

working paper

1718

Hierarchical Bank Supervision

Rafael Repullo

November 2017

Hierarchical Bank Supervision

Abstract

This paper presents a model in which a central and a local supervisor contribute their efforts to obtain information on the solvency of a local bank, which is then used by the central supervisor to decide on its early liquidation. This hierarchical model is contrasted with the alternatives of decentralized and centralized supervision, where only the local or the central supervisor collects information and decides on liquidation. The local supervisor has a higher bias against liquidation (supervisory capture) and a lower cost of getting local information (proximity). Hierarchical supervision is the optimal institutional design when the bias of the local supervisor is high but not too high and the costs of getting local information from the center are low but not too low. With low (high) bias and high (low) cost it is better to concentrate all responsibilities in the local (central) supervisor.

JEL Codes: G21, G23, D02.

Keywords: Centralized vs decentralized supervision, strategic information acquisition, bank solvency, bank liquidation, supervisory capture, optimal institutional design.

Rafael Repullo
CEMFI
repullo@cemfi.es

Acknowledgement

Presidential Address delivered at the Symposium of the Spanish Economic Association in Bilbao in December 2016. I would like to thank the comments of Óscar Arce, Manuel Arellano, Mahmoud Elamin, Luis Garicano, Charlie Kahn, Luc Laeven, David Lucca, Simone Manganelli, David Martinez-Miera, David Pérez-Castrillo, Enrique Sentana, André Sapir, Joel Sobel, Javier Suarez, Xavier Torres, Skander Van den Heuvel, Wolf Wagner, Ansgar Walther, an anonymous referee, as well as seminar audiences at CEMFI, the Bank of Portugal, the Bank of Spain, the European Central Bank, the Federal Reserve Banks of Boston, Cleveland and New York, the University of Edinburgh, and the 2017 Congress of the European Economic Association. I am also grateful to Wenbo Meng, Álvaro Remesal, and Yixin Yao for their excellent research assistance. Financial support from the Spanish Ministry of Economy and Competitiveness, Grant No. ECO2014-59262-P, is gratefully acknowledged.

1 Introduction

Bank supervision, unlike bank regulation, has not been until recently the subject of much academic interest. As stated by Eisenbach et al. (2016), regulation involves the establishment of rules under which banks operate, while supervision involves the assessment of safety and soundness of banks through monitoring, and the use of this information to request corrective actions. In contrast to regulation, that is based on verifiable information, supervision is about supervisory actions (partly) based on nonverifiable information.

This paper contributes to the theoretical literature on bank supervision by constructing a stylized model of a supervisor that collects nonverifiable information on the solvency of a bank and, on the basis of this information, decides on its early liquidation. The quality of the information on the bank's solvency depends on the intensity of supervision (the nonverifiable costly effort of the supervisor). I assume that the supervisor is not a social welfare maximizer. In particular, its payoff function incorporates a liquidation cost, which may be associated with either reputational concerns or supervisory capture (e.g. revolving doors). The paper characterizes the effort and the liquidation decisions of the supervisor, and shows that supervision will be more intense the lower the costs of supervisory effort and the lower the supervisory bias against liquidation.

The model can be interpreted as a model of *decentralized supervision*, in which the bank is a local bank and the supervisor is a local supervisor, or as a model of *centralized supervision*, in which the bank is still a local bank but the supervisor is a central (or supranational) supervisor. It can also be used as a building block for a model of *hierarchical supervision*, in which the central and the local supervisors jointly supervise the bank in order to observe a nonverifiable signal of the bank's solvency, and then the central supervisor decides on the liquidation of the bank. Under hierarchical supervision, the central and the local supervisors simultaneously choose their efforts, so they will be playing a game. The Nash equilibrium of this game describes the outcome of the hierarchical supervision model.

The main contribution of the paper is to characterize the conditions under which one of the three institutional arrangements, namely decentralized, hierarchical, and centralized

supervision, dominates in welfare terms the other two. The analysis is based on two key assumptions: (i) *the cost of effort is higher for the central than for the local supervisor*, and (ii) *the cost of liquidating the bank is lower for the central supervisor than for the local supervisor*. The first assumption may be justified by reference to distance between the central supervisor and the local bank. In the words of Torres (2015): “The central supervisor has informational disadvantages relative to the national authorities, due to their better knowledge of banks, banking systems and regulatory frameworks, as well as their geographical and cultural proximity to them.” The second assumption may be justified by reference to the looser connections between the central supervisor and the bank. In the words of Torres (2015): “The existence of a supranational supervisor allows to increase the distance between supervisors and national lobbies and politicians, which in principle should reduce the risk of supervisors implementing excessively lax policies.”¹

The results show that hierarchical supervision dominates decentralized supervision when the bias of the local supervisor is high and the costs of getting local information from the center are low. But when these forces exceed certain threshold, it is better to concentrate all responsibilities in the central supervisor. The trade-off underlying these results is clear: Decentralized supervision is better because the local supervisor finds it cheaper to gather information, but it is worse because its objective function is biased against liquidation.² The results also show that hierarchical supervision is more likely to dominate when bank profitability is low (e.g. as a result of high competition) and when bank risk-taking is high (e.g. as a result of soft regulation).

Interestingly, the model of hierarchical supervision is isomorphic to a model in which

¹Lambert (2017) analyzes the relationship between bank lobbying and supervisory decisions using bank-level information on the universe of commercial and savings banks in the US. He concludes that “these results appear rather inconsistent with an information based explanation of bank lobbying, but consistent with the theory of regulatory capture.”

²Bian et al. (2017) analyze bailouts in German savings banks, showing that banks that are bailed out by local politicians (decentralized resolution) experience less restructuring and perform considerably worse than banks that are bailed out by state associations (centralized resolution). In comparing the two alternatives they note the following trade-off: “On the one hand, close proximity has the potential to improve the decision making process, as it provides politicians with good information about banks that get into distress. On the other hand, close proximity could imply that politicians’ personal considerations distort the decision making process, which is clearly undesirable.” Such trade-off is similar to the one considered in this paper.

the central supervisor gets a signal of the bank’s solvency, the local supervisor gets another signal which truthfully reports to the central supervisor, who then decides on the liquidation of the bank. The original model corresponds to an institutional arrangement in which the supervisors work in teams, while the alternative model corresponds to an arrangement in which the supervisors work independently, but there is no problem of strategic information transmission à la Crawford and Sobel (1982).³

The results provides a rationale for the design of the Single Supervisory Mechanism (SSM), the new structure of bank supervision in Europe that comprises the European Central Bank (ECB) and the national supervisory authorities of the participating countries. Currently, the ECB is responsible for the supervision of significant (i.e. large) banks that comprise about 80% of the banking assets of the euro area. For these banks, supervision is carried out in cooperation with the national supervisors via the so-called *Joint Supervisory Teams*. National supervisors are in charge of the less significant (i.e. small) banks. To the extent that (i) the cost advantage of local supervisors is smaller for larger, more complex banks, and (ii) supervisory capture is more relevant for larger banks, the results of the model are consistent with the design of the SSM.

The model can also shed light on issues related to the organization of supervision in jurisdictions in which multiple agencies are involved in supervising banks. For example, state-chartered banks in the US are under a dual supervisory framework involving both federal and local supervisors.⁴ Also, the Federal Reserve is responsible for the supervision of bank holding companies, but subsidiaries may be supervised by other federal agencies.

Two extensions of the model are discussed. The first one examines what happens when the supervisory bias is linked to being responsible for the liquidation decision. This means that moving from decentralized to hierarchical supervision implies a reduction in the liq-

³See Sobel (2013) for a recent survey. Holthausen and Rønde (2004) contain an application of the results in Crawford and Sobel (1982) to the modeling of cooperation between host and home country supervisors of a multinational bank.

⁴Agarwal et al. (2014) show that federal supervisors are systematically tougher than state supervisors, downgrading supervisory ratings almost twice as frequently as do state supervisors. However, they do not find support for supervisor self-interest, which includes “revolving doors,” as a reason for leniency of state supervisors.

liquidation cost of the local supervisor and an increase in the liquidation cost of the central supervisor. In this case, the results show that there will be an increase in the (cheaper) effort of the local supervisor that will compensate the reduction in the effort of the central supervisor, so hierarchical supervision is more likely to dominate.

A second extension shows that limiting the size of the central supervisor in the hierarchical supervision setup is welfare improving. The result is closely related to the well-known result that a Stackelberg leader that optimizes over the reaction function of the other agent does better than by playing its Nash equilibrium strategy. The intuition is that putting the central supervisor in a situation of overload forces the local supervisor to increase its (cheaper) effort in a way that compensates the reduction in the effort of the central supervisor.⁵

Literature review Carletti et al. (2016) explore the working of a supervisory structure in which a centralized agency has legal power over decisions regarding banks, but has to rely on biased local supervisors to collect the information necessary to act. They focus on how this institutional design affects supervisors' incentives to collect information and on how this, in turn, influences bank behavior, showing that when the agency problem between the central and the local supervisor is severe, centralized supervision leads to lower information collection and increased risk-taking.⁶

Colliard (2015) considers a model in which local supervisors are more lenient, so that banks also have weaker incentives to hide information from them. These two forces can make a joint supervisory architecture optimal. However, more centralized supervision encourages banks to integrate more cross-border. Due to this complementarity, the economy can be trapped in an inferior equilibrium with both too little central supervision and too little financial integration.

Calzolari et al. (2016) study the impact of shared liability and deposit insurance arrangements on supervisors' incentives to acquire information on the activities of multinational

⁵This is reminiscent of results in the literature on the theory of organizations. For example, Aghion and Tirole (1997) write: "It is always optimal for the firm to be in a situation of overload so as to credibly commit to rewarding initiative."

⁶Gopalan et al. (2017) examine the effects of the closure of a US bank regulator's field offices, concluding that distancing supervisors from banks to prevent capture can be costly as it increases bank risk.

banks, showing that centralized supervision can induce these banks to expand abroad through branches rather than subsidiaries.

In an earlier contribution, Boyer and Ponce (2012) analyze whether banking supervision responsibilities should be concentrated in the hands of a single supervisor, showing that splitting supervisory powers among different supervisors is a superior arrangement in terms of welfare when the capture of supervisors by bankers is a concern.

The model in the paper may be considered as a special case of models in the literature on the theory of organizations.⁷ As noted in the seminal paper by Jensen and Meckling (1992), “The assignment of decision rights influences incentives to acquire information. (...) Determining the optimal level of decentralization requires balancing the costs of bad decisions owing to poor information and those owing to inconsistent objectives.” It is also closely related to the recent literature on strategic information acquisition, which departs from the literature on strategic information transmission by assuming that information is not exogenously given, but is obtained through costly effort; see Argenziano et al. (2016).

Structure of the paper Section 2 presents the basic model of bank supervision in which a supervisor collects information on the solvency of a bank and, on the basis of this information, decides on its early liquidation. This setup may be interpreted as a model of decentralized supervision in which supervisory responsibilities are allocated to a local supervisor or a model of centralized supervision in which supervisory responsibilities are allocated to a central supervisor. Building on this setup, Section 3 presents the model of hierarchical supervision in which a central and a local supervisor jointly collect information and then the central supervisor decides on liquidation. Section 4 compares in terms of welfare three possible institutional arrangements: decentralized, hierarchical, and centralized supervision. Section 5 contains the extensions, and Section 6 the concluding remarks. The proofs of the analytical results are in the Appendix.

⁷See Jensen and Meckling (1992), Aghion and Tirole (1997), Dessein (2002), Alonso et al. (2008), and Dessein et al. (2010), among many others.

2 Model Setup

Consider an economy with three dates ($t = 0, 1, 2$) and two agents: a bank and a supervisor. The bank raises a unit amount of deposits at $t = 0$, and invests them in an asset that has a random *final return* R at $t = 2$. The asset can be liquidated at $t = 1$, in which case it yields a random *liquidation return* L . Deposits are insured and the deposit rate is normalized to zero.

It is assumed that

$$\begin{bmatrix} L \\ R \end{bmatrix} \sim N\left(\bar{R} \begin{bmatrix} a \\ 1 \end{bmatrix}, \sigma^2 \begin{bmatrix} b & c \\ c & 1 \end{bmatrix}\right), \quad (1)$$

where $\bar{R} > 1$, $a < 1$, $b < 1$, and $c > 0$. Moreover, to ensure that the covariance matrix is positive-definite it must be the case that $c^2 < b$, so $b < 1$ implies $c < 1$. Thus, the expected final return $E(R) = \bar{R}$ is greater than the unit face value of the deposits, and it is also greater than the expected liquidation return $E(L) = a\bar{R}$. Moreover, the final return has a higher variance than the liquidation return, and both returns are positively correlated. Except for the assumption of normality, which is made for tractability, the rest of the assumptions capture realistic features of the distribution of banks' returns.

The supervisor chooses at $t = 0$ the *intensity of supervision* e (nonverifiable effort of the supervisor), which leads to the observation at $t = 1$ of a *nonverifiable signal*

$$s = R + \varepsilon \quad (2)$$

on the final return of the bank's investment. The noise term ε is independent of L and R , and has a distribution $N(0, \sigma^2/e)$.⁸ Thus, the effort e of the supervisor increases the precision (inverse of the variance) of the noise term.

From here it follows that

$$\begin{bmatrix} L \\ R \\ s \end{bmatrix} \sim N\left(\bar{R} \begin{bmatrix} a \\ 1 \\ 1 \end{bmatrix}, \sigma^2 \begin{bmatrix} b & c & c \\ c & 1 & 1 \\ c & 1 & 1 + e^{-1} \end{bmatrix}\right). \quad (3)$$

⁸This implies that $L = (a - c)\bar{R} + cR + u$, where $u \sim N(0, \sigma^2(b - c^2))$ and $Cov(u, R) = Cov(u, \varepsilon) = 0$. Thus, signal s contains information about L only inasmuch as it contains information about R .

By the properties of normal distributions we have

$$E(L \mid s) = a\bar{R} + \frac{c(s - \bar{R})}{1 + e^{-1}}, \quad (4)$$

$$E(R \mid s) = \bar{R} + \frac{s - \bar{R}}{1 + e^{-1}}. \quad (5)$$

Since $c < 1$ the slope of $E(L \mid s)$ is smaller than the slope of $E(R \mid s)$, which implies

$$E(L \mid s) > E(R \mid s) \text{ if and only if } s < \bar{s},$$

where

$$\bar{s} = \left[1 - \frac{1-a}{1-c}(1 + e^{-1}) \right] \bar{R} \quad (6)$$

is the *efficient liquidation threshold*. Note that higher supervisory effort e increases the efficient liquidation threshold $\bar{s}$. Also, note that for $e = 0$ we have $\bar{s} = -\infty$, so with a completely noisy signal the bank should never be liquidated.

Substituting $\bar{s}$ from (6) into (4) and (5) gives

$$E(L \mid \bar{s}) = E(R \mid \bar{s}) = \frac{a-c}{1-c} \bar{R}.$$

I will assume that parameter values satisfy

$$\frac{a-c}{1-c} \bar{R} \leq 1. \quad (7)$$

This means that the efficient liquidation threshold $\bar{s}$ is such that the corresponding expected final return is less than or equal to the face value of the deposits, so efficient liquidation occurs only if the bank is effectively bankrupt.⁹

The supervisor chooses its effort e at $t = 0$, observes the signal s at $t = 1$, and based on this observation decides on the liquidation of the bank at this date. Supervisory effort is costly. Specifically, I assume that the *cost function* takes the simple quadratic form

$$c(e) = \gamma_0 + \frac{\gamma}{2} e^2, \quad (8)$$

where $\gamma_0 > 0$ is a fixed cost of setting up the supervisory arrangements, and $\gamma > 0$ is the key cost of effort parameter.

⁹Interestingly, $E(L \mid \bar{s}) = E(R \mid \bar{s})$ does not depend on supervisory effort e .

The supervisor liquidates the bank at $t = 1$ if

$$E(L | s) - \delta > E(R | s), \quad (9)$$

where $\delta > 0$ is a *liquidation cost*. This cost may be associated with either reputational concerns or supervisory capture (e.g. revolving doors). Thus, the supervisor liquidates the bank when the social benefits of liquidation, which are $E(L | s) - E(R | s)$, are greater than the liquidation cost δ .

Substituting (4) and (5) into (9) implies that the bank will be liquidated by the supervisor when $s < \hat{s}$, where

$$\hat{s} = \bar{s} - \frac{\delta}{1 - c}(1 + e^{-1}) \quad (10)$$

is the *supervisor's liquidation threshold*.¹⁰

Figure 1 shows the determination of the efficient liquidation threshold $\bar{s}$ by the intersection of the lines $E(L | s)$ and $E(R | s)$, and of the supervisor's liquidation threshold $\hat{s}$ by the intersection of the lines $E(L | s) - \delta$ and $E(R | s)$. For signals in the range between $\hat{s}$ and $\bar{s}$ the supervisor does not liquidate the bank when it would be efficient to do so.

The supervisor's effort decision at $t = 0$ is obtained by maximizing its expected payoff at $t = 1$, given by $E[\max\{E(L | s) - \delta, E(R | s)\}]$, net of the cost of effort $c(e)$, that is

$$\hat{e} = \arg \max_e v(e),$$

where

$$v(e) = \int_{-\infty}^{\hat{s}} [E(L | s) - \delta] dF(s) + \int_{\hat{s}}^{\infty} E(R | s) dF(s) - c(e), \quad (11)$$

$F(s)$ denotes the cdf of the signal s , and the liquidation threshold $\hat{s}$ is given by (10).

The following result provides a closed form expression for $v(e)$. For the result it is convenient to write $s = E(s) + SD(s)x$, where $x \sim N(0, 1)$, and define $\hat{x}$ to be such that $\hat{s} = E(s) + SD(s)\hat{x}$.

¹⁰Notice that, by (5) and (7), $\hat{s} < \bar{s}$ implies $E(R | \hat{s}) < E(R | \bar{s}) \leq 1$, so the supervisor liquidates the bank only if it is effectively bankrupt.

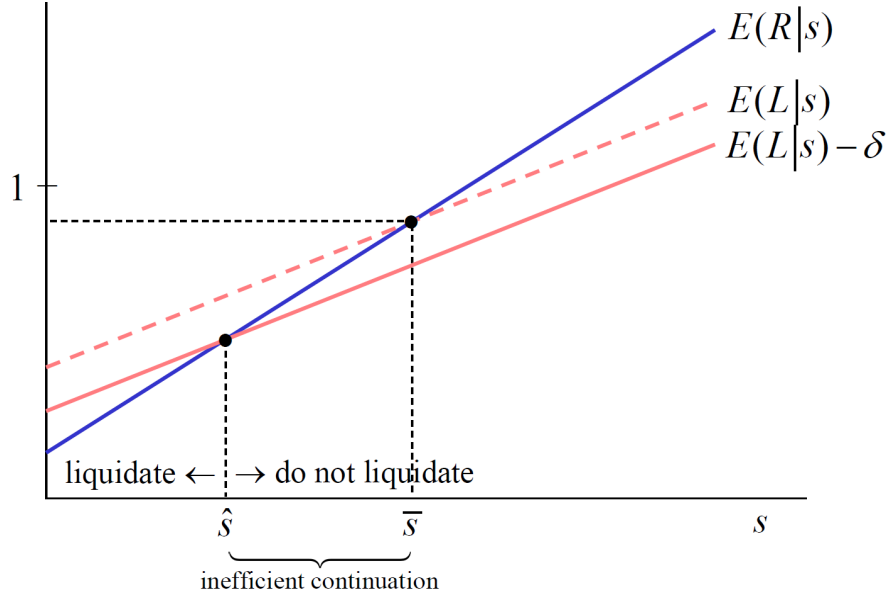


Figure 1. Supervisor's liquidation threshold

Proposition 1 *The supervisor's payoff function may be written as*

$$v(e) = \bar{R} - [(1-a)\bar{R} + \delta] \left[\Phi(\hat{x}) + \frac{\phi(\hat{x})}{\hat{x}} \right] - c(e), \quad (12)$$

where $\Phi(x)$ is the normal cdf and $\phi(x)$ is the normal density, and

$$\hat{x} = -\frac{[(1-a)\bar{R} + \delta]}{(1-c)\sigma} (1 + e^{-1})^{1/2}. \quad (13)$$

Figure 2 plots the payoff function of the supervisor $v(e)$ for the parameters that will be used in the numerical analysis below.¹¹ For $e = 0$ we have $\hat{x} = -\infty$, so the bank is never liquidated, and $v(0) = \bar{R} - \gamma_0$ (the expected final return minus the fixed cost of supervision). For $e = \infty$ we have $v(e) = -\infty$, since the expected payoff at $t = 1$ is clearly bounded while the cost of supervisory effort goes to infinity. The function $v(e)$ is initially convex, and then becomes concave, so there may be corner solutions with $\hat{e} = 0$ or interior solutions with $\hat{e} > 0$.¹²

¹¹These values are $\bar{R} = 1.25$, $a = 0.9$, $c = 0.5$, $\sigma = 0.25$, $\delta = 0.05$, and $\gamma_0 = \gamma = 0.001$. Notice that these parameters satisfy assumption (7).

¹²The nonconcavity of the function $v(e)$ is a special case of the result in Radner and Stiglitz (1984).

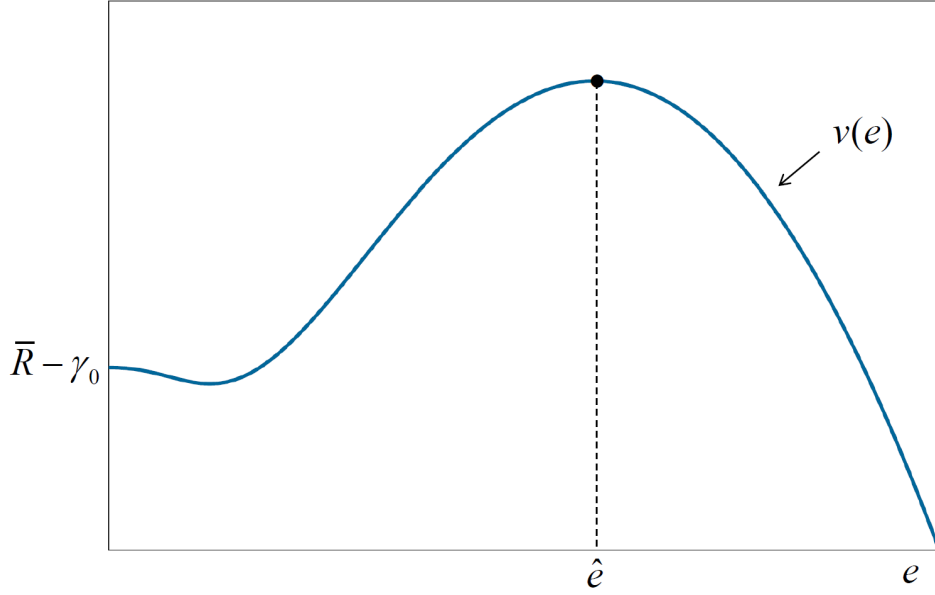


Figure 2. Supervisor's payoff function

The following result presents some comparative statics for the case where the solution is interior. In particular, it shows that supervisory effort $\hat{e}$ is decreasing in the cost of effort parameter γ and in the liquidation cost δ incurred by the supervisor. It also shows that $\hat{e}$ is decreasing in the expected return $\bar{R}$ and increasing in the standard deviation σ of the bank's investment return.

Proposition 2 *Whenever the supervisor chooses a positive level of effort $\hat{e}$ we have*

$$\frac{\partial \hat{e}}{\partial \gamma} < 0, \quad \frac{\partial \hat{e}}{\partial \delta} < 0, \quad \frac{\partial \hat{e}}{\partial \bar{R}} < 0, \quad \text{and} \quad \frac{\partial \hat{e}}{\partial \sigma} > 0.$$

Figure 3 illustrates these results. Panels A and B show that increases in the cost of effort parameter γ and in the liquidation cost δ reduce supervisory effort $\hat{e}$, which jumps to zero for sufficiently high values of γ and δ . Panel C shows that increases in the expected return $\bar{R}$ reduce supervisory effort $\hat{e}$, which jumps to zero for sufficiently high values of $\bar{R}$. Finally, Panel D shows supervisory effort $\hat{e}$ is zero for sufficiently low values of the standard deviation σ , jumping to positive and increasing levels for values of σ beyond a critical point.

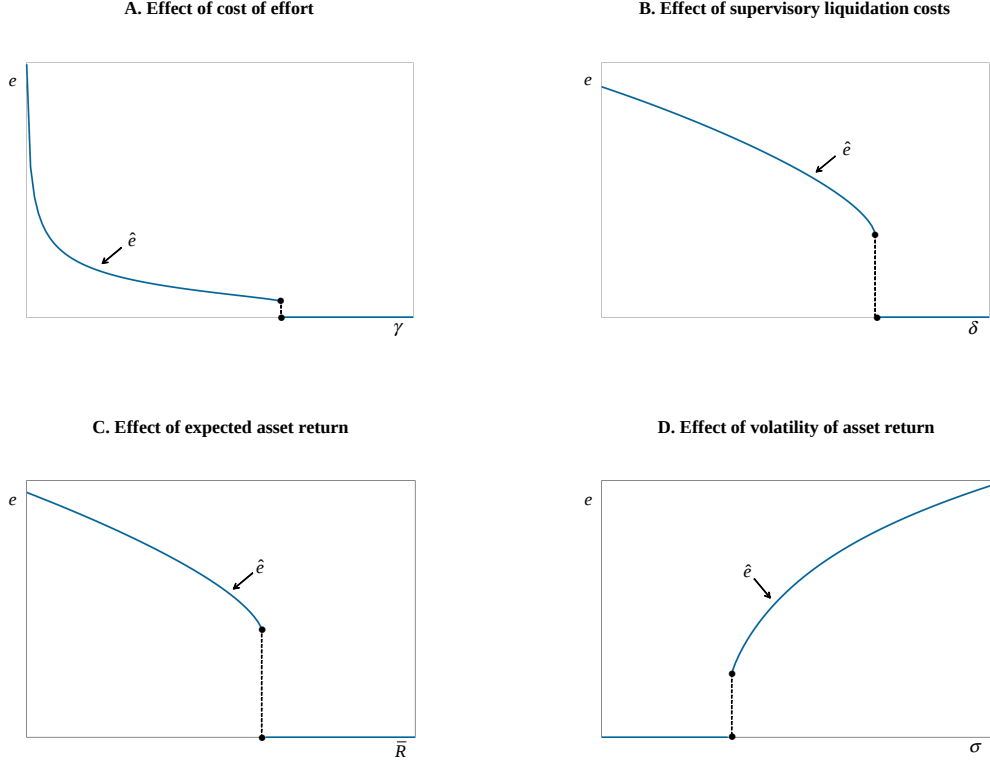


Figure 3. Comparative statics

The comparative statics results are not surprising. Supervision will be more intense when banks are easier to supervise (lower γ), or less profitable (lower $\bar{R}$), or riskier (higher σ). And it will be less intense when the supervisor is closer to lobbies and pressure groups (higher δ) that always prefer delaying intervention and gambling for resurrection.

Summing up, I have set up a simple model of a bank and a supervisor in which the supervisor exerts costly effort in order to observe a nonverifiable signal of the bank's solvency, which is used to decide the bank's early liquidation. Importantly, the supervisor is not a social welfare maximizer—it has a bias against liquidation. I have characterized the supervisor's effort decision, and derive some comparative statics results on its determinants.

The model may be interpreted as a model of decentralized supervision, in which the bank is a local bank and the supervisor is a local supervisor, or as a model of centralized supervision, in which the bank is still a local bank but the supervisor is a central (or supranational) supervisor. In this setup, it would be reasonable to assume that the cost of effort is higher

for the central supervisor than for the local supervisor, an assumption that may be justified by reference to geographical as well as cultural distance between the central supervisor and the local bank. It may also be reasonable to assume that the cost of liquidating the bank is lower for the central supervisor than for the local supervisor, an assumption that may be justified by reference to the looser connections between the central supervisor and national lobbies and pressure groups. The trade-off between the higher costs of supervision and the lower incidence of supervisory capture will be examined in Section 4. But before doing this, the following section presents a model of hierarchical supervision in which the central and the local supervisors jointly supervise the bank in order to obtain a nonverifiable signal of the bank's solvency, and then the central supervisor decides on the bank's early liquidation.

3 Hierarchical Supervision

Consider an economy with a local bank and two supervisors, a central and a local supervisor, denoted by subindices c and l . The supervisors independently choose at $t = 0$ nonverifiable efforts e_c and e_l , respectively, which leads to the observation at $t = 1$ of a single nonverifiable signal

$$s = R + \varepsilon \quad (14)$$

on the final return of the bank's investment.¹³ The noise term ε is independent of L and R , and has a distribution $N(0, \sigma^2(e_c + e_l)^{-1})$. Thus, the higher the efforts e_c and e_l the higher the precision of the noise term.

From here it follows that

$$\begin{bmatrix} L \\ R \\ s \end{bmatrix} \sim N \left(\bar{R} \begin{bmatrix} a \\ 1 \\ 1 \end{bmatrix}, \sigma^2 \begin{bmatrix} b & c & c \\ c & 1 & 1 \\ c & 1 & 1 + (e_c + e_l)^{-1} \end{bmatrix} \right). \quad (15)$$

Compared to the case of a single supervisor presented in Section 2, the only difference is that now the precision of the signal s depends on the sum $e_c + e_l$ of the efforts of the two

¹³As noted above, this setup is intended to capture the working of the Joint Supervisory Teams of the Single Supervisory Mechanism of the European Central Bank.

supervisors instead of on the effort e of the single supervisor.¹⁴

Supervisory effort is costly, and the cost of effort is assumed to be higher for the (distant) central supervisor than for the (close) local supervisor. Specifically, I assume that parameter γ in the cost function (8) takes the value γ_c for the central supervisor and γ_l for the local supervisor, where $\gamma_c > \gamma_l > 0$. The corresponding cost functions will be written $c_c(e_c)$ and $c_l(e_l)$, respectively.

The central supervisor decides whether to liquidate the bank based on the observation of the signal s at $t = 1$. I assume that the liquidation cost δ_c of the central supervisor is lower than that of the local supervisor δ_l . Moreover, to simplify the presentation I will assume that the central supervisor does not have a bias against early liquidation, so $\delta_c = 0$.¹⁵ Thus, the central supervisor liquidates the bank at $t = 1$ if

$$E(L \mid s) > E(R \mid s).$$

By the properties of normal distributions we have

$$E(L \mid s) = a\bar{R} + \frac{c(s - \bar{R})}{1 + (e_c + e_l)^{-1}}, \quad (16)$$

$$E(R \mid s) = \bar{R} + \frac{s - \bar{R}}{1 + (e_c + e_l)^{-1}}. \quad (17)$$

Hence, it follows that

$$E(L \mid s) > E(R \mid s) \text{ if and only if } s < s^*,$$

where

$$s^* = \left[1 - \frac{1-a}{1-c} [1 + (e_c + e_l)^{-1}] \right] \bar{R} \quad (18)$$

is the *efficient liquidation threshold*.

¹⁴A more general setup would be one in which the precision of the signal is $\alpha_c e_c + \alpha_l e_l + \alpha_{cl} e_c e_l$, where α_c and α_l capture the productivity of the effort of the two supervisors, and α_{cl} possible complementarities in their efforts. For simplicity, I assume $\alpha_c = \alpha_l = 1$ and $\alpha_{cl} = 0$.

¹⁵Section 5.1 analyzes a variation of the model where supervisory responsibilities increase the liquidation cost by an amount $\Delta > 0$, in which case hierarchical supervision entails a positive liquidation cost for the central supervisor.

As before, I assume that parameter values satisfy assumption (7). This implies that the threshold s^* is such that $E(L | s^*) = E(R | s^*) \leq 1$, so liquidation always takes place when the bank is effectively bankrupt.

At $t = 0$ the central and the local supervisors independently choose their efforts e_c and e_l , so they will be playing a game. I will characterize the Nash equilibrium of this game, and show some comparative static results.

The *payoff function of the central supervisor* is

$$v_c(e_c, e_l) = \int_{-\infty}^{s^*} E(L | s) dF(s) + \int_{s^*}^{\infty} E(R | s) dF(s) - c_c(e_c), \quad (19)$$

and the *payoff function of the local supervisor* is

$$v_l(e_c, e_l) = \int_{-\infty}^{s^*} [E(L | s) - \delta_l] dF(s) + \int_{s^*}^{\infty} E(R | s) dF(s) - c_l(e_l), \quad (20)$$

where $F(s)$ denotes the cdf of the signal s , and the liquidation threshold s^* is given by (18). It should be noted that in these expressions the central (local) supervisor does not take into account the cost of effort of the local (central) supervisor.¹⁶ More importantly, since the threshold s^* depends on the efforts e_c and e_l of the two supervisors, it must be the case that each supervisor observes the level of effort chosen by the other supervisor.¹⁷ This may be rationalized by assuming that the efforts of the supervisors are related to the quality of the staff that they independently choose ex-ante, but which can be observed ex-post.

The *reaction functions* of the two supervisors are given by

$$e_c(e_l) = \arg \max_{e_c} v_c(e_c, e_l), \quad (21)$$

$$e_l(e_c) = \arg \max_{e_l} v_l(e_c, e_l). \quad (22)$$

The intersection of these functions is a *Nash equilibrium* of the game played by the supervisors.

¹⁶This is without loss of generality, given that subtracting $c_l(e_l)$ and $c_c(e_c)$ in the payoff functions of the central and the local supervisor, respectively, would not change their behavior.

¹⁷What is key is that the central supervisor observes e_l , so in choosing the liquidation threshold s^* it can take into account the (out of equilibrium) deviations of the local supervisor. In contrast, the local supervisor need not observe e_c , since the (out of equilibrium) deviations of the central supervisor do not affect its behavior.

The following result provides closed form expressions for $v_c(e_c, e_l)$ and $v_l(e_c, e_l)$. As before it is convenient to write $s = E(s) + SD(s)x$, where $x \sim N(0, 1)$, and define x^* to be such that $s^* = E(s) + SD(s)x^*$.

Proposition 3 *The supervisors' payoff functions may be written as*

$$v_c(e_c, e_l) = \bar{R} - (1 - a)\bar{R} \left[\Phi(x^*) + \frac{\phi(x^*)}{x^*} \right] - c_c(e_c), \quad (23)$$

$$v_l(e_c, e_l) = \bar{R} - (1 - a)\bar{R} \left[\Phi(x^*) + \frac{\phi(x^*)}{x^*} \right] - \delta_l \Phi(x^*) - c_l(e_l), \quad (24)$$

where

$$x^* = -\frac{(1 - a)\bar{R}}{(1 - c)\sigma} [1 + (e_c + e_l)^{-1}]^{1/2}. \quad (25)$$

The analysis in the previous section shows that the supervisors' payoff functions are not everywhere concave. For this reason, the reaction functions (21) and (22) may have corner or interior solutions. In what follows, I will restrict attention to parameter values for which the solutions are interior.

Figure 4 shows the Nash equilibrium of the game played by the two supervisors, denoted (e_c^*, e_l^*) . Notice that for the chosen parameter values the reaction functions satisfy $e'_c(e_l) < 0$, $e'_l(e_c) < 0$, and $e'_c(e_l)e'_l(e_c) < 1$.¹⁸ That is, they are both downward sloping (strategic substitutes), and the slope of the reaction function of the local supervisor is (in absolute value) lower than that of the central supervisor, so the Nash equilibrium is stable. A sufficient condition for this to obtain is that the cost of effort parameters γ_c and γ_l are not too large.¹⁹

Figure 5 illustrates some comparative statics results of the game between the two supervisors. Panel A shows that increases in the cost of effort of the central supervisor γ_c shifts to the left its reaction function, leading to a reduction in the equilibrium effort e_c^* of the central

¹⁸In this figure the cost of effort parameters are $\gamma_c = 0.001$ and $\gamma_l = 0.00075$, and the liquidation cost parameters are $\delta_c = 0$ and $\delta_l = 0.025$.

¹⁹By the second order conditions we have $\partial^2 v_c / \partial^2 e_c^2 = h(e_c + e_l) < 0$ and $\partial^2 v_l / \partial^2 e_l^2 = g(e_c + e_l) < 0$. Since $e'_c(e_l) = -[h(e_c + e_l) + \gamma_c] / h(e_c + e_l)$ and $e'_l(e_c) = -[g(e_c + e_l) + \gamma_l] / g(e_c + e_l)$, a sufficient condition for $e'_c(e_l) < 0$ is that $\gamma_c < -h(e_c + e_l)$, and a sufficient condition for $e'_l(e_c) < 0$ is that $\gamma_l < -g(e_c + e_l)$. Moreover, these conditions imply $|e'_c(e_l)| < 1$ and $|e'_l(e_c)| < 1$, so $e'_c(e_l)e'_l(e_c) < 1$.

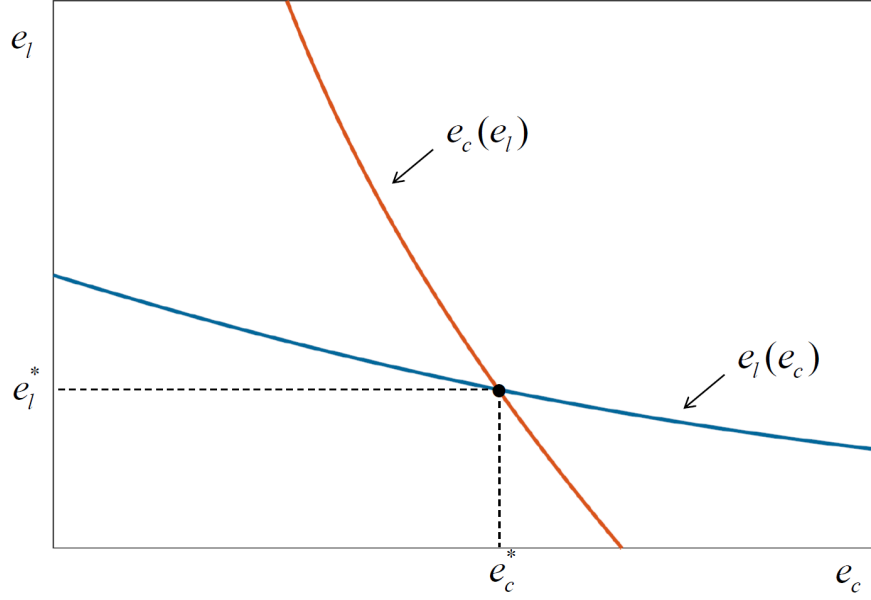


Figure 4. Nash equilibrium

supervisor and an increase in the equilibrium effort e_l^* of the local supervisor. Panel B shows that an increase in the liquidation cost δ_l of the local supervisor shifts down its reaction function, leading to a reduction in the equilibrium effort e_l^* of the local supervisor and an increase in the equilibrium effort e_c^* of the central supervisor. Panel C shows that an increase in the expected return $\bar{R}$ of the bank's investment shifts to the left the reaction function of the central supervisor and shifts down the reaction function of the local supervisor, leading to a reduction in the equilibrium effort of at least one of the two supervisors—although in the numerical results both e_c^* and e_l^* go down. Finally, Panel D shows that a decrease in the standard deviation σ of the bank's investment return shifts to the left the reaction function of the central supervisor and shifts down the reaction function of the local supervisor, leading to a reduction in the equilibrium effort of at least one of the two supervisors—although in the numerical results both e_c^* and e_l^* go down.

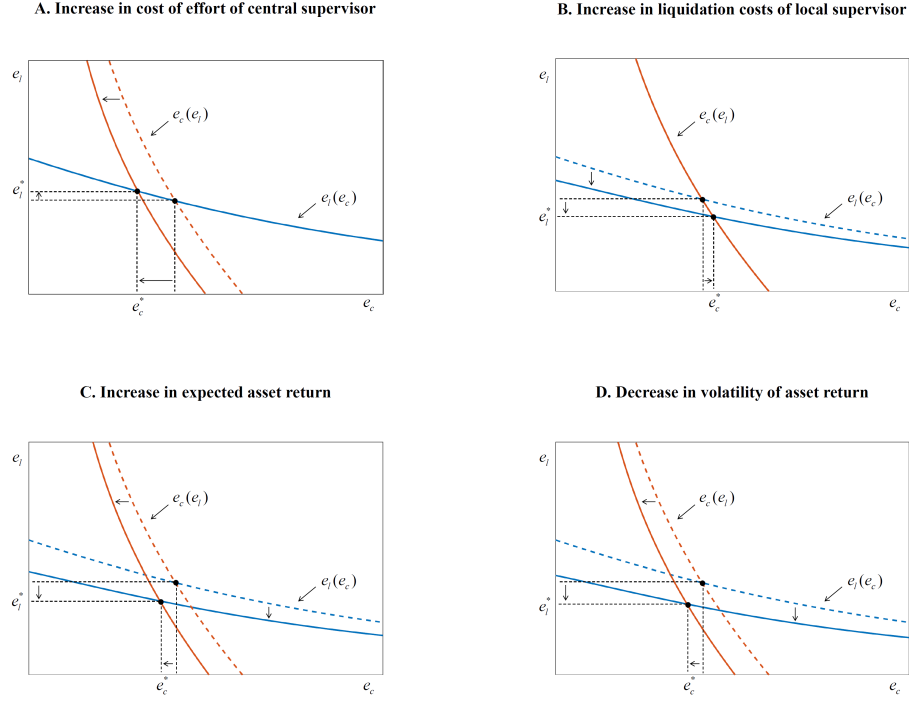


Figure 5. Comparative statics of Nash equilibrium

These comparative static results are in line with the ones obtained in the model with a single supervisor, where supervisory effort is decreasing in the cost of effort γ , the liquidation cost δ , and the expected return $\bar{R}$, and is increasing in the standard deviation σ . But in the game between the two supervisors the first two changes lead to an increase in the equilibrium effort of the supervisor not affected by the parameter changes.

It is interesting to note that the model of hierarchical supervision presented in this section is completely isomorphic to a model in which the central supervisor gets a signal $s_c = R + \varepsilon_c$, where $\varepsilon_c \sim N(0, \sigma^2/e_c)$, the local supervisor gets a signal $s_l = R + \varepsilon_l$, where $\varepsilon_l \sim N(0, \sigma^2/e_l)$, and then it truthfully reports it to the central supervisor, who decides on the liquidation of the bank.²⁰ By the properties of normal distributions we have

$$E(L \mid s_c, s_l) = a\bar{R} + \frac{c(s_l - \bar{R})}{1 + (e_c + e_l)^{-1}}, \quad (26)$$

$$E(R \mid s_c, s_l) = \bar{R} + \frac{s_c - \bar{R}}{1 + (e_c + e_l)^{-1}}, \quad (27)$$

²⁰The noise terms ε_c and ε_l are assumed to be independent of L and R as well as from each other.

where s_{cl} is a weighted average of the two signals with weights proportional to their precision, that is

$$s_{cl} = \frac{e_c}{e_c + e_l} s_c + \frac{e_l}{e_c + e_l} s_l. \quad (28)$$

The random variable s_{cl} is normally distributed, with $E(s_{cl}) = \bar{R}$ and $Var(s_{cl}) = \sigma^2 [1 + (e_c + e_l)^{-1}]$, so it has the same distribution as the random variable s in (15). Moreover, it is also the case that $Cov(s_{cl}, L) = c$ and $Cov(s_{cl}, R) = 1$. Since s_{cl} has the same properties as s , it follows that all the results for the original model of hierarchical supervision, in which the two supervisors put in effort to get a single signal, extend to the alternative model, in which each supervisor puts in effort to get a signal and then the local supervisor truthfully sends its signal to the central supervisor. The original model corresponds to an institutional design in which the supervisors work in teams, while the alternative model corresponds to a design in which the supervisors work independently, but there is no problem of strategic information transmission, that is there are procedures in place that prevent the local supervisor from misrepresenting its signal.

4 Optimal Institutional Design

This section compares in welfare terms three possible institutional arrangements: decentralized, hierarchical, and centralized supervision. Under *decentralized supervision*, only the local supervisor collects information and decides on the liquidation of the bank. Under *hierarchical supervision*, the central and the local supervisor jointly collect information and then the central supervisor decides on the liquidation of the bank. Finally, under *centralized supervision*, only the central supervisor collects information and decides on the liquidation of the bank. The aim is to characterize the conditions under which one of the three institutional arrangements dominates the other two.

The comparison between these institutional arrangements focusses on *two key parameters* of the model: the *cost of effort of the central supervisor*, captured by parameter γ_c , which is higher than parameter γ_l corresponding to the local supervisor, and the *liquidation cost of the local supervisor*, captured by parameter δ_l , which is higher than parameter $\delta_c = 0$

corresponding to the central supervisor. As noted above, the first assumption may be justified by reference to geographical as well as cultural distance between the central supervisor and the local bank, while the second may be justified by reference to the looser connections between the central supervisor and local lobbies and pressure groups.

Social welfare has two components. First, the expected bank returns, given the effort and the liquidation decisions of the supervisors. Second, with negative sign, the costs of supervisory efforts. Supervisory liquidation costs are not taken into account, since they are assumed to be linked to supervisory capture (e.g., possible transfers from banks to supervisors that cancel out in welfare terms).²¹

By the results in Section 2, *social welfare under decentralized supervision* is given by

$$\begin{aligned} w_l &= \int_{-\infty}^{\hat{s}_l} E(L | s) dF(s) + \int_{\hat{s}_l}^{\infty} E(R | s) dF(s) - c_l(\hat{e}_l) \\ &= \bar{R} - [(1-a)\bar{R} + \delta_l] \left[\Phi(\hat{x}_l) + \frac{\phi(\hat{x}_l)}{\hat{x}_l} \right] - c_l(\hat{e}_l), \end{aligned} \quad (29)$$

where $\hat{e}_l$ is the effort chosen by the local supervisor and

$$\hat{x}_l = -\frac{[(1-a)\bar{R} + \delta_l]}{(1-c)\sigma} (1 + \hat{e}_l^{-1})^{1/2}.$$

Similarly, *social welfare under centralized supervision* is given by

$$\begin{aligned} w_c &= \int_{-\infty}^{\hat{s}_c} E(L | s) dF(s) + \int_{\hat{s}_c}^{\infty} E(R | s) dF(s) - c_c(\hat{e}_c) \\ &= \bar{R} - (1-a)\bar{R} \left[\Phi(\hat{x}_c) + \frac{\phi(\hat{x}_c)}{\hat{x}_c} \right] - c_c(\hat{e}_c), \end{aligned} \quad (30)$$

where $\hat{e}_c$ is the effort chosen by the central supervisor and

$$\hat{x}_c = -\frac{(1-a)\bar{R}}{(1-c)\sigma} (1 + \hat{e}_c^{-1})^{1/2}.$$

²¹See Lucca et al. (2014) for evidence on the revolving door between regulatory agencies and the banking industry. Their results appear to be inconsistent with a “quid pro quo view” according to which future employment opportunities in the private sector affects the strictness of actions of regulators. Instead, they find evidence in favor of a regulatory “schooling view” according to schooling in complex rules enhances regulators’ future earnings, should they transition to the private sector.

Finally, by the results in Section 3, *social welfare under hierarchical supervision* is given by

$$\begin{aligned} w_h &= \int_{-\infty}^{s^*} E(L | s) dF(s) + \int_{s^*}^{\infty} E(R | s) dF(s) - c_c(\bar{e}_c) - c_l(\bar{e}_l) \\ &= \bar{R} - (1-a)\bar{R} \left[\Phi(x^*) + \frac{\phi(x^*)}{x^*} \right] - c_c(\bar{e}_c) - c_l(\bar{e}_l), \end{aligned} \quad (31)$$

where (e_c^*, e_l^*) is the Nash equilibrium of the game played by the supervisors and

$$x^* = -\frac{(1-a)\bar{R}}{(1-c)\sigma} [1 + (e_c^* + e_l^*)^{-1}]^{1/2}.$$

I can now compare in terms of welfare the three alternative institutional arrangements by computing w_l , w_h , and w_c for different values of the cost of effort of the central supervisor γ_c and the liquidation cost of the local supervisor δ_l .

The forces driving the comparison are easy to explain. The higher the cost of effort of the central supervisor γ_c the lower the likelihood that centralized supervision will be optimal, because as shown in Figure 3.A the central supervisor will have an incentive to exert too little effort. Similarly, the higher the liquidation cost the local supervisor δ_l the lower the likelihood that decentralized supervision will be optimal, because as shown in Figure 3.B the local supervisor will have an incentive to exert too little effort. Relative to the case of a single supervisor, hierarchical supervision entails incurring twice the fixed cost γ_0 , so by the convexity of the cost function (8) it will only be optimal when both supervisors have an incentive to exert significant effort. This requires that the bias of the local supervisor is high but not too high and the costs of getting local information from the center are low but not too low.

Figure 6 shows the results.²² Decentralized supervision dominates in Region *D*, where γ_c is large and δ_l is small, that is when the cost advantage of the local supervisor is sufficiently large to compensate the bias in its liquidation decision. Hierarchical supervision dominates in Region *H*, where the cost advantage of the local supervisor is not too large or the liquidation

²²In this figure the origin corresponds to $\gamma_c = \gamma_l = 0.001$ and $\delta_l = \delta_c = 0$.

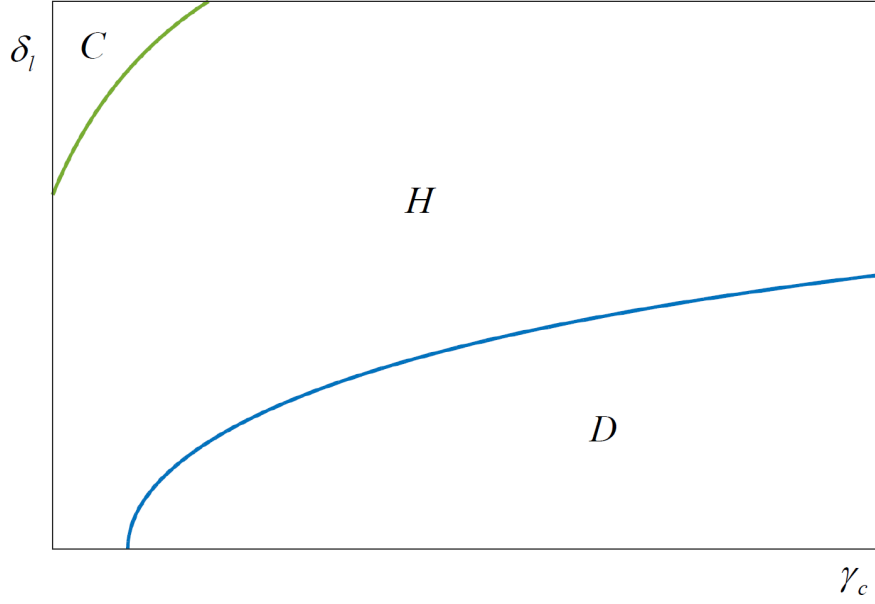


Figure 6. Optimal institutional design

cost of the local supervisor is not too small. Finally, centralized supervision dominates in Region C , where γ_c is small and δ_l is large, that is when the cost disadvantage of the central supervisor is sufficiently small to compensate the bias of the local supervisor.

Next, I analyze the effect on the regions in Figure 6 of changes in the expected return $\bar{R}$ and in the standard deviation σ of the bank's investment return. Figure 7 illustrates the results. Panel A shows that an increase in $\bar{R}$ expands Regions D and C where decentralized and centralized supervision are optimal at the expense of Region H where hierarchical supervision is optimal. Similarly, Panel B shows that a decrease in σ expands Regions D and C where decentralized and centralized supervision are optimal at the expense of the Region H where hierarchical supervision is optimal.²³ The intuition is that when the bank is more

²³Interestingly, both panels illustrate that for relatively small values of γ_c and δ_l the optimal institutional design may switch from Region D to Region C without passing through Region H . Notice also that the line that separates the two regions goes through the origin, since when $\gamma_c = \gamma_l$ and $\delta_l = \delta_c$ both supervisors have the same payoff function.

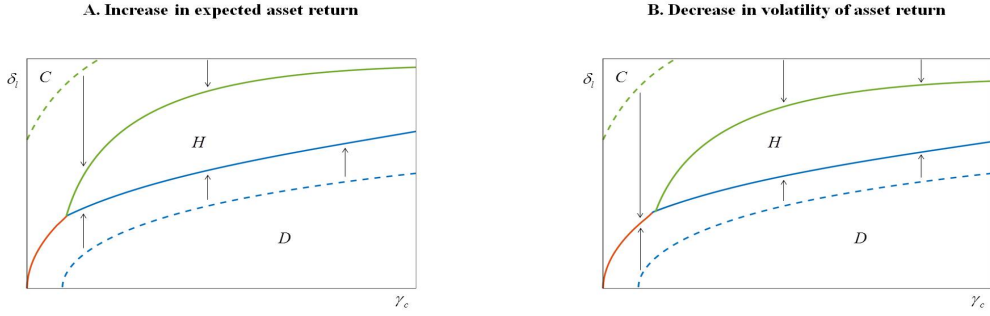


Figure 7. Comparative statics of optimal institutional design

profitable (higher $\bar{R}$) or safer (lower σ) the supervisors will exert less effort (see Figures 5.C and 5.D), Thus, given the convexity of the cost function (8), the relative advantage of having two supervisors will decrease.

Summing up, I have shown that hierarchical supervision dominates decentralized supervision when the possible capture of the local supervisor is a significant concern and the costs of getting local knowledge are not too large. But when these two forces go beyond certain threshold it is better to eliminate the local supervisor and concentrate all responsibilities in the central supervisor. Moreover, hierarchical supervision is more likely to dominate when the risk of bank failure (due to low profitability or high risk-taking) is high. These latter results illustrate the way in which banking supervision should be adjusted in response to changes in the environment, and hence may interpreted as macroprudential policies.

5 Extensions

5.1 Linking liquidation costs to supervisory responsibilities

This extension examines what happens when part of the liquidation cost incurred by the supervisors is linked to whether they are responsible for the liquidation decision.

To formalize this idea, suppose that $\delta_c = 0$ and $\delta_l > 0$ are the liquidation costs for the central and the local supervisor, respectively, in the absence of supervisory responsibilities, and that such responsibilities add Δ to these costs. Thus, the liquidation cost of the local supervisor will be $\delta_l + \Delta$ under decentralized supervision and δ_l under hierarchical supervi-

sion, whereas the liquidation cost of the central supervisor will be Δ under both centralized and hierarchical supervision.

By our previous results, under decentralized supervision the effort decision of the local supervisor is given by

$$\hat{e}_l = \arg \max_{e_l} \left\{ \bar{R} - [(1-a)\bar{R} + \delta_l + \Delta] \left[\Phi(x_l) + \frac{\phi(x_l)}{x_l} \right] - c_l(e_l) \right\},$$

where

$$x_l = x_l(e_l) = -\frac{(1-a)\bar{R} + \delta_l + \Delta}{(1-c)\sigma} (1 + e_l^{-1})^{1/2}.$$

Social welfare under decentralized supervision is

$$w_l = \bar{R} - [(1-a)\bar{R} + \delta_l + \Delta] \left[\Phi(\hat{x}_l) + \frac{\phi(\hat{x}_l)}{\hat{x}_l} \right] + (\delta_l + \Delta)\Phi(\hat{x}_l) - c_l(\hat{e}_l),$$

where $\hat{x}_l = x_l(\hat{e}_l)$. Similarly, under centralized supervision the effort decision of the central supervisor is given by

$$\hat{e}_c = \arg \max_{e_c} \left\{ \bar{R} - [(1-a)\bar{R} + \Delta] \left[\Phi(x_c) + \frac{\phi(x_c)}{x_c} \right] - c_c(e_c) \right\},$$

where

$$x_c = x_c(e_c) = -\frac{(1-a)\bar{R} + \Delta}{(1-c)\sigma} (1 + e_c^{-1})^{1/2}.$$

Social welfare under centralized supervision is

$$w_c = \bar{R} - [(1-a)\bar{R} + \Delta] \left[\Phi(\hat{x}_c) + \frac{\phi(\hat{x}_c)}{\hat{x}_c} \right] + \Delta\Phi(\hat{x}_c) - c_c(\hat{e}_c),$$

where $\hat{x}_c = x_c(\hat{e}_c)$.

Finally, since the liquidation cost of the central supervisor is now $\Delta > 0$, the payoff functions of the supervisors become

$$\begin{aligned} v_c(e_c, e_l) &= \bar{R} - [(1-a)\bar{R} + \Delta] \left[\Phi(x) + \frac{\phi(x)}{x} \right] - c_c(e_c), \\ v_l(e_c, e_l) &= \bar{R} - [(1-a)\bar{R} + \Delta] \left[\Phi(x) + \frac{\phi(x)}{x} \right] - (\delta_l - \Delta)\Phi(x) - c_l(e_l), \end{aligned}$$

where

$$x = x(e_c + e_l) = -\frac{(1-a)\bar{R} + \Delta}{(1-c)\sigma} [1 + (e_c + e_l)^{-1}]^{1/2}.$$

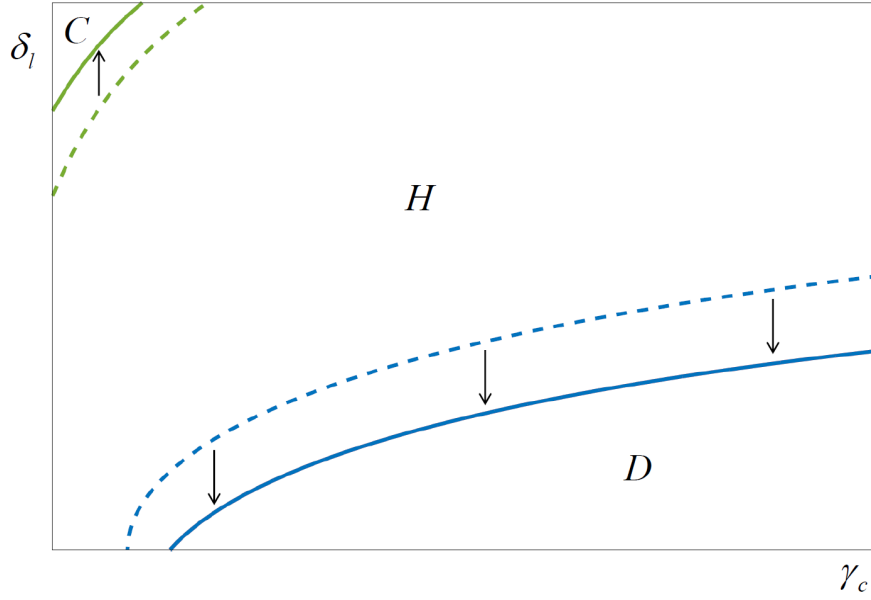


Figure 8. Linking liquidation costs to supervisory responsibilities

Let (e_c^*, e_l^*) denote the Nash equilibrium of this game. Social welfare under hierarchical supervision is then

$$w_h = \bar{R} - [(1-a)\bar{R} + \Delta] \left[\Phi(x^*) + \frac{\phi(x^*)}{x^*} \right] + \Delta \Phi(x^*) - c_c(e_c^*) - c_l(e_l^*),$$

where $x^* = x(e_c^* + e_l^*)$.

Figure 8 shows the results.²⁴ The assumption that part of the liquidation cost incurred by the supervisors is linked to whether they are responsible for the liquidation decision expands Region H where hierarchical supervision is optimal at the expense of Regions D and C where decentralized and centralized supervision are optimal.

The intuition is as follows. Consider the effect of introducing the additional cost Δ on a point in Figure 8 on the original boundary between Regions D and H . By our comparative statics results in Section 3, this will lead to a shift to the left of the reaction function of the central supervisor (due to the increase in its liquidation cost by Δ) and a shift upwards of the reaction function of the local supervisor (due to the equivalent reduction in its liquidation cost by Δ), leading to a reduction in the equilibrium effort e_c^* of the central supervisor and

²⁴In this figure $\Delta = 0.005$.

an increase in the (cheaper) equilibrium effort e_l^* of the local supervisor, which explains that the original boundary point is now inside Region H .

Similarly, consider the effect of introducing the additional cost Δ on a point in Figure 8 on the original boundary between Regions H and C . By our comparative statics results in Section 3, this will lead to a shift to the left of the reaction function of the central supervisor (due to the increase in its liquidation cost by Δ), leading to a reduction in the equilibrium effort e_c^* of the central supervisor and an increase in the (cheaper) equilibrium effort e_l^* of the local supervisor. This increase, which does not obtain under centralized supervision, explains that the original boundary point is now inside Region H .

5.2 Limiting the size of the central supervisor

This extension considers whether it would be desirable from a welfare perspective to limit (in some statutory manner) the size of the central supervisor. In terms of the model, a size limit implies an upper bound to the effort of the central supervisor. How could this be welfare improving?

To answer this question it is convenient to start considering the case where the central supervisor were a Stackelberg leader. By standard results for games with strategic substitutes, in a Stackelberg equilibrium the central supervisor would reduce its effort e_c and the local supervisor would increase its effort e_l , relative to the Nash equilibrium (e_c^*, e_l^*) . Figure 9 illustrates the result. By the definition of Nash equilibrium, the indifference curve of the central supervisor at the point (e_c^*, e_l^*) is tangent to the horizontal line $e_l = e_l^*$. But since the reaction function of the local supervisor is downward sloping, reducing the effort of the central supervisor increases its payoff.

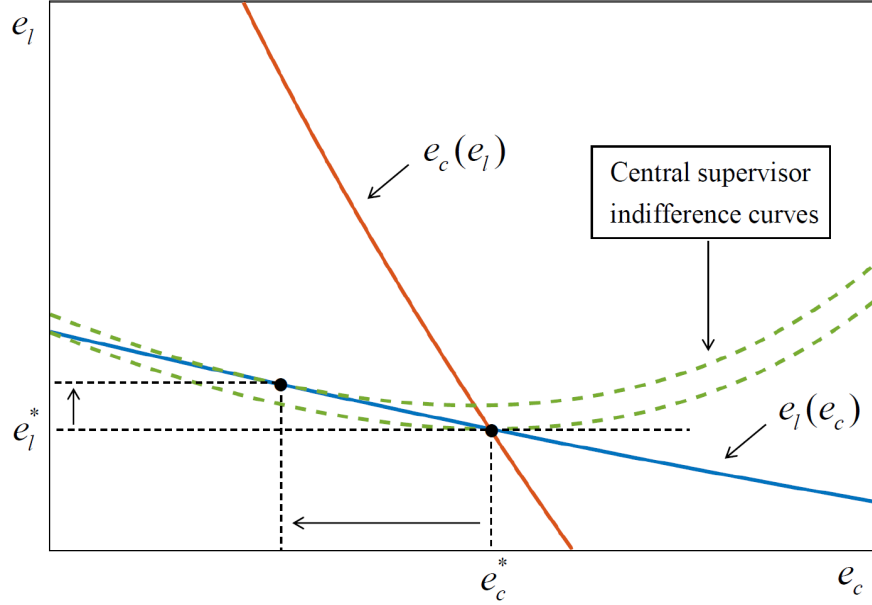


Figure 9. The central supervisor as a Stackelberg leader

The question is whether this argument also applies when we replace the indifference curve of the central supervisor by the social indifference curves. To check that this is indeed the case, notice that using the definition (23) of the payoff function of the central supervisor $v_c(e_c, e_l)$ and the definition (31) of the social welfare function $w(e_c, e_l)$, the social indifference curve going through the Nash equilibrium point (e_c^*, e_l^*) may be written as

$$w(e_c, e_l) = v_c(e_c, e_l) - c_l(e_l) = w_h^*,$$

where $w_h^* = w(e_c^*, e_l^*)$. This implies

$$\left. \frac{\partial w(e_c, e_l)}{\partial e_c} \right|_{(e_c^*, e_l^*)} = \left. \frac{\partial v(e_c, e_l)}{\partial e_c} \right|_{(e_c^*, e_l^*)} = 0,$$

where the second equality follows from the definition of Nash equilibrium. Thus, the social indifference curve at the point (e_c^*, e_l^*) is also tangent to the horizontal line $e_l = e_l^*$, so moving up the reaction function of the local supervisor increases social welfare. Figure 10 illustrates the result.

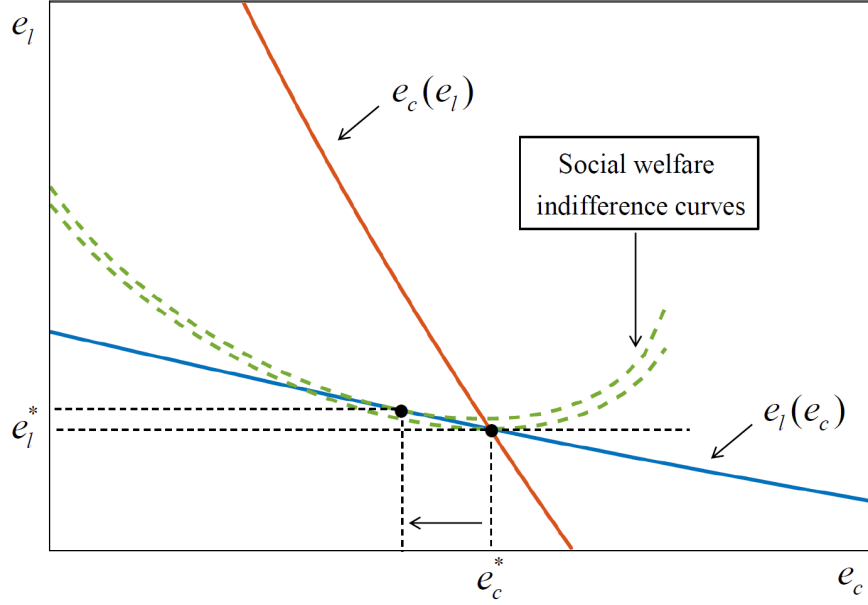


Figure 10. Limiting the size of the central supervisor

Hence, the commitment to reduce the effort of the central supervisor in the hierarchical supervision setup is welfare improving. The intuition is that a reduction in the effort of the central supervisor forces the local supervisor to increase its (cheaper) effort. However, the model assumes that the efforts of the supervisors are not verifiable, so there is an issue about how this could be implemented. One obvious way to move in this direction would be to limit the size of the central supervisor.

6 Concluding Remarks

This paper constructs a model of hierarchical supervision in which a central and a local supervisor independently choose supervisory efforts which determine the precision of a signal of the solvency of a local bank. The central supervisor uses this information to decide on the bank's early liquidation. Importantly, the local supervisor is assumed to be characterized by a lower cost of effort (due to proximity) and a higher cost of liquidating the bank (due to capture). The model of hierarchical supervision is compared in terms of welfare with two alternative arrangements, namely decentralized supervision, where the local supervisor

is fully in charge, and centralized supervision, where the central supervisor is fully in charge. The results show that moving supervision from local to hierarchical to central is more likely to be optimal when the cost of getting local knowledge is sufficiently low and/or the possible capture of the local supervisor is a significant concern.

To the extent that supervisory capture may be more relevant for large banks, the results provide support for the design of Single Supervisory Mechanism of the European Central Bank (ECB), in which significant banks are supervised by the ECB in collaboration with the National Competent Authorities (NCAs), and less significant banks are supervised by the NCAs.

I would like to conclude with a few remarks. First, the model assumes that the liquidation cost is fully driven by capture. However, it may be the case that (at least part of) the liquidation cost is also a social cost, that is internalized by the local supervisor but not by the central supervisor. In this case, having a more biased (local) supervisor could be better than having a less biased (central) supervisor, so the region in the parameter space where decentralized supervision dominates becomes larger.

Second, the model assumes that bank is completely passive. As in Carletti et al. (2016), it would be interesting to endogenize bank's choice of risk under different supervisory arrangements. Moreover, it would also be interesting to introduce bank capital regulation, and to analyze the possible trade-offs between regulation and supervision. In particular, one may conjecture that the tougher the regulation, and hence the lower the bank's risk-taking, the less valuable supervision will be, which as noted above reduces the region of the parameter space where hierarchical supervision dominates.

Third, the model is static, but one could easily think of some dynamic implications. For example, it may be the case that in good times supervisors gradually reduce their capabilities, so they may not be able to increase them quickly when bad times arrive, in which case we could end up having (involuntary) supervisory forbearance.

Finally, the model focuses on supervision of a domestic bank. It would be interesting to explore the case of an international bank that operates across several local jurisdictions. In fact, one could argue that the rationale for the creation of the Single Supervisory Mechanism

of the ECB was the fact that in a multi-country system, like the euro area, the supervisory decisions of a biased national authority may entail negative externalities on other countries. But these potential externalities also appear in the simple setup of this paper. In particular, since deposits are assumed to be insured, a bank failure may impact on the solvency of the country in which the bank is located, and through a fiscal channel affect other countries in the area.

Appendix

Proof of Proposition 1 By (3) one can write $s = \bar{R} + \sigma (1 + e^{-1})^{1/2} x$, where $x \sim N(0, 1)$. Substituting this expression into (4) and (5) gives

$$\begin{aligned} E(L \mid s) &= a\bar{R} + c\sigma (1 + e^{-1})^{-1/2} x, \\ E(R \mid s) &= \bar{R} + \sigma (1 + e^{-1})^{-1/2} x. \end{aligned}$$

Hence, one can write

$$v(e) = \int_{-\infty}^{\hat{x}} \left[a\bar{R} + c\sigma (1 + e^{-1})^{-1/2} x - \delta \right] \phi(x) dx + \int_{\hat{x}}^{\infty} \left[\bar{R} + \sigma (1 + e^{-1})^{-1/2} x \right] \phi(x) dx - c(e),$$

where $\hat{x}$ is implicitly defined by the equation

$$\bar{R} + \sigma (1 + e^{-1})^{1/2} \hat{x} = \hat{s}.$$

Solving for $\hat{x}$ and using the definition (10) of $\hat{s}$ gives (13). Then, using the fact that $x\phi(x) = -\phi'(x)$ and the definition (13) of $\hat{x}$, one concludes

$$\begin{aligned} v(e) &= (a\bar{R} - \delta) \Phi(\hat{x}) - c\sigma (1 + e^{-1})^{-1/2} \phi(\hat{x}) + \bar{R} (1 - \Phi(\hat{x})) + \sigma (1 + e^{-1})^{-1/2} \phi(\hat{x}) - c(e) \\ &= \bar{R} - [(1 - a)\bar{R} + \delta] \Phi(\hat{x}) + (1 - c)\sigma (1 + e^{-1})^{-1/2} \phi(\hat{x}) - c(e) \\ &= \bar{R} - [(1 - a)\bar{R} + \delta] \left[\Phi(\hat{x}) + \frac{\phi(\hat{x})}{\hat{x}} \right] - c(e). \quad \square \end{aligned}$$

Proof of Proposition 2 Given that

$$\frac{d}{d\hat{x}} \left[\Phi(\hat{x}) + \frac{\phi(\hat{x})}{\hat{x}} \right] = -\frac{\phi(\hat{x})}{\hat{x}^2}$$

and

$$\frac{\partial \hat{x}}{\partial e} = -\frac{\hat{x}}{2e(1 + e)},$$

the first-order condition that characterizes an interior solution $\hat{e} > 0$ is

$$v'(e) = -\frac{[(1 - a)\bar{R} + \delta]\phi(\hat{x})}{2e(1 + e)\hat{x}} - \gamma e = 0,$$

and the second-order condition is $v''(e) < 0$. Hence, by the implicit function theorem one has

$$\frac{\partial \hat{e}}{\partial \gamma} = -\frac{1}{v''(e)} \frac{\partial v'(e)}{\partial \gamma} = \frac{e}{v''(e)} < 0.$$

Next, $v''(e) < 0$ and $\hat{x} < 0$ imply

$$\begin{aligned}
\frac{\partial \hat{e}}{\partial \delta} &= -\frac{1}{v''(e)} \frac{\partial v'(e)}{\partial \delta} \\
&= \frac{1}{v''(e)} \frac{1}{2e(1+e)} \left[\frac{\phi(\hat{x})}{\hat{x}} + [(1-a)\bar{R} + \delta] \frac{d}{d\hat{x}} \left(\frac{\phi(\hat{x})}{\hat{x}} \right) \frac{\partial \hat{x}}{\partial \delta} \right] \\
&= \frac{1}{v''(e)} \frac{1}{2e(1+e)} \left[\frac{\phi(\hat{x})}{\hat{x}} - \frac{(1+\hat{x}^2)\phi(\hat{x})}{\hat{x}} \right] \\
&= -\frac{1}{v''(e)} \frac{1}{2e(1+e)} \hat{x} \phi(\hat{x}) < 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \hat{e}}{\partial \bar{R}} &= -\frac{1}{v''(e)} \frac{\partial v'(e)}{\partial \bar{R}} \\
&= \frac{1}{v''(e)} \frac{1}{2e(1+e)} \left[(1-a) \frac{\phi(\hat{x})}{\hat{x}} + [(1-a)\bar{R} + \delta] \frac{d}{d\hat{x}} \left(\frac{\phi(\hat{x})}{\hat{x}} \right) \frac{\partial \hat{x}}{\partial \bar{R}} \right] \\
&= \frac{1}{v''(e)} \frac{1-a}{2e(1+e)} \left[\frac{\phi(\hat{x})}{\hat{x}} - \frac{(1+\hat{x}^2)\phi(\hat{x})}{\hat{x}} \right] \\
&= -\frac{1}{v''(e)} \frac{1-a}{2e(1+e)} \hat{x} \phi(\hat{x}) < 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \hat{e}}{\partial \sigma} &= -\frac{1}{v''(e)} \frac{\partial v'(e)}{\partial \sigma} \\
&= \frac{1}{v''(e)} \frac{(1-a)\bar{R} + \delta}{2e(1+e)} \frac{d}{d\hat{x}} \left(\frac{\phi(\hat{x})}{\hat{x}} \right) \frac{\partial \hat{x}}{\partial \sigma} \\
&= \frac{1}{v''(e)} \frac{(1-a)\bar{R} + \delta}{2e(1+e)} \frac{(1+\hat{x}^2)\phi(\hat{x})}{\sigma \hat{x}} > 0. \quad \square
\end{aligned}$$

Proof of Proposition 3 By (15) one can write $s = \bar{R} + \sigma[1 + (e_c + e_l)^{-1}]^{1/2}x$, where $x \sim N(0, 1)$. Substituting this expression into (16) and (17) gives

$$E(L \mid s) = a\bar{R} + c\sigma[1 + (e_c + e_l)^{-1}]^{-1/2}x,$$

$$E(R \mid s) = \bar{R} + \sigma[1 + (e_c + e_l)^{-1}]^{-1/2}x.$$

Hence, one can write

$$\begin{aligned}
v_c(e_c, e_l) &= \int_{-\infty}^{x^*} [a\bar{R} + c\sigma[1 + (e_c + e_l)^{-1}]^{-1/2}x] \phi(x)dx \\
&\quad + \int_{x^*}^{\infty} [\bar{R} + \sigma[1 + (e_c + e_l)^{-1}]^{-1/2}x] \phi(x)dx - c_c(e_c), \\
v_l(e_c, e_l) &= \int_{-\infty}^{x^*} [a\bar{R} + c\sigma[1 + (e_c + e_l)^{-1}]^{-1/2}x - \delta_l] \phi(x)dx \\
&\quad + \int_{x^*}^{\infty} [\bar{R} + \sigma[1 + (e_c + e_l)^{-1}]^{-1/2}x] \phi(x)dx - c_l(e_l),
\end{aligned}$$

where x^* is implicitly defined by the equation

$$\bar{R} + \sigma[1 + (e_c + e_l)^{-1}]^{1/2}x^* = s^*.$$

Solving for x^* gives (25). Then, following the same steps as in the proof of Proposition 1 one gets the expressions for $v_c(e_c, e_l)$ and $v_l(e_c, e_l)$. $\square$

References

- Agarwal, S., D. Lucca, A. Seru, and F. Trebbi (2014), “Inconsistent Regulators: Evidence from Banking,” *Quarterly Journal of Economics*, 129, 889-938.
- Aghion, P., and J. Tirole (1997), “Formal and Real Authority in Organizations,” *Journal of Political Economy*, 105, 1-29.
- Alonso, R., W. Dessein, and N. Matouschek (2008), “When Does Coordination Require Centralization?,” *American Economic Review*, 98, 145-179.
- Argenziano, R., S. Severinov, and F. Squintani (2016), “Strategic Information Acquisition and Transmission,” *American Economic Journal - Microeconomics*, 8, 119-155.
- Bian, B., R. Haselman, T. Kick, and V. Vig (2017), “The Political Economy of Bank Bailouts: Evidence,” Working paper.
- Boyer, P., and J. Ponce (2012), “Regulatory Capture and Banking Supervision Reform,” *Journal of Financial Stability*, 8, 206-217.
- Calzolari, G., J.-E. Colliard, and G. Loranth (2016), “Multinational Banks and Supranational Supervision,” <https://ssrn.com/abstract=2912802>.
- Carletti, E., G. Dell’Ariccia, and R. Marquez (2016), “Supervisory Incentives in a Banking Union,” IMF Working Paper WP/16/186.
- Colliard, J.-E. (2015), “Optimal Supervisory Architecture and Financial Integration in a Banking Union,” ECB Working Paper No. 1786.
- Crawford, V., and J. Sobel (1982), “Strategic Information Transmission,” *Econometrica*, 50, 1431-1451.
- Dessein, W. (2002), “Authority and Communication in Organizations,” *Review of Economic Studies*, 69, 811-838.

- Dessein, W., L. Garicano, and R. Gertner (2010), “Organizing for Synergies,” *American Economic Journal - Microeconomics*, 2, 77-114.
- Eisenbach, T., D. Lucca, and R. Townsend (2016), “The Economics of Bank Supervision,” Federal Reserve Bank of New York Staff Report No. 769.
- Gopalan, Y., A. Kalda, and A. Manela (2017), “Hub-and-Spoke Regulation and Bank Leverage,” Working paper.
- Holthausen, C., and T. Rønde (2004), “Cooperation in International Banking Supervision,” ECB Working Paper No. 316.
- Jensen, M., and W. Meckling (1992), “Specific and General Knowledge and Organizational Structure,” in L. Werin and H. Wijkander (eds.), *Contract Economics*, Blackwell: Oxford, 251-274.
- Lambert, T. (2017), “Lobbying on Regulatory Enforcement Actions: Evidence from U.S. Commercial and Savings Banks,” *Management Science*, forthcoming.
- Lucca, D., A. Seru, and F. Trebbi (2014), “The Revolving Door and Worker Flows in Banking Regulation,” *Journal of Monetary Economics*, 65, 17-32.
- Radner, R., and J. Stiglitz (1984), “A Nonconcavity in the Value of Information,” in M. Boyer and R. Kihlstrom (eds.), *Bayesian Models in Economic Theory*, Elsevier Science: Amsterdam, 33-52.
- Sobel, J. (2013), “Giving and Receiving Advice,” in D. Acemoglu, M. Arellano, and E. Dekel (eds.), *Advances in Economics and Econometrics*, Tenth World Congress, Volume I: Economic Theory, Cambridge University Press: Cambridge, 305-341.
- Torres, X. (2015), “El Mecanismo Único de Supervisión y el papel de las autoridades nacionales,” Banco de España, *Revista de Estabilidad Financiera*, 29, 11-28.