NBER WORKING PAPER SERIES

THE ECONOMICS OF BANK SUPERVISION

Thomas M. Eisenbach
David O. Lucca
Robert M. Townsend

Working Paper 22201
http://www.nber.org/papers/w22201

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2016

For valuable comments, we thank Doug Diamond and Mark Flannery, as well as audience members at the New York Fed Conference on Supervising Large & Complex Financial Institutions. Emily Eisner provided excellent research assistance. Any errors are our own. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of New York, anyone else associated with the Federal Reserve System, or the National Bureau of Economic Research.

The Economics of Bank Supervision
Thomas M. Eisenbach, David O. Lucca, and Robert M. Townsend
NBER Working Paper No. 22201
April 2016
JEL No. D82,G21,G28

ABSTRACT

We study bank supervision by combining a theoretical model distinguishing supervision from regulation and a novel dataset on work hours of Federal Reserve supervisors. We highlight the trade-offs between the benefits and costs of supervision and use the model to interpret the relation between supervisory efforts and bank characteristics observed in the data. More supervisory resources are spent on larger, more complex, and riskier banks. However, hours increase less than proportionally with bank size, suggesting the presence of technological scale economies in supervision. The data also show reallocation of supervisory hours at times of stress and in the post-2008 enhanced supervisory framework for large banks, providing evidence of constraints on supervisory resources. Finally, we show theoretically limits to assessing supervisory success based on ex-post outcomes, as well as benefits of ex-ante commitment policies.

Thomas M. Eisenbach
Federal Reserve Bank of New York
33 Liberty St
New York, NY 10045
thomas.eisenbach@ny.frb.org

David O. Lucca
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
david.lucca@ny.frb.org

Robert M. Townsend
Department of Economics, E52-538
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
rtownsen@mit.edu

1 Introduction

Much policy debate in recent years has focused on bank supervision and regulation. In order "[t]o promote the financial stability of the United States," the 2010 Dodd Frank Act (DFA) introduced a number of provisions aimed in particular at the largest and most complex banking organizations.[1] While the academic literature has paid increasing attention to bank regulation, work on bank supervision remains limited, partly reflecting the opacity of supervisory activities that stems from supervisors' reliance on confidential information. In this paper, we first develop an abstract framework characterized by incentive problems and asymmetric information between banks and banking authorities. Observability and verifiability of bank information and actions distinguish regulatory from supervisory activities. Second, we use information from a unique dataset of hours spent by Federal Reserve bank examiners overseeing U.S. bank holding companies matched to banks' characteristics. We use the data to illustrate details of the supervisory process and interpret key empirical relations using the theoretical model.

Prudential supervision is closely related to, but distinct from, regulation of banking organizations. Regulation involves the development and promulgation of the rules under which banking organizations operate as well as their enforcement in the court of law. The distinguishing features of bank supervision are: (i) the assessment of the safety and soundness of banks through monitoring and exams, and (ii) the use of this information to request corrective actions from banks should their conditions or practices be deemed unsafe or unsound.

In our theoretical framework, the incentive problem between authorities and banks, which creates the need for regulation and supervision, is rooted in the bank's limited liability. Banks only value payoffs conditional on being solvent while banking authorities value asset payoffs in all states of the world (micro-prudential objectives) and take into account spillovers (macro-prudential, or financial stability, objectives).

In differentiating between regulation and supervision, we assume that regulation is "coarse" and can only be contingent on verifiable information while supervision is "discretionary" and can be contingent on non-verifiable information. Regulation can therefore restrict banks' activities ex-ante (for example, a ban on the trading of phys-

[1] The DFA contains a large number of regulatory changes both in the context of financial stability and consumer protection. In terms of financial stability, the DFA established the Financial Stability Oversight Committee (FSOC) with the authority to designate systemically important institutions. The DFA also mandated the Federal Reserve Board and other federal banking regulatory authorities to issue tighter regulations for the largest institutions related to the level of regulatory capital, proprietary trading and stress testing, among others.

ical commodities) and respond to asset payoffs ex-post (for example, the imposition of penalties after a breach of a capital requirement). Supervision can respond to interim signals about bank actions and influence asset payoffs before they are realized. In practice, a gray area can exist between supervision and regulation, as exemplified by stress testing and the fact that some supervisory activities involve compliance with regulations (see, for example Goldsmith-Pinkham, Hirtle, and Lucca, 2016). In its purest form, regulation is written into law and enforced through courts, and therefore requires verifiability (Maskin, Laffont, and Hildenbrand, 1982).

In our setting, a bank's balance sheet and returns are verifiable. Regulation can impose capital requirements that increase loss absorption and provide skin in the game ex-ante as well as provide discipline in risk management by imposing costs ex-post should the requirements be breached. In contrast to regulation, supervision can collect information about bank's actions through costly monitoring, leading to an interim signal, which in the data we proxy with the bank's supervisory rating. The interim signal is informative to banking authorities and allows improved outcomes compared to when only regulation is present. However, we assume that the signals are noisy and can prove to be incorrect. After accounting for the informativeness of the signal, supervision can respond to its realization and influence asset payoffs before they are realized. In practice these supervisory interventions are taken by communicating to banks "matters requiring attention" (MRAs) and "matters requiring immediate attention" (MRIAs) as discussed in Eisenbach et al. (2015). Although the supervisory signal is informative on average, it can point to the wrong interim action on occasions, in which case the supervisory intervention would also prove incorrect.

We assume that both supervisory monitoring and intervention are costly for the banking authority. The optimally chosen supervisory strategy therefore has to trade off benefits and costs of supervision. Even though more supervision may always be better when only considering its benefits, optimal supervision will be limited overall with higher intensity warranted only if it results in higher marginal benefits than marginal costs.

Given these trade-offs, our model predicts that intervention efforts are higher for riskier banks, which are those with a worse observed signal. Consistent with this predictions, we find empirically that Fed supervisors' efforts increase between 70 and 140 percent for banking organizations where the confidential supervisory rating indicates moderate to extreme levels of concerns (rating of 3, 4 or 5).

In light of the large size heterogeneity of U.S. banking organizations, we study in detail how supervisory efforts scale with bank size. Our model reveals that two opposite effects are at play. If the optimal intensity of supervision is independent of bank size, then technological scale economies of supervision would suggest that supervisory hours increase less than proportionally with bank size. However, the optimal intensity of supervision may increase with bank size, for example if spillovers are sufficiently large for large banks. Such increasing optimal intensity competes with potential technological scale economies such that supervisory hours may increase more than proportionally with bank size.

In the empirical analysis, we confirm that hours are increasing in bank size. Although, as discussed in the paper, we do not necessarily measure hours of all supervisory staff, the data at our disposal suggest that institutions with assets below a (trend-adjusted) threshold of $10 billion of consolidated assets are allocated about 100 hours per calendar quarter compared to about 1,500 for institutions above that asset threshold. While increasing with size, the elasticity of supervisory hours with respect to asset size is always close to or less than 1 across a number of different specifications. This includes the post-2008 subsample, which has been characterized by heightened concerns about spillovers from stress at large and complex banking institutions. This evidence points to the existence of economies of scale in bank supervision that are sufficiently strong to outweigh the effect of enhanced supervision for larger banks. The result also suggests that, in terms of realized hour allocations, banks in our sample do not appear to having grown to be "too large to be supervised." A direct implication of this finding is that aggregate supervisory costs would increase if large banks were broken up, although these costs are clearly not the only factor to be considered when assessing bank size from a financial stability perspective.

Because measurement of supervisory hours are imperfect, we compare estimates on Fed supervisory hours to those implied by schedule of assessment fees that the Office of the Comptroller of the Currency (OCC) charges banks it supervises. As stated by the OCC, the schedule reflects its estimated costs for supervising banks of different characteristics. Assessment fees are also increasing with a bank's supervisory rating and size, with an elasticity that is again less than 1, and sensitivities that are very close to those we obtain on Fed supervisory hour data.

In terms of monitoring technology, we find no evidence that efforts in the form of Fed hours are increasing in a bank's number of employees after controlling for its asset size, consistent with the model assumption that bank supervision is concerned with monitoring assets and processes as opposed to individual bank employees. Accordingly, we find evidence that supervisory efforts are increasing in bank complexity as measured by the number of legal entities within a given holding company. Finally, we observe that while for the largest institutions, supervisory efforts are stable over the course of a year, those at the smallest institutions oscillate with a peak around the time of annual full-scope examinations.

We next study the interaction between regulation and supervision by first considering, in the model, an exogenous shift in regulatory capital requirements. The model predicts that the direct effects of such a tightening of regulation is a decrease in supervisory monitoring (substitutes) but an *increase* in supervisory intervention (complements), leaving ambiguous the effect on overall supervisory efforts. We also consider an exogenous increase in the size of bank spillovers in the banking authority's preferences, which arguably are a better characterization of the post-2008 framework of higher capital requirements and enhanced supervision, and find that in this case both supervisory monitoring and intervention increase.

In the data, we study the enhanced regulatory and supervisory framework in the post-2008 sample period, where we find that both capital requirements and supervisory hours increased for the largest banking organizations. Although as noted above the elasticity of hours with respect to assets before and after 2008 is quite similar, our point estimates suggest a post-2008 level increase in average hours of about 60% at the larger institutions—after controlling for size and ratings. In contrast, we estimate a decline in the average hours spent by supervisors at the smallest institutions that we interpret as evidence of substitution due to limited available resources.

When we extend the theoretical framework to multiple banks, it predicts that—within each size group—supervisory efforts are reallocated from banks with good signals to those with bad signals. In addition, there will be significant reallocation of hours from small banks to large banks in times when bad signals are clustered at larger banks since implementing the same intensity of supervision at a large bank requires more hours. Accordingly, we find in the data that when the fraction of bank assets in distress (rating of 3, 4 or 5) in a Fed district increases, supervisory hours are allocated away from smaller institutions.

Finally, we use our model to understand how supervisory efforts can be evaluated based on ex-post outcomes. In particular, we show that based on only observing a bank default it is not obvious whether bank misbehavior and/or supervisory mistakes are a more likely cause than the inherent risk of a bank's activities. In addition, we discuss issues of supervisory incentives as well as time consistency problems that may impede the effectiveness of supervision both ex-ante and ex-post.

Related literature: The existing literature typically derives the need for supervision and regulation of banks from the special nature of banks as opposed to other firms and the resulting frictions. As for any debt-financed firm, a bank's limited liability raises basic moral hazard issues (Jensen and Meckling, 1976). However, there are limits to the ability of markets to provide the necessary discipline (Flannery, 1998; Rochet, 2004). The main debt of banks is in the form of deposits which impedes the functioning of market discipline through various channels. For example, the sequential service constraint inherent in deposit contracts creates strategic complementarities among depositors that can lead to inefficient runs (Diamond and Dybvig, 1983). To address the problem of runs, depositors are insured which eliminates their incentive to monitor the bank and creates a need for supervision/regulation (for example Mishkin, 2001). Alternative approaches view supervision/regulation as more effective monitoring than that of small, dispersed and uninformed depositors (Dewatripont and Tirole, 1994) or point to market incompleteness specific to banks' business models as a rationale for supervision/regulation (Merton, 1995). Since the financial crisis of 2008-09, the literature has focused more on macro-prudential as opposed to micro-prudential regulation, meaning not only on the default risk of an individual bank in isolation but also taking into account "systemic" effects such as spillovers of defaults on other banks and the economy more widely (for example Acharya, 2009; Brunnermeier et al., 2009; Tirole, 2013).

Few papers in the literature explicitly distinguish between supervision and regulation or focus on supervision. Supervision then takes on a number of different roles: auditing bank asset values to detect breaches of capital requirements (Rochet, 2007); preventing the bank from taking observable but non-verifiable actions (Dewatripont and Tirole, 1994); incentivizing the bank to take the right non-observable actions through punitive interference after bad, verifiable outcomes (Marshall and Prescott, 2001, 2006); incentivizing proper risk taking and information disclosure through ex-post intervention (Harris and Raviv, 2012); and screening different risk-types of banks by offering a menu of admissible combinations of risk taking and required bank capital (Prescott, 2004). Motivated by the inclusion of supervision as a pillar in the Basel framework, some papers study the interaction of supervision and regulation. For example, Bhattacharya et al. (2002) study the optimal combination of random supervisory exams and closure rules conditional on exam results; Decamps et al. (2004) have supervisors choosing intervention thresholds to maintain adequate incentives for bank risk taking and study the effects of ex-post liquidity assistance and forbearance.

Some empirical work studies the information produced by bank examinations, for example, when compared to off-site monitoring (Cole and Gunther, 1995), conditional on positive/negative findings (Berger and Davies, 1998) or relative to publicly available information (Hirtle and Lopez, 1999). In addition, there is a small literature taking on the challenge of identifying effects of supervision on bank outcomes (for example Hirtle, Kovner, and Plosser, 2016, and references therein).

The rest of the paper is organized as follows. Sections 2 and 3 describe the theoretical model and the data, respectively, that are used throughout the paper. Section 4 studies the main determinants of supervisory attention. Section 5 investigates the interactions of supervision and regulation and how they change over time. Section 6 focuses on the issues of supervising multiple banks with limited resources. Section 7 discusses how supervision can be evaluated based on ex-post outcomes. Section 8 concludes.

2 Model

We first propose a general setup in Section 2.1 that contains a range of elements of supervision and regulation that can be interpreted through an economic lens. The purpose of explicitly specifying a model is to translate the features of supervision and regulation in practice into concepts familiar to economists from theories of asymmetric information and contracting.

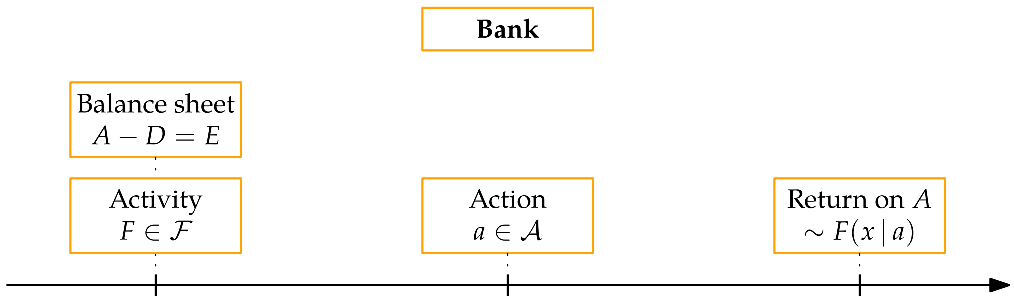
We propose an abstract framework with an incentive problem between banks and the authority responsible for regulation and supervision where there is asymmetric information and issues of observability vs. verifiability that allow us to meaningfully distinguish regulation from supervision.

We then specialize the setup in Section 2.2 to focus on supervision rather than regulation and make several simplifying assumptions for analytical tractability.

2.1 General setup

We propose a general setup involving both supervision and regulation. In particular, we define what a bank is, what projects it has and what actions it can take, and how those affect its payoff. We then define the banking authority's payoff and how it differs from the banks', providing a rational for supervision/regulation. Finally, we distinguish supervision and regulation as two sets of tools that can affect the bank in different ways.

Figure 1: **Model timeline without supervision or regulation**



Our model is essentially static and is meant to capture the interaction of banks and supervisors over a certain period, for example one year. We therefore focus on a sequence of events within a single period and leave the analysis of dynamics over multiple periods for future research. We first describe the primitives of the model without supervision/regulation.

Bank: There is a set of banks $i = 1, \ldots, I$ that start the period with assets A_i, debt D_i and equity E_i. Bank i has available a range of activities $F_i \in \mathcal{F}_i$ it can engage in where F_i represents the c.d.f. of gross return x_i on the bank's assets in activity F_i.[2] We think of activities very broadly, such as making loans, trading securities, writing derivatives and subdivisions and combinations thereof. Once the activity F is in place, the bank has to take an action $a \in \mathcal{A}$ that influences the distribution of its asset return x, that is $F(x \mid a)$. We think of the action a as akin to risk management, that is reducing the variance of x, and the relevant set of actions $\mathcal{A}$ could depend on the activity F.

The asset return x is realized at the end of the period and determines whether the bank is solvent or not. Figure 1 illustrates the timeline without supervision or regulation. Given assets A and debt with face value D, a bank is solvent at the end of the period if its equity is positive, $Ax - D \geq 0$, and insolvent otherwise. We denote by $U(c, a)$ the bank's utility which depends on its monetary payoff c at the end of the period as well as its action a, for example through effort costs. The bank's expected payoff in the absence

[2] To simplify notation, we drop the index i for a specific bank unless explicitly analyzing multiple banks.

of supervision or regulation is therefore:

$$\int_{-\infty}^{\frac{D}{A}} U(0,a)\,dF(x\,|\,a) + \int_{\frac{D}{A}}^{\infty} U(Ax-D,a)\,dF(x\,|\,a)$$
$$\text{with} \quad F \in \mathcal{F},\ \ a \in \mathcal{A} \tag{1}$$

Note that we assume that while the bank's monetary payoff is 0 in case of default, the effort cost has to be borne irrespective of the bank's solvency since the action a is taken before the realization of x. By not distinguishing between the bank's shareholders and managers, we abstract away from potential incentive problems within the bank.

Banking authority: The banking authority can differ from the banker in three ways. First, while the bank only values its cash flow conditional on being solvent, the banking authority values payoffs to the bank *and* its creditors, $(Ax - D) + D = Ax$. This component represents micro-prudential objectives, in the sense that the authority cares about the financial health, or safety and soundness, of a single bank.[3] Second, the banking authority also cares financial stability per se, representing macro-prudential objectives, in that it takes into account spillovers from banks to the rest of the economy. We capture this in reduced form with positive spillovers $P(A) > 0$ from a bank operating successfully and negative spillovers $-N(A) < 0$ of a bank failing. Note that spillovers depend on the bank's size A since a larger bank with more assets causes more spillovers. Finally, the banking authority could have a different utility $V \neq U$ over payoffs, for example with more risk aversion or with more weight on tail outcomes. The banking authority's expected payoff in the absence of supervision or regulation is therefore:

$$\int_{-\infty}^{\frac{D}{A}} V\big(Ax - N(A)\big)\,dF(x\,|\,a) + \int_{\frac{D}{A}}^{\infty} V\big(Ax + P(A)\big)\,dF(x\,|\,a) \tag{2}$$

Regulation and supervision: For the purposes of our model, we link the distinction between regulation and supervision to the difference between verifiability and observability.[4] In its purest form, regulation is written into law and enforced through courts,

[3]In general equilibrium and with complete markets, the pricing of the bank's debt would force it to take into account the creditor's payoff. We follow the banking literature in assuming that markets are incomplete, for example due to deposit insurance or coordination problems among small creditors (see the discussion in the literature review).

[4]See Maskin, Laffont, and Hildenbrand (1982) for a discussion of the concepts of verifiability and observability.

and therefore requires *verifiability*. In our model, we assume that the bank's initial balance sheet variables A, D, E are verifiable so regulation can impose ex ante constraints on the balance sheet, such as a capital requirement ρ requiring a minimum ratio between equity capital and assets, $E/A \geq \rho$. In economic terms, this provides both loss absorption capacity and incentives through "skin in the game." We further assume that the activities a bank is engaged in are verifiable so regulation can limit ex-ante the available activities to a subset $\widetilde{\mathcal{F}} \subset \mathcal{F}$, for example allow lending and trading but not insurance. The 1933 Glass-Steagall Act and the Volcker Rule in the 2010 DFA are examples of such broad regulatory restrictions. In economic terms, this imposes some bounds on the risks available but can't rule out risk entirely. Finally, we assume that the realization of the bank's asset return x at the end of the period is verifiable so regulation can be active ex-post, contingent on the realization of x. Since the capital requirement ρ has to be satisfied at the end of the period as well, or $Ax - D \geq \rho Ax$, it is possible for the bank to violate the capital requirement without defaulting, $Ax - D \in (0, \rho Ax)$. In this case, regulation can impose ex-post costs R on the bank even if it remains solvent. In economic terms, the combination (ρ, R) provides incentives through a disciplining effect.[5]

In contrast to regulation, we assume supervision has discretion and only requires *observability*.[6] We assume that while none of the bank's possible actions $a \in \mathcal{A}$ are verifiable, some are observable by supervisors while some are neither verifiable nor observable. For example, supervisors may be able to observe if a bank's risk managers participate in important meetings but not whether they are diligent in raising concerns. The supervisors can therefore prevent risk managers from not doing their work at all but still have to rely on incentives to ensure the work is done properly. Formally, this means that supervision can prevent some observable actions, that is restrict the action set to some subset $\widetilde{\mathcal{A}} \subset \mathcal{A}$. In economic terms, this makes for easier incentivizing by eliminating some "tempting" actions.[7]

In areas where even observability fails, we assume that supervision can collect infor-

[5] In principle, regulation could condition on x more generally, that is with an ex-post cost/transfer $R(x)$. We abstract away from such issues since they have already been explored in the literature on regulation.

[6] An important part of supervision is to ensure compliance with regulations (see, for example Rochet, 2007; Eisenbach et al., 2015; Goldsmith-Pinkham et al., 2016). We focus on the elements of supervision that are independent from regulation.

[7] In principle, it is not clear where to draw the line between an activity F and an action a. For the purposes of our model, we implicitly draw the line between verifiability and observability. Everything verifiable corresponds to F while everything observable (and unobservable) corresponds to a.

mation through monitoring effort s_1, for example an on-site examination that involves gathering information and distilling it into a supervisory rating. We can think of this rating as an interim signal r that is informative about the bank's action a and therefore its riskiness, the distribution of the final return x conditional on a. In economic terms, to the extent that the signal is a sufficient statistic for a with respect to x, supervision is more efficient in incentivizing the bank than regulation is contingent on final cash-flow x.

We then assume that after observing the signal r about the bank's action a but before the realization of the return x, supervision can react to the interim information with supervisory actions s_2. For example, most Fed supervisory actions today are accounted for by "matters requiring attention" and "matters requiring immediate attention," or MRAs and MRIAs, through which which Fed supervisors mandate banks remediation of issues or activities that have been deemed unsafe and unsound.[8] In terms of the model, these remedial actions affect the distribution of the final payoff, $F(x \mid a, s_2)$, in a way similar to a, that reduces the variance of x. In addition, we assume that supervisory intervention s_2 also affects the bank's utility directly, $U(z, a, s_2)$, for example because the supervisory remedies involve extra work for the bank. Importantly, both supervisory monitoring s_1 and supervisory intervention s_2 are costly and therefore affect the banking authority's utility, $V(z, s_1, s_2)$. Figure 2 illustrates the timeline with supervision and regulation.

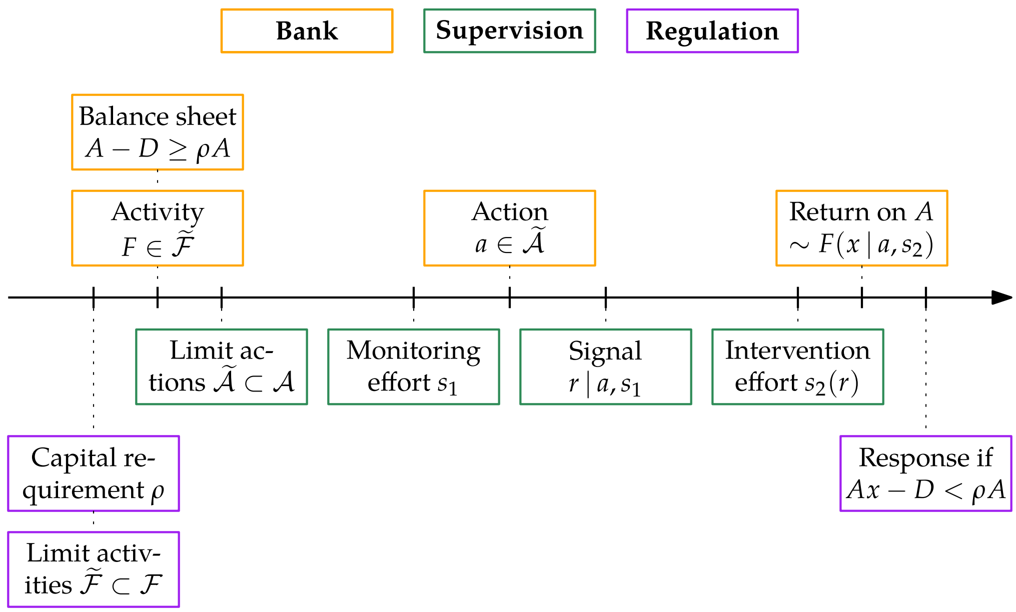
Taking into account the elements of regulation and supervision, the expected payoffs of a bank and the banking authority, (1) and (2), respectively, change as follows. For the bank, the expected payoff becomes

$$\begin{aligned}
\int_{-\infty}^{\infty} \Bigg(& \int_{-\infty}^{\frac{D}{A}} U\big(0, a, s_2(r)\big)\, dF\big(x \mid a, s_2(r)\big) \\
& + \int_{\frac{D}{A}}^{\frac{D}{(1-\rho)A}} U\big(Ax - D - R, a, s_2(r)\big)\, dF\big(x \mid a, s_2(r)\big) \\
& + \int_{\frac{D}{(1-\rho)A}}^{\infty} U\big(Ax - D, a, s_2(r)\big)\, dF\big(x \mid a, s_2(r)\big) \Bigg) \, d\kappa(r \mid a, s_1) \\
\text{with} \quad & E \geq \rho A, \;\; F \in \widetilde{\mathcal{F}}, \;\; a \in \widetilde{\mathcal{A}},
\end{aligned}$$

where $s_2(r)$ denotes the supervisory intervention conditional on a rating r and $\kappa(r \mid a, s_1)$

[8] Beside MRAs and MRIAs, supervisors have a number of different supervisory actions at their disposal ranging from informal to formal actions (see Goldsmith-Pinkham et al., 2016, for example).

Figure 2: **Model timeline with supervision and regulation**



denotes the c.d.f. of the rating r conditional on the bank's action a and the supervisory monitoring effort s_1. Comparing this expression to the one in (1), we note several differences: (i) constraints on the ex-ante balance sheet, set of activities and set of actions, (ii) costs for breaching the capital requirement ex-post, (iii) supervisory monitoring leading to a rating, and (iv) return distribution and utility affected by supervisory intervention conditional on rating.

For the banking authority, the expected payoff becomes

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\frac{D}{A}} V\big(Ax - N(A), s_1, s_2(r)\big)\, dF\big(x \,\big|\, a, s_2(r)\big)\right.$$
$$\left. + \int_{\frac{D}{A}}^{\infty} V\big(Ax + P(A), s_1, s_2(r)\big)\, dF\big(x \,\big|\, a, s_2(r)\big)\right) d\kappa(r \mid a, s_1)\, d\Delta(a),$$

where $\Delta(a)$ denotes the c.d.f. of the bank's unobservable action $a \in \widetilde{\mathcal{A}} \subset \mathbb{R}$ from the point of view of the banking authority. We assume that the penalty R is welfare-neutral.

Of course, many more elements could be added to the model. For example, we could assume that regulation is able to impose a penalty on the bank even in case of default. Alternatively, we could consider prevention of default through a bailout that benefits the bank's existing management implying a positive payoff in case of insolvency. More generally, there is a big gray area in practice between what is regulation and what is supervision. For example stress testing has elements of regulation as it specifies scenarios and clear cutoffs for banks' simulated performance. On the other hand, stress testing has elements of supervision as it involves qualitative assessments of banks' models and processes. For the purposes of this paper, we try to strike a balance between addressing what we consider key elements of supervision and maintaining a tractable model that allows for clear intuition about the main mechanisms.

2.2 Simplifying assumptions

To keep the analysis tractable, we make several simplifying assumptions about the general setup laid out in the previous section. First, we assume that both the bank and the supervisor are risk neutral and that their utilities are additively separable between payoffs and effort costs.[9] For the bank, we assume that its action a and the supervisor's intervention s_2 have effort costs C^a and C^{s_2}, respectively, so that

$$U(c, a, s_2) = c - C^a(a, A) - C^{s_2}(s_2, A).$$

Similarly, for the supervisor, we assume that its monitoring s_1 and intervention s_2 have effort costs $\mathcal{C}^1$ and $\mathcal{C}^2$, respectively, so that

$$V(c, s_1, s_2) = c - \mathcal{C}^1(s_1, A) - \mathcal{C}^2(s_2, A)$$

Note that the actions a and s_2 affect the distribution of the bank's gross return x, which does not depend on bank size. Similarly, s_1 serves to generate a signal about a. We therefore account for the effect of bank size on costs by including A as an explicit argument of the cost functions for the bank and the supervisor. For example, diligent risk management, measured by its effect on the return distribution, is more costly for assets of \$10 billion than for assets of \$1 billion. Since our data is on supervisory hours, it would be natural to limit the supervisor's costs, such as those of monitoring, to the hours necessary to achieve monitoring intensity s_1 on assets of size A and an hourly wage w, or $\mathcal{C}^1(s_1, A) = w \times h^1(s_1, A)$. However, we prefer to maintain general cost functions $\mathcal{C}^1$ and $\mathcal{C}^2$ that can also include non-monetary costs, for example discomfort experienced by the supervisor when having to confront a bank.

[9]Since our focus is on supervision, we will refer to the banking authority as the supervisor. As we discuss more below, this abstracts somewhat but not fully from incentive problems on the supervisor's part—as in "who monitors the monitors" (Hurwicz, 2007).

Next, we assume that the bank's action a and the supervisory intervention s_2 affect only the riskiness of the bank's return x but not its mean, that is they correspond to risk management. Denoting partial derivatives by subscripts, this means for $m = a, s_2$ that $F_m(x \mid a, s_2) < 0$ if $x < \mathrm{E}[x]$ and $F_m(x \mid a, s_2) > 0$ if $x > \mathrm{E}[x]$, and therefore $F_{xm} > 0$. In addition, since F is bounded between 0 and 1, we assume that both a and s_2 can reduce the bank's probability of default but at a decreasing rate, $F_{mm}(x \mid a, s_2) > 0$ if $x < \mathrm{E}[x]$ and $F_{mm}(x \mid a, s_2) < 0$ if $x > \mathrm{E}[x]$. Lastly, we assume that a and s_2 are substitutes in affecting F so the sensitivity of the default probability to supervisory intervention is higher if the bank's risk management effort is lower, $F_{as_2} > 0$.

While the supervisor is able to observe whether risk management is taking place at all—and can ensure that it is—it is not possible to perfectly observe whether risk management is too lax or not. We therefore assume that based on observability, the supervisor effectively limits the choice set $\mathcal{A}$ for the bank's action a to two actions, $\{\underline{a}, \overline{a}\} = \widetilde{\mathcal{A}} \subset \mathcal{A}$, where the high action $\overline{a}$ is the "good" action, proper risk management, while $\underline{a}$ is the "bad" action, lax risk management.

Supervisory monitoring s_1 can provide some information about which action a the bank chooses. Given the binary choice set $\{\underline{a}, \overline{a}\}$ for the bank, we assume that supervisory monitoring s_1 produces a binary signal (the supervisory rating) $r \in \{\underline{r}, \overline{r}\}$. Importantly, the signal is not perfect but has precision λ with respect to the bank's action:

$$\Pr[\overline{r} \mid \overline{a}] = \Pr[\underline{r} \mid \underline{a}] = \lambda$$

The supervisory monitoring effort s_1 influences the precision λ; for simplicity, we assume that the supervisor can choose the precision directly ($\lambda = s_1$).

Next, we make some intuitive assumptions about the shape of the cost functions $\mathcal{C}^1$ and $\mathcal{C}^2$. First, we assume that both are increasing and convex in the supervisory effort, that is $\mathcal{C}_s > 0$ and $\mathcal{C}_{ss} > 0$. For monitoring s_1, a more precise signal is costly and increasingly so since precision λ is bounded above by 1. For intervention s_2, reducing risk is costly and increasingly so since the variance of x is bounded below by 0. Second, we assume that costs are increasing and weakly concave in bank size, or $\mathcal{C}_A > 0$ and $\mathcal{C}_{AA} \leq 0$. Achieving the same signal precision or risk reduction for a larger bank incurs more costs. However, costs don't grow more than proportionally with size as reflected by the fact that for example any bank has at most one chief risk officer or that larger banks tend to have larger and not just more loans.[10] This assumption implies

[10]See Kovner, Vickery, and Zhou (2014) for evidence of similar patterns in banks' own operating costs.

technological scale economies in supervision. Third, we assume that the cross-partial derivative is positive, that is $\mathcal{C}_{As} > 0$. Here it is important to remember that s_1 and s_2 are scale-free: The total cost of achieving a given "intensity" of monitoring or risk reduction is higher for banks with more assets, $\mathcal{C}_A > 0$; therefore, the total cost of increasing the intensity is naturally higher for banks with more assets, $\mathcal{C}_{As} > 0$.

Finally, in an effort to limit the number of moving parts in the analysis, we don't explicitly consider banks' choice of initial balance sheet A, D, E or its choice of activity F. For simplicity we assume that banks' initial balance sheet composition is scale invariant by setting leverage $D/A \equiv \ell$ for all A.

3 Data description

Four data sources are used in the empirical analysis of this paper: hours data for supervisory staff at the Federal Reserve, examination information and supervisory rating information from the National Examination Database (NED), financial data for domestic bank holding companies (BHCs) and fee-schedule information of the Office of the Comptroller of the Currency (OCC). We discuss each data source and then present summary statistics for the variables included in the regressions.

We obtain information on hours spent by Federal Reserve supervisory staff from an internal database tracking information on staff activity and time allocation. The Federal Reserve supervises state member banks (SMBs) and all bank holding companies on a consolidated basis. In our analysis we only include BHCs. The hours data starts in 1998 and end in 2014. The information is self-reported by supervisory employees on a weekly basis, and the database aims at recording each employee's time allocation across activities ranging from monitoring banking institutions through examination or other activity, and other administrative and support responsibilities. Reported monitoring activity includes information on the monitored institution through its regulatory entity number (or RSSD ID).

In practice, however, hours information is not always complete. First, the Federal Reserve System consists of 12 distinct districts and the Board of Governors, and reporting practices across districts may differ. We account for these differences with district fixed effects. Second, we access information on hours pre-2000 for only a handful of districts, and have information on all districts only starting in 2006. Finally, supervisory work at the smallest institutions is often recorded using a generic bank portfolio assignment, as opposed to an institution RSSD number. By cross-checking hours information with independent information on the timing of supervisory inspection from NED, it appears that consistent monitoring information with valid supervised-entity information is only available for institutions with assets of about $750 million or more; we therefore only include institutions with at least $1 billion in assets. For each institution, we aggregate data by quarter, so that the resulting supervisory hours data is a dataset uniquely identified by a quarter and the supervised institution's RSSD ID.

We match hours information to two other data sources. First, we obtain information on bank characteristics, such as size, share of loans to total assets and employee counts, from public FR Y-9C reports, which are used to assess and monitor the financial condition of holding company organizations on a consolidated basis. The minimum asset requirement for Y-9C reporting increased from $150 million earlier in the sample to $500 million in 2006:Q1, but as discussed above we only include banks with assets greater than $1 billion.

In addition, we match supervisory hours to confidential rating information. Bank holding companies are assigned a 1-to-5 rating under the "RFI/C(D)" rating system, with lower numbers indicating fewer issues, and thus a better rating. Banks with a rating of 1 or 2 are considered in satisfactory condition and present few significant supervisory concerns. Banks with a 3, 4, or 5 rating present moderate to extreme levels of regulatory concerns. The letters in the rating system indicate different components considered in the rating assignment—"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 weighing 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 2004, BHCs received supervisory ratings known as BOPECs, an acronym which stood for five areas of supervisory concern: conditions of the BHC's bank subsidiaries, other nonbank subsidiaries, parent company, earnings, and capital adequacy. BOPECs were also assigned on a 1-to-5 scale, and BOPECs and RFI/C(D) rating levels have similar supervisory interpretations. As a result, we splice these measure together, but we include time effects in the regressions to account for possible changes in levels pre and post rating changes.

Finally, we use data on supervisory fees assessed on federally chartered commercial banks by the OCC. As the hours data discussed above can be noisy, we use this information to assess validity of the main findings obtained with the hours data. 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. Similar to the Federal Reserve, the OCC is a self-funded agency that does not depend on the congressional appropriation process. But while the Federal Reserve funds its supervisory activities through net interest margins earned on its securities portfolio (or seigniorage), the OCC 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. We obtain data on the fee schedule from the OCC's public website.[11]

Table 1 provides summary statistics for the variables included in the regression specifications by whether banks' assets are above or below a $10 billion asset threshold adjusted for the growth rate of industry assets. As shown in the first column of the table, on average, Fed supervisors allotted about 490 hours per calendar quarter to supervising a holding company in our sample. In terms of an 8-hour work day, and 63 work days per quarter, these hours can be converted into about a full-time examiner being assigned to the holding company. But large institutions have on average three times as many hours assigned (column 3), while smaller ones about a fifth.[12] As we will see in the next sections, hours at smaller institutions tend to also be heavily concentrated at times of full-scope examinations.

4 Supervisory monitoring and intervention

We first study the determinants of optimally chosen supervisory efforts s_1 and s_2 using our model and then study the key relationships in the data on Fed supervisory hours.

4.1 Theoretical analysis

We first study the determinants of supervisory efforts s_1 and s_2 at an individual bank. Analyzing our model by backwards induction, we start with the supervisory interven-

[11] 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.

[12] This calculation excludes hours (that may be substantial) that have not been booked by the examiner to a specific institution. In addition, the day-count translation would underestimate an actual headcount because it abstracts from time-off and other administrative or training activities that an examiner may be involved when not assigned to a bank.

tion s_2 conditional on the observed signal r. Then we proceed to the bank's choice of a given the supervisory monitoring s_1 and intervention strategy $s_2(r)$. Finally, we study the supervisory monitoring s_1 that determines the precision of the signal, taking into account the bank's response. Our subgame perfect approach imposes time consistency on the supervisory actions. As we discuss in Section 7.2, supervisors would be better off ex-ante if they were able to commit to a strategy that is suboptimal ex-post and therefore time-inconsistent.

Supervisory intervention: The supervisory intervention s_2 has to be conditional on the signal r received about the bank's action a. Given the signal precision λ, the probability that the bank took the bad action $\underline{a}$, and therefore has high risk, conditional on the signal realization $r \in \{\underline{r}, \overline{r}\}$ is given by Bayes' rule. We denote this conditional probability by $\kappa \in \{\underline{\kappa}, \overline{\kappa}\}$:

$$\Pr[\underline{a} \,|\, \overline{r}] = \frac{(1-\lambda)\Pr[\underline{a}]}{\lambda\left(1-\Pr[\underline{a}]\right) + (1-\lambda)\Pr[\underline{a}]} =: \overline{\kappa}$$

$$\Pr[\underline{a} \,|\, \underline{r}] = \frac{\lambda \Pr[\underline{a}]}{\lambda \Pr[\underline{a}] + (1-\lambda)\left(1-\Pr[\underline{a}]\right)} =: \underline{\kappa}$$

For a given κ, the supervisor's expected payoff when choosing s_2 is given by:

$$\begin{aligned}&(1-\kappa)\left(\mathrm{E}[Ax] + \left(1 - F(\ell \,|\, \overline{a}, s_2)\right) P(A) - F(\ell \,|\, \overline{a}, s_2)\, N(A)\right)\\ &\quad + \kappa\left(\mathrm{E}[Ax] + \left(1 - F(\ell \,|\, \underline{a}, s_2)\right) P(A) - F(\ell \,|\, \underline{a}, s_2)\, N(A)\right)\\ &\quad - \mathcal{C}^2(s_2, A)\end{aligned}$$

For either action $a \in \{\underline{a}, \overline{a}\}$, the supervisor payoff is the expected asset return plus the expected spillover, which is positive if the bank survives and negative if the bank defaults. The supervisor's first order condition for the optimal s_2^* is then:

$$-\left((1-\kappa)\, F_{s_2}(\ell \,|\, \overline{a}, s_2^*) + \kappa F_{s_2}(\ell \,|\, \underline{a}, s_2^*)\right)\left(P(A) + N(A)\right) = \mathcal{C}^2_{s_2}(s_2^*, A) \tag{3}$$

The left-hand side of equation (3) is the marginal benefit of intervention s_2: Since $\ell < 1 < \mathrm{E}[x]$, we have $F_{s_2}(\ell \,|\, a, s_2) < 0$, that is supervisory intervention reduces the probability of default; for the marginal benefit, the reduction in expected default probability is multiplied by the total spillover losses of default vs. non-default, $P(A) + N(A)$. The right-hand side is the marginal cost of intervention s_2 at a bank of size A. As discussed in Section 2.2, the costs $\mathcal{C}^2(s_2, A)$ certainly include the hours $h^2(s_2, A)$ necessary to achieve a level s_2 of intervention at a bank of size A. In addition, the costs (and therefore marginal costs) could include, for example the non-monetary discomfort experienced by the supervisor when imposing intervention measures that the bank disagrees with. If these non-monetary marginal costs are high, the trade-off in the first-order condition (3) leads to low levels of intervention—just as if the monetary costs are high or supervisory resources scarce. Since our data is on supervisory hours, we assume that total costs are proportional to hours, $\mathcal{C}^2(s_2, A) \propto h^2(s_2, A)$.

We are interested in the effect of perceived riskiness and of bank size on supervisory hours spent on intervention. Since κ is the conditional probability of high risk, the effect of riskiness on hours is given by:

$$\frac{d}{d\kappa} h^2(s_2^*, A) \propto \frac{d}{d\kappa} \mathcal{C}^2(s_2^*, A) = \mathcal{C}^2_{s_2}(s_2^*, A) \frac{ds_2^*}{d\kappa} \tag{4}$$

Since $\mathcal{C}^2_{s_2} > 0$, the sign is determined by that of $ds_2^*/d\kappa$, that is how the intensity of intervention depends on the posterior about the bank's risk. Implicit differentiation of the first-order condition (3) yields:

$$\frac{ds_2^*}{d\kappa} = \frac{\left(F_{s_2}(\ell \,|\, \bar{a}, s_2^*) - F_{s_2}(\ell \,|\, \underline{a}, s_2^*)\right) \left(P(A) + N(A)\right)}{\left((1-\kappa)\, F_{s_2 s_2}(\ell \,|\, \bar{a}, s_2^*) + \kappa F_{s_2 s_2}(\ell \,|\, \underline{a}, s_2^*)\right) \left(P(A) + N(A)\right) + \mathcal{C}^2_{s_2 s_2}(s_2^*, A)}$$

The denominator is positive since $F_{s_2 s_2} > 0$ and $\mathcal{C}^2_{s_2 s_2} \geq 0$. The numerator is positive since $F_{a s_2} > 0$, or the sensitivity of the default probability to supervisory intervention greater at $\underline{a}$ than at $\bar{a}$. We therefore have $ds_2^*/d\kappa > 0$ which implies more supervisory intervention if the bank is riskier, that is after a bad signal than after a good signal. Since there are two possible signal realizations $r \in \{\underline{r}, \bar{r}\}$, we have a supervisory intervention strategy $s_2^*(\underline{r}) = \underline{s}_2$ and $s_2^*(\bar{r}) = \bar{s}_2$ with $\underline{s}_2 > \bar{s}_2$.

Conjecture 1. *Riskier banks receive more supervisory attention because of a higher level of intervention.*

What can we say about the probability of default $F(\ell \,|\, a, s_2)$ that results from the four possible combinations of bank actions $a \in \{\underline{a}, \bar{a}\}$ and supervisory interventions $s_2 \in \{\underline{s}_2, \bar{s}_2\}$? This question is important not only because it affects the following analysis but in and of itself because it sheds light on the trade-offs involved in supervisory intervention.

Proposition 1. *Under optimally chosen supervision, the default probabilities for the four possible combinations of bank action* $a \in \{\underline{a}, \overline{a}\}$ *and supervisory intervention* $s_2 \in \{\underline{s}_2, \overline{s}_2\}$ *satisfy the following ranking:*

$$F(\ell \mid \underline{a}, \overline{s}_2) > F(\ell \mid \underline{a}, \underline{s}_2) > F(\ell \mid \overline{a}, \overline{s}_2) > F(\ell \mid \overline{a}, \underline{s}_2)$$

Proof. See Appendix A. □

There is clear intuition behind this ranking of the four default probabilities. First, it is clear that not intervening when necessary leads to the highest possible default probability $F(\ell \mid \underline{a}, \overline{s}_2)$. Second, in case the bank chooses the good action $\overline{a}$, the "right" supervisory response $\overline{s}_2$ results in a *higher* default probability than the "wrong" response $\underline{s}_2$ so that $F(\ell \mid \overline{a}, \underline{s}_2) > F(\ell \mid \overline{a}, \overline{s}_2)$. While this may seem unintuitive at first, it is due to the fact that intervention is costly: The optimal supervisory response is chosen trading off costs and benefits so high intervention is only worthwhile after a bad signal where it has high expected benefit. Third, the fact that intervention is costly also implies that $\underline{s}_2$ will never fully "offset" the effect of $\underline{a}$ so that $F(\ell \mid \underline{a}, \underline{s}_2) > F(\ell \mid \overline{a}, \overline{s}_2)$. Even though we have assumed that supervision could fully correct a bank's mistakes, the fact that it is costly means that optimal supervision will not attempt to fully correct a bank's mistakes.

We now turn to the effect of bank size on supervisory hours for intervention, we have:

$$\frac{d}{dA} h^2(s_2^*, A) \propto \frac{d}{dA} \mathcal{C}^2(s_2^*, A) = \mathcal{C}_A^2(s_2^*, A) + \mathcal{C}_{s_2}^2(s_2^*, A) \frac{ds_2^*}{dA} \tag{5}$$

In contrast to the effect of risk on supervisory hours in (4), the effect of assets has a direct and an indirect part. The first, direct part is the increase in hours when increasing the size of the bank while maintaining the same level of intervention; this part is purely technological. The second part is the change in hours if the first-order condition (3) implies different optimal levels of intervention for different bank sizes.

The second derivative tells us if the increase in hours is proportional, more than proportional, or less than proportional to the increase in assets:

$$\begin{aligned} \frac{d^2}{dA^2} h^2(s_2^*, A) \propto \frac{d^2}{dA^2} \mathcal{C}^2(s_2^*, A) = \mathcal{C}_{AA}^2(s_2^*, A) \\ + \left(\mathcal{C}_{As_2}^2(s_2^*, A) + \mathcal{C}_{s_2 s_2}^2(s_2^*, A) \frac{ds_2^*}{dA} \right) \frac{ds_2^*}{dA} \\ + \mathcal{C}_{s_2}^2(s_2^*, A) \frac{d^2 s_2^*}{dA^2} \end{aligned} \tag{6}$$

This expression has a number of terms. The first is the most straightforward, capturing the curvature of the required hours with respect to size while keeping the level of intervention constant. This part is purely technological and negative.

Conjecture 2. *Technological scale economies suggest that supervisory hours spent on intervention increase less than proportionally with bank size.*

The second and third term on the right-hand side of (6) depend on how the optimal level of intervention s_2^*—which is scale free—changes with bank size. Since we have $\mathcal{C}^2_{As_2} > 0$ and $\mathcal{C}^2_{s_2s_2} > 0$, the second term is determined by the sign of ds_2^*/dA (unless ds_2^*/dA is very negative). If optimal risk reduction is higher for larger banks ($ds_2^*/dA > 0$) the second term is positive and goes against the first term ($\mathcal{C}^2_{AA} < 0$).[13] Now consider the effect of bank size on the optimal level of intervention. Implicit differentiation of the first-order condition (3) yields:

$$\frac{ds_2^*}{dA} = \frac{-\left((1-\kappa)\,F_{s_2}(\ell\,|\,\overline{a}, s_2^*) + \kappa F_{s_2}(\ell\,|\,\underline{a}, s_2^*)\right)\left(P'(A)+N'(A)\right) - \mathcal{C}^2_{As_2}(s_2^*, A)}{\left((1-\kappa)\,F_{s_2s_2}(\ell\,|\,\overline{a}, s_2^*) + \kappa F_{s_2s_2}(\ell\,|\,\underline{a}, s_2^*)\right)\left(P(A)+N(A)\right) + \mathcal{C}^2_{s_2s_2}(s_2^*, A)}$$

As before, the denominator is positive so the sign depends on the numerator:

$$\begin{aligned}&\frac{ds_2^*}{dA} > 0 \quad \Leftrightarrow \\ &-\left(\kappa F_{s_2}(\ell\,|\,\overline{a}, s_2^*) + (1-\kappa)\,F_{s_2}(\ell\,|\,\underline{a}, s_2^*)\right)\left(P'(A)+N'(A)\right) > \mathcal{C}^2_{As_2}(s_2^*, A)\end{aligned}$$

The left-hand side is the increase in spillovers P and N for larger banks while the right-hand side is the increase in marginal cost of intervention for larger banks. We see that stronger intervention at larger banks ($ds_2^*/dA > 0$) is more likely the more the two spillovers, P and N, increase with bank size.

Conjecture 3. *If spillovers increase sufficiently strongly with bank size, then the optimal intensity of intervention increases with bank size, suggesting that supervisory hours spent on intervention increase more than proportionally with bank size.*

How total supervisory hours for intervention scale with bank size therefore depends on two potentially competing effects: technological scale economies of supervision and the effect of scale on the trade-off between marginal benefits and costs of supervision.

[13] The third term depends on whether s_2^* is convex or concave in A.

Supervisory monitoring: Given the intervention strategy $s_2(r)$ derived above, we now consider the bank's choice of a and then the supervisor's choice of monitoring intensity s_1. We denote the bank's expected utility from taking action a and then being subject to supervisory intervention s_2 by $\mathcal{U}(a, s_2)$:

$$\begin{aligned}\mathcal{U}(a, s_2) = \big(1 - F(\ell \,|\, a, s_2)\big)\, \mathrm{E}[Ax - D \,|\, x > \ell] \\ - \Big(F\big(\tfrac{\ell}{1-\rho} \,|\, a, s_2\big) - F(\ell \,|\, a, s_2)\Big)\, R \\ - C^a(a, A) - C^{s_2}(s_2, A)\end{aligned}$$

Given the signal $r \in \{\underline{r}, \overline{r}\}$ and corresponding supervisory interventions $s_2 \in \{\underline{s}_2, \overline{s}_2\}$, the bank has an incentive to take the good action $\overline{a}$ if:

$$\lambda \mathcal{U}(\overline{a}, \overline{s}_2) + (1-\lambda)\,\mathcal{U}(\overline{a}, \underline{s}_2) \geq \lambda \mathcal{U}(\underline{a}, \underline{s}_2) + (1-\lambda)\,\mathcal{U}(\underline{a}, \overline{s}_2) \tag{7}$$

The left-hand side of condition (7) is the bank's expected payoff from choosing $\overline{a}$: With probability λ, the supervisor will observe the right signal $\overline{r}$ (low risk) and choose the appropriate low level of intervention $\overline{s}_2$; with probability $1-\lambda$, however, the supervisor will observe the wrong signal $\underline{r}$ (high risk) and mistakenly choose the high level of intervention $\underline{s}_2 > \overline{s}_2$. The right-hand side of condition (7) is the bank's expected payoff from choosing the bad action $\underline{a}$, where $\underline{s}_2$ is the correct supervisory response and $\overline{s}_2$ means the supervisors don't catch the bank and therefore don't intervene appropriately.

Solving condition (7) for λ, we see that the bank chooses the good action as long as the signal is sufficiently precise, that is the supervisor is sufficiently likely to correctly observe the bank's riskiness:[14]

$$\lambda \geq \Lambda := \frac{\mathcal{U}(\underline{a}, \overline{s}_2) - \mathcal{U}(\overline{a}, \underline{s}_2)}{\mathcal{U}(\overline{a}, \overline{s}_2) - \mathcal{U}(\overline{a}, \underline{s}_2) + \mathcal{U}(\underline{a}, \overline{s}_2) - \mathcal{U}(\underline{a}, \underline{s}_2)}$$

We assume that the supervisor has some uncertainty about the difference in the bank's effort cost C^a between the actions $\overline{a}$ and $\underline{a}$. This can also be uncertain from the bank's perspective—the only important feature is that the bank knows its payoff difference at the time of choosing a while the supervisor doesn't so that there remains some uncertainty about which action the bank chooses. This implies that the supervisor doesn't

[14]For this cutoff to be well-behaved, we need (i) $\mathcal{U}(\underline{a}, \overline{s}_2) > \mathcal{U}(\overline{a}, \underline{s}_2)$, that is getting away with shirking is better than being disciplined when behaving, (ii) $\mathcal{U}(\overline{a}, \overline{s}_2) > \mathcal{U}(\overline{a}, \underline{s}_2)$, that is when behaving it's better if the supervisor knows, and (iii) $\mathcal{U}(\underline{a}, \overline{s}_2) > \mathcal{U}(\underline{a}, \underline{s}_2)$, that is when shirking it's better if the supervisor doesn't know.

know the cutoff Λ precisely but views it as distributed according to a c.d.f. G.[15]

For a given precision λ, the supervisor knows that the bank will choose $\overline{a}$ if $\Lambda \leq \lambda$, that is with probability $G(\lambda)$, and $\underline{a}$ if $\Lambda > \lambda$, that is with probability $1 - G(\lambda)$. In either case, the supervisor will sometimes make mistakes due to observing the wrong signal: If the bank chooses $\overline{a}$, the supervisor observes the wrong signal and mistakenly intervenes with probability $1 - \lambda$ while if the bank shirks, the supervisor mistakenly doesn't intervene with probability $1 - \lambda$.

Similarly to $\mathcal{U}$, we denote the supervisor's expected payoff conditional on the bank's action a and the supervisory intervention s_2 by $\mathcal{V}(a, s_2)$:[16]

$$\mathcal{V}(a, s_2) = \mathrm{E}[Ax] + \big(1 - F(\ell \,|\, a, s_2)\big)\, P(A) - F(\ell \,|\, a, s_2)\, N(A) - \mathcal{C}^2(s_2, A)$$

The supervisor's expected payoff when choosing the effort s_1 to set λ, therefore is:

$$\begin{aligned} &G(\lambda)\, \big(\lambda \mathcal{V}(\overline{a}, \overline{s}_2) + (1-\lambda)\, \mathcal{V}(\overline{a}, \underline{s}_2)\big) \\ &\quad + \big(1 - G(\lambda)\big)\, \big(\lambda \mathcal{V}(\underline{a}, \underline{s}_2) + (1-\lambda)\, \mathcal{V}(\underline{a}, \overline{s}_2)\big) \\ &\quad\quad - \mathcal{C}^1(\lambda, A) \end{aligned}$$

The first part corresponds to the states of the world where the bank chooses $\overline{a}$, which happens with probability $G(\lambda)$; the supervisor then ends up choosing $\overline{s}_2$ with probability λ and $\underline{s}_2$ with probability $1 - \lambda$. The second part corresponds to when the bank chooses $\underline{a}$, which happens with probability $1 - G(\lambda)$; the supervisor then ends up choosing $\overline{s}_2$ with probability $1 - \lambda$ and $\underline{s}_2$ with probability λ.

The supervisor's first order condition for the optimal $s_1^* = \lambda^*$ is then:

$$\begin{aligned} &G'(\lambda^*)\, \big(\lambda^* \mathcal{V}(\overline{a}, \overline{s}_2) + (1-\lambda^*)\, \mathcal{V}(\overline{a}, \underline{s}_2) - \lambda^* \mathcal{V}(\underline{a}, \underline{s}_2) - (1-\lambda^*)\, \mathcal{V}(\underline{a}, \overline{s}_2)\big) \\ &\quad + G(\lambda^*)\, \big(\mathcal{V}(\overline{a}, \overline{s}_2) - \mathcal{V}(\overline{a}, \underline{s}_2)\big) + \big(1 - G(\lambda^*)\big)\, \big(\mathcal{V}(\underline{a}, \underline{s}_2) - \mathcal{V}(\underline{a}, \overline{s}_2)\big) = \mathcal{C}^1_\lambda(\lambda^*, A) \end{aligned} \tag{8}$$

Similar to the first-order condition (3) for intervention s_2, this first-order condition for s_1 equalizes the marginal benefit of increased monitoring (left-hand side) with its marginal cost (right-hand side). The marginal benefit has two parts: First, a more precise signal makes it more likely that $\Lambda \leq \lambda$ so that the bank chooses the right action $\overline{a}$. Second,

[15] Suppose from the supervisor's perspective $C^a(\overline{a}, A) - C^a(\underline{a}, A) = \Delta C + \varepsilon$ where ε is uncertain with some c.d.f. $\widehat{G}(\varepsilon)$. Then ε drops out of the denominator of Λ but remains in the numerator so that Λ is uncertain with c.d.f. $G(\Lambda) = 1 - \widehat{G}\big(\Lambda\, (\mathcal{U}(\overline{a}, \overline{s}_2) - \mathcal{U}(\overline{a}, \underline{s}_2) + \mathcal{U}(\underline{a}, \overline{s}_2) - \mathcal{U}(\underline{a}, \underline{s}_2)) - \mathcal{U}(\underline{a}, \overline{s}_2) + \mathcal{U}(\overline{a}, \underline{s}_2)\big)$.

[16] Note that $\mathcal{V}(a, s_2)$ contains the cost of s_2 but not the cost of s_1.

a more precise signal makes it more likely that the supervisor receives the right signal and implements the correct response.[17]

Similar to the case of intervention, we are interested in the effect of bank size on supervisory hours spent on monitoring. For the first derivative, we have:

$$\frac{d}{dA}h^1(\lambda^*, A) \propto \frac{d}{dA}\mathcal{C}^1(\lambda^*, A) = \mathcal{C}^1_A(\lambda^*, A) + \mathcal{C}^1_\lambda(\lambda^*, A)\,\frac{d\lambda^*}{dA}$$

Analogous to equation (5), the expression has two parts, the increase in hours when increasing the size of the bank while maintaining the same level of precision and the change in hours if the first-order condition (8) implies different optimal levels of precision for different bank sizes.

The second derivative tells us if the increase is proportional or not:

$$\begin{aligned}\frac{d^2}{dA^2}h^1(\lambda^*, A) \propto \frac{d^2}{dA^2}\mathcal{C}^1(\lambda^*, A) = \mathcal{C}^1_{AA}(\lambda^*, A) \\ + \left(\mathcal{C}^1_{A\lambda}(\lambda^*, A) + \mathcal{C}^1_{\lambda\lambda}(\lambda^*, A)\,\frac{d\lambda^*}{dA}\right)\frac{d\lambda^*}{dA} \\ + \mathcal{C}^1_\lambda(\lambda^*, A)\,\frac{d^2\lambda^*}{dA^2}\end{aligned} \tag{9}$$

This expression has the same terms as the analogous expression (6) for $\mathcal{C}^2$. The first captures the curvature of the cost function with respect to size giving us the same technological scale economies as for supervisory intervention.

Conjecture 4. *Technological scale economies suggest that supervisory hours spent on monitoring increase less than proportionally with bank size.*

Also analogous to supervisory intervention, the second and third term on the right-hand side of (9) depend on how the optimal level of monitoring λ^*—which is scale free—changes with bank size.[18] However, in contrast to the analysis for intervention above, the effect of bank size A on the optimal level of supervisory monitoring precision λ^* is not straightforward. We can implicitly differentiate the first-order condition (8) to

[17] For a unique optimal λ^*, we want the left-hand side of (8) to be decreasing in λ^*. Jointly sufficient conditions for this are (i) $G''(\lambda^*) \leq 0$ and (ii) $\mathcal{V}(\bar{a}, \bar{s}_2) - \mathcal{V}(\bar{a}, \underline{s}_2) < \mathcal{V}(\underline{a}, \underline{s}_2) - \mathcal{V}(\underline{a}, \bar{s}_2)$. For a symmetric distribution G, the first condition is satisfied if $G(\lambda^*) \geq 1/2$, that is if in equilibrium the bank is more likely to choose $\bar{a}$ than $\underline{a}$. The second condition says that choosing the correct supervisory intervention is more important when the bank chooses $\underline{a}$ than when it chooses $\bar{a}$. We assume both conditions to be satisfied.

[18] The third term depends on whether λ^* is convex or concave in A.

get:

$$\frac{d\lambda^*}{dA} = \frac{\frac{\partial}{\partial A}\text{LHS}(8) - \mathcal{C}^1_{A\lambda}(\lambda^*, A)}{-\frac{\partial}{\partial \lambda^*}\text{LHS}(8) + \mathcal{C}^1_{\lambda\lambda}(\lambda^*, A)}$$

Note that the left-hand side of (8) is decreasing in λ^* and that we have $\mathcal{C}^1_{A\lambda} > 0$ and $\mathcal{C}^1_{\lambda\lambda} > 0$. The remaining term, the derivative of the left-hand side of (8) with respect to A, is the effect of bank size on the marginal benefit of precision to the supervisor, that is of increasing the likelihood of the bank choosing $\overline{a}$ *and* of the supervisor choosing the right intervention s_2. Given that the marginal cost of more precision is increasing in scale ($\mathcal{C}^1_{A\lambda} > 0$), we see that to warrant higher precision at larger banks, the marginal benefit not only has to be increasing in A, it has to increase *more* than the marginal cost:

$$\frac{d\lambda^*}{dA} > 0 \quad \Leftrightarrow \quad \frac{\partial}{\partial A}\text{LHS}(8) > \mathcal{C}^1_{A\lambda}(\lambda^*, A)$$

Conjecture 5. *If spillovers increase sufficiently strongly with bank size, then the optimal intensity of monitoring increases with bank size, suggesting that supervisory hours spent on monitoring increase more than proportionally with bank size.*

Summary of theoretical results: Our model points to two competing effects that determine whether hours spent on monitoring (s_1) and intervention (s_2) should increase more or less than proportionally with a bank's asset size. On the one hand, the existence of technological scale economies would imply that supervisory hours increase less than proportionally with size. On the other hand, higher optimal intensity of supervision at larger banks would imply that hours increase more than proportionally. The effect of technological scale economies is unambiguous. In contrast, the competing effect of changing intensity requires that banks' spillovers to the economy are sufficiently greater for larger banks in order to outweigh the higher costs of changing intensity.

4.2 Empirical evidence

Fed supervisory hours: Table 2 presents our baseline regression specification of log supervisory hours on log bank assets and a range of controls. The coefficient on log assets captures the elasticity of supervisory hours with respect to bank assets, that is if hours increase proportionally with assets, the coefficient would be equal to 1 and if less than proportionally, less than 1.

The elasticity of supervisory hours to bank assets is 0.96 (column 1) and 0.68 (column

2) when including bank fixed effects. In other words, the pooled and within variation in the data suggests an elasticity close to, or less, than 1. This means that a bank with $10 billion in assets receives less than twice the amount of attention than a bank with $5 billion in assets. Interpreting this terms of our model, the fact that hours increase less than proportionally with size is evidence of technological scale economies in supervision, that is achieving a certain intensity of supervision requires less supervisory resources per dollar of assets at a large bank than at a small bank (Conjectures 2 and 4). The empirical finding of an overall size-coefficient less than 1 does not rule out the possibility that larger banks warrant and receive a higher intensity of supervision, which would—on its own—imply a size coefficient greater than one (Conjectures 3 and 5). However, if this effect exists, it appears quantitatively dominated by the technological scale economies.

To assess the linearity of the estimated relation between supervisory hours and bank size, Figure 4 shows binned scatter plots and the relationship implied by the linear specification without (column 1, left-figure-panel) and including bank fixed effects (column 2, right-figure-panel). As is evident from the figure, the estimated relation between hours and assets is approximately linear in log-scale.

Riskier banks, as measured by their supervisory rating, also display increased supervisory attention, as predicted by Conjecture 1 (column 1 pooled, column 2 within). For example, compared to a bank with the best possible rating of 1, which is the left-out category in the regression, a bank of the same size but with a rating of 3 receives 70% more hours (within variation), an effect roughly equivalent to doubling the size of the bank. A bank rated the worst-possible 5 receives 136% additional hours, roughly equivalent to the baseline hours of a bank three times its size.

In practice, an important component of supervision that is also reflected in the hours measure, is ensuring that a bank is in compliance with regulations as well as assessing its internal processes. We expect these components of supervision, as well as the monitoring and intervention captured by our model, to demand more resources the more complex a banking organization is. Because complexity and size are often related, the elasticity of supervisory hours to size may be measuring the joint effect of banks' size and complexity. We attempt to control for the effect of organizational complexity by controlling the log number of legal entities within each bank holding company (see Appendix B for an exact definition). Hours display a statistically significant elasticity with respect to bank complexity of 0.26 (column 3, pooled variation) and 0.14 (column 4, within variation), consistent with the idea that some possible economies of scale are lost as banks become more complex. Also, including the complexity measures lowers the estimated size-elasticity to 0.77 in the pooled variation regression. As bank complexity does not vary much and is thus mostly accounted for by the fixed effect, the size-elasticity of hours is largely unaffected by the inclusion of the complexity measure.

In sum, the baseline regression specification of Table 2 implies the elasticity of supervisory hours with respect to assets is less than one, and that riskier banks as well as more complex banks require additional attention.

OCC assessment fees: To validate further the conclusions drawn from the hours data, we compare the estimated elasticities of supervisory hours with respect to bank size and risk to those of assessment fees collected by the Office of the Comptroller of the Currency (OCC) on its supervised entities. 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. However, because of potential cross-subsidies across different bank-size or risk categories, the assessment schedule may not be directly informative of the supervisory production function at an institution level. We find that size and risk elasticities of assessment fees turn out to be very similar to those estimated on Federal Reserve supervisory hours.

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 3 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.[19]

[19]With the exception of the addition of the $250 billion asset bracket, asset brackets and base/marginal

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 compute the implied scale economies pre- and post-2008 by regressing log fees on log assets and controls in Table 4. The elasticity of OCC fees to assets is 0.7 which is nearly identical to the within-bank estimate using Fed hours data (see Table 2, column 2). The increase in OCC fees with respect to bank risk is similar although not as steep as the estimated increase in Fed hours. Relative to a 1-rated institutions, fees increase by about 40% on average for 3-rated institutions and by about 70% for 4- or 5-rated institutions.

In Figure 5, we present a binned scatter plot and the linear relationship obtained when regressing OCC fees on bank assets (in log scale). Comparing these results with those obtained using Fed supervisory hours (Figure 4), we again see a very similar pattern. We discuss changes in the OCC fee schedule post-2008 in the next sections.

4.3 Additional determinants of supervisory hours

To investigate further the determinants of supervision in practice, we first consider the sensitivity of supervisory hours with respect to alternative measures of bank size and complexity. We then study seasonal clustering of supervisory hours.

The model assumes that supervisory efforts are devoted to acquiring information (monitoring), and enforcing corrective actions (intervention), related to the riskiness of a bank's assets. Consistently in our main specification (Table 2), we relate hours to a bank's risk and size in terms of its assets. Alternatively, one could consider a model where supervisors are monitoring bank employees to ensure that they exert reasonable effort. We directly consider this alternative by expanding the set of regressors in the baseline regression specification of Table 2 to include the log of a bank's total employees. We see a statistically significant elasticity of about 0.2 in the pooled estimation (Table 5, column 1), which is no longer significant when considering within-bank variation (column 2). This indicates that the sensitivity to a bank's head count is either relatively small or not significant, justifying the focus on assets in our model and empirical analysis.

Besides the organizational complexity in our baseline specification (number of legal entities), the complexity of the consolidated balance sheet may also be important.

fee schedules prior to 2007 were stable over time, except for an annual inflation adjustment. Both inflation adjustment and rating surcharges were capped at $20 billion, prior to 2014, and at $40 billion thereafter.

To account for this possibility, we consider two measures of asset complexity: (i) the share of loans to total assets, which is higher for traditional, less complex banking organizations, and the concentration of the bank's activities as measure by the Herfindahl index of the share of the asset categories. We find that neither measure of balance sheet complexity enters significantly in our estimated regressions (columns 3 and 4).

Finally, we consider in more detail the temporal distribution of supervisory attention at an individual bank over the course of a year. In the extreme, supervision could be done either "continuously" such that hours are stable over the year, or in a "punctuated" fashion such that hours spike at certain points during the year. While continuous supervision would seem more effective, resource constraints would be a rationale for using a punctuated approach. This would rotate focused attention from bank to bank, concentrating resources mainly on the full scope examinations, which are mandated by law to be completed annually.

To study this question empirically, we use information on whether a bank is undergoing a full-scope examination and include a dummy variable for whether such an event occurs in a given quarter. We interact this indicator variable with another for large and smaller institutions. The point estimates (Table 5, columns 5 and 6) imply that supervisory efforts increase during a full scope examination, with the larger banks experiencing an increase of about 100%, and the smaller banks of about 200%. Figure 6 uses information at monthly frequency to provide additional detail on the temporal distribution of supervisory attention. For each asset-size group, the chart plots average monthly hours (orange, right scale) against time measured relative to the month in which a full scope examination ends. Average hours for each bank and examination event are relative to those at time t-6. The blue bars in each chart denote (right scale) denote the probability with which a bank is subject to a full-scope or other examination. As shown by these bars, the probability increases (and reaches 1) as time approaches to 0 and then drops. For small BHCs, hours increase by a factor of about 4.5, as opposed to large banks for which the increase is by a factor close to 1. This evidence suggests that supervision for the larger BHCs is conducted mainly continuously, while for the smaller institutions attention is punctuated and displays periodic patterns.

5 Interaction of supervision and regulation over time

We now turn to the interaction between bank supervision and regulation—in the form of capital requirements. We first study whether, in our model, regulation is a complement or substitute to the two components of supervision: monitoring and intervention. Motivated by the changes in both regulation and supervision following the financial crisis of 2008-09, we then consider how changes in model parameters—spillovers, in particular—affect optimal supervision and its dependence on bank characteristics such as size.

5.1 Theoretical analysis

While we take regulation in our setup as exogenous, it can interact with optimal supervision. The bank has to fulfill the capital requirement ρ both ex-ante ($A - D \geq \rho A$) as well as ex-post ($Ax - D \geq \rho Ax$). In addition, regulation imposes a cost R conditional on violating the ex-post capital requirement. To understand whether supervision is a complement or substitute to regulation, we consider the effect of a change in regulation on the optimal level of supervision.

Suppose banks fulfill the ex-ante requirement exactly, that is $A - D = \rho A$ which implies $\ell = 1 - \rho$.[20] Then changes in the capital requirement ρ have direct effects on the supervisor payoff because the change in bank capitalization affects the probability of default $F(\ell \mid a, s_2) = F(1 - \rho \mid a, s_2)$. In addition, changes in ρ have indirect effects on the supervisor payoff through changes in the bank's behavior. Both the direct and indirect effect imply that optimal supervisory efforts would respond to a change in ρ.

Consider first the effect of ρ on the bank's incentives. For $\ell = 1 - \rho$, the bank's expected payoff from taking action a and facing supervisory intervention s_2 becomes:

$$\begin{aligned} \mathcal{U}(a, s_2) = & \left(1 - F(1 - \rho \mid a, s_2)\right) \mathrm{E}[Ax - D \mid x > 1 - \rho] \\ & - \left(F(1 \mid a, s_2) - F(1 - \rho \mid a, s_2)\right) R \\ & - C^a(a, A) - C^{s_2}(s_2, A) \end{aligned}$$

Raising the capital requirement increases the bank's initial equity and therefore reduces the probability of default and the bank's exposure to low returns, providing for more "skin in the game." In addition, raising ρ increases the probability of violating the requirement ex-post *conditional* on not defaulting.[21]

Since with a higher capital requirement the bank has more "skin in the game" and

[20] Allowing for a buffer β above the ex-ante capital requirement, that is $A - D = (\rho + \beta) A$ or $\ell = 1 - (\rho + \beta)$, does not affect our results.

[21] Since ρ is both the ex-ante (binding) and ex-post requirement, the unconditional probability of violating it ex-post is simply the probability of an asset return $x < 1$ which doesn't depend on ρ.

a higher conditional probability of facing the penalty R, the condition (7) is relaxed, increasing the probability that the bank chooses the good action $\overline{a}$. From the point of view of the supervisor this corresponds to a decrease in the left-hand side of the first-order condition (8) so the optimal level of monitoring λ^* decreases. Supervisory monitoring and regulatory capital requirements are therefore *substitutes*.

We now turn to the interaction of capital requirements and supervisory intervention s_2. Consider the first-order condition for intervention (3) with a binding ex-ante capital requirement, that is $\ell = 1 - \rho$:

$$-\left(\kappa F_{s_2}(1-\rho \,|\, \overline{a}, s_2^*) + (1-\kappa)\, F_{s_2}(1-\rho \,|\, \underline{a}, s_2^*)\right)\left(P(A) + N(A)\right) = \mathcal{C}^2_{s_2}(s_2^*, A) \tag{10}$$

Since $F_{xs_2} > 0$, the direct effect of raising the capital requirement is to increase the left-hand side of the first-order condition (10) so the optimal level of intervention s_2^* increases. Intuitively, raising the capital requirement moves the default threshold $1-\rho$ further into the left tail of the distribution F, where the effect of s_2 is stronger. The marginal benefit of supervisory intervention therefore increases, warranting a higher level of intervention. In contrast to supervisory monitoring, supervisory intervention and regulatory capital are therefore *complements*.

Conjecture 6. *The direct effects of a tightening of regulation in the form of a higher capital requirement are a decrease in supervisory monitoring (substitutes) but an increase in supervisory intervention (complements).*

The difference in the interaction of capital requirements and monitoring vs. intervention are surprising but intuitive. Capital requirements and monitoring both improve the bank's incentives which makes them substitutes. In contrast, intervention works mainly on the probability of default. Since tighter regulation makes intervention more effective, the two are complements. Whether the substitutability or the complementarity dominates when considering total supervisory hours is ambiguous.

Changes since 2008: One interpretation of the concurrent changes to regulation and supervision following the financial crisis of 2008-09 is a change in the actual or perceived spillovers of a bank operating or failing, $N(A)$ and $P(A)$, respectively. In fact, the total difference in surplus between a successful and a failed bank, $P(A) + N(A)$, appears as part of the marginal benefit of both supervisory monitoring and intervention in the first-order conditions (3) and (8).

Conjecture 7. *An increase in supervisory concern about spillovers suggests higher supervisory monitoring and intervention.*

For supervisory intervention, the marginal benefit is a reduction in the expected spillover loss due to a reduced probability of default. If the spillovers are larger, the marginal benefit of intervention is greater and warrants an increase in the level of intervention to rebalance marginal benefit and marginal cost. For supervisory monitoring, the marginal benefit is an increase in both the likelihood that the bank chooses the "good" action $\bar{a}$ and the likelihood that the supervisor correctly observes the bank's action and therefore chooses the correct intervention. Both are more important, if the spillovers are larger, and therefore warrant more supervisory monitoring.[22]

5.2 Empirical evidence

For the empirical analysis we use the Tier 1 capital ratio—one of the most important tools for bank regulation— to measure a bank's regulatory capital. The observed level of regulatory capital held by a bank is typically above the minimum by an additional buffer chosen by the bank. In addition, changes in the observed capital are not only due to changes in the minimum requirement but also due to the bank's realized profits and losses. This complicates the analysis of the interaction between regulation and supervision since banks making losses and having low capital levels are more likely to face supervisory intervention. In other words, a negative relation between supervisory hours and capital levels may not be an indication of substitution between the two policies.

Splitting the analysis into small and large banks, we find significant coefficients of opposite signs for the relationship between supervisory hours and bank capital: the coefficient for small banks is negative, while that for large banks is positive (Table 6, column 1). The negative coefficient for small banks is consistent both with substitution between regulation and supervision and with losses signaling trouble at the bank and therefore higher supervisory attention. However, when controlling for supervisory ratings—a more precise measure of banks in trouble—the negative coefficient for small banks loses its significance (column 2). For small banks, therefore, our empirical results are similarly ambivalent as the theoretical analysis (Conjecture 6).

[22]This result is similar to Bhattacharya et al. (2002) where exam frequency and capital requirements are substitutes in preventing bank risk taking but higher fundamental risk requires higher capital requirements and/or more frequent exams.

For large banks, the *positive* coefficient on Tier 1 capital remains significant even after controlling for ratings (column 2). All regression specifications in Table 6 include bank fixed effects, meaning that the coefficients are only identified by within-bank variation rather than cross-sectional differences. The positive coefficient for large banks could therefore be evidence either of overall complementarity between regulation and supervision (Conjecture 6) or of an exogenous change that leads to an increase in both supervisory attention and regulatory capital levels at the largest banks (Conjecture 7).

A key change in capital requirements took place after 2008 as the largest banks were required to hold more capital because of stress testing (CCAR and DFAST), as well as various mandatory capital buffers and planned surcharges such as those for "global systemically important bank holding companies" (12 CFR 217). As shown in Figure 7, larger banks used to hold significantly less capital than their smaller counterparts pre-2008, but this difference has essentially disappeared post-2008 as the largest banks have increased their capital ratios. Splitting the sample into pre- and post-2008 illustrates this effect (Table 6, columns 3 and 4) as the positive coefficient on the Tier-1 ratio for large banks is entirely driven by the post-2008 sample.

Given the increase in supervisory efforts post-2008, it is interesting to study the extent to which the elasticity of hours to bank size increased over that sample. To that end, we extend the baseline regression specification to include an interaction term between log assets and a dummy variable marking the post-2008 sample. Hours sensitivity increased by a relatively small amount of about 0.1 in absolute terms in the post-2008 period (Table 7, columns 1 and 2). The level increase in hours for the largest banks is notable and on the order of 60% (columns 3 and 4). Interestingly, however, hours spent at the smallest institutions ($\leq$ \$10 billion) declined by about 20-30%. In Figure 8, we illustrate this effect graphically. We first estimate residuals from our baseline specification without bank fixed effects (Table 2, column 1). We then average residuals of log-hours by quarter and size-group and plot their evolution over time. Residuals for the largest banks (blue) are about 50% larger in the post-2008 sample, as opposed to the smaller institutions, for which we observe a decline.

While it is possible to conceive that post-2008 supervisors became less concerned about the smallest institutions at the same time that attention increased at the largest banks, an alternative explanation is that supervisory resources are scarce and relatively inelastic in the short run. An increase in the optimal relative attention to large banks can then only be implemented by an absolute reduction of attention at small banks. To shed light on this question, we compare these results on hours with the pre/post-2008 elasticity of OCC fees to size.

As discussed in Section 4.2, the overall elasticities of OCC fees to bank size and risk are very similar to those of Fed hours. If the decline in Fed supervisory hours at small banks is due to lower concerns, we would expect to see a similar decline in the size-elasticity of OCC fees for small banks. If, however, the decline is due to limited resources, we would not expect any change in OCC fees for small banks. In 2008, the OCC introduced an additional fee bracket for banks with assets greater than \$250 billion to cover increasing costs associated with the supervision of large and complex banks. The elasticity of fees to bank assets is modestly affected by this additional bracket (Table 4, column 2), but on average, fees are about 50% higher in 2014 than in 2007 for banks with assets greater than \$10 billion (column 3), while for smaller banks fees are essentially unchanged. In addition, the increase in fees for large banks matches very closely the increase in hours of Fed supervisors post-2008.

6 Supervisory efforts when dealing with multiple banks

6.1 Theoretical analysis

In the theoretical analysis so far, we have considered the optimal level of a supervisory action $s \in \{s_1, s_2\}$ for a single bank without explicitly taking into account the fact that the supervisors have to allocate resources across multiple banks. For both monitoring and intervention, the supervisory action has some benefit and cost when the cost is proportional to hours. Denoting the benefit of action s with a bank of size A generically as $\mathbb{V}(s, A)$, the supervisor's problem for an individual bank has the structure

$$\max_s \left\{ \mathbb{V}(s, A) - wh(s, A) \right\},$$

which yields first-order conditions (3) and (8) of the form

$$\mathbb{V}_s(s, A) = wh_s(s, A). \tag{11}$$

In practice, supervisors have to deal with multiple banks and are allocating a given budget of available hours $\overline{H}$. Note that we can invert the hourly cost function h to infer the level of supervisory action $s = h^{-1}(H, A)$ that results from an allocation of H hours to a bank of size A. Then we can state the problem of allocating hours across a set of banks as:

$$\max_{\{H_i\}} \Big\{ \textstyle\sum_i \mathbb{V}\big(h^{-1}(H_i, A_i), A_i\big) \Big\} \quad \text{subject to} \quad \textstyle\sum_i H_i \le \overline{H}$$

With a Lagrange multiplier μ on the budget constraint, this yields first-order conditions of the form

$$\mathbb{V}_s(s_i, A_i)\,\frac{1}{h_s(s_i, A_i)} = \mu \quad \text{for all } i. \tag{12}$$

Comparing the first-order conditions (11) and (12) we see that they are structurally equivalent and that the wage w in the individual-bank version is simply the "shadow" wage represented by the Lagrange multiplier μ in the multiple-banks version. This means that all the comparative statics we derive above using the individual-bank version carry over to the multiple-bank version.

In addition, the first-order conditions require that the ratio of marginal benefit to marginal cost be equalized across banks:

$$\frac{\mathbb{V}_s(s_i, A_i)}{h_s(s_i, A_i)} = \frac{\mathbb{V}_s(s_j, A_j)}{h_s(s_j, A_j)} \quad \text{for all } i, j \tag{13}$$

Since the two actions s_1 and s_2 happen sequentially—first monitoring to collect information, then intervention conditional on the information—satisfying condition (13) for both requires reallocating supervisory hours after the signals $\{r_i\}$ are realized.

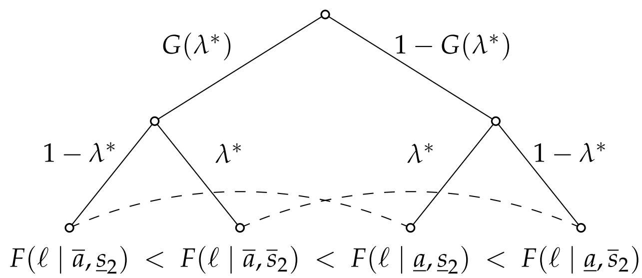
Conjecture 8. *With limited resources, supervisory attention has to be conditionally reallocated towards banks with bad signals and away from banks with good signals.*

Conditional on bank size A, more intervention is required after a bad signal than after a good signal, $\underline{s}_2 > \overline{s}_2$. With limited resources, this means that supervisory hours are reallocated from banks with good signals to banks with bad signals.

6.2 Empirical evidence

So far we have presented two empirical results that are consistent with limited resources and reallocation of attention across banks. Table 5 provided evidence that supervisory efforts are clustered during scheduled examinations and are lower at other times, especially for the smaller banks. In addition, in Table 7, we provided evidence that while supervisory efforts have increased at the largest institutions post-2008, they have declined at smaller institutions. Here we provide an additional test, motivated by Conjecture 8,

Figure 3: **Event tree leading to conditional default probabilities**



that attention is reallocated from banks with good ratings to banks with bad ratings and that this effect is particularly strong if the stressed banks are large.

We assume that that, in the short run, resources are relatively fixed within each of the twelve Federal Reserve districts, allowing us to study the impact of supervisory attention at one bank when other banks in the same district warrant relatively more attention because they are under stress. At the district level, we therefore compute the share of assets in distress (assets of banks rated 3, 4 or 5) and study its effect on hours at large and small banks.[23] As shown in columns 3 and 4 of Table 8, we find that an increase in stress at other banks in the district implies a decline of attention at a given small bank, while we do not find a significant effect at large banks (after controlling for the post-2008 change in attention). This provides evidence of resource constraints forcing reallocation of attention, as predicted by Conjecture 8. The point estimates imply that attention at a small bank would decline by between 30 and 60% if all remaining banks in a district became stressed.

7 Assessing supervisory outcomes ex-post

7.1 How can things go wrong?

If we observe a bank default, what can we infer about the likely cause? Was it just the result of a tail realization in asset returns? Did, instead, the bank misbehave, the supervisor not notice, or a combination of the two? Figure 3 lists the possible combinations of the bank's action $a \in \{\underline{a}, \bar{a}\}$ with supervisory response $s_2 \in \{\underline{s}_2, \bar{s}_2\}$ and the resulting probability of default $F(\ell \mid a, s_2)$. For example, given optimal supervisory monitoring λ^*, the bank chooses the bad action $\underline{a}$ with probability $1 - G(\lambda^*)$. The supervisor no-

[23] When calculating the share, we always exclude the bank under observation.

tices this and responds with $\underline{s}_2$ with probability λ^*, resulting in a default probability $F(\ell \mid \underline{a}, \underline{s}_2)$.

The default probabilities in the figure are ranked left-to-right from smallest to largest (see Section 4.1 for the ranking). Since the default probability is highest if the bank chooses $\underline{a}$ and the supervisor mistakenly chooses $\bar{s}_2$ one may conclude that this combination is the likeliest cause when observing a bank failure. However, this reasoning would ignore the fact that $F(\ell \mid \underline{a}, \bar{s}_2)$ is a *conditional* probability, that is the probability of default conditional on $\underline{a}$ and $\bar{s}_2$. Since the likelihood of the event sequence $\underline{a}$ and $\bar{s}_2$ is $\left(1 - G(\lambda^*)\right)(1 - \lambda^*)$, we know that the *unconditional* probability of observing a bank failure due to the bank taking high risk and the supervisor not noticing is given by:

$$\left(1 - G(\lambda^*)\right)(1 - \lambda^*)\, F(\ell \mid \underline{a}, \bar{s}_2)$$

While it the conditional default probabilities $F(\ell \mid a, s_2)$ can clearly be ranked from smallest to largest (Proposition 1), it is much harder to clearly rank the unconditional default probabilities

$$\mathcal{P}(a, s_2) \equiv \Pr[a] \times \Pr[s_2 \mid a] \times F(\ell \mid a, s_2)$$

Proposition 2. *Under optimal supervision, only an incomplete ranking of unconditional default probabilities is possible:*

$$\mathcal{P}(\underline{a}, \underline{s}_2) > \mathcal{P}(\bar{a}, \underline{s}_2) \tag{14}$$

$$\text{and} \quad \mathcal{P}(\bar{a}, \bar{s}_2) > \mathcal{P}(\bar{a}, \underline{s}_2) \tag{15}$$

The first inequality, (14), states that we are more likely to observe a default where the bank chose the "bad" action $\underline{a}$ and the supervisor intervened appropriately with $\underline{s}_2$ than a default where the bank chose the "good" action $\bar{a}$ and the supervisor mistakenly intervened with $\underline{s}_2$. The second inequality, (15), states that a default where the bank chose the "good" action and the supervisor intervened appropriately with $\bar{s}_2$ is also more likely than one with $(\bar{a}, \underline{s}_2)$. Both inequalities therefore involve the probability of a bank failure where the bank chose the low-risk action $\bar{a}$ but the supervisor received the wrong signal and chose to intervene with $\underline{s}_2$, further reducing the bank's risk. As discussed in Section 4.1, this combination leads to the lowest conditional probability of default. Inequalities (14) and (15) tell us that such a bank failure is less likely than either case of correct supervisory response, $(\underline{a}, \underline{s}_2)$ and $(\bar{a}, \bar{s}_2)$, respectively.

The one unconditional default probability missing from the inequalities (14) and (15) is the one corresponding to the most troublesome scenario, where the bank takes high risk but it goes undetected by the supervisor, $(\underline{a}, \overline{s}_2)$.

Corollary 1. *Without making further assumptions on parameters and functional forms, we cannot infer from a bank failure whether the scenario* $(\underline{a}, \overline{s}_2)$ *was a more or less likely than any of the other three possibilities,* $(\overline{a}, \underline{s}_2)$*,* $(\underline{a}, \underline{s}_2)$*, and* $(\overline{a}, \overline{s}_2)$*.*

7.2 Supervisory incentives and time inconsistency

Our analysis of the sequence of supervisory choice s_1, bank choice a and supervisory choice s_2 has been subgame perfect: The intervention strategy $s_2(r)$ is optimal conditional on the signal, the bank's action a is optimal given s_1 and taking into account the response s_2, and, finally, supervisory monitoring s_1 is optimal taking into account both its effect on a and s_2. We have therefore imposed time consistency on the supervisor. Instead, we could imagine a supervisor able to commit to an intervention strategy $s_2(r)$ that may be suboptimal ex-post in order to improve the bank's incentives ex-ante.

In the subgame perfect analysis, the intervention s_2 takes as given the bank's action and only trades off the benefit of more intervention—a reduction in the probability of default—against the costs of intervention in terms of hours (equation (3)). In particular, the supervisor does not take into account the effect of s_2 on the bank's incentives (7) where more intervention creates slack and therefore makes it more likely that the bank chooses the good action $\overline{a}$.

Conjecture 9. *Given the ability to commit to an intervention strategy that may be suboptimal ex-post, the supervisor could do better than under a constraint of time consistency.*

With commitment, the supervisor would choose a strategy $\tilde{s}_2(r)$ taking into account the effects on the bank's incentives and therefore a which is an additional benefit of more intervention and therefore worth additional supervisory costs $\mathcal{C}^2$. However, once the bank's action is taken and the supervisor receives the signal r, the level of intervention s_2 no longer has incentive effects and the supervisor's cost-benefit trade-off is purely about affecting the probability of default. The higher level of intervention that was ex-ante optimal will no longer be optimal ex-post.[24]

[24] For a similar result, see Decamps et al. (2004) who find that the effectiveness of market discipline is reduced if supervisors lack commitment against forbearance.

8 Conclusion

This paper provides a new perspective on bank supervision by combining a theoretical framework and an empirical investigation that speak to the objectives and resource constraints of supervision. Our model uses building blocks from the theory of contracts and incomplete information to derive key implications of an optimally chosen supervision strategy. Taking into account that any supervisory strategy is subject to resource constraints, we highlight the trade-off between the costs and benefits of supervision and derive comparative statics that are confirmed in the data.

As is to be expected, we find that larger banks receive more attention in the form of supervisory hours than smaller banks. However, hours increase less than proportionally with size. Through the lens of the model, this is evidence of technological scale economies in supervision that outweigh the effect of increased concerns about the largest banks.

When studying resource allocation with multiple banks, we also find evidence of substitution effects of supervisory efforts indicating binding resource constraints. We also directly measure how the post-2008 new regulatory and supervisory framework for large and complex banks has resulted in an increase in resources allocated to the largest institutions.

Much of the literature studying the role of regulation and supervision in the 2008 financial crisis has focused on possible policy distortions arising from institutional design (Agarwal et al., 2014; Carletti et al., 2015), incentive problems (Lucca et al., 2014) or time inconsistency, for example in the context of "too big to fail." This paper takes a complementary approach by focusing on the limits and trade-offs of supervision based on limited resources. But the model can speak to some issues of supervisory incentives as well. For example, the non-verifiability of the supervisors' information opens up the possibility of incentive problems for the supervisors themselves. Our model allows supervisory costs to include non-monetary costs, for example opportunity costs of foregone private benefits, and therefore implicitly allows for a wedge between the strategy chosen by the supervisor and one chosen by a social planner. However, there are potentially much richer incentive issues in the spirit of "who monitors the monitors" (Hurwicz, 2007) that could be considered but are beyond the scope of this paper.[25]

[25]See, for example, Masciandaro and Quintyn (2013) for an extensive survey on the governance of supervision.

References

Acharya, V. V. (2009). A theory of systemic risk and design of prudential bank regulation. *Journal of Financial Stability 5*(3), 224–255.

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

Berger, A. N. and S. M. Davies (1998). The information content of bank examinations. *Journal of Financial Services Research 14*(2), 117–144.

Bhattacharya, S., M. Plank, G. Strobl, and J. Zechner (2002). Bank capital regulation with random audits. *Journal of Economic Dynamics and Control 26*(7), 1301–1321.

Brunnermeier, M. K., A. Crockett, C. A. Goodhart, A. Persaud, and H. S. Shin (2009). The fundamental principles of financial regulation.

Carletti, E., G. Dell'Ariccia, and R. Marquez (2015). Supervisory incentives in a banking union. Working Paper.

Cetorelli, N. and S. Stern (2015). Same name, new businesses: Evolution in the bank holding company. *Liberty Street Economics Blog* (September 28, 2015).

Cole, R. A. and J. W. Gunther (1995). A camel rating's shelf life.

Decamps, J.-P., J.-C. Rochet, and B. Roger (2004). The three pillars of basel ii: optimizing the mix. *Journal of Financial Intermediation 13*(2), 132 – 155. Bank Capital Adequacy Regulation under the New Basel Accord.

Dewatripont, M. and J. Tirole (1994). *The prudential regulation of banks*. MIT Press.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Eisenbach, T. M., A. Haughwout, B. Hirtle, A. Kovner, D. O. Lucca, and M. C. Plosser (2015). Supervising large, complex financial institutions: What do supervisors do? *FRB of New York Staff Report* (729).

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.

Goldsmith-Pinkham, P., B. Hirtle, and D. O. Lucca (2016). Parsing the content of bank supervision. Working Paper.

Harris, M. and A. Raviv (2012). How to get banks to take less risk and disclose bad news. *Chicago Booth Research Paper* (12-55).

Hirtle, B., A. Kovner, and M. Plosser (2016). The impact of supervision on bank performance. Working Paper.

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

Hurwicz, L. (2007). But who will guard the guardians? *Nobel Prize Lecture*.

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.

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

Kovner, A., J. Vickery, and L. Zhou (2014). Do big banks have lower operating costs? *Federal Reserve Bank of New York Economic Policy Review 20*(2), 1–27.

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.

Marshall, D. A. and E. S. Prescott (2006). State-contingent bank regulation with unobserved actions and unobserved characteristics. *Journal of Economic Dynamics and Control 30*(11), 2015–2049.

Masciandaro, D. and M. Quintyn (2013). The evolution of financial supervision: the continuing search for the holy grail. *SUERF 50th Anniversary Volume Chapters*, 263–318.

Maskin, E., J. Laffont, and W. Hildenbrand (1982). The theory of incentives: An overview. In *Advances in Economic Theory (invited lectures from the 4th World Congress of the Econometric Society)*. Cambridge University Press.

Merton, R. C. (1995). A functional perspective of financial intermediation. *Financial Management*, 23–41.

Mishkin, F. S. (2001). Prudential supervision: Why is it important and what are the issues? In *Prudential Supervision: What Works and What Doesn't*, pp. 1–30. University of Chicago Press.

Prescott, E. S. (2004). Auditing and bank capital regulation. *Federal Reserve Bank of Richmond Economic Quarterly 90*(4), 47–63.

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.

Tirole, J. (2013). Financial regulation. In *Rethinking Macro Policy II: First Steps and Early Lessons*.

Table 1: **Summary statistics.** This table presents summary statistics for the variables included in the regression specifications. Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets. For detailed variable definitions see Section 3 and Appendix B.

	All		**Small BHCs**		**Large BHCs**	
	Mean	StDev	Mean	StDev	Mean	StDev
Hours	489	1368	98	177	1447	2255
Assets ($ millions)	36040	188494	2541	1754	118050	336230
Log(Hours)	4.15	2.16	3.35	1.74	6.11	1.82
Log(Assets)	8.44	1.50	7.66	0.57	10.34	1.37
Rating	1.99	0.78	2.01	0.82	1.92	0.64
Log(N BHC Subsidiaries)	2.48	1.30	1.92	0.77	3.83	1.36
Log(N BHC Employees)	6.99	1.47	6.24	0.65	8.82	1.31
HHI, asset concentration	0.51	0.10	0.53	0.10	0.47	0.10
Loan Share	0.66	0.12	0.67	0.11	0.62	0.13
Ongoing Full Exam	0.28	0.45	0.24	0.43	0.37	0.48
Large BHC	0.29	0.45	0	0	1	0
Tier 1 Capital Ratio	11.83	3.23	12.16	3.26	11.03	3
Share of District Assets in Distress	0.18	0.27	0.19	0.28	0.14	0.24
Observations	16372		11624		4748	

Table 2: **Baseline regression.** This table presents estimates of supervisory hours on bank size (assets), complexity (number of subsidiaries) and dummy variables for supervisory ratings. Sample is 1998Q1-2014Q4. Standard errors clustered by bank reported in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	Log(Hours)			
	(1)	(2)	(3)	(4)
Log(Assets)	0.96***	0.68***	0.77***	0.62***
	[0.02]	[0.11]	[0.04]	[0.12]
Log(N BHC Subsidiaries)			0.26***	0.14**
			[0.04]	[0.07]
Rating = 2	0.23***	0.15**	0.22***	0.13**
	[0.05]	[0.06]	[0.05]	[0.06]
Rating = 3	0.94***	0.70***	0.89***	0.66***
	[0.09]	[0.09]	[0.08]	[0.09]
Rating = 4	1.30***	1.07***	1.25***	1.03***
	[0.11]	[0.11]	[0.11]	[0.11]
Rating = 5	1.61***	1.36***	1.52***	1.29***
	[0.16]	[0.16]	[0.15]	[0.16]
District, Date FEs?	Yes	Yes	Yes	Yes
Bank FEs?	No	Yes	No	Yes
Adj. R2	0.49	0.56	0.51	0.57
Obs.	17969	17969	17235	17235
Distinct BHCs	785	785	727	727

Table 3: **OCC general assessment fee schedule.** This table shows the OCC assesment 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 4: **OCC general assessment semiannual fee regressions.** This table shows the relation between OCC general assessments as a function of commercial banks' assets and ratings. 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 3 and rating surcharges discussed in Section 4.2. Assets are actual, while ratings are generated from a uniform distribution. The $10 billion asset threshold is expressed in nominal terms.

	Log(Fees)		
	(1)	(2)	(3)
Log(Assets)	0.70***	0.69***	0.68***
	[0.00]	[0.00]	[0.00]
Post-2008 × Log(Assets)		0.01***	
		[0.00]	
Post-2008 × (Assets ≥ $10bn)			0.47***
			[0.03]
Post-2008 × (Assets < $10bn)			0.03***
			[0.00]
Rating = 2	-0.01*	-0.01*	-0.01*
	[0.01]	[0.01]	[0.01]
Rating = 3	0.40***	0.40***	0.40***
	[0.01]	[0.01]	[0.01]
Rating = 4	0.68***	0.68***	0.68***
	[0.01]	[0.01]	[0.01]
Rating = 5	0.69***	0.69***	0.69***
	[0.01]	[0.01]	[0.01]
Constant	-0.03	-0.02	0.04*
	[0.02]	[0.02]	[0.02]
Adj. R2	0.99	0.99	0.99
Obs.	2866	2866	2866
Distinct NAs	1772	1772	1772

Table 5: **Supervisory production function regression.** This table presents estimates of supervisory hours on alternative measures of bank size (employees) and complexity (loan share and asset concentration) as well as dummy variables indicating whether a banking institution is undergoing a full-scope examination. Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets. Sample is 1998Q1-2014Q4. Standard errors clustered by bank reported in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	Log(Hours)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Assets)	0.59***	0.53***	0.80***	0.63***	0.79***	0.65***
	[0.08]	[0.16]	[0.03]	[0.12]	[0.04]	[0.12]
Rating = 2	0.23***	0.13**	0.22***	0.13**	0.14***	0.05
	[0.05]	[0.06]	[0.05]	[0.06]	[0.05]	[0.06]
Rating = 3	0.90***	0.66***	0.90***	0.67***	0.71***	0.51***
	[0.08]	[0.09]	[0.09]	[0.09]	[0.08]	[0.08]
Rating = 4	1.26***	1.02***	1.26***	1.03***	1.03***	0.84***
	[0.11]	[0.11]	[0.11]	[0.11]	[0.10]	[0.11]
Rating = 5	1.54***	1.29***	1.54***	1.29***	1.23***	1.03***
	[0.15]	[0.16]	[0.15]	[0.17]	[0.15]	[0.16]
Log(N BHC Subsidiaries)	0.24***	0.13**	0.23***	0.14**	0.26***	0.16**
	[0.04]	[0.07]	[0.04]	[0.07]	[0.04]	[0.06]
Loan Share			-0.27	0.48		
			[0.52]	[0.77]		
Asset Conc. (HHI)			0.44	-0.50		
			[0.56]	[0.76]		
Log(N BHC Employees)	0.21**	0.12				
	[0.08]	[0.16]				
Ongoing Exam × (Large BHC)					0.98***	0.83***
					[0.07]	[0.07]
Ongoing Exam × (Small BHC)					1.98***	2.08***
					[0.04]	[0.04]
District, Date, Ongoing FEs?	Yes	Yes	Yes	Yes	Yes	Yes
Bank FEs?	No	Yes	No	Yes	No	Yes
Adj. R2	0.51	0.57	0.49	0.55	0.62	0.68
Obs.	17234	17234	16846	16846	17235	17235
Distinct BHCs	727	727	716	716	727	727

Table 6: **Supervisory efforts and capital.** This table extends the baseline regression specification in Table 2 to include controls for a bank's tier 1 capital ratio interacted with a dummy for asset size. Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets. Sample is 1998Q1-2014Q4. Standard errors clustered by bank reported in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	Log(Hours)			
	(1)	(2)	(3)	(4)
Log(Assets)	0.34***	0.50***	0.35***	0.51***
	[0.11]	[0.10]	[0.12]	[0.16]
Rating = 2		0.08	0.02	0.15
		[0.06]	[0.07]	[0.11]
Rating = 3		0.50***	0.51***	0.54***
		[0.09]	[0.13]	[0.13]
Rating = 4		0.81***	1.48***	0.75***
		[0.11]	[0.33]	[0.14]
Rating = 5		1.13***	2.00***	1.10***
		[0.16]	[0.35]	[0.20]
T1 Ratio × (Small BHC)	-0.03***	-0.01	-0.00	-0.01
	[0.01]	[0.01]	[0.02]	[0.01]
T1 Ratio × (Large BHC)	0.05***	0.06***	0.03	0.09***
	[0.01]	[0.01]	[0.02]	[0.02]
District, Date, Ongoing FEs?	Yes	Yes	Yes	Yes
Bank FEs?	Yes	Yes	Yes	Yes
Adj. R2	0.68	0.69	0.69	0.73
Obs.	17329	17329	9559	7770
Distinct BHCs	772	772	590	610
Dates	19981.20144	19981.20144	19981.20084	20091.20144

Table 7: **Changes in supervisory efforts by size over time.** This table extends the baseline regression specification in Table 2 to include interactions of asset size and size dummies with a post-2008 dummy. Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets. Sample period is reported below each column. Standard errors clustered by bank reported in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	Log(Hours)			
	(1)	(2)	(3)	(4)
Log(Assets)	0.93***	0.65***	0.90***	0.62***
	[0.03]	[0.12]	[0.02]	[0.07]
Post-2008 × Log(Assets)	0.10***	0.13***		
	[0.03]	[0.03]		
Rating = 2	0.16***	0.08	0.14***	0.08
	[0.05]	[0.06]	[0.05]	[0.06]
Rating = 3	0.74***	0.54***	0.64***	0.48***
	[0.08]	[0.08]	[0.07]	[0.08]
Rating = 4	1.09***	0.93***	1.01***	0.87***
	[0.11]	[0.11]	[0.10]	[0.10]
Rating = 5	1.34***	1.16***	1.33***	1.16***
	[0.16]	[0.15]	[0.16]	[0.15]
Post-2008 × (Large BHC)			0.61***	0.65***
			[0.09]	[0.10]
Post-2008 × (Small BHC)			-0.32***	-0.19***
			[0.05]	[0.05]
District, Ongoing FEs?	Yes	Yes	Yes	Yes
Date FEs?	Yes	Yes	No	No
Bank FEs?	No	Yes	No	Yes
Adj. R2	0.61	0.68	0.61	0.68
Obs.	17969	17969	17969	17969
Distinct BHCs	785	785	785	785
Dates	19981.20144	19981.20144	19981.20144	19981.20144

Table 8: **Substitution effects.** This table extends the baseline regressions specification in Table 2 to include each bank's district share of other assets in distress (banks rated 3, 4 or 5). Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets. Sample is 1998Q1-2014Q4. Standard errors clustered by bank reported in brackets. Significance: $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

	Log(Hours)			
	(1)	(2)	(3)	(4)
Log(Assets)	0.92***	0.66***	0.89***	0.61***
	[0.02]	[0.10]	[0.02]	[0.10]
Rating = 2	0.17***	0.07	0.17***	0.08
	[0.05]	[0.06]	[0.05]	[0.06]
Rating = 3	0.74***	0.51***	0.72***	0.52***
	[0.08]	[0.08]	[0.08]	[0.08]
Rating = 4	1.08***	0.87***	1.10***	0.92***
	[0.10]	[0.11]	[0.11]	[0.11]
Rating = 5	1.40***	1.14***	1.43***	1.20***
	[0.16]	[0.15]	[0.16]	[0.15]
Share Distress $\times$ (Small BHC)	-0.62***	-0.54***	-0.40***	-0.31***
	[0.10]	[0.10]	[0.10]	[0.10]
Share Distress $\times$ (Large BHC)	0.87***	0.72***	0.12	0.12
	[0.17]	[0.15]	[0.18]	[0.16]
Post-2008 $\times$ (Large BHC)			0.76***	0.69***
			[0.11]	[0.12]
District, Date, Ongoing FEs?	Yes	Yes	Yes	Yes
Bank FEs?	No	Yes	No	Yes
Adj. R2	0.61	0.68	0.62	0.69
Obs.	17943	17943	17943	17943
Distinct BHCs	780	780	780	780

Figure 4: **Relation between supervisory hours and assets.** This figure presents a binned scatter plot and the fitted line of supervisory hours on BHC size obtained when controlling for rating, district and date-quarter dummies (column 1 of Table 2, left-panel) as well as bank fixed effects (column 2 of Table 2, right-panel).

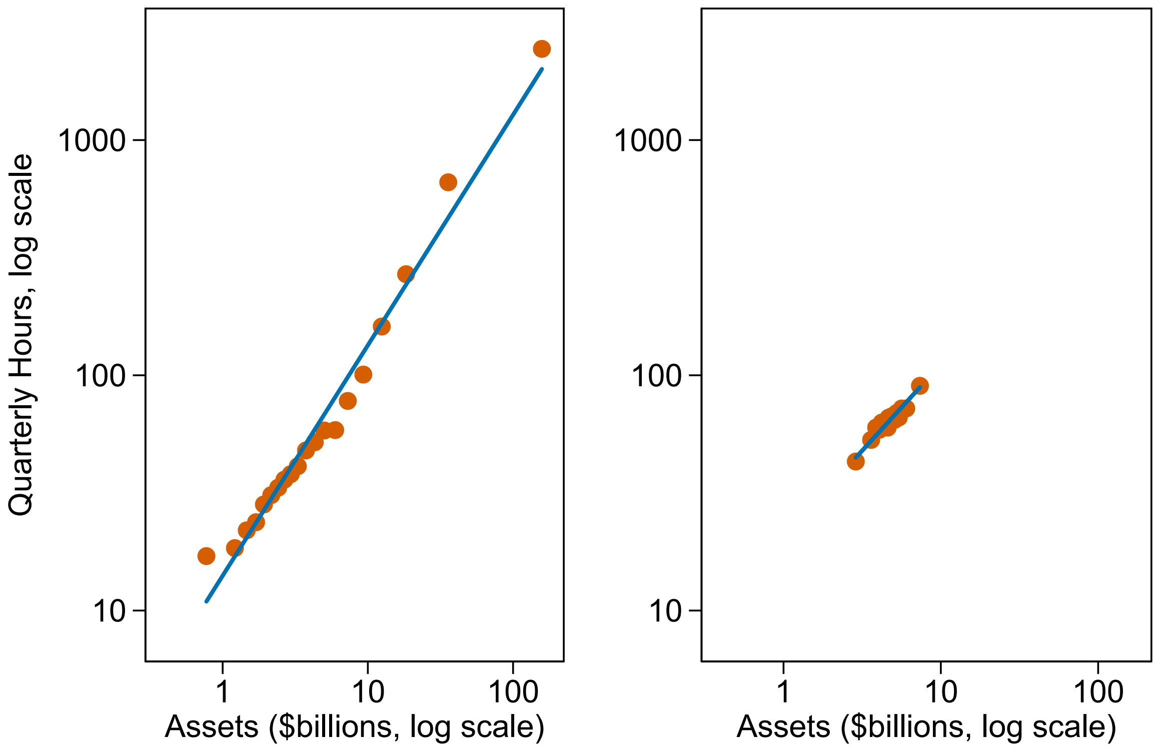


Figure 5: **Relation between OCC semiannual fees and assets.** This figure presents a binned scatter plot and the fitted line of OCC fees on commercial bank assets as computed in Table 4 on assets.

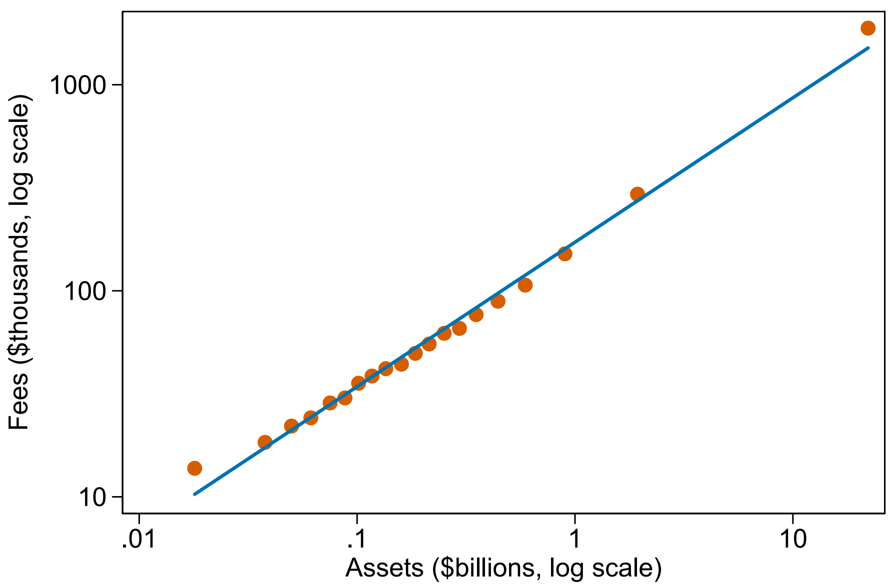
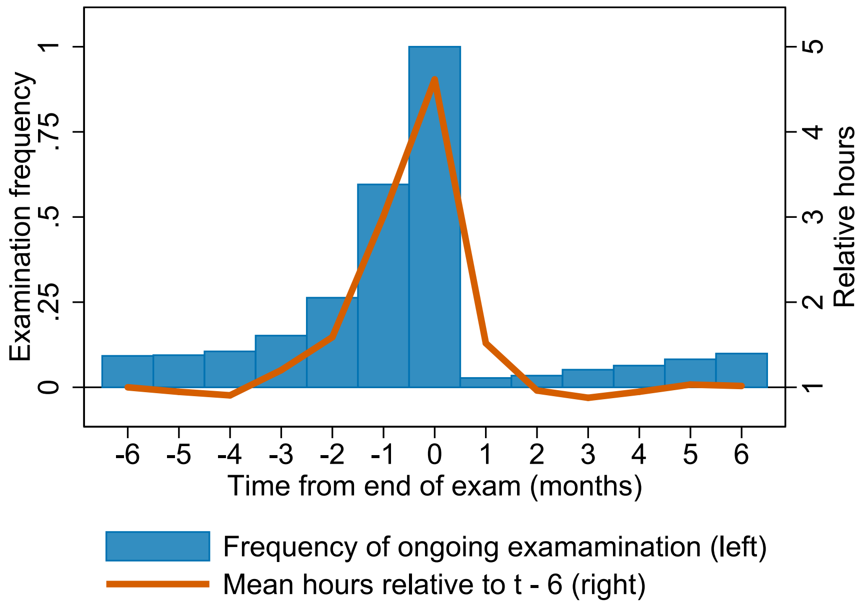


Figure 6: **Hours around examinations.** This figure shows the average time pattern of supervisory hours around the month when a bank's full-scope examination ends. Hours (orange line) are plotted relative to their level at month t-6. Vertical bars denote the frequency of full-scope or other supervisory examinations. Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets.

(a) Small BHCs



(b) Large BHCs

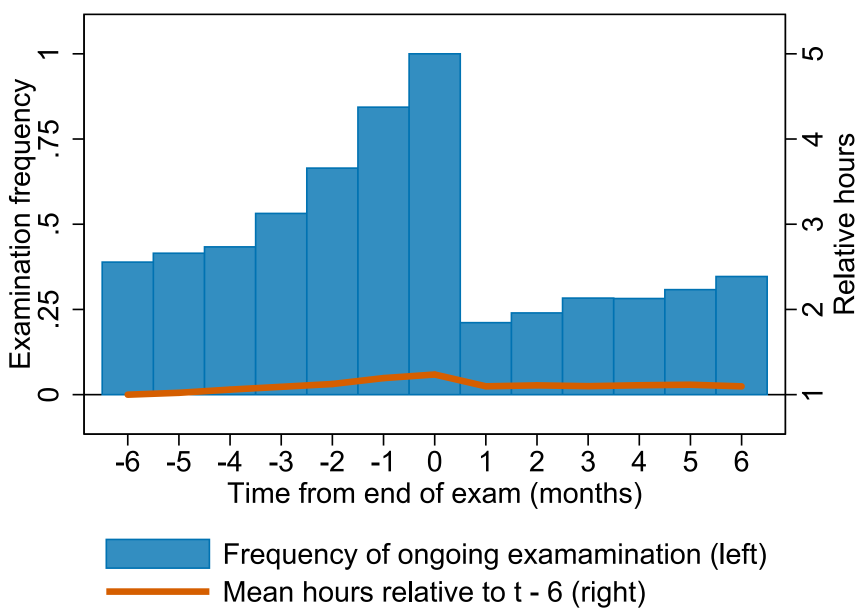


Figure 7: **Tier 1 ratio against bank assets.** This figure presents a binned scatter plot and the fitted line of tier 1 capital ratio (percent) to BHC assets for the subsample pre-2009 (orange) and post-2009 (blue).

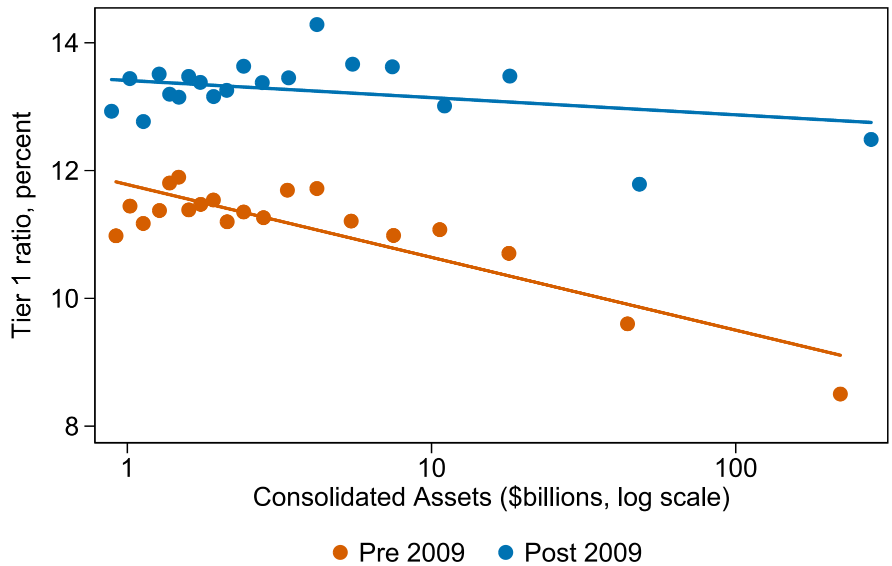
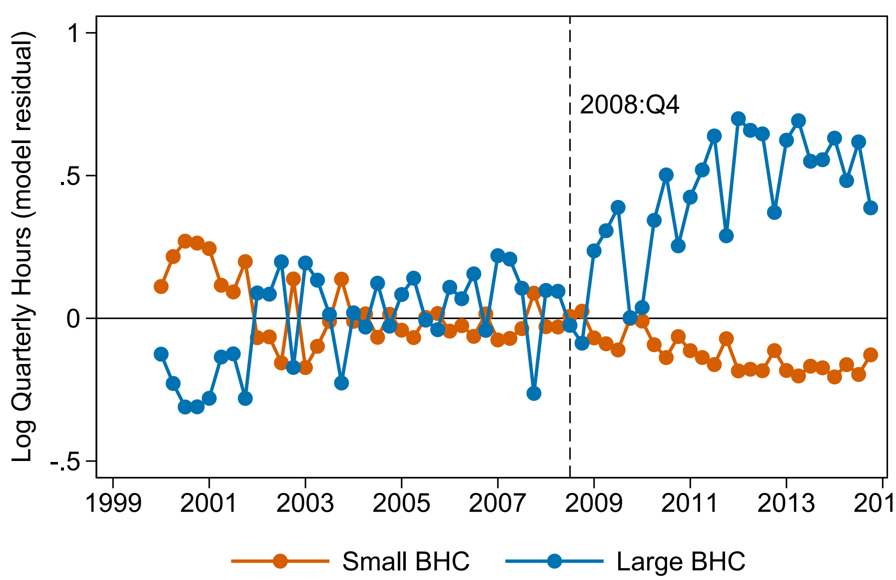


Figure 8: **Average supervisory hour residuals over time.** This figure shows the time series of residuals of the baseline regression (column 1, Table 2) of the log of supervisory hours on the log(assets) and other controls. The residuals are averaged by date and whether the BHC is large or small on that date. Large (small) BHCs are defined in terms of assets above (below) a $10 billion asset threshold adjusted for the growth rate of industry assets.



A Proofs and additional material

Proof of Proposition 1. Since $\underline{a} < \overline{a}$ and $\overline{s}_2 < \underline{s}_2$ and F is decreasing in a, s_2 for $x = \ell$, we have:

$$F(\ell \,|\, \underline{a}, \overline{s}_2) = \max_{a, s_2^*} F(\ell \,|\, a, s_2)$$

Given $\underline{s}_2 > \overline{s}_2$, we have:

$$F(\ell \,|\, \overline{a}, \underline{s}_2) < F(\ell \,|\, \overline{a}, \overline{s}_2)$$

Next, consider the case of a perfectly informative signal. Then the first-order condition (3) yields:

$$-F_{s_2}(\ell \,|\, \underline{a}, \underline{s}_2) \left(P(A) + N(A)\right) = \mathcal{C}^2_{s_2}(\underline{s}_2, A) \tag{16}$$

$$\text{and} \quad -F_{s_2}(\ell \,|\, \overline{a}, \overline{s}_2) \left(P(A) + N(A)\right) = \mathcal{C}^2_{s_2}(\overline{s}_2, A) \tag{17}$$

With $\underline{s}_2 > \overline{s}_2$, we have $\mathcal{C}^2_{s_2}(\underline{s}_2, A) > \mathcal{C}^2_{s_2}(\overline{s}_2, A)$ so the left-hand sides of (16) and (17) imply:

$$-F_{s_2}(\ell \,|\, \underline{a}, \underline{s}_2) > -F_{s_2}(\ell \,|\, \overline{a}, \overline{s}_2) \tag{18}$$

Since a and s_2 are perfect substitutes, we have $F(x \,|\, a, s_2) = \hat{F}(x \,|\, a + s_2)$. Then (18) implies that $\underline{a} + \underline{s}_2 < \overline{a} + \overline{s}_2$ and therefore

$$F(\ell \,|\, \underline{a}, \underline{s}_2) > F(\ell \,|\, \overline{a}, \overline{s}_2)$$

With a less than perfect signal, $\underline{s}_2$ is slightly lower and $\overline{s}_2$ slightly higher which strengthens the inequality. □

Detail on Conjecture 5. To get a sense of how the marginal benefit of more precision varies with bank size, note that the left-hand side of the first-order condition (8) is increasing in four differences in expected utility for the supervisor:

$$\begin{aligned}\mathcal{V}(\overline{a}, \overline{s}_2) - \mathcal{V}(\underline{a}, \underline{s}_2) = &\left(F(\ell \,|\, \underline{a}, \underline{s}_2) - F(\ell \,|\, \overline{a}, \overline{s}_2)\right) \left(P(A) + N(A)\right) \\ &+ w\left(h^2(\underline{s}_2, A) - h^2(\overline{s}_2, A)\right)\end{aligned} \tag{19}$$

$$\begin{aligned}\mathcal{V}(\overline{a}, \underline{s}_2) - \mathcal{V}(\underline{a}, \overline{s}_2) = &\left(F(\ell \,|\, \underline{a}, \overline{s}_2) - F(\ell \,|\, \overline{a}, \underline{s}_2)\right) \left(P(A) + N(A)\right) \\ &- w\left(h^2(\underline{s}_2, A) - h^2(\overline{s}_2, A)\right)\end{aligned} \tag{20}$$

$$\begin{aligned}\mathcal{V}(\overline{a}, \overline{s}_2) - \mathcal{V}(\overline{a}, \underline{s}_2) = &-\left(F(\ell \,|\, \overline{a}, \overline{s}_2) - F(\ell \,|\, \overline{a}, \underline{s}_2)\right) \left(P(A) + N(A)\right) \\ &+ w\left(h^2(\underline{s}_2, A) - h^2(\overline{s}_2, A)\right)\end{aligned} \tag{21}$$

$$\begin{aligned}\mathcal{V}(\underline{a}, \underline{s}_2) - \mathcal{V}(\underline{a}, \overline{s}_2) = &\left(F(\ell \,|\, \underline{a}, \overline{s}_2) - F(\ell \,|\, \underline{a}, \underline{s}_2)\right) \left(P(A) + N(A)\right) \\ &- w\left(h^2(\underline{s}_2, A) - h^2(\overline{s}_2, A)\right)\end{aligned} \tag{22}$$

All four utility differences consist of a difference in expected spillovers and a difference in supervisory costs. The first two differences, (19) and (20), capture the utility impact of a higher likelihood of action $\overline{a}$. This always leads to a lower default probability and therefore lower spillover losses; if the signal is right, (19), it also means cost savings while if the signal is wrong it is reduced by unnecessary additional costs, (20). The second two differences, (21) and (22), capture the utility impact of fewer supervisory mistakes. Note that "getting it right" leads to a *higher* default probability for $\overline{a}$ and a *lower* default probability for $\underline{a}$. Overall, we see that whether the benefit of precision (the left-hand side of (8)) is increasing in bank size, and sufficiently so that $d\lambda^*/dA > 0$, is not straightforward. Ceteris paribus, it is more likely the more the two spillovers, $N(A)$ and $P(A)$, increase with A. □

Proof of Proposition 2. We have assumed that $\lambda^* > 1/2$ and $G(\lambda^*) > 1/2$, and we know that $\Pr[\underline{a} \,|\, \underline{r}] > \Pr[\overline{a} \,|\, \underline{r}]$; therefore we have:

$$\big(1 - G(\lambda^*)\big)\, \lambda^* > G(\lambda^*)\, (1 - \lambda^*)$$

With the ranking of conditional default probabilities in Proposition 1 we can therefore rank

$$\begin{aligned} &\big(1 - G(\lambda^*)\big)\, \lambda^* F(\ell \,|\, \underline{a}, \underline{s}_2) > G(\lambda^*)\, (1 - \lambda^*)\, F(\ell \,|\, \overline{a}, \underline{s}_2) \\ \text{and} \quad & G(\lambda^*)\, \lambda^* F(\ell \,|\, \overline{a}, \overline{s}_2) > G(\lambda^*)\, (1 - \lambda^*)\, F(\ell \,|\, \overline{a}, \underline{s}_2). \end{aligned}$$

All other comparisons involve a mismatched ranking between the probabilities $\Pr[a]$, $\Pr[s_2 \,|\, a]$, and/or $F(\ell \,|\, a, s_2)$, making a ranking of the products $\Pr[a] \times \Pr[s_2 \,|\, a] \times F(\ell \,|\, a, s_2)$ impossible without further assumptions. □

B Detailed variable definitions

Hours and Rating: See the discussion in Section 3.

Ongoing Full Exam: A dummy indicator for whether a full-scope exam is ongoing for a holding company in a given quarter. Source: NED confidential data.

Assets $\geq$ 10bn: An indicator for whether a bank has total assets greater than \$10 billion. The size threshold is deflated by the aggregate growth rate in BHC assets to account for growing size of BHCs. Asset data from FR-Y9C, item BHCK2170.

Tier 1 Capital Ratio: Tier 1 Capital Ratio from FR-Y9C. Tier 1 Risk-based capital divided by risk-weighted assets from FR-Y9C. Basel I (pre-2014) BHCK8274/BHCKa223 Basel III (post-2014; including 2014 for advanced-approaches firms) BHCA8274/BHCAA223

N Bank Employees: The number of employees. Employee data from FR-Y9C, item BHCK4150.

HHI Asset Concentration: Measure of business concentration using a Herfindahl-Hirschman Index (HHI) for asset concentration. See Kovner, Vickery and Zhou(2014). Data from FR-Y9C. Sum of squares of the following asset types (all as proportion of total assets, item BHCK2170):

- Total loans (BHCK2122)
- Total trading assets (BHCK3545)
- Fed funds and repo assets (pre-2002: BHCK1350; post 2002: BHDMB987 + BHCKB989)
- Investment securities (BHCK1754 + BHCK1773)
- The book value (not to exceed fair value), less accumulated depreciation, if any, of all real estate other than bank premises actually owned by the bank and its consolidated subsidiaries (pre-2000: BHCK2744 + BHCK2745; post-2000: BHCK2150)
- Premises and other fixed assets (BHCK2145)
- Investments in unconsolidated subsidiaries (BHCK2130)
- Direct and indirect investments in real estate ventures (BHCK3656)
- Cash (BHCK0081 + BHCK0395 + BHCK0397)
- All other assets (pre-2001: BHCK3164 + BHCKb026 + BHCK5507 + BHCK2160; post-2001 and pre-2006: BHCK0426 + BHCK2160 + BHCK2155; post-2006: BHCK0426 + BHCK2160)

Loan Share: Proportion of total BHC assets that are loans. Data from FR-Y9C, BHCK2122, BHCK2170

N Bank Subsidiaries: The count of all legal subsidiaries of a given BHC, computed using NIC data. This series was first contructed by Nicola Cetorelli and Samuel Stern at the NY Fed. More detailed is available in Cetorelli and Stern (2015) "Same Name, New Businesses: Evolution in the Bank Holding Company" published in Liberty Street Economics on September 28, 2015.