Resource Allocation in Bank Supervision: Trade-offs and Outcomes

Thomas M. Eisenbach, David O. Lucca, and Robert M. Townsend*

April 2021

Abstract

We estimate a structural model of resource allocation on work hours of Federal Reserve bank supervisors to disentangle how supervisory technology, preferences and resource constraints impact bank outcomes. We find a significant effect of supervision on bank risk and large technological scale economies with respect to bank size. Consistent with macro-prudential objectives, revealed supervisory preferences disproportionately weight larger banks, especially post-2008 when a resource reallocation to larger banks increased risk on average across all banks. Shadow cost estimates show tight resources around the financial crisis and counterfactuals indicate that binding constraints have large effects on the distribution of bank outcomes.

Keywords: bank supervision, bank regulation, monitoring, time use
JEL Classification: D82, G21, G28

*Eisenbach and Lucca: Federal Reserve Bank of New York, thomas.eisenbach@ny.frb.org and david.lucca@ny.frb.org; Townsend: Massachusetts Institute of Technology, rtownsend@mit.edu. This paper was previously circulated under the title "The Economics of Bank Supervision". The views expressed in this paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. For valuable comments, we thank Stefan Nagel (Editor), an Associate Editor, the anonymous referees, Doug Diamond, Richard Crump, Mark Flannery, Xavier Freixas, Paul Goldsmith-Pinkham, Daniel Lewis, Gyongyi Loranth, and Rodney Ramcharan, as well as audience members at the New York Fed Conference on Supervising Large & Complex Financial Institutions, the Wharton Conference on Liquidity and Financial Crises, the BAFFI CAREFIN Annual International Banking Conference, the Jackson Hole Finance Conference, the Banco de España–CEMFI Conference on Financial Stability, the Western Finance Association Meeting, and the Chicago–Minnesota Theory Conference, as well as at Hong Kong University, Hong Kong University of Science and Technology, Singapore Management University, Nanyang Technological University, National University of Singapore, University of Mannheim, Bundesbank, Frankfurt School of Finance, and the European Central Bank. We thank Emily Eisner, Kyra Frye, Eilidh Geddes, Helene Hall, and Alice Liang for excellent research assistance. Any errors are our own.

1 Introduction

Previous literature on bank supervision and regulation mostly focused on distorted incentives of supervisors and lax regulation as contributing factors to past financial crises (Admati and Hellwig, 2013; Barth et al., 2012; Carpenter and Moss, 2013). In contrast, this paper studies the importance of the availability and allocation of supervisory resources for the level and distribution of risk in the banking system. Anecdotal evidence highlights several instances when supervisory resources became very scarce, often coinciding with times of major banking distress. For example, the FDIC "faced severe challenges, such as the volatility of workload [and] fluctuating staffing levels" in managing the banking crisis of the 1980s and early 1990s (FDIC, 1997). After the 2007–09 financial crisis, supervisory staff at the Federal Reserve increased by roughly 50% (Figure 1, left panel) and their allocation shifted, with the share devoted to smaller banks roughly cut in half (right panel). We provide new insights on the importance of supervisory resources using a structural model of bank supervision estimated on a unique dataset of work hours spent by Federal Reserve staff supervising the universe of U.S. bank holding companies (BHCs).[1] The model makes use of the observed resource allocation and bank outcomes to disentangle three main factors at play in bank supervision: the supervisory technology that maps supervisory efforts into reduced bank distress, supervisory preferences that weight banks of difference size and risk, and the overall scarcity of supervisory resources.

In the model, supervisors allocate a fixed total amount of resources to minimize banks' distress probabilities, weighted by supervisory preference loadings. Given banks' size and risk, the optimal allocation of supervisory resources depends on the interaction of two technological parameters (the impact of supervision and economies of scale), supervisory preferences, and the shadow cost of resources. With scarce overall resources, increasing supervisory attention to one bank requires less attention be paid to other banks, as suggested by the post-2008 reallocation of resources from small to large banks (Figure 1).

We find that binding resource constraints have quantitatively large effects on bank outcomes. Shadow cost estimates indicate that the 50% increase in resources post-2008 (Figure 1) was too gradual to keep up with the sudden increases in the size and risk of banks under Fed supervision. Had the 2014 resources been available already in 2007, bank distress would have been 10% lower during and in the immediate aftermath of the financial crisis. Estimates also indicate that the post-2008 reallocation of resources involved a quantitatively large and previously unreported trade-off. While the reallocation lowered risk at larger banks, it increased risk at smaller banks and, on net, across the universe of banks. In

[1] Our analysis is at the level of BHCs; we refer to them interchangeably as "banks."

general, we find that the shadow cost of supervisory resources is not equalized across Federal Reserve districts and that supervisory preferences tilt attention toward safer banks, both resulting in a higher average distress probability.

A key challenge in estimating the effect of supervisory hours on bank distress is that the allocation of hours naturally depends on bank risk. Without accounting for this endogeneity, a true negative effect of additional supervision on bank distress may appear weaker or the coefficient may flip sign, incorrectly suggesting that additional supervision increases the probability of distress. We show that, under a linear approximation of the first-order condition, our model can be expressed as a standard instrumental variable probit specification. This makes our structural approach transparent in terms of identification and interpretation.

We consider three different instruments that are conceptually consistent with the model framework and exploit variation across different sets of banks but yield consistent results. The first instrument is the shadow cost of supervisory resources, which negatively affects the supervisory attention paid to a bank, conditional on its size and level of risk. Each of the twelve Reserve Banks in the Federal Reserve System supervises BHCs located in its own district with its own supervisory staff. Consistent with resource constraints binding at the district level, we show that, after controlling for a bank's own characteristics, the bank receives less supervisory attention when more resources are needed at other banks in the same district.

The other two instruments are separate preference shifters that affect the amount of supervisory attention directed toward a given bank. The first shifter draws from Hirtle, Kovner, and Plosser (2020), who find that the largest banks in each district receive more supervisory attention than similarly sized banks in other districts where they are not the largest. This disproportionate attention indicates that supervisors are most concerned with the performance of the largest banks in their district. The second shifter is based on the lower minimum exam frequency for banks that are non-complex, have assets below $10 billion, and have a satisfactory supervisory rating. As these banks do not have to be examined every year but only once every two years, supervisory hours predictably cluster in examination years, which are predetermined by exams in prior years.

As outcome variables, we consider three separate measures of future bank distress, increasing in severity: low return on assets (below the 10th percentile, which is about zero), "severe stress" (failure or a failing rating), and outright failure. Based on our estimates, a 100% increase in supervisory hours in year t leads to a reduction in the probability of severe stress in year $t+1$ of 2.3 percentage points, which is close to half of the baseline probability of 5.3%. This average marginal effect masks considerable heterogeneity; for the riskiest banks (with the worst possible ratings of 4 or 5), the marginal effect is 8 percentage points. Effects for the other outcome variables are similarly large and we find consistent effects using each of the three instruments separately.

The structural model allows us to make two types of contributions. First, it enables us to decompose the empirical loadings of supervisory hours on bank size and risk into the effects of supervisory technology and preferences. Regarding technology, we estimate a unit elasticity of supervisory hours with respect to bank size but uncover significant economies of scale in supervision, with an estimated elasticity of the cost function with respect to bank size of about 0.6, i.e. supervising a 10% larger bank requires only 6% more resources. Regarding supervisors' preference weights, we find a competing size elasticity significantly greater than 1, meaning that supervisors place disproportionately more weight on larger banks. In other words, while, from a technological perspective, larger banks are easier to supervise, hours nevertheless increase one-for-one with bank size because supervisors care much more about distress at larger banks, consistent with systemic risk concerns (e.g. Hanson, Kashyap, and Stein, 2011). In contrast, even though hours allocated to riskier banks (rated 4 or 5) are more than three times greater than hours allocated to safe banks (rated 1), our estimates reveal that supervisors place *less* weight on riskier banks. Thus, fewer resources are allocated to risky banks than predicted by the technological trade-off in the model. This preference tilt can be interpreted as supervisors overweighting the very small distress probability of safe banks (e.g. Kahneman and Tversky, 1979) or underestimating the larger distress probability of risky banks (e.g. Gennaioli and Shleifer, 2018).

Second, the structural model enables us to study counterfactual allocations to quantify how resource scarcity and supervisory preferences affect the overall level and the distribution of risk across banks. We first study the role of scarce resources during and after the financial crisis of 2007–09, by positing that the Federal Reserve enters the financial crisis "fully staffed" at 2014 levels; we find that bank distress would have been about 10% lower over the whole period with considerably larger effects in 2007 and 2008. We then study the effect of the Federal Reserve's decentralized approach to bank supervision; we find that average bank distress is meaningfully higher both because the scarcity of resources differs across districts and because of the disproportionate attention to the largest banks within districts. Next, we show that the preference tilt toward safer banks is quantitatively meaningful, as it attenuates the supervisory response to risk by about 30%. Finally, we find that, following the shift in supervisory preferences toward larger banks after the financial crisis, risk at large banks decreased but less than it increased at small banks. The resulting net increase of risk across all banks shows supervisors trading off micro- and macro-prudential objectives (e.g. Borio, 2011).

While we report the results of these counterfactuals in terms of average bank outcomes, judging their efficiency ultimately depends on the appropriate social welfare function. We thus see our analysis as a positive, as opposed to a normative, one. We conclude with a simple back-of-the-envelope calculation that suggests net benefits to increasing overall supervisory resources when comparing the costs of additional supervisory staff against the benefits of reduced bank distress.

Related literature. The 2007–09 financial crisis has spurred renewed attention to bank regulation and supervision and their role in the buildup of risk leading up to the crisis. Duffie (2019) argues that the financial system was "prone to fail" because of a combination of weak regulation and supervision. Some argue that the banking sector was too levered and relied too much on unstable short-term funding (Greenwood, Stein, Hanson, and Sunderam, 2017; Aikman, Bridges, Kashyap, and Siegert, 2019). Others stress that regulators placed too much faith in market discipline, which was distorted by expectations that some institutions were "too big to fail" (Admati and Hellwig, 2013), or that they underestimated the probability of a severe shock (Gennaioli and Shleifer, 2018). The role of supervisory resources has not been previously discussed, with the exception of Duffie (2019), who notes significant differences in staffing at the Securities and Exchange Commission and the Federal Reserve. Lucca, Seru, and Trebbi (2014) show pro-cyclical net flows of staff from banking authorities to the private sector, consistent with scarce supervisory resources at the end of economic expansions.

The model of resource allocation underlying our analysis is in the neoclassical tradition of Becker (1965) and Radner and Rothschild (1975), whereby the allocation of time or effort maximizes an objective function subject to a resource constraint. In estimating the effect of supervision on bank risk, we do not explicitly specify the channel through which supervision operates or why supervision by a banking authority is necessary. Prior contributions show, for example, that limited liability raises moral hazard issues leading to excessive risk taking (Jensen and Meckling, 1976) and that there are limits to the ability of markets to provide the necessary discipline for banks (Flannery, 1998; Rochet, 2004). Theoretical channels through which supervision can counteract these issues include auditing bank asset values to detect breaches of capital requirements (Rochet, 2007); preventing banks from taking observable but non-verifiable actions (Dewatripont and Tirole, 1994); incentivizing banks through punitive interference after verifiable outcomes (Marshall and Prescott, 2001; Harris and Raviv, 2014); and taking corrective action to affect banks' risk-return trade-off before outcomes realize (Carletti, Dell'Ariccia, and Marquez, 2019). The analysis in this paper focuses on supervision as opposed to other pillars of banking policy, such as bank capital regulation, which have been the focus of an extensive literature (e.g. Repullo and Suarez, 2012).

In terms of the empirical literature, earlier contributions show that supervision produces valuable information (Hirtle and Lopez, 1999; Peek, Rosengren, and Tootell, 1999). Our paper is most closely related to Hirtle, Kovner, and Plosser (2020), hereafter HKP, who contribute to the growing literature that studies the effect of bank supervision on bank outcomes such as Bisetti (2018), Altavilla et al. (2019), and Granja and Leuz (2019). HKP use data on supervisory hours to show that the largest banks in each of the twelve Federal Reserve districts receive disproportionate attention. HKP then show that bank supervision lowers bank risk in a propensity score matching setting that compares "district top" banks to similar institutions in other districts that are, however, not ranked largest. We rely on HKP's "district top" instrument as one of the three instruments we use to estimate the parameters of our model. While we also use data on supervisory hours, we study the allocation of supervisory hours in a structural model. This structural approach allows us to estimate new parameters in supervision associated with both technology and preferences, via a revealed preference approach, and to conduct counterfactual experiments. Because supervisory policy depends on multiple parameters and a joint resource allocation across a whole set of banks, the structural approach is needed to study how the scarce resources and trade-offs in their allocation affect bank outcomes. A number of existing papers indirectly exploit policy distortions in resource allocation to study the effect of supervision on bank outcomes, as in HKP. Rezende and Wu (2014) exploit examination frequency requirements that aim to economize on limited resources. Kandrac and Schlusche (2019) exploit a natural experiment in which total resources declined because supervisory staff unexpectedly quit when the local supervisory office was relocated. Finally, Passalacqua et al. (2020) exploit randomized examinations by the Bank of Italy, which are also due to scarcity of resources. Our paper provides an encompassing framework for these studies.

To the best of our knowledge, our paper is the first to estimate supervisory preference weights and how they depend on bank characteristics. A new element of post-crisis analysis has been macro-prudential regulation that considers externalities (e.g. Borio, 2011), consistent with our finding of disproportionate preference weights for large banks. Starting with the Stigler (1971) rent-seeking theory of regulation, a large literature also considers incentive issues for the supervisors themselves. For example, Kroszner and Strahan (1999) investigate the role of rent-seeking in bank branching restrictions. Agarwal et al. (2014) find differences in supervision across federal and state supervisors. Similarly, Granja and Leuz (2019) find differences between the Office of Thrift Supervision and the agencies that replaced it, the Office of the Comptroller of the Currency, and the Federal

Deposit Insurance Corporation. Kisin and Manela (2014) study the effect of fee structures on supervisory incentives.

Finally, our analysis of the allocation of supervisory hours is related to the literature on time-use of private households (Aguiar and Hurst, 2007; Blundell, Pistaferri, and Saporta-Eksten, 2018), which also takes a neoclassical approach in the spirit of Becker (1965). Our use of a structural model and counterfactuals to analyze changes in supervision are similar to the analysis of deposit fragility and the effects of changes in capital regulation of Egan, Hortacsu, and Matvos (2017) or the analysis of stress tests of Corbae et al. (2018).

The rest of the paper is organized as follows. Section 2 presents our economic model of supervision and the econometric specification. Section 3 discusses the data and Section 4 studies basic determinants of supervisory hours. Section 5 presents the instrumental variables used for supervisory hours. Section 6 provides the main results of our analysis: estimates for the effect of supervision and for supervisory preference and technology parameters. Section 7 uses the estimates to conduct counterfactual experiments and Section 8 presents our conclusion.

2 Economic model of supervisory resource allocation

This section presents the model of supervisory resource allocation and characterizes the optimal allocation. The model determines the allocation of supervisory resources as a function of bank characteristics, supervisory preferences, and the availability of overall resources. We first discuss technology, i.e. the determinants of bank distress and how it is impacted by supervisory hours. Then we add preferences and specify the supervisory objective function and characterize the resulting allocation of supervisory hours. To estimate the model parameters, we linearize the first-order condition and use it as the first stage in an instrumental variable probit specification. While the first stage of the model depends on both technological and preference parameters, the second stage depends only on technological parameters; combining the estimates from the first and the second stage allows us to untangle technological and preference parameters. The econometric specification explicitly accounts for the fact that supervisors have more information about the likelihood of future bank distress than the econometrician, requiring an instrumental variable approach.

2.1 Bank distress and supervisory technology

Let $y_{idt+1} \in \{0,1\}$ be an indicator variable for bank i in district d becoming distressed in year $t+1$. We use different measures of distress in the data: negative return on assets, extreme supervisory concerns (a supervisory rating of 4 or 5), or outright failure. Distress y_{idt+1} is determined by a continuous latent variable y^*_{idt+1}, such that $y_{idt+1} = \mathbb{I}\left[y^*_{idt+1} > 0\right]$ where $\mathbb{I}[\cdot]$ is the indicator function and $y^*_{idt+1} = D_{idt} + u_{idt+1}$. Here, D_{idt} is a distress threshold determined by the bank's and the supervisors' actions at t and $u_{idt+1} \sim \mathcal{N}\left(0, \sigma_u^2\right)$ is a shock realized at $t+1$. The distress threshold is given by $D_{idt} = \varrho_{idt} - \gamma s_{idt}$, where ϱ_{idt} denotes the riskiness of bank i due to its own actions in year t and s_{idt} denotes the intensity of supervision at bank i in year t that has impact γ.

The resulting probability of distress at $t+1$ is given by

$$\Pr\left[y^*_{idt+1} > 0 \,\middle|\, \varrho_{idt}, s_{idt}\right] = \Phi\left(\frac{\varrho_{idt} - \gamma s_{idt}}{\sigma_u}\right), \tag{1}$$

where Φ denotes the c.d.f. of the standard normal distribution. The parameter $\gamma \geq 0$ captures the effectiveness of the supervisory intensity s_{idt} in reducing the probability of distress. In practice, supervision affects outcomes through so-called corrective supervisory actions that, e.g. impose restrictions on the bank's asset growth and set of activities as well as mandated divestitures of certain assets (see Eisenbach et al., 2017 for more detail).

Bank riskiness ϱ_{idt} is partly reflected in the supervisory rating $r_{idt} \in \{1, \ldots, 5\}$, with lower numbers indicating lower risk, which we measure in the data as of the end of year t. If supervisors had no additional information on bank risk than is summarized in the rating, then γ could be identified with a standard maximum likelihood estimation of the probit model in (1), using data on ratings and supervisor hours. In practice, however, supervisors' information set is larger than the econometrician's, e.g. because ratings are granular and not updated continuously.[2] In our model, we formally account for this information asymmetry by positing that only the supervisor observes an additional signal η_{idt}, which is informative about future distress. Denoting by $\varrho(r_{idt})$ the component of bank risk reflected in the rating r_{idt}, the bank's total risk from the supervisors' perspective is $\varrho(r_{idt}) + \eta_{idt}$. From the econometrician's perspective, the latent distress variable is then

$$y^*_{idt+1} = \varrho(r_{idt}) - \gamma s_{idt} + u_{idt+1}, \tag{2}$$

[2] Our model does not spell out separately the technology by which supervisors monitor for risk and intervene against risk (see Eisenbach et al., 2017 for more detail). The model is used to interpret the data and the data do not distinguish between monitoring and intervention. As we show below, even the safest banks (rating 1) receive resources, consistent with the notion that some resources are indeed used to monitor banks.

where the error u_{idt+1} is

$$u_{idt+1} = \eta_{idt} + \varepsilon_{idt+1}. \tag{3}$$

While $\varepsilon_{idt+1} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is observed neither by the supervisors nor the econometrician, η_{idt} is observed by the supervisors at t. Because the rating is meant to capture supervisors' best assessment of a bank's risk at the time it is assigned, we largely think of the signal η_{idt} as accumulating ahead of a rating assignment. That is, conditioning on rating r_{idt}, we assume that $\eta_{idt} \sim \mathcal{N}(0, \sigma_\eta^2)$ and that it is independent of ε_{idt}. This means that the end-of-year rating r_{idt} incorporates prior realizations $\{\eta_{idt-\tau}\}_{\tau>0}$ but not private information that supervisors observe after they assign the rating. Instruments are therefore necessary to break the contemporaneous dependence of the endogenous variable s_{idt} on η_{idt} but not on its lags.

Economies of scale in supervision. The latent variable y^*_{idt+1} and its components do not directly depend on bank size. In the data, however, bank size varies across several orders of magnitude and strongly affects the allocation of supervisory resources. Banks with larger balance sheets have larger loan portfolios and engage in more activities, both of which require more supervisory resources. How much more resources depends on whether there are economies or diseconomies of scale in supervision, a question we want to address empirically. Let the function h determine how many hours are needed to achieve the supervisory intensity s_{idt} at a bank of asset size A_{idt}:

$$h(s_{idt}, A_{idt}) = \exp(s_{idt}) \times A_{idt}^\alpha$$

The function h is increasing and convex in the supervisory intensity s_{idt}, has elasticity $\alpha > 0$ with respect to bank size A_{idt}, and implies a log-linear relation $s_{idt} = \log H_{idt} - \alpha \log A_{idt}$. For our concept of scale economies, we consider the output of supervision to be achieving a certain distress probability (represented by a threshold D_{idt}) at a bank of risk ϱ_{idt} and size A_{idt}. The supervisory hours cost function can be written accordingly as

$$H_{idt} = \exp\left(\frac{\varrho_{idt} - D_{idt}}{\gamma}\right) \times A_{idt}^\alpha. \tag{4}$$

We say that there are economies of scale in supervision if the average cost, i.e. the ratio H/A, is decreasing in A, which is the case if $\alpha < 1$.[3] The supervisory cost function (4) is the

[3]The distress threshold D_{idt} could also depend directly on bank size, e.g. if larger banks are inherently less risky because of diversification benefits in a way not captured by the supervisory rating. We show in Appendix A.1 that our evidence is valid also under this more general formulation. Ultimately, scale economies mean that larger banks can be supervised with proportionately fewer resources, and this can be due to a combination of larger banks being inherently less risky and supervisors being more effective at supervising

key expression that determines the technology of supervision in our model, represented by the impact parameter γ and the scale economy parameter α.

2.2 Supervisory objective and allocation of hours

As discussed in more detail below, supervision is implemented at the level of each of the twelve Federal Reserve districts. We therefore consider supervisors in a district d allocating resources to a set of banks $i \in \mathcal{I}_{dt}$ with an available budget of supervisory hours $\overline{H}_{dt}$. The supervisors' objective is to minimize a weighted sum of the banks' distress probabilities:

$$\min_{\{s_{idt}\}} \sum_{i \in \mathcal{I}_{dt}} \Pr\left[y^*_{idt+1} > 0 \,\middle|\, \varrho_{idt}, s_{idt}\right] W_{idt} \quad \text{subject to} \quad \sum_{i \in \mathcal{I}_{dt}} h(s_{idt}, A_{idt}) \leq \overline{H}_{dt} \tag{5}$$

We use the resource allocation model to disentangle the role that technology, preferences, and overall resources play in shaping the observed allocation and, ultimately, bank outcomes. The supervisory technology given by (4) is determined by the parameters α and γ. Supervisory preferences are captured by the preference weight W_{idt} given to bank i, which we infer from the data via a revealed preference approach. Supervisory weights are the result of an interplay of law, regulation, guidance, and other factors including push-and-pull with regulated entities. We therefore do not assume that the supervisory objective function accurately reflects a social welfare function. We conduct counterfactual policy experiments to quantify how much supervisory preferences affect the likelihood of bank distress as well as the role played the level overall resources $\overline{H}_{dt}$.

We assume that preferences are the product of three factors: $W_{idt} = A^{\widetilde{\alpha}}_{idt} \times f(\varrho_{idt}) \times \exp w_{idt}$. The dependence of the preference weight on bank size reflects the cost of bank distress, as viewed by the supervisors. It can account for both micro-prudential objectives (the cost of distress only to the individual bank) and macro-prudential objectives (e.g. the spillover costs of distress on the wider economy). Under systemic risk concerns, we would expect $\widetilde{\alpha} > 1$, i.e. the weight to increase more than proportionally with bank size.[4]

Conceptually, the cost of distress conditional on the distress event occurring should not depend on the probability of distress itself. The dependence of the preference weight on bank risk allows for the possibility that supervisors weight events differently than the objective probability measure. This is captured by $f(\varrho_{idt})$, which relates the supervisors'

larger banks.

[4]The objective in (5) is similar to the credit risk framework (probability of default $\times$ loss given default) used to calibrate regulatory capital requirements (Basel Committee on Banking Supervision, 2010). For example, in setting capital surcharges for global systemically important banks (GSIBs), the Federal Reserve considers a bank's "systemic loss given default," which explicitly includes externalities to the overall stability of the financial system (Board of Governors of the Federal Reserve System, 2015).

subjective distress probability to the objective one, as in a Radon–Nikodym derivative. If supervisors overweight small probabilities, then the preference weight is tilted toward safer banks, i.e. $f'(\varrho_{idt}) < 0$. Finally, w_{idt} captures anything that affects the allocation of hours but not (directly) the probability of distress; these pure preference shifters are key to our identification strategy and discussed in detail in Section 5.

The supervisors' first-order condition in terms of the intensity s_{idt} is

$$-\frac{\partial}{\partial s}\Pr\left[y^*_{idt+1} > 0 \,\middle|\, \varrho_{idt}, s_{idt}\right] \times W_{idt} = \frac{\partial}{\partial s} h(s_{idt}, A_{idt}) \times \Lambda_{dt}. \tag{6}$$

The left-hand side is the benefit of additional supervision at bank i: the reduction in distress probability multiplied by the preference weight. The right-hand side is the cost of additional supervision at bank i, which combines (i) the marginal hours cost given the bank's size and (ii) the shadow cost of hours, given by the Lagrange multiplier Λ_{dt} on the hours budget constraint. This is because increasing hours at bank i requires reducing hours at other banks in district d, an important source of identification in our empirical application. The first-order condition (6) makes clear predictions about the comparative statics of the allocation of supervisory hours with respect to bank size, riskiness, preference weight, and overall resource scarcity:

Proposition 1. *Supervisory hours at bank i are increasing in bank i's preference weight W_{idt} and decreasing in the shadow cost of hours Λ_{dt}. Holding constant W_{idt}, supervisory hours are increasing in bank size A_{idt} and risk ϱ_{idt}. If W_{idt} is increasing (decreasing) in size and/or risk then the respective effect on hours is strengthened (attenuated).*

Proof. See Appendix A.2. □

Proposition 1 shows that supervisory hours are increasing in both bank size and risk, even when the distress probability is weighted equally across banks (constant preference weight) — because large banks require additional resources for the same intensity of supervision and because the marginal impact of supervision is higher at riskier banks.

In the main empirical specification we measure bank riskiness with supervisory rating dummies, $\varrho(r_{idt}) = \rho_1 + \sum_{r=2}^{5}\rho_r \mathbb{I}[r_{idt} = r]$, as of the end of year t; hours H_{idt} are total supervisory hours at bank i in year t and assets A_{idt} are total assets at the end of year t. After combining model equations (1)–(4), the probability of distress conditional on the supervisors' information set, including η_{idt}, is

$$\Pr\left[y^*_{idt+1} > 0 \mid r_{idt}, H_{idt}, A_{idt}, \eta_{idt}\right] \\ = \Phi\left(\frac{\rho_1 + \sum_{r=2}^{5}\rho_r \mathbb{I}[r_{idt} = r] - \gamma \log H_{idt} + \alpha\gamma \log A_{idt} + \eta_{idt}}{\sigma_\varepsilon}\right), \quad (7)$$

and the first-order condition (6) is

$$\phi\left(\frac{\rho_1 + \sum_{r=2}^{5}\rho_r \mathbb{I}[r_{idt} = r] - \gamma \log H_{idt} + \alpha\gamma \log A_{idt} + \eta_{idt}}{\sigma_\varepsilon}\right)\frac{\gamma}{\sigma_\varepsilon} W_{idt} = H_{idt}\Lambda_{dt}. \quad (8)$$

We parameterize the preference weight W_{idt} as a log linear function,

$$\log W_{idt} = \widetilde{\rho}_1 + \sum_{r=2}^{5}\widetilde{\rho}_r \mathbb{I}[r_{idt} = r] + \widetilde{\alpha} \log A_{idt} + w_{idt}, \quad (9)$$

where tildes denote preference parameters as opposed to the corresponding technological parameters in the probability of distress (7), and w_{idt} are preference shifters that affect the allocation of hours but are exogenous to the probability of distress.

With our functional forms, the comparative statics of optimal hours with respect to bank characteristics takes a particularly simple form.

Proposition 2. *For any conditioning variable x_{idt} with loading κ in the distress variable y^*_{idt+1} and loading $\widetilde{\kappa}$ in the supervisory preference weight $\log W_{idt}$, the local effect of x_{idt} on optimal hours $\log H_{idt}$ is a convex combination of the loadings $\kappa, \widetilde{\kappa}$ and the impact of supervision γ,*

$$\frac{d \log H_{idt}}{dx_{idt}} = \pi_{idt}\frac{\kappa}{\gamma} + (1 - \pi_{idt})\,\widetilde{\kappa},$$

where the local weight π_{idt} is given by

$$\pi_{idt} = \frac{\Phi^{-1}\left(\Pr\left[y^*_{idt+1} > 0 \mid r_{idt}, H_{idt}, A_{idt}, \eta_{idt}\right]\right)}{\Phi^{-1}\left(\Pr\left[y^*_{idt+1} > 0 \mid r_{idt}, H_{idt}, A_{idt}, \eta_{idt}\right]\right) - \frac{\sigma_\varepsilon}{\gamma}},$$

with $\pi_{idt} \in (0, 1)$ for $\Pr\left[y^*_{idt+1} > 0 \mid r_{idt}, H_{idt}, A_{idt}, \eta_{idt}\right] < 1/2$.

Proof. See Appendix A.3. □

Proposition 2 shows how the hours allocation observed in the data depends on both supervisory technology and preferences, with the empirical loading being a convex combination of the respective parameters. For example, the average elasticity of hours with respect to bank size is a convex combination of the size elasticity α of the technological hours cost function (4) and the size elasticity $\widetilde{\alpha}$ of the preference weight (9),

$$E\left[\frac{d\log H_{idt}}{d\log A_{idt}}\right] = \overline{\pi}\alpha + (1-\overline{\pi})\,\widetilde{\alpha},$$

with $\overline{\pi} = E[\pi_{idt}]$. Based on this expression, an estimated unit elasticity of hours with respect to size could arise from a combination of economies of scale, $\alpha < 1$, and preference weights increasing more than proportional with size, $\widetilde{\alpha} > 1$. Without the structural approach, the two effects could not be separately identified.

2.3 Econometric specification and identification challenge

To estimate the parameters of the model, we linearize the first-order condition (8) as shown in Appendix A.4.[5] Together with the latent variable y^*_{idt+1} in (2) and the evolution of the binary variable y_{idt+1}, we obtain a probit model with an endogenous regressor, which can be estimated with maximum likelihood:

$$\begin{cases} y_{idt+1} = \mathbb{I}\left[y^*_{idt+1} > 0\right] & (10) \\ y^*_{idt+1} = \beta_0 + \beta_H \log H_{idt} + \beta_A \log A_{idt} + \sum_{r=2}^{5}\beta_r \mathbb{I}[r_{idt} = r] + u_{idt+1} & (11) \\ \log H_{idt} = \delta_0 + \delta_A \log A_{idt} + \sum_{r=2}^{5}\delta_r \mathbb{I}[r_{idt} = r] + \delta_w w_{idt} + \delta_\Lambda \Lambda_{dt} + v_{idt} & (12) \end{cases}$$

The error terms in the second stage (11) and the first stage (12) are $u_{idt+1} = \eta_{idt} + \varepsilon_{idt+1}$ and $v_{idt} = \delta_\eta \eta_{idt}$, respectively, with $\delta_\eta > 0$ a coefficient of the linearization. The shock η_{idt} is the source of the econometric bias. Since $\text{cov}(\log H_{idt}, \eta_{idt}) = \delta_\eta \sigma^2_\eta$, it is also the case that $\text{cov}(\log H_{idt}, u_{idt+1}) = \delta_\eta \sigma^2_\eta$. Intuitively, the identification challenge is that supervisors allocate more hours to riskier banks, so an un-instrumented regression of future bank distress on supervisory hours cannot identify the effect of supervision on bank distress. The endogeneity of hours with respect to risk attenuates the estimated effect of hours in the probit equation (11), so an un-instrumented estimate of γ would be downwardly biased (in absolute terms) or even switch sign. Controlling for bank risk with the observable supervisory rating r_{idt} only partially resolves the endogeneity problem if, as our model assumes and our empirical analysis confirms, supervisors use additional private information η_{idt}. We thus use instrumental variables to study variation in supervisory hours that is exogenous to bank risk. The first stage (12) directly informs on variables that can be used as

[5] Our model could also be estimated via non-linear GMM and, in fact, the first-order condition (8) is a quadratic equation in log hours. However, we want to draw a tight connection to other contributions in this literature that use linear approaches (e.g. Hirtle, Kovner, and Plosser, 2020). The more transparent IV probit approach permits a direct comparison between our study and other papers in the field.

instruments to identify the parameters: supervisory preference shifters w_{idt} and the Lagrange multiplier Λ_{dt}, which measures the overall scarcity of supervisory resources. As in standard probit models, the parameters in the second stage (11) cannot be separately identified from the variance of the error term and we therefore normalize $\text{var}(u_{idt+1}) = 1$.

From regression coefficients to model parameters. The reduced-form coefficients of the instrumental-variable probit (10)–(12) are functions of the structural parameters in the model equations (7)–(9),

$$\begin{array}{llll} \text{Coef. 2}^{\text{nd}}\text{ stage:} & \beta_H = -\gamma, & \beta_A = \alpha\gamma, & \beta_r = \rho_r \\ \text{Coef. 1}^{\text{st}}\text{ stage:} & & \delta_A = \overline{\pi}\alpha + (1-\overline{\pi})\,\widetilde{\alpha}, & \delta_r = \overline{\pi}\frac{\rho_r}{\gamma} + (1-\overline{\pi})\,\widetilde{\rho}_r, \end{array} \tag{13}$$

with $\overline{\pi} = E[\pi_{idt}]$ from Proposition 2; β_0 and δ_0 constants; $\delta_w > 0$ and $\delta_\Lambda < 0$ coefficients of the linearization.

Note that the second-stage coefficients on log hours, β_H, log assets, β_A, and ratings, β_r, directly yield parameters that describe the technology of supervision and the evolution of bank risk. Specifically, the coefficient on log hours yields an estimate of (the negative of) γ, i.e. the loading of the distress variable y^*_{idt} on the intensity of supervision s_{idt}. The marginal effect of supervision on the probability of distress is then equal to the product of γ and the density function, $\gamma\phi(\cdot)$. For a given estimate of γ from the coefficient β_H, the second-stage coefficient on log assets, β_A, yields yields an estimate of the parameter measuring economies (or diseconomies) of scale in supervision, i.e. the elasticity of hours cost with respect to size, $-\beta_A/\beta_H = \alpha$.[6] Finally, the coefficient on rating $r = 2, \ldots, 5$ yields an estimate of ρ_r, i.e. the loading of D_{idt} on the dummy for rating r relative to the left-out rating 1.

In contrast, the first-stage coefficients on log assets, δ_A, and ratings, δ_r, yield linear combinations of the respective preference parameters — $\widetilde{\alpha}$ and $\widetilde{\rho}_r$ — with the corresponding technological parameters — α and ρ_r (Proposition 2). This is because the first stage is derived from the first order condition of the resource allocation problem and thus depends both on technology and preferences. While we cannot separately identify the two effects from the first stage alone, with help from the second-stage coefficients and an estimate of $\overline{\pi}$ from its sample analog, we can untangle the preference parameters relative to the technological parameters.

The Lagrange multiplier Λ_{dt} enters the first stage (12) with a coefficient $\delta_\Lambda < 0$ due to the linearization. However, since δ_Λ is constant across districts and time, variation in Λ_{dt}

[6] We show in Appendix A.1 that our estimation identifies economies of scale even in a more general formulation where bank risk depends directly on size.

is the same as variation in $\lambda_{dt} \equiv -\delta_\Lambda \Lambda_{dt}$; for brevity, we also refer to λ_{dt} as the "Lagrange multiplier."

3 Data

We use four types of data: (i) institution-level work hours for supervisory staff at the Federal Reserve; (ii) examination information and supervisory ratings from the National Examination Database (NED); (iii) balance sheet information, asset quality, and profitability from Y-9C regulatory filings; and (iv) bank structure information from the National Information Center (NIC). We discuss each data source and then present summary statistics for the variables included in the regressions. Exact variable definitions are provided in Appendix B.

Hours spent by Federal Reserve supervisory staff are from an internal database. The Federal Reserve supervises state member banks (SMBs) as well as all bank holding companies on a consolidated basis. Because we do not observe hours at non-SMB banks, which are supervised by other agencies, we focus on BHCs, which are exclusively under the purview of Federal Reserve supervisors. The hours data starts in 1998 and ends in 2014. The information is reported by supervisory employees on a weekly basis and includes information on the supervised BHC through its regulatory entity number (RSSD ID). Supervisory work at the smallest institutions is often recorded using a generic bank portfolio assignment, as opposed to an institution RSSD ID. By cross-checking hours information with independent information on the timing of supervisory inspection from NED, we find that consistent hours information with valid supervised-entity information is only available for institutions with assets of about \$750 million or more; we therefore exclude institutions with less than \$1 billion in assets. For each institution, we aggregate data by year, so that the resulting supervisory hours data is a dataset uniquely identified by a year and the supervised institution's RSSD ID.[7]

We match hours information to two other data sources. First, we obtain information on bank characteristics, including balance sheet and income statement, from public Y-9C reports, which are used to assess the financial condition of BHCs on a consolidated basis. In addition, we match supervisory hours to confidential information on supervisory ratings and exams from NED. Bank holding companies are assigned a rating from 1 to 5 under the "RFI/C(D)" rating system, with lower ratings indicating fewer supervisory concerns. The acronym indicates the different components considered in constructing the

[7] We have information on pre-2000 hours for only a handful of districts and have information on all districts only starting in 2006.

rating.[8] We also obtain the yearly count of examinations from NED.

In terms of outcomes, we use three variables to measure the distress of a BHC by degree of severity. We count as outright failures whenever a termination of a BHC is recorded in the regulatory NIC data due to a failure of the holding company or when a subsidiary fails within one quarter of a BHC termination, for example because the holding company is acquired or merged. Because of the low incidence of actual failures during normal times, we additionally identify banks under "severe stress" that fail or have a rating of 4 or 5 at some point over the course of a year (officially referred to as "problem banks"). Finally, we use a realization of the return on assets (ROA) below the 10th percentile of the pooled distribution (precisely it is 10 basis points).[9] Outcomes are measured in the year after supervisory hours are recorded.

Summary statistics. Table 1 provides summary statistics for the variables included in the regression specifications, split into small and large banks ($10 billion threshold). The (unbalanced) panel is composed of about 770 unique BHCs located in the twelve Federal Reserve districts over the 1998–2014 time interval. As shown in panel (a), the average probability of failure is 0.5%, severe stress occurs about 5% of the time, and, consistent with its definition, the ROA falls below the 10th percentile about 10% of the time. As shown in the first column of the panel (b), Federal Reserve supervisors allotted about 1,500 hours per year, on average, to supervising a BHC in our sample. Based on an eight-hour work day and 48 work weeks per year, a full-time supervisor works about 1,900 hours per year, so these recorded hours can be converted into about three supervisors for every four bank holding companies. However, large institutions receive on average three times as many hours, while smaller ones receive only about one fifth.[10] About 20% of our sample is composed of BHCs with assets greater than $10 billion, and roughly 15% are also among the largest BHCs within their respective districts (as defined in Hirtle et al., 2020). The average supervisory rating in our sample is 2 and 15% of the sample is composed of banks with a

[8] Specifically, "R" is for risk management, "F" is for financial condition, "I" is for potential impact of the non-depository entities in the holding company on the depository institution(s) in the holding company, "C" is for the composite rating (that is, the overall rating considering and weighting the ratings on "R", "F", and "I"), and "D" is the rating assigned to the depositories (for example commercial banks or thrifts) owned by the holding company. Prior to 2014, BHCs received ratings known as BOPECs, an acronym that stood for five areas of supervisory concern. Despite some differences, BOPECs and RFI/C(D) rating levels have similar supervisory interpretations and we splice these measures together in our analysis.

[9] From a supervisory perspective, failures of supervised banks are the most meaningful events, but these events are also rare (only 0.5% in our data), and we therefore consider the alternative outcomes as well. Severe stress carries negative consequences such as inclusion on the FDIC's list of "problem banks" and occurs with roughly 5% probability. Low ROA occurs, by definition, with 10% probability.

[10] This calculation excludes hours that have not been booked by the supervisor to a specific institution. In addition, the day-count translation would underestimate an actual headcount because it doesn't account for other administrative or training activities that a supervisor may be involved in when not assigned to a bank.

"stressed" rating of 3 or worse.

Bank supervisors make a complexity assessment annually for each BHC (RSSD 9057) using a number of criteria, such as material credit-extending activities; significant and risky nonbank activities, such as securities broker-dealer activities or insurance underwriting; and subsidiaries that issue significant debt to the general public. About 30% of our sample is composed of complex BHCs, most of which fall into the large BHC category. In Section 5, we use information on rating, size, and complexity to construct a "high examination frequency" dummy variable that identifies BHCs that receive examinations at least once per year year. On average, BHCs receive about 1.5 examinations each year; large BHCs, which all have "high examination frequency," receive about four examinations per year, while smaller BHCs receive less than one per year.

4 Basic determinants of supervisory hours and post-2008 reallocation to large banks

To gain intuition on how supervisory hours are related to observable bank characteristics, Table 2 presents OLS parameter estimates from a linear regression of log supervisory hours on log bank assets (expressed in constant 2012 dollars) and bank risk (measured in terms of confidential supervisory ratings). Bank size and risk appear prominently in guidance to examiners in the Fed supervisory manual.[11] To document potential differences in supervision of large and small banks and the post-2008 resource reallocation from small to large banks, we then augment the baseline regression with dummy variables for the post-2008 period, for bank size buckets, and their interaction. The model specifications shown in the table differ in terms of inclusion of year fixed effects (odd columns).[12] We discuss robustness to additional risk controls in Section 6.5. In all regression tables, standard errors clustered at the bank level are reported in square brackets.

The coefficient on log assets captures the elasticity of supervisory hours with respect to bank size. If hours increase proportionally with assets, the coefficient is equal to 1, and if less than proportionally, less than 1. We find an elasticity of supervisory hours to bank assets of slightly less than 1 (columns 1 & 2) and of about 0.8 when including dummies for large banks (columns 3–6). This size-elasticity estimate corresponds to δ_A in (13), which commingles the economies of scale in supervision, α, and the size elasticity of the super-

[11] Available at https://www.federalreserve.gov/publications/supervision_bhc.htm.

[12] Results are robust to including bank fixed effects. However, we cannot use bank fixed effects in our IV probit estimation because of the incidental parameter problem — roughly 700 parameters estimated in a probit with only about 7 (yearly) observations per bank.

visor's preference weight, $\widetilde{\alpha}$. The structural estimation in the next section separates these two effects.

Because supervisory ratings of bank risk are ordinal rather than cardinal measures, we use separate dummy variables for each rating, leaving out the best rating category of 1. The effect of bank risk is very significant both statistically (p-vals < 0.01) and economically. For example, as compared to a bank with the best possible rating of 1, which is the omitted category in the regression, a bank of the same size but with a rating of 3 receives close to three times more hours ($\exp(1.314) - 1 \approx 2.7$, column 1),[13] an effect comparable to more than doubling the size of the bank. Similar to the coefficient on size, the estimated coefficients on bank risk commingle the loading ρ_r of future bank distress on current risk with the loading $\widetilde{\rho}_r$ of supervisor preference weights on banks of different risk.[14]

The Federal Reserve's mandates require enhanced supervision at banks with assets above \$10 billion. Banks with assets over \$50 billion are also subject to the Federal Reserve's "Consolidated Supervision Framework for Large Financial Institutions," which implies further enhanced supervisory practices. Furthermore, the mandated supervision of large banks intensifies after the financial crisis. We investigate these threshold effects starting with column 3 of Table 2. Controlling for log assets and rating, the point estimate of 0.328 on the assets$\geq$\$10b dummy implies that banks above the size threshold receive close to 40% more hours than banks below the threshold (p-val < 0.05). Figure 1 in the introduction shows that supervisory resources expand significantly post-2008 and that the expansion differs between large and small banks. We characterize this expansion by studying the estimated effect of an interaction term between a post-2008 dummy and the large bank dummy. The coefficient of 0.655 implies that hours at large banks increased by about 90% (p-val < 0.01, column 3) after the financial crisis.

The fact that the increase in hours at large banks is considerably larger than the post-2008 expansion in total supervisory resources of roughly 50% (Figure 1, left panel) points to a gross reallocation from small to large banks as a result of binding resource constraints. Column 4 drops the year fixed effects in order to include a post-2008 dummy that captures the change in hours at small banks. Consistent with a reallocation, we find that, controlling for size and risk, hours at banks smaller than \$10 billion decrease by over 20% post-2008,

[13] Note that in regressions with log hours as the dependent variable, the coefficient δ_d on a dummy variable d has to be transformed as $\exp(\delta_d) - 1$ to calculate the percentage change in hours for a dummy value of 1 vs. 0: $\delta_d = \log H|_{d=1} - \log H|_{d=0}$ implies $\exp(\delta_d) - 1 = (H|_{d=1} - H|_{d=0}) / H|_{d=0}$ (Halvorsen and Palmquist, 1980).

[14] As an alternative measure of supervisory efforts, in Appendix C, we use supervisory fees assessed on federally chartered commercial banks by the OCC, which reflect supervisory costs at institutions as a function of risk and size. Overall, the sensitivities of Federal Reserve supervisory hours and OCC assessment fees to size and risk are similar.

despite the increase in total supervisory resources at the Federal Reserve. After controlling for log assets, ratings, and the \$10 billion size dummy, banks larger than \$50 billion do not have measurably higher supervisory hours (column 5). Analogous to this baseline effect, controlling for the interacted break at \$10 billion, the interaction at \$50 billion is not statistically significant (column 5).

Figure 2 shows this reallocation by displaying the relation between log hours and log assets — after conditioning on rating — with a break at \$10 billion. The log-linear relation fits the data quite well in the pre-crisis period (1998–2008) but there is a clear break in the post-crisis period (2009–2014). Large banks ($\geq$ \$10 billion) seem to receive discretely more attention than small banks in the post-2008 sample. These results suggest a significant shift in the focus of supervisors toward larger banks. Our revealed preference framework captures this effect as a post-2008 increase in the preference weight on large banks, consistent with changing mandates and new macroprudential objectives.[15] To capture the differential treatment of large banks and the post-2008 reallocation, we include the assets$\geq$\$10b dummy and its interaction with the post-2008 dummy in our specifications. Despite the fact that this reallocation is consistent with a preference shift, we do not exploit this variation for identification of resource allocation in the next section because other post-2008 regulatory changes affecting large banks may invalidate the exclusion restriction. We, instead, include the size dummy and interaction with post-2008 in the second stage of the IV probit estimation, i.e. we do not impose an exclusion restriction for the interaction term.

5 Identification of the model parameters

In the structural model described by equations (10)–(12), the preference shifters w_{idt} and the Lagrange multiplier λ_{dt} are valid instruments "z_{idt}" for log hours. Because they enter the first but not the second stage, they satisfy the exclusion restriction, $\mathrm{cov}(\eta_{idt}, z_{idt}) = 0$, and the relevance condition, $\mathrm{cov}(\log H_{idt}, z_{idt}) \neq 0$. We measure λ_{dt} as resource scarcity for a given Federal Reserve district and year and consider two separate proxies for the preference shifter w_{idt}, excess attention paid to the top banks in each district and variation in minimum exam frequency for small banks. The three instruments exploit different sources of variation in the data, enhancing the external validity of our estimates. Identification from the "shadow cost" instrument uses variation in the allocation of resources at banks of all sizes, the "district top" instrument largely uses variation for large banks, and

[15] SR letter 12-17 explicitly bases the post-crisis supervisory framework for large banks on the crisis experience and cites macroprudential concerns, which are captured in our framework by the supervisory preference weight. The continuity of the cost function specification in (4) with respect to size implies that any discontinuities in the allocation of hours are not attributed to technology.

the "examination frequency" instrument relies on variation for small banks over time. We discuss the empirical validity of each instrument in turn.

5.1 Resource scarcity within districts

From Proposition 1, supervisory resources allocated to a bank not only depend on the bank's characteristics and supervisors' preference weight for the bank, but also on the shadow cost of supervisory resources, as measured by the Lagrange multiplier λ_{dt}. The shadow cost is a valid instrument for log hours if, after conditioning on bank characteristics, it only affects the probability of distress through supervisory hours.[16]

The Federal Reserve Board of Governors has the authority and responsibility for supervising BHCs on a consolidated basis. In practice, each of the twelve Reserve Banks in the Federal Reserve System supervises the BHCs that are located within its own district under delegated authority from the Board of Governors. While supervisory activities are coordinated at the Federal Reserve System level via committees, each Reserve Bank employs dedicated supervisory staff ("examiners") and determines its own hiring, performance assessments, and staff allocations. We consider the relevant budget constraint to be at the district–year level, consistent with the data where nearly all hours allocated to a bank (95%) are from staff at its district's Reserve Bank. We also show below that hours allocated to a bank in a given district are unaffected by changes in the shadow cost in other districts.

Correlated shocks within districts are a potential threat to the exclusion restriction. In this case a regional shock would not only lower supervisory hours through λ_{dt} but also imply that conditions at the observed bank may have worsened. To control for this possibility, all specifications include a bank's current rating, which, as a supervisory summary of the bank's risk, also accounts for the effect of regional shocks. In robustness checks (Section 6.5), we find similar results when including additional bank-level risk controls as well as accounting for differences across districts with district fixed effects. All our specifications include year fixed effects.[17]

[16]More generally, the shadow cost of supervisory resources would also increase under the less stringent assumption that total resources can be partially, but not fully, adjusted at a given point in time, for example due to budgetary processes or simply the time it takes to hire additional staff.

[17]Another potential concern with the exclusion restriction is that district supervisors could request additional supervisory resources based on the district average realization of the supervisors' private information η_{idt}. While η_{idt} is unobserved, if local supervisors can successfully lobby for additional resources then current changes in total hours in a district should predict future changes in district-level ratings or our three distress outcomes because, by its definition, η_{idt} is informative about future distress. In unreported results we do not find this to be the case. Furthermore, the ability of district supervisors to lobby for additional resources would imply that total Federal Reserve resources matter for the hours allocation as opposed to

The Lagrange multiplier is not directly observable. However, because it is common to all banks in a district and year, λ_{dt} can be estimated based on the first stage (12) from district–year averages of supervisory hours, bank characteristics, and preference shifters up to the unknown coefficients. Estimates of the coefficients, in turn, also depend on the estimates of the Lagrange multipliers. We therefore instrument for log hours using a "plug-in" estimator of λ_{dt} by directly including district–year averages of supervisory hours and bank characteristics in the first stage, thereby jointly estimating all parameters. Following work on judicial outcomes identified through variation in judge leniency (Dahl et al., 2014; Dobbie et al., 2018), we use leave-out averages excluding bank i, denoted by $\overline{x}_{-idt}$ for variable x. As noted by Dobbie et al. (2018), leave-out averages are equivalent to leave-out fixed effect estimators and can be interpreted as reduced-form jackknife IV estimators (Angrist et al., 1999). These leave-out averages are recommended in our setting and lead to more conservative estimates because, without leaving out, measurement error in the dependent variable would also appears in the independent variable. The plug-in estimator is then obtained by replacing λ_{dt} in the first stage with the leave-out averages $\overline{\log H}_{-idt}$, $\overline{\log A}_{-idt}$ and of the five $\overline{\mathbb{I}[r]}_{-idt}$.[18]

Column (1) of Table 3 shows regression estimates of log hours at bank i on within-district leave-out averages of log hours, log assets, and each of the rating indicators. As noted at the bottom of the table, the regression also includes bank i's own log assets and supervisory rating indicators as well as the reallocation controls and year fixed effects (coefficients omitted from the table). The estimated coefficients on the leave-out averages have the expected signs: log hours at bank i are increasing in average log hours of other banks in the district, declining in average log assets, and declining in the rating indicators with larger effects for worse ratings. The loadings on average log assets and average log hours have similar magnitudes because the elasticity of hours with respect to assets is close to 1 in the data. As shown at the bottom of the table, the F-statistic for the null that the coefficients on average log hours, log assets, and rating indicators are zero is 38.8.

In contrast to the leave-out averages in our baseline specification, column (2) of Table 3 uses averages including bank i. In this case, the coefficient on average log hours is (by

district ones. But we find the opposite as discussed below.

[18]The judge leniency literature exploits quasi-random assignment of cases to judges who vary in their leniency. Variation in leniency is constructed from leave-out averages of the judges' other decisions after controlling for other covariates in a supplementary regression. Similarly here, the Lagrange multipliers are measured from supervisory hours allocations at other banks. But differently from the judge leniency literature, the variation in the Lagrange multiplier is determined from the leave-out average of the first stage itself because the structural model fully characterizes the hours allocation. Intuitively, in our model, variation in the shadow cost of hours is measured by the gap between the average size and risk of other supervised banks relative to average hours assigned to them, which can be interpreted as a measure of "supervisory workload."

construction) approximately equal to 1 and the F-statistic increases to 67.2, suggesting a "stronger" instrument than when we use leave-out averages. However, we use leave-out averages as our preferred specification because of the likely bias when using leave-in averages, as noted above. In column (3), we augment the regression with averages for banks in all districts other than the one where bank i is located. We find mostly insignificant coefficients, consistent with the assumption that the relevant resource constraint is at the level of a Federal Reserve district, as opposed to at the level of the entire Federal Reserve System. In column (4), we augment the specification with averages of the other first-stage covariates, including the preference shifters discussed in Section 5.2 below, and find similar results. In Appendix A.5, we characterize cross- and within-district variation in the shadow cost instrument and show that their contribution to its variation has about equal magnitude.

5.2 Supervisory preference shifters

We consider two supervisory preference shifters w_{idt}: an indicator for the largest banks within each Federal Reserve district and the number of supervisory examinations in the year prior for non-complex banks with assets $<$ \$10 billion and a satisfactory rating (equal to 1 or 2). Before we discuss why these variables act as preference shifters and their validity as instruments, it is useful to note that they exploit different sources of variation. The "district top" instrument, which was proposed by Hirtle et al. (2020), relies on cross-sectional variation among large banks; the lagged examination count for small, non-complex, well-rated banks exploits full panel variation among small banks.

District top. Hirtle et al. (2020) show that the largest banks in each Federal Reserve district receive additional supervisory attention, after accounting for their size and risk, suggesting that regional supervisors are most concerned with the performance of the largest banks in their district. As in Hirtle et al. (2020), we construct a "district top" dummy for the largest five banks in each district, as well as the banks within 25% of the size of the fifth ranked bank in their district. Column (1) of Table 4 corroborates their finding in our longer sample. The highly significant coefficient of 0.619 on "district top" implies that these banks receive 86% more hours than other banks ($\exp(0.619) - 1 \approx 0.86$, p-val < 0.01, F-statistic of 20.7). As noted at the bottom of the table, this is true after controlling for a bank's supervisory rating and, most importantly, the size of the bank as measured by log assets.[19]

[19] As in all of our first-stage specifications, column 1 also includes the \$10 billion threshold dummy and its interaction with the post-2008 dummy. However, since "district top" banks naturally tend to be large banks, there is a fair amount of overlap with the assets$\geq$\$10b dummy (82% of "district top" banks are larger than \$10 billion). To separately assess the strength of the "district top" instrument, we exclude "district top"

The exclusion restriction for this instrument is that "district top" banks differ from other banks with the same size and rating only because the ordinal ranking within their district implies additional supervisory resources. Hirtle et al. (2020) provide support for this assumption by including a large set of additional controls and by matching "district top" banks with similar non-top banks in other districts.

Exam frequency. The second preference shifter instrument is based on minimum mandated examination frequencies which depend on size, supervisory rating, and complexity. Differences in mandated examination frequencies generate variation in supervisory hours that is uncorrelated with the shock η_{idt}, as we explain next. According to the Federal Reserve supervisory manual, among BHCs with assets below \$10 billion, those deemed non-complex and rated 1 or 2 have to be examined at least once every two years while those deemed complex or rated worse than 2 have to be examined at least once every single year. BHCs above \$10 billion also have to be examined at least once every year. Appendix Table A4 lists in detail the exam frequency requirements. We accordingly classify banks in our sample into "low exam frequency" ($<$\$10b in assets, non-complex, and rated 1 or 2) and "high exam frequency" (all other banks). We find that supervisors tend to stick to the minimum frequency stipulated by their policy manual, but not perfectly: for example, about 20% of the time supervisors do more than the minimum for low-frequency banks (more than one exam in two years).

As for the shadow cost instrument, the logic here is that resources are scarce and supervisors economizing on resources generate predictable variation in hours that is exogenous to bank risk. But now the identification relies on variation over time in the amount of supervision at smaller banks: when a "low exam frequency" bank has just been examined in year $t-1$ it is less likely to receive resources again in year t and vice versa.[20] We therefore use lagged exam count and its interaction with a dummy for "high exam frequency" banks as shifters w_{idt} for log hours in equation (12). In addition to the usual controls such as size and the current year's rating, all specifications including the lagged exam count also condition on the lagged supervisory rating, which is up to date as of the last of the previous year's exams. The exclusion restriction for this instrument is that, conditional on the rating as of the end of the prior year, the number of examinations in the prior year does not contain information about the shock η_{idt} observed by the supervisor in the current year. While bank risk is persistent over time, η_{idt} is an information residual after controlling for

banks from the assets$\geq$\$10b dummy.

[20] A similar logic of using exam frequency as natural experiments has been used in two other contributions. Rezende and Wu (2014) apply the logic to legally mandated exam frequency for U.S. commercial banks. Similarly, Passalacqua et al. (2020) use the fact that exams of smaller Italian banks are randomized (conditional on some observables) to save on resources.

contemporaneous and lagged supervisory ratings. Thus, conditional on these controls, the number of examinations in the year prior should not contain information about the information residual η_{idt} in the current period.

Column (2) of Table 4 shows that an additional exam in the year prior predicts 63% lower hours at banks that are neither stressed, nor large or complex (p-val < 0.01). The interacted coefficients of lagged exam count and the dummies for large, complex, and stressed banks have very similar magnitudes ranging from 68% to 88% (p-vals < 0.01) and roughly offset the coefficient on lagged exam count. In other words, for these "high exam frequency" banks, the lagged number of examinations is not negatively correlated with current hours.

In column (3), we pool all exam count interaction dummies from column (2) in a single "high exam frequency" dummy. We consistently find that "low exam frequency" banks have 63% lower hours for an additional exam in the year prior (p-val < 0.01). The more parsimonious specification that uses a single dummy for high exam frequency increases the F-statistic to 35.6 from 17.9 in column (2).

The last column of Table 4 combines both preference shifters. Consistent with the fact that, by construction, these preference shifters exploit different variation in the data, the coefficient point estimates are essentially unchanged and remain highly statistically significant (p-vals < 0.01) with a joint F-statistic of 45.7.

6 Estimation results

We estimate the coefficients of the IV probit in equations (10)–(12) and obtain estimates of the underlying structural parameters measuring the technology of supervision, which is given by the effect of supervision γ and economies of scale parameter, α, and systematic supervisory preference loadings on size $\widetilde{\alpha}$ and risk $\widetilde{\rho}$. We then study the evolution of the shadow cost of supervisory resources and provide some intuition on the mechanism by which supervision affects bank distress based on the response to supervision of banks' regulatory ratios, asset quality, and major categories of income and expense. Finally, we present robustness tests of the main specification.

Tables 5 and 6 present second-stage estimates for the three bank distress outcomes: severe stress, failure, and low ROA. Outcomes are measured in the year after supervisory hours are recorded. Each table reports coefficient estimates, standard errors clustered by bank in brackets, and estimated average marginal effects in curly braces. In the previous section, we presented estimates of the first-stage regression, which are the same, up to sample differences, to first stages estimates in the IV probits. For brevity, we do not revisit those estimates and report results in Appendix Tables A5, A6, and A7. Each second-stage table in the main text reports effective F-statistics from the first stages and critical values for the weak instrument test of Olea and Pflueger (2013), which is robust to heteroskedasticity, autocorrelation, and clustering. Depending on the specification, the critical values range between 11 and 23.[21]

The different columns in the tables show parameter estimates using different instrumental variables. We first instrument log supervisory hours with variation in the shadow cost of resources using the leave-out district–year averages of hours and bank characteristics. We then instrument log hours with preference shifters: the "district top" indicator as well as the "exam frequency" instrument.[22] We combine all instruments in column (4) of Table 5 for outcome severe stress and columns (3) and (8) of Table 6 for outcomes failure and low ROA, respectively. To understand the importance of the IV approach, we also report probit estimates without instrumenting (column 5 of Table 5 and columns 5 and 9 of Table 6). In addition to log hours, the second stage includes log assets and rating indicators. As additional controls, the specifications include year fixed effects, the assets$\geq$\$10b dummy, and its interaction with the post-2008 dummy; for the exam frequency instrument based on interaction terms, the uninteracted high exam frequency dummy replaces the assets$\geq$\$10b dummy and lagged rating indicators are also included (noted at the bottom of the table, coefficients omitted).

6.1 Effect of supervision

The effect of supervision on the probability of bank distress is measured by the coefficient on log hours in Table 5, which represents $-\gamma$, the loading of the distress variable y^*_{idt} on the intensity of supervision s_{idt}. The marginal effect of supervision on the probability of distress is then equal to the product of $-\gamma$ and the density function, $-\gamma\phi(\cdot)$. We first discuss estimates of $-\gamma$ and then of the average marginal effect; finally, we present estimates of the marginal effects evaluated at different ratings that show non-linear effects of supervision on banks with different risk.

We find statistically and economically significant effects of supervision when using the probability of severe stress as the outcome (Table 5, first rows of columns 1 to 4). All p-values are smaller than 5% and, in terms of economic significance, as shown by the estimates of the average marginal effect in curly braces, an increase in supervisory hours of

[21] We use the critical value for the commonly used 5% significance level for the test that approximate asymptotic bias does not exceed 10%.

[22] When bank failure is the outcome variable (Table 6), we only consider the exam frequency preference shifter, because only one "district top" bank fails in the sample.

100% lowers the future probability of distress by 2.3 percentage points on average when including all instruments (column 4). Other columns show similar magnitudes, with "district top" showing the largest, though still broadly similar, effect (–5.9 percentage points). These estimated average marginal effects suggest large economic effects of supervision. For example, doubling supervisory hours implies a reduction in the probability of severe stress by about half the unconditional probability of 5.3% (Table 1). In contrast, the un-instrumented probit regression in column (5) shows a coefficient much smaller in absolute magnitude, not significantly different from zero. This is consistent with attenuation bias due to an omitted variable as posited in our econometric model. Specifically, the choice of the supervisory hours, H_{idt}, depends on news about the future distress, η_{idt}, which is observed by the supervisor but not the econometrician.

Table 6 presents estimates for the year-ahead probability of failure (columns 1–4) and the year-ahead probability of low ROA (columns 5–9). For failure, we only consider the "exam frequency" preference shifter because only one "district top" bank fails in the sample. We find the expected negative (although insignificant) effect of log hours on failure probability when using the shadow cost instrument (column 1) and a strong negative impact (p-val < 0.01) when using exam frequency (column 2) or when including both instruments (column 3). A doubling of supervisory hours reduces the probability of failure by about 1 percentage point, twice the unconditional probability of 0.5%. For the probability of a low ROA, we find a significant effect of supervisory hours for the shadow cost instrument (column 5) and the "district top" instrument (column 6). When combining all instruments, the estimated marginal effect of log hours is –2 percentage points (column 8). Again, the un-instrumented probit regressions in columns (4) and (9) show a coefficient much smaller in absolute magnitude and less, or even not, significant, consistent with attenuation bias due to the endogeneity of H_{idt}.

Hirtle et al. (2020) find that supervision lowers risk at "district top" banks compared to a set of similar banks matched via propensity scores. By relying on two additional instruments, our findings corroborate theirs and, in addition, shows similar effects of supervision on bank distress by making use of alternative sources of identification. In fact, within each outcome variable, the estimated effect of supervision is quite similar across instruments and the differences in the estimated effects are never statistically significant at conventional levels. This is in spite of the fact that the three instruments exploit variation for different sets of banks. While identification from the "district top" instrument involves large institutions, the "exam frequency" instrument largely exploits variation for small banks, and the "shadow cost" instrument uses variation in the allocation of resources at banks of all size.

Effect conditional on risk. The probability of future distress strongly depends on current supervisory risk assessments. Relative to a rating of 1, a rating of 3 increases the probability of severe stress by 18 percentage points (Table 5, column 4). Due to the non-linearity of the probit specification, the marginal effect of supervision at a particular rating can differ from the average marginal effect. Figure 3 shows marginal effects conditional on the five rating categories compared to the unconditional marginal effect for each of the three outcome variables. The marginal effect of supervision varies considerably across banks of different riskiness. For outcome severe stress, the effect increases (in absolute value) from about 0 at a bank rated 1 to about 8 percentage points at a bank rated 3 or worse.

Mechanism of effect. To provide intuition on how increased supervision lowers bank distress, we estimate IV probits for low (high) realizations of the major income (expense) sub-categories of ROA (scaled by assets), and of the year-ahead realization of a high non-performing loans ratio and a low tier-1 capital ratio (Appendix Table A8). For each measure, a low (high) realization is defined as in the 10th (90th) percentile. We estimate the impact of log hours on each dependent variable using all instruments. With respect to the ROA subcategories, the lower probability of a low ROA realization (column 1 replicates column 8 of Table 6) is driven by lower probabilities of low non-interest income, high loan loss provisioning, low net interest income, and low realized gains on securities.[23] We also see that higher supervision increases the likelihood of high non-interest expenses, consistent with higher compliance costs. As shown in column 7, increased supervision has the expected effect of lowering the likelihood of a high nonperforming loan ratio, although the effect is not significant at conventional levels. Overall, the results suggest that banks de-risk when supervision increases: lower probabilities of high nonperforming loan ratio and high loan loss provisioning point to safer loan portfolios; a lower probability of low realized gains on securities indicates less-risky security holdings; and a lower probability of low non-interest income suggests more conservative positions in cash and derivative instruments. In unreported results, we also use averages, rather than tail events, for each outcome variable and find either smaller or insignificant effects. These results suggest that supervision has a greater effect on the tails of the distributions of bank performance and risk than it has on the averages.

[23] The income-to-asset categories are: non-interest income, interest income, and realized gains on securities not held to maturity. The expense-to-asset categories we consider are non-interest expense and loan loss provisioning.

6.2 Economies of scale

The hours cost function in (4) has a size elasticity of α, meaning that to achieve the distress probability at a bank with double the assets requires 2α hours, and economies of scale in supervision exist for $\alpha < 1$. To measure α, we can divide the coefficient β_A on log assets in the second stage (11), which is an estimate of $\alpha\gamma$, by the coefficient β_H on log hours, which is an estimate of $-\gamma$. This implies point estimates for the size elasticity of hours cost α of 0.52 for outcome severe stress, 0.60 for failure, and 0.76 for low ROA, indicative of large scale economies (significantly less than 1 for severe stress and failure, p-val < 0.01, as obtained via the delta method).

Our finding of economies of scale in supervision is also valid in a more general model where size directly affects bank risk (see Appendix A.1 for details on how our evidence covers this more general case). Ultimately, scale economies mean that larger banks can be supervised with proportionately fewer resources, and this can be due to a combination of larger banks being inherently less risky and supervisors being more effective at larger banks.

In other words, economies of scale in supervision imply that supervising one large bank requires less resources than supervising two banks whose size adds up to the same total. This can inform the post-2008 debate about constraining the size of large banks or actually splitting them up (Johnson and Kwak (2010), Bair (2012), Barth, Caprio, and Levine (2012), Admati and Hellwig (2013)). The debate has mostly been about the benefits in terms of systemic risk vs. the potential costs in terms of bank efficiency. From a technological perspective, our findings show that breaking up banks would significantly increase risk absent a significant boost in supervisory resources.

6.3 Supervisory preferences

We study how the preference weight W_{idt} loads on log assets and ratings by combining coefficient estimates of our first and second stages. The second stage coefficients only depend on the technological parameters. In contrast, the first stage is derived from the first order condition, which determines the optimal allocation as a function of both supervisory technology and preferences. Proposition 2 shows that the coefficients in the first stage (12) are estimates of convex combinations of the technological and preference parameters, i.e. the loadings of the distress variable y^*_{idt} and the preference weight W_{idt} on log assets and on ratings. For any variable x_{idt} entering the distress variable with loading κ and the supervisory preference weight with loading $\widetilde{\kappa}$, the second-stage coefficient β_x is an estimate of κ while the first-stage coefficient δ_x is an estimate of

$$\widehat{\delta}_x \mathrel{\widehat{=}} \overline{\pi}\,\frac{\kappa}{\gamma} + (1-\overline{\pi})\,\widetilde{\kappa}. \tag{14}$$

To obtain point estimates of $\widetilde{\alpha}$ and $\widetilde{\rho}_2, \ldots, \widetilde{\rho}_2$, we construct the sample analog of the weight $\overline{\pi}$ from Proposition 2, $\overline{\pi} = E[\pi_{idt}]$, using fitted values $\widehat{\mathrm{PD}}_{idt}$ and estimates of γ and σ_ε.[24]

After accounting for the inclusion of several size-related dummy variables in the first stage, which lower the first-stage coefficient on log assets, we obtain an estimate for $\widetilde{\alpha}$ of 1.18 for outcome severe stress, 1.4 for failure, and 1.00 for low ROA, with the first two significantly greater than 1 (p-val < 0.05).[25] These estimates suggest that the preference weight increases at least proportionally with bank size as would be expected if supervisors perceive larger banks to have disproportionately larger distress costs. Thus bank size affects the supervisory technology and preferences in opposite ways. While economies of scale in the supervisory technology ($\alpha < 1$) suggest that it takes fewer resources to supervise larger banks, supervisors weight larger banks more than smaller banks that sum to the same size ($\widetilde{\alpha} > 1$), consistent with greater, macro-prudential concerns about larger institutions.

Based on our model, non-zero loadings of the preference weight on risk ($\widetilde{\rho}_r \neq 0$) mean that supervisors use a stochastic discount factor when calculating the expectation in their objective function so that their subjective distress probabilities differ from the objective ones. Figure 4 shows the estimates for the preference weight loadings on rating, $\widetilde{\rho}_2, \ldots \widetilde{\rho}_5$, relative to rating 1. The significantly negative values (p-vals < 0.01) imply that supervisors weight the distress probability of the higher-risk banks rated $2, \ldots, 5$ *less* than the distress probability of a low-risk bank rated 1. As with bank size, because the hours allocation depends both on bank riskiness (loadings ρ_r) and on the preference weights (loadings $\widetilde{\rho}_r$), weighting safer banks more than riskier banks does not necessarily imply that safer banks receive more attention in absolute terms. Indeed, under equal-weighting, riskier banks would receive more supervisory hours because the marginal effect of supervision on their distress is higher (Proposition 1). Given that, empirically, riskier banks receive considerably more hours (Table 2), the inverse preference weighting in Figure 4 only attenuates the cost-benefit effect.

[24] We obtain an estimate of $\sigma_\varepsilon = \sqrt{\mathrm{var}(\varepsilon_{idt+1})}$ from the IV probit estimates of $\mathrm{var}(v_{idt})$ and $\mathrm{corr}(u_{idt+1}, v_{idt})$, using the fact that $v_{idt} = \delta_\eta \eta_{idt}$ and $u_{idt+1} = \eta_{idt} + \varepsilon_{idt+1}$, and the normalization $\mathrm{var}(u_{idt+1}) = 1$. In Appendix A.6, we derive bounds on the preference parameters that do not depend on $\overline{\pi}$ and yield consistent results.

[25] We use as the estimate of α the size elasticity of hours, controlling only for ratings, from Table 2, column 1. We calculate p-values and construct confidence intervals for the preference parameters based on bootstrapped samples clustered at the bank level (20,000 replications).

It is somewhat surprising that supervisors weight less the banks that are already in relatively bad condition because, as shown by the second stage coefficients on ratings, worse rated banks are much more likely to fail. One way the higher subjective distress probability for safer banks could arise is from supervisors overweighting tail risk, as in probability weighting, where very small probabilities receive disproportionate weight (Kahneman and Tversky, 1979). Alternatively, the evidence could stem from a subjective supervisory assessment that distress probabilities of riskier banks are lower than estimated from historical data (Gennaioli and Shleifer, 2018). To quantify how much supervisory preferences tilt attention toward safer banks, in Section 7, we compare outcomes under the actual allocation to those under two extreme counterfactuals in which the allocation either does not depend on bank risk at all, or when no preference tilt exists.

6.4 Shadow cost of supervisory resources

The resource allocation model can be used to estimate the shadow cost of resources at the district–year level, allowing us to study how resource scarcity varies across districts and over time.[26] Decomposing the variance of the shadow cost, we find that 38% of the variance is across districts while 63% is within district (and a slight negative covariance). We discuss the economic implications of variation in the shadow cost next.

The large degree of cross-sectional variation in the shadow cost suggests that there are persistent differences in resources available across district and thus potential gains from a centralized resource allocation instead of district level allocation. These gains could in principle be offset by other losses if, for example, the Federal Reserve assigns different weights to different districts.[27] From a measurement perspective it is nonetheless important to establish if these resource asymmetries exist in the first place, and we explore the quantitative significance of shifting the resource constraint to the national level in Section 7. Similar considerations apply also internationally, for example, with the establishment of the European's Single Supervisory Mechanism in 2012, which our framework could be directly applied to as well.

The within-district variation owes to the fact that Federal Reserve supervisory staff increased only gradually when the financial crisis hit (Figure 1). In fact, the financial crisis increased the need for resources for two reasons. First, new institutions came under Federal Reserve supervision (prominent examples are Goldman Sachs and Morgan Stan-

[26] We estimate the shadow cost at the district–year level using a specification matching our baseline first-stage regression but including district–year fixed effects instead of year fixed effects and the district averages of the shadow-cost instrument.

[27] Conceptually, unequal shadow costs across districts imply benefits only under an objective function that equally weights each district's objective function.

ley). Figure 5 (left panel) shows total supervisory hours starting in 2006, split by whether resources are spent at existing or new banks. Consistent with the headcount numbers in Figure 1, total supervisory hours increased gradually by about 50% following the financial crisis, but new banks accounted for a significant fraction of the increase in total hours (blue area). Second, the risk profile of the supervised banks changed. For example, the Federal Reserve's inaugural Supervision and Regulation Report (2018, Figures 14 and 19) shows that ratings for the bank universe started deteriorating in 2007, and by 2010 almost a third of banks had unsatisfactory ratings of 3 or worse, including about half of the largest banks (over $50 billion).

We can use our estimates to trace out changes in the shadow cost over time to compare the scale and speed of the expansion in the supply of resources to the increase in the demand for resources due to the greater size and riskiness of supervised banks. The right panel of Figure 5 shows the estimated shadow cost of supervisory resources averaged across districts. Resource scarcity increases by over 50% by 2009 and stays at that elevated level through 2011. Only starting in 2012 do resources begin to catch up such that (with improving ratings), the shadow cost returns roughly to its 2006 level by 2014. This means that Federal Reserve supervisors faced tight resource constraints throughout the crisis and post-crisis period and it took until 2014 for the expansion to resolve the resource crunch. In Section 7, we assess the effects of an earlier and faster expansion of resources.

6.5 Robustness checks

The baseline results in Tables 5 and 6 present estimated effects of supervisory hours on bank outcomes using three outcome variables and four different instruments, both independently and jointly. In Appendix D, we consider four additional robustness exercises: (i) including district fixed effects and an additional set of risk controls, (ii) excluding the largest banks with assets greater than $50 billion, (iii) excluding the crisis years 2008–2009, and (iv) running the estimation in a linear probability instrumental variable setting. We find that baseline IV probit effects are robust to each of these robustness checks.

7 Effect of resources and preferences on outcomes

So far, we used data on supervisory hours, bank characteristics, and outcomes to learn about the parameters of the supervisory technology (α and γ) and the loadings of supervisory preferences ($\widetilde{\alpha}$ and $\widetilde{\rho}$). In this section, we use counterfactual experiments to quantify how resource availability and supervisory objectives affect the distribution of outcomes across banks. Because resources are allotted contemporaneously across many banks, the counterfactual analysis not only relies on the estimated coefficients but also on the resource constraint, which the estimation did not require. Each experiment takes the supervisory technology as given and alters the resource allocation to derive implications for bank outcomes. With respect to resource availability, we first study the role of scarce resources in 2007–09, by positing that resources are boosted to their post-crisis peak already as of 2007. In other words, this experiment posits that the Federal Reserve enters the financial crisis "fully staffed." We then study the effect of a centralization of supervisory resources that equalizes the shadow cost across districts. We next turn to three experiments that undo key features of the supervisory weights: the post-2008 shift toward large banks, the disproportionate focus on "district top" banks, and the tilt toward safer banks.

We alter the allocation of supervisory resources and then trace out the implications on distress probabilities. Specifically, we construct a counterfactual hours allocation $\widehat{\log H^*_{idt}}$ that differs from the one predicted by the first stage (12). We then use the second stage (11) to predict a counterfactual distress probability, $\widehat{\mathrm{PD}}\left(\widehat{\log H^*_{idt}}, A_{idt}, r_{idt}\right)$, and the resulting change in distress probability:[28]

$$\widehat{\mathrm{PD}}\left(\widehat{\log H^*_{idt}}, A_{idt}, r_{idt}\right) - \widehat{\mathrm{PD}}\left(\widehat{\log H_{idt}}, A_{idt}, r_{idt}\right)$$

We impose the resource constraint on total hours under the counterfactual allocation and construct lump-sum transfers such that overall resources are kept constant at the level of the respective counterfactual experiment.

As noted by the formula above, the impact on the distress probability is constructed at the level of each bank–year observation in the sample. We summarize these results along key bank characteristics, such as size and risk, as well as in the aggregate.[29] Translating these averages into welfare effects requires taking a stance on a social welfare function which is a point of disagreement for long-established literature in public economics, such as tax policy (see, e.g., Chetty (2009) for a review of the debate). Rather than imposing a specific welfare function, we report effects using equal- and size-weighted averages (panels (b)–(g) of Table 7; panel (a) shows baseline distress probabilities). While we do not suggest that either of these is the correct welfare function, the averages have intuitive properties. The equal-weighted average can be interpreted as a purely micro-prudential objective in which all banks are treated the same. The size-weighted average is closer to

[28]We use predicted hours $\widehat{\log H_{idt}}$ as a baseline — instead of actual hours $\log H_{idt}$ — to not conflate the effects of the prediction error with the effects of the counterfactual.

[29]We calculate p-values and construct confidence intervals for the reported statistics based on bootstrapped samples clustered at the bank level (20,000 replications).

a more macro-prudential objective since larger banks have disproportionately larger distress spillovers (Hanson, Kashyap, and Stein, 2011). From the difference between the equal- and size-weighted averages, one can extrapolate (or interpolate) to consider a welfare function with greater than (or less than) unit size elasticity.

The main findings of the counterfactual analysis are that average distress probabilities are higher as a result of the decentralized resource allocation at the district level, both because the availability of resources across districts differs and due to the disproportionate attention being paid to "district top" banks. The tilt in supervisory preferences toward safer banks and toward larger banks after 2008 similarly results in higher bank distress on average. Judging the efficiency of these outcomes ultimately depends on whether the average distress probabilities that we measure are a good proxy for the social welfare function, which is unknown. We thus see our analysis as a positive, as opposed to a normative, one. Lastly, we find that the significant scarcity of supervisory resources between 2007 and 2012 increased average distress probability about 10% relative to if resources had been as ample as prior to the financial crisis.

7.1 Counterfactual resource constraints

Earlier expansion of resources in the 2007–09 crisis. A large literature studies how bank supervision and regulation may have contributed factors to the 2007–09 financial crisis. Duffie (2019) sees the financial system as "prone to fail" due to a combination of weak regulation and supervision; others similarly stress that regulators placed too much faith in market discipline, which was distorted by expectations that some institutions were "too big to fail" (Admati and Hellwig, 2013), or that they underestimated the probability of a severe shock (Gennaioli and Shleifer, 2018). The role of supervisory resources has not been previously discussed, with the exception of Duffie (2019), who notes significant differences in staffing at the Securities and Exchange Commission and the Federal Reserve. The right panel of Figure 5 shows that the cross-district average shadow cost of resources increases by over 50% from 2006 to 2009 and only starts decreasing in 2012. Even though supervisory resources increase post-2008 (Figure 1, left panel), the slow pace of the expansion and concurrent increase in the size and riskiness of supervised banks left supervisors facing a resource crunch during and after the financial crisis. We therefore consider a counterfactual where district-level resources increase already in 2007 all the way to their 2014 level. This implies resources at bank i in district d and year t are given by

$$\widehat{\log H^*_{idt}} = \widehat{\log H_{idt}} + \hat{\tau}^{\text{early}}_{dt} \quad \text{for} \quad t \geq 2007,$$

where $\widehat{\tau}_{dt}^{\text{early}}$ is a lump-sum percentage increase, constant across banks in district d and year t, such that total hours in district d are at their 2014 level for all years $t = 2007, \ldots, 2014$ (see Appendix A.7 for details on the calculation of τ).

Table 7, panel (b) shows that this earlier expansion would have lowered distress probabilities quite significantly, both statistically and economically. Comparing the change in distress probabilities across columns to the baseline levels in panel (a) shows that an earlier resource expansion lowers the average distress probability across the treated years 2007–2013 by about 10% of its baseline. Figure 6a shows that, year by year, effects are largest in 2007 and 2008, reducing the probability of severe distress, failure, and low ROA, respectively, by about 2.5, 0.75, and 3 percentage points in absolute terms. In sum, these estimates suggest that the scarcity of supervisory resources during the financial crisis limited the effect of supervision and that an earlier and faster expansion would have had quantitatively meaningful effects on bank distress.

Centralized allocation. As discussed in Section 5.1, the decentralized resource allocation at the district level results in unequal shadow costs across districts. This implies that a centralized allocation at the national level that equally weights each district's objective function would reallocate resources across districts. We assess the quantitative importance of the decentralized allocation with a counterfactual that equalizes the shadow cost across districts within each year via a centralized allocation. We construct the counterfactual allocation by offsetting the estimated effect $\widehat{\lambda}_{dt}$ of the district-specific Lagrange multiplier (which enters with a negative sign) and then applying a lump-sum transfer $\widehat{\tau}_t^{\text{central}}$ to keep total hours (across all districts) constant each year:

$$\widehat{\log H_{idt}^*} = \widehat{\log H_{idt}} + \widehat{\lambda}_{dt} + \widehat{\tau}_t^{\text{central}} \tag{15}$$

Compared to the realized allocation, this counterfactual reallocates hours from districts with low shadow costs to districts with high shadow costs.

Averaging across districts, Table 7, panel (c) shows that this reallocation reduces distress probabilities on net with effects that are considerably smaller than those seen for the earlier-expansion counterfactual. Intuitively, a centralized allocation only implies a redistribution of resources as opposed to an actual expansion. However, we see consistently larger effects for riskier than for larger banks and for size-weighted averages (even columns) than for equal-weighted averages (odd columns), suggesting that districts with larger banks have systematically fewer resources. Compared to these averages across districts, Figure 6b shows that the effects are orders of magnitude larger at the districts receiving and losing the most resources.

7.2 Counterfactual supervisory objectives

No reallocation to large banks. Controlling for bank size and risk, point estimates in column (3) of Table 2 imply that hours at small banks (< \$10 billion) dropped roughly 20% post-2008 while hours at large banks increased roughly 90%. In this shift toward a more macroprudential objective (Hanson et al., 2011), supervisors appear willing to tolerate additional risk at smaller banks for the reduced risk at larger banks. We quantify the effect of this policy on bank distress by assuming that resources are not reallocated to large banks in the post-2008 period,

$$\widehat{\log H^*_{idt}} = \widehat{\log H_{idt}} - \widehat{\delta}_{\text{post-large}} \mathbb{I}[t > 2008]\,\mathbb{I}[A_{idt} > \$10\text{b}] + \widehat{\tau}^{\text{noreall}}_{dt},$$

where $\widehat{\tau}^{\text{noreall}}_{dt}$ is a transfer to keep total hours constant at the district–year level. Compared to the actual allocation, this counterfactual first removes the additional hours large banks receive post-2008. This creates slack in the budget constraint, implying a drop in the shadow cost λ_{dt}. To make use of the slack, the counterfactual then increases all banks' hours by the change in the shadow cost, represented by $\widehat{\tau}^{\text{noreall}}_{dt}$.

Table 7, panel (d) and Figure 6c show that undoing this reallocation has the expected effect of increasing the distress probability at large banks while decreasing it at small banks. But the quantitative analysis allows us to tell that the magnitude of these effects are much bigger at smaller than at large banks. Across all banks, the distress probability declines but the effect is much larger with equal weighting. In turn, this implies that as supervisory weights shifted toward macroprudential objectives, overall distress probabilities increased.

No disproportionate supervision of district top banks. As first documented by Hirtle et al. (2020), the top banks within each district receive disproportionate attention when compared to similar-sized banks in other districts that are not "top". To quantify this tilt of district supervisors toward their largest banks, we consider a counterfactual without the "district top" effect,

$$\widehat{\log H^*_{idt}} = \widehat{\log H_{idt}} - \widehat{\delta}_{\text{top}} \mathbb{I}[i \in \text{top}_{dt}] + \widehat{\tau}^{\text{notop}}_{dt},$$

where $\widehat{\tau}^{\text{notop}}_{dt}$ is a lump-sum transfer to keeps total hours constant at the district–year level. While Table 7, panel (e) shows that the reallocation away from the largest banks in each district leaves the size-weighted average distress probability almost unchanged, the reallocation has quite large effects on small banks (< \$10 billion) and on risky banks (rating $\geq$ 3). Even among large banks ($\geq$ \$10 billion), sufficiently many receive additional hours to decrease their equal-weighted average distress probability.

No/full response to bank risk. As noted in Section 4, supervision responds strongly to bank risk, with poorly rated banks receiving multiple times the hours of a bank with the best possible rating of 1. But the estimated supervisory preference weights $\widetilde{\rho}$ are larger for safer banks, meaning that supervisory preferences attenuate the response of the allocation to risk. To quantify this attenuation in terms of outcomes, we consider two extreme counterfactuals. In the first, resources do not respond to risk at all, while in the second, they respond fully.

When the allocation does not respond to risk at all, i.e. an extreme case of attenuation, counterfactual hours are

$$\widehat{\log H^{*}_{idt}} = \widehat{\log H_{idt}} - \sum_{r=2}^{5} \widehat{\delta}_r \mathbb{I}[r_{idt} = r] + \widehat{\tau}^{\text{noresp}}_{dt}, \tag{16}$$

where $\widehat{\tau}^{\text{noresp}}_{dt}$ is a transfer to keep total hours constant at the district–year level. When, instead, the supervisory preferences weight all ratings equally, i.e. there is no attenuation at all ($\widetilde{\rho}_r = 0$), we have a counterfactual response of log hours to rating r given by $\delta_r = \overline{\pi}\,\frac{\rho_r}{\gamma}$. Counterfactual hours are then given by

$$\widehat{\log H^{*}_{idt}} = \widehat{\log H_{idt}} - \sum_{r=2}^{5} \widehat{\delta}_r \mathbb{I}[r_{idt} = r] + \sum_{r=2}^{5} \overline{\pi}\,\frac{\widehat{\rho}_r}{\widehat{\gamma}}\,\mathbb{I}[r_{idt} = r] + \widehat{\tau}^{\text{fullresp}}_{dt}, \tag{17}$$

where $\widehat{\tau}^{\text{fullresp}}_{dt}$ is a transfer to keep total hours constant at the district–year level.

The "no response" and "full response" counterfactuals reallocate hours away from safer and toward riskier banks, respectively, and Table 7, panels (f) and (g) show that the results have the expected opposite signs. In terms of absolute magnitude, the effects are quite similar, with the "no response" effects slightly larger and more significant. For example, relative to the actual allocation, "no response" increases the average probability of severe stress by 38.8 basis points, and "full response" reduces it by 17.2 basis points. This implies that the higher supervisory preference weights for safer banks attenuate the response about 30% of the way from the "full response" allocation toward the "no response" allocation. In sum, while the empirical sensitivity of supervisory attention with respect to bank risk appears large, it is actually attenuated considerably through the effect of supervisory preferences on the allocation.

8 Conclusion and overall supervisory resources

Our structural model provides new insights on how the technology of bank supervision, latent supervisory preferences, and resource scarcity shape bank outcomes. We find that supervision has an economically large effect in lowering bank distress. Parameter estimates also suggest that larger banks are easier to supervise thanks to economies of scale in the supervisory technology, but that supervisory efforts scale up almost proportionally with bank size in the data because supervisors care disproportionately about distress at the largest banks. Despite these macro-prudential concerns, the technological estimates suggest that banks have not become too big to supervise. In contrast, our estimates reveal that supervisors place less weight on riskier banks even though hours allocated to riskier banks (rated 4 or 5) are more than three times greater than hours allocated to safe banks (rated 1). We can interpret this as evidence that supervisors overweight tail risk in the form of very small probabilities.

Parameter estimates and counterfactual experiments suggest that the shadow cost of supervisory resources is not equalized across Federal Reserve districts, meaning that resource are not allocated evenly across regions. In addition, shadow cost estimates indicate that the very large increase in resources post-2008 was too gradual to keep up with the sudden increases in the size and composition of banks under Fed supervision. We also find that the post-2008 reallocation of resources lowered risk at large banks but increased it at smaller banks and, on net, across the universe of banks.

The model is used to clarify the identification approach, to characterize the resource allocation across a portfolio of banks, and to uncover deep, and previously unknown, parameters related to supervisory preferences and technology. Yet the model does not speak to other central issues in supervision, such as why supervision is necessary in the first place and what is its optimal level. While the model takes the overall level of resources as given, we can consider a simple back-of-the-envelope calculation of increasing the overall supervisory budget by 1% and comparing the marginal expected benefit to the marginal cost. According to the Federal Reserve's 2017 annual report, total operating expenses for supervision and regulation were \$1.6 billion. Thus, increasing the budget by 1% would cost \$16 million. Assuming that total supervisory hours grow at the same rate as budgeted costs, and abstracting from estimation uncertainty, our estimates indicate that the resulting 1% increase in supervisory hours would lower the probability of failure by 0.012 percentage points on average. With bankruptcy cost estimates in the literature of about 12% and total bank holding company assets under Federal Reserve supervision of \$19 trillion in 2017, this calculation implies a reduction in expected bankruptcy costs of roughly

$270 million.

The back-of-the-envelope calculation suggests that the marginal benefit of $270 million outweighs the marginal cost of $16 million. But this calculation omits increases in resources at other agencies that Federal Reserve supervisors rely upon. For example, the FDIC had 2017 operating expenditures of $1.9 billion, and total expenses in 2017 at the OCC were $1.2 billion. Including 1% increases in the budgets of both agencies increases the marginal cost to $47 million, which is still only a fraction of the marginal benefit. Thus, the simple back-of-the-envelope calculation suggests that there may be net benefits to increasing overall supervisory resources.

References

Admati, A. and M. Hellwig (2013). *Bankers' New Clothes: What's Wrong with Banking and What to Do About It*. Princeton University Press.

Agarwal, S., D. Lucca, A. Seru, and F. Trebbi (2014). Inconsistent regulators: Evidence from banking. *Quarterly Journal of Economics 129*(2), 889–938.

Aguiar, M. and E. Hurst (2007). Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics 122*(3), 969–1006.

Aikman, D., J. Bridges, A. Kashyap, and C. Siegert (2019). Would macroprudential regulation have prevented the last crisis? *Journal of Economic Perspectives 33*(1), 107–30.

Altavilla, C., M. Boucinha, J.-L. Peydró, and F. Smets (2019). Banking supervision, monetary policy and risk-taking: Big data evidence from 15 credit registers. Working Paper.

Angrist, J. D., G. W. Imbens, and A. B. Krueger (1999). Jackknife instrumental variables estimation. *Journal of Applied Econometrics 14*(1), 57–67.

Bair, S. (2012). Why it's time to break up the 'too big to fail' banks. *Fortune* (February 6).

Barth, J., G. Caprio, and R. Levine (2012). *Guardians of Finance: Making Regulators Work for Us*. MIT Press.

Basel Committee on Banking Supervision (2010). *Calibrating Regulatory Minimum Capital Requirements and Capital Buffers: A Topdown Approach*.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal 75*(299), 493–517.

Bisetti, E. (2018). The value of regulators as monitors: Evidence from banking. Working Paper.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2018). Children, time allocation, and consumption insurance. *Journal of Political Economy 126*(S1), S73–S115.

Board of Governors of the Federal Reserve System (2015). *Calibrating the GSIB Surcharge*.

Board of Governors of the Federal Reserve System (2018). *Supervision and Regulation Report*.

Borio, C. (2011). Implementing the macro-prudential approach to financial regulation and supervision. In C. J. Green, E. J. Pentecost, and T. G. Weyman-Jones (Eds.), *The Financial Crisis and the Regulation of Finance*, pp. 101–117. Edward Elgar Cheltenham.

Carletti, E., G. Dell'Ariccia, and R. Marquez (2019). Supervisory incentives in a banking union. *Management Science* (forthcoming).

Carpenter, D. and D. A. Moss (2013). *Preventing regulatory capture: Special interest influence and how to limit it*. Cambridge University Press.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ. 1*(1), 451–488.

Corbae, D., P. D'Erasmo, S. Galaasen, A. Irarrazabal, and T. Siemsen (2018). Structural stress tests. Working Paper.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family welfare cultures. *Quarterly Journal of Economics 129*(4), 1711–1752.

Dewatripont, M. and J. Tirole (1994). *The Prudential Regulation of Banks*. MIT Press.

Dobbie, W., J. Goldin, and C. S. Yang (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review 108*(2), 201–40.

Duffie, D. (2019). Prone to fail: The pre-crisis financial system. *Journal of Economic Perspectives 33*(1), 81–106.

Egan, M., A. Hortacsu, and G. Matvos (2017). Deposit competition and financial fragility: Evidence from the US banking sector. *American Economic Review 107*(1), 169–216.

Eisenbach, T. M., A. Haughwout, B. Hirtle, A. Kovner, D. O. Lucca, and M. C. Plosser (2017). Supervising large, complex financial institutions: What do supervisors do? *Federal Reserve Bank of New York Economic Policy Review 23*(1), 57–77.

FDIC (1997). *History of the Eighties – Lessons for the Future*.

Flannery, M. J. (1998). Using market information in prudential bank supervision: A review of the U.S. empirical evidence. *Journal of Money, Credit and Banking 30*(3), 273–305.

Gennaioli, N. and A. Shleifer (2018). *A Crisis of Beliefs: Investor Psychology and Financial Fragility*. Princeton University Press.

Granja, J. and C. Leuz (2019). The death of a regulator: Strict supervision, bank lending and business activity. Working Paper.

Greenwood, R., J. C. Stein, S. G. Hanson, and A. Sunderam (2017). Strengthening and streamlining bank capital regulation. *Brookings Papers on Economic Activity 2017*(2), 479–565.

Halvorsen, R. and R. Palmquist (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review 70*(3), 474–475.

Hanson, S. G., A. K. Kashyap, and J. C. Stein (2011, March). A macroprudential approach to financial regulation. *Journal of Economic Perspectives 25*(1), 3–28.

Harris, M. and A. Raviv (2014). How to get banks to take less risk and disclose bad news. *Journal of Financial Intermediation 23*(4), 437 – 470.

Hirtle, B., A. Kovner, and M. Plosser (2020). The impact of supervision on bank performance. *Journal of Finance 75*(5), 2765–2808.

Hirtle, B. J. and J. A. Lopez (1999). Supervisory information and the frequency of bank examinations. Working Paper.

Jensen, M. C. and W. H. Meckling (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics 3*(4), 305–360.

Johnson, S. and J. Kwak (2010). *13 Bankers: The Wall Street Takeover and the Next Financial Meltdown*. Pantheon.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Kandrac, J. and B. Schlusche (2019). The effect of bank supervision and examination on risk taking: Evidence from a natural experiment. *Review of Financial Studies* (forthcoming).

Kisin, R. and A. Manela (2014). Funding and incentives of regulators: Evidence from banking. Working Paper.

Kroszner, R. S. and P. E. Strahan (1999). What drives deregulation? Economics and politics of the relaxation of bank branching restrictions. *Quarterly Journal of Economics 114*(4), 1437–1467.

Lucca, D., A. Seru, and F. Trebbi (2014). The revolving door and worker flows in banking regulation. *Journal of Monetary Economics 65*(0), 17–32.

Marshall, D. A. and E. S. Prescott (2001). Bank capital regulation with and without state-contingent penalties. *Carnegie-Rochester Conference Series on Public Policy 54*(1), 139–184.

Olea, J. L. M. and C. Pflueger (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Passalacqua, A., P. Angelini, F. Lotti, and G. Soggia (2020). The real effects of bank supervision. Working Paper.

Peek, J., E. S. Rosengren, and G. M. Tootell (1999). Is bank supervision central to central banking? *Quarterly Journal of Economics 114*(2), 629–653.

Radner, R. and M. Rothschild (1975). On the allocation of effort. *Journal of Economic Theory 10*(3), 358–376.

Repullo, R. and J. Suarez (2012). The procyclical effects of bank capital regulation. *Review of Financial Studies 26*(2), 452–490.

Rezende, M. and J. J. Wu (2014). The effects of supervision on bank performance: Evidence from discontinuous examination frequencies. Working Paper.

Rochet, J.-C. (2004). Macroeconomic shocks and banking supervision. *Journal of Financial Stability 1*(1), 93–110.

Rochet, J.-C. (2007). *Why Are there So Many Banking Crises? The Politics and Policy of Bank Regulation*. Princeton University Press.

Stigler, G. J. (1971). The theory of economic regulation. *Bell Journal of Economics and Management Science 2*(1), 3–21.

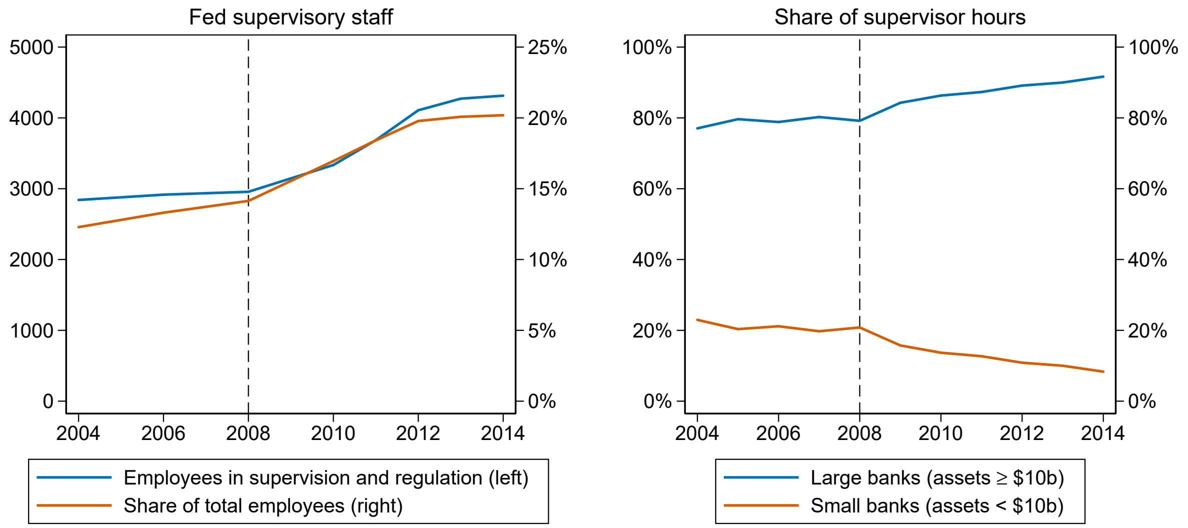


Figure 1: Federal Reserve supervisory staff and allocation of resources. The data on employees is from Federal Reserve Annual reports. The data on resources is from internal hours data for supervisory examiners at the Federal Reserve. The hours data in the right panel exclude resources allocated to institutions that were not under Federal Reserve supervision pre-2008.

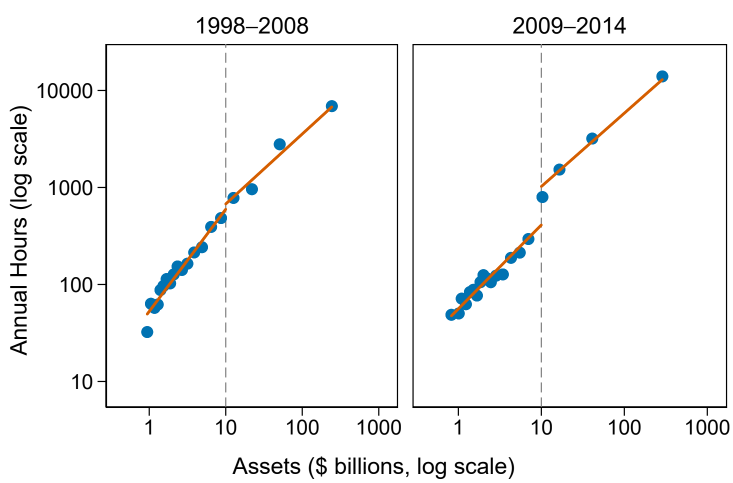


Figure 2: Increased attention to large banks. The figure presents binned scatter plots and the fitted lines of regressing log supervisory hours on log assets and supervisory ratings for different bank size categories ($10 billion asset thresholds) before and after 2008.

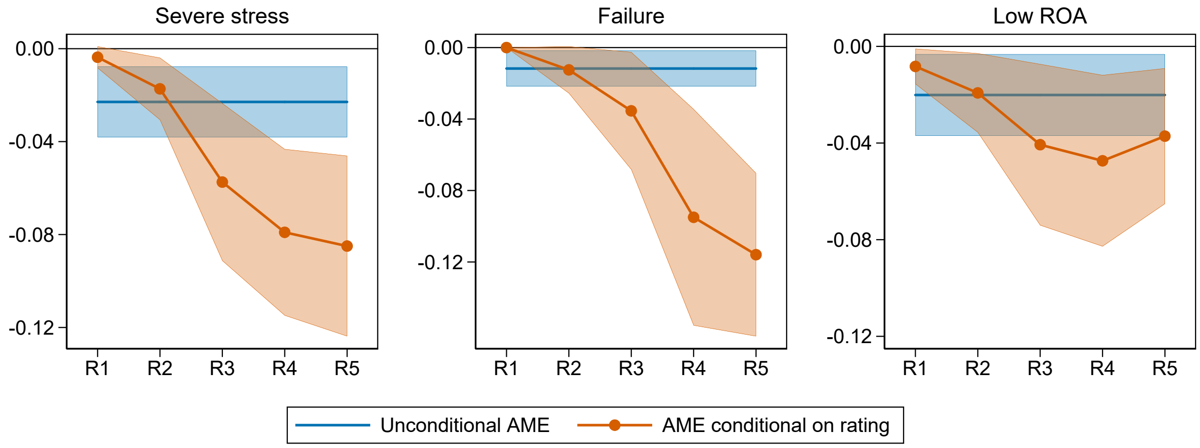


Figure 3: Marginal effect of log hours on probability of distress conditional on different ratings. The figure shows the average marginal effect (AME) of log hours on a bank's next-year probability of distress unconditionally (blue) and evaluated at each of the five supervisory ratings (orange). The three panels correspond to the three variables measuring distress. Shaded areas represent 95% confidence intervals computed based on standard errors clustered at the bank level via the delta method.

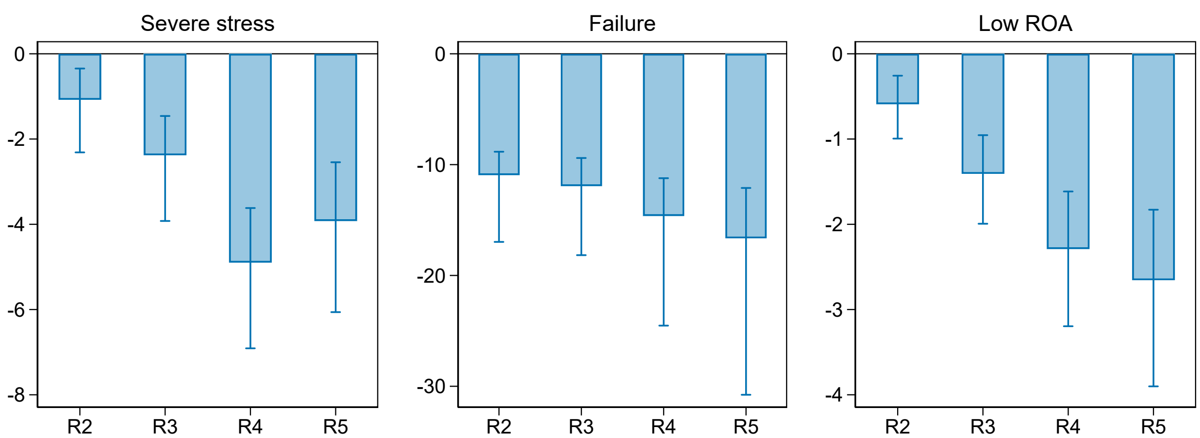


Figure 4: Implied supervisory preference loadings relative to rating 1. The figure shows the implied loadings of the supervisory preference weight W_{idt} on rating 2 to rating 5. The three panels correspond to the three variables measuring distress. Whiskers represent bootstrapped 95% confidence intervals clustered at the bank level (20,000 replications).

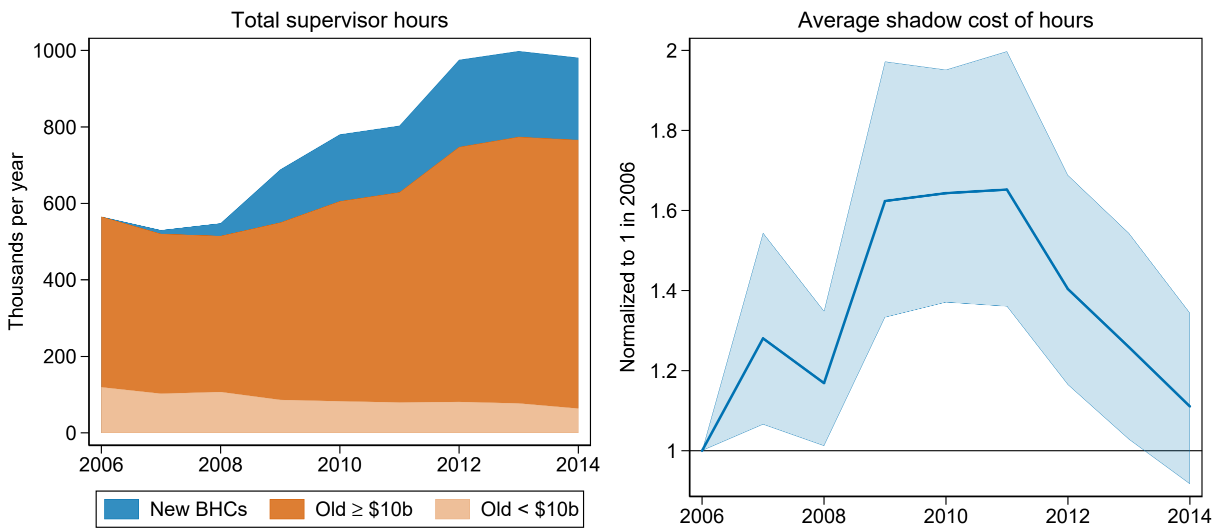
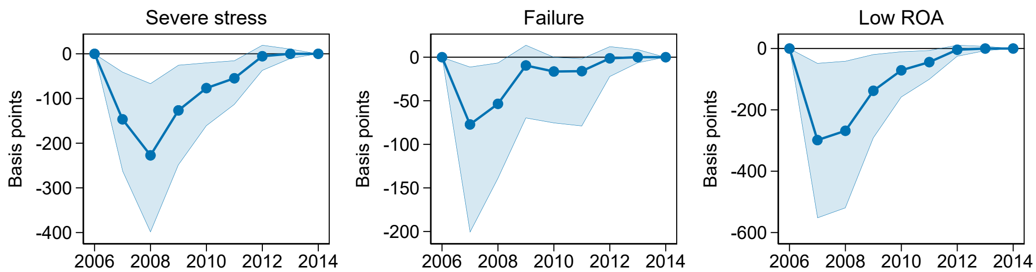
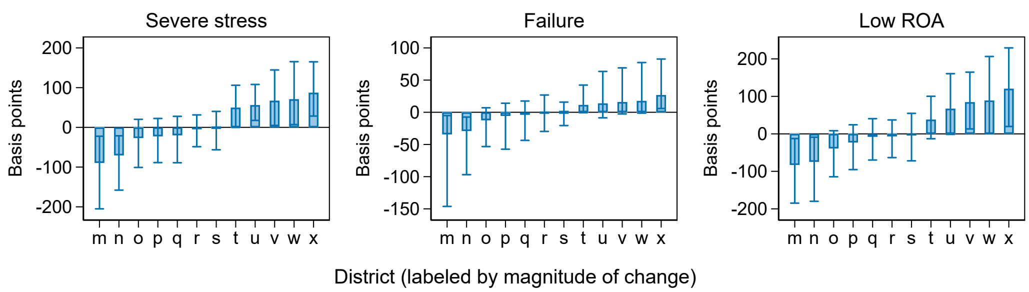


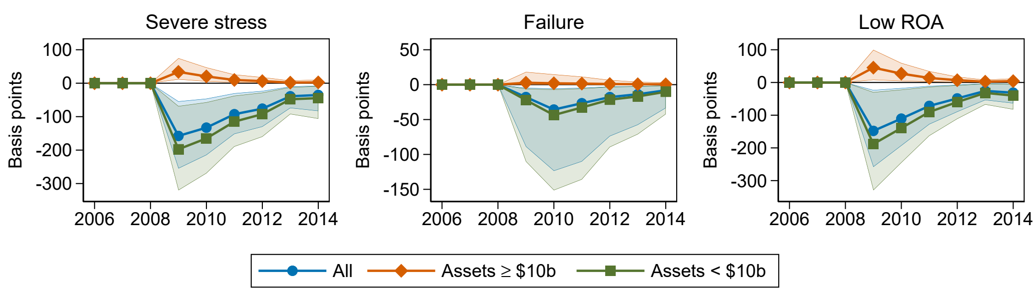
Figure 5: Expansion of supervisory resources and average shadow cost of hours. The left panel shows the allocation of total supervisor hours across BHCs under Federal Reserve supervision as of 2006, split between large BHCs (≥\$10b) and small BHCs (<\$10b), as well as new BHCs coming under Federal Reserve supervision in the following years. The right panel shows the average across districts of the estimated shadow cost of supervisor hours, estimated using a specification matching our baseline first-stage regression but including district–year fixed effects instead of year fixed effects and the district averages of the shadow-cost instrument. Shaded areas represent bootstrapped 95% confidence intervals clustered at the bank level (20,000 replications).



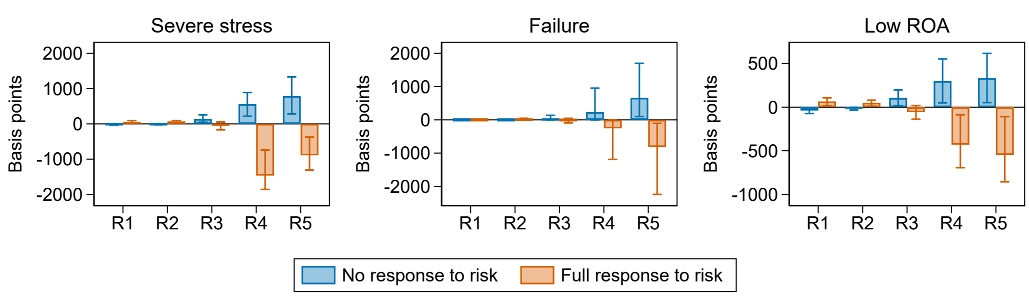
(a) Earlier expansion of resources. Average change by year.



(b) Centralized allocation. Average change by district.



(c) No reallocation to large banks. Average change by year.



(d) No/full response to bank risk. Average change by rating.

Figure 6: Change in distress probability under counterfactual experiments. The figure shows equal-weighted averages of the change in bank-level distress probability implied by the counterfactuals. See Section 7 for details on the counterfactuals. The three panels correspond to the three variables measuring distress. Shaded areas and whiskers represent bootstrapped 95% confidence intervals clustered at the bank level (20,000 replications).

Table 1: Summary statistics. The table presents summary statistics for the variables included in the regression specifications. For detailed variable definitions see Section 3 and Appendix B. Sample is 1998–2014.

(a) Explanatory variables.

	All banks			Assets<\$10b		Assets≥\$10b	
	Mean	StDev	Obs	Mean	StDev	Mean	StDev
Hours (thousands)	1.544	4.942	5900	0.314	0.498	6.868	9.704
Assets (real, 2012 $ bil.)	32.970	174.620	5900	2.961	2.262	162.754	376.441
Log Hours	5.336	2.082	5900	4.742	1.698	7.908	1.572
Log Assets (real)	8.347	1.443	5900	7.772	0.629	10.832	1.333
Rating	1.978	0.782	5900	1.969	0.810	2.016	0.643
Assets≥\$10b	0.188	0.391	5900	0.000	0.000	1.000	0.000
Post-2008	0.459	0.498	5900	0.468	0.499	0.420	0.494
Assets≥\$50b	0.075	0.263	5900	0.000	0.000	0.399	0.490
District top	0.165	0.371	5900	0.040	0.197	0.701	0.458
Complex	0.276	0.447	5900	0.184	0.388	0.674	0.469
Stressed (rating≥3)	0.150	0.357	5900	0.147	0.354	0.162	0.369
Exam count	1.469	2.386	5900	0.843	0.776	4.178	4.323
High exam frequency	0.506	0.500	5900	0.392	0.488	1.000	0.000
Return on assets	0.009	0.007	5451	0.009	0.007	0.009	0.007
Non-perf. loans ratio	0.016	0.016	5542	0.015	0.016	0.016	0.015
Tier-1 capital ratio	0.120	0.032	5565	0.122	0.032	0.110	0.031

(b) Outcome variables.

	All banks			Assets<\$10b		Assets≥\$10b	
	Mean	StDev	Obs	Mean	StDev	Mean	StDev
Failure	0.005	0.073	5445	0.006	0.076	0.003	0.054
Severe stress	0.053	0.225	5445	0.059	0.235	0.030	0.172
Low ROA	0.103	0.305	5405	0.104	0.305	0.103	0.304
High NPL ratio	0.102	0.303	5465	0.101	0.301	0.107	0.310
Low tier-1 capital ratio	0.103	0.304	5470	0.081	0.273	0.197	0.398
Low noninterest income	0.095	0.294	5502	0.103	0.304	0.059	0.236
High noninterest expense	0.103	0.304	5498	0.097	0.295	0.133	0.339
High loan-loss provisions	0.102	0.303	5452	0.095	0.293	0.136	0.343
Low net interest income	0.099	0.299	5528	0.081	0.272	0.187	0.390
Low real. gains on securities	0.102	0.303	5404	0.097	0.296	0.122	0.328

Table 2: Basic determinants of supervisory hours. The table shows estimates from linear regressions of log supervisory hours on the listed controls. For detailed variable definitions see Section 3 and Appendix B. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Log(Hours)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Assets (real)	0.959***	0.966***	0.820***	0.823***	0.827***	0.833***
	[0.030]	[0.030]	[0.044]	[0.044]	[0.051]	[0.051]
Rating = 2	0.460***	0.419***	0.465***	0.447***	0.467***	0.449***
	[0.069]	[0.069]	[0.068]	[0.069]	[0.068]	[0.069]
Rating = 3	1.314***	1.153***	1.281***	1.193***	1.281***	1.194***
	[0.103]	[0.093]	[0.100]	[0.097]	[0.100]	[0.097]
Rating = 4	1.686***	1.489***	1.673***	1.557***	1.675***	1.561***
	[0.124]	[0.115]	[0.130]	[0.127]	[0.130]	[0.127]
Rating = 5	1.858***	1.642***	1.913***	1.787***	1.915***	1.790***
	[0.174]	[0.169]	[0.175]	[0.180]	[0.175]	[0.180]
Assets≥$10b			0.328**	0.325**	0.339**	0.335**
			[0.134]	[0.134]	[0.145]	[0.146]
Post-2008×(Assets≥$10b)			0.655***	0.659***	0.620***	0.625***
			[0.117]	[0.117]	[0.139]	[0.140]
Post-2008				-0.270***		-0.269***
				[0.065]		[0.065]
Assets≥$50b					-0.084	-0.106
					[0.221]	[0.222]
Post-2008×(Assets≥$50b)					0.088	0.089
					[0.216]	[0.216]
Year FEs	Yes	No	Yes	No	Yes	No
Adj. R^2	0.50	0.49	0.50	0.50	0.50	0.50
Observations	5900	5900	5900	5900	5900	5900
Distinct BHCs	769	769	769	769	769	769

Table 3: Instruments for supervisory hours: shadow cost. The table shows estimates from linear regressions of log supervisory hours on the listed controls. Other controls and year fixed effects are noted at the bottom; abbreviations: "a" is log assets, "r" is rating = 2,…,5, "g" is assets≥\$10b, "pg" is post-2008×(assets≥\$10b). District averages either leave out or leave in bank i, as noted at the bottom. National averages leave out bank i's district. For detailed variable definitions see Section 3 and Appendix B. F-statistics are for the test that the coefficients on the instruments are zero. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Log(Hours)			
	(1)	(2)	(3)	(4)
District avg. Log Hours	0.725***	0.995***	0.723***	0.692***
	[0.053]	[0.052]	[0.051]	[0.060]
District avg. Log Assets	-0.781***	-0.976***	-0.784***	-0.726***
	[0.089]	[0.091]	[0.088]	[0.132]
District avg. (Rating = 2)	0.110	-0.364	-0.062	0.203
	[0.223]	[0.224]	[0.208]	[0.247]
District avg. (Rating = 3)	-0.463	-1.174***	-0.563	-0.252
	[0.366]	[0.367]	[0.351]	[0.422]
District avg. (Rating = 4)	-1.301*	-1.832***	-1.684**	-1.232*
	[0.685]	[0.681]	[0.708]	[0.709]
District avg. (Rating = 5)	-2.210**	-1.780**	-2.254**	-2.264***
	[0.869]	[0.868]	[0.895]	[0.864]
National avg. Log Hours			-0.273*	
			[0.153]	
National avg. Log Assets			0.105	
			[0.272]	
National avg. (Rating = 2)			-1.102*	
			[0.570]	
National avg. (Rating = 3)			-0.202	
			[1.103]	
National avg. (Rating = 4)			-3.421	
			[3.680]	
National avg. (Rating = 5)			-0.692	
			[3.515]	
Dist. avg. Post-2008×(Assets≥\$10b)				-0.416
				[0.637]
Dist. avg. District top				0.693
				[0.469]
Dist. avg. Lagged exam count				0.302
				[0.237]
Dist. avg. Lag exam ct.×(Hi. exam freq.)				-0.335
				[0.258]
Dist. avg. High exam frequency				0.396
				[0.260]
Other controls	a r g pg	a r g pg	a r g pg	a r g pg
Year FEs	Yes	Yes	No	Yes
Avg. calculated as leave-out	Yes	No	Yes	Yes
F-statistic	38.8	67.2	24.9	21.9
Adj. R^2	0.54	0.56	0.54	0.55
Observations	5900	5900	5900	5188
Distinct BHCs	769	769	769	722

Table 4: Instruments for supervisory hours: preference shifters. The table shows estimates from linear regressions of log supervisory hours on the listed controls and year fixed effects. Other controls are noted at the bottom; abbreviations: "a" is log assets, "r" is rating = 2,...,5, "g" is assets≥\$10b, "pg" is post-2008×(assets≥\$10b). For detailed variable definitions see Section 3 and Appendix B. F-statistics are for the test that the coefficients on the instruments are zero. Standard errors clustered by bank reported in brackets; significance: $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$. Sample is 1998–2014.

	Log(Hours)			
	(1)	(2)	(3)	(4)
District top	0.619***			0.737***
	[0.136]			[0.127]
Small (assets<\$10b), complex		0.477***		
		[0.106]		
Small (assets<\$10b), stressed		0.127		
		[0.150]		
Lagged exam count		-0.628***	-0.628***	-0.653***
		[0.119]	[0.119]	[0.117]
Lag exam ct.×(Small, complex)		0.884***		
		[0.126]		
Lag exam ct.×(Small, stressed)		0.683***		
		[0.153]		
Lag exam ct.×(Assets≥\$10b)		0.720***		
		[0.122]		
High exam frequency			0.493***	0.456***
			[0.088]	[0.084]
Lag exam ct.×(Hi. exam freq.)			0.727***	0.755***
			[0.122]	[0.119]
Other controls	a r g pg	a r g pg	a r pg	a r pg
F-statistic	20.7	17.9	35.6	45.7
Adj. R^2	0.50	0.58	0.57	0.58
Observations	5900	5188	5188	5188
Distinct BHCs	769	722	722	722

Table 5: Second stage of IV probit with outcome variable 1y-ahead probability of severe stress. The table shows estimates from the second stages of IV probit regressions of distress probability on the listed controls and year fixed effects where log hours are instrumented for (columns 1–4, corresponding first stages in Table A5) as well as from a non-instrumented probit regression (column 5). The instruments used and other controls are noted at the bottom. Instrument abbreviations: "SC" is shadow cost, "DT" is district top, "EF" is exam frequency. Other controls abbreviations: "g" is assets≥\$10b, "pg" is post-2008×(assets≥\$10b), "h" is high exam frequency, "lr" is lagged rating = 2,…,5. For detailed variable definitions see Section 3 and Appendix B. The effective F-statistic and critical value are for the weak-instrument test of Olea and Pflueger (2013), robust to heteroskedasticity, autocorrelation, and clustering, from the respective first stage. Average marginal effects reported in curly braces. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Severe stress$_{t+1}$				
	(1)	(2)	(3)	(4)	(5)
Log Hours	-0.216**	-0.442**	-0.307***	-0.252***	0.001
	[0.095]	[0.184]	[0.067]	[0.065]	[0.032]
	{-0.019}	{-0.059}	{-0.030}	{-0.023}	{0.000}
Log Assets (real)	0.095	0.337*	0.132**	0.131**	-0.083*
	[0.096]	[0.199]	[0.062]	[0.064]	[0.048]
	{0.008}	{0.045}	{0.013}	{0.012}	{-0.006}
Rating = 2	0.622***	0.636***	0.895***	0.876***	0.812***
	[0.165]	[0.142]	[0.191]	[0.196]	[0.209]
	{0.053}	{0.085}	{0.089}	{0.080}	{0.062}
Rating = 3	1.708***	1.748***	1.972***	1.951***	1.805***
	[0.187]	[0.180]	[0.209]	[0.212]	[0.226]
	{0.146}	{0.235}	{0.195}	{0.178}	{0.137}
Rating = 4	3.183***	3.010***	3.478***	3.487***	3.382***
	[0.187]	[0.392]	[0.230]	[0.230]	[0.245]
	{0.272}	{0.404}	{0.344}	{0.318}	{0.256}
Rating = 5	2.950***	2.870***	3.203***	3.185***	3.014***
	[0.239]	[0.335]	[0.257]	[0.262]	[0.280]
	{0.253}	{0.386}	{0.317}	{0.290}	{0.228}
Instrument	SC	DT	EF	All	None
Other controls	g pg	g pg	pg h lr	pg h lr	pg h lr
F-statistic	42.5	15.8	31.7	40.3	
Critical value	15.3	23.1	11.6	16.6	
Observations	4924	4924	4290	4290	4290
Distinct BHCs	744	744	698	698	698

Table 6: Second stage of IV probit with outcome variables 1y-ahead probability of failure and 1y-ahead probability of low ROA. The table shows estimates from the second stages of IV probit regressions of distress probability on the listed controls and year fixed effects where log hours are instrumented for (columns 1–3 and 5–8, corresponding first stages in Tables A6 and A7) as well as from non-instrumented probit regressions (columns 4 and 9). The outcome variable is noted at the top. The instruments used and other controls are noted at the bottom. Instrument abbreviations: "SC" is shadow cost, "DT" is district top, "EF" is exam frequency. Other controls abbreviations: "g" is assets≥\$10b, "pg" is post-2008×(assets≥\$10b), "h" is high exam frequency, "lr" is lagged rating = 2,…,5. For detailed variable definitions see Section 3 and Appendix B. The effective F-statistic and critical value are for the weak-instrument test of Olea and Pflueger (2013), robust to heteroskedasticity, autocorrelation, and clustering, from the respective first stage. Average marginal effects reported in curly braces. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Failure$_{t+1}$				Low ROA$_{t+1}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Hours	-0.151	-0.385***	-0.366***	-0.092**	-0.162**	-0.401***	-0.114	-0.147**	0.001
	[0.180]	[0.062]	[0.082]	[0.044]	[0.078]	[0.155]	[0.092]	[0.059]	[0.022]
	{-0.003}	{-0.013}	{-0.012}	{-0.002}	{-0.022}	{-0.069}	{-0.015}	{-0.020}	{0.000}
Log Assets (real)	-0.007	0.231***	0.218***	0.037	0.073	0.358**	0.039	0.112*	-0.043
	[0.182]	[0.075]	[0.075]	[0.060]	[0.079]	[0.156]	[0.073]	[0.062]	[0.038]
	{-0.000}	{0.008}	{0.007}	{0.001}	{0.010}	{0.062}	{0.005}	{0.015}	{-0.006}
Rating = 2	0.406*	3.658***	3.617***	3.539***	0.727***	0.729***	0.707***	0.709***	0.666***
	[0.220]	[0.286]	[0.236]	[0.278]	[0.090]	[0.092]	[0.119]	[0.116]	[0.117]
	{0.009}	{0.128}	{0.116}	{0.059}	{0.099}	{0.126}	{0.095}	{0.097}	{0.087}
Rating = 3	0.805**	4.391***	4.344***	4.125***	1.699***	1.757***	1.710***	1.717***	1.619***
	[0.377]	[0.355]	[0.332]	[0.374]	[0.126]	[0.119]	[0.155]	[0.148]	[0.147]
	{0.017}	{0.154}	{0.139}	{0.069}	{0.230}	{0.303}	{0.230}	{0.234}	{0.212}
Rating = 4	1.744***	5.495***	5.465***	5.187***	2.551***	2.586***	2.716***	2.726***	2.593***
	[0.475]	[0.478]	[0.443]	[0.524]	[0.182]	[0.186]	[0.217]	[0.209]	[0.206]
	{0.038}	{0.192}	{0.175}	{0.086}	{0.346}	{0.446}	{0.365}	{0.372}	{0.340}
Rating = 5	2.477***	6.166***	6.126***	5.875***	3.245***	3.289***	3.287***	3.304***	3.135***
	[0.509]	[0.540]	[0.503]	[0.571]	[0.291]	[0.303]	[0.306]	[0.300]	[0.290]
	{0.054}	{0.216}	{0.196}	{0.098}	{0.440}	{0.566}	{0.441}	{0.451}	{0.411}
Instrument	SC	EF	All	None	SC	DT	EF	All	None
Other controls	g	h lr	h lr	h lr	g pg	g pg	pg h lr	pg h lr	pg h lr
F-statistic	31.4	30.0	29.5		47.3	21.5	24.7	38.5	
Critical value	15.0	11.3	17.0		15.4	23.1	14.0	17.3	
Observations	3418	2860	2860	2860	5405	5405	4717	4717	4717
Distinct BHCs	715	645	645	645	750	750	678	678	678

Table 7: Change in distress probability under counterfactual experiments. The table shows averages of the bank–year level distress probabilities in panel (a) and of changes in distress probabilities under counterfactual policy experiments in panels (b)–(g), in basis points. See Section 7 for details on the counterfactuals. Distress outcome variables are noted at the top. "EW" denotes equal-weighted and "SW" size-weighted averages. Averages in panels (b) and (d) include only the affected years. Significance based on bootstrapped samples clustered at the bank level (20,000 replications): * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

		Severe stress		Failure		Low ROA	
		(1) EW	(2) SW	(3) EW	(4) SW	(5) EW	(6) SW
(a) Baseline probability	All banks	534.4***	337.5***	53.3***	8.0***	1034.2***	1125.9***
	Assets<\$10b	588.0***	540.6***	58.8***	55.2***	1035.1***	913.7***
	Assets≥\$10b	303.0***	321.5***	29.3***	4.2***	1030.4***	1141.6***
	Rating<3	162.5***	40.0***	43.3***	6.3***	750.1***	735.8***
	Rating≥3	2602.4***	1188.3***	108.4***	12.7***	3035.7***	2061.2***
(b) Earlier expansion	All banks	-78.0***	-48.6**	-21.1**	-11.4**	-100.0**	-101.1**
	Assets<\$10b	-82.4***	-79.4***	-22.4**	-21.8**	-99.4**	-100.3**
	Assets≥\$10b	-57.5***	-46.5**	-15.0**	-10.7*	-103.1**	-101.1**
	Rating<3	-54.6***	-38.9***	-14.7**	-15.0**	-101.4**	-134.5**
	Rating≥3	-154.1***	-61.4	-41.9**	-6.7	-95.5**	-57.5
(c) Centralized allocation	All banks	-4.4	-24.3***	-3.3	-5.5**	-2.0	-32.3**
	Assets<\$10b	-3.0	-4.8	-3.1	-3.6	-0.6	-1.8
	Assets≥\$10b	-10.5	-25.8***	-4.3	-5.7**	-8.2	-34.7**
	Rating<3	-4.1	-5.9**	-1.6	-5.9*	-3.5	-17.5**
	Rating≥3	-6.0	-68.0***	-13.0	-4.7*	6.1	-67.6**
(d) No post-2008 realloc.	All banks	-88.3***	-2.4**	-20.0***	-1.1**	-72.0**	1.0
	Assets<\$10b	-110.1***	-99.0***	-24.6***	-21.8***	-91.1**	-87.6**
	Assets≥\$10b	12.0***	4.0***	1.4***	0.3***	16.1**	6.8**
	Rating<3	-31.2***	-2.2**	-5.5***	-0.5	-41.8**	0.4
	Rating≥3	-237.7***	-2.5	-57.9***	-1.6**	-150.9**	1.5
(e) No district-top effect	All banks	-57.1***	0.9	-18.1***	0.4	-59.3**	5.8**
	Assets<\$10b	-68.5***	-60.1***	-21.8***	-20.4***	-71.3**	-65.1**
	Assets≥\$10b	-10.6*	5.4**	-2.8	2.0	-10.4	11.0**
	Rating<3	-30.6***	-0.5	-11.0***	1.2	-47.1**	5.9**
	Rating≥3	-192.2***	4.0	-54.1***	-1.4	-121.5**	5.5
(f) No response to risk	All banks	38.8***	21.3***	22.9***	3.9***	9.2**	6.2**
	Assets<\$10b	43.5***	34.3***	27.1***	22.0***	10.3**	4.4
	Assets≥\$10b	19.7***	20.3***	6.0***	2.6**	4.7	6.3*
	Rating<3	-12.5***	-6.1***	-0.7	-0.6	-22.4**	-19.4**
	Rating≥3	300.5***	83.9***	143.2***	14.2***	170.5**	64.7**
(g) Full response to risk	All banks	-17.2	-3.4	-20.3**	-1.4	13.3*	18.2**
	Assets<\$10b	-21.6	-6.6	-24.7**	-18.8**	11.9	19.4**
	Assets≥\$10b	0.7	-3.2	-2.3	-0.1	19.2**	18.1**
	Rating<3	64.9***	35.4***	5.3	3.2	53.4**	49.7**
	Rating≥3	-436.0***	-92.1***	-151.0**	-11.8**	-191.1**	-53.6**

Internet Appendix

A Proofs and details

A.1 Direct effect of size on risk

In the model of Section 2, we assume that bank size enters the distress threshold only indirectly through the hours required for a certain intensity of supervision s_{idt}:

$$\begin{aligned} D_{idt} &= \varrho(r_{idt}) - \gamma s_{idt} \\ &= \varrho(r_{idt}) - \gamma \log H_{idt} + \alpha\gamma \log A_{idt}. \end{aligned}$$

Instead, one could assume that bank size also affects risk directly. For example, larger banks could be inherently less risky because of better asset diversification. Suppose log assets enter the distress threshold directly with impact $\omega > 0$, then the distress threshold becomes

$$\begin{aligned} D_{idt} &= \varrho(r_{idt}) - \omega \log A_{idt} - \gamma s_{idt} \\ &= \varrho(r_{idt}) - \gamma \log H_{idt} + (\alpha\gamma - \omega) \log A_{idt}. \end{aligned} \tag{18}$$

The total hours cost of achieving a distress threshold D_{idt} is then given by

$$H_{idt} = \exp\left(\frac{\varrho(r_{idt}) - D_{idt}}{\gamma}\right) \times A_{idt}^{\alpha - \frac{\omega}{\gamma}},$$

and we say there are economies of scale in supervision if $\alpha - \omega/\gamma < 1$. Under this more general formulation, the coefficient on log assets in the second stage (11) includes the parameter ω, that is, $\beta_A = \alpha\gamma - \omega$. However, the presence of scale economies would still be determined by the ratio of the coefficients on log assets and on log hours, $-\beta_A/\beta_H = \alpha - \omega/\gamma \lesseqgtr 1$.

Our evidence for economies of scale in supervision is based on finding that the ratio of coefficients satisfies $-\beta_A/\beta_H < 1$ (Section 6.2). This inference is valid even if bank risk depends directly on bank size. Ultimately, scale economies mean that larger banks can be supervised with proportionately fewer resources, and this can be due to a combination of larger banks being inherently less risky ($\omega > 0$) and supervisors being more effective at larger banks ($\alpha < 1$).

A.2 Proof of Proposition 1

All the comparative static results in Proposition 1 result from implicit differentiation of the first-order condition (6). Dropping subscripts and denoting $\text{PD}(\varrho, s) \equiv \Pr\left[y^*_{t+1} > 0 \,\middle|\, \varrho, s\right]$

for clarity, the implicit derivatives are:

$$\frac{dH}{dA} = \frac{-\frac{\partial^2}{\partial s^2}\mathrm{PD}(\varrho, s)\,\frac{\partial}{\partial A}h^{-1}(H, A)\,W - \frac{\partial}{\partial s}\mathrm{PD}(\varrho, s)\,\frac{dW}{dA} - \frac{\partial^2}{\partial A \partial s}h(s, A)\,\Lambda}{\frac{\partial^2}{\partial s^2}\mathrm{PD}(\varrho, s)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}W + \frac{\partial^2}{\partial s^2}h(s, A)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}\Lambda}$$

$$\frac{dH}{d\varrho} = \frac{-\frac{\partial^2}{\partial \varrho \partial s}\mathrm{PD}(\varrho, s)\,W - \frac{\partial}{\partial s}\mathrm{PD}(\varrho, s)\,\frac{dW}{d\varrho}}{\frac{\partial^2}{\partial s^2}\mathrm{PD}(\varrho, s)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}W + \frac{\partial^2}{\partial s^2}h(s, A)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}\Lambda}$$

$$\frac{dH}{dW} = \frac{-\frac{\partial}{\partial s}\mathrm{PD}(\varrho, s)}{\frac{\partial^2}{\partial s^2}\mathrm{PD}(\varrho, s)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}W + \frac{\partial^2}{\partial s^2}h(s, A)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}\Lambda}$$

$$\frac{dH}{d\Lambda} = \frac{-\frac{\partial}{\partial s}h(s, A)}{\frac{\partial^2}{\partial s^2}\mathrm{PD}(\varrho, s)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}W + \frac{\partial^2}{\partial s^2}h(s, A)\,\frac{1}{\frac{\partial}{\partial s}h(s,A)}\Lambda}$$

For the various partial derivatives, we have:[30]

$$\frac{\partial \mathrm{PD}}{\partial s} = -\phi\left(\frac{\varrho - \gamma s}{\sigma_u}\right)\frac{\gamma}{\sigma_u} < 0$$

$$\frac{\partial^2 \mathrm{PD}}{\partial s^2} = \phi'\left(\frac{\varrho - \gamma s}{\sigma_u}\right)\frac{\gamma^2}{\sigma_u^2} > 0$$

$$\frac{\partial^2 \mathrm{PD}}{\partial \varrho \partial s} = -\phi'\left(\frac{\varrho - \gamma s}{\sigma_u}\right)\frac{\gamma}{\sigma_u^2} < 0$$

$$\frac{\partial h}{\partial s} = \exp(s)\,A^{\alpha} > 0$$

$$\frac{\partial^2 h}{\partial s^2} = \exp(s)\,A^{\alpha} > 0$$

$$\frac{\partial^2 h}{\partial s \partial A} = \alpha \exp(s)\,A^{\alpha-1} > 0$$

$$\frac{\partial h^{-1}}{\partial A} = -\alpha / A < 0$$

The denominator is the same in all four implicit derivatives and is positive. The sign of the comparative statics is therefore determined by the numerator.

- Larger banks receive more attention, $dH/dA > 0$, if

$$-\frac{\partial^2}{\partial s^2}\mathrm{PD}(\varrho, s)\,\frac{\partial}{\partial A}h^{-1}(H, A)\,W - \frac{\partial}{\partial s}\mathrm{PD}(\varrho, s)\,\frac{dW}{dA} > \frac{\partial^2}{\partial A \partial s}h(s, A)\,\Lambda. \qquad (19)$$

  This holds for $dW/dA = 0$ and is strengthened (attenuated) for $dW/dA > 0$ ($dW/dA < 0$).

[30] Note that the decreasing marginal impact $\partial^2\mathrm{PD}/(\partial s)^2 > 0$ requires the distress threshold $D_{idt} = \varrho_{idt} - \gamma s_{idt}$ to be in the left tail of the distribution of u_{idt+1} where the density ϕ is increasing. This requires $\mathrm{PD}_{idt} < 0.5$, which is satisfied in the data.

- Riskier banks receive more attention, $dH/d\varrho > 0$, if

$$-\frac{\partial^2}{\partial\varrho\partial s}\mathrm{PD}(\varrho, s)\, W - \frac{\partial}{\partial s}\mathrm{PD}(\varrho, s)\,\frac{dW}{d\varrho} > 0.$$

  This holds for $dW/d\varrho = 0$ and is strengthened (attenuated) for $dW/d\varrho > 0$ ($dW/d\varrho < 0$).

- Banks with a higher preference weight receive more attention, $dH/dW > 0$, since

$$-\frac{\partial}{\partial s}\mathrm{PD}(\varrho, s) > 0.$$

- A higher shadow cost reduces attention, $dH/d\Lambda < 0$, since

$$-\frac{\partial}{\partial s}h(s, A) < 0.$$

A.3 Proof of Proposition 2

For any variable x_{idt} that enters the distress variable y^*_{idt+1} with loading κ and the log supervisory preference weight $\log W_{idt}$ with loading $\widetilde{\kappa}$, implicit differentiation of the first-order condition (8) yields

$$\frac{d\log H_{idt}}{dx_{idt}} = \frac{\phi'(z_{idt})\,\frac{\kappa}{\sigma_\varepsilon}\frac{\gamma}{\sigma_\varepsilon}\,W_{idt} + \phi(z_{idt})\,\frac{\gamma}{\sigma_\varepsilon}\,W_{idt}\widetilde{\kappa}}{\phi'(z_{idt})\,\frac{\gamma}{\sigma_\varepsilon}\frac{\gamma}{\sigma_\varepsilon}\,W_{idt} + H_{idt}\Lambda_{dt}},$$

where

$$z_{idt} = \frac{\rho_1 + \sum_{r=2}^{5}\rho_r\mathbb{I}[r_{idt} = r] - \gamma\log H_{idt} + \alpha\gamma\log A_{idt} + \eta_{idt}}{\sigma_\varepsilon}.$$

Substituting in the LHS of the first-order condition for $H_{idt}\Lambda_{dt}$ yields

$$\begin{aligned}\frac{d\log H_{idt}}{dx_{idt}} &= \frac{\phi'(z_{idt})\,\frac{1}{\sigma_\varepsilon}\,\kappa + \phi(z_{idt})\,\widetilde{\kappa}}{\phi'(z_{idt})\,\frac{\gamma}{\sigma_\varepsilon} + \phi(z_{idt})} \\ &= \pi_{idt}\frac{\kappa}{\gamma} + (1 - \pi_{idt})\,\widetilde{\kappa},\end{aligned}$$

where the local weight π_{idt} is given by

$$\pi_{idt} = \frac{\phi'(z_{idt})\,\frac{\gamma}{\sigma_\varepsilon}}{\phi'(z_{idt})\,\frac{\gamma}{\sigma_\varepsilon} + \phi(z_{idt})}.$$

Making use of the fact that $\phi'(z) = -\phi(z)\, z$ and the definition of the distress probability in (7) yields

$$\begin{aligned}\pi_{idt} &= \frac{-z_{idt}\frac{\gamma}{\sigma_\varepsilon}}{-z_{idt}\frac{\gamma}{\sigma_\varepsilon}+1} \\ &= \frac{\Phi^{-1}(\mathrm{PD}_{idt})}{\Phi^{-1}(\mathrm{PD}_{idt})-\frac{\sigma_\varepsilon}{\gamma}}.\end{aligned}$$

For $\mathrm{PD}_{idt} < 1/2$, we know that $z_{idt} < 0$ and therefore $\pi_{idt} \in (0,1)$.

A.4 Linearization of first-order condition

The first-order condition for the supervisor's problem 5 is given by

$$\begin{aligned}\sigma_\varepsilon^{-1}\phi\Big(\left(\varrho(r_{idt}) - \gamma \log H_{idt} + \alpha\gamma \log A_{idt} + \eta_{idt}\right)\sigma_\varepsilon^{-1}\Big)\,\gamma \exp(\widetilde{\varrho}(r_{idt}) + \widetilde{\alpha}\log A_{idt} + w_{idt}) \\ = \Lambda_{dt}\exp(\log H_{idt})\end{aligned}$$

Suppose we linearize around a point

$$(r_{idt}, \log H_{idt}, \log A_{idt}, \eta_{idt}, w_{idt}, \Lambda_{dt}) = \left(\overline{r}, \overline{\log H}, \overline{\log A}, \overline{\eta}, \overline{w}, \overline{\Lambda}\right),$$

where the first-order condition holds:

$$\phi\left(\frac{\varrho(\overline{r}) - \gamma\overline{\log H} + \alpha\gamma\overline{\log A} + \overline{\eta}}{\sigma_\varepsilon}\right)\frac{\gamma}{\sigma_\varepsilon}\exp\left(\widetilde{\varrho}(\overline{r}) + \widetilde{\alpha}\overline{\log A} + \overline{w}\right) = \overline{\Lambda}\exp\overline{h}.$$

Solving for $\log H_{idt}$ then yields

$$\log H_{idt} \approx \overline{Y} + \frac{\overline{\pi}}{\gamma}\varrho(r_{idt}) + (1-\overline{\pi})\,\widetilde{\varrho}(r_{idt}) + \left(\overline{\pi}\alpha + (1-\overline{\pi})\,\widetilde{\alpha}\right)\log A_{idt} + \delta_w w_{idt} + \delta_\Lambda \Lambda_{dt} + \delta_\eta \eta_{idt}$$

with coefficients

$$\begin{aligned}\overline{\pi} &= \frac{\sigma_\varepsilon^{-1}\phi'(\overline{y})\,\gamma}{\sigma_\varepsilon^{-1}\phi'(\overline{y})\,\gamma + \phi(\overline{y})}, \\ \delta_\Lambda &= -\frac{\exp\left(\overline{\log H}\right)}{\sigma_\varepsilon^{-1}\phi'(\overline{y})\,\gamma + \phi(\overline{y})}, \\ \delta_w &= \frac{\phi(\overline{y})}{\sigma_\varepsilon^{-1}\phi'(\overline{y})\,\gamma + \phi(\overline{y})}, \\ \delta_\eta &= \frac{\sigma_\varepsilon^{-1}\phi'(\overline{y})}{\sigma_\varepsilon^{-1}\phi'(\overline{y})\,\gamma + \phi(\overline{y})},\end{aligned}$$

and a constant

$$\overline{Y} = \overline{\log H} - \frac{\overline{\pi}}{\gamma}\varrho(\overline{r}) - (1-\overline{\pi})\,\widetilde{\varrho}(\overline{r}) - (\overline{\pi}\alpha + (1-\overline{\pi})\,\widetilde{\alpha})\,\overline{\log A} - \delta_w\overline{w} + \delta_\Lambda\overline{\Lambda} - \delta_\eta\overline{\eta},$$

where

$$\overline{y} = \left(\varrho(\overline{r}) - \gamma\overline{\log H} + \alpha\gamma\overline{\log A} + \overline{\eta}\right)\sigma_\varepsilon^{-1},$$
$$\overline{W} = \exp\left(\widetilde{\varrho}(\overline{r}) + \widetilde{\alpha}\overline{\log A} + \overline{w}\right).$$

Since we estimate the linearized first order condition, the point around which the linearization is evaluated is determined by the regression coefficients. The regression fits a linear relationship to the potentially nonlinear relationship between log hours and the covariates minimizing the mean squared error. Since the relationship between log hours and log assets is approximately linear and we control for ratings as well as most shifters non-linearly with dummies, the approximation error of our linearization is likely to be low.

A.5 Variation in shadow cost instrument

We can characterize across- and within-district variation in the shadow cost instrument, calculated as $-\overline{\delta}_H\overline{\log H}_{dt} + \overline{\delta}_A\overline{\log A}_{dt} + \sum_{r=2}^{5}\overline{\delta}_r\overline{\mathbb{I}[r]}_{dt}$, where $\overline{x}_{dt}$ denotes the average of variable x within district d and year t,[31] and where $\overline{\delta}_H$, $\overline{\delta}_A$ and $\overline{\delta}_r$ are the coefficients on average log hours, log assets, and rating indicators from Table 3, column (1). The overall variation of the total shadow cost is close to equally split into variation between and within districts (Appendix Table A3). To determine the source of variation, we decompose total shadow cost into two additive components, the "assets/hours component" $\overline{\delta}_A\overline{\log A}_{dt} - \overline{\delta}_H\overline{\log H}_{dt}$ (because assets and hours co-vary strongly) and the "ratings component" $\sum_{r=2}^{5}\overline{\delta}_r\overline{\mathbb{I}[r]}_{dt}$.[32] The majority of the overall variation in shadow cost is due to variation in the assets/hours component, which varies both between and within districts. In contrast, the ratings component varies mostly within districts.

A.6 Bounds on supervisory preference parameters

In the convex combination (14), the structural parameter κ appears normalized by γ, which corresponds to the second stage coefficient on log hours, β_H. Without relying on estimates of $\overline{\pi}$, equation (14) implies bounds for the unobserved preference parameter: $\widetilde{\kappa}$ is *greater* than the first-stage coefficient $\widehat{\delta}_x$ if the normalized second stage coefficient $\widehat{\beta}_x/\widehat{\beta}_H$ is *smaller* than the first-stage coefficient $\widehat{\delta}_x$ and vice versa. Figure A1 shows the normalized second-

[31] Even though we construct leave-out averages $\overline{x}_{-idt}$ to instrument for each individual bank's hours, the average across banks of leave-out averages equals the overall leave-in average used to illustrate the sources of variation, $\frac{1}{|\mathcal{I}_{dt}|}\sum_{i\in\mathcal{I}_{dt}}\overline{x}_{-idt} = \frac{1}{|\mathcal{I}_{dt}|}\sum_{i\in\mathcal{I}_{dt}}x_{idt} = \overline{x}_{dt}$.

[32] The squared overall standard deviations of the two components sum to about the squared overall standard deviation of the total shadow cost, indicating only limited covariance between the two components.

stage coefficients and the first-stage coefficients for log assets and the dummies for ratings 2 to 5 for each of the three distress outcomes. For log assets, the normalized second-stage coefficient is weakly smaller than the first-stage coefficient and both are significantly greater than zero for all distress outcomes. This implies $\widetilde{\alpha} > 0$, that is, supervisors give more weight to the distress probability of larger banks. In contrast, for ratings 2 to 5, the normalized second-stage coefficient is significantly larger than the first-stage coefficient for all distress outcomes (p-val < 0.05). This implies $\widetilde{\rho}_r \ll \rho_r$, that is, the additional weight $\widetilde{\rho}_r$ that supervisors place on a bank rated r (compared to a bank rated 1) is considerably *lower* than the relative effect of rating r on the probability of distress.

A.7 Details on counterfactuals

The transfers of the different counterfactual allocations are as follows:

- Earlier expansion:

$$\widehat{\tau}_{dt}^{\text{early}} = \log\left(\frac{\sum_{i\in\mathcal{I}_{d2014}} \exp\left(\widehat{\log H}_{id2014}\right)}{\sum_{i\in\mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt}\right)}\right) \mathbb{I}[t \geq 2007]$$

- Centralized allocation:

$$\widehat{\tau}_{t}^{\text{central}} \equiv \log\left(\frac{\sum_{i\in\bigcup_d \mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt}\right)}{\sum_{i\in\bigcup_d \mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt} + \widehat{\lambda}_{dt}\right)}\right)$$

- No post-2008 reallocation:

$$\widehat{\tau}_{dt}^{\text{noreall}} = \log\left(\frac{\sum_{i\in\mathcal{I}_{d2008}} \exp\left(\widehat{\log H}_{id2008}\right)}{\sum_{i\in\mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt} - \widehat{\delta}_{\text{post-large}}\mathbb{I}(t > 2008)\,\mathbb{I}(A_{idt} > \$10\text{b})\right)}\right) \mathbb{I}[t > 2008]$$

- No district-top effect:

$$\widehat{\tau}_{dt}^{\text{notop}} \equiv \log\left(\frac{\sum_{i\in\mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt}\right)}{\sum_{i\in\mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt} - \widehat{\delta}_{\text{top}}\mathbb{I}(i \in \text{top}_{dt})\right)}\right)$$

- No response to risk:

$$\widehat{\tau}_{dt}^{\text{norisk}} = \log\left(\frac{\sum_{i\in\mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt}\right)}{\sum_{i\in\mathcal{I}_{dt}} \exp\left(\widehat{\log H}_{idt} - \sum_{r=2}^{5}\widehat{\rho}_{Hr}\mathbb{I}[r_{idt} = r]\right)}\right)$$

- Full response to risk:

$$\hat{\tau}_{dt}^{\text{fullrisk}} = \log\left(\frac{\sum_{i\in\mathcal{I}_{dt}}\exp\left(\widehat{\log H_{idt}}\right)}{\sum_{i\in\mathcal{I}_{dt}}\exp\left(\widehat{\log H_{idt}} - \sum_{r=2}^{5}\hat{\rho}_{Hr}\mathbb{I}[r_{idt}=r] + \sum_{r=2}^{5}\hat{\rho}_{r}\mathbb{I}[r_{idt}=r]\right)}\right)$$

B Detailed variable definitions

This appendix provides definitions of the variables used. Codes such as "BHCK2170" refer to form FR Y-9C.

Hours and ratings: See the discussion in Section 3.

Assets: Total assets (BHCK2170).

Assets (real): Total assets expressed in 2012 dollars, using the implicit price deflator (GDPDEF) from FRED.

Assets≥\$10b: An indicator for BHCs with assets greater than \$10 billion (nominal).

Post-2008: An indicator for all years 2009 and later.

Assets≥\$50b: An indicator for BHCs with assets greater than \$50 billion (nominal).

District top: An indicator for BHCs ranked 1st to 5th in their district by assets or within 25% of the 5th ranked BHC's assets.

Complex: An indicator for BHCs considered a "complex organization" based on supervisory judgment and updated at least annually (RSSD 9057); a complex BHC is defined as one with material credit-extending nonbank subsidiaries or debt outstanding to the general public (see SR letter 13-21).

Stressed: An indicator for BHCs with supervisory ratings 3, 4, or 5.

Exam count: The total number of supervisory exams of the BHC in a year.

High exam frequency: An indicator for BHCs that are large (assets≥\$10b) and/or complex and/or stressed.

Return on assets: The ratio of net income (BHCK4340) to assets.

Non-performing loans ratio: The ratio of total non-performing loans (total loans and leases, 90+ days past due [BHCK5525 net of BHCK3506], and nonaccrual [BHCK5526 net of BHCK3507]) to total loans net of unearned income (BHCK2122).

Tier-1 capital ratio: Tier-1 risk-based capital divided by risk-weighted assets. Basel I (pre-2014) BHCK8274/BHCKa223; Basel III (post-2014; including 2014 for advanced-approaches firms) BHCA8274/BHCAA223.

Failure: An indicator for BHCs that fail. A BHC fails when it terminates and the reason for its termination or the termination of a subsidiary within one quarter is failure (RSSD 9061).

Severe stress: An indicator for BHCs that have a rating of 4 or 5 and/or fail.

Low ROA: An indicator for BHCs with returns on assets in the bottom 10th percentile of the pooled distribution.

High NPL ratio: An indicator for BHCs with non-performing loans ratio in the top 10th percentile of the pooled distribution.

Low tier-1 capital ratio: An indicator for BHCs with tier-1 capital ratio in the bottom 10th percentile of the pooled distribution.

Low noninterest income: An indicator for BHCs with noninterest income (BHCK4079) relative to assets in the bottom 10th percentile of the pooled distribution.

High noninterest expense: An indicator for BHCs with noninterest expense (BHCK4093) relative to assets in the top 10th percentile of the pooled distribution.

High loan-loss provisions: An indicator for BHCs with loan-loss provisions (BHCK4230) relative to assets in the top 10th percentile of the pooled distribution.

Low net interest income: An indicator for BHCs with net interest income (BHCK4074) relative to assets in the bottom 10th percentile of the pooled distribution.

Low realized gains on securities: An indicator for BHCs with realized gains on securities (BHCK3521 and BHCK3196) relative to assets in the bottom 10th percentile of the pooled distribution.

C OCC assessment fees

One potential concern in using the Federal Reserve supervisory hours data as a measure of supervisory efforts is that the quality of hours are unaccounted for because we do not measure price information. In addition, the hours data may not be representative of banking supervisors other than the Federal Reserve. To validate the hours data, we compare the estimated elasticities of supervisory hours with respect to bank size and risk just discussed to those of assessment fees collected by the Office of the Comptroller of the Currency (OCC) on its supervised entities.

We use data on supervisory fees assessed on federally chartered commercial banks by the OCC, which we obtain from the OCC's public website.[33] The OCC supervises nationally chartered commercial banks as well as federal savings associations (FSAs) since 2011 following the integration of the Office of Thrift Supervision (OTS) into the OCC. The OCC

[33] www.occ.treas.gov/topics/examinations/assessments-and-fees/index-assessments-fees.html. See also Kisin and Manela (2014) for another work using this same information.

levies assessments, fees, and other charges on federally chartered banks to meet the expenses of carrying out its supervisory activities. The OCC assesses semi-annual fees on its supervised entities under 12 U.S.C. 13 and 12 CFR 8. The fee schedule is adjusted by the OCC each year and determines fees as a function of bank size and bank risk, as measured by confidential supervisory ratings. As the hours data discussed above could in principle be very noisy measures, we use this information to compare how Federal Reserve supervisory hours and OCC fees vary as a function of bank assets and supervisory rating. In contrast to the hours data, this fee data is a more direct measure of the supervisory cost function as fees are expressed in dollar terms (see Appendix C.1 for details). However, because of potential cross-subsidies across different size or risk categories, the assessment schedule may not be directly informative of the supervisory production function at an institution level.

Bearing these caveats in mind, we apply the fee structure to the universe of nationally chartered commercial banks using asset information as of 2006:Q4 and 2013:Q4 (relevant periods for fee calculations in 2007 and 2014) and regress log fees on log assets and bank rating in Table A2, column (1). The elasticity of OCC fees to assets is 0.70, which is close to the 0.75 estimate for the within-bank estimate using Federal Reserve hours data (Table 2, column 6). The increase in OCC fees with respect to bank risk is similar although not as steep as the estimated increase in Federal Reserve hours. Relative to 1-rated institutions, fees increase by about 50% on average for 3-rated institutions ($\exp(0.4) - 1 \approx 0.5$) and by about 100% for 4- or 5-rated institutions. Overall, we find that size and risk elasticities of assessment fees are similar to those estimated on Federal Reserve supervisory hours, suggesting that federal supervisors display similar sensitivities and that hours sensitivities capture cost sensitivities reasonably well. Among other instruments discussed in Section 5, we use pre- and post-2008 asset discontinuities. Columns (2) and (3) of Table A2 extend the baseline OCC specification to include asset thresholds at $10 billion and $50 billion, and interactions of each threshold with a post-2008 dummy variable. In the next Section we provide evidence that Federal Reserve supervisory hours at the largest banks increased after 2008. Consistently, OCC assessment fees for these banks also increased in the post-2008 sample after controlling for log assets (discontinuities are present both at $10 billion and $50 billion). Overall, the OCC assessment fee data shows that the sensitivities of supervisory hours and assessment fees to size and risk are similar.

C.1 OCC fee data

The OCC's base assessment is calculated using a table with eleven categories, or brackets, each of which comprises a range of asset-size values. In addition to the base amount, which is the same for every bank in its asset-size bracket, the fee includes a marginal amount, which is computed by applying a marginal assessment rate to the assets in excess of the lower bound of the asset-size bracket. The marginal assessment rate declines as asset size increases, "reflecting economies of scale in bank examination and supervision" (Federal Register Vol. 79, No. 81, April 28, 2014).

Table A1 provides summaries for semiannual assessments (meaning that annual fees are twice as large) as a function of assets in 2007 and 2014 that we obtain from OCC bulletins. The 2014 fee structure includes a new bracket for the largest banks, with assets greater than $250 billion. This additional bracket was introduced to help the OCC recover additional costs associated with supervising large and complex banks. Starting in 2001, the OCC began imposing a surcharge of 25% on their original (size-based) assessment for national banks with a 3, 4, or 5 rating, to "reflect the increased cost of supervision" (OCC 2000-30). By 2004, the size of the surcharge had been increased to 50% for 3-rated banks and to 100% for 4- or 5-rated banks.[34]

D Robustness checks

In Tables A9 and A10, we consider additional robustness exercises: (i) including district fixed effects and an additional set of risk controls, (ii) excluding the largest banks with assets greater than $50 billion, (iii) excluding the crisis years 2008–2009, and (iv) running the estimation in a linear probability instrumental variable setting. For brevity, we do not revisit the first-stage regressions, which are the same, up to sample differences, to first stages reported previously. However, each second-stage table reports effective F-statistics from the respective first stages and critical values for the weak instrument test of Olea and Pflueger (2013).

District fixed effects and additional risk controls. Table A9 shows IV probit estimates for severe stress, failure, and low ROA when also controlling for district fixed effects and additional bank risk controls, separately and together, when including all instruments. Although the sample size is reduced, results are similar to the baseline specification in terms of statistical significance and economic magnitudes. And we find similar effects when separately including year fixed effects or additional risk controls, especially for our main outcome variable, severe stress.

Exclude largest banks. In the first three columns of Table A10, we estimate the main specification for each outcome variable but exclude the largest banks with assets greater than $50 billion. We see that excluding the largest banks has negligible effects on our estimates.

Exclude crisis years. In the middle three columns of Table A10, we estimate the main specification for each outcome variable but exclude the crisis years 2008–2009. We see that excluding the crisis years also has negligible effects on our estimates.

Linear IV. In the last three columns of Table A10, we estimate the main specification for each outcome variable as a linear instrumental variables regression. We see that the estimated linear effect of log hours on the probability of all three distress outcomes is statistically significant. The loading on log hours tends to be smaller than the IV-probit estimated average marginal effect, indicating that nonlinearity plays a role.

[34]With the exception of the addition of the $250 billion asset bracket, asset brackets and base/marginal fee schedules prior to 2007 were stable over time, except for an annual inflation adjustment. Both inflation adjustments and rating surcharges were capped at $20 billion, prior to 2014, and at $40 billion thereafter.

E Additional figures and tables

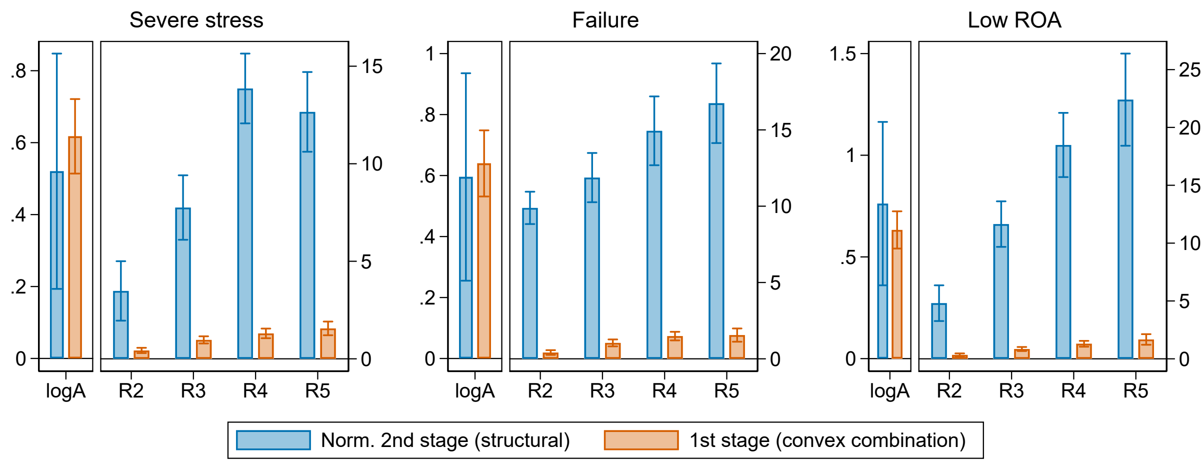


Figure A1: Comparison of coefficients on size and ratings from first and second stage. The figure shows the coefficients on log assets and the dummies for rating 2 to rating 5 from the second-stage equation (11) and the first-stage equation (12). The second stage coefficients are normalized by the second stage coefficient on log hours. The three panels correspond to the three variables measuring distress. Whiskers represent 95% confidence intervals computed based on standard errors clustered at the bank level via the delta method.

Table A1: OCC general assessment fee schedule. The table shows the OCC assessment fee schedule on federally chartered commercial banks and savings association as a function of asset size. Source: 12 CFR 8 and OCC bulletins.

If the amount of the total balance sheet assets (consolidated domestic and foreign subsidiaries) is: ($ millions)		The Semiannual Assessment will be:		
Year 2007				
Over	But Not Over	This Amount ($)	Plus	Of Excess Over ($ millions)
0	2	5,480	0	0
2	20	5,480	0.000227454	2
20	100	9,574	0.000181963	20
100	200	24,131	0.000118274	100
200	1,000	35,958	0.000100078	200
1,000	2,000	116,020	0.000081883	1,000
2,000	6,000	197,903	0.000072785	2,000
6,000	20,000	489,043	0.000061932	6,000
20,000	40,000	1,356,091	0.000050403	20,000
40,000		2,364,151	0.000033005	40,000
Year 2014				
Over	But Not Over	This Amount ($)	Plus	Of Excess Over ($ millions)
0	2	5,997	0	0
2	20	5,997	0.000236725	2
20	100	10,258	0.000189379	20
100	200	25,408	0.000123092	100
200	1,000	37,717	0.000104156	200
1,000	2,000	121,041	0.000085218	1,000
2,000	6,000	206,259	0.000075749	2,000
6,000	20,000	509,255	0.000064454	6,000
20,000	40,000	1,411,611	0.000048553	20,000
40,000	250,000	2,382,671	0.000033132	40,000
250,000		9,340,391	0.0000328	250,000

Table A2: OCC general assessment fees. The table shows estimates from linear regressions of log OCC general assessments on the listed controls. The fees are calculated for the universe of all federally chartered commercial banks that filed Call Reports in 2006:Q4 and 2013:Q4 using the fee schedule in Table A1 and rating surcharges discussed in Section C. Assets are actual, while ratings are generated from a uniform distribution. For detailed variable definitions see Section 3 and Appendix B. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Log(Fees)		
	(1)	(2)	(3)
Log(Assets)	0.697***	0.664***	0.663***
	[0.004]	[0.004]	[0.004]
Rating = 2	0.002	-0.002	-0.002
	[0.008]	[0.006]	[0.006]
Rating = 3	0.394***	0.395***	0.395***
	[0.007]	[0.006]	[0.006]
Rating = 4	0.702***	0.699***	0.699***
	[0.007]	[0.006]	[0.006]
Rating = 5	0.693***	0.689***	0.689***
	[0.007]	[0.005]	[0.006]
Post-2008		0.040***	0.041***
		[0.002]	[0.002]
Assets≥\$10b		0.353***	0.393***
		[0.033]	[0.032]
Post-2008×(Assets≥\$10b)		0.204***	0.091***
		[0.044]	[0.028]
Assets≥\$50b			-0.102*
			[0.058]
Post-2008×(Assets≥\$50b)			0.284***
			[0.091]
Constant	-0.031	0.118***	0.121***
	[0.024]	[0.022]	[0.021]
Adj. R^2	0.99	0.99	0.99
Obs.	2866	2866	2866
Distinct NAs	1772	1772	1772

Table A3: Variation of shadow cost across and within districts. The table shows the source of across- and within-district variation in the shadow cost instrument calculated as $-\overline{\delta}_H \overline{\log H}_{dt} + \overline{\delta}_A \overline{\log A}_{dt} + \sum_{r=2}^{5} \overline{\delta}_r \overline{\mathbb{I}[r]}_{dt}$, where $\overline{x}_{dt}$ denotes the average of variable x within district d and year t, and where $\overline{\delta}_H$, $\overline{\delta}_A$, and $\overline{\delta}_r$ are the coefficients on average log hours, log assets, and rating indicators from Table 3, column (1). The "assets/hours component" is $\overline{\delta}_A \overline{\log A}_{dt} - \overline{\delta}_H \overline{\log H}_{dt}$ and the "ratings component" is $\sum_{r=2}^{5} \overline{\delta}_r \overline{\mathbb{I}[r]}_{dt}$. Because the panel is unbalanced, the squared across and within standard deviations may not sum exactly to the squared overall standard deviation.

	Standard deviation		
	Overall	Across districts	Within districts
Total shadow cost	0.432	0.319	0.296
• Assets/hours component	0.416	0.313	0.280
• Ratings component	0.143	0.049	0.136
Observations	$N = 178$	$n = 12$	$\overline{T} = 14.83$

Table A4: Examination frequency requirements. The table shows examination frequency requirements by bank size, complexity, and rating. Sources: SR letter 13-21 for banks $<$ \$10 billion; Board policy statement, October 7, 1985 (as cited in BHC Supervision Manual, Section 5000.0.2) for banks $>$ \$10 billion.

		Rating 1 or 2	Rating 3, 4 or 5
$<$ \$10b	Non-complex	At least *every two years*: targeted off-site exam required every two years; additional follow-up and interim exams may be required.	At least *every year*: full-scope off-site exam required annually; additional follow-up and interim exams may be required.
	Complex	At least *every year*: full-scope exam required annually; additional follow-up and interim exams may be required.	At least *every year*: full-scope exam required annually; additional follow-up and interim exams may be required.
$\geq$ \$10b		At least *every year*: full-scope exam required annually; additional limited-scope or targeted exam presumed annually.	At least *twice every year*: full-scope exam required annually; one additional limited-scope or targeted exam required annually.

Table A5: First stage of IV probit with outcome variable 1y-ahead probability of severe stress. The table shows estimates from IV probit first-stage regressions of log hours on the listed controls and year fixed effects (corresponding second stages in Table 5). Other controls are noted at the bottom; abbreviations: "g" is assets≥\$10b, "pg" is post-2008×(assets≥\$10b), "lr" is lagged rating = 2,…,5. For detailed variable definitions see Section 3 and Appendix B. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Log(Hours)			
	(1)	(2)	(3)	(4)
Log Assets (real)	0.842***	0.834***	0.580***	0.618***
	[0.042]	[0.047]	[0.054]	[0.053]
Rating = 2	0.378***	0.483***	0.484***	0.429***
	[0.064]	[0.072]	[0.075]	[0.069]
Rating = 3	1.227***	1.365***	0.988***	0.970***
	[0.100]	[0.109]	[0.100]	[0.094]
Rating = 4	1.622***	1.694***	1.266***	1.302***
	[0.127]	[0.131]	[0.132]	[0.126]
Rating = 5	1.927***	1.874***	1.453***	1.559***
	[0.183]	[0.180]	[0.175]	[0.179]
District avg. Log Hours	0.731***			0.672***
	[0.055]			[0.052]
District avg. Log Assets	-0.768***			-0.728***
	[0.094]			[0.095]
District avg. (Rating = 2)	0.093			0.025
	[0.241]			[0.227]
District avg. (Rating = 3)	-0.526			-0.628*
	[0.392]			[0.367]
District avg. (Rating = 4)	-1.290*			-1.106
	[0.711]			[0.702]
District avg. (Rating = 5)	-2.065**			-2.047**
	[0.866]			[0.823]
District top		0.571***		0.454***
		[0.143]		[0.136]
High exam frequency			0.578***	0.468***
			[0.091]	[0.086]
Lagged exam count			-0.668***	-0.690***
			[0.114]	[0.110]
Lag exam ct.×(Hi. exam freq.)			0.776***	0.791***
			[0.117]	[0.113]
Other controls	g pg	g pg	pg lr	pg lr
Observations	4924	4924	4290	4290
Distinct BHCs	744	744	698	698

Table A6: First stage of IV probit with outcome variable 1y-ahead probability of failure. The table shows estimates from IV probit first-stage regressions of log hours on the listed controls and year fixed effects (corresponding second stages in Table 6). Other controls are noted at the bottom; abbreviations: "g" is assets≥\$10b, "lr" is lagged rating = 2,...,5. For detailed variable definitions see Section 3 and Appendix B. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Log(Hours)		
	(1)	(2)	(3)
Log Assets (real)	0.867***	0.585***	0.640***
	[0.049]	[0.061]	[0.055]
Rating = 2	0.300***	0.456***	0.411***
	[0.069]	[0.085]	[0.079]
Rating = 3	1.300***	1.053***	1.043***
	[0.120]	[0.122]	[0.116]
Rating = 4	1.777***	1.441***	1.487***
	[0.133]	[0.147]	[0.144]
Rating = 5	1.932***	1.439***	1.553***
	[0.217]	[0.214]	[0.222]
District avg. Log Hours	0.701***		0.668***
	[0.063]		[0.063]
District avg. Log Assets	-0.761***		-0.777***
	[0.105]		[0.104]
District avg. (Rating = 2)	0.154		-0.051
	[0.264]		[0.239]
District avg. (Rating = 3)	-0.242		-0.646
	[0.548]		[0.512]
District avg. (Rating = 4)	-1.090		-0.180
	[1.017]		[1.022]
District avg. (Rating = 5)	-2.788**		-2.694**
	[1.214]		[1.208]
High exam frequency		0.554***	0.463***
		[0.103]	[0.099]
Lagged exam count		-0.788***	-0.824***
		[0.135]	[0.136]
Lag exam ct.×(Hi. exam freq.)		0.914***	0.932***
		[0.139]	[0.140]
Other controls	g	lr	lr
Observations	3418	2860	2860
Distinct BHCs	715	645	645

Table A7: First stage of IV probit with outcome variable 1y-ahead probability of low ROA. The table shows estimates from IV probit first-stage regressions of log hours on the listed controls and year fixed effects (corresponding second stages in Table 6). Other controls are noted at the bottom; abbreviations: "g" is assets$\geq$\$10b, "pg" is post-2008$\times$(assets$\geq$\$10b), "lr" is lagged rating = 2,…,5. For detailed variable definitions see Section 3 and Appendix B. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Log(Hours)			
	(1)	(2)	(3)	(4)
Log Assets (real)	0.838***	0.827***	0.601***	0.633***
	[0.039]	[0.045]	[0.047]	[0.047]
Rating = 2	0.355***	0.456***	0.386***	0.342***
	[0.061]	[0.069]	[0.070]	[0.065]
Rating = 3	1.196***	1.333***	0.846***	0.846***
	[0.089]	[0.098]	[0.098]	[0.090]
Rating = 4	1.730***	1.781***	1.246***	1.298***
	[0.141]	[0.147]	[0.138]	[0.129]
Rating = 5	2.266***	2.168***	1.567***	1.672***
	[0.231]	[0.223]	[0.212]	[0.234]
District avg. Log Hours	0.734***			0.667***
	[0.054]			[0.051]
District avg. Log Assets	-0.777***			-0.705***
	[0.090]			[0.091]
District avg. (Rating = 2)	0.100			-0.031
	[0.231]			[0.217]
District avg. (Rating = 3)	-0.454			-0.457
	[0.387]			[0.360]
District avg. (Rating = 4)	-1.377**			-1.428**
	[0.694]			[0.700]
District avg. (Rating = 5)	-2.732***			-2.529***
	[0.888]			[0.858]
District top		0.645***		0.518***
		[0.139]		[0.128]
High exam frequency			0.602***	0.473***
			[0.095]	[0.085]
Lagged exam count			-0.602***	-0.623***
			[0.123]	[0.116]
Lag exam ct.×(Hi. exam freq.)			0.702***	0.713***
			[0.126]	[0.118]
Other controls	g pg	g pg	pg lr	pg lr
Observations	5405	5405	4717	4717
Distinct BHCs	750	750	678	678

Table A8: Detail on channels of supervisory effects. The table shows estimates from the second stages of IV probit regressions of 1y-ahead probability of tail realizations of balance sheet and income statement items on the listed controls and year fixed effects where log hours are instrumented for (using all instruments). Tail realizations are in the top ("H") or bottom ("L") 10th percentile of the distribution of the variable listed at the top of each column ("I" is income, "E" is expense). Other controls are noted at the bottom; abbreviations: "pg" is post-2008×(assets≥$10b), "h" is high exam frequency, "lr" is lagged rating = 2,…,5. For detailed variable definitions see Section 3 and Appendix B. The effective F-statistic and critical value are for the weak-instrument test of Olea and Pflueger (2013), robust to heteroskedasticity, autocorrelation, and clustering, from the respective first stage. Average marginal effects reported in curly braces. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	(1) Low ROA	(2) L non-int. I	(3) H non-int. E	(4) H LL prov.	(5) L net int. I	(6) L real. gains	(7) High NPL	(8) L tier-1 cap.
Log Hours	-0.147**	-0.316***	0.145*	-0.053	-0.198**	-0.093**	-0.114	0.020
	[0.059]	[0.076]	[0.077]	[0.080]	[0.091]	[0.046]	[0.071]	[0.082]
	{-0.020}	{-0.054}	{0.025}	{-0.007}	{-0.034}	{-0.015}	{-0.016}	{0.003}
Log Assets (real)	0.112*	0.107	-0.108	0.123*	0.380***	0.139***	0.136*	0.184**
	[0.062]	[0.081]	[0.073]	[0.072]	[0.083]	[0.046]	[0.071]	[0.084]
	{0.015}	{0.018}	{-0.018}	{0.016}	{0.066}	{0.023}	{0.019}	{0.027}
Rating = 2	0.709***	0.533***	0.032	0.730***	0.151	0.153	0.611***	0.522***
	[0.116]	[0.108]	[0.103]	[0.128]	[0.115]	[0.101]	[0.118]	[0.117]
	{0.097}	{0.091}	{0.005}	{0.096}	{0.026}	{0.026}	{0.087}	{0.076}
Rating = 3	1.717***	1.159***	0.183	1.180***	0.528***	0.226	1.338***	0.955***
	[0.148]	[0.130]	[0.157]	[0.167]	[0.166]	[0.151]	[0.145]	[0.167]
	{0.234}	{0.198}	{0.031}	{0.156}	{0.091}	{0.038}	{0.191}	{0.138}
Rating = 4	2.726***	1.574***	0.692***	1.795***	1.342***	0.460**	2.247***	1.910***
	[0.209]	[0.177]	[0.210]	[0.243]	[0.218]	[0.209]	[0.213]	[0.211]
	{0.372}	{0.269}	{0.118}	{0.236}	{0.231}	{0.077}	{0.321}	{0.277}
Rating = 5	3.304***	2.142***	0.928***	1.946***	1.961***	0.101	2.313***	2.639***
	[0.300]	[0.250]	[0.271]	[0.317]	[0.252]	[0.299]	[0.315]	[0.323]
	{0.451}	{0.366}	{0.159}	{0.256}	{0.338}	{0.017}	{0.330}	{0.382}
Other controls	pg h lr	pg h lr	pg h lr	pg h lr	pg h lr	pg h lr	pg h lr	pg h lr
F-statistic	38.5	37.4	37.4	38.6	39.2	38.8	39.2	38.7
Critical value	17.3	17.3	17.3	17.2	17.6	17.2	17.3	17.5
Observations	4717	4826	4820	4761	4843	4727	4779	4790
Distinct BHCs	678	684	681	686	689	694	687	685

Table A9: Robustness of IV probit: district fixed effects and additional risk controls. The table shows estimates from the second stages of IV probit regressions of distress probability (outcome variable noted at the top) on the listed controls and year fixed effects where log hours are instrumented for (all instruments). Additional risk controls are return on assets, non-performing loans ratio, and tier-1 capital ratio. Other controls are noted at the bottom; abbreviations: "pg" is post-2008$\times$(assets$\geq$\$10b), "h" is high exam frequency, "lr" is lagged rating = 2,...,5. For detailed variable definitions see Section 3 and Appendix B. The effective F-statistic and critical value are for the weak-instrument test of Olea and Pflueger (2013), robust to heteroskedasticity, autocorrelation, and clustering, from the respective first stage. Average marginal effects reported in curly braces. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	District fixed-effects			Risk controls			Both		
	Sev. stress$_{t+1}$	Failure$_{t+1}$	L ROA$_{t+1}$	Sev. stress$_{t+1}$	Failure$_{t+1}$	L ROA$_{t+1}$	Sev. stress$_{t+1}$	Failure$_{t+1}$	L ROA$_{t+1}$
Log Hours	-0.280***	-0.412***	-0.152**	-0.289***	-0.312***	-0.150**	-0.330***	-0.375***	-0.140*
	[0.066]	[0.070]	[0.071]	[0.069]	[0.097]	[0.061]	[0.070]	[0.089]	[0.075]
	{-0.026}	{-0.018}	{-0.020}	{-0.019}	{-0.008}	{-0.018}	{-0.024}	{-0.014}	{-0.016}
Log Assets (real)	0.175***	0.285***	0.121*	0.132*	0.096	0.081	0.182**	0.159*	0.072
	[0.067]	[0.085]	[0.071]	[0.074]	[0.079]	[0.064]	[0.074]	[0.096]	[0.077]
	{0.016}	{0.012}	{0.016}	{0.009}	{0.002}	{0.010}	{0.013}	{0.006}	{0.008}
Rating = 2	0.861***	3.562***	0.707***	0.705***	4.281***	0.475***	0.675***	3.977***	0.469***
	[0.190]	[0.237]	[0.117]	[0.222]	[0.202]	[0.125]	[0.216]	[0.191]	[0.127]
	{0.080}	{0.156}	{0.094}	{0.048}	{0.103}	{0.058}	{0.049}	{0.146}	{0.055}
Rating = 3	1.924***	4.276***	1.676***	1.468***	4.630***	1.096***	1.445***	4.294***	1.076***
	[0.208]	[0.337]	[0.151]	[0.275]	[0.495]	[0.180]	[0.275]	[0.398]	[0.184]
	{0.179}	{0.187}	{0.223}	{0.099}	{0.111}	{0.133}	{0.105}	{0.158}	{0.126}
Rating = 4	3.412***	5.472***	2.656***	3.044***	6.223***	1.734***	3.004***	5.939***	1.701***
	[0.228]	[0.453]	[0.213]	[0.373]	[0.791]	[0.332]	[0.379]	[0.797]	[0.345]
	{0.317}	{0.239}	{0.353}	{0.205}	{0.150}	{0.211}	{0.218}	{0.219}	{0.200}
Rating = 5	3.123***	6.023***	3.325***	2.798***	11.825***	1.186**	2.754***	11.147***	1.210**
	[0.261]	[0.538]	[0.311]	[0.538]	[1.014]	[0.584]	[0.541]	[0.987]	[0.583]
	{0.290}	{0.264}	{0.441}	{0.189}	{0.285}	{0.144}	{0.200}	{0.410}	{0.142}
Other controls	pg h lr	h lr	pg h lr	pg h lr	h lr	pg h lr	pg h lr	h lr	pg h lr
District FEs	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Risk controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	27.4	17.6	24.8	38.8	21.0	37.7	25.9	11.4	25.2
Critical value	17.2	17.0	17.7	17.1	18.8	17.6	17.4	18.9	18.0
Observations	4290	2292	4717	3821	2229	4375	3653	1607	4375
Distinct BHCs	698	526	678	673	544	664	648	399	664

Table A10: Robustness of IV probit: exclude largest banks ($\geq$ \$50 billion), exclude crisis years (2008–2009), and linear IV. The table shows estimates from the second stages of IV probit regressions of distress probability (outcome variable noted at the top) on the listed controls and year fixed effects where log hours are instrumented for (all instruments). Model and other controls are noted at the bottom; abbreviations: "pg" is post-2008×(assets≥\$10b), "h" is high exam frequency, "lr" is lagged rating = 2,…,5. For detailed variable definitions see Section 3 and Appendix B. The effective F-statistic and critical value are for the weak-instrument test of Olea and Pflueger (2013), robust to heteroskedasticity, autocorrelation, and clustering, from the respective first stage. Average marginal effects reported in curly braces. Standard errors clustered by bank reported in brackets; significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sample is 1998–2014.

	Exclude largest banks			Exclude crisis years			Linear IV		
	Sev. stress$_{t+1}$	Failure$_{t+1}$	L ROA$_{t+1}$	Sev. stress$_{t+1}$	Failure$_{t+1}$	L ROA$_{t+1}$	Sev. stress$_{t+1}$	Failure$_{t+1}$	L ROA$_{t+1}$
Log Hours	-0.242***	-0.318***	-0.138**	-0.200**	-0.321***	-0.137**	-0.014***	-0.004***	-0.016**
	[0.069]	[0.099]	[0.064]	[0.084]	[0.097]	[0.067]	[0.005]	[0.001]	[0.007]
	{-0.023}	{-0.009}	{-0.019}	{-0.013}	{-0.010}	{-0.015}			
Log Assets (real)	0.167**	0.363***	0.067	0.105	0.212**	0.101	0.009**	0.003**	0.013
	[0.075]	[0.106]	[0.075]	[0.076]	[0.090]	[0.071]	[0.004]	[0.001]	[0.008]
	{0.016}	{0.010}	{0.009}	{0.007}	{0.006}	{0.011}			
Rating = 2	0.866***	3.737***	0.702***	0.812***	3.727***	0.578***	0.033***	0.006*	0.072***
	[0.196]	[0.234]	[0.116]	[0.283]	[0.193]	[0.149]	[0.009]	[0.003]	[0.015]
	{0.081}	{0.105}	{0.096}	{0.051}	{0.114}	{0.062}			
Rating = 3	1.946***	4.441***	1.697***	1.918***	4.279***	1.647***	0.172***	0.017***	0.282***
	[0.214]	[0.367]	[0.152]	[0.320]	[0.567]	[0.191]	[0.022]	[0.007]	[0.029]
	{0.182}	{0.125}	{0.232}	{0.121}	{0.131}	{0.176}			
Rating = 4	3.527***	5.588***	2.659***	3.837***	5.516***	2.841***	0.695***	0.041**	0.613***
	[0.234]	[0.461]	[0.216]	[0.375]	[0.653]	[0.254]	[0.042]	[0.018]	[0.055]
	{0.330}	{0.157}	{0.364}	{0.243}	{0.169}	{0.304}			
Rating = 5	3.187***	6.315***	3.268***	4.031***	6.890***	3.328***	0.616***	0.073***	0.831***
	[0.265]	[0.529]	[0.302]	[0.410]	[0.723]	[0.338]	[0.064]	[0.028]	[0.087]
	{0.298}	{0.177}	{0.447}	{0.255}	{0.211}	{0.356}			
Model	IV probit	IV probit	IV probit	IV probit	IV probit	IV probit	Linear IV	Linear IV	Linear IV
Other controls	pg h lr	h lr	pg h lr	pg h lr	h lr	pg h lr	pg h lr	h lr	pg h lr
Exclude banks≥\$50b	Yes	Yes	Yes	No	No	No	No	No	No
Exclude 2008–2009	No	No	No	Yes	Yes	Yes	No	No	No
F-statistic	40.6	36.3	39.0	34.2	25.3	34.5	43.1	45.1	38.5
Critical value	16.7	17.2	17.4	16.5	17.0	17.3	16.4	16.7	17.3
Observations	3951	2633	4331	3470	2040	4078	4764	4764	4717
Distinct BHCs	666	612	645	691	618	673	704	704	678