Designing Agile Banking Supervision*

Jeong Ho (John) Kim† Kyungmin Kim‡ Victoria Liu§ Noam Tanner¶

July 6, 2023

Abstract

The supervisor guides the bank's risk-taking decisions by providing communication based on his private information about the economy. The bank takes on high risk provided that, based on its private information and the supervisor's communication, it think the supervisor is unlikely to object to high risk. Ultimately, the supervisor allows high risk only if he thinks the economy is likely to be strong. We show that welfare can deteriorate when there is more information on the bank's side: the supervisor is too eager to respect the bank's risk-taking decisions when it is relatively more informed, which in turn dilutes his ability to induce the bank to reveal its information. We propose two methods to mitigate this problem: (i) err on the side of giving the supervisor too much power in case the bank does not meet supervisory expectations and (ii) give the supervisor ex-ante commitment power over his supervisory ruling.

JEL Classification: D82, D83, G21. *Keywords:* Banking supervision, Disclosure, Information design, Mechanism design.

*We are grateful for comments from our discussants, Andrea Passalacqua, Keeyoung Rhee, and Anthony Lee Zhang. We are also grateful for comments from Nicolas Inostroza, Yaron Leitner, Stephen Morris, conference participants at the FMA Annual Meeting, the KIF-KAEA-KAFA Symposium, the MFA Annual Meeting, and seminar participants at Boston Fed. The views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Boston or the Federal Reserve system.

†Department of Economics, Emory University, *Email:* jkim230@emory.edu

‡Department of Economics, Emory University, *Email:* kyungmin.kim@emory.edu

§Federal Reserve Bank of Boston, *Email:* victoria.liu@bos.frb.org

¶Federal Reserve Bank of Boston, *Email:* noam.tanner@bos.frb.org

1 Introduction

Banking supervision would be easy, if risks within the economy are publicly known by the supervisor and the banks. Difficulty level of the challenge increases when only banks know this information perfectly. Supervisor faces the obstacle of incentives mismatch. However, the cruel reality is that despite how much effort is put into learning about the underlying risks, neither the supervisor nor the banking industry has perfect insight. Banks have very detailed views of their own portfolios, but they cannot look into the business lines of their peers, not to mention market as a whole. In the recent event of Archegos, dealers admitted to not having the ability to obtain complete information on even their client portfolios [1]. Supervisor, on the other hand, is able to probe into portfolios across all of her supervised institutions, despite a lack of finer details. Information is imperfect for all players. However, as if double sided information uncertainty is not challenging enough, dynamics within the information structure rarely stay unchanged. Absolute and relative strengths of private signals vary greatly as conditions of the market change. To achieve socially desirable outcome, supervisor needs to actively take this information dynamics into account while designing her public message to guide and monitor bank behaviors. While cost of supervisory objection might stay constant given the infrastructure setup, supervision communication should be an agile response to the fluid information conditions.

Our main contribution is to formally model this two-sided information uncertainty, and allow supervisor to learn from actions taken by the banks. We solve for the optimal disclosure rule within this framework. In our model, socially desirable outcome is when banks take the appropriate risk management approach depending on the state of the world. If economic fundamental is strong, then supervisor allows banks to be aggressive in risk management practices, as economy is more likely to withstand idiosyncratic events. If economic fundamental is weak, then supervisor prefers banks to be conservative, so that any idiosyncratic event will not trigger systemic fragility. Therefore, supervisor wants to induce separation of banks' behavior conditional on its private signals and minimize the probability she has to costly object to banks' actions. This notion of socially desirable outcome leads to three main results. First, optimal disclosure strategy depends on the information structure between supervisor and banks relative to rejection cost. There is no single

[1] Since Archegos broke up their trading relationships across multiple dealers, individual dealer was unable to figure out the actual leverage the firm has put on in aggregate. Excessive leverage is believed to have caused the unravel of Archegos. This inability to get aggregate risk information and manage accordingly has cost some dealers billions in losses.

optimal disclosure strategy that supervisor can follow in all situations. While cost of rejection cannot change, information disclosure should be nimbly adjusted as economic conditions turn. Second, while optimal persuasion rule can Pareto improve on the cases of no disclosure and full disclosure, it cannot achieve first best outcome. By introducing just enough modification on supervisor's private signal, supervisor can induce more separation on the banks. However, there is still going to be some pooling left, resulting in some welfare loss. Lastly, improvements in supervisory and bank signals do not necessarily lead to more separation or an easier mechanism design problem on cost of rejection.

In our model, banks always prefer to take the aggressive risk management approach regardless of state of the economy. This ranges from calibrating model parameters to benign market conditions to having flexible policy and procedures. After all, capital set aside by risk is capital that is not generating returns. However, this aggressive action might have costly consequences when supervisor has sufficient reasons to believe that economic conditions are not good enough to support such practices. Dividend payments will be capped if supervisor objects to a capital plan during stress testing. Findings, that will ultimately impact bank's supervisory ratings, will be issued if bank fails to meet expectations during a regulatory exam. Banks will then have to spend significant amount of resources to remediate the matter requiring attention (MRA), or in worse situation matter requiring immediate attention (MRIA). When belief is not optimistic enough to overcome such costs, banks will reconsider and consequently take the conservative approach. Therefore, banks do not take the aggressive risk management approach regardless of potential supervisor action. In fact, banks will evaluate their chances of acceptance while making decisions.

Before banks take their action, supervisor has a chance to signal to the bank what she knows or some modification of what she knows. This information is implicit in the scenarios she hands out in a stress test, exam scope she delivers in a first day letter, and guidance she issues on topics of interest to the industry. When supervisor has very accurate information, she should fully disclose what she knows. It is better for banks to pool on action permissible under supervisory signal. This eliminates inefficiencies from banks trying to use their own information that might result in needless rejection or lower capital utilization. Things become complicated when supervisor and bank have imperfect information that neither is certain enough on its own to make accurate inference about the underlying risks. This leaves banks a chance to change supervisory opinion on the economy with their collective actions. Banks want to behave in ways that boost supervisor's confidence about the economy, and consequently result in higher likelihood for their aggressive actions to pass. Supervisor has to disclose her information in such a way that she disciplines banks' beliefs to be thoughtfully conservative, but does not completely deter banks from appropriate risk taking if separation could occur. Her goal can be achieved by introducing some messaging noise on the good state of the world, when a positive public message is sent. The result here is a bit counter-intuitive, that a good public message does not guarantee a good private signal for the supervisor. It is precisely this uncertainty that urges banks to be mindful and separate, as it correctly infers a positive probability with which supervisor received a bad signal and therefore will not allow aggressive action sparingly. The increase in utilization of banks' private information will result in more separation. However, persuasion is unable to induce full separation. Banks still pool with positive probability.

The final piece of banking supervision is setting the cost for when bank's aggressive action is objected by the supervisor. Cost setting is more rigid than the previously mentioned information exchange, as legislature requires these consequences to be established ahead of time and stay fixed until legislative changes occur. Some criticism around stress testing is that capping dividend and suspending share repurchases are too severe as disciplinary measures. However, it is not at discretion of the supervisor to reset this cost from one period to the next. We show that, if cost of rejection cannot be set just right and swiftly change to stay so, supervisor can use corresponding optimal disclosure rule to mitigate some of the welfare losses. The excessive amount of caution caused by high rejection cost, and aggression caused by low rejection cost can be partially mitigated through a carefully designed disclosure rule. While it is not possible to achieve the most socially desirable outcome, strategic ambiguity introduced by supervisory disclosure will restore some separation and hence improve welfare. We also show that improvement in signal precision does not monotonically induce more separation.

2 Related Literature

We explore the optimal disclosure strategy for the supervisor under a setting of two-sided information uncertainty. In traditional setup, agent has perfect information and biased preference. However, in this principal-agent problem, both the principal and the agent have private information. Principal can choose to disclose a public message. Then the agent takes an action. Lastly, the principal takes a follow up action, which might incur some cost to the agent.

This paper uses concepts of Bayesian persuasion as outlined in Kamenica and Gentzkow

(2011). Supervisor commits to a disclosure rule, and induces a Bayesian updated distribution of beliefs on the bank which advantages the supervisor. In addition, information structure in this paper also relates to literature on information design Taneva (2019); Bergemann and Morris (2013). Like this literature, this paper explores not only the impact of changing payoffs and allocations in order to incentivize actions, but also the impact of changing information structure in order to alter actions of the players. Information and persuasion design have previously been applied to voting Alonso and Câmara (2016) and stress testing Goldstein and Leitner (2018) and Inostroza and Pavan (2021). This paper applies the design setting to a new environment of banking supervision. Instead of focusing on bank runs or stress testing in particular, we investigate how to incentivize best risk management practice in a generalized setting. We also eliminated perfect information regarding the state of the world. Both the supervisor and bank receive imperfect private information. The new ingredient of the model is allowing supervisor an attempt to learn from the bank and increase information utilization.

The most relevant paper is that of Leitner and Williams (2023), where the authors provide a general characterization for the optimal disclosure rule for stress testing in the case of one bank and one supervisor. Our paper differs from that in that it solves a less general model, but addresses a setting where there are multiple banks not perfectly observing the state variable. Thus, there is two-sided learning (banks learn from the supervisor and the supervisor learns from each individual bank and the profile of bank risk management choices). As in Leitner and Williams (2023), we find that full disclosure can yield lower payoffs to the supervisor. However, our mechanism is different. In our model, while banks may take suboptimal action, the dominance of no disclosure over full disclosure is the result of the incentive constraints and the properties of beliefs of the banks. In addition, we find that full disclosure may dominate no disclosure (privacy) in certain settings. In addition, as in Leitner and Williams (2023), we find that partial disclosure in the form of a persuasion rule may be optimal. However, we find that this optimality of persuasion arises only in cases when full disclosure yields higher payoff than no disclosure and induces partial separation.

This paper also considered the possibility of private feedback. Recent literature such as Eső and Szentes (2007) and Bergemann et al. (2022) allow for different feedback responses depending on agents and types. For example, Li and Shi (2017) showed that in auctions settings, releasing different amounts of additional information to different buyer types dominates full disclosure in terms of seller revenue. However, we keep away from discriminatory information disclosure. While there could be improvements from tailoring information to recipients, it is against the reality of supervision faced by supervisors. On one hand, supervisors deliver consistent message across institutions for fairness. On the other hand, industry participants talk to each other through formal industry forums (such as ISDA for uncleared margins rule) and revolving doors. Informational discrepancies are very short lived before they are challenged and leveled. For these reasons, we decide to skip this channel and focus on informational tradeoff in a simpler setting.

There is also empirical literature studying the impact various sources of economic uncertainty have on investment and output. Papers such as Baker et al. (2016) document an adverse effect of policy uncertainty on key economic activity measures, such as investment and employment. These findings are consistent with our results when supervisor has superior information. Optimal outcome in this case calls for supervisor to fully disclose her information and minimize uncertainty for the bank to just what is embedded in its private signal. Supervisor does not prefer separation under all conditions. However, we show that in the case where information for the supervisor is not good enough on its own, optimal disclosure deviates from full disclosure. Leaving some signaling uncertainty in the public message actually induces separation in bank's action conditional on receiving different private signals, and therefore increases social welfare.

3 The Model

There is a bank and a supervisor. The bank decides whether to take high risks with weak risk management practices ("Aggressive") or take low risks with strong risk management practices ("Conservative"). The payoff from high risk endeavors potentially depends on the realization of a random variable $\omega \in \{G, B\}$. We think of ω as the state of the economy, which can be good or bad. The bank's payoff is u_ω and the supervisor's payoff is v_ω. We think of v_ω as the value of high risk endeavors to society. Without loss of generality, the payoff from conservative risk management practices does not depend on ω and is in turn normalized to zero for both the bank and the supervisor.[2]

We focus on the case where the following inequalities hold: $u_G = u_B > 0$ and $v_G > 0 > v_B$. These assumptions encode the following two economic intuitions: (i) the bank prefers high risk endeavors to conservative risk management practices in every state of the economy; but (ii) the supervisor prefers high risk endeavors only if the state of the economy is good. In other words, there is a conflict of interest between the bank and the supervisor. In this case, the payoff from aggressive risk-taking when $\omega = G$ can be further normalized to one for both the bank and the supervisor: $u_G = v_G = 1$. Given our assumptions on payoffs, $u_B = 1$ and $v_B = -d$, where $d > 0$.

Both the bank and the supervisor do not observe ω, but each of them receives a private signal about the state of the economy. In other words, there is incomplete information.

On the one hand, the bank's signal s takes one of two values, g or b. We let γ denote the probability that the bank receives signal $s = g$ (respectively, $s = b$) when $\omega = G$ (respectively, $\omega = B$). We assume that $\gamma > \frac{1}{2}$, so that the signal is indeed informative about the state. We also assume that $\gamma < 1$, so that the bank does not perfectly observe ω.

On the other hand, the supervisor observes the probability t that $\omega = G$. We refer to t, which itself is distributed as F on $[0, 1]$, as the supervisor's "type." Note that this formulation is equivalent to the standard one in which we would specify the prior probability t_0 that $\omega = G$ and the supervisor's signal $s' \in [\underline{s}, \overline{s}]$ that has CDF F_ω conditional on ω.

After the bank decides on its risk level, the supervisor assesses the bank's risk-management practices, in which he observes its risk level and decides whether to allow or object to its risk-management practices. If the supervisor allows the bank's risk-management practices, it keeps the

[2]For our purposes, all that matters is the *relative gains* from taking on high risk endeavors, compared to holding safe assets.

risk level intact as it chose. If the supervisor objects to the bank's risk-management practices, it is forced to readjust its risk level to be low. In this case, the bank incurs a cost $c > 0$, which could represent the fact that when it is forced to sell its high-risk assets, such assets will likely sell at fire sale prices. Another way to interpret the cost c is that it reflects the bank's cost of reputation loss. An implicit assumption here is that the supervisor never objects to a conservative bank.

Before the bank decides on its risk level, the supervisor discloses information about his type t. The supervisor's communication strategy is modeled following the recent literature on information design. Formally, it consists of an arbitrary finite set M of messages and a function $\pi : [0, 1] \to M$, where $\pi(t)$ denotes the message that the supervisor of type t picks to send. We let $F(\cdot|m)$ represent the bank's posterior belief distribution about the supervisor's type t after observing $m \in M$. Without loss of generality, we assume $\Pr(m) > 0$ for all $m \in M$. We let $\delta(m) \in \{0, 1\}$ denote the bank's (observed) risk level following message $m \in M$ (where 1 and 0 stand for "Aggressive" and "Conservative", respectively).

In summary, the timing of the game is as follows. The supervisor publicly commits to his communication strategy (M, π). Nature chooses ω, the bank observes s, and the regulator observes t. The supervisor discloses information about t according to his communication strategy. The bank decides whether to take high risks ("Aggressive") or take low risks ("Conservative"). The supervisor assesses the bank's risk-management practices, i.e., if the bank is aggressive, he decides whether to accept or object to its risk-management practices. Finally, the payoffs are realized.

3.1 Preliminaries

We begin by demonstrating that, insofar as we consider a pure-strategy equilibrium, there are at most three different outcomes after the supervisor discloses information, which in turn implies that we can restrict attention to his communication strategy that involves a ternary message space.

3.1.1 Supervisor's Policy

Let q denote the probability that the supervisor thinks the state of the economy is good ($\omega = G$). If the bank is aggressive, the supervisor's expected payoff is $q - (1 - q)d$. If the bank is conservative, the supervisor's payoff is zero. Hence, he allows the bank's aggressive risk-taking if and only if

$$q \geq \widehat{t} := \frac{d}{1 + d}. \tag{1}$$

In other words, based solely on his private information, the supervisor allows aggressive risk-taking if and only if $t \geq \widehat{t}$.[3]

Let $\sigma_S(t, m)$ denote the supervisor's final (equilibrium) action, whether to allow the bank's aggressive risk-taking ($\sigma_S(t, m) =$ "Allow") or not ($\sigma_S(t, m) =$ "Object"), depending on his type $t \in [0, 1]$ and his message $m \in M$.

Lemma 1 *Suppose that, for a given communication strategy* (M, π)*, there are more than one* t *in the support of* $F(\cdot|m)$*. Then there exists* $t^* \in (0, 1)$ *such that* $\sigma_S(t, m) =$ *"Accept" if* $t \geq t^*$*, while* $\sigma_S(t', m) =$ *"Object" if* $t < t^*$.

Proof. By Bayes' rule,

$$\begin{aligned}\Pr(\omega = G|t, \delta(m) = 1) &= \frac{\Pr(\delta(m) = 1|\omega = G)t}{\Pr(\delta(m) = 1|\omega = G)t + \Pr(\delta(m) = 1|\omega = B)(1 - t)} \\ &\leq \Pr(\omega = G|t', \delta(m) = 1) = \frac{\Pr(\delta(m) = 1|\omega = G)t'}{\Pr(\delta(m) = 1|\omega = G)t + \Pr(\delta(m) = 1|\omega = B)(1 - t')}\end{aligned}$$

The desired result then follows from the fact that the supervisor allows aggressive risk-taking if and only if $Pr(\omega = G|t', \delta(m) = 1) \geq \widehat{t}$. ■

The lemma captures the intuition that the supervisor allows (object to) aggressive risk-taking if a supervisor of more pessimistic (optimistic) type would allow (object to) it.

3.1.2 Bank's Risk-Taking

Let p denote the probability that the bank thinks the supervisor will allow its aggressive risk-taking. If the bank is aggressive, its expected payoff is $p - (1 - p)c$. If the bank is conservative, its payoff is zero. Hence, the bank is aggressive only if

$$p \geq \widehat{p} := \frac{c}{1 + c}.\text{[4]} \tag{2}$$

[3] Note that we assume that the supervisor allows aggressive risk-taking if his perceived probability q of $\omega = G$ is exactly equal to $\widehat{t}$. This assumption is benign for our characterization below, as the event occurs with probability zero under the supervisor's optimal communication strategy.

[4] We assume that the bank plays a pure strategy (aggressive or conservative, but not both) if its perceived probability of $\sigma_S(m) =$ "Allow" is exactly $\widehat{p}$. This assumption can be relaxed to allow for a mixed strategy by the bank. XXX asdf XXX

Let $\sigma_B(s,m)$ denote the bank's choice of risk-taking, whether to be aggressive ($\sigma_B(s,m)=1$) or not ($\sigma_B(s,m)=0$), conditional on receiving signal $s \in \{g,b\}$ and observing message $m \in M$.

Lemma 2 *Suppose that, for a given communication strategy (M,π), there are more than one t in the support of $F(\cdot|m)$. Then $\sigma_B(l,m)=1$ implies $\sigma_B(h,m)=1$, while $\sigma_B(h,m)=0$ implies $\sigma_B(l,m)=0$.*

Proof. Let $F(\cdot|s,m)$ denote the bank's posterior belief distribution about the supervisor's type conditional on receiving $s \in \{g,b\}$ and observing $m \in M$. By Bayes' rule, we have

$$F(t|g,m) = \frac{\int_0^t [\gamma x + (1-\gamma)(1-x)]dF(x|m)}{\int_0^1 [\gamma x + (1-\gamma)(1-x)]dF(x|m)}$$

and

$$F(t|b,m) = \frac{\int_0^t [(1-\gamma)x + \gamma(1-x)]dF(x|m)}{\int_0^1 [(1-\gamma)x + \gamma(1-x)]dF(x|m)}$$

Notice that $\gamma x + (1-\gamma)(1-x) = (2\gamma-1)x + 1 - \gamma$ is strictly increasing in x, while $(1-\gamma)x + \gamma(1-x) = \gamma - (2\gamma-1)x$ is strictly decreasing in x, so there exists $t^* \in (0,1)$ such that

$$\frac{\gamma x + (1-\gamma)(1-x)}{\int_0^1 [\gamma x + (1-\gamma)(1-x)dF(x|m)} \leq \frac{(1-\gamma)x + \gamma(1-x)}{\int_0^1 [(1-\gamma)x + \gamma(1-x)]dF(x|m)}$$

if and only if $t \leq t^*$. Since $F(\cdot|g,m)$ and $F(\cdot|b,m)$ have the same support, this implies that $F(t|g,m) \leq F(t|b,m)$ for all t, with strict inequalities holding for any t in the interior of their support. This means that $\Pr(t \geq t^*|g,m) > \Pr(t \geq t^*|b,m)$ for any t^* in the interior of the support of $F(\cdot|g,m)$ and $F(\cdot|g,m)$. The desired result then follows from Lemma 1. ■

The lemma captures the idea that, if the bank would be willing to be aggressive (conservative) after receiving a bad (good) signal and observing a message, it will be aggressive (conservative) after receiving a good (bad) signal and observing the same message.

3.1.3 Simplifying the Supervisor's Communication Strategy

Lemma 2 shows that there are at most three different outcomes after the bank observes $m \in M$: (i) it will be aggressive regardless of its signal; (ii) it will be aggressive if its signal was good ($s = g$)

and be conservative if its signal was bad ($s = b$); and (iii) it will be conservative regardless of its signal. Importantly, this implies that the message space M can be partitioned into three subsets:

$$M^{1,1} := \{m \in M | \sigma_B(g, m) = \sigma_B(b, m) = 1\}$$
$$M^{1,0} := \{m \in M | \sigma_B(g, m) = 1, \sigma_B(b, m) = 0\}$$
$$M^{0,0} := \{m \in M | \sigma_B(g, m) = \sigma_B(b, m) = 0\}$$

Note that the supervisor cannot benefit by using multiple messages that lead to the same outcome; in what follows, we can restrict attention, with no loss of generality, to strategy profiles in which the supervisor sends (says) $m \in M^* := \{(1, 1), (1, 0), (0, 0)\}$, where the bank is aggressive regardless of its signal after observing $m = (1, 1)$, only if its signal is good ($s = g$) after observing $m = (1, 0)$, and never after $m = (0, 0)$.

4 Equilibrium Characterization

In this section, we characterize the unique equilibrium of the dynamic supervision game. We begin by observing that the supervisor updates his perceived probability q of $\omega = G$ when he observes the bank's risk level only after having sent the message $m = (1, 0)$. In this case, the supervisor learns that the bank's signal was good (bad) by observing that it is aggressive (conservative) in its risk-taking. Let $\bar{t}$ be the supervisor's type that would be indifferent between allowing and objecting to aggressive risk-taking after learning that the bank received signal $s = b$:

$$\frac{(1-\gamma)\bar{t}}{(1-\gamma)\bar{t} + \gamma(1-\bar{t})} = \widehat{t} \iff \frac{\bar{t}}{1-\bar{t}} = \frac{\gamma}{1-\gamma}\frac{\widehat{t}}{1-\widehat{t}} = \frac{\gamma}{1-\gamma}d \tag{3}$$

Similarly, let $\underline{t}$ be the supervisor's type that would be indifferent between allowing and objecting to aggressive risk-taking after learning that the bank received signal $s = g$:

$$\frac{\gamma\underline{t}}{\gamma\underline{t} + (1-\gamma)(1-\underline{t})} = \widehat{t} \iff \frac{\underline{t}}{1-\underline{t}} = \frac{1-\gamma}{\gamma}\frac{\widehat{t}}{1-\widehat{t}} = \frac{1-\gamma}{\gamma}d \tag{4}$$

Note that $\underline{t} < \widehat{t} < \bar{t}$: (i) the supervisor of type $t \in [\underline{t}, \widehat{t})$, who would object to aggressive risk-taking a priori, is swayed to allow it by the bank's good news, and (ii) the supervisor of type $t \in [\widehat{t}, \bar{t})$, who would allow aggressive risk-taking a priori, is swayed to object to it by the bank's bad news.

On the other hand, the supervisor learns nothing about the bank's signal after having sent the message $m = (1, 1)$, since in this case the bank is always aggressive regardless of its signal, so the supervisor allows aggressive risk-taking if and only if $t \geq \widehat{t}$.

In summary, a type-t supervisor obtains an expected payoff of $\gamma t - (1-\gamma)(1-t)d$ if $t \geq \underline{t}$ and zero otherwise from sending the message $m = (1, 0)$, whereas he obtains an expected payoff of $t - (1-t)d$ if $t \geq \widehat{t}$ and zero otherwise from sending the message $m = (1, 1)$. Finally, note the less interesting case, where the supervisor sends the message $m = (0, 0)$ and the bank is always conservative, so the supervisor takes no further action and his payoff is always zero. Of course, all these statements implicitly assume that each message induces the bank to behave as expected. These assumptions can formally be expressed as a group of incentive compatibility constraints for the bank when it receives signal $s \in \{g, b\}$ and message $m \in M^*$.

4.1 Supervisor's Problem

We are now ready to write down the supervisor's problem:

$$\max_{T^{(1,0)},T^{(1,1)} \subset [0,1]} \int_{T^{(1,0)} \cap [\underline{t},1]} [\gamma t - (1-\gamma)(1-t)d] dF(t) + \int_{T^{(1,1)} \cap [\widehat{t},1]} [t - (1-t)d] dF(t) \tag{5}$$

subject to $T^{(1,0)} \cap T^{(1,1)} = \emptyset$ and

$$\int_{T^{(1,0)} \cap [\underline{t},1]} [\gamma t + (1-\gamma)(1-t)] dF(t) \geq c \int_{T^{(1,0)} \cap [0,\underline{t})} [\gamma t + (1-\gamma)(1-t)] dF(t) \qquad (IC_g^{(1,0)})$$

$$\int_{T^{(1,0)} \cap [\underline{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t) \leq c \int_{T^{(1,0)} \cap [0,\underline{t})} [(1-\gamma)t + \gamma(1-t)] dF(t) \qquad (IC_b^{(1,0)})$$

$$\int_{T^{(1,1)} \cap [\widehat{t},1]} [\gamma t + (1-\gamma)(1-t)] dF(t) \geq c \int_{T^{(1,1)} \cap [0,\widehat{t})} [\gamma t + (1-\gamma)(1-t)] dF(t) \qquad (IC_g^{(1,1)})$$

$$\int_{T^{(1,1)} \cap [\widehat{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t) \geq c \int_{T^{(1,1)} \cap [0,\widehat{t})} [(1-\gamma)t + \gamma(1-t)] dF(t) \qquad (IC_b^{(1,1)})$$

$$\int_{T^{(0,0)} \cap [\widehat{t},1]} [\gamma t + (1-\gamma)(1-t)] dF(t) \leq c \int_{T^{(0,0)} \cap [0,\widehat{t})} [\gamma t + (1-\gamma)(1-t)] dF(t) \qquad (IC_g^{(0,0)})$$

$$\int_{T^{(0,0)} \cap [\widehat{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t) \leq c \int_{T^{(0,0)} \cap [0,\widehat{t})} [(1-\gamma)t + \gamma(1-t)] dF(t) \qquad (IC_b^{(0,0)})$$

where $T^{(0,0)} := [0, 1] \setminus (T^{(1,0)} \cup T^{(1,1)})$.

First, $T^{(1,0)}$ and $T^{(1,1)}$ denote disjoint subsets of $[0,1]$ such that the supervisor chooses to send the message $m=(1,0)$ if he is of type $t \in T^{(1,0)}$ and $m=(1,1)$ if he is of type $t \in T^{(1,1)}$; the supervisor must be choosing to sending the message $m=(0,0)$ otherwise, i.e., if he is of type $t \in [0,1] \setminus (T^{(1,0)} \cup T^{(1,1)})$. This subset of $[0,1]$ is naturally denoted by $T^{(0,0)}$.

Second, (IC_s^m) represents the bank's incentive constraint when it receives signal $s \in \{g, b\}$ and message $m \in M^*$. To understand these IC constraints, suppose, for example, that the bank has received signal $s = l$ and message $m = (1,0)$. In this case, the bank is willing to be conservative if and only if it thinks the supervisor will allow aggressive risk-taking with probability less than or equal to $\widehat{p}$. Recalling that only the supervisor of type $t \geq \underline{t}$ allows aggressive risk-taking, this condition is equivalent to

$$\Pr(t \geq \underline{t} | s = b, m = (1,0)) = \frac{\int_{T^{(1,0)} \cap [\underline{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t)}{\int_{T^{(1,0)}} [(1-\gamma)t + \gamma(1-t)] dF(t)} \leq \widehat{p}$$

Substituting the expression for $\widehat{p}$ and rearranging terms, we have ($IC_b^{(1,0)}$). One can analogously derive the other IC constraints. Perhaps surprisingly, however, the following three lemmas establish that none of the IC constraints, except for ($IC_b^{(1,0)}$), are binding, substantially simplifying the supervisor's problem.

Lemma 3 *Suppose that $T^{(1,0)}, T^{(1,1)} \subset [0,1]$ satisfy incentive constraints ($IC_b^{(1,1)}$) and ($IC_g^{(0,0)}$). Then they satisfy incentive constraints ($IC_g^{(1,1)}$) and ($IC_b^{(0,0)}$).*

Proof. We show that ($IC_b^{(1,1)}$) implies ($IC_g^{(1,1)}$). To this end, first observe that

$$\int_{T^{(1,1)} \cap [\widehat{t},1]} [(1-\gamma)\widehat{t} + \gamma(1-\widehat{t})] dF(t) \geq \int_{T^{(1,1)} \cap [\widehat{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t)$$
$$\geq c \int_{T^{(1,1)} \cap [0,\widehat{t})} [(1-\gamma)t + \gamma(1-t)] dF(t) \geq c \int_{T^{(1,1)} \cap [0,\widehat{t})} [(1-\gamma)\widehat{t} + \gamma(1-\widehat{t})] dF(t)$$

where the first and the last inequalities hold because $(1-\gamma)t + \gamma(1-t) = \gamma - (2\gamma - 1)t$ is strictly decreasing in t for $\gamma > \frac{1}{2}$, while the second inequality is ($IC_b^{(1,1)}$) itself. Therefore, we have

$$\int_{T^{(1,1)} \cap [\widehat{t},1]} dF(t) \geq c \int_{T^{(1,1)} \cap [0,\widehat{t})} dF(t)$$

In addition, observe that

$$\int_{T^{(1,1)}\cap[\widehat{t},1]}[\gamma t+(1-\gamma)(1-t)]dF(t) \geq \int_{T^{(1,1)}\cap[\widehat{t},1]}[\gamma\widehat{t}+(1-\gamma)(1-\widehat{t})]dF(t)$$

and

$$c\int_{T^{(1,1)}\cap[0,\widehat{t})}[\gamma\widehat{t}+(1-\gamma)(1-\widehat{t})]dF(t) \geq c\int_{T^{(1,1)}\cap[0,\widehat{t})}[\gamma t+(1-\gamma)(1-t)]dF(t)$$

because $\gamma t+(1-\gamma)(1-t)=(2\gamma-1)t+1-\gamma$ is strictly increasing in t for $\gamma>\frac{1}{2}$. Combined with the inequality above, it follows that $(IC_g^{(1,1)})$ holds.

Proceeding similarly as above, one can conclude that $(IC_g^{(0,0)})$ implies $(IC_b^{(0,0)})$. ■

The lemma captures the idea that, if the bank would be willing to be aggressive (conservative) after receiving a bad (good) signal and observing $m=(1,1)$ ($m=(0,0)$), it will be aggressive (conservative) after receiving a good (bad) signal and observing $m=(1,1)$ ($m=(0,0)$).

Lemma 4 *At the optimal solution to the supervisor's problem, $T^{(1,1)}\subset[\widehat{t},1]$ and $T^{(0,0)}\subset[0,\widehat{t})$, so incentive constraints $(IC_b^{(1,1)})$ and $(IC_g^{(0,0)})$ are not binding.*

Proof. Suppose that $V:=T^{(1,1)}\cap[0,\widehat{t})\neq\emptyset$ at the optimal solution. Consider $\tilde{T}^{(1,1)}=T^{(1,1)}\backslash V$ and $\tilde{T}^{(0,0)}=T^{(0,0)}\cup V$. This change relaxes $(IC_b^{(1,1)})$ and $(IC_g^{(0,0)})$ by lowering the right-hand side of $(IC_b^{(1,1)})$, while raising the right-hand side of $(IC_g^{(0,0)})$. However, it leaves both $(IC_g^{(1,0)})$ and $(IC_b^{(1,0)})$ unaffected. For any $t\in V$, the supervisor obtains zero payoff if he were to send $m=(0,0)$, which strictly exceeds $t-(1-t)d$, i.e., what he obtains if he sends $m=(1,1)$. Therefore, this change raises the supervisor's expected payoff, which is a contradiction.

Proceeding similarly as above, one can prove that $T^{(0,0)}\subset[0,\widehat{t})$, so $(IC_g^{(0,0)})$ is not binding. ■

The lemma captures the intuition that the supervisor will learn nothing about ω from observing the bank's risk level after having sent $m=(1,1)$ or $m=(0,0)$, so it is optimal for the supervisor to recommend the bank to take a certain action regardless of its signal only if he is willing to allow that action based solely on his prior information.

Lemma 5 *Suppose that $T^{(1,0)}\subset[0,1]$ satisfies the constraint $(IC_b^{(1,0)})$. Then there exists $\tau\in[0,\underline{t}]$ such that $\mathcal{T}^{(1,0)}:=[0,\tau)\cup(T^{1,0}\setminus[0,\underline{t}))$ satisfies both $(IC_g^{(1,0)})$ and $(IC_b^{(1,0)})$.*

Proof. Let $\tau \in [0, \underline{t}]$ be the unique value such that

$$\int_{T^{(1,0)} \cap [\underline{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t) = c \int_0^{\tau} [(1-\gamma)t + \gamma(1-t)] dF(t)$$

Note that τ is well defined since the right-hand side is continuous and strictly increasing in τ, and

$$\begin{aligned} c\int_0^0 [(1-\gamma)t + \gamma(1-t)] dF(t) &\leq \int_{T^{(1,0)} \cap [\underline{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t) \\ &\leq c \int_{T^{(1,0)} \cap [0,\underline{t})} [(1-\gamma)t + \gamma(1-t)] dF(t) \leq c \int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)] dF(t) \end{aligned}$$

where the second inequality holds because $T^{(1,0)}$ satisfies the constraint ($IC_b^{(1,0)}$), while the last inequality holds because $T^{(1,0)} \cap [0, \underline{t}) \subset [0, \underline{t}]$. By construction, $\mathcal{T}^{(1,0)} := [0, \tau) \cup (T^{1,0} \setminus [0, \underline{t}))$ satisfies ($IC_b^{(1,0)}$) with equality.

We now show that $\mathcal{T}^{(1,0)}$ satisfies ($IC_g^{(1,0)}$). To this end, first observe that

$$\begin{aligned} \int_{\mathcal{T}^{(1,0)} \cap [\underline{t},1]} [(1-\gamma)\underline{t} + \gamma(1-\underline{t})] dF(t) &\geq \int_{\mathcal{T}^{(1,0)} \cap [\underline{t},1]} [(1-\gamma)t + \gamma(1-t)] dF(t) \\ &= c \int_0^{\tau} [(1-\gamma)t + \gamma(1-t)] dF(t) \geq c \int_0^{\tau} [(1-\gamma)\underline{t} + \gamma(1-\underline{t})] dF(t) \end{aligned}$$

where the two inequalities hold because $(1-\gamma)t + \gamma(1-t) = \gamma - (2\gamma - 1)t$ is strictly decreasing in t for $\gamma > \frac{1}{2}$ and because $\tau \leq \underline{t}$, while the equality holds by construction. Therefore, we have

$$\int_{\mathcal{T}^{(1,0)} \cap [\underline{t},1]} dF(t) \geq c \int_0^{\tau} dF(t)$$

In addition, observe that

$$\begin{aligned} \int_{\mathcal{T}^{(1,0)} \cap [\underline{t},1]} [\gamma t + (1-\gamma)(1-t)] dF(t) &\geq \int_{\mathcal{T}^{(1,0)} \cap [\underline{t},1]} [\gamma \underline{t} + (1-\gamma)(1-\underline{t})] dF(t) \\ c \int_0^{\tau} [\gamma \underline{t} + (1-\gamma)(1-\underline{t})] dF(t) &\geq c \int_0^{\tau} [\gamma t + (1-\gamma)(1-t)] dF(t) \end{aligned}$$

because $\gamma t + (1-\gamma)(1-t) = (2\gamma - 1)t + 1 - \gamma$ is strictly increasing in t for $\gamma > \frac{1}{2}$. Combined with the inequality above, it follows that ($IC_g^{(1,0)}$) holds. ■

Given that the bank is aggressive after receiving the message $m = (1,0)$ only when its signal was good, the supervisor is willing to allow aggressive risk-taking if and only if his type is $t \geq \underline{t}$. Were the supervisor to send $m = (1,0)$ only when $t \geq \underline{t}$, however, the bank would prefer to deviate to be aggressive regardless of its signal after receiving that message. Hence, the supervisor needs to commit to sending $m = (1,0)$ sometimes when $t < \underline{t}$, in which case he will object to the bank's risk management practice if it is aggressive.

By pooling his types that would accept aggressive risk-taking (i.e., $T^{(1,0)}_{\text{Allow}} := T^{(1,0)} \cap [\underline{t}, 1] \neq \emptyset$) with those that would object (i.e., $T^{(1,0)}_{\text{Object}} := T^{(1,0)} \cap [0, \underline{t}) \neq \emptyset$) in sending the message $m = (1,0)$, the supervisor induces the bank to act on its own private information, discouraging (encouraging) the bank to be aggressive if its signal was bad (good). The main thrust of Lemma 5 is to show that the most efficient way to maintain the bank's incentives for different realizations of its signal is pooling $T^{(1,0)}_{\text{Allow}}$ with the extreme types of the supervisor that would object, i.e., those that are the most sure of the bad state of the economy, so $T^{(1,0)}_{\text{Object}} = [0, \tau)$ for some $\tau \in [0, \underline{t}]$.

The above lemmas establish that none of the IC constraints, except for $(IC_b^{(1,0)})$, are binding, so the supervisor's problem reduces to the following:

$$\max_{T^{(1,0)}_{\text{Allow}} \subset [\underline{t},1], T^{(1,1)} \subset [\widehat{t},1], \tau \in [0,\underline{t}]} \int_{T^{(1,0)}_{\text{Allow}}} [\gamma t - (1-\gamma)(1-t)d] dF(t) + \int_{T^{(1,1)}} [t - (1-t)d] dF(t) \tag{6}$$

subject to $T^{(1,0)}_{\text{Allow}} \cap T^{(1,1)} = \emptyset$ and

$$\int_{T^{(1,0)}_{\text{Allow}}} [(1-\gamma)t + \gamma(1-t)] dF(t) \leq c \int_0^{\tau} [(1-\gamma)t + \gamma(1-t)] dF(t) \tag{7}$$

4.2 Unconstrained Optimum

If we ignore the constraint (7) for a moment, that the optimal solution to (6) involves $T^{(1,0)}_{\text{Allow}} = [\underline{t}, \bar{t})$ and $T^{(1,1)} = [\bar{t}, 1]$ is immediately clear from three facts: (i) $\gamma t - (1-\gamma)(1-t)d \geq 0$ for all $t \geq \underline{t}$, (ii) $t - (1-t)d \geq 0$ for all $t \geq \widehat{t}$, and (iii)

$$\gamma t - (1-\gamma)(1-t)d \geq t - (1-t)d$$

if and only if $t \geq \bar{t}$. This is intuitive and expected. On the one hand, the supervisor of type $t \in [\underline{t}, \bar{t})$ is a priori sufficiently uncertain about the state of the economy such that he would be influenced by the bank's signal in determining his supervisory action, so the supervisor prefers the bank to act on its own private information if he could dictate its behavior, i.e., he sends the message $m = (1,0)$. On the other hand, the supervisor of type $t \geq \bar{t}$ is a priori sufficiently confident that the state is good such that he would allow aggressive risk-taking regardless of the bank's signal, so the supervisor prefers the bank to be aggressive regardless of its own private information if he could dictate its behavior, i.e., he sends the message $m = (1,1)$.

The following proposition summarizes the above discussion, as well as providing the necessary and sufficient condition for the supervisor's unconstrained optimum to be feasible.

Proposition 1 *The unconstrained optimal solution to (6) involves $T^{(1,0)}_{Allow} = [\underline{t}, \bar{t})$ and $T^{(1,1)} = [\bar{t}, 1]$ (so $T^{(1,0)} = [0,\tau) \cup [\underline{t}, \bar{t})$ for some $\tau \in [0, \underline{t}]$). The supervisor's unconstrained optimum is feasible, or IC, if and only if*

$$\int_{\underline{t}}^{\bar{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) \leq c \int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) \tag{8}$$

Proof. If (8) holds, then there exists $\tau \in [0, \underline{t}]$ such that

$$\int_{\underline{t}}^{\bar{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) = c \int_0^{\tau} [(1-\gamma)t + \gamma(1-t)]dF(t)$$

so $T^{(1,0)}_{\text{Allow}} = [\underline{t}, \bar{t})$ and τ satisfy the incentive constraint (7). Similarly, if $T^{(1,0)}_{\text{Allow}} = [\underline{t}, \bar{t})$ satisfies (7) for some $\tau \in [0, \underline{t}]$, (8) holds because $\int_0^{\tau}[(1-\gamma)t+\gamma(1-t)]dF(t) \leq \int_0^{\underline{t}}[(1-\gamma)t+\gamma(1-t)]dF(t)$. ■

4.3 Constrained Solution

Now suppose that the supervisor's unconstrained optimum is *not* feasible, i.e., (8) does *not* hold.

As a preliminary, note that it is optimal to set $\tau = \underline{t}$ from the fact that the incentive constraint (7) is binding. In this case, the supervisor would like to, but cannot, send the message $m = (1,0)$ for every $t \in [\underline{t}, \bar{t}]$ without violating the constraint. Increasing $\tau(< \underline{t})$ relaxes (7) by raising the right-hand side of (7). However, it does not directly affect the objective function (6). Therefore, this change strictly raises the supervisor's payoff by allowing him to send the message $m = (1,0)$ for a wider range of $t \in [\underline{t}, \bar{t})$ without violating the constraint.

We now form the Lagrangian of the supervisor's problem:

$$
\begin{aligned}
\mathcal{L} = & \int_{T^{(1,0)}_{\text{Allow}}} [\gamma t - (1-\gamma)(1-t)d] dF(t) + \int_{T^{(1,1)}} [t - (1-t)d] dF(t) \\
& - \lambda \{ \int_{T^{(1,0)}_{\text{Allow}}} [(1-\gamma)t + \gamma(1-t)] dF(t) - c \int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)] dF(t) \} \\
= & \int_{T^{(1,0)}_{\text{Allow}}} \{\gamma t - (1-\gamma)(1-t)d - \lambda[(1-\gamma)t + \gamma(1-t)]\} dF(t) \\
& + \int_{T^{(1,1)}} [t - (1-t)d] dF(t) \\
& + \lambda c \int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)] dF(t)
\end{aligned}
\tag{9}
$$

where λ is the Lagrange multiplier on (7). Note that the last term in (9) affects $\mathcal{L}$ only through λ and so can be regarded as a constant.

Let $\tau_*(\lambda)$ be the unique value of t such that

$$\gamma t - (1-\gamma)(1-t)d - \lambda[(1-\gamma)t + \gamma(1-t)] = 0 \iff \tau_*(\lambda) := \frac{(1-\gamma)d + \lambda\gamma}{\gamma + (1-\gamma)d + \lambda(2\gamma - 1)}.$$

It is easy to see that $\tau_*(\lambda)$ is strictly increasing in γ from $\underline{t}$ to $\frac{\gamma}{2\gamma-1}$ as λ rises from 0 to ∞. In addition, let $\tau^*(\lambda)$ be the unique value of t such that

$$\gamma t-(1-\gamma)(1-t)d-\lambda[(1-\gamma)t+\gamma(1-t)] = t-(1-t)d \iff \tau^*(\lambda) := \frac{\gamma d - \lambda\gamma}{1-\gamma+\gamma d - \lambda(2\gamma-1)}.$$

It is easy to see that $\tau^*(\lambda)$ is strictly decreasing in γ from $\bar{t}$ to $-\infty$ as λ rises from 0 to $\frac{1-\gamma+\gamma d}{2\gamma-1}$. Finally, let Λ be the unique value of λ such that

$$\tau^*(\Lambda) = \tau_*(\Lambda) \iff \Lambda := \frac{(2\gamma-1)d}{\gamma + (1-\gamma)d}.$$

It is easy to verify that $\tau^*(\Lambda) = \tau_*(\Lambda) = \widehat{t}$.

Given the discussion above, $\lambda \in [0, \Lambda]$ guarantees that $\underline{t} \leq \tau_*(\lambda) \leq \tau^*(\lambda) \leq \bar{t}$. Thus, $[\tau_*(\lambda), \tau^*(\lambda))$ can be interpreted as the optimal subinterval of $[\underline{t}, \bar{t})$ for which the supervisor chooses to send the message $m = (1, 0)$ taking as given the marginal cost λ of violating the constraint (7) from the fact that, for every $t \in [\tau_*(\lambda), \tau^*(\lambda))$, $\gamma t - (1-\gamma)(1-t)d - \lambda[(1-\gamma)t + \gamma(1-t)] \geq 0$ and $\gamma t - (1-\gamma)(1-t)d - \lambda[(1-\gamma)t + \gamma(1-t)] > t - (1-t)d$. We further note that increasing λ would shrink $[\tau_*(\lambda), \tau^*(\lambda))$, because the supervisor cannot send $m = (1,0)$ as often as he used to when $t \in [\underline{t}, \overline{t})$ without violating the constraint (7), and $[\tau_*(\lambda), \tau^*(\lambda))$ eventually collapses to $\widehat{t}$. This is because the supervisor is the most uncertain about the optimal supervisory action when he is of type $t = \widehat{t}$, which is why he is indifferent between allowing and objecting to aggressive risk-taking, so the supervisor would optimally choose to send $m = (1,0)$ in a region concentrated around $\widehat{t}$ if he had to select a particular subinterval of $[\underline{t}, \overline{t})$.

The optimal value of $\lambda \in [0, \Lambda]$, which we denote by λ^*, is the value such that

$$\int_{\tau_*(\lambda)}^{\tau^*(\lambda)} [(1-\gamma)t + \gamma(1-t)]dF(t) = c\int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) \tag{10}$$

Note that λ^* is well defined since the left-hand side is continuous and strictly decreasing in λ, and

$$\begin{aligned}\int_{\tau_*(0)}^{\tau^*(0)} [(1-\gamma)t+\gamma(1-t)]dF(t) &= \int_{\underline{t}}^{\overline{t}} [(1-\gamma)t+\gamma(1-t)]dF(t) > c\int_0^{\underline{t}} [(1-\gamma)t+\gamma(1-t)]dF(t) \\ &\geq \int_{\widehat{t}}^{\widehat{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) = \int_{\tau_*(\Lambda)}^{\tau^*(\Lambda)} [(1-\gamma)t + \gamma(1-t)]dF(t)\end{aligned}$$

where the first inequality holds because (8) does *not* hold.

The following proposition summarizes the preceding discussion.

Proposition 2 *Suppose that (8) does not hold. Then the IC constraint (7) is binding (i.e., it holds with equality), and the constrained optimal solution to (6) involves* $T^{(1,0)}_{\mathit{Allow}} = [\tau_*(\lambda^*), \tau^*(\lambda^*))$, $T^{(1,1)} = [\tau^*(\lambda^*), 1]$, *where* $\tau_*(\lambda) \in [\underline{t}, \widehat{t}]$, $\tau^*(\lambda) \in [\widehat{t}, \overline{t}]$, *and* $\lambda^* \in [0, \Lambda]$ *is the solution to (10) (so* $T^{(1,0)} = [0, \underline{t}) \cup [\tau_*(\lambda^*), \tau^*(\lambda^*))$*).*

4.4 Comparative Statics

The following proposition, which is our main result in this paper, stresses that the informativeness of the bank's signal has a non-monotonic effect on welfare.

Proposition 3 *If* γ *is sufficiently close to* $\frac{1}{2}$*, then the supervisor's expected payoff increases in* γ*. For* γ *sufficiently close to* 1*, the supervisor's expected payoff decreases in* γ*.*

Proof. First, we show that there exists $\overline{\gamma} > \frac{1}{2}$ such that (8) holds if $\gamma < \overline{\gamma}$, which follows from

$$\lim_{\gamma \to \frac{1}{2}} \int_{\underline{t}}^{\overline{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) = 0$$

$$< c \int_0^{\widehat{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) = \lim_{\gamma \to \frac{1}{2}} c \int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)]dF(t)$$

where the two equalities hold because $\lim_{\gamma \to \frac{1}{2}} \underline{t} = \lim_{\gamma \to \frac{1}{2}} \overline{t} = \widehat{t}$. By Proposition 1, whenever $\gamma < \overline{\gamma}$, the supervisor's unconstrained optimum is feasible, so he obtains an expected payoff of $\int_{\underline{t}}^{\overline{t}} [\gamma t - (1-\gamma)(1-t)d]dF(t) + \int_{\overline{t}}^{1} [t - (1-t)d]dF(t)$, which can be easily shown to be strictly increasing in γ. The first part of the lemma follows.

Similarly, there exists $\gamma_* < 1$ such that (8) does *not* hold if $\gamma > \gamma_*$, which follows from

$$\lim_{\gamma \to 1} \int_{\underline{t}}^{\overline{t}} [(1-\gamma)t + \gamma(1-t)]dF(t) = \lim_{\gamma \to \frac{1}{2}} \int_0^1 [(1-\gamma)t + \gamma(1-t)]dF(t)$$

$$> 0 = \lim_{\gamma \to 1} c \int_0^{\underline{t}} [(1-\gamma)t + \gamma(1-t)]dF(t)$$

where the two equalities hold because $\lim_{\gamma \to 1} \underline{t} = 0$ and $\lim_{\gamma \to 1} \overline{t} = 1$. By Proposition 2, whenever $\gamma > \gamma_*$, the IC constraint is binding, so the pair (τ_*, τ^*) is uniquely determined by (10). Since the right-hand side approaches 0 as γ tends to 1, both τ^* and τ_* converge to $\widehat{t}$. Therefore, the supervisor's expected payoff approaches $\int_{\widehat{t}}^{1} [t - (1-t)d]dF(t)$ as $\gamma \to 1$. Finally, note that the supervisor's indirect utility as a function of γ is minimized at $\gamma = 1$, which concludes the proof. ■

Proposition 3 accentuates that an increase in γ has two distinct effects, the information effect and control dilution. The former refers to the fact that an increase in γ (the informativeness of the bank's signal) enables the supervisor to make more informed supervisory decisions when he can induce the bank to reveal its signal and, therefore, improves welfare, i.e., increases the supervisor's expected payoff. If the supervisor is of type $t \in [\underline{t}, \overline{t})$, then his expected payoff attainable with the bank acting on its signal rises from an increase in γ because it implies a lower probability of the bank erring. An increase in γ has another effect of expanding the *ideal* information-acquisition region $[\underline{t}, \overline{t})$, so the supervisor would benefit from observing the bank's signal in determining the supervisory action for a wider range of priors about the state of the economy.

Of course, the potential welfare improvement is only attainable insofar as the supervisor can induce ideal behavior in the bank by satisfying the incentive constraint (7). When the bank's signal becomes more informative, this task becomes more challenging, because the supervisor relies less on his prior beliefs. In particular, an increase in γ contracts the default-objection region $[0, \underline{t})$, which reduces the probability that the bank thinks the supervisor will object to its aggressive risk-taking. When its perceived probability is sufficiently low, the bank has strong incentives to be aggressive regardless of its signal, restraining which forces the supervisor to shrinking the *actual* information-acquisition region $[\tau_*, \tau^*)$. This immediately implies that the bank reveals information less frequently and, therefore, welfare deteriorates, i.e., the supervisor's expected payoff decreases. We refer to this effect as control dilution.

To see this clearly, suppose that γ is so close to 1 that $[0, \underline{t})$ is vanishingly small. Were the supervisor to send $m = (1, 0)$ over some strictly positive-length subinterval of $[\underline{t}, \bar{t})$ (which is feasible because $[\underline{t}, \bar{t})$ expands to fill the entire unit interval), the bank can infer that he is almost surely of type $t \geq \underline{t}$, so it would prefer to deviate to be aggressive regardless of its signal after receiving that message. Hence, the supervisor needs to commit to sending $m = (1, 0)$ over a vanishingly small subinterval of $[\underline{t}, \bar{t})$ for him to have plausible deniability. In essence, an increase in γ (the informativeness of the bank's signal) implies the supervisor's greater eagerness to persuade the bank to act on its signal, which ironically can compromise his ability to do so.

The preceding discussion indicates that whether the IC constraint for the bank is binding or not plays a key role in determining which of the two opposing effects of γ (the information effect and control dilution) dominates. This explains why, in Proposition 3, the information effect dominates when γ is relatively small, while the control-dilution effect dominates when γ is relatively large. If γ is small, then the supervisor mostly relies on his prior belief to make the supervisory decision, so the IC constraint for the bank is not binding and the information effect dominates, improving welfare as γ increases. If γ is large, then the supervisor is already dependent enough on the bank's private information so that the IC constraint for the bank is binding and the control-dilution effect dominates, reducing welfare as γ increases.

4.4.1 An Example

To help visualize Proposition 3 and the mechanism behind this result, suppose the supervisor's type t is uniform on $[0, 1]$. Figure 1 illustrates how the supervisor's equilibrium expected payoff depends on γ (the informativeness of the bank's signal).

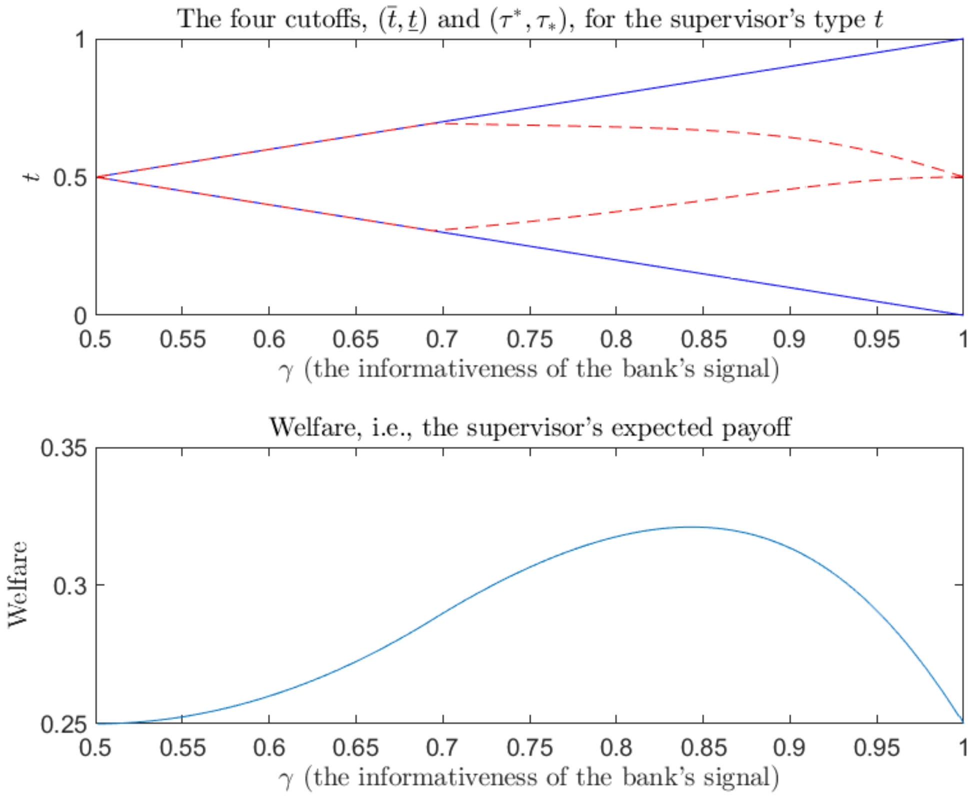


Figure 1: Optimal values of the key quantities as functions of the informativeness parameter γ assuming the supervisor's type t is uniform on $[0, 1]$. The parameter values used for this figure are $c = 1$ and $d = 1$.

The top panel of Figure 1 shows that $\bar{t}$ (the upper portion of the blue solid curve) is upward-sloping, while $\underline{t}$ (the lower portion of the blue solid curve) is downward-sloping, consistent with an increase in γ expanding the *ideal* information-acquisition region $[\underline{t}, \bar{t})$ as explained above. The top panel shows another interesting feature: τ^* (the upper portion of the red dashed curve) is initially upward-sloping but eventually downward-sloping, while τ_* (the lower portion of the red dashed curve) is initially downward-sloping but eventually upward-sloping, ergo an increase in γ is initially expanding but eventually shrinking the *actual* information-acquisition region $[\tau_*, \tau^*)$. As γ increases, notice that the default-objection region $[0, \underline{t})$ contracts, which makes it more challenging to satisfy the incentive constraint (7). This forces the supervisor to eventually shrinking the actual information-acquisition region despite the ever expanding ideal information-acquisition region.

The bottom panel of Figure 1 plots welfare, i.e., the supervisor's expected payoff, as γ increases. It starts at 0.25 utils when $\gamma = 0.5$, peaks at 0.321 utils when $\gamma = 0.844$, and finishes at 0.25 utils when $\gamma = 1$. On the one hand, an increase in γ initially expands (eventually shrinks) the actual information-acquisition region $[\tau_*, \tau^*)$, which in turn implies that the bank reveals information more (less) frequently and, therefore, welfare improves (deteriorates). On the other hand, an increase in γ has another effect of enabling the supervisor to make more informed supervisory decisions within the actual information-acquisition region and, therefore, welfare continues to improve beyond the point at which $[\tau_*, \tau^*)$ starts to shrink.

4.4.2 Comparison to the Cheap-Talk Equilibrium

Our model gives the supervisor the commitment power over his choice of a message as a function of his type. To investigate the welfare benefits of this power, it is instructive to compare our results to cheap-talk equilibria.

As usual, there are multiple equilibria when the supervisor communicates with the bank via cheap talk. First, babbling, which means no information is transmitted, is always an equilibrium outcome. A more efficient cheap-talk equilibrium, in which some information is communicated, can be immediately constructed, as formally stated in the following proposition.

Proposition 4 *There always exists a cheap-talk equilibrium in which the supervisor sends the message $m = (1, 1)$ if he is of type $t \geq \widehat{t}$ and $m = (0, 0)$ otherwise; the bank is aggressive if and only if it receives message $m = (1, 1)$, in which case the supervisor will allow it to be aggressive.*

Proof. Recall that, based solely on his private information, the supervisor will allow aggressive risk-taking if and only if $t \geq \widehat{t}$. Given that the bank knows that $t \geq \widehat{t}$ ($t < \widehat{t}$) conditional on observing $m = (1, 1)$ ($m = (0, 0)$), it is straightforward that the optimal action is to be aggressive (conservative) regardless of its signal.[5] Given what he expects the bank to do, the supervisor of type $t \geq \widehat{t}$ ($t < \widehat{t}$) prefers ex ante aggressive (conservative) risk-taking, ergo sending $m = (1, 1)$ ($m = (0, 0)$) as the bank expects him to do. ■

Notice that in the limit as $\gamma \to \frac{1}{2}$ or $\gamma \to 1$, the supervisor's expected payoff shrinks to his expected payoff in the cheap-talk equilibrium presented in Proposition 4. In the former case, the

[5]The bank can be aggressive when it is not supposed to be (after observing $m = (0, 0)$) off the equilibrium path. We restrict attention to equilibria in which the supervisor's off-equilibrium-path belief does not change from his type (i.e., he expects the state to be good with probability t).

supervisor chooses not to induce the bank to act on its own information ($[\underline{t}, \bar{t})$ shrinks toward $\{\widehat{t}\}$), while in the latter case, he cannot induce the bank to act on its own information ($[\tau_*, \tau^*)$ shrinks toward $\{\widehat{t}\}$). Compared to this cheap-talk equilibrium, commitment power on the supervisor's side improves welfare as long as the bank has some but not perfect information about the state.

In fact, however, the supervisor can do better than the equilibrium presented in Proposition 4 even when he can send only a cheap-talk message. Suppose that (8) holds, so the supervisor's unconstrained optimum is feasible with commitment power on the supervisor's side. In this case, the supervisor's unconstrained optimum is feasible even when he can only engage in cheap talk precisely because it can implement the supervisor's ideal outcome for every one of his type: while the bank could be aggressive for some $t < \underline{t}$, the supervisor can *costlessly* object to this, forcing the bank to become conservative.

In contrast, if (8) does *not* hold, the informative cheap-talk equilibrium presented in Proposition 4 is unique. In this case, the supervisor cannot incentivize the bank to act on its information for all $t \in [\underline{t}, \bar{t})$, so he ought to send $m \neq (1,0)$ for some of these types. Without commitment power, the supervisor of such types would prefer to deviate ex post, sending the message $m = (1,0)$. For example, the constrained optimal solution involves the supervisor committing to send the message $m = (1,1)$ if $t \geq \tau^*$ and $m = (0,0)$ if $t \in [\underline{t}, \tau_*)$; without this commitment, the supervisor would prefer to deviate ex post if he turns out to be of type $t \in [\underline{t}, \tau_*) \cup [\tau^*, \bar{t})$ by sending $m = (1,0)$.

The following proposition summarizes the preceding discussion.

Proposition 5 *Suppose that (8) holds. Then the supervisor's unconstrained optimum is a cheap-talk equilibrium: the supervisor sends the message $m = (1,1)$ if he is of type $t \geq \bar{t}$, $m = (1,0)$ if he is of type $t \in [0, \tau) \cup [\underline{t}, \bar{t})$ for some $\tau \in [0, \underline{t}]$, and $m = (0,0)$ otherwise.*

In light of this proposition, we conclude that a sufficient condition for the supervisor's commitment power to improve welfare is that $\gamma \in (\gamma_*, 1)$, where $\gamma_* \in (\frac{1}{2}, 1)$ such that $\gamma > \gamma_*$ implies (8) does not hold:[6] $\gamma > \gamma_*$ ensures that the supervisor is actually using the commitment power vested in him, while $\gamma < 1$ ensures that the supervisor has plausible deniability he will not always allow aggressive risk-taking even if he is of type $t \in [\underline{t}, \bar{t})$. As depicted in Figure 1, welfare improvements from commitment power is non-monotonic in γ: welfare is increasing at $\gamma = \gamma_*$ (the point at which $[\tau_*, \tau^*)$ starts to shrink) before it eventually shrinks to the level of welfare in the unique informative cheap-talk equilibrium in the limit as $\gamma \to 1$.

[6] We show that such γ_* does exist in the proof of Proposition 3.

4.5 Bank's Cost in Case of Supervisory Objection

One important lesson from our analysis is that whether the IC constraint for the bank is binding or not plays a key role in determining the welfare implications of more information on the bank's side. An increase in γ (the informativeness of the bank's signal) enables the supervisor to make more informed supervisory decisions and, therefore, improves welfare primarily when he can induce the bank to reveal its signal, i.e., the IC constraint for the bank is *not* binding. Looking at (8), it is immediate that increasing c (the bank's cost in case of supervisory objection) relaxes the IC constraint, ergo improving welfare until the IC constraint is no longer binding.

Proposition 6 *Let $v(c)$ denote the supervisor's maximal attainable payoff when the bank is faced with a cost c in case of supervisory objection. Then $v(c)$ is strictly increasing for $c \in [0, c^*)$ and is equal to $v(c^*)$ for $c \geq c^*$, where $c^* > 0$ is the value of c such that (8) holds with equality.*

Intuitively, the bank is worried not only about how frequently the supervisor will object to its aggressive risk-taking ($\int_0^{\underline{t}}[(1-\gamma)t + \gamma(1-t)]dF(t)$), but also how costly those supervisory objections will be (c), so increasing c can offset the control-dilution effect of increased γ. Our analysis taking c as given reflects the fact that the supervisor can be agile in his communication strategy, but he cannot freely adjust the bank's cost in case of supervisory objection. Yet the supervisor does have the power to occasionally change such costs for the bank by passing legislation to promote financial stability. For example, the Dodd-Frank Act made all banks with assets above $50 billion subject to a much more aggressive supervisory regime, effectively raising c for mid-sized banks; in 2018, Congress scaled back Dodd-Frank, raising the threshold for increased scrutiny of banks from $50 billion to $250 billion, effectively reducing c for mid-sized banks.

To the extent that the supervisor has some control over the bank's cost, Proposition 6 has an important policy implication. It is optimal to err on the side of giving the supervisor too much power *in case he finds that the bank does not meet supervisory expectations*, i.e., err on the side of setting c too high. If c is too high (i.e., $c > c^*$), the supervisor could simply scale back how frequently he will object to aggressive risk-taking after having sent $m = (1, 0)$—$T^{(1,0)}_{\text{Object}} = [0, \tau)$ for some $\tau \in [0, \underline{t}]$ that is decreasing in c.[7] If c is too low (i.e., $c < c^*$), not only is the supervisor's unconstrained optimum infeasible (leaving welfare on the table), but the economy is exposed to experiencing a welfare loss in case the bank experiences a sudden boost in its private information.

[7] We show that such τ does exist in the proof of Proposition 1.

5 Additional Commitment to the Supervisory Ruling

Our baseline model does not give the supervisor commitment power over his follow-up supervisory ruling. In particular, the supervisor allows aggressive risk-taking only when it is ex post efficient: he will allow the bank's aggressive risk-taking if and only if he is of type $t \geq \underline{t}$ ($t \geq \widehat{t}$) after having sent $m = (1, 0)$ ($m = (1, 1)$). We now turn attention to the case where the supervisor also has commitment power over his follow-up supervisory ruling. Specifically, the supervisor can commit a priori to allowing (objecting to) aggressive risk-taking even if it is ex post inefficient: he will object to the bank's aggressive risk-taking, for example, if he is of type $t \in T^{(1,0)}_{\text{Object}} \cap [\underline{t}, \overline{t})$ although he prefers ex post to allow it. Importantly, we will find that this additional commitment enables the supervisor to largely offset the negative welfare effect of more information on the bank's side such that welfare is now monotonically increasing with the informativeness of the bank's signal.

Following similar steps as in the baseline model, it is straightforward to show that it is still the case that $T^{(1,1)} = T^{(1,1)}_{\text{Allow}} \subset [\widehat{t}, 1]$ and $T^{(0,0)} = T^{(0,0)}_{\text{Object}} \subset [0, \widehat{t})$ (and $T^{(1,1)}_{\text{Object}} = T^{(1,1)}_{\text{Allow}} = \emptyset$). It is also straightforward to show that it is still the case that $T^{(1,0)} = [0, \tau) \cup T^{(1,0)}_{\text{Allow}}$, where $T^{(1,0)}_{\text{Object}} = [0, \tau)$ and $T^{(1,0)}_{\text{Allow}} \cap T^{(1,0)}_{\text{Object}} = \emptyset$. The most crucial difference from the baseline case is that τ can exceed $\underline{t}$.

Again, the supervisor's problem reduces to the following:

$$\max_{T^{(1,0)}_{\text{Allow}}, T^{(1,1)} \subset [\widehat{t}, 1], \tau \in [0, \widehat{t}]} \int_{T^{(1,0)}_{\text{Allow}}} [\gamma t - (1-\gamma)(1-t)d] dF(t) + \int_{T^{(1,1)}} [t - (1-t)d] dF(t) \tag{11}$$

subject to $T^{(1,0)}_{\text{Allow}} \cap T^{(1,1)} = T^{(1,0)}_{\text{Allow}} \cap [0, \tau) = \emptyset$ and

$$\int_{T^{(1,0)}_{\text{Allow}}} [(1-\gamma)t + \gamma(1-t)] dF(t) \leq c \int_0^{\tau} [(1-\gamma)t + \gamma(1-t)] dF(t) \tag{12}$$

Clearly, if the supervisor's unconstrained optimum is feasible in the baseline model, it is also feasible with this additional commitment power. Given this observation, it suffices to consider the case where (8) does *not* hold. As in the baseline model, it continues to hold that $T^{(1,0)}_{\text{Allow}} = [\tau_{**}, \tau^{**})$ for some $\tau_{**} \in (\underline{t}, \widehat{t})$ and $\tau^{**} \in (\widehat{t}, \overline{t})$. In contrast to our baseline model, it is straightforward to prove that $\tau = \tau_{**}$—$T^{(1,0)}_{\text{Object}} = [0, \tau_{**})$ and $T^{(1,0)} = [0, \tau^{**})$. Recall that, in the baseline model, $T^{(1,0)}_{\text{Object}} = [0, \underline{t})$ and $T^{(0,0)} = [\underline{t}, \tau_*)$. Intuitively, the supervisor of type $t \in [\underline{t}, \tau_*)$ is tempted to respect the bank's decisions if they were reflective of its signal: sending $m = (1, 0)$ in this region would make the IC constraint even more binding, so he resorted to sending $m = (0, 0)$ instead.

Now, the supervisor is able to put this region to good use with the additional commitment power vested in him. The supervisor can overcome the temptation to respect the bank's risk-taking decision if t turns out to be in $[\underline{t}, \tau_*)$ by committing to object to aggressive risk-taking in this region even after having sent $m = (1, 0)$. That is, the bank will still end up with conservative risk-taking when $t \in [\underline{t}, \tau_*)$, but now, $[\underline{t}, \tau_*)$ can be annexed to the default-objection region $[0, \underline{t})$, which raises the probability that the supervisor will object to aggressive risk-taking after having sent $m = (1, 0)$. Increasing τ beyond $\underline{t}$ to τ_* relaxes (8) by raising the right-hand side of (8). This immediately implies the supervisor can expand the actual information-acquisition region $[\tau_*, \tau^*)$ beyond what is feasible in the baseline model, helping to fight off control dilution.

The following lemma summarizes the preceding discussion.

Lemma 6 *Suppose that (8) does not hold. Then the IC constraint (12) is binding (i.e., it holds with equality), and the optimal solution to (11) involves* $T^{(1,0)}_{Allow} = [\tau_{**}, \tau^{**})$, $T^{(1,0)}_{Object} = [0, \tau_{**})$, *where* $\tau_{**} \in (\underline{t}, \widehat{t})$, $\tau^{**} \in (\widehat{t}, \overline{t})$ *(so* $T^{(1,0)} = [0, \tau^{**})$ *and* $T^{(0,0)} = \emptyset$*).*

This lemma allows us to simplify the supervisor's problem even further to the following:

$$\max_{\tau_{**} \in (\underline{t}, \widehat{t}), \tau^{**} \in (\widehat{t}, \overline{t})} \int_{\tau_{**}}^{\tau^{**}} [\gamma t - (1-\gamma)(1-t)d] dF(t) + \int_{\tau^{**}}^{1} [t - (1-t)d] dF(t) \tag{13}$$

subject to

$$\int_{\tau_{**}}^{\tau^{**}} [(1-\gamma)t + \gamma(1-t)] dF(t) = c \int_{0}^{\tau_{**}} [(1-\gamma)t + \gamma(1-t)] dF(t) \tag{14}$$

Note that the only difference from our baseline model is that the integral on the right-hand side is taken over $[0, \tau_{**}]$, not over $[0, \underline{t}]$. For any given set of parameters (c, d) and initial beliefs of the supervisor (F), the right-hand side of (14) is larger than that of (10). This implies that $[\tau_{**}, \tau^{**})$ can be larger than $[\tau_*, \tau^*)$, which is the case when λ is smaller (equivalently, less binding IC constraint). This is precisely how commitment power to the supervisory ruling can benefit the supervisor.

In fact, this additional commitment suffices to ensure that welfare is monotonically increasing with the informativeness of the bank's signal.

Proposition 7 *The supervisor's expected payoff with additional commitment is strictly monotone-increasing in* γ *on the interval* $(\frac{1}{2}, 1)$.

Proof. Fix $\gamma > \frac{1}{2}$, and consider the optimal solution (τ_{**}, τ^{**}) to (13). This solution is feasible for any $\gamma' > \gamma$ if and only if $\int_{\tau_{**}}^{\tau^{**}} [(1-\gamma')t + \gamma'(1-t)]dF(t) \leq c \int_0^{\tau_{**}} [(1-\gamma')t + \gamma'(1-t)]dF(t)$, which holds if and only if

$$\int_{\tau_{**}}^{\tau^{**}} (1-2t)dF(t) \leq c \int_0^{\tau_{**}} (1-2t)dF(t)$$

because $\gamma' > \gamma$ and (14) is satisfied. Rearranging the terms in (14) to obtain an expression for c, and plugging this into the inequality above, tedious but straightforward simplification shows that the inequality above is equivalent to

$$\int_0^{\tau_*} tdF\left(t \,|t \in [0, \tau_{**})\right) \leq \int_{\tau_*}^{\tau^*} tdF\left(t \,|t \in [\tau_{**}, \tau^{**})\right),$$

which is obviously true. Therefore, (τ_{**}, τ^{**}) is still a feasible solution for any $\gamma' > \gamma$.

It is easy to check that this solution gives a strictly higher expected payoff to the supervisor when γ increases to $\gamma' > \gamma$. Therefore, the supervisor's optimal expected payoff must be strictly monotone-increasing in γ on the interval $(\frac{1}{2}, 1)$. ■

Proposition 7 shows that, unlike in our baseline model, the supervisor can do strictly better than with cheap talk even in the limit as $\gamma \to 1$. At the same time, note that Lemma 6 implies that it still is the case that the supervisor cannot attain his unconstrained optimum. But more importantly, the proposition shows that, with additional commitment to the supervisory ruling, more information does result in higher welfare. The policy implication is, thus, that it is important to give the supervisor enough commitment power, particularly in the form of supervisory ruling. Commitment power over how much he discloses about his own information alone can be impotent in case the bank experiences a sudden boost in its private information.

5.1 An Example

To help visualize Proposition 7, suppose the supervisor's type t is uniform on $[0, 1]$. Figure 2 illustrates how the supervisor's equilibrium expected payoff in the case of additional commitment over the final ruling depends on γ (the informativeness of the bank's signal) in the bottom panel, as well as how the actual information-acquisition region $[\tau_{**}, \tau^{**})$ in this case depends on γ in the top panel. We plot all of these quantities with magenta dash-dotted lines. In the figure, for clearer comparison, we also plot the corresponding quantities in the baseline model with red dashed lines and the corresponding quantities in the unconstrained optimum with blue solid lines.

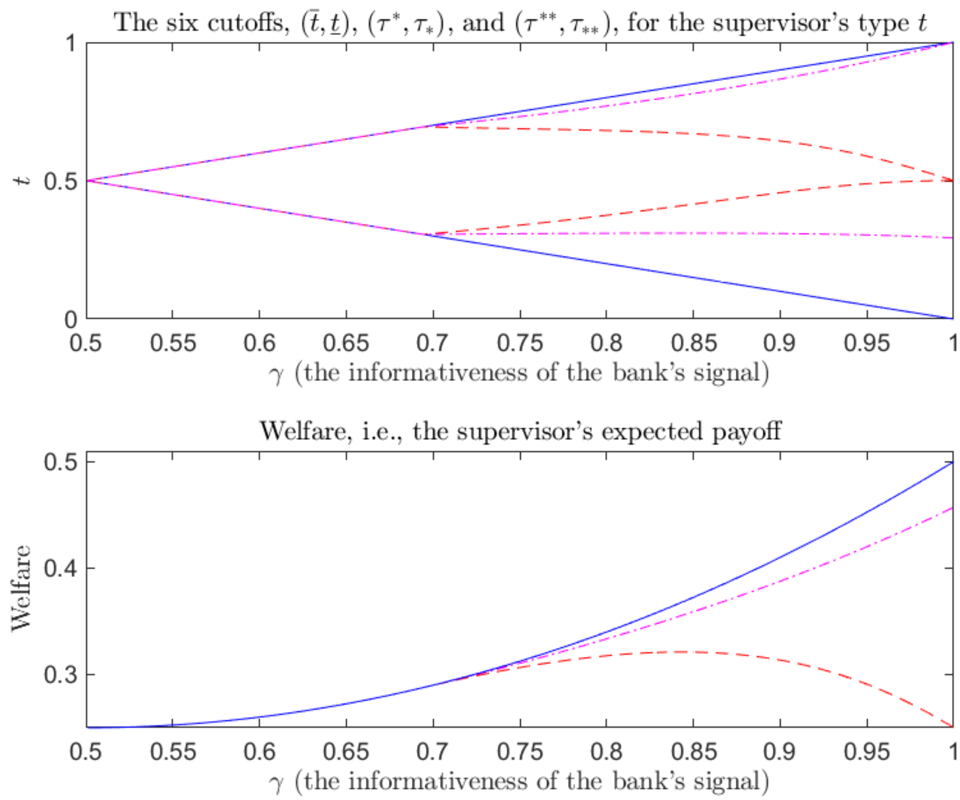


Figure 2: Optimal values of the key quantities when the supervisor also has commitment power over his final ruling (magenta dash-dotted lines) as functions of the informativeness parameter γ assuming the supervisor's type t is uniform on $[0, 1]$. The parameter values used for this figure are $c = 1$ and $d = 1$.

As discussed above, Figure 2 confirms that, unlike in the baseline model, the supervisor can do strictly better than with cheap talk even in the limit as $\gamma \to 1$. While it still is the case that the supervisor cannot attain his unconstrained optimum, the figure shows that the supervisor can do surprisingly well even in the limit as $\gamma \to 1$—the welfare gap from the unconstrained optimum is visibly small. In fact, it is easy to check that, as shown in the figure, $\tau^{**} \to 1$ in this limit, so the supervisor can induce the bank to reveal its information whenever he is optimistic enough. Thus, the figure vividly reinforces the policy implication of Proposition 7: it is important to give the supervisor enough commitment power, particularly in the form of supervisory ruling.

6 Discussion

[John: Please discuss the policy implications here. Feel free to go wild as we'll tame this section together later anyway. Ideally, let's write a page or two.]
[John: Can we say anything about the Collapse of SVB?]

7 Conclusion

We characterize optimal disclosure rules for the supervisor under various information settings. The framework we present allows for informational uncertainty on both sides of the table, which is the reality of banking supervision. Supervisor's optimal messaging strategy depends on the information structure. We show that a carefully designed message is vitally important when neither the supervisor nor the bank has clear advantage in what they know. Moreover, agile banking supervision should react to not only information structure, but also the design of punishment regime. Since change in punishment regime occurs less frequently than changes in information, supervisor should reassess what she and banks know as much as it is needed. Her optimal disclosure strategy is determined by signal informativeness relative to rejection cost.

The channel of interbank information competition makes individual bank no longer pivotal. Critical number of banks is needed to convince supervisor of a good state. Supervisor therefore gives less consideration to action of any single bank. Peer discipline gives individual bank less incentive to act aggressively. Therefore, informational design can no longer correct flawed mechanism design to the extent that gets equilibrium back to first best outcome, as was the case for 1 bank. In limited scenarios (B and F), where persuasion does better than both full and no disclosure, it only achieves more separation but not full separation. For all the rest scenarios, persuasion cannot improve upon no disclosure. This channel also creates discontinuity in supervisory action given banks' actions. Thus improvement in signal precision does not monotonically induce more separation and increase welfare.

References

Alonso, Ricardo and Odilon Câmara, "Persuading voters," *American Economic Review*, 2016, *106* (11), 3590–3605. 4

Austen-Smith, David and Jeffrey S Banks, "Information aggregation, rationality, and the Condorcet jury theorem," *American political science review*, 1996, *90* (1), 34–45.

— **and Timothy J Feddersen**, "Deliberation, preference uncertainty, and voting rules," *American political science review*, 2006, *100* (2), 209–217.

Baker, Scott R, Nicholas Bloom, and Steven J Davis, "Measuring economic policy uncertainty," *The quarterly journal of economics*, 2016, *131* (4), 1593–1636. 5

Bergemann, Dirk and Stephen Morris, "The comparison of information structures in games: Bayes correlated equilibrium and individual sufficiency," *Economic Theory Center Working Paper*, 2013, (054-2013). 4

— **, Tibor Heumann, Stephen Morris, Constantine Sorokin, and Eyal Winter**, "Optimal information disclosure in auctions," 2022. 4

Blackwell, David A and Meyer A Girshick, *Theory of games and statistical decisions*, Courier Corporation, 1979.

Brocas, Isabelle and Juan D Carrillo, "Influence through ignorance," *The RAND Journal of Economics*, 2007, *38* (4), 931–947.

Crawford, Vincent P. and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 1982, *50* (6), 1431–1451.

Dogra, Keshav and Keeyoung Rhee, "Stress Tests and Model Monoculture," 2022. Working Paper.

Doraszelski, Ulrich, Dino Gerardi, and Francesco Squintani, "Communication and voting with double-sided information," *Contributions in Theoretical Economics*, 2003, *3* (1).

Dworczak, Piotr and Alessandro Pavan, "Preparing for the worst but hoping for the best: Robust (bayesian) persuasion," 2020.

Eső, Péter and Balazs Szentes, "Optimal information disclosure in auctions and the handicap auction," *The Review of Economic Studies*, 2007, *74* (3), 705–731. 4

Faria-e-Castro, Miguel, Joseba Martinez, and Thomas Philippon, "Runs versus Lemons: Information Disclosure and Fiscal Capacity," *The Review of Economic Studies*, 2017, *84* (4), 1683–1707.

Feddersen, Timothy and Wolfgang Pesendorfer, "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting," *American Political science review*, 1998, *92* (1), 23–35.

Gerardi, Dino and Leeat Yariv, "Deliberative voting," *Journal of Economic theory*, 2007, *134* (1), 317–338.

_ **and** _ , "Information acquisition in committees," *Games and Economic Behavior*, 2008, *62* (2), 436–459.

Goldstein, Itay and Yaron Leitner, "Stress tests and information disclosure," *Journal of Economic Theory*, 2018, *177*, 34–69. 4

Guo, Yingni and Eran Shmaya, "The Interval Structure of Optimal Disclosure," *Econometrica*, 2019, *87* (2), 653–675.

Hirtle, Beverly, Anna Kovner, and Matthew Plosser, "The Impact of Supervision on Bank Performance," *The Journal of Finance*, 2020, *75* (5), 2765–2808.

Inostroza, Nicolas and Alessandro Pavan, "Persuasion in Global Games with Application to Stress Testing," 2021. CEPR Discussion Paper. 4

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615. 3

Leitner, Yaron and Basil Williams, "Model Secrecy and Stress Tests," *The Journal of Finance*, 2023, *78* (2), 1055–1095. 4

_ **and Bilge Yilmaz**, "Regulating a model," *Journal of Financial Economics*, 2019, *131* (2), 251–268.

Li, Hao and Xianwen Shi, "Discriminatory information disclosure," *American Economic Review*, 2017, *107* (11), 3363–85. 4

Mathis, Jerome, "Deliberation with evidence," *American Political Science Review*, 2011, *105* (3), 516–529.

McLennan, Andrew, "Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents," *American political science review*, 1998, *92* (2), 413–418.

Milgrom, Paul and John Roberts, "Relying on the information of interested parties," *The RAND Journal of Economics*, 1986, pp. 18–32.

Persico, Nicola, "Committee design with endogenous information," *The Review of Economic Studies*, 2004, *71* (1), 165–191.

Shapiro, Joel and Jing Zeng, "Stress Testing and Bank Lending," 2022. Working Paper.

Spargoli, Fabrizio, "Bank Recapitalization and the Information Value of a Stress Test in a Crisis," 2013. Working Paper.

Taneva, Ina, "Information design," *American Economic Journal: Microeconomics*, 2019, *11* (4), 151–85. 4

A Allowing for Mixed Strategy in the Bank's Risk-Taking

Consider the following trinary signal: $M = \{-, \sigma_B, +\}$ for some $\sigma_B \in [0,1]$, where σ_B represents the probability that B chooses $a = 1$ conditional on $m = \sigma_B$ and $s_B = \ell$. (It can be shown that the case where B chooses $a = 1$ with an interior probability conditional on $s_B = h$ can never be optimal.)

The subsequent analysis proceeds as follows. First, given σ_B, we characterize the optimal signal. Then, we optimize over σ_B.

Optimal signal given σ_B**.** Let $\underline{t}(\sigma_B) \in [\underline{t}, \theta]$ be the value such that

$$\frac{\underline{t}(\sigma_B)}{1-\underline{t}(\sigma_B)} = \frac{\theta}{1-\theta}\frac{1-\gamma+\gamma\sigma_B}{\gamma+(1-\gamma)\sigma_B}.$$

It is easy to see that $\underline{t}(\sigma_B)$ increases from $\underline{t}$ to σ_B as σ_B rises from 0 to 1.

S obtains her ideal outcome given σ_B if and only if

$$(1-\kappa)\int_{\underline{t}(\sigma_B)}^{\bar{t}} [t(1-\gamma)+(1-t)\gamma]\, dF(t) \leq \kappa \int_0^{\underline{t}(\sigma_B)} [t(1-\gamma)+(1-t)\gamma]\, dF(t).$$

The left-hand side is decreasing in σ_B, while the right-hand side is increasing in σ_B. Therefore, there exists $\overline{\sigma}_B$ such that S obtains her ideal outcome given σ_B if and only if $\sigma_B \geq \overline{\sigma}_B$.

If the above condition fails then S faces the following problem (after all reductions):

$$\max \int_{\tau^*}^{1} [t\cdot 1+(1-t)(-d)]\, dF(t)+\int_{\tau_*}^{\tau^*} [t(\gamma+(1-\gamma)\sigma_B)+(1-t)(1-\gamma+\gamma\sigma_B)(-d)]\, dF(t)$$

subject to

$$(1-\kappa)\int_{\tau_*}^{\tau^*} [t(1-\gamma)+(1-t)\gamma]\, dF(t) \leq \kappa \int_0^{\underline{t}(\sigma_B)} [t(1-\gamma)+(1-t)\gamma]\, dF(t).$$

and $\tau_* \geq \underline{t}(\sigma_B)$. The case studied just above corresponds to the case where neither constraint binds.

We can first ignore the second constraint, namely, that $\tau_* \geq \underline{t}(\sigma_B)$. Then, for each $\lambda > 0$ let

$\tau_*(\lambda)$ be the value of t such that

$$t(\gamma + (1-\gamma)\sigma_B) + (1-t)(1-\gamma+\gamma\sigma_B)(-d) = \lambda\left[t(1-\gamma) + (1-t)\gamma\right].$$

Similarly, let $\tau^*(\lambda)$ be the value of t such that

$$t \cdot 1 + (1-t)(-d) = t(\gamma + (1-\gamma)\sigma_B) + (1-t)(1-\gamma+\gamma\sigma_B)(-d) - \lambda\left[t(1-\gamma) + (1-t)\gamma\right].$$

Then it suffices to find λ that satisfies

$$(1-\kappa)\int_{\tau_*(\lambda)}^{\tau^*(\lambda)} \left[t(1-\gamma) + (1-t)\gamma\right] dF(t) \leq \kappa \int_0^{\underline{t}(\sigma_B)} \left[t(1-\gamma) + (1-t)\gamma\right] dF(t).$$

According to the numerical example, the second constraint does not bind (i.e., $\tau_* \geq \underline{t}(\sigma_B)$) if σ_B is sufficiently small. Let $\underline{\sigma}_B$ be the largest value of σ_B such that $\tau_*(\lambda(\sigma_B)) \geq \underline{t}(\sigma_B)$.

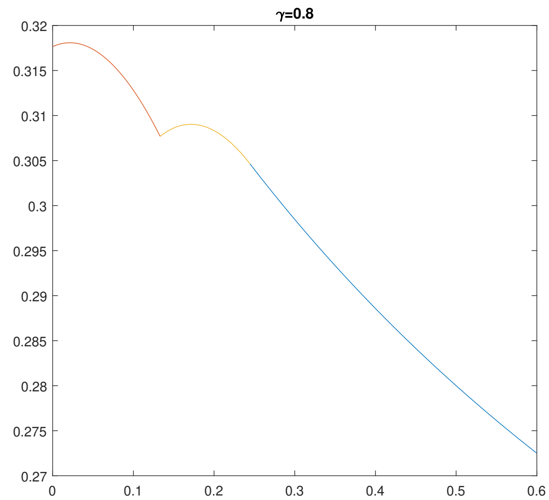
If $\sigma_B \in (\underline{\sigma}_B, \overline{\sigma}_B)$ then $\tau_* = \underline{t}(\sigma_B)$. Therefore, γ^* can be found from

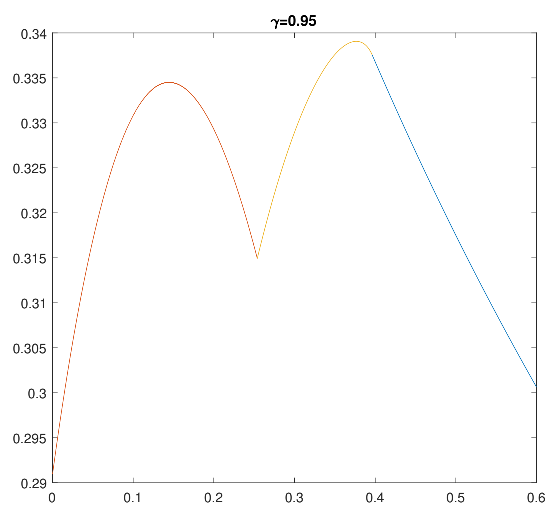
$$(1-\kappa)\int_{\underline{t}(\sigma_B)}^{\tau^*(\lambda)} \left[t(1-\gamma) + (1-t)\gamma\right] dF(t) \leq \kappa \int_0^{\underline{t}(\sigma_B)} \left[t(1-\gamma) + (1-t)\gamma\right] dF(t).$$

In short, there seem to be the following three cases.

- $\sigma_B \geq \overline{\sigma}_B$: S obtains her ideal outcome (so $\tau_* = \underline{t}(\sigma_B)$ and $\tau^* = \bar{t}$).
- $\sigma_B \in [\underline{\sigma}_B, \overline{\sigma}_B)$: In this case, (perhaps) $\tau_* = \underline{t}(\sigma_B)$ and $\tau^* < \bar{t}$.
- $\sigma_B < \underline{\sigma}_B$: In this case, $\tau_* > \underline{t}(\sigma_B)$ and $\tau^* < \bar{t}$.

Optimizing over σ_B. According to some numerical examples, the optimal solution can be in $(0, \underline{\sigma}_B)$ or in $(\underline{\sigma}_B, \overline{\sigma}_B)$. See the following two figures.


γ=0.8
0.32
0.315
0.31
0.305
0.3
0.295
0.29
0.285
0.28
0.275
0.27
0
0.1
0.2
0.3
0.4
0.5
0.6


γ=0.95
0.34
0.335
0.33
0.325
0.32
0.315
0.31
0.305
0.3
0.295
0.29
0
0.1
0.2
0.3
0.4
0.5
0.6