Hierarchical Bank Supervision

Rafael Repullo
CEMFI and CEPR

February 2017

Preliminary and incomplete

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 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. Moreover, whenever hierarchical supervision is optimal, limiting the size of the central supervisor is always welfare increasing.

JEL Classification: G21, G23, D02.

Keywords: Bank supervision, bank liquidation, supervisory capture, strategic information acquisition, optimal institutional design.

Presidential Address delivered at the Symposium of the Spanish Economic Association in Bilbao in December 2016. I would like to thank Luis Garicano, David Martinez-Miera, Enrique Sentana, Joel Sobel, and Javier Suarez for helpful comments, and Wenbo Meng and Álvaro Remesal for their excellent research assistance. Financial support from the Spanish Ministry of Economy and Competitiveness, Grant No. ECO2014-59262-P, is gratefully acknowledged. E-mail: repullo@cemfi.es.

1 Introduction

Bank supervision, unlike bank regulation, has not been until recently the subject of much academic interest. As noted 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 nonpecuniary 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."

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. 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 completely isomorphic to a model in which the central supervisor gets a signal of the bank's solvency, the local supervisor gets another signal, and then it truthfully reports it to the central supervisor, who 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.

The paper 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 127 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–such as the state-chartered banks in the US; see Agarwal et al. (2014).[1]

A somewhat surprising result that follows from the model is that reducing the effort of the central supervisor in the hierarchical setup is always 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 this change forces the local supervisor to increase its (cheaper) effort. The way to implement this result is to limit the capacity of the central supervisor.

TBC

Literature review TBC

Structure of the paper Section 2 presents the 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. Section 3 presents the model of hierarchical supervision in which a central and a

[1] They show that federal supervisors are systematically tougher that 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.

local supervisor 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 shows that limiting the size of the central supervisor in the hierarchical supervision setup is always welfare improving. Section 6 presents some concluding remarks. The proofs of the analytical results are in the Appendix.

2 Model Setup

Consider an economy with three dates ($t = 0, 1, 2$) and two agents: a bank and a bank 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

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

where $\overline{R} > 1$, $0 < a < 1$, $c > 0$, and $c^2 < b < 1$. Thus, the expected final return $E(R) = \overline{R}$ is greater than the unit face value of the deposits, and it is also greater than the expected liquidation return $E(L) = a\overline{R}$. Moreover, the final return has a higher variance than that of the liquidation return, and both returns are positively correlated.[2]

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

$$s = R + \varepsilon \tag{2}$$

on the final return of the bank's investment, where the noise term ε is independent of L and R, and has a distribution $N(0, \sigma^2/e)$.[3] Parameter e is proportional to the precision (inverse

[2] Assumption $c^2 < b$ ensures that the covariance matrix is positive-definite.

[3] Notice that we can write $L = (a - c)\overline{R} + cR + u$, where $Cov(u, R) = Cov(u, s) = 0$. Thus, signal s contains information about L only inasmuch as it contains information about R.

of the variance) of the noise term. From here it follows that

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

By the properties of normal distributions we have

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

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

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

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

where

$$s^* = \overline{R} - \frac{1-a}{1-c}(1+e^{-1})\overline{R} \tag{6}$$

is the *efficient liquidation threshold* (for a given value of e). Thus, a higher precision e of the supervisory signal s increases the efficient liquidation threshold s^*.[4]

Substituting (6) into (4) and (5) we get

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

I will assume that parameter values satisfy

$$\frac{a-c}{1-c}\overline{R} \leq 1. \tag{7}$$

This means that the efficient liquidation threshold s^* is such that the corresponding expected final return is less than or equal to the face value of the deposits, so the bank is effectively bankrupt.[5]

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

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

[4] Notice that for $e = 0$ we have $s^* = -\infty$, so the bank would never be liquidated.

[5] Interestingly, $E(L \mid s^*) = E(R \mid s^*)$ does not depend on supervisory effort e.

where $\gamma_0 > 0$ and $\gamma > 0$.

The supervisor decides whether to liquidate the bank based on the observation of the signal s at $t = 1$. I assume that the supervisor liquidates the bank at $t = 1$ if

$$E(L \mid s) - \delta > E(R \mid s), \tag{9}$$

where $\delta > 0$ is a nonpecuniary 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 $E(L \mid s) - E(R \mid s)$, are greater than the supervisor's private cost δ of liquidation. Substituting (4) and (5) into (9) implies that the bank will be liquidated by the supervisor when $s < \widehat{s}$, where

$$\widehat{s} = s^* - \frac{1 + e^{-1}}{1 - c}\delta \tag{10}$$

is the *supervisor's liquidation threshold* (for a given value of e). Hence, the higher the cost δ of liquidating the bank the more lenient the supervisor will be.

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

[FIGURE 1]

The supervisor's effort decision at $t = 0$ is obtained by maximizing its expected payoff at $t = 1$ net of the cost of effort $c(e)$, which gives

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

where

$$\begin{aligned} v(e) &= E\left[\max\left\{E(L \mid s) - \delta, E(R \mid s)\right\}\right] - c(e) \\ &= \int_{-\infty}^{\widehat{s}} \left[E(L \mid s) - \delta\right] dF(s) + \int_{\widehat{s}}^{\infty} E(R \mid s) dF(s) - c(e), \end{aligned} \tag{11}$$

and $F(s)$ denotes the cdf of the signal s. The following result provides a closed form expression for $v(e)$.

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

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

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

$$\widehat{x} = \frac{\widehat{s} - E(s)}{Var(s)} = -\frac{\left[(1-a)\overline{R} + \delta\right]}{(1-c)\sigma} \left(1 + e^{-1}\right)^{1/2}. \tag{12}$$

Figure 2 plots the payoff function of the supervisor $v(e)$ for the parameter that will be used in the numerical analysis below.[6] For $e = 0$ we have $\widehat{x} = \widehat{s} = -\infty$, so the bank is never liquidated and $v(0) = \overline{R}$. For $e = \infty$ we have $v(e) = -\infty$, since the cost of supervisory effort goes to infinity. The function $v(e)$ is initially convex, and then becomes concave, so we may have corner solutions with $\widehat{e} = 0$ or interior solutions with $\widehat{e} > 0$.

[FIGURE 2]

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

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

$$\frac{\partial \widehat{e}}{\partial \gamma} < 0, \ \frac{\partial \widehat{e}}{\partial \delta} < 0, \ \frac{\partial \widehat{e}}{\partial \overline{R}} < 0, \ \textit{and} \ \frac{\partial \widehat{e}}{\partial \sigma} > 0.$$

Figure 3 illustrates these results. Panel A shows that increases in γ reduce supervisory effort $\widehat{e}$, and Panel B shows that increases in δ also reduce supervisory effort $\widehat{e}$, which jumps to zero for sufficiently high values of the liquidation cost. Panel C shows that increases in

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

$\overline{R}$ reduce supervisory effort $\widehat{e}$, while Panel D shows that increases in σ increase supervisory effort $\widehat{e}$ for sufficiently high values of the standard deviation σ of the bank's investment return.

[FIGURE 3]

The comparative statics results are not surprising. Supervision will be more intense when banks are easier to supervise (lower γ), or less profitable (lower $\overline{R}$), 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 bank 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 early liquidation. I have characterized the supervisor's effort decision, and derive some comparative statics results on its determinants.

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. 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 reputational 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 local lobbies and pressure groups. The trade-off between the higher costs of supervision against the lower 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 observe 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 nonverifiable signal

$$s = R + \varepsilon \tag{13}$$

on the final return of the bank's investment. As before, 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 lower the variance of ε.[7] From here it follows that

$$\begin{bmatrix} L \\ R \\ s \end{bmatrix} \sim N\left(\begin{bmatrix} a\overline{R} \\ \overline{R} \\ \overline{R} \end{bmatrix}, \sigma^2 \begin{bmatrix} b & c & c \\ c & 1 & 1 \\ c & 1 & 1 + (e_c + e_l)^{-1} \end{bmatrix}\right). \tag{14}$$

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

$$E(L \mid s) = a\overline{R} + \frac{c(s - \overline{R})}{1 + (e_c + e_l)^{-1}}, \tag{15}$$

$$E(R \mid s) = \overline{R} + \frac{s - \overline{R}}{1 + (e_c + e_l)^{-1}}, \tag{16}$$

[7] As 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.

so we conclude that

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

where

$$s^* = \overline{R} - \frac{1-a}{1-c}[1 + (e_c + e_l)^{-1}]\overline{R} \tag{17}$$

is the efficient liquidation threshold (for given values of e_c and e_l). As before, I assume that parameter values are such that (7) holds. Thus, the efficient liquidation threshold s^* is such that $E(L \mid s^*) = E(R \mid s^*) \leq 1$.

At $t = 0$ the central and the local supervisors simultaneously 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 given by

$$v_c(e_c, e_l) = \int_{-\infty}^{s^*} E(L \mid s) dF(s) + \int_{s^*}^{\infty} E(R \mid s) dF(s) - c_c(e_c), \tag{18}$$

where $F(s)$ denotes the cdf of the signal s. Similarly, The payoff function of the local supervisor is given by

$$v_l(e_c, e_l) = \int_{-\infty}^{s^*} [E(L \mid s) - \delta_l] dF(s) + \int_{s^*}^{\infty} E(R \mid s) dF(s) - c_l(e_l). \tag{19}$$

It should be noticed that in these expressions the central (local) supervisor does not take into account the cost of effort of the local (central) supervisor, and the local supervisor anticipates the liquidation threshold s^* used by the central supervisor.

The reaction functions of the two supervisors are given by

$$e_c(e_l) = \arg\max_{e_c} v_c(e_c, e_l), \tag{20}$$

$$e_l(e_c) = \arg\max_{e_l} v_l(e_c, e_l). \tag{21}$$

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

The following result provides closed form expressions for $v_c(e_c, e_l)$ and $v_l(e_c, e_l)$.

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

$$v_c(e_c, e_l) = \overline{R} - (1-a)\overline{R}\left[\Phi(x^*) + \frac{\phi(x^*)}{x^*}\right] - c_c(e_c), \tag{22}$$

$$v_l(e_c, e_l) = \overline{R} - (1-a)\overline{R}\left[\Phi(x^*) + \frac{\phi(x^*)}{x^*}\right] - \delta_l\Phi(x^*) - c_l(e_l), \tag{23}$$

where

$$x^* = \frac{s^* - E(s)}{Var(s)} = -\frac{(1-a)\overline{R}}{(1-c)\sigma}[1 + (e_c + e_l)^{-1}]^{1/2}. \tag{24}$$

The analysis in the previous section shows that the supervisors' payoff functions are not everywhere concave. For this reason, the reaction functions (20) and (21) may have corner or interior solutions. But even if we restrict attention to the case where solutions are interior, deriving results on the shape of the reaction functions requires restricting parameter values. For this reason, in what follows I will resort to numerical solutions for a set of 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 $e_c'(e_l)e_l'(e_c) < 1$. 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.

[FIGURE 4]

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 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 $\overline{R}$ 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]

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 $\overline{R}$, 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 $\varepsilon_c \sim N(0, \sigma^2/e_l)$, the local supervisor gets a signal $s_l = R + \varepsilon_l$, where $\varepsilon_l \sim N(0, \sigma^2/e_l)$,[8] and then it truthfully reports it to the central supervisor, who decides on the liquidation of the bank. By the properties of normal distributions we have

$$E(L \mid s_c, s_l) = a\overline{R} + \frac{c[s_{cl} - \overline{R}]}{Var(s_{cl})}, \tag{25}$$

$$E(R \mid s_c, s_l) = \overline{R} + \frac{s_{cl} - \overline{R}}{Var(s_{cl})}, \tag{26}$$

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

$$s_{cl} = \frac{e_c}{e_c + e_l} s_c + \frac{e_l}{e_c + e_l} s_l. \tag{27}$$

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

The random variable s_{cl} is normally distributed, with $E(s_{cl}) = \overline{R}$ and $Var(s_{cl}) = \sigma^2 \left[1 + (e_c + e_l)^{-1}\right]$, so it has the same distribution as the random variable s in (14). Moreover, it is also the case that $Cov(s_{cl}, L) = Cov(s_{cl}, R) = 1$. 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 send 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 the three 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 cost of bank liquidation incurred by the local supervisor, captured by parameter δ_l (which is 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. Note that 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).

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

$$\begin{aligned} w_l &= \int_{-\infty}^{\widehat{s}_l} E(L \mid s)dF(s) + \int_{\widehat{s}_l}^{\infty} E(R \mid s)dF(s) - c_l(\widehat{e}_l) \\ &= \overline{R} - \left[(1-a)\overline{R} + \delta\right] \left[\Phi(\widehat{x}_l) + \frac{\phi(\widehat{x}_l)}{\widehat{x}_l}\right] - c_l(\widehat{e}_l), \end{aligned} \tag{28}$$

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

$$\widehat{x}_l = -\frac{\left[(1-a)\overline{R} + \delta_l\right]}{(1-c)\sigma} \left(1 + \widehat{e}_l^{\,-1}\right)^{1/2}.$$

Similarly, social welfare under centralized supervision is given by

$$\begin{aligned} w_c &= \int_{-\infty}^{\widehat{s}_c} E(L \mid s)dF(s) + \int_{\widehat{s}_c}^{\infty} E(R \mid s)dF(s) - c_c(\widehat{e}_c) \\ &= \overline{R} - (1-a)\overline{R} \left[\Phi(\widehat{x}_c) + \frac{\phi(\widehat{x}_c)}{\widehat{x}_c}\right] - c_c(\widehat{e}_c), \end{aligned} \tag{29}$$

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

$$\widehat{x}_c = -\frac{(1-a)\overline{R}}{(1-c)\sigma} \left(1 + \widehat{e}_c^{\,-1}\right)^{1/2}.$$

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

$$\begin{aligned} w_h &= \int_{-\infty}^{s^*} E(L \mid s)dF(s) + \int_{s^*}^{\infty} E(R \mid s)dF(s) - c_c(e_c^*) - c_l(e_l^*) \\ &= \overline{R} - (1-a)\overline{R} \left[\Phi(x^*) + \frac{\phi(x^*)}{x^*}\right] - c_c(e_c^*) - c_l(e_l^*), \end{aligned} \tag{30}$$

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

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

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. Figure 6 shows the results.[9] 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 (small) bias in its liquidation decision. Hierarchical supervision dominates in Region H, where the cost advantage of the local supervisor is not so large or the liquidation cost of the local supervisor is not so 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, which does not make it worthwhile its participation in the joint supervision of the bank.[10]

[FIGURE 6]

Next, I analyze the effect on the region in Figure 6 of changes in the expected return $\overline{R}$ and in the standard deviation σ of the bank's investment return. Figure 7 illustrates these results. Panel A shows that an increase in $\overline{R}$ 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. Interestingly, both panels illustrate that for relatively small values of γ_c and δ_l that the optimal institutional design may switch from decentralized to centralized supervision without passing through the region where hierarchical supervision is optimal.[11]

[FIGURE 7]

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

[10] A key element of this result is the saving of the fixed cost γ_0 in the quadratic cost function (8).

[11] Notice 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.

Summing up, I have shown that hierarchical supervision dominates decentralized supervision when the possible capture of the local supervisor is a 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 bank profitability is low and bank risk-taking is high.

5 Limiting the Size of the Central Supervisor

This section 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 $\overline{e}_c$ 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 8 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, it follows that moving up the reaction function increases the central supervisor's payoff (in the case of Figure 8 until the corner where the central supervisor exerts no effort).

[FIGURE 8]

The question now 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 (22) and (30) imply that the social indifference curve at the point (e_c^*, e_l^*) may be written as

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

which 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 9 illustrates the result.

[FIGURE 9]

Summing up, reducing the effort provided by the central supervisor in the hierarchical supervision setup is always welfare improving. The intuition is that this change forces the local supervisor to increase its (cheaper) effort.[12] However, the model assumes that the efforts of the supervisors are not verifiable, so there is an issue about how this could this be implemented. In this regard, limiting the size of the central supervisor may be a second best way of increasing welfare.

6 Concluding Remarks

TBC

[12] This result 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."

Appendix

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

$$E(L \mid s) = a\overline{R} + c\sigma\left(1 + e^{-1}\right)^{-1/2} x,$$
$$E(R \mid s) = \overline{R} + \sigma\left(1 + e^{-1}\right)^{-1/2} x.$$

Hence, one can write

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

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

$$\overline{R} + \sigma\left(1 + e^{-1}\right)^{1/2} \widehat{x} = \widehat{s}.$$

Solving for $\widehat{x}$ and using the definition (10) of $\widehat{s}$ gives (12). Then, by the properties of normal densities,[13] and using the definition (12) of $\widehat{x}$, one concludes

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

Proof of Proposition 1 Given that

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

and

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

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

$$v'(e) = -\frac{(1 - a)\overline{R} + \delta}{2e(1 + e)} \frac{\phi(\widehat{x})}{\widehat{x}} - \gamma e = 0,$$

[13] In particular, the fact that $x\phi(x) = -\phi'(x)$.

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

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

Similarly, $v''(e) < 0$ and $\widehat{x} < 0$ imply

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

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

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

Finally, $v''(e) < 0$ and $\widehat{x} < 0$ also imply

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

given that $v''(e) < 0$ and $\widehat{x} < 0$. □

Proof of Proposition 3 By (14) one can write $s = \overline{R} + \sigma[1 + (e_c + e_l)^{-1}]^{1/2}x$, where $x \sim N(0,1)$. Substituting this expression into (15) and (16) gives

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

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

Hence, one can write

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

where x^* is implicitly defined by the equation

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

Solving for x^* and using the definition (17) of s^* gives (24). 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.

Argenziano, R., S. Severinov, and F. Squintani (2016), "Strategic Information Acquisition and Transmission," *American Economic Journal - Microeconomics*, 8, 119-155.

Boyer, P., and J. Ponce (2012), "Regulatory Capture and Banking Supervision Reform," *Journal of Financial Stability*, 8, 206-217.

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.

Eisenbach, T., D. Lucca, and R. Townsend (2016), "The Economics of Bank Supervision," Federal Reserve Bank of New York Staff Report No. 769.

Holmström, B. (1982), "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324-340.

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.

Repullo, R., and J. Suarez (2004), "Venture Capital Finance: A Security Design Approach," *Review of Finance*, 8, 75-108.

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.

Figures

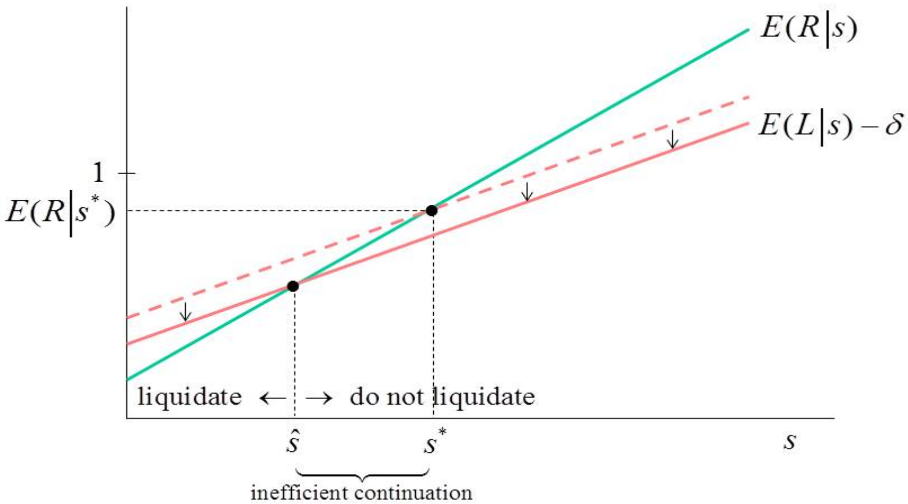

Figure 1. Supervisor's liquidation threshold

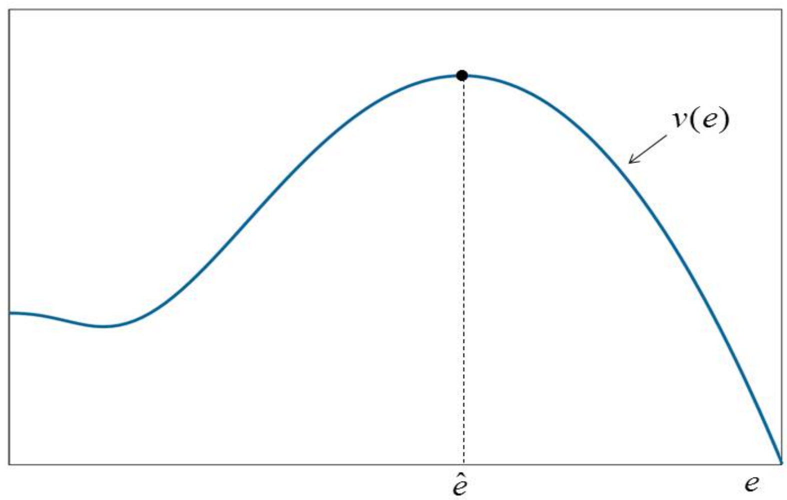

Figure 2. Supervisor's payoff function

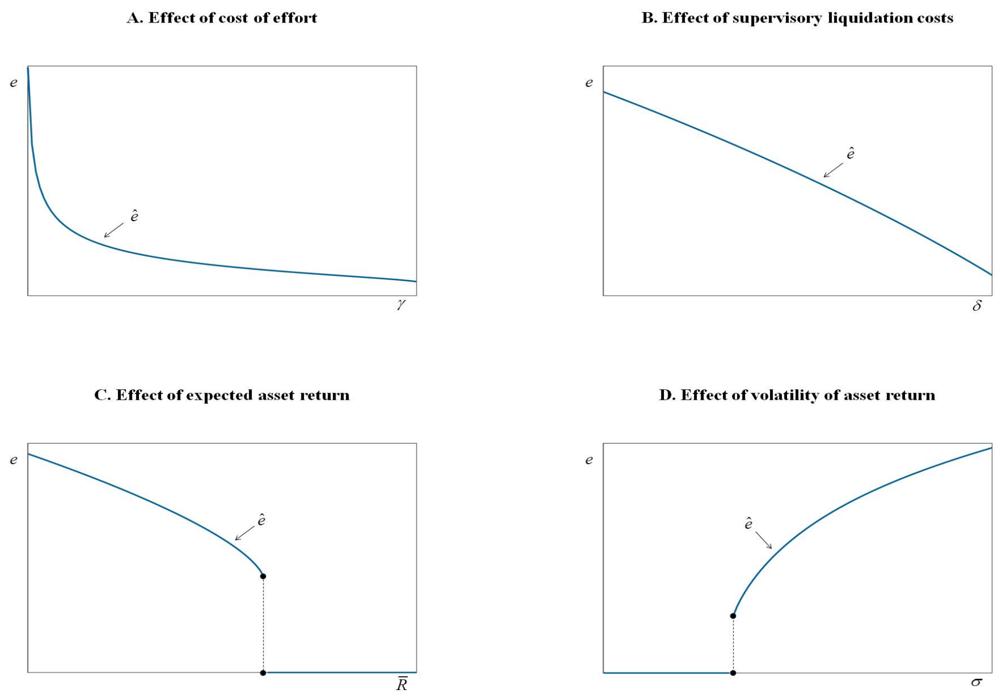

Figure 3. Comparative statics

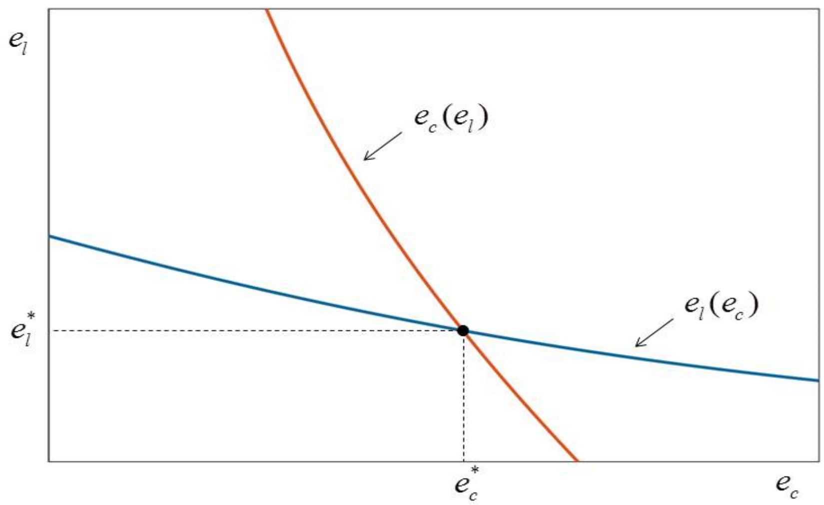

Figure 4. Nash equilibrium

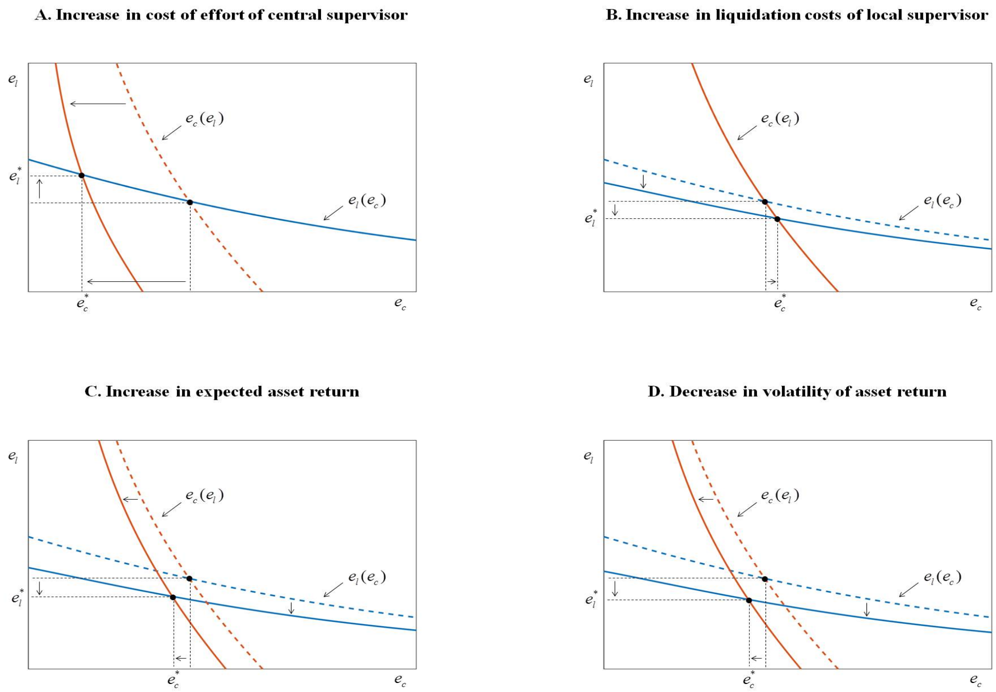

Figure 5. Comparative statics of Nash equilibrium

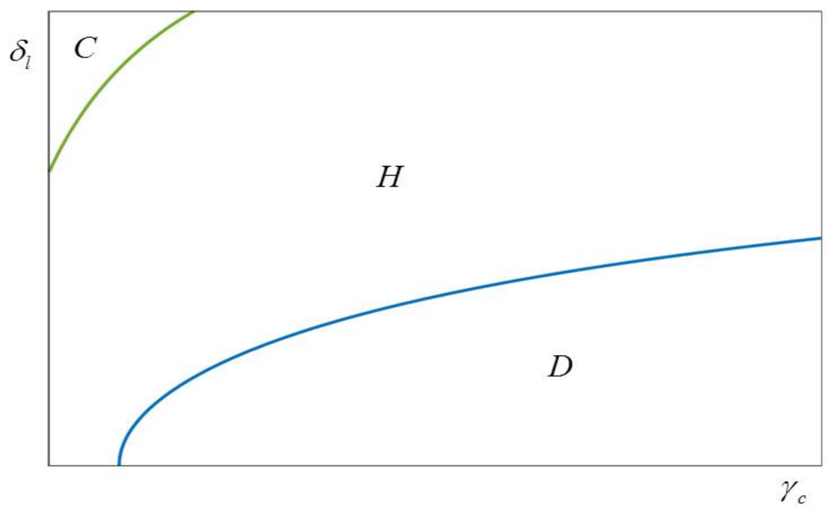

Figure 6. Optimal institutional design

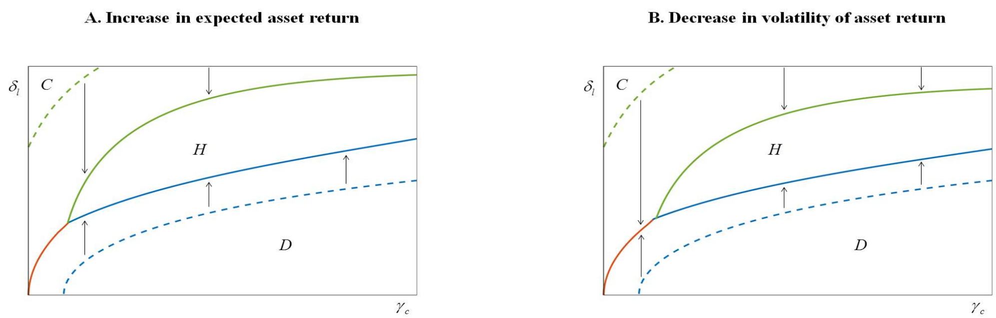

Figure 7. Comparative statics of optimal institutional design

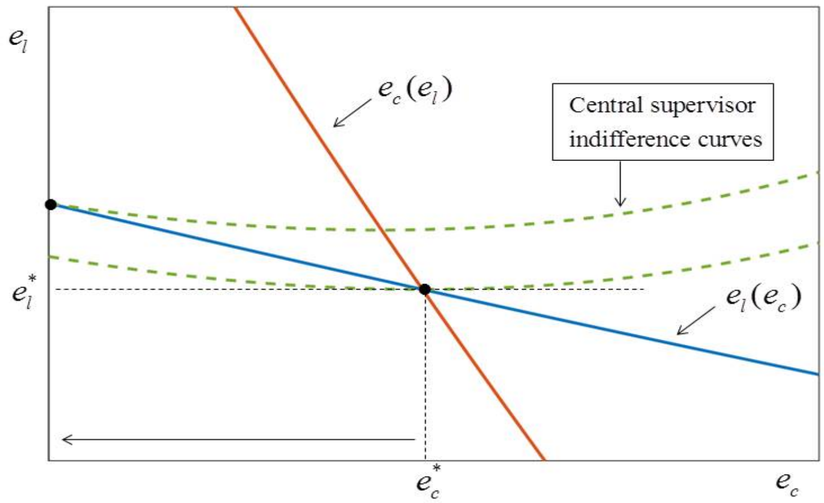

Figure 8. Central supervisor as Stackelberg leader

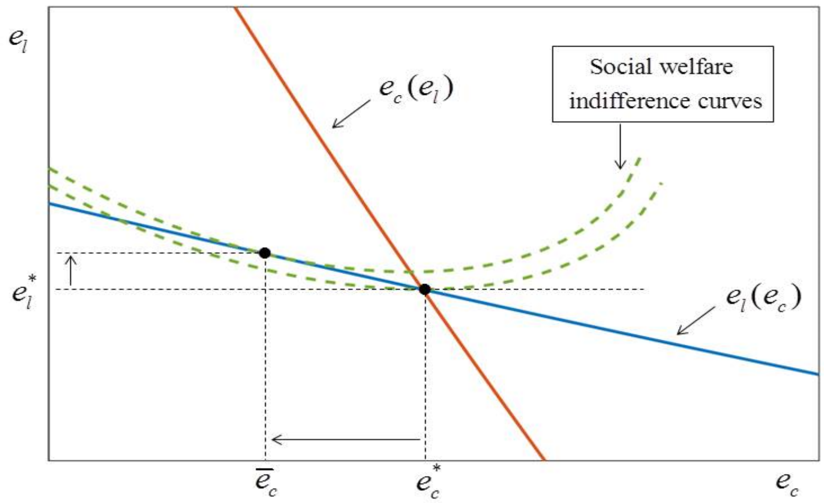

Figure 9. Limiting size of central supervisor