FTG Working Paper Series

Model Secrecy and Stress Tests

by

Yaron Leitner
Basil Williams

Working Paper No. 00061-01

Finance Theory Group

www.financetheory.com

Model Secrecy and Stress Tests*

Yaron Leitner† Basil Williams‡

September 1, 2021

Abstract

Should regulators reveal the models they use to stress test banks? In our setting, revealing leads to gaming, but secrecy can induce banks to underinvest in socially desirable assets for fear of failing the test. We show that although the regulator can solve this underinvestment problem by making the test easier, some disclosure may still be optimal (e.g., if banks have high appetite for risk or if capital shortfalls are not very costly). Cutoff rules are optimal within monotone disclosure rules, but more generally, optimal disclosure is single-peaked. We discuss policy implications and offer applications beyond stress tests.

JEL Classification: D82, G01, G28

Keywords: stress tests, Bayesian persuasion, information design, delegation, bank incentives, Fed models

---

*We thank Dong Choi, Pierre Chaigneau, Daniel Quigley, Anjan Thakor, Jing Zeng, Zhen Zhou, two anonymous referees, an anonymous associate editor, and the editor (Wei Xiong). We also thank seminar and conference participants at Bar Ilan University, Boston Fed, CEMFI, Federal Reserve Board, Office of Financial Research, NY Fed, NYU, Penn State, Philadelphia Fed, UNC Chapel-Hill, University of Houston, Virtual Finance Theory Seminar, Washington University in St. Louis, Tilburg University, Yeshiva University, CEPR Regulatory Forbearance Workshop, Econometric Society Asian meetings and World Congress, Einaudi Junior Finance Conference, Federal Reserve Stress Testing Conference, EFA, FIRS, UECE Lisbon Meetings in Game Theory and Applications, Midwest Macro Meetings, SAET, SED, and Tsinghua Theory and Finance Workshop.

†Olin Business School, Washington University in St. Louis, yaronleitner@wustl.edu

‡New York University, basil.williams@nyu.edu

1 Introduction

From the early days of bank stress tests in the wake of the 2008 financial crisis, disclosure has been a key issue of discussion among practitioners, academics, and regulators. Most of the academic discussion has centered around disclosure of the test results to the public (e.g., Goldstein and Sapra (2014), Goldstein and Leitner (2018)). However, academics have paid less attention to another important issue: should regulators disclose the models they use to project bank capital when conducting the test? This issue has recently gained momentum among policy makers and practitioners, leading to a change in the Fed's policy. Under the old policy, the Federal Reserve provided only a broad description of its stress test models. Under the new policy, it provides more information on certain equations and key variables, and illustrates how its models work on hypothetical loan portfolios. Yet, even under the new regime, the Federal Reserve does not fully reveal its models.[1]

An important reason for not revealing the models to the banks is to prevent banks from gaming the test—i.e., taking actions that enable them to pass the test without reducing risk. Indeed, in a speech on September 26, 2016, Former Fed Governor, Daniel Tarullo, said that "Full disclosure would permit firms to game the system—that is, to optimize portfolio characteristics based on the parameters of the model and take risks in areas not well-captured by the stress test just to minimize the estimated stress losses."[2] However, banks have constantly complained about model secrecy, claiming that even their best efforts to prepare for a test could result in unexpected and costly failure.[3] These claims cannot be ignored, particularly given evidence that regulatory uncertainty causes banks to reduce lending (Gissler, Oldfather, and Ruffino, 2016).

We present a stylized framework that allows us to examine the effects of revealing the regulator's stress test models to banks before the test. Our setting has two main forces.

[1]See `https://www.federalreserve.gov/newsevents/pressreleases/bcreg20190205a.htm`.
[2]See `https://www.federalreserve.gov/newsevents/speech/tarullo20160926a.htm`.
[3]See "Fed 'Stress Tests' Still Pose Puzzle to Banks," *Wall Street Journal*, March 12, 2015.

Not revealing reduces gaming, but it can also induce banks to reduce investment in socially desirable assets.

In our model, the bank has better capacity than the regulator to identify and measure risk, but there is a conflict of interest between the bank and the regulator: the bank wants to take more risk than is socially desirable. To be concrete, the bank can invest in a safe asset or a risky asset. The bank knows what the value of the risky asset will be during a crisis, but the regulator observes only a noisy signal of that value. This signal could represent the asset value predicted by the regulator's model, or alternatively it could represent a model parameter. The bank prefers to invest in the risky asset regardless of its true value during a crisis, but the regulator prefers the risky asset only if this value is sufficiently high. If the bank invests in the safe asset, it always passes the test. If the bank invests in the risky asset, it passes only if the regulator's signal is above some threshold. A bank that fails the test is required to reduce risk, which is costly to the bank. In the baseline model, the bank reduces risk by replacing the risky asset with the safe asset, but our setting also extends to the case in which the bank's portfolio is given and the bank submits a capital plan that needs to be approved by the regulator.

Our main focus is on whether the regulator should reveal his private signal to the bank before the bank makes its investment decision. Crucially, we allow the regulator to choose not only the disclosure policy but also the passing threshold. The regulator can adjust this passing threshold by choosing minimum capital requirements or the severity of the announced stress scenario.

We first compare between a transparent regime, in which the regulator reveals his signal, and a secret regime, in which the regulator does not reveal his signal. Under the transparent regime, the bank games the test in the sense that when the regulator reveals a passing signal, the bank invests in the risky asset even if it knows that the true value is low. Secrecy mitigates this problem. In particular, fear of failing the test incentivizes the bank to act more cautiously, investing in the risky asset only if its value exceeds some threshold. However, secrecy can open the door to a new problem: the bank avoids the risky asset not only when it is bad for society but also in some cases when it is good. Our first main result is that if the regulator can freely adjust the passing threshold, then despite this tradeoff, secrecy is always preferred. Intuitively, secrecy prevents gaming, and by setting a sufficiently low passing threshold (an easy-to-pass test), the regulator can also prevent the underinvestment that could result from secrecy.

We then analyze more flexible disclosure rules. Our second main result is that even if the regulator can set the passing threshold optimally, some disclosure may be optimal. The logic behind this result is as follows. The regulator has two tools to induce the bank to reduce risk. First, he can make the test harder by increasing the passing threshold. Second, he can provide partial information. In particular, he can commit to a cutoff disclosure rule, under which he sends a high message if his private signal is above some threshold and a low message if his private signal is below the threshold. The benefit from this disclosure policy is that if the regulator sends the low message, the bank infers that the risky asset is more likely to fail the test, and so it reduces investment in this asset.

However, each tool has a social cost. Partial disclosure leads to excessive risk if the regulator sends the high message, while a high passing threshold commits the regulator to sometimes fail the bank even if the regulator's model indicates the asset is good. In some cases, full secrecy requires a very high probability of failure to incentivize the bank. But then the regulator can gain by passing the bank more often and mitigating the worsening bank incentives via partial disclosure.

We use the intuition above to derive comparative statics (Section 5.3). For example, we show that under some regularity conditions, the regulator discloses more information if the bank's cost of failing the test is lower or if the bank's appetite for the risky asset is higher (e.g., if it can charge higher fees for originating risky loans). Intuitively, in this case, a hard test is not a very effective tool in inducing the bank to take less risk, and so, the regulator combines it with partial disclosure. In addition, we show that the regulator discloses more information when the social cost of capital shortfalls is low (e.g., if contagion is unlikely) or when the probability of a crisis is lower. Regarding bank leverage, two effects go in opposite directions. On the one hand, higher leverage makes capital shortfalls more likely, which pushes toward less disclosure. On the other hand, higher leverage can increase the bank's payoff from the risky asset (risk shifting), which pushes towards more disclosure. We also discuss the effect of the regulator's model accuracy on the optimal disclosure regime (Section 6).

On a technical level, we show that the single cutoff disclosure rule described above is optimal even if the regulator can choose multiple cutoffs. However, for some parameter values, the regulator can obtain a better outcome via a nonmonotone disclosure rule, in which messages pool signals from disconnected intervals. Essentially, very low signals are pooled with very high signals, less low with less high, etc. As we explain in Section 5.4, this disclosure rule can reduce the cost of providing incentives to the bank.

We discuss practical limitations on the regulator's ability to implement the two tools above. One example is when the regulator cannot commit to act according to a prespecified disclosure rule. In this case, the regulator may not be able to implement partial disclosure, and so the relevant comparison might be between a fully transparent regime and a fully secret regime. Another example is when the regulator faces heterogeneous banks but must apply the same passing threshold for everyone. We show that if banks are sufficiently different from one another, then in contrast to our first result, full transparency is preferred to secrecy. We also provide the following policy implications. First, greater model transparency does not necessarily require increased capital requirements. Second, illustrating how the Fed's model works on hypothetical loan portfolios could lead to increased correlation in bank asset holdings. (See Section 6.)

Finally, we offer applications of our theory beyond stress tests. One application is a firm's board of directors approving a manager's strategic plan. Another application is an investor approving an investment recommendation by a financial advisor (See Section 6.).

2 Related Literature

Our paper is related to several strands of literature. The first strand studies stress test disclosure. This literature has focused on disclosure of the test results to the public (e.g., Goldstein and Leitner (2018)).[4] In contrast, we focus on disclosure of the regulator's stress test models to the banks before the test. To our knowledge, we are the first paper to offer a formal analysis of this problem. An informal discussion, which includes additional effects that are not studied in our paper, is provided by Goldstein and Leitner (2020). In particular, they distinguish between revealing the models to the public vs. revealing them to the bank.[5] Recent papers have also explored other issues that relate to stress tests, besides disclosure. Colliard (2019) and Leitner and Yilmaz (2019) study the extent to which regulators should rely on banks' internal risk models. Shapiro and Zeng (2019) show that regulators' reputational concerns could lead to inefficiently tough stress tests. Parlatore and Philippon (2018) study the design of stress scenarios.

On the empirical front, there is growing evidence on the effect of stress tests on bank credit supply and the allocations of credit between safe and risky loans (e.g., Acharya, Berger, and Roman (2018) and Cortés et al. (2020)). However, these papers do not discuss welfare implications or the effect of regulatory uncertainty. There is also a large literature documenting the effects of political and regulatory uncertainty on the real economy, including reduced investment.[6] In particular, Gissler, Oldfather, and Ruffino (2016) offer evidence suggesting that uncertainty about the regulation of qualified mortgages caused banks to reduce mortgage lending. This literature is consistent with the idea that model secrecy could induce banks to reduce investment.

[4] A partial list of this growing literature includes Bouvard, Chaigneau, and Motta (2015), Faria-e-Castro, Martinez, and Philippon (2017), Williams (2017), Inostroza and Pavan (2017), Orlov, Zryumov, and Skrzypacz (2018), Corona, Nan, and Zhang (2019), and review papers by Goldstein and Sapra (2014), Leitner (2014), and Goldstein and Leitner (2020). More recent papers include Dogra and Rhee (2018), Quigley and Walther (2020), Inostroza (2019), and Huang (2019).

[5] See also Flannery (2019).

[6] See, for example, Julio and Yook (2012), Fernández-Villaverde et al. (2015), and Baker, Bloom, and Davis (2016).

Our paper also relates to the Bayesian persuasion and information design literature (e.g., Kamenica and Gentzkow (2011) and Bergemann and Morris (2019)), and in particular to the literature on Bayesian persuasion of a privately informed receiver. We contribute to this literature by providing a new application, in which a regulator (sender) discloses information about a parameter in his stress test model (his private signal) to induce a bank (receiver) to make socially desirable investments. Crucially, in our setting, the regulator not only discloses information but also controls the test difficulty. Moreover, our results on general disclosure (Section 5.4) provide an example of negative assortative disclosure that is single peaked (Kolotilin and Wolitzky, 2020). In particular, if the passing threshold is set optimally, it is optimal to give the bank higher action recommendations upon observing moderate signals, and lower action recommendations upon observing more extreme signals. Another example of single-peaked disclosure is provided by Guo and Shmaya (2019). In their setting, a sender discloses information about a product's quality to persuade a privately informed receiver to buy the product. The sender always wants to sell, but the receiver wants to buy only if the quality is sufficiently high. In contrast, in our setting the receiver always wants to invest, while the sender wants to invest only if the asset value is sufficiently high. Negative assortative disclosure is also optimal in Goldstein and Leitner (2018), but in their settings, optimal disclosure is single *dipped*. That is, more extreme states lead to *higher* action recommendations. Finally, Kolotilin, Mylovanov, et al. (2017) and Kolotilin (2018) study persuasion problems with a privately informed receiver, but because they focus on linear payoffs, the optimal disclosure rule in their settings can have a simple form and need not be nonmonotone[7]

Another related literature is that of delegation of authority in organizations, in which a principal can delegate authority to an informed but biased agent but cannot design monetary transfers.[8] In this literature, delegation is preferred if the gain from incorporating more

[7]See Proposition 1 and Examples 1 and 2 in Kolotilin and Wolitzky (2020).

[8]A partial list includes Holmstrom (1982), Aghion and Tirole (1997), Dessein (2002), Harris and Raviv (2008), Alonso and Matouschek (2008), Grenadier, A. Malenko, and N. Malenko (2016), and Chakraborty

information from the agent outweighs the losses due to the agent's bias. In our setting, if the regulator (principal) reveals his signal, he effectively restricts the bank's actions to those that will surely pass the test. So effectively, the regulator keeps authority. If the regulator does not reveal his signal, he gives the bank more freedom to choose an action, but in contrast to the delegation literature above, the regulator responds to the bank's action using an evaluation process that is based on the regulator's private information. Hence, we can think of our secrecy regime as "delegation with hidden evaluation." In our setting, this sort of delegation is powerful because by controlling the passing threshold, the regulator can indirectly control the bank's bias, namely the bank's equilibrium investment threshold. Our main results can be interpreted in light of this. The regulator is more likely to keep authority (reveal a passing signal) if it's more costly to control the agent's bias (e.g., if the bank's cost of failing the test is low, so a hard test is an ineffective tool in providing incentives).

Finally, in a different context, MacLeod (2003), Levin (2003), and Fuchs (2007) study settings of hidden evaluation, but these papers focus on optimal contracting rather than optimal disclosure. Levit (2020) studies a setting in which an informed principal can take a follow-up action, but in his setting the agent is uninformed and communication with the agent is only via cheap talk. Jehiel (2015) provides conditions under which a principal should remain silent about a payoff relevant variable, but in his setting, the agent is uniformed. Ederer, Holden, and Meyer (2018) provide conditions under which a principal can gain by randomizing between two incentives schemes, but in their setting, the principal is uninformed. Finally, Lazear (2006) studies a setting in which a principal can monitor only a limited number of actions that an agent can take. He shows that if agents do not respond much to penalties, the principal can gain by preannouncing the actions that will be monitored.

---

and Yılmaz (2017). See also Leitner and Yilmaz (2019), in which a regulator allocates authority to a bank based on the realization of a signal that the bank produces endogenously.

3 Model

There is a bank and a regulator. The bank can take one of two actions: invest in a safe asset or invest in a risky asset. The payoff from the risky asset depends on the realization of a random variable $\omega \in [\underline{\omega}, \bar{\omega}]$, which represents the value of the risky asset during a crisis. We refer to ω as the state of nature. The bank's payoff is $u(\omega)$ and the regulator's payoff, which represents the value to society, is $v(\omega)$. Both u and v are increasing in ω ($u' > 0$, $v' > 0$) and incorporate the probability of crisis, resulting losses, payoffs during normal times, etc. (For microfoundations of the payoff functions and all other model components, see Section 3.1.) The payoff from investing in the safe asset does not depend on ω and is normalized to zero for both the bank and regulator. Hence, u and v are the *relative gains* from investing in the risky asset, compared to the safe asset. To save on notation, we use the same letter to denote both a random variable and its realization.

There is a conflict of interest between the bank and the regulator. The bank prefers the risky asset to the safe asset in every state ω, but the regulator prefers the risky asset only if $\omega \geq \omega_r$, where $\omega_r \in (\underline{\omega}, \bar{\omega})$. Formally:

Assumption 1. $u(\omega) \geq 0$ *for all* $\omega \in [\underline{\omega}, \bar{\omega}]$

Assumption 2. $v(\omega) \geq 0$ *if and only if* $\omega \in [\omega_r, \bar{\omega}]$

The conflict of interest captures the idea that the bank does not internalize the social cost associated with risk. For our results, it is not crucial that the bank prefers the risky asset in every state. What matters is that there are states in which the bank prefers the risky asset but the regulator does not.[9]

There is also information asymmetry. The bank has superior information about the value of the risky asset during a crisis, and for simplicity, we assume the bank perfectly observes ω. The regulator does not observe ω, but he observes the realization of a noisy signal $s \in [\underline{s}, \bar{s}]$

[9] If there were states ω for which $u(\omega) < 0$ (in contrast to Assumption 1), the bank would never invest in those states, regardless of the regulator's disclosure policy.

of ω. The bank privately observes ω and the regulator privately observes s before the bank makes its investment decision. Everything else is common knowledge. The random variable ω has a cumulative distribution function (CDF) G and density g. Conditional on ω, s has CDF $F(\cdot|\omega)$ and density $f(\cdot|\omega)$. Both $g(\cdot)$ and $f(\cdot|\omega)$ have full support. We also assume:

Assumption 3 (MLRP)**.** *If $\omega' > \omega$, the ratio $f(s|\omega')/f(s|\omega)$ is strictly increasing in s.*

Assumption 3 implies that $1 - F(s|\omega)$ is strictly increasing in ω. That is, the regulator is more likely to observe higher signals when the state ω is higher.[10]

After the bank makes its investment decision, the regulator observes the bank's investment and decides whether to pass or fail the bank—i.e., the regulator conducts a stress test. If the bank chooses the safe asset, it always passes the test. If the bank chooses the risky asset, it passes only if the regulator's signal s is above some threshold, which we denote by s_p. The regulator chooses and publicly announces s_p before the bank invests.

If the bank fails the test, the regulator forces it to replace the risky asset with the safe asset. In this case, the bank incurs a cost $c > 0$, which reflects a cost to the bank's managers from loss of reputation or a decline in the stock price. Alternatively, c could reflect a fixed cost that the bank needs to incur before investing in the risky asset, and which is already included in $u(\omega)$. If c is a transfer to other economic agents, then this cost affects the bank but not the regulator.

Hence, the final payoffs are as follows (see Table 1). If the bank invests in the safe asset, both the bank and the regulator end up with a final payoff of zero. If the bank invests in the risky asset and passes the test, the bank's final payoff is $u(\omega)$ and the regulator's final payoff is $v(\omega)$. Finally, if the bank invests in the risky asset and fails the test, the bank's final payoff is $-c$, and the regulator's final payoff is zero. Our main results do not depend on the exact specification of final payoffs above. For example, c could depend on ω, u could be flat (and positive), and the regulator's payoff after failing the bank need not be zero (see Remark 1 and item 5 of Section 6).

[10] See Milgrom (1981).

Action	Test result	Bank's Payoff	Regulator's payoff
Safe	Always pass	0	0
Risky	Pass ($s \geq s_p$)	$u(\omega)$	$v(\omega)$
	Fail ($s < s_p$)	$-c$	0

Table 1: Final payoffs to the bank and the regulator.

The focus of our paper is whether the regulator should reveal his signal s to the bank. We start with the case in which the regulator can only reveal or not reveal s (Section 4). Then we explore more general disclosure rules (Section 5). In both cases, the regulator publicly commits to the disclosure policy and to a pass/fail rule, assumptions that we discuss in Section 6. We refer to investment in the risky asset simply as "investing" and investment in the safe asset as "not investing."

The sequence of events is as follows: (i) the regulator publicly commits to a disclosure policy about s and to a passing threshold s_p; (ii) nature chooses ω, the bank observes ω, and the regulator observes s; (iii) the regulator discloses information about s in accordance with his disclosure policy; (iv) the bank chooses the risky asset ("invest") or safe asset ("not invest"); (v) the regulator performs a stress test, and final payoffs are realized.

We solve the game backwards. We first characterize the bank's investment decision, for a given passing threshold and disclosure regime. Then we solve for the optimal threshold and regime. If the bank is indifferent between two actions, we assume that it chooses the one that is preferred by the regulator, and if also the regulator is indifferent, we assume that the bank invests. If the regulator is indifferent between multiple passing thresholds, he picks the highest one. Our main results do not depend on these assumptions.

To simplify the exposition, we focus on the more interesting case in which (i) it is optimal to sometimes fail the bank ($s_p > \underline{s}$) and (ii) secrecy induces the bank to respond by reducing investment. A sufficient condition for this is the following.[11]

Assumption 4. *$E[v(\omega)|\underline{s}] < 0$ and $u(\underline{\omega}) = 0$.*

[11] We provide more details in the proofs of Lemma 1 and Lemma 3.

3.1 Microfoundations

We now fill in specific forms in the context of the banking stress tests application. Example 1 shows how to incorporate the probability of a crisis, social costs of capital shortfalls, and bank's leverage, into the payoff functions u and v. Example 2 shows that the regulator's signal s could represent a parameter in his stress test model. Example 3 shows that choosing a higher passing threshold s_p amounts to setting higher capital requirements and/or choosing a more severe stress scenario. Finally, Example 4 shows that our model maps into a case in which the bank's portfolio is given, and instead of choosing a portfolio, the bank submits a capital plan, which needs to be approved by the regulator.

Example 1 (payoff functions)**.** Suppose the bank has \$1 that can be invested in either a safe asset or a risky asset. The value of the risky asset is \$2 in normal times and $\omega \in (0, 1)$ during a crisis. The value of the safe asset is always \$1. A crisis occurs with probability q. Suppose the bank also has a debt liability with face value $D < 1$, and that it acts as to maximize the payoffs to its equity holders. If the bank invests in the safe asset, the debt is riskless, and the payoff to its equity holders is $1 - D$. If the bank invests in the risky asset, the debt is risky, and the expected payoff to equity holders is $(1 - q)(2 - D) + q \max\{\omega - D, 0\}$. Recall that the function u is the relative gain from investing in the risky asset compared to the safe asset. Hence, $u(\omega) = (1 - q)(2 - D) + q \max\{\omega - D, 0\} - (1 - D)$.

As for the regulator, assume that the if the bank defaults, there is a social loss $L(D - \omega)$, where $L > 0$. For example, the social loss could be due to spillovers (contagion) to other banks or the rest of the economy. Assume that the regulator acts as to maximize total surplus, which is the sum of payoffs to debt holders and equity holders less the social loss. Then if the bank invests in the safe asset, the regulator's payoff is 1. If the bank invests in the risky asset, the regulator's payoff is $2(1 - q) + q\omega - qL(D - \omega)$. The relative gain from investing in the risky asset is $v(\omega) = 2(1 - q) + q\omega - qL(D - \omega) - 1$. Note that u and v are increasing in ω, and if q, D, and L are chosen appropriately, Assumptions 1 and 2 are satisfied.

Example 2 (regulator's model)**.** Suppose the value of the risky asset during a crisis is given by $\beta_0 + \sum_{i=1}^{n} \beta_i \widetilde{X}_i$, where $\widetilde{X}_i$ is a random variable that denotes some macroeconomic variable (factor i), and β_i is the sensitivity to this factor.[12] Assume that the density of $(\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_n)$ and all the sensitivities are common knowledge, except for β_1, which is privately observed by the bank. The regulator observes only an estimate of β_1, which we denote by $\hat{\beta}_1$, and which could result from a regression that uses historical data. Before the bank invests, the regulator chooses and publicly announces a "stress scenario," namely a particular realization $(X_1, X_2, ..., X_n)$ of $(\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_n)$. The value of the risky asset predicted by the regulator's model for the assumed stress scenario is then $\beta_0 + \widehat{\beta}_1 X_1 + \sum_{i=2}^{n} \beta_i X_i$. To map this example to our baseline model, relabel β_1 as ω and $\hat{\beta}_1$ as s.[13]

Example 3 (passing threshold)**.** Continue Example 2. Suppose the bank has existing debt with face value D, and let κ denote minimum capital requirements. Assume that $1 - D \geq \kappa$, so if the bank chooses the safe asset, it always passes the test. If instead the bank chooses the risky asset, it passes if and only if it has sufficient capital according to the regulator's model:

$$\frac{\beta_0 + \widehat{\beta}_1 X_1 + \sum_{i=2}^{n} \beta_i X_i - D}{\beta_0 + \widehat{\beta}_1 X_1 + \sum_{i=2}^{n} \beta_i X_i} \geq \kappa \tag{1}$$

This reduces to

$$\hat{\beta}_1 \geq \frac{\frac{D}{1-\kappa} - \beta_0 - \sum_{i=2}^{n} \beta_i X_i}{X_1}. \tag{2}$$

The right-hand side in equation (2) maps to the passing threshold s_p in our baseline model. Hence, a higher s_p amounts to setting higher capital requirements κ or choosing a more severe stress scenario, namely, lower values for X_i.

Example 4 (capital plans)**.** Suppose the bank has already invested in the risky asset from

[12] The example easily extends to the case in which the value of the risky asset is $f(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)$ for some nonlinear function f. For example, the Federal Reserve uses similar formulas to project credit losses (Board of Governors of the Federal Reserve System, 2021).

[13] Assume that $E(\widetilde{X}_i) > 0$. Then the payoffs from investing in the risky asset are increasing in β_1 for both the bank and the regulator.

Example 1, and that it also has $1 + \delta$ dollars in excess cash. The bank can submit one of two capital plans: retain the cash (safe action) or pay it as dividends (risky action). Assume that retaining cash leads to a deadweight loss δ, which could reflect wasteful investment resulting from a free cash flow problem (Jensen, 1986) or that investors could channel the money to better use. However, retaining too little cash increases the probability of default and financial distress, which can increase the bank's future borrowing costs. To capture this, specify the final payoffs as in Table 1, where the relative gains of the risky action compared to the safe action are $u(\omega) = \delta - qr(D - \omega)$ to the bank and $v(\omega) = \delta - qL(D - \omega)$ to the regulator. The first term in each function is the gain from avoiding wasteful investment, and the second term is the expected cost of financial distress and contagion ($L > r > 0$). Note that u and v are increasing in ω, and if the parameter are chosen appropriately, Assumptions 1 and 2 are satisfied.[14,15]

4 Revealing vs. Not Revealing

In this section, we compare between two disclosure regimes: revealing (the regulator reveals his signal s to the bank) and not revealing (the regulator does not reveal his signal to the bank).

4.1 Bank's Investment

Let p denote the bank's perceived probability of passing the test upon investment. If the bank invests, its expected payoff is $pu(\omega) - (1 - p)c$. If the bank does not invest, its payoff

[14]Note that in Example 1, the debt level was given, and so equity holders could benefit from risk shifting as the expense of existing debt holders. More generally, we could combine both the benefit from risk shifting due to existing debt and the cost of financial distress which will affect borrowing costs on new debt.

[15]Here c could represent a cost to the bank's managers from failing the test or a fixed cost (transfer to other economic agents) that the bank needs to incur before submitting a capital plan that involves a dividend distribution.

is zero. Hence, the bank invests if and only if

$$pu(\omega) - (1-p)c \geq 0. \tag{3}$$

First, consider revealing. If the regulator reveals a passing signal $s \geq s_p$, then $p = 1$, and the bank's payoff from investing is $u(\omega) \geq 0$. If the regulator reveals a failing signal $s < s_p$, then $p = 0$, and the bank's payoff from investing is $-c < 0$. Hence, the bank invests if and only if the regulator observes a passing signal.

Next, consider not revealing. Now the perceived probability p of passing the test depends on ω:

$$p(\omega) \equiv P(s \geq s_p|\omega) = 1 - F(s_p|\omega). \tag{4}$$

From Assumption 3 (MLRP), $p(\omega)$ is increasing in ω. Since $u' > 0$, it then follows that the left-hand-side in equation (3) is increasing in ω. Hence, the bank follows a cutoff rule, investing if and only if the state ω is above some threshold, which we denote by ω_{NR} ("NR" stands for "not revealing"). Intuitively, a higher ω indicates that the bank is more likely to pass the test, and it also indicates that its payoff conditional on passing is higher. Later, we also use $\omega_{NR}(s_p)$ to denote the dependence of ω_{NR} on s_p. For some parameter values, it is optimal for the bank not to invest at all. In this case, we let $\omega_{NR} = \bar{\omega}$, which implies that the bank invests with probability 0.

The next lemma summarizes the preceding discussion.

Lemma 1. *1. If the regulator reveals his signal s to the bank, the bank invests if and only $s \geq s_p$.*

2. If the regulator does not reveal his signal s to the bank, there exists $\omega_{NR} > \underline{\omega}$, such that the bank invests if and only if $\omega \geq \omega_{NR}$. The investment threshold ω_{NR} is continuous and increasing in both s_p and c.

The first part in Lemma 1 captures the idea that revealing the regulator's model (which is captured by the signal s) could lead to gaming. In particular, if the regulator reveals a passing signal, the bank invests even if it knows the risky asset will perform poorly in a crisis—i.e., the bank games the test. This is consistent with regulator concerns about gaming, discussed in the introduction. Regulators have also expressed concerns that revealing the regulator's models will cause banks to rely too heavily on them rather than their own models. Consistent with these concerns, Lemma 1 shows that under revealing, the bank's investment relies on the regulator's signal rather than its own information about ω (see also the discussion of endogenous information production in Section 6).

The second part in Lemma 1 captures the idea that not revealing makes the bank more cautious, leading it to avoid investment if $\omega < \omega_{NR}$. The fact that ω_{NR} increases in both s_p and c reflects that the bank becomes more cautious if the test is more difficult or if the cost of failing is higher. The fact that $\omega_{NR} > \underline{\omega}$ follows from Assumption 4.

Remark 1. Our results do not depend on the exact specification of bank characteristics, namely, the payoff function u and the cost of failing the test c. Any specification such that the left-hand-side in equation (3) is increasing in ω will imply that the bank follows a cutoff investment rule and will hence yield similar results.

4.2 Regulator's Payoff

We use V_R and V_{NR} to denote the regulator's payoff under revealing and under not revealing, respectively; later, we also use $V_R(s_p)$ and $V_{NR}(s_p)$ to denote the dependence on s_p. We derive the regulator's payoffs as follows. Conditional on observing a failing signal $s < s_p$, the regulator's payoff is zero because the bank either does not invest or invests and fails the test. Conditional on observing a passing signal $s \geq s_p$, the regulator's payoff depends on the bank's investment decision from Lemma 1. Under revealing, the bank invests in every state $\omega \geq \underline{\omega}$ and the regulator obtains $\int_{\omega \geq \underline{\omega}} v(\omega) f(\omega|s) d\omega$. Under not revealing, the bank invests only if $\omega \geq \omega_{NR}$ and the regulator obtains $\int_{\omega \geq \omega_{NR}} v(\omega) f(\omega|s) d\omega$. Taking the expectation across all signals $s \in S$ and changing the order of integration, we obtain the following:

Lemma 2. *If the regulator reveals his signal, his payoff is*

$$V_R = \int_{\omega \geq \underline{\omega}} [1 - F(s_p|\omega)] v(\omega) g(\omega) d\omega.$$

If the regulator does not reveal his signal, his payoff is

$$V_{NR} = \int_{\omega \geq \omega_{NR}} [1 - F(s_p|\omega)] v(\omega) g(\omega) d\omega. \tag{5}$$

The payoffs under the two disclosure regimes are similar, except that the integral for V_R starts at $\underline{\omega}$, whereas the integral for V_{NR} starts at $\omega_{NR} > \underline{\omega}$. This reflects the fact that under not revealing, the bank acts more cautiously, investing in fewer states. The expression inside the integrals reflects that the regulator obtains $v(\omega)$ only if the bank passes the test, which happens with probability $1 - F(s_p|\omega)$.

4.3 Preferred Regime

The preferred regime is the one that gives a higher payoff to the regulator. We show that for a given passing threshold s_p, revealing is preferred if and only if the bank's investment threshold ω_{NR} is sufficiently high. However, if the regulator can set the passing threshold optimally, not revealing is strictly preferred.

To see why, suppose first that s_p is given. If $\omega_{NR} < \omega_r$, it follows from Lemma 2 that $V_{NR} > V_R$. In this case, not revealing is preferred, because it induces the bank not to invest in states $\omega < \omega_{NR}$, in which investment is socially undesirable. In other words, not revealing reduces *overinvestment*. However, if $\omega_{NR} > \omega_r$, the following tradeoff exists. Not revealing eliminates the bank's overinvestment in states $\omega < \omega_r$, but it also leads to *underinvestment*: the bank does not invest in states $\omega \in [\omega_r, \omega_{NR})$, in which investment is socially desirable. In this case, revealing is preferred only if the underinvestment effect dominates, namely if ω_{NR} is sufficiently high. We provide a formal statement of this result in Proposition D1 in

Appendix D.

However, the next theorem shows that if the regulator can set the passing threshold optimally, then despite the tradeoff above, not revealing is always preferred. Formally, let s_p^R denote the passing threshold that the regulator sets if he plans to reveal his signal and s_p^{NR} denote the passing threshold that he sets if he does not plan to reveal. That is, $s_p^R \in \arg\max_{s_p} V_R(s_p)$ and $s_p^{NR} \in \arg\max_{s_p} V_{NR}(s_p)$. Then:

Theorem 1. *If the regulator sets the passing threshold optimally, not revealing is strictly preferred to revealing. That is,* $V_{NR}(s_p^{NR}) > V_R(s_p^R)$.

Theorem 1 captures the following intuition. Not revealing has two effects. It prevents gaming, but it can also lead to underinvestment. However, if the regulator can freely adjust the passing threshold, he can solve the underinvestment problem, while continuing to prevent gaming, by simply making the test easier (reducing s_p).

More formally, reducing s_p has two effects on the regulator's payoff under not revealing. First, it reduces the bank's investment threshold ω_{NR}. Second, it increases the probability $1 - F(s_p|\omega)$ of passing the bank. If s_p is such that revealing is preferred, there must be underinvestment. But then both effects of reducing s_p increase the regulator's payoff. In particular, a lower s_p mitigates the bank's underinvesment, and it also allows the regulator to capture the benefits from the bank's investment more often. As a result, any optimal s_p must imply that not revealing is preferred.

The next lemma includes two observations that follow from the discussion above.

Lemma 3. *Under the optimal passing threshold (and preferred disclosure regime):*

1. The bank invests in some states in which investment is socially undesirable. That is, $\omega_{NR}(s_p^{NR}) < \omega_r$.

2. The regulator sometimes fails the bank even though it is strictly optimal to pass the bank ex post.

The first part in Lemma 3 says that although the regulator could induce his ideal investment threshold ω_r, he sets the passing threshold s_p so that some overinvestment remains. If this were not true, the regulator could gain by making the test easier, which would allow him to capture the social benefit from the bank investment more often. If $\omega_{NR} > \omega_r$, an easier test would also reduce the bank's underinvesment, thereby benefiting the regulator even more. If $\omega_{NR} = \omega_r$, an easier test would worsen the bank's incentives, but this effect is negligible compared to the benefit of approving the bank's investment more often.

The second part shows that under the optimal policy, the regulator takes advantage of his ability to commit to a passing threshold. If this were not true, the regulator could increase his payoff by raising s_p. In particular, if it were strictly optimal to fail the bank ex post whenever $s \leq s_p$, a higher s_p would both mitigate the bank's overinvestment from part 1 and allow the regulator to fail the bank in more situations in which is it ex-post optimal to do so. If at $s = s_p$, the regulator were just indifferent between passing and failing the bank, a higher s_p would reduce his ex-post payoff, but this effect is negligible compared to the benefit of improved incentives.

In Appendix C, we show that Theorem 1 and the first part in Lemma 3 continue to hold even if the regulator cannot commit to a passing threshold. The logic is as follows. First, although the inability to commit to a passing threshold makes it harder to provide incentives to the bank—which reduces the benefit from not revealing—not revealing still provides better incentives than revealing. That is, under not revealing, there is less overinvestment. Second, because of the better incentives, not revealing provides a higher payoff to the regulator conditional on any signal s, and so the regulator passes the bank more often. This means that under not revealing, the passing threshold is effectively lower, which allows the regulator to capture the social benefit from the bank's investment more often. This logic extends to the case in which the regulator follows a cutoff disclosure rule as in Section 5.

5 Optimal Disclosure

We saw that for a given passing threshold, revealing is preferred to not revealing if the latter leads the bank to act too cautiously. However, if the regulator sets the passing threshold optimally, revealing is strictly dominated. In this section, we show that once we allow for partial disclosure, revealing some information may be optimal even if the regulator sets the passing threshold optimally. We also characterize optimal disclosure and provide comparative statics.

For ease of exposition, we first focus on a simple form of partial disclosure: a cutoff rule, which is defined by a threshold s_d, such that the regulator reveals whether the signal realization s is above or below s_d (the subscript "d" stands for "disclosure threshold"). We show that even this simple rule may be preferred to no disclosure. Moreover, as we explain in Section 5.4, cutoff rules remain optimal even within the larger set of "monotone disclosure rules" under which the regulator partitions the signal space into nonoverlapping intervals and reveals the interval to which the signal belongs.

5.1 Cutoff Rules

Under a cutoff disclosure rule, the regulator sends the bank one of two messages: a "low" message upon observing a signal below s_d and a "high" message upon observing a signal above s_d. Given each message, the bank forms a posterior belief regarding the probability of passing the test (p in equation (3)). From Assumption 3 (MLRP), these posterior probabilities are increasing in ω.[16] Hence, as in Section 4.1, each message induces the bank to follow a cutoff investment rule.

Specifically, if $s_d \in (s_p, \bar{s})$, the high message fully reveals to the bank that the regulator observed a passing signal. So upon receiving this message, the bank invests in every state. The low message pools together the remaining passing signals with all the failing signals. So

[16] In particular, given any message m, the bank's posterior beliefs satisfy MLRP, and as a result, $1 - F(s_p|m, \omega)$ is increasing in ω. The proof of Lemma B1 contains more details.

upon receiving this message, the bank behaves more cautiously, investing only if ω is above some threshold, which we denote by $\omega_L(s_d, s_p)$. As s_d increases, more passing signals are pooled with the failing signals, and so the bank behaves less cautiously. Hence, $\omega_L(s_d, s_p)$ is decreasing in s_d. If instead $s_d \in (\underline{s}, s_p)$, the low message fully reveals that the regulator observed a failing signal. So upon receiving this message, the bank does not invest at all. The high message pools together the remaining failing signals with all the passing signals. So upon receiving this message, the bank behaves less cautiously, investing only if ω is above some threshold, which we denote by $\omega_H(s_d, s_p)$. As s_d increases, less failing signals are pooled, and so the bank behaves less cautiously. Hence, $\omega_L(s_d, s_p)$ is decreasing in s_d.

Essentially, if $s_d \in (s_p, \bar{s})$, the regulator commits to give the bank a green light to invest if the signal realization is particularly high. So if the regulator does not give a green light, the bank behaves more cautiously than under not revealing. Similarly, if $s_d \in (\underline{s}, s_p)$, the regulator commits to give the bank a warning not to invest if the signal realization is particularly low. So if the regulator does not give a warning, the bank behaves less cautiously than under not revealing. [17]

The next lemma summarizes the observations above.

Lemma 4. *1) If $s_d \in (s_p, \bar{s})$, then if the regulator sends the high message, the bank invests in every state, and if he sends the low message, the bank invests if and only if $\omega \geq \omega_L(s_d, s_p)$.*

2) If $s_d \in (\underline{s}, s_p)$, then if the regulator sends the high message, the bank invests if and only if $\omega \geq \omega_H(s_d, s_p)$, and if he sends the low message, the bank does not invest.

3) The investment thresholds $\omega_L(s_d, s_p)$ and $\omega_H(s_d, s_p)$ are continuous and strictly decreasing in s_d. Moreover, $\omega_L(s_d, s_p) > \omega_{NR}(s_p) > \omega_H(s_d, s_p)$.

Note that if $s_d \in (s_p, \bar{s})$, we obtain the same outcome if instead of sending the high message, the regulator reveals the actual signal realization. Hence, we refer to this case as

[17] In practice, the Fed illustrates how its models work by providing estimates of loss rates on hypothetical loan portfolios and different kinds of loans held by banks. Low estimates can be interpreted as a green light to invest in a specific loan class, while high estimates can be interpreted as a warning not to invest.

the regulator revealing some of the passing signals. Similarly, if $s_d \in (\underline{s}, s_p)$, we obtain the same outcome if instead of sending the low message, the regulator reveals the actual signal realization. Hence, we refer to this case as the regulator revealing some of the failing signals. If $s_d = s_p$, the outcome is the same as under full disclosure, and if $s_d \in \{\underline{s}, \bar{s}\}$, the outcome is the same as under no disclosure.

The next theorem shows that if the regulator sets the passing threshold optimally, it may be optimal to reveal some of the passing signals. However, it is never optimal to reveal a failing signal.

Theorem 2. *If the regulator sets the passing threshold optimally, then either no disclosure is optimal or else it is optimal to reveal some of the passing signals, namely, set $s_d \in (s_p, \bar{s})$. A sufficient condition for partial disclosure to strictly dominate no disclosure is that equation* (6) *below holds.*

The proof of Theorem 2 has two parts. First, we show that for some parameter values, setting $s_d \in (s_p, \bar{s})$ leads to a better outcome than no disclosure. Recall that under not revealing and the corresponding passing threshold $s_p = s_p^{NR}$, the bank invests in some states in which investment is socially undesirable (Lemma 3). Setting $s_d \in (s_p, \bar{s})$ can help solve this overinvestment problem because if the regulator sends the low message, the bank acts more cautiously: $\omega_L(s_d, s_p) > \omega_{NR}(s_p)$. However, this type of partial disclosure has a social cost: if the regulator sends the high message, the bank acts more recklessly, investing in every state. Hence, this type of partial disclosure is not necessarily optimal. A sufficient condition for partial disclosure to be optimal is that if the regulator does not disclose anything ($s_d = \bar{s}$) and fixes the passing threshold at the one that is optimal under no disclosure (s_p^{NR}), he can obtain a better outcome by slightly reducing s_d. In the proof, we show that this condition reduces to

$$-\int_{\underline{\omega}}^{\omega_{NR}} v(\omega) f(s_d|\omega) g(\omega) d\omega < \frac{\partial \omega_L}{\partial s_d} v(\omega_{NR}) g(\omega_{NR}) [F(s_d|\omega_{NR}) - F(s_p|\omega_{NR})], \quad (6)$$

evaluated at $(s_d, s_p) = (\bar{s}, s_p^{NR})$. The left-hand side in (6) is the marginal cost of reducing s_d, and the right-hand side is the marginal benefit. The cost is that if the regulator sends the high message, the bank's investment threshold falls from ω_{NR} to $\underline{\omega}$. That is, there is more overinvestment. The benefit is that if the regulator sends the low message, the bank's investment threshold ω_L rises above ω_{NR}. That is, there is less overinvestment. The social gain from the reduced overinvestment is that with probability $F(s_d|\omega_{NR}) - F(s_p|\omega_{NR})$, which is the probability of observing a passing signal and sending the low message, the regulator does not incur the negative payoff $v(\omega_{NR})$.[18]

In the second part of the proof, we show that setting $s_d \in (\underline{s}, s_p)$ is suboptimal. In this case, sending the high message worsens the overinvestment problem, $\omega_H(s_d, s_p) < \omega_{NR}(s_p)$, while sending the low message leads to no investment at all. Hence, $s_d \in (\underline{s}, s_p)$ cannot be optimal.

We conclude this subsection with the following observations, which are similar to those in Lemma 3.

Lemma 5. *Under the optimal policy (s_d, s_p):*

(i) Upon receiving the low message, the bank invests in some states in which investment is socially undesirable. That is, $\omega_L(s_d, s_p) < \omega_r$.

(ii) The regulator sometimes fails the bank even though it is optimal to pass the bank ex post.

5.2 Two Tools To Mitigate Overinvestment

Theorem 2 captures the idea that the regulator has two tools to mitigate the bank's overinvestment. First, he can make the test harder (increasing s_p). Second, he can reveal some of the passing signals (reducing s_d), so that the bank acts less recklessly upon receiving the

[18] Recall that $\frac{\partial \omega_L}{\partial s_d} < 0$.

low message. Each tool helps induce a lower investment threshold ω_L, but each tool has a social cost. A harder test forces the regulator to fail the bank in more situations in which it is optimal to pass the bank ex post, while revealing passing signals leads to gaming and overinvestment. Partial disclosure is optimal if, starting with no disclosure, it is less costly to provide incentives to the bank via the second tool.

Formally, by combining (6) with the first-order condition for s_p, we obtain a sufficient condition for partial disclosure in terms of the marginal costs of each tool:

$$\begin{aligned}&\left(\frac{\partial \omega_L}{\partial s_d}\right)^{-1}\int_{\underline{\omega}}^{\omega_L} v(\omega)f(s_d|\omega)g(\omega)d\omega \\ &< -v(\omega_L)g(\omega_L)[F(s_d|\omega_{NR})-F(s_p|\omega_{NR})] \\ &= \left(\frac{\partial \omega_L}{\partial s_p}\right)^{-1}\int_{\omega_L}^{\overline{\omega}} v(\omega)f(s_p|\omega)g(\omega)d\omega,\end{aligned} \tag{7}$$

evaluated at $(s_d, s_p) = (\bar{s}, s_p^{NR})$. The reciprocals of the partial derivatives capture how much of each tool is needed to implement a small change in ω_L, and the integrals capture the social cost of doing so. The first line refers to the cost of increasing ω_L by reducing s_d, namely the overinvestment in $[\underline{\omega}, \omega_L]$ due to gaming. The third line refers to the cost of increasing ω_L by increasing s_p, namely the foregone investment in $[\omega_L, \overline{\omega}]$ due to test failure.[19] The second line captures the benefit of increasing ω_L, which must be equal to the cost of raising s_p because s_p is interior. So condition (7) shows that partial disclosure is optimal if, starting from no disclosure, raising ω_L by reducing s_d is less costly than raising ω_L by raising s_p.

5.3 Comparative Statics

We now present some comparative statics with respect to parameters, such as the bank's cost of failing the test, the probability of a crisis, the social cost of capital shortfalls, and the bank's leverage. As we will see, the general principle is that any parameter changes that

[19]Technically, the negative of the integral in the first line captures the cost of reducing s_d.

increase the bank's or regulator's gain from investing in the risky asset will lead to more disclosure.

We first examine the effect of an increase in the bank's cost of failing the test c. Under some regularity conditions, the sufficient condition for partial disclosure in (7) is also necessary, and so partial disclosure is optimal if and only if c is sufficiently low. To see why, observe that ω_L solves (3), where $p = \Pr(s \geq s_p | s < s_d)$. Applying the implicit function theorem, we then obtain that

$$\frac{\partial \omega_L / \partial s_p}{-\partial \omega_L / \partial s_d} = \frac{f(s_p|\omega_L)}{f(s_d|\omega_L)} \frac{F(s_d|\omega_L)}{F(s_p|\omega_L)}. \tag{8}$$

(Appendix D contains more details.) Hence, we can rewrite (7) as follows:

$$\frac{\int_{\omega \geq \omega_L} v(\omega) f(s_p|\omega) g(\omega) d\omega}{-\int_{\underline{\omega}}^{\omega_L} v(\omega) f(s_d|\omega) g(\omega) d\omega} > \frac{f(s_p|\omega_L)}{f(s_d|\omega_L)} \frac{F(s_d|\omega_L)}{F(s_p|\omega_L)}. \tag{9}$$

From Assumption 3 (MLRP), the term $F(s_d|\omega_L)/F(s_p|\omega_L)$ on the right-hand side of (9) is increasing in ω_L,[20] and under some regularity conditions on f, it has a first-order effect on the comparative statics. In this case, we obtain that (9) holds if and only if ω_L is sufficiently low, and the result on c follows because ω_L is increasing in c.

Intuitively, if c is high the bank is very concerned about failing the test, and so making a harder test is a relatively more effective tool in providing incentives. However, if c is low, the bank is relatively less responsive to a harder test, and so the regulator combines it with partial disclosure. In this case, equation (9) holds with equality.

Extending the intuition above, we also obtain that under some regularity conditions, if c is lower, the regulator discloses more information. That is, if c is lower, the regulator reveals more passing signals (reducing s_d). Figure 1 illustrates an example in which this comparative statics holds.

[20] It is the inverse of the posterior probability of failing the test conditional on the low message.

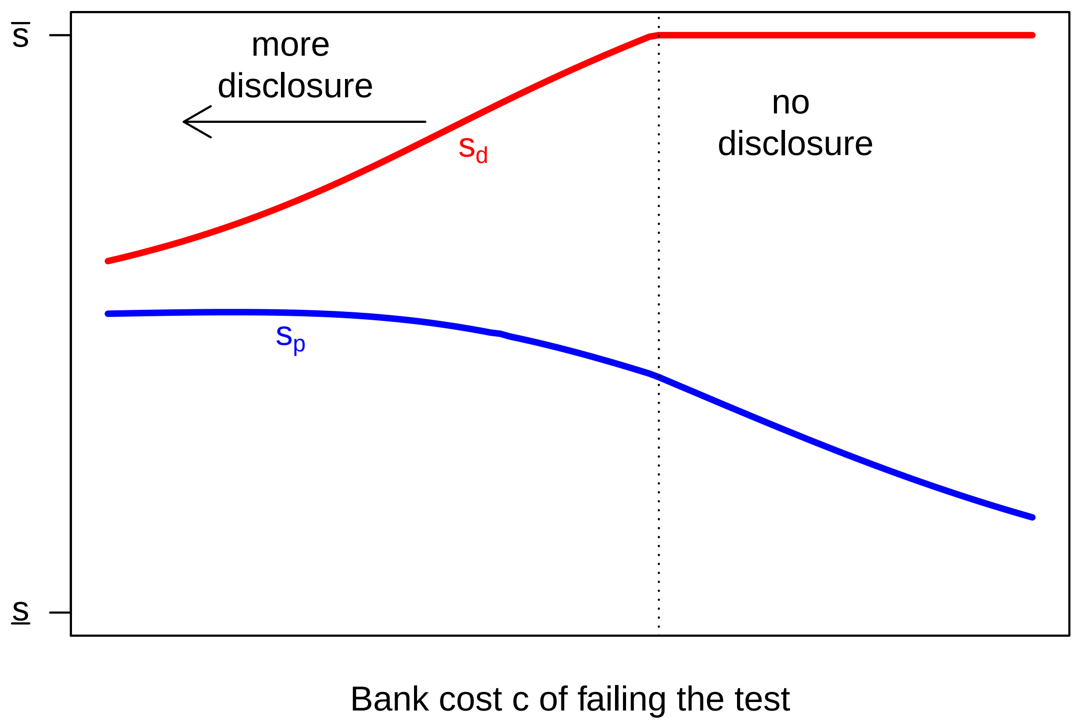

Figure 1: Optimal disclosure threshold s_d and optimal passing threshold s_p, as a function of the bank cost c of failing the test. In this example, $u(\omega) = (3\omega)^{0.1}$ and $v(\omega) = \omega - 0.5$, ω is uniform on $[0, 1]$, and $f(s|\omega) = 2(s\omega + (1 - s)(1 - \omega))$. We obtain a similar figure if $f(s|\omega)$ is a truncated normal distribution centered at ω.

Remark 2. In the comparative statics above, we assumed that the passing threshold s_p is chosen optimally. If s_p was exogenous, we would obtain that at an intermediate c, the regulator does not disclose anything, and that as c moves in *either* direction, the regulator discloses more. That is, the regulator discloses more not only when c is lower, but also when c is higher. The intuition for the second part is that for a fixed passing threshold, a very high c induces the bank to act too cautiously, leading to underinvestment. In this case, the regulator discloses some of the *failing* signals to induce the bank act less cautiously, and he discloses more failing signals when c is higher. Appendix D contains more details.

The comparative statics with respect to parameters that affect the bank's payoff from the risky asset u are opposite to that with respect to c. In particular, parameter changes that lead to an increase in $u(\omega)$ will push toward more disclosure, and changes that lead to a decrease in u push toward less disclosure. For example, higher fees from originating risky loans increase u and hence, push towards more disclosure. Intuitively, higher fees make the bank more willing to risk failing the test, and so it is less responsive to a harder test. Hence, the regulator resorts to more disclosure.

We can also do comparative statics with respect to model parameters that affect the regulator's payoff from the risky asset. One example is the parameter L in Example 1, which represents the social loss that occurs when the bank's capital falls below a threshold.[21] A higher L reduces the regulator's payoff $v(\cdot)$, which has two effects. First, the social cost of gaming rises, and so partial disclosure becomes more costly. That is, the denominator on the left-hand side in (9) increases. Second, there is less to lose if the regulator fails the bank, and so a harder test becomes less costly. That is, the numerator on the left-hand side in (9) rises. Both effects push towards less disclosure. Hence, we obtain that partial disclosure is optimal if and only if L is sufficiently low, and that under some regularity conditions, if L is lower, the regulator reveals more information. More generally, parameter changes that lead

[21] In Example 1, the threshold is the face value of debt. See also Goldstein and Leitner (2018) for additional examples.

to a decrease in $v(\omega)$ push towards less disclosure, while changes that lead to an increase in $v(\omega)$ push towards more disclosure.

With respect to the probability of a crisis (parameter q in Example 1), this parameter affects both the bank's payoff u and the regulator's payoff v. First, a higher q reduces u, which pushes toward less disclosure. Second, a higher q reduces v, which also pushes towards less disclosure. Hence, under some regularity conditions, partial disclosure is optimal only if q is sufficiently low, and if q is lower, the regulator reveals more.

Finally, the effect of leverage on disclosure can go in either direction. On the one hand, higher leverage D can reduce the regulator's payoff from the risky asset because of the social cost of contagion, namely, the term $L(D - \omega)$ in Example 1. This pushes towards less disclosure. On the other hand, more leverage can increase the bank's payoff from the risky asset because debt holders bear more risk. This pushes towards more disclosure. Specifically, in Example 1, the bank's payoff from the risky asset reduces to $u(\omega) = 2(1 - q) + q\omega - 1 + q \max\{D - \omega, 0\}$, where the last term reflects the additional gain from risk shifting, which is clearly increasing in D.[22] If L is sufficiently small, this effect will dominate, and higher leverage would lead to more disclosure.

5.4 General Disclosure Rules

In the previous section, we focused on simple cutoff rules. Can a more general rule achieve a better outcome? In Appendix B, we show that, in general, the answer is yes, but only if we allow for the possibility that messages pool signals from disconnected intervals. We provide more details below.

A general disclosure rule is defined by a set of messages and a function that maps each signal s to a distribution over these messages. In our setting, sending a message is equivalent

[22]The first three terms in $u(\omega)$ reflect the NPV of the risky asset relative to the safe asset if the bank is all equity. The last term reflects the additional gain from risk shifting: with probability q a crisis occurs, and debt holders obtain $\min\{D, \omega\}$ instead of D. That is, debt holders lose $\max\{D - \omega, 0\}$. The loss to debt holders is a gain to equity holders.

to recommending an investment threshold ω_i to the bank, such that the bank invests if and only $\omega \geq \omega_i$. Hence, a general disclosure rule maps each signal to a distribution over investment threshold recommendations.

We first show that if the passing threshold s_p is set optimally, the optimal disclosure rule must be "single-peaked." That is, the recommendations that come from failing signals $s < s_p$ are increasing in s, and the recommendations that come from passing signals $s \geq s_p$ are decreasing in s. The idea is as follows. In our setting, the purpose of disclosure is to mitigate the bank's overinvestment. Hence, the binding constraint is that if the regulator recommends investment threshold ω_i, the marginal bank that observes ω_i is indifferent between investing and not investing. The most efficient way to satisfy this constraint is to give the recommendation ω_i from passing signals that the marginal bank believes are less likely, and failing signals that the marginal bank believes are more likely. Single peak then follows from Assumption 3 (MLRP).

While such a nonmonotone rule may have no precedent in regulatory stress testing, we see no reason why such a rule could not be implemented in practice. The main assumption is the regulator's commitment to follow the disclosure rule, but this commitment is necessary even for the simple cutoff rule (see Section 6).

However, we do admit that regulators may have a preference for simple rules. Interestingly, if we restrict attention to monotone rules such that each message corresponds to a signal interval (i.e., the regulator partitions the signal space into nonoverlapping intervals and reveals the interval to which the signal belongs), then the cutoff rule we developed in Sections 5.1-5.3 remains optimal. The idea behind this last result is as follows. Without loss of generality, the regulator can merge all the messages that reveal passing signals into one message. Moreover, since disclosure aims to mitigate the bank's overinvestment, it is optimal to merge all the other messages into a second message, which induces the bank to act more cautiously. We provide more details in Appendix B.3.

6 Discussion

In this section, we discuss some of the assumptions, interpretations, policy implications, and possible extensions of the model.

1. We assumed that the regulator has full flexibility in adjusting the passing threshold. The result that not revealing is strictly preferred to revealing (Theorem 1) relies on this assumption. However, in practice, the regulator may not have such flexibility, and so revealing might be preferred. One example is when the regulator must apply the same passing threshold to banks with different characteristics. This case could arise because of practical considerations, or because the bank's characteristics are privately observed by the bank. We show that if banks are sufficiently different from one another, then for some parameter values, revealing is strictly preferred to not revealing. We provide a formal statement of this result in Appendix A, but the intuition is simple. The benefit from not revealing the regulator's signal is that by choosing the passing threshold appropriately, the regulator can affect the bank's investment threshold in his favor. But if banks are very different from one another, it is impossible to calibrate the passing threshold to induce desired investment by everyone.

2. In our basic setting, the regulator can commit to pass or fail the bank according to a prespecified passing threshold. As we discussed in Section 4, this commitment helps the regulator achieve a better outcome (Lemma 3 and Lemma 5) but is not crucial for our main results. We also assumed that the regulator can commit to act according to a prespecified disclosure rule. The analysis of partial disclosure relies on this assumption. Without this commitment, the regulator would prefer to deviate ex-post, sending the message that induces the highest threshold. We believe that a commitment to follow prespecified rules, including a prespecified disclosure rule, is reasonable in the context of annual bank stress tests that are conducted by the

regulator during normal times.[23] The commitment outcome can also arise endogenously in models of repeated interaction.[24] However, in some applications (see item 10 below) the commitment to follow prespecified rules may be less plausible. Then, the relevant comparison would be simply between a secrecy regime and a fully transparent regime, which do not require this type of commitment.[25]

3. In practice, regulators conduct not only top-down tests, which rely on the regulator's model, but also bottom-up tests, which incorporate input from the bank. An important question is whether the regulator can obtain a better outcome by committing to a passing threshold that depends on the bank's report. The answer to this question is no, because the bank will always choose a report that leads to the highest probability of passing the test. However, this assumes that the regulator's disclosure policy does not depend on the bank's report. Extending our setting to a full mechanism design in which both the disclosure policy and passing threshold depend on the bank's report is left for future research.[26] For a formal treatment of bottom-up testing, see also Colliard (2019) and Leitner and Yilmaz (2019).

4. In our model, the regulator has two tools to provide incentives to the bank: the disclosure policy and the passing threshold. In practice, the regulator may be able to use additional tools, such as imposing penalties on banks that fail the test. We can incorporate this into our setting by assuming that the regulator can affect the bank's private cost of failing the test c. If the regulator has full control over c, he can get

[23]In contrast, during a crisis, when faced with extraordinary circumstances, the regulator may renege on commitments.

[24]See Mathevet, Pearce, and Stacchetti (2019) and Best and Quigley (2020).

[25]If not revealing is preferred (for a given s_p), the regulator cannot gain by revealing, because this will worsen the overinvestment problem. If revealing is preferred, we can assign out-of-equilibrium beliefs to rule out a deviation to not revealing.

[26]Kolotilin, Mylovanov, et al. (2017) can be viewed as a starting point for studying this case. They show that if the sender's signal space is not binary, then once we go beyond the case of linear payoffs, the sender can generally obtain a better outcome by conditioning messages on reports from the sender. Note that they focus only on information design, so in their setting, the sender cannot affect final payoffs by choosing a passing threshold.

arbitrarily close to the first best by setting c close to infinity, passing the banks almost surely, and not revealing anything. However, in the more realistic case in which the regulator does not have full control over the parameter c, the main results in our paper will continue to hold. More generally, if there are multiple tools to incentivize the banks, we believe that as long as these tools cannot be fully adjusted or are costly to adjust, the result that partial disclosure may be optimal will continue to hold.

5. We assumed that if the bank fails the test, the regulator's payoff is zero. This assumption is not crucial for our main results. What is crucial is that there is some social cost of providing incentives by increasing s_p. For example, we could assume that upon failing the bank, the regulator obtains $av(\omega)$ for some $a \in (0, 1)$. This case could reflect a situation in which the risky asset is transferred to other financial institutions that are less skilled at monitoring the asset but are also less systemically important. If a is not too large, it would still be costly to provide incentives by increasing s_p alone, and so partial disclosure will continue to be optimal.[27]

6. Policy makers have suggested that if the Fed model were to be published, then to counteract gaming, the minimum capital requirement would need to materially increase.[28] Our model suggests that this conclusion is only partially correct. In particular, for some parameter values, the optimal passing threshold under revealing is lower than that under not revealing: $s_p^R < s_p^{NR}$. For example, this could happen if the bank's cost of failing c is low, so the regulator needs to set a very high s_p^{NR} to reduce overinvestment.[29] As we illustrated in Example 3, the passing threshold s_p could represent minimum capital requirements or the severity of the stress scenario. Hence, holding the severity of the stress scenario fixed, our model suggests that revealing the regula-

[27] For example, under the assumptions of Figure 1 and assuming $c = 0.1$, partial disclosure is optimal whenever $a < 0.76$.

[28] See the departing speech by Fed Governor Daniel Tarullo: https://www.federalreserve.gov/newsevents/speech/tarullo20170404a.htm.

[29] E.g., in Figure 1, $s_p^R = 0.5$, and for a sufficiently low c, $s_p^{NR} > 0.5$.

tor's model could actually necessitate a reduction in minimum capital requirements. Alternatively, holding the capital requirements fixed, revealing the regulator's model could necessitate a less severe stress scenario.

7. A widely expressed concern is that disclosing the Fed's models could increase correlations in asset holdings among banks subject to the stress tests (i.e., the largest banks), making the financial system more vulnerable to adverse financial shocks. An extension of our model would suggest that this concern is also valid if the Fed just illustrates how its models work on hypothetical loan portfolios, as under the new policy discussed in the introduction. In particular, the proposed hypothetical portfolios could serve as benchmark portfolios in which too many banks invest, leading to correlated investment. So just as in our basic model, in which the bank could underinvest in a socially valuable risky assets by choosing the safe asset for which the test results are predictable, banks could also underinvest in their idiosyncratic risky portfolios, for which the test results are unpredictable, and overinvest in the benchmark risky portfolio, for which the test results are predictable.

8. A related concern is that revealing the regulator's models will cause banks to exert less effort in developing their own models. A simple extension in which the bank needs to incur a fixed cost to obtain its private signal about ω would imply that the bank will incur this cost only if the regulator does not reveal signal. However, a complete setting that incorporates information production by the bank or by the regulator is beyond the scope of this paper. We believe that in general the conclusions will depend on how

we model information production.[30,31]

9. An interesting question is how the optimal disclosure regime changes with respect to the information in the regulator's signal. A more informative signal makes it easier to incentivize the bank, which pushes towards secrecy, but a more informative signal also pushes towards revealing, because the regulator can use his information to force actions without the need to rely on the bank's information. Hence, the relationship between the informativeness of the regulator's signal and the preferred disclosure regime need not be monotone. To illustrate this, we consider a sequence of signals that becomes less informative in the sense of Blackwell (1953). That is, each signal is a garbling of the previous signal. For a fixed passing threshold, we can construct examples in which if the level of garbling is intermediate, not revealing is preferred to revealing, but if the level of garbling is either very high or very low, revealing is preferred.[32]

10. Finally, our setting is an example of a principal-agent problem in which an informed but biased agent takes an action on behalf of a partially informed principal, who can respond to the agent's action after an evaluation process that is based on the principal's private information. In our setting, the agent is the bank and the principal is the regulator, but our setting can also fit other applications. For example, the agent could be a financial advisor and the principal could be a wealthy individual. Alternatively, the agent could be the firm's manager and the principal could be the firm's board of directors. Hence, our results suggest that in some cases, the individual could benefit by

[30] For example, one could think of a setting in which the bank can generate one of two signals: an informative signal that gives the actual realization of ω, or a less informative signal, whose only purpose is to predict the test outcome; e.g., tell whether s is above or below the passing threshold. If the cost of obtaining the second signal is sufficiently low compared to the cost of obtaining the more informative signal, the outcome might be that if the regulator does not reveal his model, the bank generates only the second signal. In this case, revealing the regulator's model could generate a better outcome by saving the inefficient information production by the bank.

[31] See also Leitner and Yilmaz (2019), who show that under some conditions, it is optimal to allow banks to produce two models: a less informative for regulation and a more informative model for their own investment decisions.

[32] See the end of Appendix D.

not always sharing his views with the financial advisor about a new investment strategy but replace the advisor if the latter suggests an investment that is deemed too risky by the individual. Similarly, the board could benefit by not expressing their opinions while the manager is working on a strategic plan but use their opinions to disapprove the plan if its value is deemed too low. Moreover, our comparative statics imply that remaining silent (revealing less) is more beneficial when the agent's or principal's payoff from the risky action is lower.

Crucially, in our setting, the agent has a safe action whose payoff does not depend on the hidden evaluation process. In particular, the regulator cannot force the bank to switch from the safe asset to the risky asset. This assumption is natural in the banking application, but may not describe situations in which a principal can force an agent to take more risk. In addition, in our setting, the principal's signal represents hard information, and so the commitment to follow a prespecified rule that is based on the signal is natural. In other applications, the signal may reflect soft information, such as the board's opinion. In this case, the commitment assumption may be less natural. As we saw in item 2 above, the commitment to follow a disclosure rule is crucial for some of our results, but the commitment to follow a prespecified pass/fail rule is not crucial.

7 Conclusion

We study whether a regulator should reveal his stress test model to banks before conducting the test. We also explore the interaction between the regulator's disclosure policy and another regulatory tool that can be used to incentivize banks: the threshold for passing the test.

We show that if the regulator has full flexibility in adjusting the passing threshold, not revealing is always preferred to revealing. However, if the regulator cannot freely adjust the passing threshold, revealing may be preferred. Finally, if the regulator can commit to act according to a disclosure policy that goes beyond just revealing or not revealing, then for some parameter values, some disclosure is optimal even if the regulator can fully adjust the passing threshold. If we restrict attention to monotone disclosure rules, a simple cutoff rule is optimal. Otherwise, optimal disclosure is single peaked. We also derive comparative statics and policy implications, and offer applications beyond stress tests.

Our paper leaves open several questions that could be explored in future work. For example, our framework is static, but because regulators continually update their models, it would be interesting to explore the optimal dynamic disclosure policy. Our framework also assumes the regulator's signal is one dimensional. It would be interesting to explore the case in which the bank can invest in multiple assets, and the regulator's model takes the form of multiple signals that predict the value of each asset.

Appendix

A Heterogeneous Banks

In this appendix, we extend our basic model to analyze the case in which the regulator must apply the same passing threshold to banks with different characteristics.

Suppose the bank's private cost of failure c is a random variable with a CDF H. The bank observes the realization of c but the regulator does not. Recall that under not revealing, the bank expects to pass the test with probability $p(\omega) = 1 - F(s_p|\omega)$. Rearranging (3), it follows that the bank invests in state ω if and only if $c \leq [F(s_p|\omega)^{-1} - 1]u(\omega)$, i.e., with probability $I(\omega, s_p) \equiv H([F(s_p|\omega)^{-1} - 1]u(\omega))$. Extending the logic of Lemma 2, we obtain that the regulator's payoff under not revealing is:

$$V_{NR} = \int_{\omega \geq \underline{\omega}} I(\omega, s_p)[1 - F(s_p|\omega)]v(\omega)g(\omega)d\omega. \tag{A1}$$

The payoff under revealing does not depend on H and is given by $V_R(s_p^R)$, as in Lemma 2. In the special case in which H has all of the mass on a particular c, H is a step function, and (A1) reduces to (5).

To formalize the idea of banks that are sufficiently different from one another, we examine a sequence of distributions H_i that are median-preserving spreads in c, with a limiting distribution that places half the mass on $c = 0$ and half the mass on $c = \infty$.[33] We show that if $V_R(s_p^R)$ is sufficiently high, as we make precise in the proof, then in the limit, revealing is preferred.

Proposition A1. *Consider a sequence $\{H_i\}_{i=1}^{\infty}$ of distribution functions satisfying (i) H_{i+1} is a median-preserving spread of H_i for all $i \in N$; and (ii) $\lim_{i\to\infty} H_i(c) = \frac{1}{2}$ for all $c > 0$.*

[33] H_b is a *median-preserving spread* of H_a if H_a and H_b have the same median m, $H_b(x) \geq H_a(x)$ for all $x \leq m$, and $H_b(x) \leq H_a(x)$ for all $x \geq m$, with a strict inequality for at least one x.

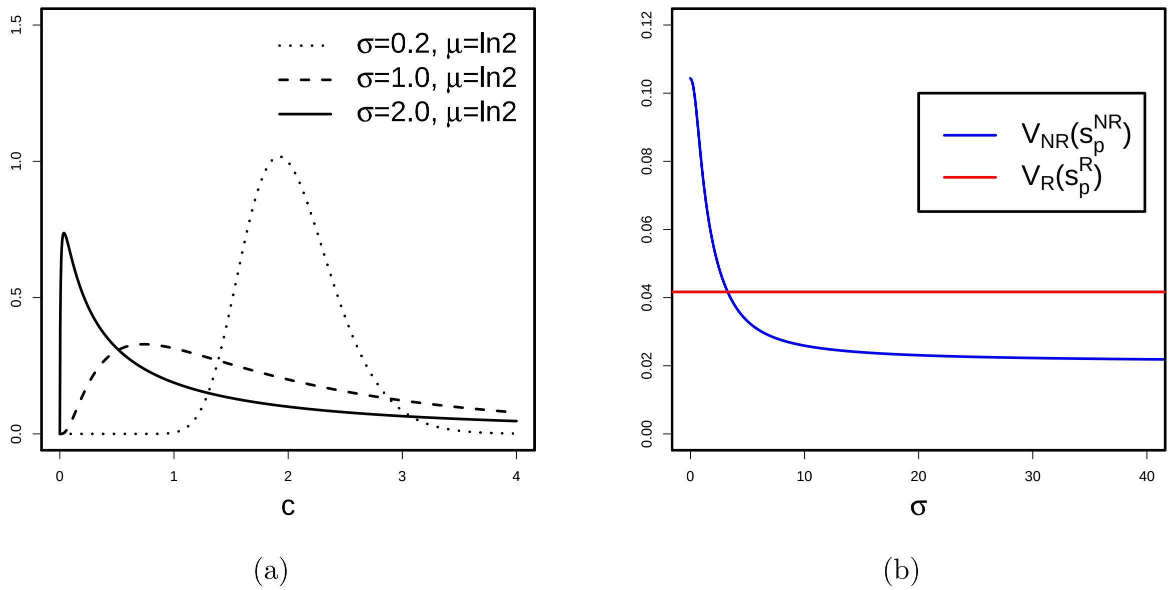

Figure A1: Panel (a) shows the density of c, when its distribution is lognormal with parameters $\mu = \ln 2$ and various values of σ. Panel (b) shows the regulator's payoff under revealing, $V_R(s_p^R)$, and under not revealing, $V_{NR}(s_p^{NR})$, as a function of σ.

Then if $V_R(s_p^R)$ is sufficiently high, revealing is strictly preferred to not revealing for high enough i.

Figure A1 illustrates the result in Proposition A1 for the case in which H is lognormal with parameters $\mu = \ln 2$ and various values of σ, which amounts to fixing the median of H at 2 and increasing uncertainty by increasing σ. For a very low level of uncertainty, not revealing is strictly optimal. For a very high level of uncertainty, revealing is strictly optimal.

B General Disclosure

In this appendix, we solve for an optimal disclosure rule. To avoid technical issues, we assume that ω and s are drawn from finite sets Ω and S. We denote the elements of Ω by $\omega_1 < \omega_2 < .... < \omega_n$, assume that $\omega_r \in \Omega$, and let i_r denote the $i \in \{1, ..., n\}$ such that

$\omega_i = \omega_r$. We use $f(s|\omega)$ and $g(\omega)$ to denote probability mass functions. A *disclosure rule* is defined by a finite set of messages M and a function h that maps each signal $s \in S$ to a distribution over messages. We let $h_m(s)$ denote the probability that the regulator sends message m upon observing s. ($\sum_{m \in M} h_m(s) = 1$ for every $s \in S$.)

B.1 Regulator's Problem

As in Section 4.1, we first show that the bank follows a cutoff rule, investing if and only if the state ω is above some threshold. We denote the decision to not invest by the threshold $\omega_{n+1} > \omega_n$ and let $\Omega' \equiv \Omega \cup \{\omega_{n+1}\}$. Formally:

Lemma B1. *For any disclosure rule* (M, h)*, there exists a function* $\omega : M \to \Omega'$ *such that if the regulator sends message* $m \in M$*, the bank invests if and only if* $\omega \geq \omega(m)$*.*

Lemma B1 implies that sending a message is equivalent to sending an investment recommendation $\omega_i \in \Omega'$ such that the bank invests if and only if $\omega \geq \omega_i$.

Using a "revelation principle" we can assume, without loss of generality, that the regulator sends only recommendations that the bank obeys.[34] The obedience constraints are that if the bank observes state ω, and the regulator recommends investment threshold ω_i, then if $\omega < \omega_i$, the bank cannot gain by investing and if $\omega \geq \omega_i$, the bank cannot gain by not investing. In a slight abuse of notation, we let $h_i(s)$ denote the probability that the regulator recommends $\omega_i \in \Omega'$ upon observing s. We let $v_i(s) \equiv \sum_{\omega \geq \omega_i} v(\omega) f(\omega|s)$.

Lemma B2. *The regulator's problem reduces to choosing a set of functions* $\{h_i : S \longrightarrow [0,1]\}_{i=1,\ldots,n+1}$ *to maximize*

$$\sum_{s \geq s_p} f(s) \sum_{i=1}^{n} v_i(s) h_i(s) \tag{B1}$$

[34] See Bergemann and Morris (2019).

such that

$$u(\omega_{i-1})\sum_{s\geq s_p} f(s|\omega_{i-1})h_i(s) - c\sum_{s<s_p} f(s|\omega_{i-1})h_i(s) \leq 0 \qquad i = 2, .., n+1 \tag{B2}$$

$$\sum_{i=1}^{n+1} h_i(s) = 1 \qquad s \in S. \tag{B3}$$

Equation (B1) is the regulator's expected payoff if the bank follows the regulator's recommendations. In particular, conditional on observing a failing signal $s < s_p$, the payoff is zero, and conditional on observing a passing signal $s \geq s_p$ and sending recommendation ω_i, the payoff is $v_i(s)$. Equation (B2) says that if the regulator recommends investment threshold ω_i, the bank cannot gain by investing upon observing the lower state ω_{i-1}. To see that, note that by Bayes' rule, the probability of passing the test conditional on state ω and recommendation ω_i is

$$p_i(\omega) \equiv \frac{\sum_{s\geq s_p} f(s|\omega)h_i(s)}{\sum_s f(s|\omega)h_i(s)}. \tag{B4}$$

So the bank cannot gain from investing in state ω_{i-1} if and only if

$$u(\omega_{i-1})p_i(\omega_{i-1}) - c[1 - p_i(\omega_{i-1}] \leq 0, \tag{B5}$$

which reduces to equation (B2). Equation (B3) simply says that conditional on observing a signal, the regulator sends a recommendation with probability 1. In the proof, we show that the solution to the linear programming problem above also satisfies the other obedience constraints.

B.2 Single-Peak Property

If the regulator sets s_p such that $\omega_{NR}(s_p) \geq \omega_r$, a cutoff rule is optimal. In particular, Lemma 4 implies that there exists $s_d < s_p$, such that if the regulator sends the high message, the bank invests if and only if $\omega \geq \omega_r$, and if the regulator sends the low message, the bank does not invest. Hence, a cutoff rule can implement the regulator's ideal investment threshold ω_r whenever the bank passes the test. However, setting s_p such that $\omega_{NR}(s_p) \geq \omega_r$ cannot be optimal because the regulator would gain by reducing s_p.[35]

The rest of this section focuses on the case in which $\omega_{NR}(s_p) < \omega_r$. In this case, no disclosure leads the bank to act too recklessly, and the purpose of disclosure is to make the bank act more cautiously.

Lemma B3. *If $\omega_{NR}(s_p) < \omega_r$, then under an optimal disclosure rule:*

1. *The incentive constraint (B2) is satisfied with equality.*
2. *The regulator never recommends the bank to underinvest. That is, $h_i(s) = 0$, for every $s \in S$ and every $i > i_r$.*

The next proposition shows that optimal disclosure must be "single peaked." That is, recommended thresholds weakly increase for failing signals $s < s_p$ and weakly decrease for passing signals $s \geq s_p$.

Proposition B1. *If $\omega_{NR}(s_p) < \omega_r$, then under an optimal disclosure rule, the following hold:*

1. *For every $\omega_i > \omega_j$ and $s < s' < s_p$, if $h_i(s) > 0$, then $h_j(s') = 0$.*
2. *For every $\omega_i < \omega_j$ and $s > s' \geq s_p$, if $h_i(s) > 0$, then $h_j(s') = 0$.*

The first part in Proposition B1 says that if the regulator recommends ω_i in some failing signal $s < s_p$, he never makes a lower recommendation in a higher failing signal. The second part says that if the regulator recommends ω_i in some passing signal $s \geq s_p$, he never makes a lower recommendation in a lower passing signal.

The idea behind Proposition B1 is as follows. To induce $\omega_i \leq \omega_r$, the regulator must pool failing signals with passing signals. From equation (B2), the most efficient way to do so is to increase the probability $h_i(s)$ in passing signals $s \geq s_p$ that have a low $f(s|\omega_{i-1})$ and failing signals $s < s_p$ that have a high $f(s|\omega_{i-1})$. In other words, the regulator recommends ω_i in

[35] In particular, if the set of signal S is sufficiently dense, we can apply similar logic as is in Lemma 3.

passing signals which a bank that observes ω_{i-1} thinks are relatively less likely, and in failing signals that a bank that observes ω_{i-1} thinks are relatively more likely. By MLRP, higher types ω_i place more weight on higher signals s. This leads to increasing recommendations in failing signals $s < s_p$ and decreasing recommendations in passing signals $s \geq s_p$. For passing signals, an additional force leads to decreasing recommendations. When the regulator observes a higher passing signal $s \geq s_p$, he is less worried about investment in low states ω, because by MLRP, these are less likely. Hence, he can recommend a lower investment threshold.

B.3 Monotone Rules

Proposition B1 implies that in general the cutoff disclosure rule from Section 5 need not be optimal. However, a cutoff rule is optimal if we restrict attention to "monotone disclosure rules" under which the regulator partitions the signal space into disjoint intervals and reveals the interval to which the signal belongs.

The idea is as follows. For any set of intervals, the regulator can obtain the same payoff by merging all the intervals that contain only failing signals into one interval, and all the intervals that contain only passing signals into a second interval. Hence, without loss of generality, there are at most three intervals: (i) an interval containing only failing signals, (ii) an interval containing only passing signals, and (iii) an interval containing both passing and failing signals.

If there are two intervals or less, a cutoff rule is optimal. Otherwise, we obtain a contradiction as follows. Suppose there are three intervals, which are defined by the cutoffs s_1 and s_2, where $s_1 < s_p < s_2$, and suppose that the message that the regulator sends upon observing $s \in (s_1, s_2)$ ("the middle message") induces the bank to invest if and only if $\omega \geq \omega'$. If $\omega' < \omega_r$, the regulator can obtain a better outcome by reducing s_1. If $\omega' > \omega_r$, the regulator can obtain a better outcome by increasing s_1. If $\omega' = \omega_r$, it is possible to obtain a better outcome by increasing s_2 and reducing s_1, so that the middle message continues to implement ω_r.

C No Commitment to Pass/Fail Rule

In this appendix, we provide more details for the case in which the regulator cannot commit to a pass/fail rule. We show that not revealing continues to be strictly preferred to revealing, and that for some parameter values partial disclosure continues to be optimal.

Revealing vs. not revealing Consider a pure strategy equilibrium in which the regulator passes the bank if and only if $s \in S_p$, where $S_p \subseteq S$. If the regulator reveals s, the bank invests if and only if $s \in S_p$. Since the bank's action conveys no additional information to the regulator about ω, the regulator passes the bank if and only if $E[v(\omega)|s] \geq 0$. By MLRP, $E[v(\omega)|s]$ is increasing in s, and so the regulator follows a cutoff rule. Moreover, the regulator's cutoff in this case is identical to the cutoff that he would follow if he could commit to a passing threshold. In both cases, the cutoff is the unique s that solves $E[v(\omega)|s] = 0$; in Section 4.1, we denoted this cutoff by s_p^R.[36]

Now suppose the regulator commits to not revealing his signal. Consider an equilibrium in which the bank invests if and only if $\omega \in \Omega_B$, where $\Omega_B \subseteq \Omega$ and $\Omega_B \neq \varnothing$ (empty set).[37] If the regulator passes the bank, he expects to get $E[v(\omega)|s, \omega \in \Omega_B]$, which is strictly increasing in s (MLRP). Hence, there exists a unique s_p', such that the regulator passes the bank if and only if $s \geq s_p'$. It then follows, as in Section 4.1, that there exists $\omega' > \underline{\omega}$, such that the bank invests if and only if $\omega \geq \omega'$. Hence, conditional on signal s, the regulator's payoff if he passes the bank is $E[v(\omega)|\omega \geq \omega', s]$, and s_p' is the unique s that solves $E[v(\omega)|\omega \geq \omega', s] = 0$. Since $E[v(\omega)|\omega \geq \omega', s] > E[v(\omega)|s]$, the cutoff under not revealing is lower than that under revealing; that is, $s_p' < s_p^R$. Moreover, $\omega' < \omega_r$, because if to the contrary $\omega' \geq \omega_r$, the regulator would always passes the bank, and the bank would

[36] To see why s_p^R solves the equation, take the first-order condition for V_R in Lemma 2.

[37] There are also some uninteresting equilibria in which the bank never invests (e.g., because the regulator never passes the bank or because the regulator passes the bank only if he observes a very high signal).

invest also if $\omega < \omega_r$. It then follows that the regulator's payoff under not revealing is higher. Formally,

$$\begin{aligned}
&\int_{s_p'} E[v(\omega)|\omega \geq \omega', s]P(\omega \geq \omega'|s)f(s)ds \qquad \text{(C1)}\\
&\qquad > \int_{s_p^R} E[v(\omega)|\omega \geq \omega', s]P(\omega \geq \omega'|s)f(s)ds\\
&\qquad = \int_{s_p^R} \int_{\omega'} v(\omega)f(\omega|s)f(s)ds\\
&\qquad > \int_{s_p^R} \int_{\underline{\omega}} v(\omega)f(\omega|s)f(s)ds\\
&\qquad = \int_{s_p^R} E[v(\omega)|s]f(s)ds.
\end{aligned}$$

The top line in (C1) is the payoff under not revealing, and the bottom line is the payoff under revealing. The second inequality follows from Assumption 2 about v and since $\omega' \in (\underline{\omega}, \omega_r)$.

Cutoff rules. We can extend the logic above to the case in which the regulator follows a cutoff disclosure rule. In this case, there exist thresholds s_H, s_L, such that the regulator follows the following strategy: after sending a high message, he passes the bank if and only if $s \geq s_H$, and after sending the low message, he passes the bank if and only if $s \geq s_L$. We can construct examples in which partial disclosure gives a better outcome than no disclosure, even without commitment to a pass/fail rule.

D Proofs for Main Text

Proof of Lemma 1. All the necessary steps for Part 1 and the cutoff rule in Part 2 are explained in the text. To see why ω_{NR} is continuous and increasing in s_p and c, apply the implicit function theorem to equation (3) with $\omega = \omega_{NR}$ and $p = 1 - F(s_p|\omega_{NR})$. Finally, Assumption 4 guarantees that $\omega_{NR} > \underline{\omega}$, because $u(\underline{\omega}) = 0$ implies that if the bank invests when $\omega = \underline{\omega}$, it obtains a negative payoff.

Proof of Lemma 2. The regulator's expected payoff under revealing is

$$\int_{s\geq s_p}\int_{\omega\geq\underline{\omega}} v(\omega)f(\omega|s)d\omega f(s)ds = \int_{\omega\geq\underline{\omega}} v(\omega)\int_{s\geq s_p} f(s|\omega)dsg(\omega)d\omega$$
$$= \int_{\omega\geq\underline{\omega}} v(\omega)[1-F(s_p|\omega)]g(\omega)d\omega$$

The payoff under not revealing is obtained in a similar fashion, but the integral starts in ω_{NR} rather than $\underline{\omega}$.

Proposition D1. *Suppose the passing threshold s_p is given and $s_p > \underline{s}$. If $V_R < 0$, then $\mathbf{V}_{NR} > \mathbf{V}_R$ for all $\omega_{NR} > \underline{\omega}$. If instead $V_R \geq 0$, let ω_I be the unique $\omega' > \underline{\omega}$ that solves*

$$\int_{\omega\geq\underline{\omega}}[1-F(s_p|\omega)]v(\omega)g(\omega)d\omega = \int_{\omega\geq\omega'}[1-F(s_p|\omega)]v(\omega)g(\omega)d\omega.$$

Then $\omega_I > \omega_r$, and the following hold: (i) if $\omega_{NR} > \omega_I$, the regulator strictly prefers to reveal; (ii) if $\omega_{NR} < \omega_I$, the regulator strictly prefers not to reveal; and (iii) if $\omega_{NR} = \omega_I$, the regulator is indifferent between revealing and not revealing.

Proof. V_{NR} is continuous in ω_{NR}. Moreover, $\frac{\partial V_{NR}}{\partial \omega_{NR}} = -v(\omega_{NR})[1-F(s_p|\omega_{NR})]$. Hence, $\frac{\partial V_{NR}}{\partial \omega_{NR}}$ has an opposite sign to $v(\omega_{NR})$. Hence, V_{NR} strictly increases from V_R over the interval $\omega_{NR} \in (\underline{\omega}, \omega_r)$ and strictly decreases to zero over the interval $(\omega_r, \bar{\omega}]$. The results follow.

Proof of Theorem 1. Suppose to the contrary that revealing is weakly preferred. Then $V_R(s_p^R) \geq V_{NR}(s_p^{NR}) \geq V_{NR}(s_p^R)$. Moreover, $V_R(s_p^R) \geq V_R(\bar{s}) = 0$. Hence, it follows from Proposition D1 that $\omega_{NR}(s_p^R) > \omega_r$. Since $\omega_{NR}(\cdot)$ is continuous, increasing, and equals $\underline{\omega}$ at $\underline{s}$, there exists $\hat{s} \in (\underline{s}, s_p^R)$ such that $\omega_{NR}(\hat{s}) = \omega_r$.[38] Using Assumptions 2 and 3, we then obtain the following contradiction:

$$V_R(s_p^R) = \int_{\omega\geq\underline{\omega}}[1-F(s_p^R|\omega)]v(\omega)g(\omega)d\omega < \int_{\omega\geq\omega_r}[1-F(s_p^R|\omega)]v(\omega)g(\omega)d\omega$$

[38]The proof also works if $\hat{s}$ is chosen such that $\omega_{NR}(\hat{s}) \in (\omega_r, \omega_I)$, where ω_I is the indifference point in Proposition D1 when $s_p = s_p^R$.

$$< \int_{\omega \geq \omega_r} [1 - F(\hat{s}|\omega)] v(\omega) g(\omega) d\omega = V_{NR}(\hat{s}) \leq V_{NR}(s_p^{NR}).$$

Proof of Lemma 3. As a preliminary, we show that $s_p^{NR} > \underline{s}$. Observe that

$$\frac{dV_{NR}}{ds_p} = \frac{\partial V_{NR}}{\partial s_p} + \frac{\partial V_{NR}}{\partial \omega_{NR}} \frac{\partial \omega_{NR}}{\partial s_p},$$

where

$$\frac{\partial V_{NR}}{\partial s_p} = -\int_{\omega \geq \omega_{NR}} f(s_p|\omega) v(\omega) g(\omega) d\omega = -f(s_p) \int_{\omega \geq \omega_{NR}} v(\omega) f(\omega|s_p) d\omega$$

and

$$\frac{\partial V_{NR}}{\partial \omega_{NR}} = -[1 - F(s_p|\omega_{NR}] v(\omega_{NR}).$$

Moreover, $\omega_{NR}(\underline{s}) = \underline{\omega}$. Hence, from Assumption 2, $\frac{\partial V_{NR}}{\partial \omega_{NR}}|_{\underline{s}} > 0$, and from Assumption 4, $\frac{\partial V_{NR}}{\partial s_p}|_{\underline{s}} > 0$. Since $\frac{\partial \omega_{NR}}{\partial s_p} \geq 0$ (Lemma 1), it then follows that $\frac{dV_{NR}}{ds_p}|_{\underline{s}} > 0$. Hence, $s_p^{NR} > \underline{s}$.

Part 1. Suppose to the contrary that $\omega_{NR}(s_p^{NR}) \geq \omega_r$. From Assumption 2, $\frac{\partial V_{NR}}{\partial \omega_{NR}}|_{s_p^{NR}} < 0$ and $\frac{\partial V_{NR}}{\partial s_p}|_{s_p^{NR}} < 0$. Hence, $\frac{dV_{NR}}{ds_p}|_{s_p^{NR}} < 0$, contradicting $s_p^{NR} > \underline{s}$.

Part 2. Suppose to the contrary that the regulator sets s_p^{NR}, such that it is ex-post (weakly) optimal to fail the bank for every $s \leq s_p$. Then $\int_{\omega \geq \omega_{NR}} v(\omega) f(\omega|s_p^{NR}) d\omega \leq 0$. Hence, $\frac{\partial V_{NR}}{\partial s_p}|_{s_p^{NR}} \geq 0$. Moreover from part 1, $\omega_{NR}(s_p^{NR}) < \omega_r$. Hence, from Assumption 2, $\frac{\partial V_{NR}}{\partial \omega_{NR}}|_{s_p^{NR}} > 0$. Hence, $\frac{dV_{NR}}{ds_p}|_{s_p^{NR}} > 0$, implying that $s_p^{NR} = \bar{s}$. But then $\omega_{NR}(s_p^{NR}) = \bar{\omega} > \omega_r$, which is a contradiction.

Proof of Lemma 4. Parts 1 and 2 are explained in the text. For part 3, note that ω_L solves equation (3) with $p = \Pr(s \geq s_p | s < s_d, \omega) = 1 - \frac{F(s_p|\omega)}{F(s_d|\omega)}$, and ω_H solves equation (3) with $p = \Pr(s \geq s_p | s \geq s_d, \omega) = \frac{1-F(s_p|\omega)}{1-F(s_d|\omega)}$. In both cases, the left-hand side in (3) is strictly increasing in ω and s_d. Hence, by the implicit function theorem, ω_L and ω_H are continuous and strictly decreasing in s_d. Finally, $\lim_{s_d \uparrow \bar{s}} \omega_L = \lim_{s_d \downarrow \underline{s}} \omega_H = \omega_{NR}$. Hence, $\omega_L(s_d, s_p) > \omega_{NR}(s_p) > \omega_H(s_d, s_p)$.

Proof of Theorem 2. For a given policy (s_d, s_p), denote the regulator's expected payoff by $V(s_d, s_p)$. Using Lemma 4, and since the regulator obtains nothing if he fails the bank, $V(s_d, s_p)$ is continuous and is given by:

$$\begin{cases} \int\limits_{s \geq s_p} \int_{\omega \geq \omega_H} v(\omega) f(\omega|s) f(s) d\omega ds & \text{if } s_d \in (\underline{s}, s_p) \\ \int\limits_{s \in (s_p, s_d)} \int_{\omega \geq \omega_L} v(\omega) f(\omega|s) f(s) d\omega ds + \int\limits_{s \geq s_d} \int_{\omega \geq \underline{\omega}} v(\omega) f(\omega|s) f(s) d\omega ds & \text{if } s_d \in (s_p, \bar{s}) \\ V_{NR}(s_p) & \text{if } s_d \in \{\underline{s}, \bar{s}\} \\ V_R(s_p) & \text{if } s_d = s_p \end{cases} \tag{D1}$$

From Theorem 1, full disclosure cannot be optimal. Hence, a sufficient condition for partial disclosure to be optimal is that it can improve on no disclosure: $\frac{\partial V(s_d, s_p)}{\partial s_d}|_{(s_d, s_p) = (\bar{s}, s_p^{NR})} < 0$. Observe that if $s_d > s_p$, then

$$\frac{\partial V(s_d, s_p)}{\partial s_d} = \int_{\omega \geq \omega_L} v(\omega) f(\omega|s_d) f(s_d) d\omega - \int\limits_{s \in (s_p, s_d)} \frac{\partial \omega_L}{\partial s_d} v(\omega_L) f(\omega_L|s) f(s) ds - \int_{\omega \geq \underline{\omega}} v(\omega) f(\omega|s_d) f(s_d) d\omega. \tag{D2}$$

Since $f(\omega|s) f(s) d\omega = f(s|\omega) g(\omega) d\omega$, this reduces to

$$\frac{\partial V(s_d, s_p)}{\partial s_d} = -\frac{\partial \omega_L}{\partial s_d} g(\omega_L) v(\omega_L) [F(s_d|\omega_L) - F(s_p|\omega_L)] - \int_{\underline{\omega}}^{\omega_L} v(\omega) f(s_d|\omega) g(\omega) d\omega. \tag{D3}$$

Since $\lim_{s_d \uparrow \bar{s}} \omega_L = \omega_{NR}$ and $F(\bar{s}|\omega_L) = 1$, the sufficient condition for partial disclosure reduces to (6). Finally, as in Lemma 2, we can show that if $s_d < s_p$, the regulator's payoff reduces to $\int_{\omega \geq \omega_H} v(\omega)[1 - F(s_p|\omega)] g(\omega) d\omega$. From Lemma 3 and Lemma 4, $\omega_H(s_d, s_p) < \omega_{NR}(s_p) < \omega_r$. It then follows from Lemma 2 and Assumption 2 that $V(s_d, s_p) < V_{NR}(s_p)$, for every (s_d, s_p). Hence setting $s_d \in (\underline{s}, s_p)$ cannot be optimal.

Proof of Lemma 5. Suppose the regulator optimally sets (s_d, s_p). From Theorem 2, we can assume without loss of generality that $s_d \geq s_p$. Hence,

$$\frac{\partial V(s_d, s_p)}{\partial s_p} = -f(s_p) \int_{\omega \geq \omega_L} v(\omega) f(\omega|s_p) d\omega - \frac{\partial \omega_L}{\partial s_p} \int_{s \in (s_p, s_d)} v(\omega_L) f(\omega_L|s) f(s) ds \tag{D4}$$

Following similar steps as in the proof of Lemma 4, we can show that $\frac{\partial \omega_L}{\partial s_p} > 0$. Moreover, we must have $s_p \in (\underline{s}, \bar{s})$, because $s_p = \underline{s}$ implies that $\omega_L = \underline{\omega}$ and $\frac{\partial V(s_d, s_p)}{\partial s_p} > 0$ (Assumptions 2 and 4), and if $s_p = \bar{s}$, the regulator's payoff is zero, which is less than what he obtains under no disclosure. In particular, from Lemma 3, the payoff conditional on the optimal s_p must be strictly positive. Hence, $s_p \in (\underline{s}, \bar{s})$ and $\frac{\partial V(s_d, s_p)}{\partial s_p} = 0$.

Part 1. Suppose to the contrary that $\omega_L(s_d, s_p) \geq \omega_r$. From Assumption 2, $v(\omega_L) > 0$. Hence, $\frac{\partial V(s_d, s_p)}{\partial s_p} < 0$, which is a contradiction.

Part 2. Suppose to the contrary that the regulator sets (s_d, s_p), such that it is always ex-post optimal to fail the bank when $s \leq s_p$. Then we must have $\int_{\omega \geq \omega_L} v(\omega) f(\omega|s_p) d\omega < 0$. From part 1 $\omega_L(s_d, s_p) < \omega_r$. Hence, $v(\omega_L) < 0$ (Assumption 2). Hence, $\frac{\partial V}{\partial s_p} > 0$, which is a contradiction.

Explanation for equation (7). From the proof of Lemma 5, $s_p \in (\underline{s}, \bar{s})$. Hence, the first-order condition for s_p reduces to

$$\int_{\omega \geq \omega_L} v(\omega) f(\omega|s_p) f(s_p) d\omega = -\frac{\partial \omega_L}{\partial s_p} v(\omega_L) g(\omega_L) [F(\omega_L|s_d) - F(\omega_L|s_p)] \tag{D5}$$

Combining (D5) with (6), we obtain (7).

Explanation for equation (8). The bank's investment threshold ω_L solves (3), where $p = \Pr(s \geq s_p | s < s_d | \omega_L) = 1 - \frac{F(s_p|\omega_L)}{F(s_d|\omega_L)}$. Rearranging terms, we obtain that $\frac{u(\omega)+c}{u(\omega)} = \frac{F(s_d|\omega_L)}{F(s_p|\omega_L)}$ and

$$H(s_d, s_p) \equiv [u(\omega) + c][F(s_d|\omega_L) - F(s_p|\omega_L)] - cF(s_d|\omega_L) = 0. \tag{D6}$$

By the implicit function theorem,

$$\frac{\partial \omega_L/\partial s_p}{-\partial \omega_L/\partial s_d}=\frac{\partial H/\partial s_p}{-\partial H/\partial s_d}=\frac{u(\omega)+c}{u(\omega)}\frac{f(s_p|\omega_L)}{f(s_d|\omega_L)}=\frac{F(s_d|\omega_L)}{F(s_p|\omega_L)}\frac{f(s_p|\omega_L)}{f(s_d|\omega_L)}. \tag{D7}$$

More details for Remark 2. Suppose s_p is given. Since ω_{NR} is decreasing in c (from $\bar{\omega}$ if $c \uparrow \infty$ to $\underline{\omega}$ if $c \downarrow 0$), there exists $\bar{c}$, such that if $c = \bar{c}$, no disclosure induces the bank to invest if and only if investment is socially optimal ($\omega_{NR} = \omega_r$). If $c < \bar{c}$, no disclosure induces the bank to act too recklessly ($\omega_{NR} < \omega_r$). In this case, if equation (6) holds for the given s_p, it is optimal to set $s_d > s_p$, so that if the regulator sends the low message, the bank acts more cautiously; and under some regularity conditions, it is optimal to disclose more if c is lower. In contrast, if $c > \bar{c}$, no disclosure induces the bank to act too cautiously ($\omega_{NR} > \omega_r$). In this case, there exists $\bar{s}_d < s_p$, such that if the regulator sets $s_d = \bar{s}_d$ and sends the high message, the bank invests if and only if $\omega \geq \omega_r$. So setting $s_d = \bar{s}_d$ is optimal. Moreover, if c increases, $\bar{s}_d$ increases, so the regulator reveals more of the failing signals.

Example (Informativeness of Regulator's Signal). The garbling we consider is a mixture between the original signal $s \in [0,1]$ that is drawn from a distribution $f(s|\omega)$ satisfying MLRP and a signal that is drawn from a uniform distribution on $[0,1]$, where the mixture weight is $\phi \in (0,1)$. Formally, define the stochastic transformation density $g(s'|s) \equiv \phi \cdot \delta(s'-s) + (1-\phi)\cdot 1$, where $\delta(\cdot)$ is the Dirac delta function. Then $\hat{f}(s|\omega) = \int_0^1 g(s|t)f(t|\omega)dt = \phi \cdot f(s|\omega) + (1-\phi)\cdot 1$ is the density of the garbled signal, and it also satisfies MLRP. If the original signal is garbled k times, the resulting density is $\hat{f}_k(s|\omega) = \phi^k \cdot f(s|\omega) + (1-\phi^k)\cdot 1$, so less weight is placed on $f(s|\omega)$ the more garbling occurs. We let $\alpha \equiv 1-\phi^k$ be a measure of uninformativeness, define $f_\alpha(s|\omega) = (1-\alpha)f(s|\omega) + \alpha \cdot 1$, and consider the regulator's disclosure policy as α increases from zero to 1.

For example, suppose $\Omega = [0,1]$, $S = [0,1]$, $u(\omega) = \omega^4$, $v(\omega) = \omega - 0.5$, $s_p = 0.5$, $c = 1.1$, and $f(s|\omega)$ is a truncated normal distribution with mean ω and standard deviation 0.1, truncated to the interval $[0,1]$. Numerical computations show in this case that not revealing is strictly preferred if uninformativeness α is in $(.07,.66)$, and revealing is strictly preferred otherwise. So the optimal disclosure regime is nonmonotonic in the uninformativeness of the signal.

E Proof for Appendix A

Proof of Proposition A1. For a given distribution H_i denote $I_i(\omega, s_p) \equiv H_i([F(s_p|\omega)^{-1} - 1]u(\omega))$. So $V_{NR}(s_p^{NR}) = \int_{\underline{\omega}}[1 - F(s_p^{NR}|\omega)]I_i(\omega, s_p^{NR})v(\omega)g(\omega)d\omega$. Also, as a preliminary, observe that from Proposition 2 and Assumption 2, $V_R(s_p^R) = \int_{\underline{\omega}}[1 - F(s_p^R|\omega)]v(\omega)g(\omega)d\omega \leq \int_{\omega_r} v(\omega)g(\omega)d\omega$. Moreover, $V_R(s_p^R) \geq V_R(\underline{s}) = E[v(\omega)]$, and $V_R(s_p^R) \geq V_R(\bar{s}) = 0$. Hence, $V_R(s_p^R) \geq \max\{E[v(\omega)], 0\}$. Hence, there exists $\nu \in [0,1]$ such that $V_R(s_p^R) = (1-\nu)\int_{\omega_r} v(\omega)g(\omega)d\omega + \nu \max\{E[v(\omega)], 0\}$.[39]

To prove the proposition, assume that $\nu < 1/2$.[40] So, $V_R(s_p^R) > 0$. Fix a small $\varepsilon > 0$. From the assumptions on $\{H_i\}_{i=1}^{\infty}$, there exists $N > 0$, such that $|I_i(\omega, s) - 1/2| < \varepsilon$ for all $i \geq N$, $\omega \in \Omega$, and $s \in [\underline{s} + \varepsilon, \bar{s} - \varepsilon]$. Suppose $i \geq N$. We will show that if ε is sufficiently small, then $V_{NR}(s_p^{NR}) < V_R(s_p^R)$ for any possible s_p^{NR}. Specifically, if $s_p^{NR} \in [\underline{s} + \varepsilon, \bar{s} - \varepsilon]$, then

$$\begin{aligned} V_{NR}(s_p^{NR}) &< \int_{\underline{\omega}}^{\omega_r} [1 - F(s|\omega)] \left(\frac{1}{2} - \varepsilon\right) v(\omega)g(\omega)d\omega + \int_{\omega_r}^{\bar{\omega}} [1 - F(s|\omega)] \left(\frac{1}{2} + \varepsilon\right) v(\omega)g(\omega)d\omega \\ &= \frac{1}{2}V_R(s_p^{NR}) + \varepsilon\left(\int_{\omega_r}^{\bar{\omega}} [1 - F(s|\omega)]v(\omega)g(\omega)d\omega - \int_{\underline{\omega}}^{\omega_r} [1 - F(s|\omega)]v(\omega)g(\omega)d\omega\right), \end{aligned}$$

[39] This equation says that $V_R(s_p^R)$ is a weighted average of two extremes: the payoff under a perfectly informative signal (weight $1-\nu$) and the payoff under a perfectly uninformative signal (weight ν). To see that, note that in the first case, the regulator can set the passing threshold so that the bank invests and passes the test if and only if $\omega \geq \omega_r$. So the regulator's payoff is $\int_{\omega \geq \omega_r} v(\omega)g(\omega)d\omega$. In the second case, the regulator either bans investment completely or always approves it. So the regulator's payoff is $\max\{E[v(\omega)], 0\}$.

[40] This assumption says that the weight on the payoff under the more informative signal, as explained in footnote 39, is at least $\frac{1}{2}$.

where the inequality follows from Assumption 2. Hence, for a small enough ε, $V_{NR}(s_p^{NR}) < V_R(s_p^{NR}) \leq V_R(s_p^R)$. Next, if $s_p^{NR} > \bar{s} - \varepsilon$, then $V_{NR}(s_p^{NR}) \leq \int_{\omega_r}^{\bar{\omega}}[1 - F(s_p^{NR}|\omega))]v(\omega)g(\omega)d\omega$, which is less than $V_R(s_p^R)$, for a small enough ε. Finally, if $s_p^{NR} < \underline{s} + \varepsilon$, then $I_i(\omega, s_p^{NR}) \geq I_i(\omega, \underline{s} + \varepsilon) > \frac{1}{2} - \varepsilon$. Moreover, if ε is small enough, $1 - F(s_p^{NR}|\omega) > \frac{2\nu}{1-2\varepsilon}$ (because $v < 1/2$ implies that $\frac{2\nu}{1-2\varepsilon} < 1$). Hence, $[1 - F(s_p^{NR}|\omega)]I_i(\omega, s_p^{NR}) > \nu$. Hence,

$$
\begin{aligned}
&= V_{NR}(s_p^{NR}) < \nu \int_{\underline{\omega}}^{\omega_r} v(\omega)g(\omega)d\omega + \int_{\omega_r}^{\bar{\omega}} v(\omega)g(\omega)d\omega \\
&= \nu \int_{\underline{\omega}}^{\bar{\omega}} v(\omega)g(\omega)d\omega + (1-\nu)\int_{\omega_r}^{\bar{\omega}} v(\omega)g(\omega)d\omega \\
&\leq \nu \max\{E[v(\omega)], 0\} + (1-\nu)\int_{\omega_r}^{\bar{\omega}} v(\omega)g(\omega)d\omega = V_R(s_p^R).
\end{aligned}
$$

This concludes the proof.

F Proofs for Appendix B

Proof of Lemma B1. Consider a disclosure rule (M, h). We first show that for any $m \in M$ such that $h_m(s) > 0$ for some $s \geq s_p$, the posterior distribution $f(s|\omega, m)$ satisfies MLRP. That is, if $\omega' > \omega$, the ratio $f(s|\omega', m)/f(s|\omega, m)$ is strictly increasing in s. To see this, observe that $h_m(s) = f(m|s) = f(m|s, \omega) = f(m|s, \omega')$. Hence, from Bayes' rule,

$$
\frac{f(s|\omega', m)}{f(s|\omega, m)} = \frac{f(m|s, \omega')f(s|\omega')}{f(m|\omega')} \frac{f(m|\omega)}{f(m|s, \omega)f(s|\omega)} = \frac{f(s|\omega')}{f(m|\omega')} \frac{f(m|\omega)}{f(s|\omega)},
$$

which is increasing in s by Assumption 3.

We now prove the lemma. A bank that observes state ω and receives message m forms posterior belief $p_m(\omega) \equiv \Pr(s \geq s_p|\omega, m)$. So the payoff from investing in the risky asset is $u_m(\omega) \equiv u(\omega)p_m(\omega) - c[1 - p_m(\omega)]$. If the bank receives a message m such that $h_m(s) > 0$ for some $s \geq s_p$, then by the result above, $p_m(\omega)$ is strictly increasing in ω. Hence, $u_m(\omega)$ is strictly increasing in ω, and the bank follows a cutoff rule. If instead the bank receives a message m such that $h_m(s) = 0$ for every $s \geq s_p$, then $p_m(\omega) = 0$. Hence, $u_m(\omega) = -c$, implying the bank does not invest regardless of the value of ω.

Proof of Lemma B2. If the bank follows the regulator's recommendations, the regulator's payoff is (B1), as explained in the text. The regulator's problem is to choose a disclosure rule to maximize (B1) such that the bank follows the recommendations. The obedience constraints are as follows. When the bank observes state $\omega \in \Omega$ and obtains recommendation $\omega_i \in \Omega'$, it expects to pass the test with probability $p_i(\omega)$. So the payoff from investing is $u_i(\omega) \equiv u(\omega)p_i(\omega) - c[1 - p_i(\omega)]$. The bank will follow recommendation ω_{n+1}, if and only if $u_{n+1}(\omega) \leq 0$ for every $\omega < \omega_i$, and it will follow recommendation $\omega_i \in \Omega$ if and only if (i) $u_i(\omega) \geq 0$ for every $\omega \in [\omega_i, \omega_n]$, and (ii) $u_i(\omega) \leq 0$ for every $\omega < \omega_i$. By the proof of Lemma B1, $u_i(\omega)$ is either strictly increasing in ω or equals to $-c$. Hence, the obedience constraints reduce to

$$u_i(\omega_i) \geq 0 \qquad \text{if } i \in \{1, ..., n\} \tag{F1}$$

$$u_i(\omega_{i-1}) \leq 0 \qquad \text{if } i \in \{2, ..., n+1\}. \tag{F2}$$

(F2) reduces to (B2), using (B4). Moreover, if the regulator never recommends ω_i (so $h_i(s) = 0$ for every $s \in S$), then (B2) is clearly satisfied Hence, a solution to the regulator's problem satisfies (B2) and (B3).

To complete the proof, we show that if $\{h_i(s)\}_{i,s}$ solves the problem in Lemma B2, then (F1) is satisfied. Suppose to the contrary that there exists $i \in \{1, ..., n\}$ such $u_i(\omega_i) < 0$. If $u_i(\omega_k) < 0$ for every $k \in \{i + 1, ..., n\}$, let $j = n + 1$. Otherwise, let j be the lowest $k \geq i + 1$ such that $u_i(\omega_k) \geq 0$. If $j \leq i_r$, we obtain a contradiction because the regulator can increase his payoff without violating the constraints by recommending ω_j instead of ω_i. If $j > i_r$, there exists a function $q(s)$ that satisfied the following: (i) $u(\omega_{i_r-1}) \sum_{s \geq s_p} f(s|\omega_{i_r-1})q(s) - c \sum_{s<s_p} f(s|\omega_{i_r-1})q(s) = 0$; (ii) for every $s < s_p$, $q(s) \leq h_i(s)$, with at least one strict inequality; and (iii) for every $s \geq s_p$, $q(s) = h_i(s)$. The regulator can increase his payoff without violating the constraints if in each state s, instead of recommending ω_i with probability $h_i(s)$, he recommends ω_r with probability $q(s)$ and ω_{n+1} with probability $h_i(s) - q(s)$.

Proof of Lemma B3. Suppose $\{h_i(s)\}_{i,s}$ solves the regulator's problem.

1. We first show there exists $s' \geq s_p$ such that $h_{i_r}(s') < 1$. If not, then

$$\begin{aligned} &u(\omega_{i_r-1}) \sum_{s \geq s_p} f(s|\omega_{i_r-1}) \cdot 1 - c \sum_{s < s_p} f(s|\omega_{i_r-1}) h_i(s) \\ &\geq u(\omega_{i_r-1}) \sum_{s \geq s_p} f(s|\omega_{i_r-1}) - c \sum_{s < s_p} f(s|\omega_{i_r-1}) > 0, \end{aligned}$$

where the strict inequality follows since $\omega_{NR} < \omega_r$. But this contradicts (B2). Hence, there also exists $j \neq i_r$ such that $h_j(s') > 0$. Next, we show that $\{h_i(s)\}_{i,s}$ satisfies (B2) with equality. For $i = i_r$, this is true because otherwise, the regulator could improve his payoff without violating the constraints by raising $h_{i_r}(s')$ by some $\Delta > 0$ and reducing $h_j(s')$ by $\Delta > 0$. Now suppose to the contrary that (B2) is slack for some $i \notin i_r$. Then there exists $s'' < s_p$ such that $h_i(s'') > 0$. The regulator can reduce $h_i(s'')$ by some $\Delta > 0$, raise $h_{i_r}(s'')$ by Δ, raise $h_{i_r}(s')$ by $\Delta' \equiv \Delta c f(s''|\omega_{i_r-1})/[u(\omega_{i_r-1}) f(s'|\omega_{i_r-1})]$, and reduce $h_j(s')$ by Δ'. If Δ is small enough, then (B2) and (B3) continue to hold, and the regulator increases his payoff by $f(s')\Delta'(v_{i_r}(s') - v_j(s')) > 0$, contradicting optimality.

2. Suppose to the contrary there exists $s \in S$ and $i > i_r$ such that $h_i(s) > 0$. Then there exists $s''' < s_p$ such that $h_i(s''') > 0$. But then the regulator can improve his payoff without violating the constraints by adjusting the disclosure rule in the manner described in part 1. Hence a contradiction.

Proof of Proposition B1. As a preliminary, observe that the Lagrangian of the regulator's problem is $\mathcal{L} = \sum_{s \geq s_p} f(s) \sum_{i=1}^{n} v_i(s) h_i(s) - \sum_{i=2}^{n+1} \lambda_i [u(\omega_{i-1}) \sum_{s \geq s_p} f(s|\omega_{i-1}) h_i(s) - c \sum_{s < s_p} f(s|\omega_{i-1}) h_i(s)] - \sum_{s \in S} \mu_s \sum_{i=1}^{n+1} h_i(s)$, where $\lambda_i \geq 0$ and μ_s are the lagrange multipliers on (B2) and (B3), respectively. From the definition of $v_i(s)$ and since $f(\omega|s)f(s) =$

$f(s|\omega)g(\omega)$, we obtain that

$$d_i(s) \equiv \frac{\partial \mathcal{L}}{\partial h_i(s)} = \begin{cases} \lambda_i c f(s|\omega_{i-1}) - \mu_s & \text{if } s < s_p \\ \sum_{\omega \geq \omega_i} v(\omega) f(s|\omega) g(\omega) - \lambda_i u(\omega_{i-1}) f(s|\omega_{i-1}) - \mu_s & \text{if } s \geq s_p. \end{cases} \tag{F3}$$

The first-order conditions imply the following: (i) if $h_i(s) = 1$, then $d_i(s) \geq 0$; (ii) if $h_i(s) \in (0,1)$, then $d_i(s) = 0$; and (iii) if $h_i(s) = 0$, then $d_i(s) \leq 0$. Also note that by Lemma B3, $h_i(s) > 0$ implies that $\omega_i \leq \omega_r$. We are now ready to prove the proposition.

Suppose $\omega_i > \omega_j$ and $h_i(s) > 0$. If $\lambda_i = 0$, we can show that the regulator never recommends ω_j. Specifically, for every $s'' \geq s_p$, $d_j(s'') - d_i(s'') = \sum_{\omega_j}^{\omega_{i-1}} v(\omega) f(s|\omega) g(\omega) - \lambda_j u(\omega_{i-1}) f(s|\omega_{i-1})$, which is negative by Assumptions 2 and 1. Hence, $d_j(s'') < d_i(s'')$, and the first-order conditions imply that $h_j(s'') = 0$ for every $s'' \geq s_p$. Lemma (B3) part 1 then implies that $h_j(s') = 0$ for every $s' < s_p$. The rest of the proof assumes $\lambda_i > 0$.

1. If $s < s' < s_p$, the first order conditions imply that $d_i(s) \geq 0 \geq d_j(s)$. Hence, $\lambda_i c f(s|\omega_{i-1}) \geq \lambda_j c f(s|\omega_{j-1})$. Hence, $\lambda_i c \frac{f(s|\omega_{i-1})}{f(s|\omega_{j-1})} \geq \lambda_j c$. From MLRP, $\frac{f(s'|\omega_{i-1})}{f(s'|\omega_{j-1})} > \frac{f(s|\omega_{i-1})}{f(s|\omega_{j-1})}$. Hence, $\lambda_i c \frac{f(s'|\omega_{i-1})}{f(s'|\omega_{j-1})} \geq \lambda_j c$. Hence, $d_i(s') > d_j(s')$. Hence, $h_j(s') = 0$.

2. Suppose $s > s' \geq s_p$. Following the logic in part 1, it is sufficient to show that $d_i(s) \geq d_j(s)$ implies that $d_i(s') > d_j(s')$. This follows from Assumptions 2 and 1, MLRP, and the observation that

$$\frac{d_i(s) - d_j(s)}{f(s|\omega_{j-1})} = \sum_{\omega_i}^{\omega_{j-1}} v(\omega) \frac{f(s|\omega)}{f(s|\omega_{j-1})} g(\omega) - \lambda_i u(\omega_{i-1}) \frac{f(s|\omega_{i-1})}{f(s|\omega_{j-1})} + \lambda_j u(\omega_{j-1}).$$

References

Acharya, Viral V, Allen N Berger, and Raluca A Roman (2018). "Lending implications of US bank stress tests: Costs or benefits?" *Journal of Financial Intermediation* 34, pp. 58–90.

Aghion, Philippe and Jean Tirole (1997). "Formal and real authority in organizations". *Journal of Political Economy* 105.1, pp. 1–29.

Alonso, Ricardo and Niko Matouschek (2008). "Optimal delegation". *Review of Economic Studies* 75.1, pp. 259–293.

Baker, Scott R, Nicholas Bloom, and Steven J Davis (2016). "Measuring economic policy uncertainty". *Quarterly Journal of Economics* 131.4, pp. 1593–1636.

Bergemann, Dirk and Stephen Morris (2019). "Information design: A unified perspective". *Journal of Economic Literature* 57.1, pp. 44–95.

Best, James and Daniel Quigley (2020). "Persuasion for the long run". *Available at SSRN 2908115*.

Blackwell, David (1953). "Equivalent comparisons of experiments". *Annals of Mathematical Statistics*, pp. 265–272.

Board of Governors of the Federal Reserve System (2021). "Dodd-Frank Act Stress Test 2021: Supervisory Stress Test Methodology".

Bouvard, Matthieu, Pierre Chaigneau, and Adolfo De Motta (2015). "Transparency in the financial system: Rollover risk and crises". *Journal of Finance* 70.4, pp. 1805–1837.

Chakraborty, Archishman and Bilge Yılmaz (2017). "Authority, consensus, and governance". *Review of Financial Studies* 30.12, pp. 4267–4316.

Colliard, Jean-Edouard (2019). "Strategic selection of risk models and bank capital regulation". *Management Science* 65.6, pp. 2591–2606.

Corona, Carlos, Lin Nan, and Gaoqing Zhang (2019). "The Coordination Role of Stress Tests in Bank Risk-Taking". *Journal of Accounting Research* 57.5, pp. 1161–1200.

Cortés, Kristle R, Yuliya Demyanyk, Lei Li, Elena Loutskina, and Philip E Strahan (2020). "Stress tests and small business lending". *Journal of Financial Economics* 136.1, pp. 260–279.

Dessein, Wouter (2002). "Authority and communication in organizations". *Review of Economic Studies* 69.4, pp. 811–838.

Dogra, Keshav and Keeyoung Rhee (2018). "Stress tests and model monoculture". *Available at SSRN 3208233*.

Ederer, Florian, Richard Holden, and Margaret Meyer (2018). "Gaming and strategic opacity in incentive provision". *RAND Journal of Economics* 49.4, pp. 819–854.

Faria-e-Castro, Miguel, Joseba Martinez, and Thomas Philippon (2017). "Runs versus lemons: information disclosure and fiscal capacity". *Review of Economic Studies* 84.4, pp. 1683–1707.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Keith Kuester, and Juan Rubio-Ramírez (2015). "Fiscal volatility shocks and economic activity". *American Economic Review* 105.11, pp. 3352–84.

Flannery, Mark J (2019). "Transparency and Model Evolution in Stress Testing". *Available at SSRN 3431679*.

Fuchs, William (2007). "Contracting with repeated moral hazard and private evaluations". *American Economic Review* 97.4, pp. 1432–1448.

Gissler, Stefan, Jeremy Oldfather, and Doriana Ruffino (2016). "Lending on hold: Regulatory uncertainty and bank lending standards". *Journal of Monetary Economics* 81, pp. 89–101.

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

— (2020). "Stress Tests Disclosure: Theory, Practice, and New Perspectives". *(to appear in) Financial Stress Testing*. Ed. by Farmer, Kleinnijenhuis, Wetzer, and Schuermann. Cambridge University Press.

Goldstein, Itay and Haresh Sapra (2014). “Should banks’ stress test results be disclosed? An analysis of the costs and benefits”. *Foundations and Trends in Finance* 8.1, pp. 1–54.

Grenadier, Steven, Andrey Malenko, and Nadya Malenko (2016). “Timing decisions in organizations: Communication and authority in a dynamic environment”. *American Economic Review* 106.9, pp. 2552–81.

Guo, Yingni and Eran Shmaya (2019). “The interval structure of optimal disclosure”. *Econometrica* 87.2, pp. 653–675.

Harris, Milton and Artur Raviv (2008). “A theory of board control and size”. *Review of Financial Studies* 21.4, pp. 1797–1832.

Holmstrom, Bengt (1982). *On the theory of delegation.* Northwestern University.

Huang, Jing (2019). “Optimal Stress Tests in Financial Networks”. *Available at SSRN 3465799.*

Inostroza, Nicolas (2019). “Persuading Multiple Audiences: Disclosure Policies, Recapitalizations, and Liquidity Provision”.

Inostroza, Nicolas and Alessandro Pavan (2017). *Persuasion in global games with application to stress testing.*

Jehiel, Philippe (2015). “On transparency in organizations”. *Review of Economic Studies* 82.2, pp. 736–761.

Jensen, Michael C (1986). “Agency costs of free cash flow, corporate finance, and takeovers”. *American Economic Review* 76.2, pp. 323–329.

Julio, Brandon and Youngsuk Yook (2012). “Political uncertainty and corporate investment cycles”. *Journal of Finance* 67.1, pp. 45–83.

Kamenica, Emir and Matthew Gentzkow (2011). “Bayesian persuasion”. *American Economic Review* 101.6, pp. 2590–2615.

Kolotilin, Anton (2018). “Optimal information disclosure: A linear programming approach”. *Theoretical Economics* 13.2, pp. 607–635.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li (2017). “Persuasion of a privately informed receiver”. *Econometrica* 85.6, pp. 1949–1964.

Kolotilin, Anton and Alexander Wolitzky (2020). "Assortative Information Disclosure".

Lazear, Edward P (2006). "Speeding, terrorism, and teaching to the test". *Quarterly Journal of Economics* 121.3, pp. 1029–1061.

Leitner, Yaron (2014). "Should regulators reveal information about banks". *Federal Reserve Bank of Philadelphia Business Review, Third Quarter.*

Leitner, Yaron and Bilge Yilmaz (2019). "Regulating a model". *Journal of Financial Economics* 131.2, pp. 251–268.

Levin, Jonathan (2003). "Relational incentive contracts". *American Economic Review* 93.3, pp. 835–857.

Levit, Doron (2020). "Words speak louder without actions". *The Journal of Finance* 75.1, pp. 91–131.

MacLeod, W Bentley (2003). "Optimal contracting with subjective evaluation". *American Economic Review* 93.1, pp. 216–240.

Mathevet, Laurent, David Pearce, and Ennio Stacchetti (2019). "Reputation and information design". New York University.

Milgrom, Paul R (1981). "Good news and bad news: Representation theorems and applications". *Bell Journal of Economics*, pp. 380–391.

Orlov, Dmitry, Pavel Zryumov, and Andrzej Skrzypacz (2018). "Design of macro-prudential stress tests".

Parlatore, Cecilia and Thomas Philippon (2018). "Designing Stress Scenarios". New York University.

Quigley, Daniel and Ansgar Walther (2020). "Inside and outside information: reverse unraveling and stress test design". *Unpublished Working Paper.*

Shapiro, Joel D and Jing Zeng (2019). "Stress testing and bank lending". *Available at SSRN 3432291.*

Williams, Basil (2017). "Stress tests and bank portfolio choice". *WP, New York University.*