Subordinates in Charge: Does Delegation Improve Bank Supervision?

Wouter Dessein, Di Gong, Thomas Lambert and Wolf Wagner*

June 24, 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 authority over a subset of bank branches to a lower level. Affected branches become 57-80% more likely to face supervisory intervention, implying substantial efficiency gains arising from improved detection of banking misconduct and more accurate assessments of its severity. The evidence is inconsistent with alternative explanations, including increased supervisory stringency or greater risk-taking by banks. Our analysis highlights decentralization benefits within supervisory hierarchies specifically, and speaks to the optimal organizational design for fraud detection more broadly.

Keywords: Communication, decentralization, financial architecture, fraud, organizational design, supervision

JEL: D23, G21, G28

*We thank our discussants, Margherita Bottero, João Rafael Cunha, Doron Levit, Guangqian Pan, Tianyue Ruan, Alessandro Scopelliti, Yang Su, Wenyu Zhu, as well as Andreas Brøgger, Bing Han, Kewei Hou, Dirk Jenter, Perrin Lefebvre, 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, FTG Conference "Bridging Theory and Empirical Research in Finance", 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. We are grateful to Guangzi Li for sharing the data on bank lending and Yanxin Yu for research assistance. A prior version of the paper 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 with Columbia Business School and the CEPR (wd2179@gsb.columbia.edu). Di Gong is with the University of International Business and Economics (d.gong@uibe.edu.cn). Thomas Lambert is with Erasmus University (t.lambert@rsm.nl). Wolf Wagner is with Erasmus University and the CEPR (wagner@rsm.nl).

1 Introduction

A central question in the design of financial architecture is how supervisory powers should be allocated between local and central authorities. Much of the debate focuses on settings in which local supervisors are *independent* of the central authority and finds that local decision-making produces less stringent supervisory outcomes.[1] This is commonly attributed to local supervisors being captured by local interests or failing to internalize spillovers across regions.

This paper shows that within a supervisory *hierarchy*, local supervision may deliver both more stringent and more efficient outcomes.[2] The reason is that divergent interests of local and central supervisors make communication between them less informative. This results in a loss of local information when decisions are made at the central level and prevents desirable interventions from being undertaken. Delegating authority to (biased) local supervisors can therefore make supervision both more stringent and more efficient.

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 of branches belonging to banks classified as "local" but not of branches belonging to "national" banks.[3] Prior to 2015 all branches were supervised in a hub-and-spoke system in which information was predominantly collected by supervisors at the prefecture-level, while decisions involved the central supervisor, requiring communication between the two. The reform fully delegated authority over branches of local banks to local supervisors. Importantly, delegation did not alter the overall objective of supervision, as local supervisors remain subordinated and accountable to the central supervisor.

We use a large, partly hand-collected, dataset covering 5,429 prefecture-level branches over a

[1] See, for example, Agarwal et al. (2014), Haselmann et al. (2022) and Beck et al. (2023). Note that less stringent supervision does not necessarily imply worse supervision. Theory offers divergent predictions on whether centralized or decentralized settings are preferred (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 United States, Fed, FDIC and OCC have headquarters and regional offices. Similar distributed structures also exist 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 commercial banks.

ten-year window around the 2015 policy reform. We measure supervisory outcomes through formal declarations of banking misconduct ("supervisory interventions"). To do so, we assemble novel data at both the branch- and supervisory-office-level on these interventions.[4] The granularity of the data allows us to exploit variation across branches within the same bank as well as within a supervisory office, comparing decisions involving branches of different banks.

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

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

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

Our model has two important implications. First, moderate levels of misconduct are now also detected. Such misconduct warrants lower penalties, so the average fine issued in a supervisory intervention declines. Second, fines become more dispersed. This is because under centralized su-

[4]In the Chinese system, a supervisory intervention is a formal measure taken when a bank is suspected of violating rules or prudent practice. It identifies the specific violation, mandates corrective action, and imposes a penalty. More than 300 local supervisory offices gained greater decision-making authority as a result of the reform.

[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.

pervision, 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. Both predictions are confirmed in the data. Our model also identifies two factors that determine the extent to which decentralization improves outcomes: central access to local information and the predictability of that information. Using proxies for both, we find the reform's effect is stronger when (incremental) local information is more valuable.

Our findings on the frequency, severity, and dispersion of supervisory interventions, as well as on the value of local information, speak to an informational interpretation of the reform. They are difficult to reconcile with alternative explanations. First, there may have been a confounding shock around the time of the reform that increased risk specifically for branches of local banks. While such a shock could generate more interventions in local branches, it should also increase average fines, contrary to our findings. The risk explanation is also inconsistent with evidence, discussed below, that local banks became more risk-averse following tighter supervision. Second, the reform may have changed local supervisors' preferences for intervening in local banks. But if supervisors had simply become stricter towards local banks, one would again expect average fines to rise rather than fall. Third, supervisory capacity (Eisenbach et al., 2022) may have increased after the reform, potentially explaining why supervisory outcomes improve. However, our results are robust to controlling for changes in supervisory capacity, proxied by new hires.

We conclude by examining whether improved supervision reduces bank risk-taking. Using data on individual lending decisions, we find that branches of local banks become more conservative following reform: they require higher compensation for risk and reduce loan sizes, thereby taking on less credit risk. At the aggregate level, prefectures with a higher share of branches of local banks experience lower loan supply following the reform.

Our paper has an important message for policy. A large part of the discussion on the design of supervisory architecture, especially in the Eurozone, has centered on the centralization of supervision relative to independent supervisors (Ampudia et al., 2019), with evidence pointing to stricter outcomes under centralization. For example, Agarwal et al. (2014) show that federal regulators are systematically tougher than state regulators, while Haselmann et al. (2022) find that large significant banks face higher standards under the Single Supervisory Mechanism (SSM).[6] Our analysis shows that in hierarchical supervisory systems of the kind observed in many countries, outcomes may be stricter (and more efficient) when the local supervisor is granted authority. The key mechanism we highlight is that local bias distorts communication upward in the hierarchy, making the central supervisor less informed about local conditions[7] and, as a result, intervene less.[8]

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 the trade-off between a setting with a single central regulator and one with several independent local regulators (Dell'Ariccia and Marquez, 2006; Kara, 2016; Foarta, 2018; Calzolari et al., 2019; Bolton and Oehmke, 2019; Colliard, 2020; Lóránth et al., 2022). This literature has primarily focused on inefficiencies arising from externalities across jurisdictions. Repullo (2018) and Carletti et al. (2021) study the incentives of local and central supervisors to acquire information within a hierarchy. Our paper instead focuses on the role of coarse communication (Crawford and Sobel, 1982) in 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 interactions among different layers of management (Harris and Raviv, 2005; Grenadier et al., 2016; Chakraborty and Yilmaz, 2017; Levit, 2020).

Our model builds directly on Dessein (2002) who analyzes the benefits of delegation relative to communication and finds that there is a trade-off between information and incentives but that delegation is preferred over a wide range of parameter values.[9] Our work extends Dessein (2002) to a setting with both an extensive and an intensive margin (the decision to intervene and the severity

[6]Our model shows that stricter supervision should not be equated with greater efficiency. Decentralization may be optimal even if it produces fewer interventions, because the interventions that do occur better reflect underlying misconduct.

[7]Several papers highlight the importance of local information in the supervisory process. For example, Lim et al. (2023) and Gopalan et al. (2021) show that closures of local offices of central supervisors are associated with greater bank risk.

[8]Another reason why, in practice, decentralization to subordinated (relative to independent) supervisors is more beneficial is that within a hierarchy a bias of the local supervisor is more easily contained (as the latter is accountable to the central supervisor).

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

of intervention). We show that due to the presence of the extensive margin, the overall trade-off disappears – that is, interventions can become both more informed and less biased. The model also maps into observable outcomes, allowing us to infer the efficiency implications of an actual decentralization reform.

Beyond supervision, the insights from our model apply more broadly to the question of which organizational level should have the power to discipline when misconduct or fraud is suspected. Lower-level managers have incentives to withhold information about misconduct for fear of punishment (of their subordinates, or themselves). While this seemingly favours making discipline a central function, our model shows a cost to that: Lower-level managers may fail to report moderate misconduct, and withhold information about the severity of larger misconduct.[10] Our model also applies to the optimal allocation of authority over investment decisions in case lower-level managers are biased against undertaking projects, and against larger projects.[11]

Empirical studies have emphasized benefits to decentralization in several contexts. Li (2017) shows that experts are more biased but still make better decisions because they posess better information. Aghion et al. (2021) document that firms that delegated more power to local plant managers prior to Great Recession outperformed more centralized counterparts in sectors that were hardest hit by the subsequent crisis. Bandiera et al. (2021) analyze the consequences of granting autonomy of procurement officers in Pakistan, and find that autonomy reduces procurement prices without reducing quality. Kala (2024) finds that granting managers in state-owned enterprises more autonomy results in greater value added but reduces investment in activities that primarily generate social benefits. Limodio et al. (2026) find that ethnic conflict in Ethiopia increases the value of local information, leading banks to appoint locally matched managers but reduce their decision-making autonomy. Vannutelli (2026) shows that reducing local control over auditor appointments strengthens alignment between municipal fiscal behavior and national fiscal goals.

[10] The ability of organizational structure to influence risk in institutions has received considerable attention in the context of banks (Laeven and Levine, 2009; Fahlenbrach and Stulz, 2011; Beltratti and Stulz, 2012; Ellul and Yerramilli, 2013) and firms more broadly (Khanna et al., 2015; Burns and Kedia, 2006; Efendi et al., 2007).

[11] Such bias may arise because undertaking new projects requires effort and exposes the manager to the risk of failure.

2 Institutional Background

2.1 An overview of China's banking sector

The Chinese banking system has experienced substantial growth and is now the world's largest. At the end of our sample period (2020) it had $38.98 trillion assets, 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 at the end of 2020, accounting for 69.5% of all commercial bank assets. In addition to the national banks, there is a 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 similar to the one present in many other countries (see footnote 2). 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 3). 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.

2.3 The decentralization reform of 2015

In January 2015, the CBRC had its first major organizational reform. In the area of supervision, the primary objective 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,[12] with the ultimate authority being with the central supervisor. The most common case regarding supervisory intervention 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, 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.[13] The local offices now independently decide on interventions, without an obligation to communicate about individual interventions the central office. 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 accountable.

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

[13] See Articles 13 and 14 of the Interim Measures for On-site Inspections of the China Banking Regulatory Commission (2015) and the Measures of the China Banking Regulatory Commission on Administrative Penalties (2015).

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 subordinated to the former 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 initial decision). The reform then fully delegated decision-making authority to the local supervisor.

This section develops a model that allows analyzing the impact of the reform on supervisory decision-making.[14] The main tension in the model is that the local supervisor has better information about banks, but faces distorted incentives as she is biased against interventions. 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 $\theta + \delta$. The component θ is observed only by the agent

[14]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).

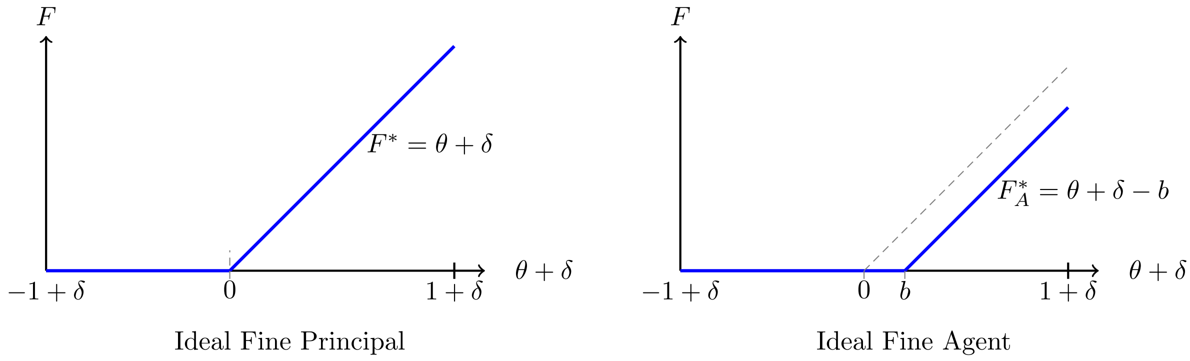


Figure 1: Ideal Fines

and is independently and uniformly distributed on $[-1, 1]$. The component δ is public information and is observed by both principal and agent. It take values

$$\delta \in \left\{-\bar{\delta}, \bar{\delta}\right\}$$

with equal probability, where $\bar{\delta} \in (0, \frac{1}{2})$ measures the importance of public information.[15] The bank type can be interpreted as the likelihood that the bank has violated supervisory guidelines, the severity of the violation, or the bank's overall risk-taking.

The central supervisor would like to impose a fine whenever $\theta + \delta > 0$. Her ideal fine is

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

The local supervisor is more lenient. He would like to impose a fine only when $\theta + \delta > b$, where $b > 0$. His ideal fine is

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

Preferences over fines are represented by absolute-loss utility. If the actual fine is $F \geq 0$, the principal's pay-off is $U_P(F, \theta, \delta) = -\left|F - F^*(\theta, \delta))\right|$ and the agent's payoff is $U_A(F, \theta, \delta) = -\left|F - F_A^*(\theta, \delta)\right|$. We refer to an outcome with a positive fine, $F > 0$, as a supervisory intervention.

[15] The restriction $\bar{\delta} < 1/2$ reduces the number of cases that must be discussed and simplifies the exposition.

3.2 Organization design

We consider two institutional settings.

First, under decentralized supervision, final authority is delegated to the local supervisor. The agent then chooses his ideal fine, so the the realized fine is

$$F = \max\{0, \theta + \delta - b\}.$$

Second, under centralized supervision, the central supervisor chooses the fine, but she can consult the local supervisor. Communication is modeled as cheap talk (Crawford and Sobel, 1982; Dessein, 2002). The agent sends a message m, and the principal chooses a fine

$$F(m) \in \arg\max_{F \geq 0} E\left[U_P(F, \theta, \delta) | G(\theta, m)\right]$$

where the conditional distribution of θ after message m is the determined in equilibrium.

3.3 Equilibrium under centralization

We now characterize the cheap-talk outcome under centralization. As in Crawford and Sobel (1982), all equilibria are characterized by a partition $\{a_0, a_1, a_2, ..., a_n\}$ of $[-1, 1]$, with $a_0 = -1$ and $a_n = 1$, and are outcome-equivalent to an equilibrium with n messages $m_k \in M = \{m_1, \ldots, m_n\}$ in which
(i) the agent sends message m_k whenever $\theta \in (a_{k-1}, a_k]$, so that $G(\theta|m_k) = U(a_{k-1}, a_k]$,
(ii) upon receiving the message m_k, the principal issues a fine

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

Note that for this to be an equilibrium, we must have that at $\theta = a_k$, the agent is indifferent between fines F_k and F_{k+1} whenever $1 < k < n$

As is standard, we focus on the most informative cheap talk equilibrium (see, e.g., Dessein, 2002), since this equilibrium is ex-ante preferred by both principal and agent. As is well known, there always exists an equilibrium in which $n = 1$ and ny communication is uninformative.

For a fixed value of δ, define

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

where $\lfloor x \rfloor$ denotes the largest integer below x/

Proposition 1. *(i) When $E(\theta+\delta) < 0$ (because $\delta = -\bar{\delta}$):*

- *If $b > \frac{1}{2} - \frac{\bar{\delta}}{2}$, there is no informative communication and $F = 0$.*
- *If $b < \frac{1}{2} - \frac{\bar{\delta}}{2}$, there is informative communication with $n^* > 1$ messages,*
  - *the principal issues no fine, $F_1 = 0$, upon receiving the message m_1,*
  - *the principal issues a fine $F_k > 0$ upon receiving message m_k with $k > 1$,*
  - *the agent sends m_1 if and only if $\theta < a_1$ where*

$$a_1 = (n^* \cdot b - \delta) \cdot \frac{2(n^*-1)}{2n^*-1} + \frac{1}{2n^*-1} > b - \delta \tag{1}$$

(ii) When $E(\theta+\delta) > 0$ (because $\delta = \bar{\delta}$):

- *If $b > 1/2$, there is no informative communication and $F = \bar{\delta} > 0$.*
- *If $b \in (\bar{b}, 1/2)$ with $\bar{b} \equiv \max\{\frac{1}{2} - \bar{\delta}, \frac{1}{3} - \frac{\bar{\delta}}{2}\}$, there is informative communication but all messages induce strictly positive fines ($F_k > 0$ for any k).*
- *If $b < \bar{b}$ there is informative communication with a no fine message: $F_1 = 0$ following message m_1 and $F_k > 0$ following message m_k with $k > 1$. As above, the equilibrium has n^* messages and a_1 is given by (1).*

Proof. See Appendix □

The above proposition shows that for moderate biases ($b < \frac{1}{2} - \frac{\bar{\delta}}{2}$ or $b < \bar{b}$), informative communication between principal and agent is possible and the principal does not issue 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 $b > \min\{\frac{1}{2} + \frac{\delta}{2}, \frac{1}{2}\}$, no informative communication is feasible and intervention depends on the sign of δ.

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

3.4 Optimality of Delegation and Likelihood of Supervisory Interventions

We now examine the circumstances under which delegation is preferred to centralization from the principal's perspective, and the corresponding implications for observable supervisory outcomes. Supervisory outcomes under delegation are given by $F = \max\{0, \theta + \delta - b\}$, while outcomes under centralization are $F = \max\{0, E(\theta + \delta | m_i)\}$.

Proposition 2. *The optimality of delegation and the relative frequency of interventions depend critically on the severity of the local supervisor's bias, b:*

(i) ***High Bias*** *($b > 1-(\sqrt{2}-1)\bar{\delta}$): Centralization is optimal, and interventions are more frequent under centralization than under delegation.*

(ii) ***Medium Bias*** *($\bar{b} < b < 1 - (\sqrt{2} - 1)\bar{\delta}$): Delegation is optimal, but interventions remain more frequent under centralization.*

(iii) ***Low Bias*** *($b < \bar{b}$): Delegation is optimal, and interventions are* ***strictly more frequent*** *under delegation. Under delegation, there is no intervention when $\theta < b - \delta$. Under centralization, there is no intervention when $\theta < a_1$, where $a_1 > b - \delta$.*

To unpack the intuition behind these results, consider first the case where the bias is sufficiently large so that the agent never intervenes. The principal, observing a positive public signal ($\delta = \bar{\delta} > 0$), optimally intervenes under centralization, making centralization both strictly preferred and more interventionist.

As the bias shrinks to an intermediate range (b close to $1/2$), the agent intervenes when $\theta + \delta > b$ and delegation becomes optimal due to its informational benefits: fines are responsive to the local state θ: $F = \theta + \delta - b$. In contrast, the principal indiscriminately intervenes whenever $\delta > 0$ setting $F = \bar{\delta}$, independently of the realization of θ as communication is un-informative (Proposition 1).

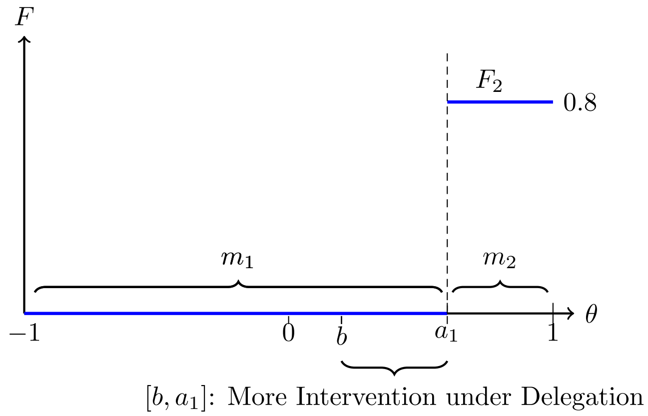


Figure 2: More Intervention under Delegation (example with $\bar{\delta} = 0; b = 0.2$). For types $\theta \in (b - \delta, a_1)$, the agent desires a small fine but sends message m_1 (yielding $F_1 = 0$) to prevent the principal from imposing a large fine ($F_2 >> 0$).

Centralization still yields a higher sheer volume of interventions, however, as the principal intervenes in at least half of all cases, compared to $(1 - b)/2$ of cases when the agent sets fines.

Note that 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 $b < 1 - (\sqrt{2} - 1)\bar{\delta}$. Intuitively, as public information becomes more relevant (a larger $\bar{\delta}$), delegation is less likely to be optimal.

The most nuanced — and empirically relevant— case emerges when the bias is small ($b < \bar{b}$). In this regime, delegation not only remains strictly optimal but also paradoxically results in *more* frequent interventions than centralization, despite the local supervisor being inherently more lenient than the principal.

This counterintuitive result hinges on strategic communication. Because the agent's bias is small, the principal defers to the agent when he sends the lowest-risk message (m_1), meaning the agent wields "real authority" over the extensive margin. However, the agent strategically expands the use of this zero-fine message to shield banks with moderate violations. Specifically, for intermediate realizations of the bank type where $\theta \in (b-\delta, a_1)$, the local supervisor observes a violation and *would* issue a small, tailored fine under delegation. But under centralization, revealing this violation would trigger the principal to impose a severe, uncalibrated fine. To prevent this disproportionate punishment, the agent conceals the information entirely, recommending a zero fine (m_1). Figure 2 illustrates this strategic omission. By delegating authority, the principal eliminates the agent's fear of excessive punishment, resulting in interventions for these moderate violations (or types).

Consequently, delegation strictly improves efficiency through both the intensive and extensive margins. On the intensive margin, there is a recognized trade-off between fines that are too small (due to the agent's bias) but perfectly tailored to θ, mirroring the classic insights of Dessein (2002). On the extensive margin, however, delegation provides an unambiguous efficiency gain: it strictly increases the number of warranted interventions. As we demonstrate in Section 3.5.1, this extensive margin effect dominates. Overall, delegation increases the total volume of appropriate fines issued.

Empirical Mapping

This theoretical framework provides a direct lens for interpreting our empirical results. Proposition 2 establishes that whenever delegation is suboptimal (cases i), the frequency of supervisory interventions decreases. This theoretical prediction is starkly inconsistent with our empirical evidence, which demonstrates an *increase* in interventions following the reform. We can therefore deduce that the supervisory system operates within the third parameter range: the local bias is sufficiently small, informative communication exists, and delegation is both optimal and intervention-maximizing.[16]

3.5 Two message equilibrium

In this section we derive several predictions of the model that can be empirically examined. We contrast them with the predictions of alternative channels, unrelated to supervisory communication, for why decentralization may affect interventions.

[16] In practice, the likelihood of an intervention also depends on risk-taking by banks. The evidence in Section 8 suggests that risk-taking actually declines following the reform. This implies that the estimated positive effect of the reform on interventions is biased downwards; the true increase in supervisory rigor is likely even stronger, further reinforcing that the system resides in this third parameter regime.

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 case of $n^* = 2$ (that is, $1 + \sqrt{\frac{1+\delta}{2b}} < 3$). 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 θ 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 (in Section 5.4 we quantify this effect based on our empirical estimates).

We can confirm that this is indeed an equilibrium. 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 Average fines and their dispersion

Under delegation, the average fine in a supervisory intervention ($E(F|F > 0)$) is lower (the average 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 averages fines are lower under decentralization, the (normalized) deviation of this 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 ($F2$).

A noteworthy implication of our model is that the total bias in interventions (combining the extensive and intensive margin) decreases. This bias can be (inversely) represented by the total disbursed fines ($E(F)$). Total fines under decentralization are $E[F] = \frac{1-b}{2}\frac{1-b}{2}$ (the product of the likelihood of intervention $E[F > 0] = \frac{1-b}{2}$ and the average fine issued $E[F|\, F > 0] = \frac{1-b}{2}$) and total fines under centralization are $E[F] = \frac{1-a_1}{2}\frac{1+a_1}{2}$ (the likelihood of intervention and the average fine are $E[F > 0] = \frac{1-a_1}{2}$ and $F_2 = \frac{1+a_1}{2}$). The latter expression strictly less than $\frac{1-b}{2}\frac{1-b}{2}$ unless we have $b = 0.2$ (in which case there is equality).

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 decisions 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, let $0 < \bar{\delta} < b$ instead of $\bar{\delta} = 0$ and assume that with certain probability $\theta = 0$ (instead of being distributed on $[-1, 1]$). In such a situation, the decentralized supervisor would never intervene, whereas the central supervisor would intervene with 50% probability (when $\delta = \bar{\delta}$). Thus, removing uncertainty reverses the effect of decentralization on interventions. Given that the overall effect of decentralization is positive in our data, higher uncertainty 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, 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 her). 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 $\frac{1+r-b}{2}$ 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 the previous subsection (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.

3.5.6 Competing channel III: Increased supervisory capacity

Suppose that following the reform local supervisors are endowed with additional resources (such as more manpower). This may again affect interventions irrespective of supervisory communication.

We model increased resources as resulting in higher precision of a signal about the true state of the bank θ. Assume (as in Section 3.5.4) that informational frictions are absent (that is, the central supervisor also observes the signal received by the local supervisor). Suppose that with certain probability this signal takes the true value θ and with remaining probability the signal is only noise. In the first case, the central supervisor observes the true state of the bank. As discussed in Section 3.5.4, this results in an intervention likelihood of $\frac{1}{2}$ and an average fine of $\frac{1}{2}$. In the second case, she is essentially indifferent to interventions. This can be interpreted as intervening with equal likelihood (probability $\frac{1}{2}$) and imposing a fine $F = 0$ upon intervention.[17]

It follows that higher supervisory capacity (interpreted as a high likelihood of observing the true state of the world) results in higher fines upon intervention. The intuition is when the supervisor has low information (low capacity), she is uninformed about the true risk of the bank. Since the priors are that average misconduct is low in each year, this results in low fines when the supervisor decides to intervene.

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, the affected institution or individual, the responsible supervisory authority, the reasons underlying the intervention, the specific laws or regulations

[17] In the general model of Section 3, this intervention behavior becomes the unique outcome when letting $\bar{\delta} \to 0$.

violated, and the penalty (or penalties) imposed.

There are five types of penalties. The by far most common one is a fine, which is a monetary penalty imposed on a bank (like the civil money penalty in the US context). Besides fines there are warnings (a formal letter issued by a supervisor stating non-compliance with laws or regulations), a disqualification barring bank managers from holding senior positions in the banking sector, a prohibition barring bank staff from working in the banking sector and a revocation of the license to operate the branch. Figure 4 provides an example of a penalty as disclosed on the CBRC/CBIRC website.[18]

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, showing that fines and warnings are the most issued type of penalties (63% and 32%). 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.

Penalties are aggregated at the prefecture branch level (containing both penalties to the branch itself and to lower-level offices) as the prefecture is the perimeter of the local supervisors. 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[19]) 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.

Table 2 presents information about local and national banks at the branch, bank and at aggregate level. One takeaway from this table is that local and national banks are fairly comparable at the branch-level: Branches of local and national banks are of similar size and have similar prof-

[18] 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 4).

[19] We exclude all other banks such as rural credit cooperatives, village banks, foreign banks.

itability, and operate in prefectures of similar market concentration, credit availability and distance to Beijing. Branches of national banks have a modestly higher market share (6.56% compared to 5.03%).

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:

$$Penalty_{ijt} = \alpha_i + \alpha_j + \alpha_t + \beta Localbank_i \times Post_t + \varepsilon_{ijt}, \tag{2}$$

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 interventions. 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 × 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.[20]

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 2. 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, and year fixed effects, implying that only the interaction between $Localbank_i \times Post_t$ is estimated. In column 3, we include prefecture × 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 × 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%.[21] Note

[20] 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.

[21] Table A2 of the Appendix shows that the result are robust to excluding state-owned banks or penalties issued by the local offices of the CBRC in Beijing (in the latter case central and local office are in the same prefecture, possibly

that the estimated effect on interventions is net of a potential, concurrent, change in risk-taking by banks as a result of the reform. As shown in Section 8, bank risk-taking declines following the reform, which by itself should lower interventions. In other words, our estimate likely understates the direct effect of the reform.

The last four columns focus on the number of interventions as outcome, confirming the results from the previous columns. In these regressions we use the log of 1 plus the number of penalties to retain observations with zero-valued outcomes. Table A3 in the Appendix reports alternatively the coefficient of a Poisson regression model.

Table 5 shows next that the results hold (qualitatively) for each type and each recipient of penalties employing a specification similar to column 4 of Table 4.[22]

In Table 6 we control for several bank characteristics and differential trends based on bank characteristics. In the first two columns of 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 lose 734 banks—mostly small banks and banks that are not obligated to publicly disclose their financial statements because they are not listed). The next two columns in addition control for a differential impact of the reform on banks with different characteristics.[23] For this we interact the various bank characteristics (averaged over the pre-reform period) with $Post_t$.

In Panel B we 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). We match bank (branch) name and city with our penalty sample and end up with 3,037 matched branches. We construct a set of branch covariates, including branch size (log of branch total asset in RMB) and ROA (net income over total asset).[24]

blurring the analysis of decentralization.

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

[23]For example, in 2015 China downgraded loan-to-deposit ratio as a supervisory tool which may have affected banks with different deposit ratios differentially.

[24]The data quality of the survey is very variable. For instance, reported equity is negative for a large fraction of

The results in Panels A and B confirm the ones 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).

5.2 Parallel trends

We now examine the dynamics of the effect. Figure 5 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. We see a delayed response in interventions to the reform, consistent with a lag between initial suspicion of misconduct and the conclusion of a full supervisory investigation, as well as a general adjustment period to new supervisory procedures. In year 2 following the reform, we observe a permanently higher incidence of penalties. 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

firms, and we have thus not included controls based on equity. For more detailed description of the NTSD, see Liu and Mao (2019).

in Figure 6. Both panels of Figure 6 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 efficiency 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 supervisors 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.

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 from inefficient interventions under centralized supervision,[25] arising in particular because now also more intermediate forms of misconduct are caught.[26] 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 fines and their dispersion

The model of communication and supervisory interventions predicts that the severity of interventions declines following decentralization (even though interventions become more likely, see Section 3.5.1). 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 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.

[25] 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%.

[26] 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.

A second prediction of the model is that the severity of punishment becomes more variable following decentralization (Section 3.5.1). 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.

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 3.5.2). In this case, decentralization entails a greater informational gain.

We examine this prediction in Table 8. 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.[27] We then examine the differential 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 $\times$ prefecture fixed effects.

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

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 $\times$ year and prefecture $\times$ year fixed effects. The estimated coefficients on the triple interaction term are again positive and statistically significant at the 1% level. 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.[28]

In columns 5 and 6 of the table, 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 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 for decision-making. 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 3.5.3. 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

[28]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 kilometers.

(year $t-3$ to year $t-1$).[29] 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 9. 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 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 $< 3y$).[30] We report the results of local political uncertainty in Panel B. Consistent with predictions, the persistently positive and significant coefficients on the local bank × post-reform × high uncertainty interaction terms indicate that decentralization reforms amplify supervisory stringency toward local banks in high-uncertainty political environments.

[29] Our findings are insensitive to alternative event window of past five years (available upon request).
[30] Alternative four-year cutoff yields consistent results.

7 Other channels

Our findings on the frequency, severity, and dispersion of supervisory interventions, as well as on the value of local 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 the evidence.

Bank risk-taking. A higher frequency of supervisory interventions may reflect higher risk-taking by banks. This could explain our main result if for some reason local banks have increased their risk-taking around the time of the reform. For example, in the year of the reform the loan-to-deposit ratio was de-emphasized as a supervisory tool, which may affect local banks differently if they undertake more lending. Our evidence is inconsistent with this explanation, for several reasons. First, as explained in Section 3.5.4, higher risk should result in higher fines imposed in interventions, the opposite of what we find. In addition, in Table 6 we have controlled for risk at the bank and branch level and for the reform to have differential effects based on bank characteristics (such as the loan-to-deposit ratio). Increased risk-taking is also inconsistent with the results in Section 8, where we find that banks have become more conservative in their lending.

Supervisory stringency. A higher frequency of supervisory interventions may also result from increased supervisory stringency. This may explain our results if following the reforms local supervisors have specifically become stricter towards local banks (in our notation, a reduction in b for intervening in local banks). For example, the reform may have made local supervisors more accountable when dealing with local banks, effectively reducing their bias. However, as explained in Section 3.5.5, in this case we would expect interventions fines to increase, not to decrease.

Supervisory capacity. The ability to collect information about misconduct depends on supervisory capacity (Eisenbach et al., 2022). Following the reform, local supervisors may have hired additional staff, providing them with more information about banks, and possibly affecting interventions.

To control for changes in supervisory capacity, we manually collect data from the CBRC/CBIRC's website about new hires at the provincial level. Column 1 and 2 in Table 10 display the baseline specifications including the (log) number of new hires. In addition, in column 3 and 4 we restrict the sample to the period before 2018. The reason is that following the merger between the CBRC and the CIRC (see Sub-section 2.2 for institutional details), it is no longer possible to identify hires specifically for bank supervision. Our main result remains robust across all specifications.

Higher supervisory capacity is also inconsistent with our findings of lower fines, as explained in Section 3.5.6. The reason is that in absence of precise information about misconduct, supervisors are uncertain about whether a fine is warranted, and will hence only impose small fines when they intervene.

Besides the ability to collect information, supervisory capacity may also affect the ability to make (many) decisions. Any increase in supervisory capacity may hence lead to more interventions. However, this type of capacity constraint is inconsistent with the aggregate pattern of interventions and fines.

8 Decentralization and lending

Better detection of misconduct should lower banks' incentives to take on risks. We test this prediction by examining individual lending decisions of branches (examining individual loans has the advantage that we can control for many factors, such as borrower fixed effects).

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 (size, leverage, tangibility, cash holdings, and ROA). We then merge the 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. Summary statistics on firm characteristics and loan terms are presented in

Panels C and D of Table 3 and the variable definitions are provided in Table A1 of the Appendix.

Table 11 reports the results. We first examine loan spreads. In column 1, the specification contains lagged borrower characteristics together with year, bank, and borrower fixed effects. 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 5% level) indicate that local banks charge higher loan spreads following the 2015 reform. The coefficients of interest are 31.3-31.5 basis points (about 29% of the standard deviation of loan spreads). This result is consistent with local banks show less aggressive lending behaviour post-reform by requiring higher compensation for identical borrowers. It should 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.

Columns 3 and 4 focus on loan quantities. The estimated coefficients on the interaction term (statistically significant at the 1% and 5% levels, respectively) show that local banks significantly reduce the amount they lend after the 2015 reform. Economically, the reform leads to a 28-32% reduction in loan sizes, which is also consistent with increased conservatism.[31]

The findings reported in Table 11 suggest that branches of local banks became more conservative in their lending because of decentralization. We next investigate whether this has real consequences at the aggregate, prefecture level. Specifically, we examine whether loan supply in prefectures with a higher presence of local banks is lower relative to other prefectures. It is not clear ex ante that such effects should be present at 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. We proxy loan supply using the ratio of credit over GDP. The interaction term of interest is between the share of local banks in the

[31] 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 11 restricting ourselves to borrowers already being listed prior to the reform and obtain similar results.

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 12 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 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 policy discussion, especially in the Eurozone, often centres around centralization of supervision relative to independent supervisors. Our analysis by contrast speas to benefit of decentralization relative to subordinated supervisors. In a hierarchical structure as present in many countries, outcomes may be stricter (and more efficient) when the local supervisor is granted authority. The reason is that local bias distorts communication upward in the hierarchy, making the central supervisor less informed about local conditions and, as a result, intervene less.

References

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi, 2014, Inconsistent regulators: Evidence from banking, *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, *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.

Beltratti, Andrea, and René M Stulz, 2012, The credit crisis around the globe: Why did some banks perform better?, *Journal of Financial Economics* 105, 1–17.

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.

Burns, Natasha, and Simi Kedia, 2006, The impact of performance-based compensation on misreporting, *Journal of Financial Economics* 79, 35–67.

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.

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

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.

Efendi, Jap, Anup Srivastava, and Edward Swanson, 2007, Why do corporate managers misstate financial statements? the role of in-the-money options and other incentives, *Journal of Financial Economics* 85, 667–708.

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.

Ellul, Andrew, and Vijay Yerramilli, 2013, Strong risk controls, lower risk: Evidence from US bank holding companies, *Journal of Finance* 68, 1757–1803.

Fahlenbrach, Rüdiger, and René M Stulz, 2011, Bank ceo incentives and the credit crisis, *Journal of Financial Economics* 99, 11–26.

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.

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.

Khanna, Vikramaditya S, E Han Kim, and Yao Lu, 2015, Ceo connectedness and corporate fraud, *Journal of Finance* 70, 1203–1252.

Laeven, Luc, and Ross Levine, 2009, Bank governance, regulation and risk-taking, *Journal of Financial Economics* 93, 259–275.

Levit, Doron, 2020, Words speak louder without actions, *The Journal of Finance* 75, 91–131.

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, 2023, The effects of regulatory office closures on bank behavior, *Journal of Money, Credit and Banking* 55, 1927–1963.

Limodio, Nicola, Luca Picariello, and Tom Schwantje, 2026, Banking under conflict: Managers and organizational design, *Working Paper* .

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 Jing 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.

Vannutelli, Silvia, 2026, From lapdogs to watchdogs: Random auditor assignment and municipal fiscal performance, *Working Paper* .

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.

A Appendix: Proofs

A.1 Proof of Proposition 1

Proof. Fix δ. We ignore knife-edge equalities, which can be handled by the usual tie-breaking conventions and do not affect the characterization. Any informative cheap-talk equilibrium can be represented by an ordered partition $-1 = a_0 < a_1 < \cdots < a_n = 1$ and messages $M = \{m_1, \ldots, m_n\}$ such that the agent sends m_k whenever $\theta \in (a_{k-1}, a_k]$.

The principal chooses F_k to minimize the expected absolute distance from her ideal fine, conditional on receiving message m_k. Because the loss function is absolute value, the optimal fine is a median of the posterior distribution of $F_P^*(\theta)$. Since θ is uniformly distributed on $(a_{k-1}, a_k]$ after message m_k, and since $F_P^*(\theta) = \max\{0, \theta + \delta\}$ is weakly increasing in θ, this implies

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

At every interior cutoff a_k, the agent must be indifferent between sending m_k and m_{k+1}. This gives

$$2(a_k + \delta) = F_k + F_{k+1} + 2b. \tag{3}$$

If two adjacent messages induce the same action, then they are equivalent to a single message. We therefore impose without loss of generality that $F_k > 0$ for all $k > 1$.

Equilibria with $F_1 = 0$. First consider equilibria in which the lowest message induces no fine. By the action formula above, this requires

$$\frac{a_1 - 1}{2} + \delta < 0,$$

or equivalently $a_1+\delta < 1-\delta$. For $k > 1$, the principal's action is interior, so $F_k = (a_{k-1}+a_k)/2+\delta$. Using (3), the first cutoff satisfies

$$a_1 + \delta = \frac{1}{2}(a_2 - a_1) + 2b,$$

and the remaining cutoffs satisfy

$$a_k - a_{k-1} = a_{k+1} - a_k + 4b, \qquad k = 2, \ldots, n-1.$$

Let $x \equiv a_n - a_{n-1}$ be the length of the highest interval. Then, for $k = 2, \ldots, n$,

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

The first cutoff is therefore pinned down by

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

Using $a_n = 1$, we obtain

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

Thus

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

A partition with n messages exists if and only if $x > 0$, or

$$2b(n-1)^2 < 1 + \delta.$$

Hence the most informative equilibrium has

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

and informative communication is feasible if and only if $n^* > 1$, which is equivalent to

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

Substituting the value of x into the expression for $a_1 + \delta$ gives

$$a_1 + \delta = b(n^* - 1)\left(\frac{2n^*}{2n^* - 1}\right) + \frac{1+\delta}{2n^* - 1}.$$

The remaining interval lengths are then determined recursively by

$$a_2 - a_1 = 2(a_1 + \delta) - 4b,$$

$$a_{k+1} - a_k = a_k - a_{k-1} - 4b, \qquad k > 1.$$

Finally, recall that the condition that $F_1 = 0$ equals $a_1 + \delta < 1 - \delta$. Substituting a_1, we have

$$b < 1/n^* - \delta/(n^* - 1).$$

One can verify that for $\delta \in (0, 1/3)$, the relevant condition is the one where, at the cut-off for b, $n^* = 2$. For $\delta \in (1/3, 1/2)$, the relevant condition has $n^* = 3$ at the cut-off. Thus, the condition for $F_1 = 0$ to be optimal following m_1 is given by

$$b < \bar{b} \equiv \max\{1/2 - \delta, 1/3 - \delta/2\}$$

We conclude that an informative equilibrium with a no-intervention message exists whenever

$$b < \min\left\{\frac{1}{2} + \frac{\delta}{2}, \bar{b}\right\}.$$

Equilibria with $F_1 > 0$. Now consider equilibria in which the principal imposes a strictly positive fine after every message. Then all actions are interior, and (3) implies

$$a_k - a_{k-1} = a_{k+1} - a_k + 4b, \qquad k = 1, \ldots, n-1.$$

Letting $x = a_n - a_{n-1}$, we have

$$a_k - a_{k-1} = x + 4(n-k)b, \qquad k = 1, \ldots, n.$$

Since the intervals sum to two,

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

Thus a positive-action partition with n messages exists if and only if

$$2bn(n-1) < 2.$$

The largest feasible number of messages $n^{\#}$ is therefore

$$n^{\#} = \left\lfloor \frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4}{b}} \right\rfloor,$$

and informative communication of this form is feasible if and only if $b < 1/2$.

For the first action to be strictly positive, we need

$$\frac{a_1 - 1}{2} + \delta > 0,$$

Since in equilibrium, $a_1 + 1 = x + 4(n-1)b = 2/n + 2(n-1)b$, this is equivalent to

$$b > \frac{1}{n^{\#}} - \frac{\delta}{n^{\#} - 1}$$

where $n^{\#}$ is a function of b. One can show that, for $\delta < 1/3$, the relevant inequality has $n^{\#} = 2$ and is thus $b > 1/2-\delta$. When $\delta \in (1/3, 1/2)$, the relevant condition has $n^{\#} = 3$ and is thus $b > 1/3-\bar{\delta}/2$. Note that this is the same threshold $\bar{b}$ we obtained as a condition for $F_1 = 0$. Therefore, when $\delta > 0$ and $b \in (\bar{b}, 1/2)$, informative communication may be feasible, but all messages induce strictly positive fines.

Combining the two cases proves the proposition. If $\delta < 0$, the binding condition for informative communication is $b < (1/2) + (\delta/2)$; when this fails, communication is uninformative and the principal chooses no fine because the unconditional median of $\theta + \delta$ is negative. If $\delta > 0$, the principal always issues a positive fine once $b > \bar{b}$. Communication may still be informative for $b < 1/2$, but it cannot affect the extensive margin; for $b > 1/2$, no informative communication is feasible and the principal still issues a fine because the unconditional median of $\theta+\delta$ is positive. $\square$

A.2 Proof of Proposition 2

Proof. 1) Frequency of Intervention: Part (i) and (ii): From Proposition 1, whenever $b > \bar{b}$, 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. Hence, interventions are more frequent under centralization. Part (iii): Under delegation, there is no intervention when $\theta < b - \delta$. From proposition 1, under centralization, there is no intervention when $\theta < a_1$, where $a_1 > b_1 - \delta$. Hence, interventions are more frequent under delegation.

2) Optimality of Intervention. We only prove that delegation is optimal for $b \in (\bar{b}, 1-(\sqrt{2}-1)\bar{\delta})$, and suboptimal for $b > 1-(\sqrt{2}-1)\bar{\delta}$. Optimality of delegation for $b < \bar{b}$ follows from the discussion following Propositin 2, and is omitted.

a) Whenever $b < 1 - \bar{\delta}$, expected pay-offs under delegation given δ equal

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

Since $\delta = -\bar{\delta}$ with 50% probability and $\delta = \bar{\delta}$ with 50% probability, we obtain

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

b) Whenever $b > 1 - \bar{\delta}$, expected pay-offs under delegation given $\delta = \bar{\delta}$ equal

$$EU_{D,\bar{\delta}} = -\frac{1}{2}b(1+\bar{\delta}) + \frac{1}{4}b^2;$$

expected pay-offs when $\delta = -\bar{\delta}$ equal

$$EU_{D,-\bar{\delta}} = -\frac{1}{2}\int_{\bar{\delta}}^{1}(\theta - \bar{\delta}) = -\frac{1}{4}(1-\bar{\delta})^2.$$

c) Consider now the 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

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

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, yielding pay-offs equal to

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

Since $\delta = -\bar{\delta}$ with 50% probability and $\delta = \bar{\delta}$ with 50% probability, expected pay-offs under centralization (and no communication) equal

$$EU_C = -1/4.$$

It follows that whenever $b < 1 - \bar{\delta}$, delegation is preferred over centralization (without communication) whenever $EU_D > EU_C$ or still:

$$2b - b^2 < 1$$

which is always satisfied for $b < 1$. Whenever $b \in (1-\bar{\delta}, 1)$, pay-offs under delegation and centralization (no communication) are identical when $\delta = -\bar{\delta}$. In contrast, when $\delta = \bar{\delta}$, $EU_{D,\bar{\delta}} > EU_{C,\bar{\delta}}$ if and only if

$$-\frac{1}{2}b(1+\bar{\delta}) + \frac{1}{4}b^2 > -\frac{1}{4}(1+2\bar{\delta}-\bar{\delta}^2)$$

or still

$$b < 1 - (\sqrt{2}-1)\bar{\delta}$$

.

□

Figure 3: Structure of China Banking Regulatory Commission

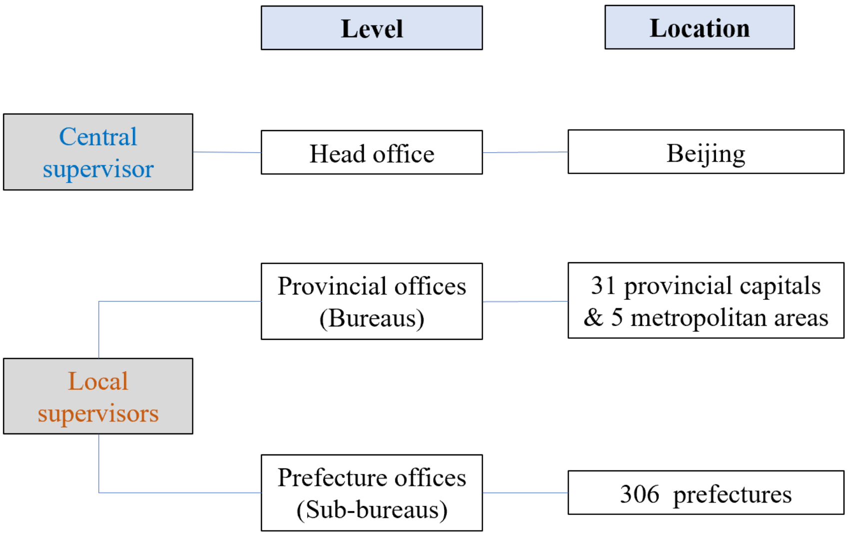


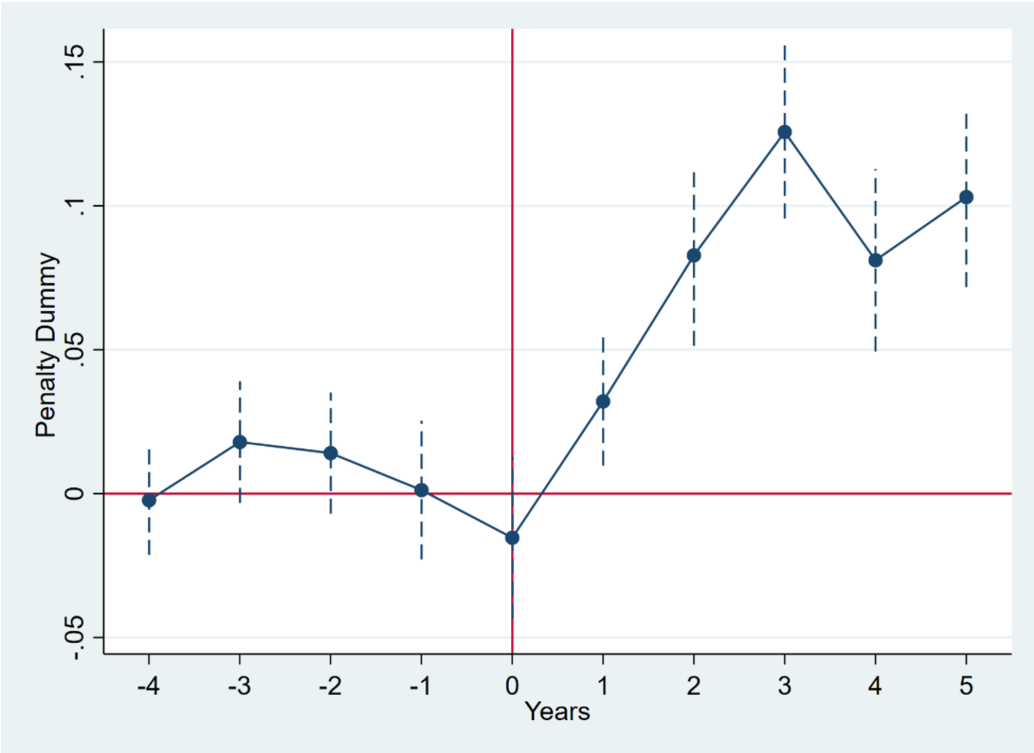
Figure 4: Example of a penalty
Note: The figure shows a snapshot of a randomly chosen penalty from the website of the CBIRC (source: www.cbirc.gov.cn/branch/hubei/view/pages/common/ItemDetail.html?docId =107940&itemId=1437&generaltype=0; last accessed: August 2023).

Information disclosure form for supervisory penalty, Yichang sub-bureau of the CBRC

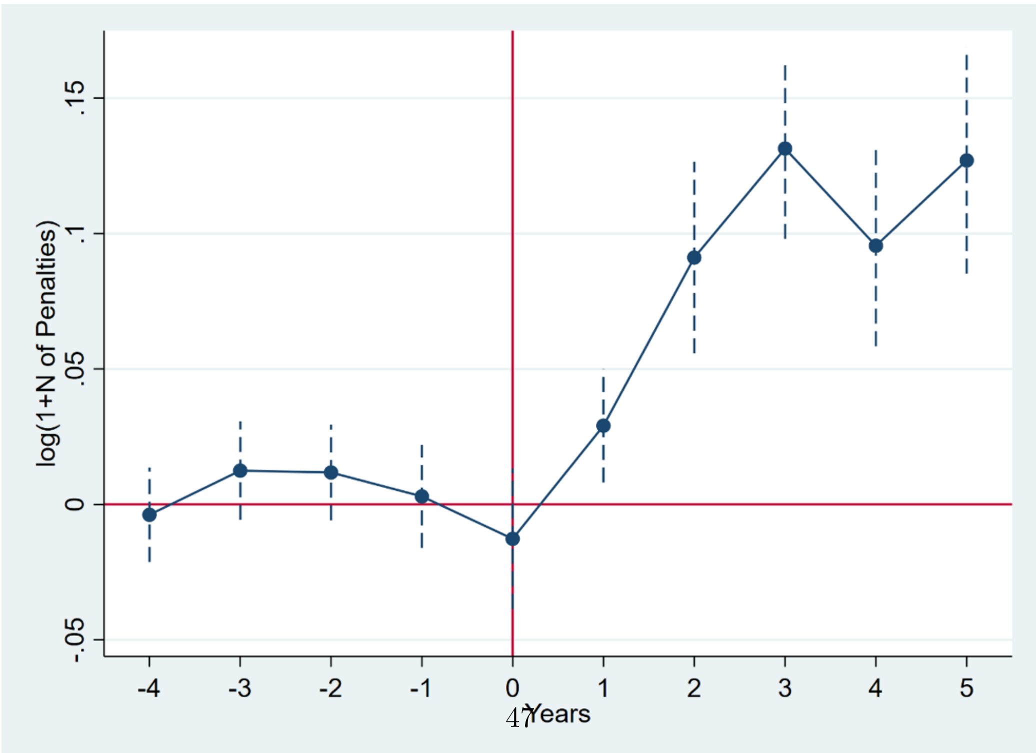
宜昌银监分局行政处罚信息公开表
（湖北银行股份有限公司宜昌分行）

Penalty ID	行政处罚决定书文号			宜银监罚决字[2015]3号	No. 3, Yichang authority 2015
Punished entity	被处罚当事人姓名或名称	个人姓名 Individual		——	
		单位	名称 Bank name	湖北银行股份有限公司宜昌分行	Yichang branch, Bank of Hubei
			法定代表人（主要负责人）姓名 Bank CEO	何青平	
Facts of misconduct	主要违法违规事实（案由）			存在贷款风险分类不准确、以贷转存吸收存款的违规行为。	Misconducts of inaccurately classifying loans based on their risk level and withholding a certain percentage of the loan as a deposit when the loan is granted.
Penalty basis	行政处罚依据			《中华人民共和国银行业监督管理法》第四十六条第（五）项	Banking Supervision and Administration Law of the PRC, Article 46(5).
Penalty decision	行政处罚决定			罚款人民币40万元	A fine of 400,000 RMB
Supervisory office	作出处罚决定的机关名称			中国银行业监督管理委员会宜昌监管分局	Yichang sub-bureau of the CBRC
Decision date	作出处罚决定的日期			2015年11月5日	November 5, 2015

Figure 5: 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.



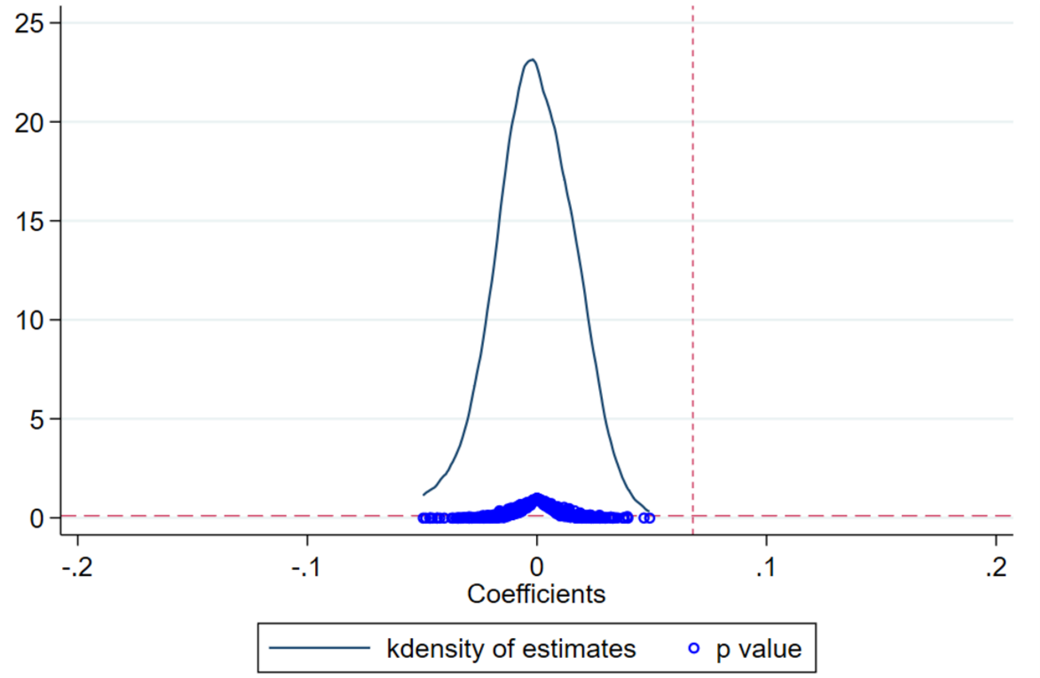
(a) Likelihood of a penalty



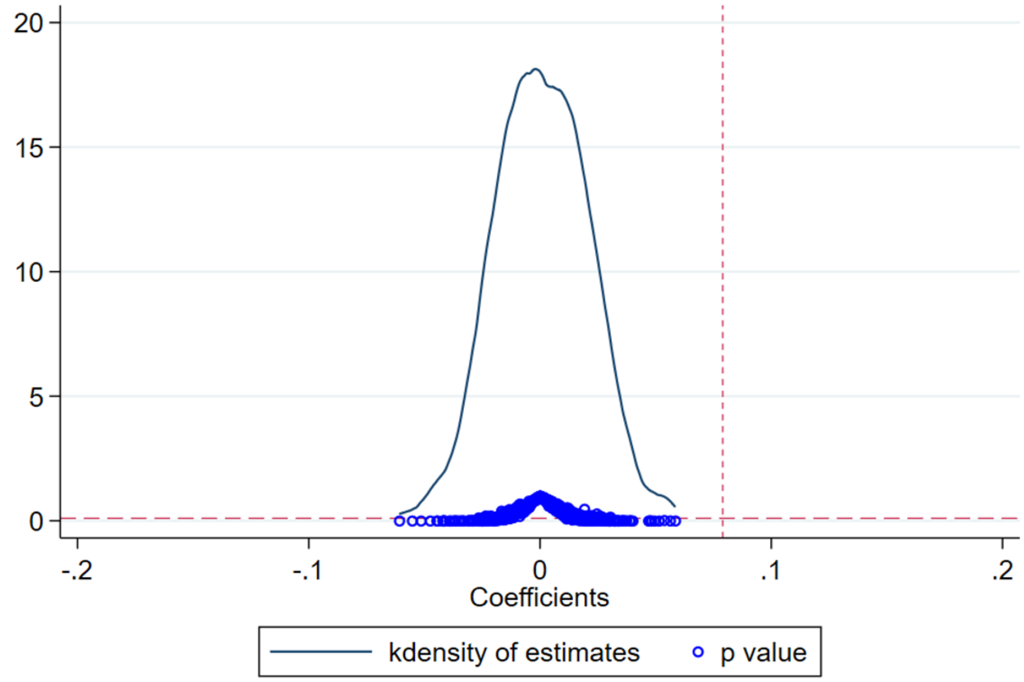
(b) Number of penalties

Figure 6: 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: Penalty type, reason and recipient

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
Fine	8,589	4,050	4,539
Warning	4,341	2,431	1,910
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	6,819	3,688	3,131
Deposit-related reasons	815	271	544
Interbank-related reasons	463	284	179
Acceptance-related reasons	1,534	666	868
Credit Card-related reasons	181	33	148
Guarantee-related reasons	358	154	204
Prudential regulation-related reasons	1,558	771	787
Internal control-related reasons	875	388	487
Governance-related reasons	294	273	21
Panel C: Recipient of penalties	**N**	**Local banks**	**National banks**
Individuals	4,670	2,737	1,933
Banks	7,858	3,513	4,345
Both individuals and banks	450	252	198

Table 2: Branches at different levels of aggregation
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.

Aggregation level	Local banks	National banks
Branch-level		
Size (log of total assets in million RMB)	11.76	11.79
ROA	0.97%	0.98%
Market share (% offices)	5.03%	6.56%
Dist. of branches' pref. capital to Beijing (km)	1101.98	1183.73
credit/GDP of branches' prefecture	108.04%	112.79%
HHI of branches' prefecture (% offices)	0.09	0.10
Number of penalties	2.80	1.86
Bank-level		
Number of branches per bank	2.03	182.56
Number of penalties per bank	5.68	337.78
Aggregate-level		
Number of banks	1056	18
Number of penalties	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	0.110	0.313	0	0	1
Number of penalties	52,773	0.229	1.050	0	0	4
Fine dummy	52,773	0.105	0.307	0	0	1
Number of fines	52,773	0.163	0.652	0	0	3
Fine amount	52,773	103.318	3,361.840	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	0.320	0	0.260	0.984
Fine dispersion (log)	3,871	0.410	0.441	0	0.266	1.478
Warning dummy	52,773	0.035	0.183	0	0	1
Number of warnings	52,773	0.082	0.670	0	0	2
Disqualification dummy	52,773	0.005	0.068	0	0	0
Number of disqualifications	52,773	0.007	0.117	0	0	0
Prohibition dummy	52,773	0.005	0.071	0	0	0
Number of prohibitions	52,773	0.007	0.118	0	0	0
Distance (km)	52,773	1,152.448	602.336	47.225	1,080.204	2,771.434
Distance (log)	52,773	6.845	0.393	3.876	6.986	7.927
Branch size (log)	18,819	11.788	1.302	9.202	11.756	14.236
Branch ROA (%)	18,823	0.979	1.433	-2.445	0.986	4.195
Panel B: Bank-level variables	**N**	**Mean**	**SD**	**P1**	**Median**	**P99**
Bank size (log)	44,722	7.636	1.970	2.014	8.675	9.134
CAR (%)	44,028	13.259	2.022	8.480	13.170	17.720
Z-score	42,671	5.413	0.909	3.338	5.390	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
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	0.500	30	1,000
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)	3,643	0.501	0.500	0	1	1
High uncertainty (CV of GDP growth)	3,643	0.501	0.500	0	1	1
High uncertainty (Vacancy>6m)	3,729	0.020	0.141	0	0	1
High uncertainty (Turnover<3y)	3,729	0.317	0.465	0	0	1
New hires (log)	387	3.169	0.824	1.099	3.401	4.419

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 2. 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	52,773	52,773	52,756	52,769	52,773	52,773	52,756	52,769
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 2. 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	Number of fines	Warning dummy	Number of warnings	Disqualification dummy	Number of disqualifications	Prohibition dummy	Number of prohibitions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local bank×Post	0.084***	0.079***	0.049***	0.057***	0.007***	0.006***	0.004***	0.004***
	(0.008)	(0.008)	(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	0.212	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 2 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	Number of penalties	Penalty dummy	Number of penalties
	(1)	(2)	(3)	(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
			(0.001)	(0.001)
Year FE	Yes	Yes	Yes	Yes
Bank×Prefecture FE	Yes	Yes	Yes	Yes

Continued on next page...

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

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	18,721	18,721	17,238	17,238
R^2	0.254	0.255	0.257	0.251

Table 7: Averages fines 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**	0.100*
	(276.252)	(0.116)	(0.041)	(0.055)
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	0.496

Table 8: 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 2. 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**				
	(0.068)	(0.074)				
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 9: 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 2. 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	0.085***	0.077***	0.100***	0.091***
	(0.008)	(0.010)	(0.010)	(0.012)
High uncertainty (Vacancy>6m)	0.029**		0.025**	
	(0.013)		(0.012)	
Local bank×High uncertainty (Vacancy>6m)	-0.026*		-0.020	
	(0.016)		(0.019)	
Post×High uncertainty (Vacancy>6m)	-0.068***		-0.062***	
	(0.024)		(0.021)	
Local bank×Post×High uncertainty (Vacancy>6m)	0.101**		0.133**	
	(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 10: 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 2 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 (1)	Number of penalties (2)	Penalty dummy (3)	Number of penalties (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 11: 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 2. 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 12: 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 2. Columns 1-4 present results using credit to GDP at the prefecture level as dependent variable. In columns 3-4, 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 prefecture. ***, **, 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
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
Firm cash holdings	Cash holdings to total assets (%) of a borrower.	CNRDS
Firm ROA	Return on assets (%) of a borrower.	CNRDS
Prefecture-level variables		

Table A1: Variable definitions and sources (continued)

Variable name	Definition	Source
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 2. 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 (1)	Number of penalties (2)	Penalty dummy (3)	Number of penalties (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 2. 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 2 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