcome are biased and may even produce the opposite sign of the true relationship (Cohn et al., 2022). An alternative to estimating equation 1 is to rely on Poisson regressions as they can accommodate count outcomes with a mass of values at zero. Poisson regressions produce consistent and reasonably efficient estimates under standard exogeneity conditions, even with multiple levels of fixed effects as we have. Table A3 in the Appendix reports the coefficient of a Poisson regression model. As can be observed in column 1, the Poisson regression confirms the finding of column 8 in Table 4.

[18] We carry out some robustness checks that are shown in Table A2 of the Appendix. First, these results are unchanged if we drop branches of state-owned banks. State-owned banks may have specific means to influence supervisory interventions, and may generally operate differently, while the remaining national banks (the joint-stock banks) operate in a more market-oriented manner as local banks. Second, these results are also unchanged if we exclude penalties issued by the local offices of the CBRC in Beijing. In this case, both central and local office are in the same prefecture, possibly blurring the analysis of decentralization.

Table 5 presents the results for each type and each recipient of penalties employing a specification similar to column 4 of Table 4. In Panel A, we examine each type of penalties, that is, fines, warnings, disqualification, and prohibition.[19] Columns 1 to 8 show that all types of penalties are more likely, and more often issued, against local banks than national banks after the reform. Regarding the fines, we use the log of the dollar amount instead of their counts. The size of fines post-reform is larger for local than national banks. In Panel B, we split penalties according to whether they are imposed on individuals or institutions. The results show that local supervisors issue more penalties towards both. Table A3 in the Appendix reports the coefficients of Poisson regression models for count outcomes. Again, the Poisson regressions confirm the findings of Table 5.

Table 6 then gauges the robustness of our results by controlling for several bank characteristics and differential trends based on bank characteristics. In Panel A, we control for bank size, capital adequacy ratio, Z-score, net interest margins, loan-to-deposit ratio, and loan-to-asset ratio. The inclusion of these bank covariates reduces sample size. We indeed lose 734 local banks—mostly small banks and banks that are not obligated to publicly disclose their financial statements because they are not listed. Panel A confirms the results reported in the first two columns of Table 4. However, these regressions should be interpreted with caution, as bank characteristics are likely endogenous to supervisory activity (Angrist and Pischke, 2009).

Adding bank covariates helps restoring randomness if banks with certain characteristics have different trends in outcome (that is, penalties). However, banks with certain characteristics could potentially be affected by the 2015 reform or other confounding events. For example, China had previously strict limits on the loan-to-deposit ratio, which capped the amount of loans that banks could make relative to their deposits. The mandatory limit was set at 75%, meaning banks could not lend more than 75% of

[19] We do not conduct regressions for license revocations as there are only three instances of such type of penalties.

their deposits. This ratio was considered as a crucial regulatory indicator, equivalent to capital adequacy ratio, with the aim of ensuring that banks had sufficient liquidity. In 2015, China removed the loan-to-deposit ratio as a mandatory requirement and turned it into a monitoring indicator instead.[20] We address this concern by interacting covariates (measuring various bank characteristics) averaged over the pre-reform period with $Post_t$. Doing so controls for non-random assignment (based on bank characteristics) and differential trends (based on bank characteristics). Panel B shows that including our interaction term of interest remains unaffected even after accounting for the differential effects of the 2015 reform on penalties across banks, which vary in size, capital adequacy ratio, Z-score, net interest margins, loan-to-deposit ratio, and loan-to-asset ratio.

Whereas so far we have controlled for bank-level controls, we next control for branch-level covariates. For this we utilize firm-level data originating from the National Tax Survey Database (NTSD), which is jointly collected by the State Administration of Taxation of China and the Ministry of Finance of China (SAT-MOF) based on the stratified random sampling method (Liu and Mao, 2019). It covers 700 thousand corporate tax payers annually from 2012-2020. A fraction of surveyed firms are banks and bank branches. We match bank (branch) name and city with our penalty sample, eventually end up with 3,037 branches of 559 banks across 333 prefecture-level cities. Based on the information available from the NTA, we construct a set of branch covariates, including branch size (log of branch total asset in RMB) and ROA (net income over total asset).[21] The results are presented in Panels C and D in Table 6 and show that our main results are unchanged.

[20] See Bloomberg, "China Moves to Scrap Rule Limiting Bank Loans to 75% of Deposits", https://www.bloomberg.com/news/articles/2015-06-24/china-moves-to-cancel-cap-limiting-bank-loans-to-75-of-deposits; and Reuters, "China to scrap commercial banks' loan-to-deposit ratio", https://www.reuters.com/article/business/china-to-scrap-commercial-banks-loan-to-deposit-ratio-idUSL3N0ZA3SA/ (last accessed: September 2024).

[21] The purposes of the dataset is for the Chinese authority to have better information on the tax base. As such, data quality is very variable. For instance, reported equity is negative for a large fraction of firms, and we have thus not included controls based on equity. For more detailed description of the NTSD, see Liu and Mao (2019).

5.2 Parallel trends

We now examine the dynamics of the effect. Figure 3 shows a version of our baseline specification that interacts the variable $Localbank_i$ with the time in years relative to the decentralization reform. Panel A plots the estimated coefficients for the likelihood of a penalty, while Panel B does the same for the number of penalties. The parallel trends assumption holds as there are no visible differences between the treated group and control group prior to the reform. At the same time, an increase is evident in the years subsequent to the reform. Table A4 in the Appendix also shows the regressions of the dynamics of the effect of decentralization (omitting the year 2015 as the benchmark). Again, no statistically significant effect exists in the years prior to the reform, and a clear increase appears following it.

5.3 Placebo tests

A potential concern in difference-in-differences analysis is that serial correlation may bias standard errors, in turn leading to over-rejection of the null hypothesis of no effect (Bertrand et al., 2004). We address this concern by performing a permutation test following Chetty et al. (2009) and Ohrn (2018), among others. We start the procedure by randomly selecting a placebo implementation year between 2010 and 2020 for each permutation. Then, we randomly designate 18 banks from the entire sample and assign them (and their branches) the status of national banks, while treating the remaining banks and their branches as local banks. The baseline regression (specification 4 of Table 4) is then re-estimated for each of our two dependent variables using the placebo treatment. Point estimates are recorded, and the procedure is repeated another 499 times to produce the plots in Figure 4. Both panels of Figure 4 display the empirical distribution of placebo effects for both dependent variables. Reassuringly, the estimated coefficients are normally distributed around 0 and are far away from the actual estimated

effects.

5.4 Interventions distortions and reform gains

We can use the estimated coefficients from Table 4 to gauge how much the bias distorts interventions and how informative communication between local and central supervisor is. For this we generalize the model in Section 3 by assuming that only with probability p ($\in (0,1)$) each year a situation materializes in which an intervention is possibly desirable (for instance, because a negative event at a bank materializes). In such a situation, the model described in Section 3 begins, and with remaining probability no interventions take place.

Denote the probability of an intervention under local and central supervision with p_L and p_C, respectively. From Table 4 column 4 we have that $p_L - p_C = 8.8\%$. The overall likelihood of intervention in the sample is 11% and the fraction of observations that are falling under local supervision throughout the sample period is 23.4%. From this we can calculate $p_L = 17.7\%$ and $p_C = 9.0\%$. We have for the likelihood of intervention that $p_L = p^{\frac{1-b}{2}}$ and $p_C = p^{\frac{1-a_1}{2}}$. From the fact that interventions increase (that is, $p_L > p_C$), we know that we are in the parameter range of part (iii) of Proposition 2. We can then use the equations for n^* and a_1 in Section 3.3 to obtain numerically the (unique) solution $b = 0.196$, $a_1 = 0.595$, $n^* = 2$ and $p = 44.1\%$.

The implied values suggest that the preferences of the local supervisor are considerably more lenient than of the central supervisor: In about 19.6% of cases where the central supervisor would want to intervene, the local supervisor does not want to intervene. Interventions are much more biased under central supervision: in about 59.5% of cases where the central supervisor would want to intervene, she ultimately does not intervene. As explained in Section 3.4 this is because strategic communication by the local supervisor reduces the informativeness of information that arrives at the central level. Notably, the resulting bias under central supervision is larger than under local

supervision.

The inefficiency of communication is also reflected in a low n^*, the number of messages that are send is equal to the minimum level ($n^* = 2$). That is, the local supervisor either sends a low message indicating that bank risk is low (following which the central supervisor does not intervene), or he sends a high message, following which there is intervention. Under this communication, significant information about the true health of the banks is withheld as within each risk classification (*low*, *high*) there is no longer information about the true underlying risk (the realization of θ). This means that supervisory decisions under central supervision do not well reflect bank conditions.

We can also calculate the model-implied welfare gain due to the reform, which is $\frac{-U_C+U_L}{-U_C} = 18.8\%$. That is, the reform eliminates 18.8% of the losses arising from inefficient interventions under centralized supervision[22], arising in particular because now also more intermediate forms of misconduct are caught.[23] Coupled with the fact that our analysis above has shown that there are large distortions under centralized supervision, this suggests that the reform has material welfare benefits.

5.5 Average fees and fee dispersion

The model of communication and supervisory interventions predicts that the severity of interventions declines following decentralization (even though interventions become more likely, see Proposition 3). Among others, this is because now more minor misconduct is punished, and thus misconduct that deserves a lower penalty.

Our data contains information on the size of fines issued, which allows us to test this prediction. Table 7 column (1) and (2) contain the results. The regression now only

[22] We can arrive at an alternative parameterization of the model using the estimated change in the fine severity instead of the change in the intervention probabilities. Using the average of the linear and log-specification estimates (column (1) and (2) in Table 7) and following similar calculations as above we obtain $b = 0.136$, $a_1 = 0.515$, $n^* = 2$ and $p = 41.1\%$ and a welfare gain of 33.6%.

[23] Efficient interventions have wider effects in terms of improving risk-taking of banks (see Section 8) and in terms of financial stability. To the extent that those are reflected in supervisory objectives, they are already incorporated in the welfare estimates

includes bank-prefecture-year observations with fines, and the dependent variable is the average fine in that bank-year (a bank may receive more than one fine in a year). As fines display a very skewed distribution, we use next to fines itself also the log of the fine. As can be observed, the interaction term is negative and statistically significant at the 10% and 5% level, respectively. This supports the prediction of, on average, more minor violations being punished following decentralization.

A second prediction of the model is that the severity of punishment becomes more variable following decentralization (Proposition 3). This is because under decentralization, more granular information is used for supervisory decision. Thus fines can be well tailored to varying levels of misconduct. By contrast, under centralization information that reaches the central supervisor is more partitioned. In the extreme, in the two-message equilibrium, there is only one information-level that leads to a punishment, and consequently there is only one fine-level issued.

Columns (3) and (4) apply this prediction to the dispersion of fine. The dependent variable is now the standard deviation of the fine using either the fine itself (column 3) or the log of the fine (column 4). The interaction effect is now positive (significant at 5% and 10%), indicating higher fine dispersion following the reform.

5.6 Local bias

Our model is predicated on local supervisors having a bias in their decision-making. As shown in Section 3, this bias distorts interventions regardless of centralized and decentralized decision-making, and does so in the same direction.

Such bias may arise from local supervisors being more susceptible to influence from local political interests. As the local supervisor is closer to the supervised bank, it has incentives to collude with local banks' stakeholders (Gopalan et al., 2021; Lim et al., 2023), resulting in lax supervisory outcomes (Correia, 2014; Lambert, 2019; Lim et al., 2019; Yue et al., 2022). Although local supervisory offices in China are in principle

independent of the government, local governments (that is, provincial-level or prefecture-level governments) might interfere in the supervisory process to protect their banks. This may particularly be so when local governments have an equity ownership, or even are the controlling shareholders of local banks.

We examine the presence of such a bias in Table 8. Our proxy is the total share of local governments among the top three shareholders of the bank. We obtain information on bank shareholders and their equity ownership from CNRDS. Shareholders are identified as being affiliated with the central or local government based on their names and registration information.[24] The estimated coefficients on local bias are negative (significant at 10% and 5% level), thus less interventions take place when local government are more likely to be subject to a bias. This is consistent with our model.

6 Local information

6.1 Access to information

Our model shows that under centralized decision-making local information is lost as the local supervisor communicates this information strategically and selectively with the central supervisor. The extent to which this worsens decision-making by the central supervisor depends on how much own information she has about local banks. Our model thus predicts that decentralization is more beneficial when it is carried out in a situation where the central supervisor had limited (own) access to local information (see section 3.6.1). In this case, decentralization entails a greater informational gain.

We examine this prediction in Table 9. We proxy the informational gain induced by the reform using the (log) distance in kilometres between the prefecture where the branch (of a local bank) is located and Beijing.[25] We then examine the differential

[24] Examples of local government entities that serve as shareholders for certain local banks on behalf of the local governments are Local Bureau of Public Finance, State-owned Assets Supervision and Administration Commission, Management Committee of Development Zones.

[25] Previous literature has offered US evidence that the distance from the firm location to the banking

effect of the reform on penalties across local banks that differ in terms of their distance to the central supervisor (proxying for informational gain). In columns 1 and 2, the specification includes year and bank × prefecture fixed effects. Whether we look at the likelihood of a penalty in column 1 or the number of penalties in column 2, the estimated coefficients on the triple interaction term are positive and statistically significant at the 1% level. In columns 3 and 4, we only exploit within-bank variation arising because different branches of a bank have different locations, and hence varying distances to the central supervisor. The specification in these two columns includes bank × year and prefecture × year fixed effects. The estimated coefficients on the triple interaction term are again positive and statistically significant at the 1% level. The implied effect is sizeable. Using the estimated coefficient on the triple interaction term of 0.022 in column 3, we find that a one-standard-deviation increase in the log distance (that is, 0.854 in Table 3) is associated with an increase of 1.9 pp in the probability of getting a penalty for local banks after the reform (that is, $0.022 \times 0.854 = 0.019$). We obtain implied effects in similar ranges once calculated from the estimated coefficients of the other columns. Together, these estimation results indicate that supervisory outcomes improve more when the central supervisor is more at an informational disadvantage, consistent with limited communication under centralization.[26]

In columns 5 and 6 of Table 8, we examine whether the informational gain is a linear one, or whether it predominantly arises for branches that are very far from Beijing. We create distance dummies (long, intermediate, and short) based on the tercile distribution that we interact with $Localbank_i \times Post_t$. The estimated coefficients on the triple interaction term are increasing from short, to intermediate, and eventually to long distance. It is noteworthy that already the coefficient on the triple interaction term for

regulator, the Department of Justice, or the Securities and Exchange Commission, affects the costs of monitoring and information acquisition (see, e.g., Kedia and Rajgopal (2011); Wilson and Veuger (2017); Ganduri (2019); Gopalan et al. (2021); Ha et al. (2023); Lim et al. (2023)).

[26] Our results are robust to using travel distance (measured by travel time between two prefectures of the supervisor and the bank) instead of distance in kilometres.

a short distance is sizeable. This speaks to the importance of eliminating even small informational frictions.

6.2 Uncertainty

The previous section considered a proxy of the amount of information that is lost under centralization. In this section we consider the importance of local information. In particular, our model predicts that information is more valuable when local conditions are more variable and hence more difficult to predict (for the central supervisor), see Section 3.6.2. In such situations, delegation should be more beneficial, resulting in a large improvement in detection of misconduct.

We measure local uncertainty through economic and political uncertainty at the prefecture-level. We approximate economic uncertainty by the standard deviation (SD) and coefficient of variation (CV) of prefecture-level city GDP growth rate over an event window of past three years (year $t-3$ to year $t-1$).[27] Then, we classify cities of high uncertainty if one has standard deviation or coefficient of variation higher than sample median each year. That is, the high uncertainty is a dummy variable equals one for higher SD or CV for a city than the sample median in each year, and 0 otherwise.

We report the results of local economic uncertainty in Panel A in Table 10. The first two columns report the results employing the penalty dummy as the dependent variable, while the last two columns employ the log of number of penalties as the dependent variable. Column 1 reports the regression with standard deviations (SD) of past GDP growth rates as the measure for local economic uncertainty. Having been interested in the triple interaction term of local bank, post, and high uncertainty, we confirm the prediction that local supervisors become tighter for local banks after the decentralization reform if local economic uncertainty is high. The results are stronger when uncertainty is measured by CV of GDP growth rates in column 2 and remain similar when using the

[27] Our findings are insensitive to alternative event window of past five years and are available upon request.

log of number of penalties as the dependent variable in columns 3 and 4.

We measure local political uncertainty using leadership vacancy or change in leadership less then mandatory term in local government. We manually collect the names and tenures of the top Party and government leaders at the prefecture-level over our sample period, including the precise dates of appointment and termination. First, if a leadership vacancy persists for more than six months for prefecture-level units within a given year, we classify it as a high-uncertainty period for a given prefecture. In specific, we define a dummy for High uncertainty (Vacancy $> 6m$). Second, following China's five-year term norm, we identify high-uncertainty episodes as leadership turnover within three years. In specific, we define a dummy for High uncertainty (Turnover $< 3y$).[28] We report the results of local political uncertainty in Panel B. Consistent with predictions, the persistently positive and significant coefficients on the local bank $\times$ post-reform $\times$ high uncertainty interaction terms indicate that decentralization reforms amplify supervisory stringency toward local banks in high-uncertainty political environments.

7 Other channels

Our findings regarding the likelihood, severity and dispersion of supervisory intervention as well as the importance of proxies for information speak to a better usage of local information under decentralization, resulting in better supervision. In this section we discuss whether alternative explanations are consistent with our evidence.

Bank risk-taking. A higher frequency of supervisory interventions may reflect higher risk-taking by banks. However, in Table 6 we have controlled for bank risk. In addition, the analysis in the following Section 8 shows that bank risk-taking has declined following the reform, not increased.

Supervisory capacity. A first alternative explanation for our results relates to

[28] Alternative four-year cutoff yields consistent results.

potential changes in the capacity of local supervisors (Eisenbach et al., 2022). Following the 2015 reform, local supervisors may have hired additional staff. More staff at local supervisory office may have allowed to improve risk detection irrespective of an informational benefit due to the reform. To examine this channel, we include as a control the (log) number of new hires of supervisory officials at the provincial level. We manually collected the data from the CBRC/CBIRC's official website. Table 11 displays the results. In columns 1 and 2, we can observe that our main effect remains after controlling for potential changes in supervisory capacity. We also note that the number of new hires does not enter significantly in the regression model, speaking against a "supervisory-capacity" channel in our setting. In addition, we restrict the sample to the period before 2018. The reason is the 2018 merger between the CBRC and the CIRC (see Sub-section 2.2 for institutional details), which does not allow identifying post-2018 hires specifically for bank supervision. In columns 3 and 4, we show that our results are unchanged when restricting the sample prior to 2018.

Accountability. Another explanation for our results could be that the 2015 reform clearly reassigned oversight responsibilities and powers to local supervisors, making them more accountable for their actions. This could lead to a fear of punishment by the central supervisor if problems at supervised banks arise, incentivizing local supervisors to adopt stricter oversight. While we cannot rule out the possibility of a fear-of-punishment motive, we observe that increased strictness contradicts our findings on average fines reported in Sub-section 5.5. In fact, fines decline rather than increase, which is consistent with improved risk detection following the 2015 reform.

Confounding reforms. Confounding policy reforms introduced in 2015 may have also affected local and national banks differently. Specifically, in 2015, the loan-to-deposit ratio ceiling was downgraded from a mandatory requirement to a voluntary monitoring indicator, marking a significant reform in China's banking regulation (cf. Footnote 18). If branches of local and national banks display different loan-to-deposit ratio prior to the

reform, they may be differentially impacted by it. In Table 5, Panel B, we controlled for the loan-to-deposit ratio (measured prior to the shock) interacted with the post-reform dummy. As explained in Sub-section 5.1, this approach allows us to effectively control for differential trends based on various bank characteristics, including the loan-to-deposit ratio. From Panel B, we saw that the interaction of interest between $Localbank_i$ and $Post_t$ keeps its positive sign and high statistical significance, suggesting that the ceiling downgrade on loan-to-deposit ratio does not affect our key results.

8 Decentralization and lending

A better supervisory ability to detect misconduct should lower banks' incentives to take on risks. We test this prediction by examining individual lending decisions of branches.[29] Specifically, we consider loan announcements of all listed firms in China over our sample period. We use textual analysis to extract information on the identity of borrowers, the loan origination date, the loan amount, the loan spread, and the entity of loan issuing bank branches. We start with 16,184 loans taken out by 1,678 firms from 327 banks. We obtain borrower financial information using CNRDS, namely size, leverage, tangibility, cash holdings, and ROA. We then merge this borrower financial information with information on branches, leaving us with 13,358 loans.

We use loan spreads (i.e., interest rates) and loan quantities (log of loan amounts) to measure conservativism in lending. Conditional on borrower characteristics, a more conservative branch is expected to charge a higher compensation for taking on credit risk, and to issue smaller loans, that is, take on less risk. All the variable definitions are provided in Table A1 of the Appendix and summary statistics on firm characteristics

[29] We cannot examine the impact on the riskiness of branches themselves as balance sheet information is not available. However, studying their loan extensions offers arguably a more suitable empirical setting as the granularity of the individual loan data allows to control for many factors, such as borrower fixed effects. In addition, this more directly measures risk-taking behavior as it focuses on new lending decisions.

and loan terms are presented in Panels C and D of Table 3.

We report the loan-level estimation results in Table 12. We first examine loan spreads. In column 1, the specification contains lagged borrower characteristics together with year, bank, and borrower fixed effects (identification thus arises from different banks lending to the same borrower). In column 2, we add prefecture fixed effects. In both columns, the number of observations is relatively small given that loan spreads are often missing. The estimated coefficients on the interaction term (statistically significant at the 1% level) indicate that local banks charge higher loan spreads following the 2015 reform. The coefficients of interest are 31.3-31.5 basis points, a remarkable effect of about 29% of the standard deviation of loan spreads. Thus, local banks show less aggressive lending behaviour post-reform by requiring higher compensation for identical borrowers. This result is consistent with our prediction that tighter supervision following decentralization is effective at reducing risk-taking incentives at banks. It should also be noted that most borrower controls are insignificant, suggesting that borrower risk is fairly time-invariant and hence well captured by the borrower fixed effects.

In columns 3 and 4, we look at loan quantities. The estimation results provide further support to our prediction. The estimated coefficients on the interaction term (statistically significant at the 1% level) show that local banks significantly reduce the amount they lend after the 2015 reform. The coefficients of interest are 3.3-3.9, that is, about 9% of the standard deviation of loan amounts. Since smaller loan size implies lower risk, this result is consistent with our prediction.[30]

The findings reported in Table 12 suggest that branches of local banks became more conservative in their lending because of decentralization. We next investigate whether such conservative lending at the bank level has real consequences for the aggregate, prefecture level. Specifically, we test whether loan supply in prefectures with a higher

[30]The number of listed firms has increased sharply during our sample period, suggesting that the sample of borrowers before and after the reform may differ substantially. To address this concern, we run the regressions in Table 12 restricting ourselves to borrowers already being listed prior to the reform and obtain similar results.

presence of local banks is lower relative to other prefectures. It is not clear ex ante that such effects happen in the aggregate. First, more conservative lending by local banks post-reform might be compensated by more accommodating lending by national banks. Second, as our analysis exploits variation across prefectures, we may fail to empirically identify a significant effect if prefectures do not vary significantly regarding the importance of local banks.

We use prefecture-level information on loan supply, GDP, and fiscal balance to construct a panel of 287 prefectures over the same sample period as before. We capture loan supply using the ratio of credit over GDP. The interaction term of interest is here between the share of local banks in the prefecture (measured by their share of offices) and the dummy $Post_t$. Our specification includes GDP growth and the fiscal balance of the municipal governments as controls. Robust standard errors are clustered at the prefecture level.

Table 13 displays the estimation results. In column 1, the estimated coefficient on the interaction term (statistically significant the 1% level) is negative, meaning a reduction in loan supply in prefectures with a higher share of local banks. In column 2, we obtain a similar result when including province fixed effects.

To deal with endogeneity, we rely on an instrumental variable (IV) that captures exogenous variations in the actual share of local banks. We closely follow Gilje et al. (2016) by using the predetermined share of local banks in 2010 as an IV. We first show that our IV is powerful, easily passing tests for weak instruments. It also meets the exclusion restriction because in 2010 banks (or prefectures) could not plausibly have anticipated the decentralization reform of 2015 and therefore adjust the structure of local banking markets. The IV results are presented in columns 3 and 4. We observe that they are very similar to the ones in the previous two columns.

Overall, the findings in this section imply that tighter supervision resulting from the decentralization reform has real aggregate effects. This is noteworthy as the change in

supervisory stringency is only applied to a subset of bank branches: the local ones.

9 Conclusion

This paper studies the effect of decentralization in the world's largest banking sector, China. In 2015, China shifted responsibilities and powers for undertaking supervisory interventions from central (national) to local (prefecture) supervisors. We find that local supervisors are more likely to intervene into branches of local banks following the decentralization reform. Economically, the likelihood of a penalty increases by 57% to 80%. This result, and a large set of others, are consistent with the prediction of a model of supervisory communication in the presence of local information, and suggests improved supervision due to better usage of local information.

Our paper has an important message. The current policy discussion (for example, in the Eurozone) often centres around centralization of policies vis-à-vis independent supervisors. Our paper by contrast speaks to benefits of decentralization vis-à-vis subordinated supervisors. In a hierarchical structure such as present in many countries, biases of local supervisors are less likely to play a role, favouring decentralization. Lower biases of local supervisors also speak towards a benefit of hierarchical supervision generally. Among others, this informs the debate in the Eurozone where frictions arise due to (independent) national supervisors having different objectives than the central supervisor.

References

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi, 2014, Inconsistent regulators: Evidence from banking, *The Quarterly Journal of Economics* 129, 889–938.

Aghion, Philippe, Nicholas Bloom, Brian Lucking, Raffaella Sadun, and John Van Reenen, 2021, Turbulence, firm decentralization, and growth in bad times, *American Economic Journal: Applied Economics* 13, 133–169.

Aghion, Philippe, and Jean Tirole, 1997, Formal and real authority in organizations, *Journal of Political Economy* 105, 1–29.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek, 2008, When does coordination require centralization?, *American Economic Review* 98, 145–179.

Ampudia, Miguel, Thorsten Beck, Andreas Beyer, Jean-Edouard Colliard, Agnese Leonello, Angela Maddaloni, and David Marques-Ibanez, 2019, The architecture of supervision, Working Paper 2287, European Central Bank.

Angrist, Joshua D, and Jörn-Steffen Pischke, 2009, *Mostly harmless econometrics: An empiricist's companion* (Princeton university press).

Baldenius, Tim, Nahum Melumad, and Xiaojing Meng, 2014, Board composition and ceo power, *Journal of Financial Economics* 112, 53–68.

Bandiera, Oriana, Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat, 2021, The allocation of authority in organizations: A field experiment with bureaucrats, *The Quarterly Journal of Economics* 136, 2195–2242.

Beck, Thorsten, C. Silva-Buston, and Wolf Wagner, 2023, The economics of supranational bank supervision, *Journal of Financial and Quantitative Analysis* 58, 324–351.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, How much should we trust differences-in-differences estimates?, *Quarterly Journal of Economics* 119, 249–275.

Bolton, Patrick, and Mathias Dewatripont, 1994, The firm as a communication network, *The Quarterly Journal of Economics* 109, 809–839.

Bolton, Patrick, and Martin Oehmke, 2019, Bank resolution and the structure of global banks, *The Review of Financial Studies* 32, 2384–2421.

Calzolari, Giacomo, Jean-Edouard Colliard, and Gyongyi Lóránth, 2019, Multinational banks and supranational supervision, *The Review of Financial Studies* 32, 2997–3035.

Carletti, Elena, Giovanni Dell'Ariccia, and Robert Marquez, 2021, Supervisory incentives in a banking union, *Management Science* 67, 455–470.

Chakraborty, Archishman, and Bilge Yilmaz, 2017, Authority, consensus, and governance, *The Review of Financial Studies* 30, 4267–4316.

Chetty, Raj, Adam Looney, and Kory Kroft, 2009, Salience and taxation: Theory and evidence, *American Economic Review* 99, 1145–1177.

Cohn, Jonathan B, Zack Liu, and Malcolm I Wardlaw, 2022, Count (and count-like) data in finance, *Journal of Financial Economics* 146, 529–551.

Colliard, Jean-Edouard, 2020, Optimal supervisory architecture and financial integration in a banking union, *Review of Finance* 24, 129–161.

Correia, Maria M., 2014, Political connections and sec enforcement, *Journal of Accounting and Economics* 57, 241–262.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, 2020, Fast poisson estimation with high-dimensional fixed effects, *Stata Journal* 20, 95–115.

Crawford, Vincent P, and Joel Sobel, 1982, Strategic information transmission, *Econometrica* 1431–1451.

Dell'Ariccia, Giovanni, and Robert Marquez, 2006, Competition among regulators and credit market integration, *Journal of Financial Economics* 79, 401–430.

Dessein, Wouter, 2002, Authority and communication in organizations, *The Review of Economic Studies* 69, 811–838.

Eisenbach, Thomas M., David O. Lucca, and Robert M. Townsend, 2022, Resource allocation in bank supervision: Trade-offs and outcomes, *Journal of Finance* 77, 1685–1736.

Foarta, Dana, 2018, The limits to partial banking unions: A political economy approach, *American Economic Review* 108, 1187–1213.

Ganduri, Rohan, 2019, Too far to regulate?, Working Paper.

Gilje, Erik P., Elena Loutskina, and Philip E. Strahan, 2016, Exporting liquidity: Branch banking and financial integration, *Journal of Finance* 71, 1159–1184.

Gopalan, Yadav, Ankit Kalda, and Asaf Manela, 2021, Hub-and-spoke regulation and bank leverage, *Review of Finance* 25, 1499–1545.

Gormley, Todd A, and David A Matsa, 2014, Common errors: How to (and not to) control for unobserved heterogeneity, *The Review of Financial Studies* 27, 617–661.

Grenadier, Steven R., Andrey Malenko, and Nadya Malenko, 2016, Timing decisions in organizations: Communication and authority in a dynamic environment, *American Economic Review* 106, 2552–2581.

Ha, Sehoon, Fei Ma, and Alminas Zaldokas, 2023, Motivating collusion, Working Paper.

Harris, Milton, and Artur Raviv, 2005, Allocation of decision-making authority, *Review of Finance* 9, 353–383.

Harris, Milton, and Artur Raviv, 2008, A theory of board control and size, *The Review of Financial Studies* 21, 1797–1832.

Haselmann, Rainer, Shikhar Singla, and Vikrant Vig, 2022, Supranational supervision, Technical report, LawFin Working Paper.

Kala, Namrata, 2024, The impacts of managerial autonomy on firm outcomes, *Econometrica* 92, 1777–1800.

Kara, Gazi, 2016, Systemic risk, international regulation, and the limits of coordination, *Journal of International Economics* 99, 192–222.

Kedia, Simi, and Shivaram Rajgopal, 2011, Do the sec's enforcement preferences affect corporate misconduct?, *Journal of Accounting and Economics* 51, 259–278.

Lambert, Thomas, 2019, Lobbying on regulatory enforcement actions: Evidence from us commercial and savings banks, *Management Science* 65, 2545–2572.

Li, Danielle, 2017, Expertise versus bias in evaluation: Evidence from the nih, *American Economic Journal: Applied Economics* 9, 60–92.

Lim, Ivan, Jens Hagendorff, and Seth Armitage, 2019, Is the fox guarding the henhouse? bankers in the federal reserve, bank leverage and risk-shifting, *Journal of Corporate Finance* 58, 478–504.

Lim, Ivan, Jens Hagendorff, and Seth Armitage, 2023, The effects of regulatory office closures on bank behavior, *Journal of Money, Credit and Banking* 55, 1927–1963.

Liu, Yongzheng, and Jie Mao, 2019, How do tax incentives affect investment and productivity? firm-level evidence from china, *American Economic Journal: Economic Policy* 11, 261–291.

Lóránth, Gyöngyi, Anatoli Segura, and Jan Zeng, 2022, Voluntary support and ring-fencing in cross-border banks, Working Paper.

Malenko, Nadya, 2014, Communication and decision-making in corporate boards, *The Review of Financial Studies* 27, 1486–1532.

Malenko, Nadya, 2024, Information flows, organizational structure, and corporate governance, in David J. Denis, ed., *Handbook of Corporate Finance*, 511–546 (Edward Elgar Publishing).

Melumad, Nahum D., and Stefan Reichelstein, 1987, Centralization versus delegation and the value of communication, *Journal of Accounting Research* 1–18.

Ohrn, Eric, 2018, The effect of corporate taxation on investment and financial policy: Evidence from the dpad, *American Economic Journal: Economic Policy* 10, 272–301.

Repullo, Rafael, 2018, Hierarchical bank supervision, *SERIEs* 9, 1–26.

Wilson, Karen, and Stephanie Veuger, 2017, Information frictions in uncertain regulatory environments: Evidence from us commercial banks, *Oxford Bulletin of Economics and Statistics* 79, 205–233.

Yue, Heng, Li Zhang, and Qiong Zhong, 2022, The politics of bank opacity, *Journal of Accounting and Economics* 73, 101452.

10 Appendix: Proofs

10.1 Proof of Proposition 1.

Proof. Following Crawford and Sobel (1982) and Dessein (2002), any cheap talk equilibrium is characterized by a partition $(a_0, a_1, a_2, ..., a_n)$ of $[-1, 1]$ with $a_0 = -1, a_n = 1$, and $a_{k-1} < a_k$ for $k \in \{1, .., n\}$, and a set of messages $M = \{m_1, ...m_n\}$ such that:

(i) the agent sends message m_k whenever $\theta \in (a_{k-1}, a_k)$;

(ii) upon receiving the message m_k, the principal issues a fine F_k given by

$$F_k = \arg\max E(U(F, \theta)|\theta \in (a_{k-1}, a_k))$$

or still

$$F_k = \max\left\{0, \frac{a_{k-1} + a_k}{2} + \delta\right\}$$

(iii) $U_A(F_k, a_k) = U_A(F_{k+1}, a_k)$ or still

$$2\left(a_k + \delta\right) = F_k + F_{k+1} + 2b \tag{2}$$

Off-equilibrium, we assume the principal issues a fine $F \geq F_n$ whenever the agent sends a message $m \notin M$. It is further without loss to impose that $F_k > 0$ for $k > 1$. Indeed, if $F_1 = F_2 = 0$, then the agent is indifferent between sending messages m_1 and m_2 for $\in (a_0, a_2)$ as both messages induce the same action. But then this equilibrium is equivalent (in terms of induced actions) to the agent always sending message m_1 for $\theta \in (a_0, a_2)$.

A key difference with Dessein (2002) and the leading example in Crawford and Sobel (1982) is that fines (and desired fines) are bounded below by 0. As a result, it matters for our characterization of the cheap talk equilibrium whether $F_1 = 0$ (corner solution

for F or $F_1 > 0$. We have that

$$E(U(F_1,\theta)|m_1) = -\frac{1}{2}\left[\int_{-1}^{-\delta} F_1 d\theta + \int_{-\delta}^{F_1-\delta} (F_1 - (\theta+\delta))d\theta + \int_{F_1-\delta}^{a_1} (\theta+\delta-F_1)d\theta\right],$$

where one can verify that $\arg\max_{F_1} E(U(F_1,\theta)|m_1)$ has a corner solution of $F_1 = 0$, if and only if $\frac{a_1-1}{2} + \delta < 0$, or still if and only if $a_1 + \delta < 1 - \delta$.

Equilibria where $F_1 = 0$. Consider first all cheap talk equilibria where $F_1 = 0$, which will require that $a_1 + \delta < 1 - \delta$.

Since $F_k = \frac{a_{k-1}+a_k}{2} + \delta$ for $k > 1$, it follows from (2) that

$$\begin{aligned} 2a_1 &= \frac{a_1 + a_2}{2} - \delta + 2b \\ 2a_k &= \frac{a_{k-1} + a_k}{2} + \frac{a_k + a_{k+1}}{2} + 2b \text{ for } n^* > 2, k > 1, \end{aligned}$$

from which

$$\begin{aligned} a_1 + \delta &= \frac{1}{2}(a_2 - a_1) + 2b \\ a_k - a_{k-1} &= a_{k+1} - a_k + 4b, \text{ for } n^* > 2, k > 1 \end{aligned}$$

Let $x \equiv a_n - a_{n-1}$, then

$$a_k - a_{k-1} = x + 4(n-k)b$$

and

$$\begin{aligned} a_1 + \delta &= 2b + \frac{1}{2}(a_2 - a_1) \\ &= 2b + 2(n-2)b + \frac{x}{2} \\ &= 2(n-1)b + \frac{x}{2} \end{aligned}$$

It follows that

$$\begin{aligned} 1+\delta &= \sum_{k=2}^{k=n}(a_k - a_{k-1}) + a_1 + \delta \\ &= (n-1)x + \sum_{k=2}^{k=n} 4(n-k)b + 2(n-1)b + \frac{x}{2} \\ &= 2b\,(n-1)^2 + (n - \frac{1}{2})x \end{aligned}$$

As in Dessein (2002), we will focus on the equilibrium with the largest number of messages subject to $F_2 > F_1$, denoted by $n = n^*$. Since $x = a_n - a_{n-1} > 0$, we must have that $2b\,(n-1)^2 < 1+\delta$. It follows that n^* is the largest integer smaller than

$$1 + \sqrt{\frac{1+\delta}{2b}}$$

Hence, informative communication ($n^* > 1$) can be supported if and only if

$$b < \frac{1}{2} + \frac{\delta}{2}$$

It further follows that

$$x = \frac{(1+\delta) - 2b\,(n-1)^2}{n - 1/2}$$

from which

$$a_1 + \delta = 2(n-1)b + \frac{x}{2}$$

Since $F_1 = 0$ is optimal if and only $a_1 + \delta < 1 - \delta$,

$$F_1 = 0 \iff b \cdot (n^* - 1)\left(\frac{2n^*}{2n^* - 1}\right) + \frac{1+\delta}{2n^* - 1} < 1 - \delta$$

For $n^* = 2$, this is equivalent to

$$F_1 = 0 \iff b < \frac{1}{2} - \delta.$$

More generally, one can show that[31]

$$a_1 + \delta \in \left(\frac{1+\delta}{n^*}, \frac{1+\delta}{n^*-1} \right)$$

Hence, $b < \frac{1}{2} - \delta$ is also a sufficient condition for $F_1 = 0$ when $n^* > 2$.

When $\delta < 0$, we have that $\frac{1}{2} + \frac{\delta}{2} < \frac{1}{2} - \delta$. Hence, for $b < \frac{1}{2} + \frac{\delta}{2}$ there is informative communication ($n^* > 1$) with $F_1 = 0$. For $b > \frac{1}{2} + \frac{\delta}{2}$, there is no informative communication, and the principal never intervenes, since $E(\theta + \delta) < 0$.

In contrast, whenever $\delta > 0$, we have that $\frac{1}{2} - \delta < \frac{1}{2} + \frac{\delta}{2}$. There is then only informative communication with $F_1 = 0$, whenever $b < \frac{1}{2} - \delta$. As we show in the next subsection, for $b \in (\frac{1}{2} - \delta, 1/2)$, there is still informative communication, but with $F_1 > 0$ (so the principal always intervenes). Finally, for $b > 1/2$, there then no communication (and the principal always intervenes as $E(\theta + \delta > 0)$

Equilibria where $F_1 > 0$. Consider now all partition equilibria where $F_1 > 0$ and hence the principal always intervenes. Note that this requires that $a_1 + \delta > 1 - \delta$. As we show below, in equilibrium, this will be equivalent to $b > \frac{1}{2} - \delta$. For $F_1 > 0$ and $n \geq 2$, we must have that for $k \geq 1$

$$2a_k = \frac{a_{k-1} + a_k}{2} + \frac{a_k + a_{k+1}}{2} + 2b$$

from which

$$a_k - a_{k-1} = a_{k+1} - a_k + 4b$$

[31] Indeed, $1 + \delta = 2b\,(n-1)^2 + (n - \frac{1}{2})x$ implies that $\frac{(1+\delta)}{2n^2} \leq b \leq \frac{(1+\delta)}{2(n-1)^2}$.

Let $x \equiv a_n - a_{n-1}$, then

$$a_k - a_{k-1} = x + 4(n-k)b$$

It follows that

$$\begin{aligned} 2 &= \sum_{k=1}^{k=n}(a_k - a_{k-1}) \\ &= nx + \sum_{k=1}^{k=n} 4(n-k)b \\ &= nx + 2bn(n-1) \end{aligned}$$

Hence, n^* is the largest integer such that $2bn(n-1) < 2$, or still, n^* is the largest integer smaller than

$$\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4}{b}}$$

It follows that informative communication ($n^* > 1$) is feasible if and only if $b < \frac{1}{2}$.

For $n^* = 2$, we further have that $a_1 + 1 = 1 - a_k + 4b$ from which $a_1 = 2b$. It follows that the condition for $F_1 > 0$, namely $a_1 + \delta > 1 - \delta$, will be satisfied if and only if $b > 1 - \delta$.

We conclude that an equilibrium with $n^* \geq 2$ and $F_1 > 0$ exists if and only if $1/2 - \delta < b < 1/2$. Whenever $\delta > 0$ and $b > \frac{1}{2}$, no informative communication is possible and the principal always issues a fine since $E(\theta + \delta) > 0$ □

10.2 Proof of Proposition 2

Proof. We only prove that delegation is optimal for $b \in (\frac{1}{2} - \bar{\delta}, 1 - \sqrt{\bar{\delta}\left(1 - \bar{\delta}\right)})$. The remainder of proposition 2 follows directly from results in proposition 1 and/or the discussion following proposition 2, and is omitted. Expected pay-off under delegation given δ :

$$-\frac{1}{2}\int_{-\delta}^{b-\delta}(\theta+\delta)\,d\theta-\frac{1}{2}\int_{b-\delta}^{1}bd\theta=-\frac{1}{2}b(1+\delta)+\frac{1}{4}b^2$$

Since $\delta=-\bar{\delta}$ with 50% probability and $\delta=\bar{\delta}$ with 50% probability, we thus have that expected pay-off equals

$$EU_D=-\frac{1}{2}b+\frac{1}{4}b^2$$

Expected pay-off under centralization and no communication when $\delta=-\bar{\delta}<0$. Since $\delta<0$, the principal issues no fine under no communication and its expected pay-off equals

$$\frac{1}{2}\left(-\int_{\bar{\delta}}^{1}(\theta-\bar{\delta})d\theta\right)=-\frac{1}{4}(1-\bar{\delta})^2$$

Expected pay-off under centralization, no communication, when $\delta=\bar{\delta}>0$. Since $\delta>0$, the principal issues a fine $F=\bar{\delta}$ under no communication (can be shown to be optimal), resulting in the following expected pay-off

$$\begin{aligned}
&\frac{1}{2}\left(\int_{-1}^{-\bar{\delta}}(-F)d\theta\right)+\frac{1}{2}\left(\int_{-\bar{\delta}}^{0}(\theta+\bar{\delta}-F)d\theta\right)-\frac{1}{2}\left(\int_{0}^{1}(\theta+\bar{\delta}-F)d\theta\right)\\
=\;&\frac{1}{2}\left(\int_{-1}^{-\bar{\delta}}(-\bar{\delta})d\theta\right)+\frac{1}{2}\left(\int_{-\bar{\delta}}^{0}\theta d\theta\right)-\frac{1}{2}\left(\int_{0}^{1}\theta d\theta\right)\\
=\;&-\frac{1}{4}(1+2\bar{\delta}-\bar{\delta}^2)
\end{aligned}$$

Hence, since $\delta=-\bar{\delta}$ with 50% probability and $\delta=\bar{\delta}$ with 50% probability, we thus have that expected pay-off under centralization equals

$$\begin{aligned}
EU_C&=-\frac{1}{4}(1-\bar{\delta})^2-\frac{1}{4}\bar{\delta}\\
&=-\frac{1}{4}(1-\bar{\delta}+\bar{\delta}^2)
\end{aligned}$$

It follows that delegation is preferred over communication (without communication)

whenever $EU_D > EU_C$ or still:

$$2b - b^2 < 1 - \bar{\delta} + \bar{\delta}^2$$

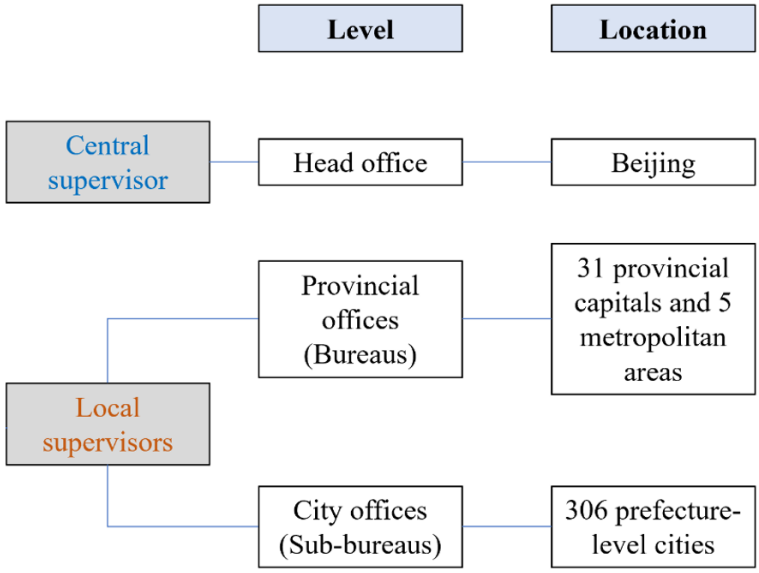
or still (since $b < 1$), delegation is optimal provided that

$$b < \bar{b} \equiv 1 - \sqrt{\bar{\delta}\left(1 - \bar{\delta}\right)}$$

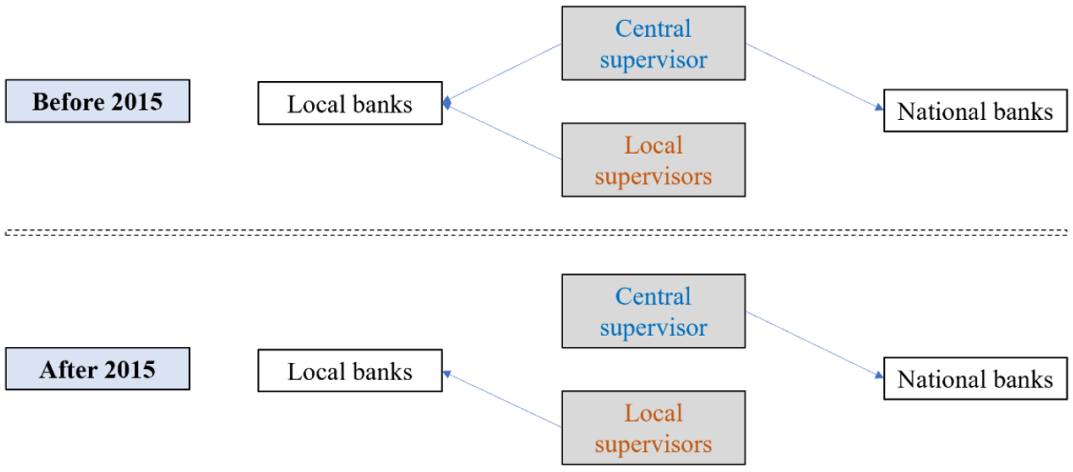
Since $\bar{\delta} < 1/2$, we have that $\bar{b} \in (1/2, 1)$.

□

Figure 1: Supervisory architecture
Note: The figures show the structure of China Banking Regulatory Commission (Panel A) and the allocation of responsibilities and powers before and after the decentralization reform of 2015 (Panel B).



(a) Structure of China Banking Regulatory Commission (CBRC)



(b) Responsibilities and powers for supervisory interventions before and after 2015

Figure 2: Example of a penalty

Note: The figure shows a snapshot of a randomly chosen penalty from the website of the CBIRC (source: www.cbirc.gov.cn/branch/hubei/view/pages/common/ItemDetail.html?docId=107940&itemId=1437&generaltype=0; last accessed: August 2023).

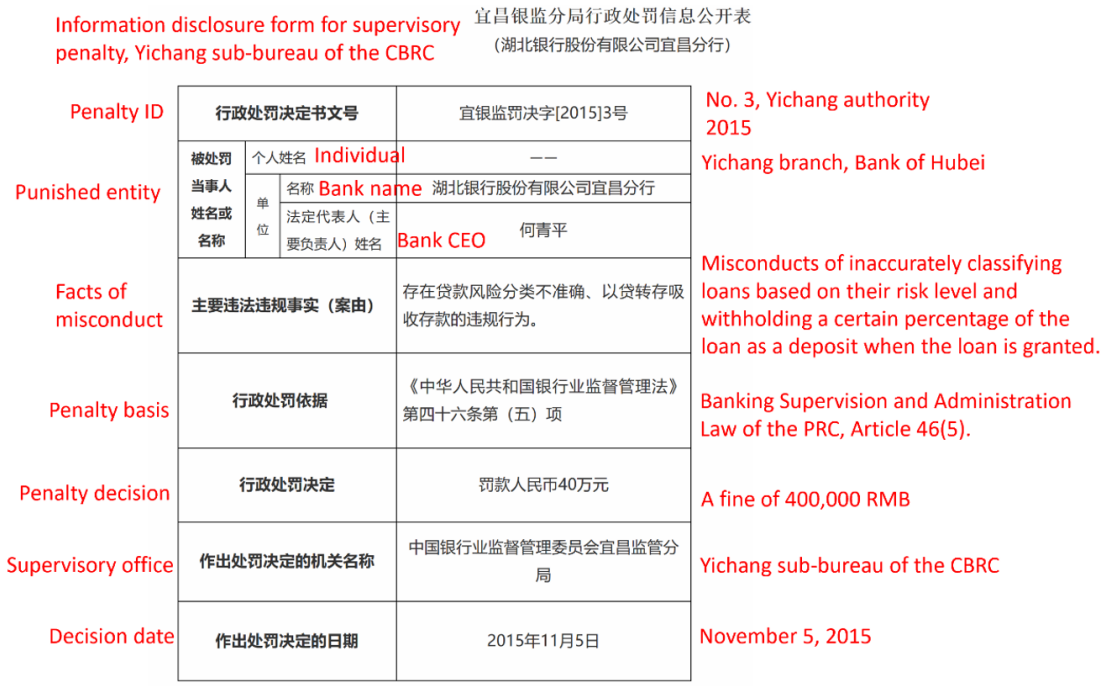
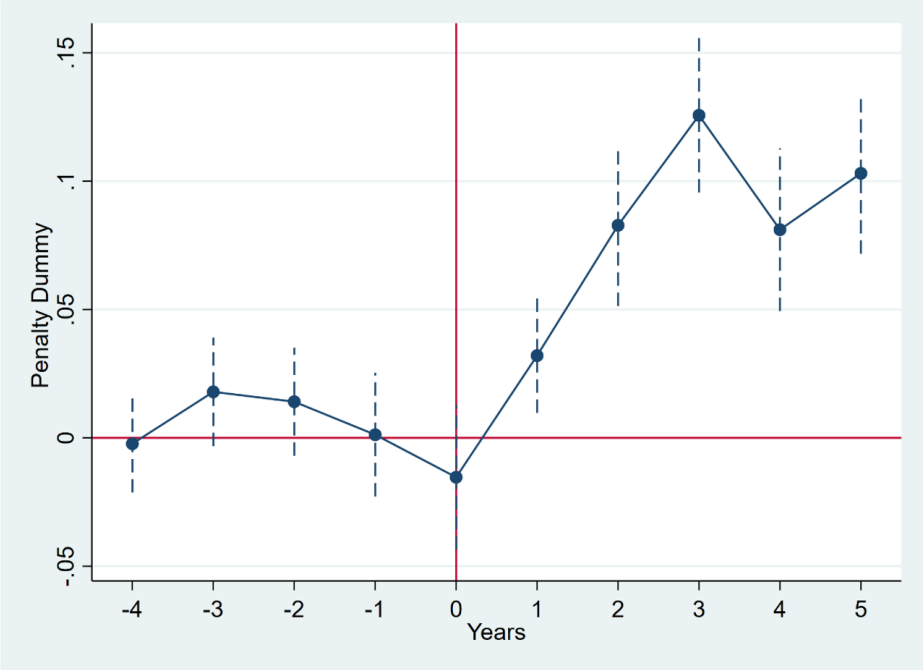
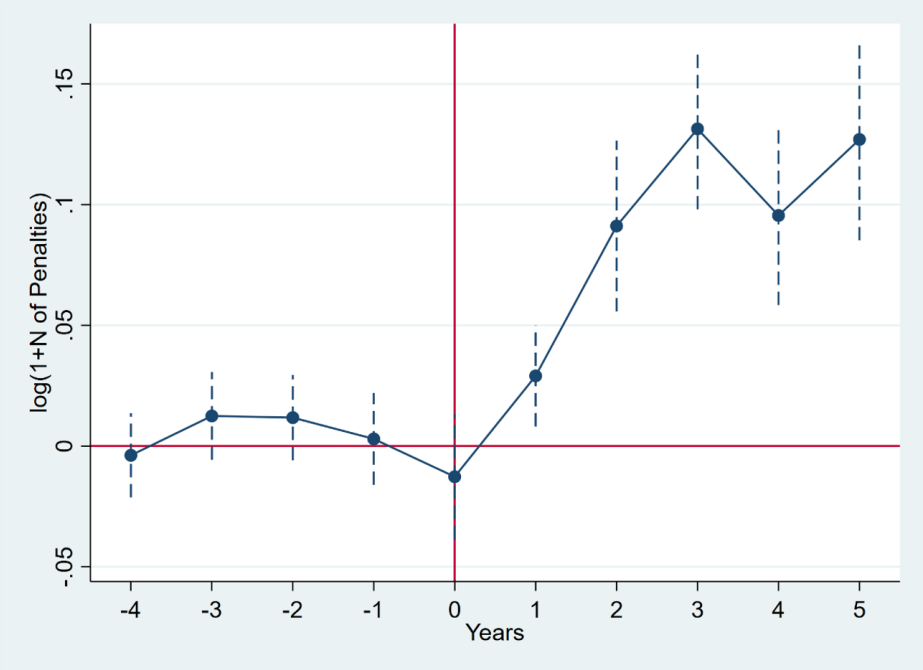


Figure 3: Parallel trends

Note: The figures show the parallel trends for the penalty dummy (Panel A) and number of penalties (log) (Panel B) over our the period surrounding the 2015 decentralization reform. Year 0 refers to year 2015.



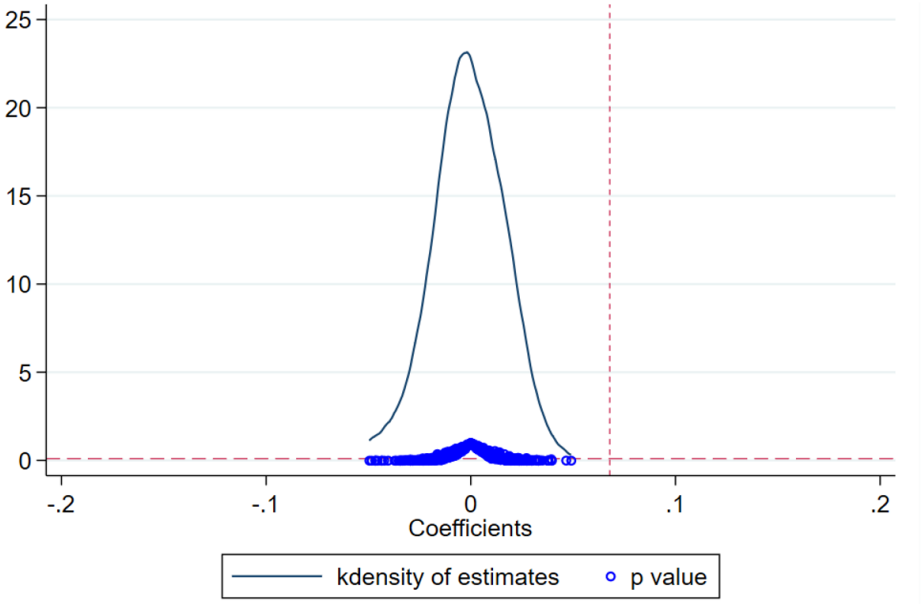
(a) Likelihood of a penalty



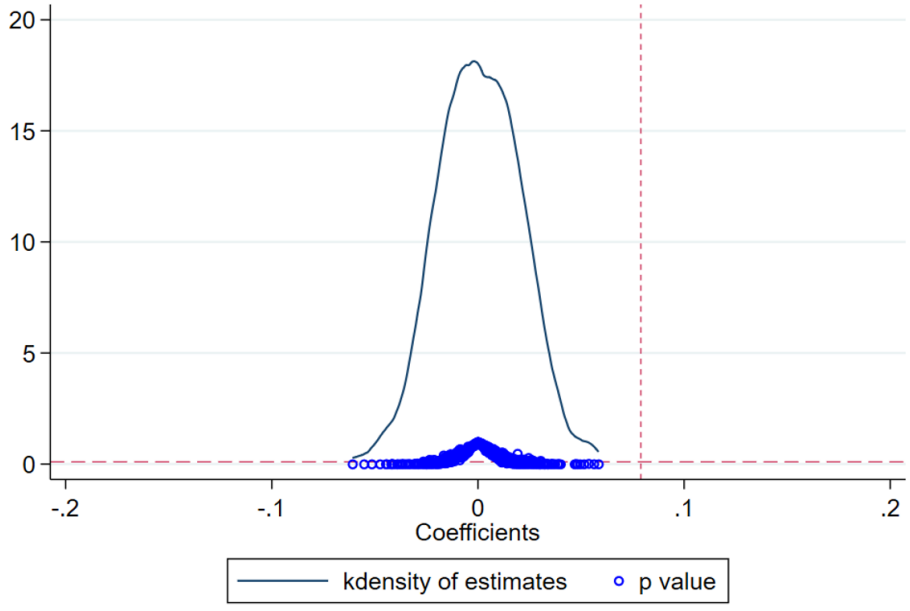
(b) Number of penalties

Figure 4: Distribution of placebo estimates

Note: The figures show the empirical distributions of placebo effects for each of the two dependent variables of interest as described in Sub-section 5.3. The vertical red dash line is the actual estimated coefficient of specificaiton 4 in Table 4, the horizonal red dash line is p-value of 10%, and the solid blue line (blue circles) is the distribution of coefficients (p-values) resulting from the 500 placebo tests for the penaly dummy (Panel A) and number of penalties (log) (Panel B).



(a) Likelihood of a penalty



(b) Number of penalties

Table 1: Summary statistics for the penalty sample

Note: The table presents summary statistics on penalties imposed on local and national banks. Panel A shows the breakdown by type, Panel B by reason, and Panel C by recipient. All variables are defined in Table A1 of the Appendix.

Panel A: Type of penalties	N	Local banks	National banks
Warning	4,341	2,431	1,910
Fine	8,589	4,050	4,539
Disqualification	365	192	173
Prohibition	376	194	182
License revocation	3	0	3
Panel B: Reason of penalties	**N**	**Local banks**	**National banks**
Loan-related reasons	6,872	3,638	3,234
Deposit-related reasons	816	266	550
Interbank-related reasons	464	284	180
Acceptance-related reasons	1,536	664	872
Credit Card-related reasons	182	27	155
Guarantee-related reasons	358	149	209
Prudential regulation-related reasons	1,577	751	826
Internal control-related reasons	973	289	684
Governance-related reasons	298	269	29
Panel C: Recipient of penalties	**N**	**Local banks**	**National banks**
Individuals	4,749	2,693	2,056
Banks	7,919	3,430	4,489
Both individuals and banks	457	240	217

Table 2: Sample composition
Note: The table presents the composition of national and local banks in our sample and summarizes some information at the branch, bank, and aggregate levels.

Bank type	Local banks	National banks
Branch-level statistics		
Average branch size (log of total assets in billion RMB)	11.76	11.79
Average branch ROA	0.98%	1.00%
Average market share (% offices)	5.03%	6.56%
Average number of penalties per branch	2.8	1.85
Average dist. of branches' pref. capital to Beijing (km)	1101.98	1183.73
Average credit/GDP of branches' prefecture	108.04%	112.79%
Average HHI of branches' prefecture (% offices)	0.09	0.1
Bank-level statistics		
Average number of branches per bank	2.03	182.56
Average number of penalties per bank	5.68	337.78
Aggregate-level statistics		
Number of banks in the full sample	1056	18
Number of penalties in the full sample	5,998	6,080
Total fine amount (million)	1761.7	3690.69

Table 3: Summary statistics for the full sample

Note: The table presents summary statistics for the full sample. Panel A reports them for branch-level variables, Panel B for bank-level variables, Panel C for loan-level variables, and Panel D for prefecture- and province-level variables. All variables are defined in Table A1 of the Appendix.

Panel A: Branch-level variables	N	Mean	SD	P1	Median	P99
Local bank	52,773	0.383	0.486	0	0	1
Penalty dummy	52,773	0.11	0.313	0	0	1
Number of penalties	52,773	0.229	1.05	0	0	4
Fine dummy	52,773	0.105	0.307	0	0	1
Fine amount	52,773	103.318	3,361.84	0	0	1,200
Average fine	3,871	679.844	8704.643	26	300	4513.333
Average fine (log)	3,871	5.696	0.887	3.258	5.704	8.415
Fine dispersion	3,871	0.337	0.32	0	0.26	0.984
Fine dispersion (log)	3,871	0.41	0.441	0	0.266	1.478
Warning dummy	52,773	0.082	0.67	0	0	1
Number of warnings	52,773	0.005	0.068	0	0	2
Disqualification dummy	52,773	0.007	0.117	0	0	0
Number of disqualifications	52,773	0.005	0.071	0	0	0
Prohibition dummy	52,773	0.007	0.118	0	0	0
Number of prohibitions	52,773	1,152.45	602.336	0	0	0
Distance (km)	52,773	6.845	0.854	47.225	1,080.20	2,771.43
Distance (log)	52,773	0.783	0.393	3.876	6.986	7.927
Share of offices outside the prefecture	52,773	0.383	0.486	0	0.993	1
Branch size (log)	18,819	11.777	1.294	9.14	11.756	14.1
Branch leverage ratio (%)	18,823	8.699	22.002	-1.999	1.48	90.837
Branch ROA (%)	18,823	0.996	1.371	-2.03	0.986	4.095
Panel B: Bank-level variables	**N**	**Mean**	**SD**	**P1**	**Median**	**P99**
Bank size (log)	44,722	7.636	1.97	2.014	8.675	9.134
CAR (%)	44,028	13.259	2.022	8.48	13.17	17.72
Z-score	42,671	5.413	0.909	3.338	5.39	7.471
NIM (%)	44,595	2.115	0.582	0.225	2.099	4.075
Loan-to-deposit (%)	44,594	71.796	15.761	33.163	73.046	111.223
Loan-to-asset (%)	44,459	49.486	8.955	25.112	51.812	64.152
Local ownership (%)	37,793	1.503	4.964	0	0	21.03
Panel C: Loan-level analysis	**N**	**Mean**	**SD**	**P1**	**Median**	**P99**
Loan spread	7,474	26.625	106.298	-335	15	369
Loan amount	13,358	106.177	623.442	0.5	30	1,000.00
Firm size	13,358	8.653	1.15	6.347	8.542	11.716
Firm leverage (%)	13,358	51.253	18.436	10.709	51.239	93.989
Firm tangibility (%)	13,358	20.812	16.264	0.138	17.268	66.911
Firm cash holdings (%)	13,358	15.162	9.592	1.177	13.091	49.742
Firm ROA (%)	13,358	2.349	6.618	-38.937	2.807	14.848
Panel D: Prefecture- and province-level variables	**N**	**Mean**	**SD**	**P1**	**Median**	**P99**
Share of local banks	3,136	0.381	0.154	0.008	0.382	0.688
Credit-to-GDP (%)	3,136	99.935	57.851	30.199	82.423	335.777
GDP growth (%)	3,136	8.704	4.164	-4.8	8.4	18.2
Fiscal balance (%)	3,136	12.345	10.279	-0.083	9.787	56.609
High uncertainty (SD of GDP growth)	3,653	0.501	0.5	0	1	1
High uncertainty (CV of GDP growth)	3,653	0.501	0.5	0	1	1
High uncertainty (Vacancy¿6m)	3,718	0.02	0.142	0	0	1
High uncertainty (Turnover¡3y)	3,718	0.317	0.465	0	0	1
Province NPL (%)	52,773	1.677	0.875	0.54	1.46	4.57
New hires (log)	52,773	3.561	0.624	1.099	3.664	4.431

Table 4: Decentralization and penalties: Basic results

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1. Columns 1-4 present results using the likelihood of a penalty as dependent variable, while columns 5-8 present results using the number of penalties (log) as dependent variable. Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Penalty dummy				Number of penalties			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local bank	-0.015***				-0.012***			
	(0.003)				(0.002)			
Post	0.105***				0.110***			
	(0.007)				(0.008)			
Local bank×Post	0.067***	0.075***	0.063***	0.088***	0.078***	0.088***	0.073***	0.104***
	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)	(0.009)	(0.010)
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Bank FE	No	Yes	Yes	No	No	Yes	Yes	No
Prefecture FE	No	Yes	No	No	No	Yes	No	No
Prefecture×Year FE	No	No	Yes	No	No	No	Yes	No
Bank×Prefecture FE	No	No	No	Yes	No	No	No	Yes
Observations	52773	52773	52756	52769	52773	52773	52756	52769
R^2	0.047	0.144	0.245	0.213	0.046	0.141	0.239	0.208

Table 5: Decentralization and penalties by type and recipient

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1. Panel A presents results on each type of penalties, while Panel B presents results on each recipient of penalties. Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Type of penalties	Fine dummy	Fine amount (log)	Warning dummy	Number of warnings	Disqualification dummy	Number of Number of	Prohibition dummy	Number of prohibitions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local bank×Post	0.084***	0.491***	0.049***	0.057***	0.007***	0.006***	0.004***	0.004***
	(0.008)	(0.047)	(0.005)	(0.007)	(0.002)	(0.001)	(0.002)	(0.001)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,769	52,769	52,769	52,769	52,769	52,769	52,769	52,769
R^2	0.212	0.222	0.148	0.149	0.102	0.104	0.110	0.110

Panel B: Recipient of penalties	Individual penalty dummy	Number of individual penalties	Bank penalty dummy	Number of bank penalties
	(1)	(2)	(3)	(4)
Local bank×Post	0.049***	0.060***	0.081***	0.069***
	(0.005)	(0.007)	(0.008)	(0.007)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	52,769	52,769	52,769	52,769
R^2	0.155	0.153	0.210	0.212

Table 6: Decentralization and penalties: Bank / branch covariates and differential trends

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1 further controlling for bank characteristics and differential trends. Columns 1 and 3 present results using the likelihood of a penalty as dependent variable, while columns 2 and 4 present results using the number of penalties (log) as dependent variable. Columns 1-2 in Panel A presents results controlling for several bank characteristics, while columns 3-4 present results controlling for the average pre-2015 values of various bank characteristics interacted with the post dummy. Columns 1-2 in Panel B present results controlling for several branch characteristics, while columns 3-4 present results controlling for the average pre-2015 values of various branch characteristics interacted with the post dummy. Observations are bank-prefecture-years from 2010 to 2020 for Panel A, while branch-years from 2012 to 2020 for Panel B. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Bank covariates and differential trends	Penalty dummy (1)	Number of penalties (2)	Penalty dummy (3)	Number of penalties (4)
Local bank×Post	0.023***	0.033***	0.037**	0.065***
	(0.009)	(0.009)	(0.017)	(0.019)
Bank size (log)	0.012	0.004		
	(0.009)	(0.009)		
CAR	-0.004***	-0.006***		
	(0.002)	(0.002)		
Zscore	0.004	0.006**		
	(0.002)	(0.003)		
NIM	0.009	0.003		
	(0.006)	(0.007)		
Loan-to-deposit	0.000	-0.000		
	(0.000)	(0.000)		
Loan ratio	-0.001	-0.000		
	(0.001)	(0.001)		
Pre Bank size (log)×Post			0.005	0.009**
			(0.004)	(0.005)
Pre CAR×Post			0.001	0.002
			(0.001)	(0.002)
Pre Zscore×Post			-0.032***	-0.033***
			(0.009)	(0.010)
Pre NIM×Post			-0.002	0.003
			(0.009)	(0.011)
Pre Loan-to-deposit×Post			-0.002*	-0.002
			(0.001)	(0.001)
Pre Loan ratio×Post			0.002*	0.001

Continued on next page...

Table 6: Decentralization and penalties: Bank / branch covariates and differential trends (Continued)

			(0.001)	(0.001)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	42,556	42,556	42,098	42,098
R^2	0.219	0.214	0.215	0.211
Panel B: Branch covariates and differential trends	Penalty dummy	Number of penalties	Penalty dummy	Number of penalties
	(1)	(2)	(3)	(4)
Local bank×Post	0.074***	0.106***	0.075***	0.107***
	(0.013)	(0.014)	(0.012)	(0.014)
Branch size (log)	-0.015***	-0.019***		
	(0.005)	(0.006)		
Branch ROA	-0.002	-0.003		
	(0.002)	(0.003)		
Pre Branch size (log)×Post			0.043***	0.049***
			(0.006)	(0.007)
Pre Branch ROA×Post			-0.000	0.004
			(0.004)	(0.004)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	17,238	17,238	18,721	18,721
R^2	0.257	0.251	0.254	0.255

Table 7: Fine severity and dispersion

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on fines. The dependent variable in column (1) and (2) are average fines (and log of fines) in a bank-prefecture-year with fines. The dependent variable in column (3) and (4) are the absolute value of fines (and log of fines) relative to the mean, normalized by the mean, in a bank-prefecture-year with fines. Observations are bank-prefecture-years with positive fines from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Average fine (1)	Average fine (log) (2)	Fine dispersion (3)	Fine dispersion (log) (4)
Local bank×Post	-465.083*	-0.272**	0.107**	0.068*
	(276.252)	(0.116)	(0.041)	(0.040)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	3,871	3,871	3,871	3,871
R^2	0.333	0.545	0.505	0.470

Table 8: Local bias

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1 Columns 1-2 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and further interacting Local bank×Post by the bank-level Local ownership. Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Penalty dummy	Number of penalties
	(1)	(2)
Local bank×Post	0.023***	0.034***
	(0.008)	(0.009)
Local ownership	-0.002*	-0.003**
	(0.001)	(0.002)
Year FE	Yes	Yes
Bank×Prefecture FE	Yes	Yes
Observations	37,763	37,763
R^2	0.227	0.226

Table 9: Access to local information

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1. Columns 1-4 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and further interacting Local bank×Post by the log distance (in km) between the prefecture of the branch and Beijing. Columns 5-6 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and further interacting Local bank×Post dummies taking the value of 1 for short, intermediate, and long distance, respectively. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Distance (log)				Distance dummies	
	Penalty dummy (1)	Number of penalties (2)	Penalty dummy (3)	Number of penalties (4)	Penalty dummy (5)	Number of penalties (6)
Local bank×Post	-0.118*	-0.189**			0.082***	0.100***
	(0.068)	(0.074)			(0.010)	(0.012)
Post×Distance	-0.019*	-0.022				
	(0.011)	(0.014)				
Local bank×Post×Distance	0.030***	0.043***	0.022***	0.032***		
	(0.010)	(0.011)	(0.005)	(0.007)		
Local bank×Post×Long distance					0.094***	0.122***
					(0.014)	(0.021)
Local bank×Post×Intermediate distance					0.094***	0.111***
					(0.010)	(0.013)
Local bank×Post×Short distance					0.077***	0.079***
					(0.011)	(0.014)
Bank FE	No	No	No	No	No	No
Year FE	Yes	Yes	No	No	Yes	Yes
Bank×Year FE	No	No	Yes	Yes	No	No
Prefecture×Year FE	No	No	Yes	Yes	No	No
Bank×Prefecture FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,769	52,769	43,064	43,064	52,769	52,769
R^2	0.213	0.209	0.379	0.370	0.213	0.209

Table 10: Local uncertainty

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1. Panel A present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and further interacting Local bank×Post by the proxies for local economic uncertainty, measured by standard deviation (SD) and coefficient of variation (CV) of prefecture-level GDP growth rates. Panel B present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and further interacting Local bank×Post by the proxies for local political uncertainty, measured by either vacancy of leadership exceeding 6 months (Vacancy>6m) or changes in leadership less than 3 years (Turnover<3y). Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Economic uncertainty	Penalty dummy		Number of penalties	
	(1)	(2)	(3)	(4)
Local bank×Post	0.069***	0.067***	0.083***	0.080***
	(0.012)	(0.012)	(0.014)	(0.014)
High uncertainty (SD of GDP growth)	0.025***		0.033***	
	(0.009)		(0.010)	
Local bank×High uncertainty (SD of GDP growth)	-0.028***		-0.039***	
	(0.010)		(0.013)	
Post×High uncertainty (SD of GDP growth)	-0.031**		-0.034**	
	(0.014)		(0.014)	
Local bank×Post×High uncertainty (SD of GDP growth)	0.039**		0.041**	
	(0.016)		(0.019)	
High uncertainty (CV of GDP growth)		0.019**		0.031***
		(0.009)		(0.010)
Local bank×High uncertainty (CV of GDP growth)		-0.021**		-0.028**
		(0.010)		(0.013)
Post×High uncertainty (CV of GDP growth)		-0.021*		-0.029**
		(0.012)		(0.013)
Local bank×Post×High uncertainty (CV of GDP growth)		0.042***		0.048***
		(0.014)		(0.018)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	51,999	51,999	51,999	51,999
R^2	0.213	0.213	0.209	0.209
Panel B: Political uncertainty	Penalty dummy		Number of penalties	
	(1)	(2)	(3)	(4)
Local bank×Post	0.086***	0.077***	0.101***	0.091***
	(0.008)	(0.010)	(0.010)	(0.012)
High uncertainty (Vacancy>6m)	0.019		0.017	
	(0.013)		(0.012)	
Local bank×High uncertainty (Vacancy>6m)	-0.012		-0.003	
	(0.014)		(0.016)	
Post×High uncertainty (Vacancy>6m)	-0.058**		-0.054**	
	(0.024)		(0.021)	
Local bank×Post×High uncertainty (Vacancy>6m)	0.088**		0.116*	
	(0.042)		(0.064)	
High uncertainty (Turnover<3y)		0.017**		0.020***
		(0.007)		(0.007)
Local bank×High uncertainty (Turnover<3y)		-0.017**		-0.017**
		(0.008)		(0.008)
Post×High uncertainty (Turnover<3y)		-0.028**		-0.028**
		(0.011)		(0.012)
Local bank×Post×High uncertainty (Turnover<3y)		0.033**		0.035**
		(0.014)		(0.016)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	52,553	52,553	52,553	52,553
R^2	0.213	0.213	0.209	0.208

Table 11: Supervisory capacity

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1 further controlling for local supervisory capacity. Columns 1-2 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable for the full sample. Columns 3-4 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable for the truncated sample before 2018 merger of the CBRC and CIBC. Observations are bank-prefecture-years from 2010 to 2020 for the full sample and 2010-2017 in the truncated sample. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Full sample		Before 2018 merger	
	Penalty dummy	Number of penalties	Penalty dummy	Number of penalties
	(1)	(2)	(3)	(4)
Local bank×Post	0.088***	0.104***	0.068***	0.075***
	(0.008)	(0.010)	(0.008)	(0.009)
Log of new hires	-0.004	-0.003	-0.002	0.001
	(0.006)	(0.006)	(0.006)	(0.005)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	52,769	52,769	41,856	41,856
R^2	0.213	0.208	0.240	0.233

Table 12: Decentralization and lending: Loan-level analysis

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on loan conditions based on the model similar to equation 1. Columns 1-2 present results using loan spreads as dependent variable, while columns 3-4 present results using loan amounts (log) as dependent variable. Observations are loan-branch-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by firm. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Loan spreads		Loan Amounts	
	(1)	(2)	(3)	(4)
Local bank×Post	31.477**	31.257**	-0.385***	-0.329**
	(13.315)	(12.448)	(0.139)	(0.139)
Firm size	4.595	6.563	-0.022	-0.051
	(6.432)	(5.934)	(0.074)	(0.059)
Firm leverage	0.278	0.102	0.002	0.002
	(0.304)	(0.366)	(0.002)	(0.002)
Firm tangibility	-0.388**	-0.523***	0.001	0.000
	(0.193)	(0.183)	(0.003)	(0.002)
Firm cash holdings	0.042	0.095	0.002	-0.001
	(0.344)	(0.406)	(0.003)	(0.003)
Firm ROA	0.040	-0.373	0.001	0.001
	(0.356)	(0.371)	(0.004)	(0.003)
Year FE	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Prefecture FE	No	Yes	No	Yes
Observations	7,229	7,203	13,012	12,986
R^2	0.602	0.657	0.396	0.458

Table 13: Decentralization and lending: Prefecture-level analysis

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on loan supply at the prefecture level based on a (OLS and IV) model similar to equation 1. Columns 1-4 present results using credit to GDP at the prefecture level as dependent variable. In columns 2-3, the IV is the predetermined share of local banks in 2010 (as in Gilje et al. (2016)). Observations are prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by firm. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

City-level	Credit/GDP			
	OLS		IV	
	(1)	(2)	(3)	(4)
Local bank share	-70.729***	-41.114**	-73.159***	-35.427*
	(17.093)	(18.554)	(17.874)	(20.028)
Local bank share×Post	-31.245**	-35.624***	-43.634***	-32.137**
	(15.558)	(12.240)	(16.437)	(14.063)
GDP growth	-0.427	-0.561	-0.387	-0.572
	(0.706)	(0.349)	(0.706)	(0.350)
Fiscal balance	-0.122	-0.792**	-0.100	-0.824**
	(0.303)	(0.381)	(0.302)	(0.397)
Year FE	Yes	Yes	Yes	Yes
Province FE	No	Yes	No	Yes
Observations	3,136	3,136	3,136	3,136
R^2	0.153	0.440	0.061	0.060
F statistics			201.85	195.47

Internet Appendix

Table A1: Variable definitions and sources

Variable name	Definition	Source
Branch-level variables		
Local bank	Dummy variable equals to one if a bank is not a state-owned bank or joint-stock bank.	CBRC/CBIRC
Penalty dummy	Dummy variable equals to one if a bank receives a penalty. A penalty is punitive measure on a bank enforced by a supervisor as a consequence of significant non-compliance with laws or regulations.	CBRC/CBIRC, authors' calculation
Number of penalties	Log of 1 plus number of penalties.	CBRC/CBIRC, authors' calculation
Fine dummy	Dummy variable equals to one if a bank receives a fine. A fine includes a monetary penalty imposed on a bank and confiscation of its illegal proceeds.	CBRC/CBIRC, authors' calculation
Fine amount	Log of 1 plus fine amount (in thousand RMB).	CBRC/CBIRC, authors' calculation
Average fine	Aggregate fine amount divided by the number of fine incidences.	CBRC/CBIRC, authors' calculation
Fine dispersion	Absolute value of (fine amount - sample mean of fine amount) / sample mean of fine amount. The sample mean is defined as average fine amount for local/national banks in a prefecture-year.	CBRC/CBIRC, authors' calculation
Fine dispersion (log)	Absolute value of (log of fine amount – log of sample mean of fine amount).	CBRC/CBIRC, authors' calculation
Warning dummy	Dummy variable equals to one if a bank receives a warning. A warning is a formal notification letter issued by a supervisor, alerting a bank of its non-compliance with laws or regulations.	CBRC/CBIRC, authors' calculation
Number of warnings	Log of one plus number of warnings.	CBRC/CBIRC, authors' calculation
Disqualification dummy	Dummy variable equals to one if a bank receives a penalty that its manager is disqualified and barred from holding positions of senior managers in the banking industry for a specified period or permanently.	CBRC/CBIRC, authors' calculation

Table A1: Variable definitions and sources (continued)

Variable name	Definition	Source
Number of disqualifications	Log of 1 plus number of disqualifications.	CBRC/CBIRC, authors' calculation
Prohibition dummy	Dummy variable equals to one if a bank receives a penalty that its staff is prohibited or banned from working in the banking industry for a specified period or indefinitely.	CBRC/CBIRC, authors' calculation
Number of prohibitions	Log of 1 plus number of prohibitions.	CBRC/CBIRC, authors' calculation
Distance	Log of distance (in kilometers) between the bank branch and Beijing.	Baidu Map, authors' calculation
Share of offices outside the prefecture	The proportion of the number of bank offices located outside the prefecture where the bank branch operates.	CBRC/CBIRC, authors' calculation
Branch size	Log of a bank's total assets (in million RMB).	NTSD
Branch Leverage ratio	Equity over total assets (%).	NTSD
Branch ROA	Net income over total assets (%).	NTSD
Bank-level variables		
Bank size	Log of a bank's total assets (in billion RMB).	CNRDS
CAR	Capital adequacy ratio (%).	CNRDS
Z-score	Sum of equity to asset ratio and ROA divided by standard deviation of ROA. We use 3-year rolling window when calculating the standard deviation of ROA.	CNRDS
NIM	Net interest margin (%).	CNRDS
Loan-to-deposit	Gross loans to total deposits (%).	CNRDS
Loan-to-asset	Gross loans to total assets (%).	CNRDS
Local ownership	The sum of equity ownership (%) by local (provincial-level or prefecture-level) governments in top three shareholders.	CNRDS
Loan-level variables		
Loan spread	Loan spread over the benchmark interest rate in basis points.	Authors' collection
Loan amount	Log of loan amount (in million RMB).	Authors' collection
Firm-level variables		
Firm size	Log of total assets (in million RMB) of a borrower.	CNRDS
Firm leverage	Total liabilities to total assets (%) of a borrower.	CNRDS
Firm tangibility	Total property, plant, and equipment to total assets (%) of a borrower.	CNRDS

Table A1: Variable definitions and sources (continued)

Variable name	Definition	Source
Firm cash holdings	Cash holdings to total assets (%) of a borrower.	CNRDS
Firm ROA	Return on assets (%) of a borrower.	CNRDS
Prefecture-level variables		
Share of local banks	Share of local banks in terms of number of bank offices in a prefecture.	CBRC/CBIRC, Authors' collection
Credit to GDP	Private credit to GDP (%) of a prefecture.	CNRDS
GDP growth	Growth rate of GDP (%) of a prefecture.	CNRDS
Fiscal balance	A municipal (prefecture) government's revenue minus its expenditure, divided by its GDP.	CNRDS
High uncertainty (SD of GDP growth)	Dummy variable equals to one if the standard deviation of GDP growth rates over the past three years is greater than sample median in a year.	CNRDS
High uncertainty (CV of GDP growth)	Dummy variable equals to one if the coefficient of variation of GDP growth rates over the past three years is greater than sample median in a year.	CNRDS
High uncertainty (Vacancy>6m)	Dummy variable equals to one if the local leadership if vacant for over six months.	Manually collected and calculated
High uncertainty (Turnover<3y)	Dummy variable equals to one if the local leadership if leadership turnover within three years.	Manually collected and calculated
Province-level variables		
Regional NPL	Province-level nonperforming loan ratios (%).	CSMAR
New hires	Province-level log of number of annual new hires.	CBRC/CBIRC

Table A2: Additional robustness checks

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1. Columns 1-2 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and excluding from the sample state-owned banks. Columns 3-4 present results using either the likelihood of a penalty or the number of penalties (log) as dependent variable and excluding from the sample penalties issued by the local offices of the CBRC in Beijing. Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Excluding state-owned banks		Excluding penalties in Beijing	
	Penalty dummy	Number of penalties	Penalty dummy	Number of penalties
	(1)	(2)	(3)	(4)
Local bank×Post	0.063***	0.086***	0.090***	0.106***
	(0.012)	(0.014)	(0.008)	(0.010)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes
Observations	31,914	31,914	52,439	52,439
R^2	0.223	0.217	0.211	0.206

Table A3: Poisson regressions

Note: The table presents Poisson estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1. Columns 1-8 present results using the number of each type or recipient of penalties as dependent variable and implementing Poisson pseudo maximum likelihood regressions with (multiple levels of) fixed effects as described by Correia et al. (2020). Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Number of penalties	Number of warnings	Number of disqualifications	Number of prohibitions	Number of individual penalties	Number of bank penalties
	(1)	(2)	(3)	(4)	(5)	(6)
Local bank×Post	1.173***	1.537***	1.268***	1.518	1.162***	1.032***
	(0.141)	(0.279)	(0.406)	(1.186)	(0.381)	(0.130)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,906	15,573	2,515	2,218	14,790	32,720
Pseudo R^2	0.321	0.350	0.136	0.299	0.392	0.242

Table A4: Decentralization and penalties: Dynamic effects

Note: The table presents difference-in-differences estimates of the effect of the 2015 decentralization reform on penalties based on the model in equation 1 with $Localbank_i$ interacted with $Year_{(2015-[+]1)}$, that is, each year before and after 2015 (excluding 2015). Column 1 presents results using the likelihood of a penalty as dependent variable, while column 2 presents results using the number of penalties (log) as dependent variable. Observations are bank-prefecture-years from 2010 to 2020. All variables are defined in Table A1 of the Appendix. Robust standard errors are clustered by prefecture. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Penalty dummy	Number of penalties
	(1)	(2)
Local bank×year 2010	-0.000	-0.005
	(0.015)	(0.013)
Local bank×year 2011	-0.000	-0.005
	(0.014)	(0.013)
Local bank×year 2012	0.019	0.011
	(0.014)	(0.013)
Local bank×year 2013	0.019	0.015
	(0.014)	(0.013)
Local bank×year 2014	0.012	0.011
	(0.015)	(0.013)
Local bank×year 2016	0.053***	0.050***
	(0.014)	(0.013)
Local bank×year 2017	0.120***	0.129***
	(0.016)	(0.018)
Local bank×year 2018	0.158***	0.167***
	(0.018)	(0.018)
Local bank×year 2019	0.114***	0.132***
	(0.015)	(0.017)
Local bank×year 2020	0.149***	0.180***
	(0.019)	(0.023)
Year FE	Yes	Yes
Bank×Prefecture FE	Yes	Yes
Observations	52,769	52,769
R^2	0.217	0.213