Inside and Outside Information

Daniel Quigley
University of Oxford

Ansgar Walther
Imperial College London and CEPR*

April 2023

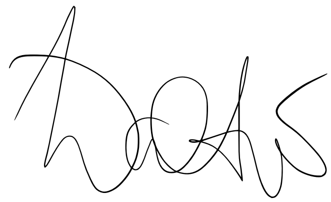
Abstract

We study an economy with financial frictions in which a regulator designs a test that reveals *outside information* about a firm's quality to investors. The firm can also disclose verifiable *inside information* about its quality. We show that the regulator optimally aims for "public speech and private silence", which is achieved with tests that give insiders an incentive to stay quiet. We fully characterize optimal tests by developing tools for Bayesian persuasion with incentive constraints, and use these results to derive novel guidance for the design of bank stress tests, as well as benchmarks for socially optimal corporate credit ratings.

*Quigley: Exeter College and Department of Economics, University of Oxford, Oxford OX1 3DP, UK. Email: daniel.quigley@economics.ox.ac.uk. Walther: Imperial College Business School, London SW7 2AZ, UK, and CEPR. Email: a.walther@imperial.ac.uk. We are grateful to our discussants Pierre Chaigneau, Nadya Malenko, Tim MacQuade, and Alessio Piccolo for their feedback. We also thank Heski Bar-Isaac, James Best, Vince Crawford, Peter Eso, Nicolas Inostroza, Ian Jewitt, Emir Kamenica, Alex Kohlhas, Paul Klemperer, Meg Meyer, Ines Moreno de Barreda, David Myatt, Larry Samuelson, Joel Shapiro, Anton Tsoy, Victoria Vanasco, Adrien Vigier, Selma Walther, Lucy White, Peyton Young, and audiences at various seminars and conferences, for their comments. A previous version of this paper was circulated under the title "Crowding out disclosure".

Disclosure Statement: Ansgar Walther

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.



Disclosure Statement: Daniel Quigley

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

During the financial crisis of 2007-9, regulators such as the Federal Reserve used stress tests to alleviate investors' uncertainty about the banking system. These stress tests are a form of *outside* information about firms that is released to investors. In addition, banks and other firms have wide discretion over their public corporate disclosures, which reveal *inside information* to investors. This potential interplay between inside and outside information appears frequently in financial markets. For instance, credit ratings agencies provide outside information about borrowers, who can also make their own inside disclosures.

In this paper, we analyze how outside information should be designed in an economy with financial frictions. Unlike the growing literature on this question (e.g., Goldstein and Leitner, 2018; Orlov et al., 2021; Inostroza and Pavan, 2022) we take into account that inside disclosures may respond endogenously to the design of outside information. We show that accounting for these responses changes the optimal design of outside information both qualitatively and quantitatively.

To motivate our analysis, imagine the Fed is designing a stress test to reveal outside information about the unobserved quality θ of a bank's assets. Investors will run on the bank if their expectation of asset quality, after observing the test result, drops below a critical threshold t. Suppose the Fed wants to minimize the probability of runs. Using tools from the literature on Bayesian persuasion, it is easy to show that—in the absence of inside information—the optimal policy is a simple pass/fail test, which splits banks around a quality threshold $\theta^\star$. Banks with quality $\theta < \theta^\star$ publicly "fail", resulting in a run with probability one. Banks with assets of quality $\theta \geq \theta^\star$ "pass" and avoid a run: The failure threshold $\theta^\star$ is chosen so that the pass grade is just credible enough to save the bank, that is $E\left[\theta|\theta \geq \theta^\star\right] = t$. Intuitively, this test achieves higher welfare than, say, full disclosure, because it pools strong types $\theta \geq t$ with as many weak types as possible, thus avoiding runs on vulnerable banks.[1]

Suppose now that banks can verifiably disclose θ at some cost in anticipation of the test result, and consider the incentives of strong banks. If they do not disclose, they will avoid a run but investors will regard their assets as worth only t. As long as banks care about market perceptions beyond the immediate need to avoid a run, disclosure may be a best response for the strong, so long as disclosure costs are not too large. Once strong types disclose, investors' expectations conditional on seeing a "pass" result without inside disclosure drop below the threshold t, because the strongest types of bank are no longer pooled into the

[1] In richer models of bank runs (e.g., Inostroza, 2019), multiple failing grades can be optimal, but the unique passing grade is still defined and motivated in the same way.

passing grade. Thus, all non-disclosing banks now face a run. This is an instance of the classical "unraveling" of inside information, as first described by Grossman and Hart (1980). Indeed, whenever the simple pass/fail test incentivizes any inside disclosures among passing banks, our results show it cannot be optimal. In this sense, naive application of pass/fail tests do not tackle financial instability in an efficient way.

We study a class of models with financial frictions, deriving tests that optimally account for firms' incentives to make disclosures. As we demonstrate in a series of examples, our general model nests models of bank runs along the lines of our example above (e.g., Diamond and Dybvig, 1983; Morris and Shin, 2000), a model of corporate financing with contracting frictions as in Holmstrom and Tirole (1997), in which outside information can be interpreted as the release of credit ratings, and richer treatments of financial instability such as Inostroza (2019). The important common feature of these applications is that financial constraints begin to bind whenever investors' expectations about an indicator θ of underlying quality drop below a critical threshold. Moreover, welfare is typically a convex, increasing function of investors' expectations when financial constraints are binding.

In our model, a regulator can commit to an arbitrary test that reveals outside information about θ, as in the literature on Bayesian persuasion (e.g., Kamenica and Gentzkow, 2011; Dworczak and Martini, 2019). After learning θ, firms / banks can either stay quiet or make additional inside disclosures in anticipation of the test result. For simplicity, we present the case where the firm may verifiably disclose its type θ at a cost, but the main insights extend to richer message spaces.

A key result in our characterization of optimal tests is that it is always optimal for the regulator to encourage *public speech and private silence*: without loss, the regulator can focus on designing tests that sustain "opaque" equilibria, in which firms make no inside disclosures. This result is reminiscent of a revelation principle, but requires a slightly different proof. In addition to arguing that the regulator can replicate any disclosures that firms would otherwise make, we must verify that opacity is a (perfect Bayesian) equilibrium of the subsequent messaging game between firms and investors. This result contributes an interesting economic intuition in its own right, but also makes the design problem substantially more tractable. Indeed, the regulator's problem now boils down to a Bayesian persuasion problem subject to a series of incentive compatibility constraints, which ensure that no type of firm wants to deviate from silence.

While this finding helps to simplify the search for an optimal test, the regulator's problem is still non-trivial and cannot be solved with existing methods for Bayesian persuasion. We show how existing techniques can be extended to our model with incentive compatibility constraints. After outlining this methodological contribution, we provide a full characterization of optimal tests.

We show that optimal tests still have a threshold property, in the following sense: all firms with $\theta < \theta^\star$ fail the test, meaning that they face binding financial frictions with probability one; all remaining firms pass. However, a variety of passing grades are employed: the regulator designs a standard "pass" grade that lifts firms just above the threshold where financial constraints bind, but also a range of novel grades that pass strong firms "with flying colors" and raise investors' expectations discretely above the critical threshold. These novel grades are released to preserve strong firms' incentives to stay quiet. The regulator trades off the cost of releasing these grades, which force her to raise the failure threshold $\theta^\star$ relative to the benchmark without inside disclosures, against firms' incentive compatibility constraint. We show that the optimal way to resolve this trade-off is to assign a stochastic grade, by which strong types are revealed with just enough probability to discourage inside disclosures.

While there are typically multiple implementations of optimal tests in our model, we prove that the key economic properties we have highlighted are necessary features of *any* optimal test: all optimal tests have the threshold property described above and, moreover, assign stochastic grades to types of firms with a binding incentive constraint. That is, in an optimally designed test, strong firms should be confident of passing the test, but unsure about the precise grade they will receive.

We supplement this characterization with a set of comparative statics. We show that the incentive constraints associated with private silence become tighter from the regulator's perspective if i) inside disclosures become less costly; ii) disclosure messages become "more verifiable" in a natural sense;[2] or iii) the payoffs to firms' managers become more sensitive to investors' beliefs. In each of these cases, the regulator is more constrained, which means that she must ultimately raise the threshold $\theta^\star$ below which firms fail. We also show that, in each of these scenarios, the regulator must make the optimal tests more informative, in that investors' posterior expectations undergo a mean-preserving spread. Finally, we show that, under some mild regularity conditions, the failure threshold also increases when investors' prior beliefs become more pessimistic.

In an instructive special case of our model, we show that private silence can be the *only* motivation for releasing outside information. Consider the case where investors' prior beliefs are benign, in the sense that financial constraints would not bind under the prior. In this scenario, strong firms can still have a strict incentive to make disclosures in order to separate themselves in investors' expectations. The unraveling argument again applies, and a positive

[2]We derive this comparative static in an extension with general messages spaces. We say that messages are "more verifiable" if a (weakly) smaller set of types is able to send each possible message – see Section 3.5 for details.

mass of firms can face binding frictions in equilibrium. We show that a regulator can do better by designing a test that fails a small proportion of weak firms but provides strong firms with an incentive to stay quiet. Therefore, while the regulator would prefer to release no information at all, and to preserve prior beliefs, she optimally decides to release some information that is geared towards preventing unraveling. This example strongly conveys the idea that inside disclosures motivate the release of more accurate outside information.

Our baseline model focuses on *pre-emptive* disclosures, which the firm must choose before the test result is realized. As we discuss in Section 2.1, the available evidence suggests that this case is relevant in practice. Indeed, firms have an apparent incentive to disclose in advance of public news, so as to prevent reputational and legal costs, as well as adverse market responses (e.g., Skinner, 1994; Bischof et al., 2021; Kays, 2022). However, it is also interesting to explore the robustness of our results under alternative timing assumptions. On one hand, we demonstrate that optimal tests in our baseline model continue to be optimal when verifiable disclosures need to be pre-prepared before the test result is known, but can be released ex post. On the other hand, we show that the principle of *public speech and private silence* extends to models in which firms can respond with flexible ex-post disclosures, and derive a full characterization of optimal tests in this case. Optimal tests with ex post disclosures are deterministic, and characterized by a combination of *fine-tuned* pooling of weak with strong types, according to a negative assortative matching function, and *coarse* pooling of a continuum of intermediate types into a single grade.

Finally, we can also consider a broader design problem in which the regulator can privately communicate with firms, showing them a preview of the test result before firms decide on inside disclosures. We show that it is always optimal for the regulator not to send a preview – another dimension of private silence. In terms of our model environment, this result also means that the timing of the game we assume, namely, that firms have to prepare disclosures before observing the test result, is an optimal arrangement from the policy-maker's perspective.

We further discuss the applied insights that arise from our theoretical analysis. In the headline application to bank runs and stress tests, we provide a full characterization of optimal, granular stress tests. In practice, regulators in the crisis of 2007-9 released information about banks' losses in adverse macroeconomic scenarios. Details of these losses were published even for banks that "passed" the test in the sense that the regulator did not find a capital shortfall. One distinctive feature of optimal stress tests in our model is that they maintain some constructive ambiguity, assigning stochastic grades to strong banks; interestingly, such uncertainty has been a characteristic of the Federal Reserve's stress tests.[3] By

[3] See "Fed 'Stress Tests' Still Pose Puzzle to Banks", *Wall Street Journal*, March 12, 2015.

contrast, weak banks face certainty about the grades they will receive, and never enjoy more surplus than is needed to marginally avoid a bank run. Our results also imply that it is strictly suboptimal to give banks a preview of stress test results, because this only serves to increase their incentives to make private disclosures to investors.

Our comparative statics provide additional guidance for regulators in practice. For instance, we demonstrate that optimal tests must be more informative (and more granular) when bank managers have high-powered incentives that depend on current market perceptions, and when it is relatively easy for the bank to make verifiable disclosures. Moreover, the special case described above suggests that at the "eve of a crisis", when investors have an intermediate degree of pessimism, it can be valuable to run stress tests with the sole purpose of disincentivizing inside disclosures.

In the context of credit ratings, our results are the first to show how a broad spectrum of optimal ratings emerges endogenously as a result of financial constraints. Our characterization of optimal tests can be interpreted as an ideal set of credit ratings, which involve "junk" grades that effectively fail the firm and lead to binding financing constraints, bunching of firms at "investment grade" ratings that just avoid financial constraints, and "premium" grades that transfer additional surplus to the firm in order to prevent excessive corporate disclosures. The mechanism behind this design is related to Daley et al. (2020), who show that accurate ratings can crowd out intermediaries' incentives to acquire costly signals of credit quality. However, our characterization is novel in that it emphasizes the granularity of optimally *designed* ratings, while the typical signal structure in the existing literature is an exogenously structured, binary test. Moreover, a range of recent work emphasizes additional incentive problems arising from the interaction of firms, ratings agencies and investors (e.g., Skreta and Veldkamp, 2009; Bolton et al., 2012; Manso, 2013). We view our result as an interesting benchmark describing the type of ratings that maximize overall surplus, which may be useful to future studies on the efficiency properties of ratings markets.

In addition to the work we have cited above, this paper relates to the broader literature on verifiable disclosures (e.g., Milgrom, 1981; Verrecchia, 1983; Shin, 2003; Acharya et al., 2011; Bond and Zeng, 2021). We focus on an information designer's goal of discouraging or "crowding out" private disclosures. Similar crowding-out effects have been studied in a few related papers, including Goldstein and Yang (2017), Einhorn (2018) and Frenkel et al. (2018) in models with verifiable disclosures, as well as Feltovich et al. (2002) and Daley and Green (2014) in the context of signaling games and Gigler and Hemmer (1998) in a cheap talk environment, but these papers do not contain an analysis of optimal information design.

In the literature on Bayesian persuasion and information design, Gentzkow and Kamenica (2017) and Li and Norman (2021) study multi-sender persuasion. As they assume all senders have ex ante commitment power, these models cannot address the novel incentive constraints introduced by interim disclosures. Our work is also connected to Guo and Shmaya (2019), who derive the optimal design of (outside) information subject to the constraints implied by the receiver's private information, but our methodology and solution take a very different shape.

In the accounting literature, Friedman et al. (2015) study the interaction between persuasion and disclosure in a model with binary signals, which precludes most of the effects we emphasize, while DeMarzo et al. (2019) and Bertomeu et al. (2021) study the incentives for firms to block the release of outside information in a setting without additional inside disclosures. We contribute a novel characterization of optimal, granular outside information that anticipates the associated incentives to disclose in a general framework.

The remainder of the paper is structured as follows: Section 2 describes our general model environment and provides a series of examples that highlight potential applications in finance. Section 3 characterizes optimal tests – it is designed so that the reader interested in applications can skip the technical Subsection 3.2, which develops our solution method. Section 4 discusses the applied insights arising from our results, and Section 5 concludes.

2 Model

We begin by describing a general information design problem in the presence of inside disclosures. We then present a series of examples to highlight the applications of our framework to banking and corporate finance.

2.1 General Environment

A firm and a regulator interact with a population of (risk-neutral) investors. The firm has assets with fundamental value $\theta \in \left[\underline{\theta}, \bar{\theta}\right] \equiv \Theta$, where $\bar{\theta} > \underline{\theta} \geq 0$. All agents have a prior belief that θ is drawn from a continuous cumulative distribution $F(\theta)$.

Inside and Outside Information: The regulator designs a *test* which reveals a public signal s of θ about the firm's assets. We refer to this signal as *outside information.* A test is defined by a set of possible realizations $s \in S$ and a family $\{G(s \mid \theta)\}_{s \in S, \theta \in \Theta}$ of conditional cumulative distributions of test results.

The firm, after privately observing θ, sends investors a message $m \in \{\theta, \emptyset\}$. We refer to this message as *inside information.* The message $m = \theta$ denotes full and verifiable disclosure of θ, since it is impossible for an agent with type $\theta' \neq \theta$ to send this message. The message

$m = \emptyset$ stands for non-disclosure or silence. Silence does not convey any verifiable information because it is feasible for all types.

Investors' beliefs about θ, conditional on inside information m and outside information m, determine the posterior expected value of the firm's assets, which we define as

$$z = E_G [\theta | s, m] . \tag{1}$$

We present our analysis for the simple "all-or-nothing" message space because it conveys most of the novel economics in our analysis, and because it makes for a particularly clear discussion of equilibrium selection, which we provide below. However, the proofs of all our main results extend to more general message spaces, which we introduce in Appendix A.1.

Regulator's Preferences: The regulator's preferences are described by the function

$$w(z, \theta) = \theta (1 - \lambda (z)) , \tag{2}$$

where $\lambda(z) \geq 0$ is a decreasing, concave function on the interval $[\underline{\theta}, t)$, and equal to zero on the interval $\left[t, \bar{\theta}\right]$. As we demonstrate in more detail below, this functional form naturally measures social surplus in economies with *financial frictions*, which begin to bind when market valuations fall below a critical threshold t. The function $\lambda(z)$ is interpreted as the fraction of assets lost due to financial frictions. We assume throughout the paper that we are not in the trivial case where $\lambda(z) \equiv 0$.

For our analysis of test design, it is also useful to define the regulator's (interim) expected utility given investors' updated belief z about the firm's assets. Indeed, taking expectations of (2), and noting that $E[\theta|z] = z$ by definition, we can define

$$v(z) \equiv E[w(z, \theta)|z] = z (1 - \lambda(z)) . \tag{3}$$

We show below that, when designing information, the regulator focuses on maximizing the expected value of $v(z)$. Importantly, whenever financial frictions bind, this objective function is convex in market expectations z and, otherwise, it is linear. Figure 1 illustrates the shape of $v(z)$ in three different parametric cases, which also map to our applied examples below.

Firm's Preferences: The firm's preferences are described by the utility function

$$u(z, \theta) - c(m) . \tag{4}$$

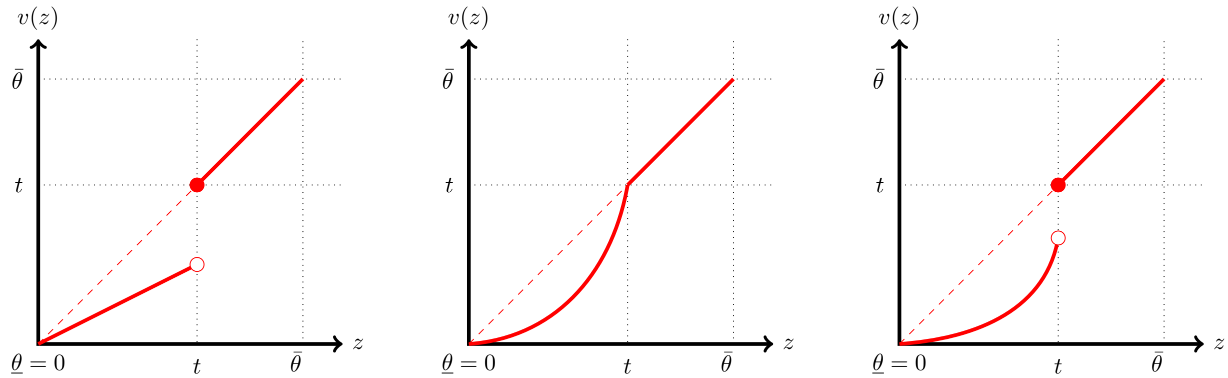


Figure 1: **Regulator's Expected Utility.** Each panel shows the regulator's expected utility $v(z)$ as a function of investors' expectation $z = E[\theta|s, m]$, as defined in Equation (3). We have set the lowest type $\underline{\theta}$ equal to zero for illustrative purposes. Panel (a) shows piecewise linear payoffs, with $\lambda(z) = \bar{\lambda} * 1\{z < t\}$, where $\bar{\lambda}$ is a constant. Notice that $\lambda(z)$ has a jump at $z = t$, which makes $v(z)$ discontinuous. Our model of bank runs in Example 1 induces payoffs of this type with $\bar{\lambda} = 1$. Panel (b) shows payoffs for which $\lambda(z)$ is a strictly convex and continuous function on the interval $[0, t]$, and equal to zero on the interval $\left[t, \bar{\theta}\right]$. These payoffs are induced by our model of corporate investment in Example 2. Panel (c) shows a possible combination of the previous two cases, with both convex and discontinuous payoffs on the interval $[0, t]$, which arises in Example 3.

The first term in the firm's payoff is defined as

$$u(z, \theta) = [\phi z + (1 - \phi)\theta](1 - \lambda(z)). \tag{5}$$

This expression again accounts for the fraction $\lambda(z)$ of asset values lost due to financial frictions. The firm's utility is given by a weighted average of investors' expected value z of the remaining assets, and their true fundamental value θ, with $\phi > 0$ denoting the weight on current market beliefs. For example, this formulation measures the profits of a firm who sells a fraction ϕ of its assets to outside investors and retains remaining fraction until maturity.[4] Alternatively, Equation (5) could be interpreted as the utility of a manager whose compensation depends on both current and future asset values.

The second term in the firm's payoff captures the direct cost $c(m)$ of the firm's message to investors. We normalize the cost of silence to $c(\emptyset) = 0$ and assume that verifiable disclosure is costly with $c(\theta) > 0$ for all θ. Notice that by making the verifiable disclosure $m = \theta$, the

[4]Daley et al. (2020) have characterized the case in which ϕ is endogenous and serves as a signal of asset quality. By contrast, we treat ϕ as a constant and focus on direct disclosures m about asset quality.

firm can always secure the following payoff with probability one:

$$\bar{u}(\theta) \equiv u(\theta, \theta) - c(\theta). \tag{6}$$

The full-disclosure payoff $\bar{u}(\theta)$ plays the role of the firm's outside option in our analysis. Intuitively, regardless of the test that the regulator designs, the firm can always guarantee itself a payoff of $\bar{u}(\theta)$ after learning its type. Therefore, in any equilibrium that the regulator can induce with any test, the firm's utility cannot be less than $\bar{u}(\theta)$.

Equilibrium Definition: For a given test G,[5] a perfect Bayesian equilibrium is defined by a messaging strategy for the firm and an expected value z for all realizations $\{s, m\}$. In equilibrium, the firm maximizes its expected utility after observing θ, and investors compute the expectation in Equation (1) using Bayes' rule on the equilibrium path.

Notice that our baseline model focuses on pre-emptive disclosures: The firm chooses its message m before observing the realization s of outside information. We focus on the case of pre-emptive disclosures because it captures frictions that are relevant in our applications. In reality, attempts to delay material disclosures can be costly. For instance, in the US, the Exchange Act Section 10(b) and the SEC's Rule 10b-5 require that firms disclose material information in a timely manner. Skinner (1994) finds that firms do indeed make pre-emptive disclosures ahead of mandatory earnings reports, arguing that failure to do so would expose them to costly fines (as well as other sanctions), litigation risks, and even reputational damage.[6] Similarly, Kays (2022) finds evidence that firms voluntarily disclose information that pre-empts and supplements mandatory tax disclosures, arguing that this mitigates reputational risk. For these reasons, we find it reasonable to analyze situations in which firms find it suboptimal to delay publication of material facts available to them.

Likewise, it may also be costly to wait until after a regulatory release to start acquiring evidence that rebuts it. Unlike "cheap talk", verifiable information is often difficult to prepare at short notice, for example, when it needs to be certified or audited by third parties.[7] On the

[5]With a small abuse of notation, we write G to concisely describe the regulator's choice of test, which consists of the signal space S and the conditional distributions $G(s|\theta)$ described above.

[6]In the US, public companies report material facts (outside of regular annual and quarterly filings) via the SEC's Form 8-K disclosures, which must be submitted within 4 business days of a triggering event. While litigation costs are particularly acute for 'bad news', Skinner (1994) notes that failures to disclose good news have resulted in litigation too. He finds pre-emption of both bad and good news in the data, though the latter is less surprising.

[7]Perhaps for this reason, the Exchange Act grants special status to "periodic" disclosures in its timely disclosure rules. For instance, public companies may publish their annual reports to shareholders within 4 months of fiscal year end (https://www.sec.gov/rules/sro/nysemkt/2017/34-80619-ex5.pdf). Similarly, the FCA allows additional time for periodic disclosures, where "immediate public disclosure ... would impact on the orderly production and release of the report and could result in the incorrect assessment of the

other hand, financial markets respond very quickly to bad news, such as credit downgrades for a firm, or an announcement that a bank has failed its stress test. Indeed, Bischof et al. (2021) conclude from an empirical analysis of banks' disclosures that "markets do not wait for bank disclosures when a crisis starts to unfold and market conditions deteriorate." As a result, a firm concerned about the outcome of a regulatory test may need to prepare evidence in advance. As material disclosure rules likely apply to these findings, such a strategy is effectively a pre-emptive decision to disclose. These arguments notwithstanding, it is interesting to understand how test design must respond when firms can make verifiable disclosures after the realization of the test result. In section 3.7, we discuss extensions of the model in which *(i)* firms are able to wait until after the test to make a costly disclosure, and *(ii)* firms must pre-emptively decide whether to prepare verifiable information (i.e. pay a cost to acquire it), but may delay its release until after the test (or even suppress it altogether).

Optimal Tests: The regulator's problem is to choose a test G that maximizes her expected utility subject to the constraint that the firm's message m, and hence investor's expectations z, will be determined endogenously in equilibrium. We assume that the regulator can fully commit to the test design G. Best and Quigley (2020) show that such commitment can be motivated by repetition over time or by applying the same test to many firms; in reality, stress testing authorities may rely on both to some degree. Beyond this, we remark in section 4 that the regulator's optimal test is robust when she can commit to the test itself, but may be tempted to supplement it with additional information ex post

We assume that, if a test G induces multiple equilibria, the regulator can select her preferred one. While this is common in the literature on mechanism and information design, it is important nonetheless to confirm that the regulator-preferred equilibrium is not 'unreasonable'—for instance, in the sense of failing to satisfy refinements such as the Intuitive Criterion and D1 (Cho and Kreps, 1987a; Banks and Sobel, 1987). In the baseline model, the regulator-preferred equilibrium trivially survives these tests. As we will see, the regulator prefers equilibria with "private silence", in which all types of firm choose $m = \emptyset$. In such an equilibrium, the only "off-path" events are full disclosures, $m = \theta$ for $\theta \in [\underline{\theta}, \overline{\theta}]$. Beliefs in such events are uniquely defined, since only one type can send such a message. Hence, the silent equilibrium automatically passes the Intuitive and D1 criteria. Similarly, the regulator's preferred equilibrium is undefeated in the sense of Mailath et al. (1993).[8] In Online Appendix C.2, we derive conditions under which the regulator's preferred equilibria

information by the public" (https://www.fca.org.uk/publication/ukla/tn-506-2.pdf).

[8]The regulator's optimal test gives each type a payoff weakly greater than $\bar{u}(\theta)$, which is pinned down by (6) regardless of the equilibrium played. As the corresponding beliefs are similarly pinned down, the silent equilibrium is trivially undefeated.

also survive equilibrium selection criteria in more general message spaces.

2.2 Applied Examples

We now illustrate the applied relevance of our model in three examples, each of which maps our general framework to a specific model of financial frictions. Moreover, after we characterize optimal tests, we return to these examples to extract practical insights from our theoretical results.

Example 1. [**Bank runs – Morris and Shin (2000)**] In this example, we interpret the firm as a bank who interacts with a population of depositors. Each depositor can withdraw a fixed claim of $1 immediately, or roll over their claim until the bank's assets mature and yield a net return θ. We adapt the model of bank runs in Morris and Shin (2000) to a setting with additional inside and outside information θ. In this setting, which we describe in detail in Appendix D, depositors observe any disclosure m by the bank and the test result s, which can be interpreted as a regulatory stress test. For simplicity, we assume that the worst type of bank still has positive net present value $\underline{\theta} \geq 0$, so that the regulator always has an incentive to prevent runs. This means that all banks in the model are illiquid, yet solvent.[9] After observing $\{m, s\}$, every depositor decides whether to withdraw early. We show that there is a run, meaning that all depositors withdraw early, if and only if

$$z < \frac{\kappa}{2} \equiv t \tag{7}$$

where z is depositors' expectation of θ, defined as in Equation (1), and κ is a parameter measuring the illiquidity of the bank's assets. A regulator wants to maximize depositors' joint utility, which is given by θ if there is no bank run, and by a constant (normalized to zero) if there is a bank run. Thus, the regulator's objective function is described by Equation (2), and the function $\lambda(z)$ capturing financial frictions in this example is a step function:

$$\lambda(z) = \begin{cases} 1, & z < t, \\ 0, & z \geq t. \end{cases} \tag{8}$$

A similar specification arises in Bouvard et al. (2015) and Goldstein and Leitner (2018).[10]

[9] It is easy to show, in addition to our characterizations below, that the optimal test design for insolvent banks would be to fail them, i.e., to reveal their true type, with probability one.

[10] In Goldstein-Leitner, the loss implied by a run is a constant, while here it is a fraction of assets. With minor notational changes, all our results go through when we adopt the bank's objective function from their paper. As we explain in Section 3.7, optimal tests differ markedly across the papers due to differences in constraints. Relative to those studied in Goldstein-Leitner, our tests strictly improve the regulator's payoff.

The bank's managers choose inside disclosures to maximize the utility function in Equation (5). As in the general model, $\phi > 0$ in this expression can reflect either a bank that sells a fraction of the its assets to outside investors, or managers who care about short-term results in the sense of their bank's current market valuation.

Example 2. [Corporate investment – Holmstrom and Tirole (1997)] Consider a firm who has an endowment of $A < 1$ units of cash and is able to make investments $I \in [0,1]$ in a productive technology, which yields $(1+\theta)I$ at a future date. We again assume that all firms have positive net present value, with $\underline{\theta} \geq 0$. The firm sells a security to risk-neutral investors, which entitles them to a fraction ϕ of the firm's future returns. Investors observe any disclosure m by the firm and the result of a credit rating s before buying the security. Investor's willingness to pay for the firm's security is then $\phi\,(1+z)\,I$, where z is defined in Equation (1). The firm's budget constraint for investment is

$$A + \phi\,(1+z)\,I \geq I.$$

Define the firm's maximum investment capacity $I\,(z)$ as:

$$I\,(z) = \min\left\{1, \frac{A}{1-\phi\,(1+z)}\right\}.$$

If $I(z) < 1$, then financial constraints bind, in the sense that the firm cannot invest up to its technological capacity. This occurs when market expectations satisfy

$$z < \frac{1-A-\phi}{\phi} \equiv t.$$

Whenever this inequality holds, the fraction of valuable investments that is lost is $\lambda\,(z) = 1 - I\,(z)$, a concave function of z. Notice that financial constraints are never binding when $\phi = 1$ or if $A \geq 1$. Thus, frictions in this model are driven by the firm's need to attract outside financing while also retaining at least some of its assets. This requirement is a common motivation for frictions in corporate finance, for example, when retention is necessary to maintain appropriate managerial incentives as in Holmstrom and Tirole (1997). With these definitions, we obtain a special case of our general model. The regulator, who designs credit ratings to maximize the utility function in Equation (2), can be interpreted as a ratings agency who wants to maximize the joint surplus of firms and investors.

Example 3. [Cash flow and Liquidity Constraints – Inostroza (2019)] Inostroza (2019) studies the interaction between firms' cash flows and liquidity constraints in an environment motivated by bank runs. We discuss the key qualitative effects in his model, which generates objective functions that are a special case of our general framework. In Inostroza's model, a bank raises cash by issuing securities against an investment project with unknown value θ and, at a future date, faces the threat of investor runs that can reduce the value of its initial investment. This generates a feedback loop: Any increase in investors' expectation z of θ raises their willingness to provide the bank with cash. This cash injection reduces the probability of a run, which further raises investors' willingness to pay. As a result, social surplus is strictly convex in z up to a threshold $z = t$, above which liquidity constraints stop binding and surplus becomes linear in z. More broadly, this intuition highlights an interaction between cash flow and liquidity constraints in corporate financing. One can view this environment as a third possible micro-foundation of our model.

3 Characterization of Optimal Tests

In this section, we derive necessary and sufficient conditions for an optimal test G. We present our analysis as follows: In Section 3.1, we simplify the regulator's problem by showing that, without loss, she can focus on tests that discourage all inside disclosures. This allows us to develop the idea that the regulator optimally aims for *public speech and private silence*, in the sense that the firm stays quiet and investors rely only on outside regulatory information. The regulator's desire to implement private silence introduces an incentive compatibility constraint into the regulator's problem, which ensures that firms have no incentive to make inside disclosures. In Section 3.2, we extend existing methods from the literature on Bayesian persuasion in order to solve the constrained version of the regulator's problem. This subsection highlights a methodological contribution, but readers who are interested mostly in the applied implications can skip it. In Sections 3.3 and 3.4, we use our method to fully characterize optimal tests, and we discuss their economic implications in Sections 3.5 and 3.6. Section 3.7 studies the effect of alternative assumptions on the timing of disclosures on optimal test design.

3.1 Public Speech and Private Silence

The regulator's problem in our model is complex because she designs the test G while anticipating that firms will choose their disclosure strategies endogenously in perfect Bayesian equilibrium. In order to make this problem tractable, we show that the regulator can focus on tests that discourage all inside disclosures in equilibrium. In other words, it is always optimal for the regulator to aim for public speech and private silence:

Lemma 1. ***[Public Speech and Private Silence]*** *For any test G, there exists an alter-*

native test $\hat{G}$ such that

1. *For any equilibrium under G, there is an equilibrium under $\hat{G}$ in which the regulator obtains the same expected utility, and*
2. *The firm makes no verifiable disclosures in the equilibrium under $\hat{G}$, choosing $m = \emptyset$ with probability* 1*.*

Lemma 1 shows that, given any test G and any associated equilibrium disclosures, the regulator can always do equally well with a test that induces the firm to stay quiet after learning θ. The first step of the argument is similar to a classical revelation principle: Since the regulator has full flexibility in designing the set of test results S and the conditional distributions $G(s|\theta)$, she can always construct a modified test $\hat{G}$, which replicates whatever disclosure the firm would have made in any equilibrium under G. In principle, one might worry that the new test may introduce new opportunities for the firm to affect investors' beliefs via (partially verifiable) disclosures.[11] The second step, which goes slightly beyond the standard argument, shows that $\hat{G}$ in fact does induce a perfect Bayesian equilibrium without any verifiable disclosures.

In the remainder of the paper, we utilize Lemma 1 and focus on tests that induce silence. An interesting property of these tests is that disclosure costs are zero on the equilibrium path, since the firm always sends the "quiet" message $m = \emptyset$ and the associated costs are $c(\emptyset) = 0$.[12]

Given Lemma 1, we can reframe the regulator's problem in a much more tractable way. First, we can simplify the regulator's expected payoff using the fact that she focuses on tests that induce private silence. Let G be such a test. Given that firms stay quiet and so are uninformative, investors' expectations under G are therefore given by the random variable $z = E_G[\theta|s]$. As the regulator is the sole source of information in any such test, from now on we will refer to z as the *grade* that the firm achieves under the test, and G its distribution.

Taking the conditional expectation $E_G[w(z,\theta)|z]$ of the regulator's payoff in Equation (2), and then applying the law of iterated expectations, we find that the regulator's unconditional expected utility under G is given by

$$\begin{aligned} E_G[w(z,\theta)] &= E_G[z(1-\lambda(z))] \\ &= E_G[v(z)]. \end{aligned}$$

[11] This issue is particularly relevant when we move beyond the "all-or-nothing" message space; we discuss it in more depth following Lemma 2.

[12] Notice that Lemma 1 is strengthened if the regulator's utility function in Equation (2) includes the firm's costs of disclosure $c(m)$. Therefore, our results below apply regardless of whether the costs of disclosure are private or social costs.

Therefore, the regulator's problem is to design a test that keeps firms quiet but also maximizes the expected value of $v(z)$. Notice that, under our general assumptions and in all our examples, this function is convex on the interval $[\underline{\theta}, t)$ and linear on the interval $\left[t, \overline{\theta}\right]$, with a potential upward jump at the point $z = t$ (e.g., in the bank run application – see Figure 1, panel (a)). The following result builds on this derivation to provide a useful formulation of the regulator's problem:

Lemma 2. ***[Regulator's Problem]*** *The regulator finds an optimal test G by solving the following problem:*

$$V = \max_{G} \quad E_G\left[v\left(z\right)\right] \quad \textit{subject to}$$
$$E_G\left[u\left(z,\theta\right) \mid \theta\right] \geq \bar{u}\left(\theta\right), \ \textit{for all } \theta, \tag{9}$$

where $\bar{u}(\theta)$ is the worst-case disclosure payoff, defined in Equation (19).

Lemma 2 employs the above derivation of the regulator's objective function. Furthermore, it contains a tractable characterization of the incentive constraint determining "private silence". In the baseline model, the incentive compatibility constraint immediately reduces to the requirement that the firm's expected payoff $E_G\left[u\left(z,\theta\right) \mid \theta\right]$ if it stays quiet, is at least equal to $\bar{u}(\theta)$. Were a test G to violate this inequality for some θ, then that type would have an incentive to pre-emptively disclose itself.

When we allow the message space to be more general (see Appendix A.1), a new complication arises. If a disclosure m is only partially informative, then the firm's outside option may depend on both m and the test result s. This effect arises because the regulator's test influences equilibrium play and, hence, the inferences drawn by investors following different disclosures. Fortunately, the proofs of Lemmas 1 and 2 for general message spaces demonstrate that one can focus on a class of tests in which these feedback effects are absent. For any test that induces silence, private disclosures are off the equilibrium path. In the regulator's preferred constellation of off-path beliefs, investors act skeptically following any disclosure, attaching probability one to the worst type that can make this disclosure. Hence, within the class of silence-inducing signals on which we focus, the value of firms' outside option (i.e., of deviating to verifiable disclosure) is independent of the regulator's information design.

The regulator's problem in Lemma 2 further highlights the role of timing in our model. The constraint in problem (9) captures the idea that the firm has to decide on disclosures before observing the result s of the regulator's test G. As discussed in Section 2.1, this timing is supported in the empirical corporate disclosure literature, which has identified several plausible costs of delaying disclosures.[13] Nonetheless, it is interesting to consider the

[13] Daley et al. (2020) further discuss the relevance of pre-emptive signaling in the context of credit ratings.

implications of reactive disclosures for test design; Section 3.7 provides a full comparison.

In Online Appendix C.3, we extend the model to allow the regulator to show an arbitrary, potentially noisy preview of s to the firm *before* it makes its disclosure m. We show she can never gain from this. Since reactive disclosures are formally equivalent to a full preview, one can therefore view the timing in section 2 as property of optimal design. In this sense, the idea that the regulator wants public speech and private silence extends further: not only does she prefer that the firm stay quiet when facing investors, but also that no information is privately communicated to the firm in advance of the regulatory test result.[14] Practically, this means that regulators should aim to choose the timing and institutional detail of tests so as to have the "last word" before investors respond (where possible).[15]

This insight is broadly related to revelation principles in mechanism design: In a general stochastic mechanism in Myerson (1982), for example, the designer can refrain from revealing any information about the outcome to agents, except for recommended actions, without loss of optimality. In our context, the regulator effectively recommends silence to all firms, and also does not benefit from releasing further information about the test result. This insight is also related to results in Ederer et al. (2018), who demonstrate that strategic opacity can be an optimal feature of mechanisms with gaming incentives.

3.2 Bayesian Persuasion with Private Silence

While we have considerably simplified the regulator's problem in Lemmas 1 and 2, the solution is still not trivial. A common approach to information design is to boil the problem down to a choice of cumulative distribution function $G(z)$ for grades/posterior expectations z. The regulator can then construct a test that induces this distribution of posteriors, so long as the prior distribution F is a mean-preserving spread of the chosen distribution of the regulator's chosen distribution G (e.g., Dworczak and Martini, 2019). Unfortunately, this method cannot be applied to our model, where the regulator needs to consider the incentive

[14]The ECB has on past occasions shared partial, preliminary findings with banks under its purview, ahead of the wider publication of its stress test results. Our results indicate that this can be costly. Interestingly, in 2014 the ECB proposed that banks sign non-disclosure declarations regarding the information shared in these "supervisory dialogues"—apparently concerned these meetings might prompt pre-emptive disclosures. Yet, German and Italian banks objected to the proposal, citing their obligations under timely disclosure rules. Hence, this episode also illustrates that bank managers perceive significant costs associated with attempts to delay disclosures. See https://www.reuters.com/article/us-ecb-banks-tests-results-idUSKCN0HK1X720140925 .

[15]Between 2011 and 2015, the Fed published its stress test results each March, during the annual reporting window of major US banks (cf. footnote 7); since 2016, publication of the results has been pushed back to June. Before 2016, Goldman Sachs had waited until April to publish its annual report, but has since switched its publication to March. While only anecdotal, this is perhaps suggestive. To the extent that annual reports offer a good opportunity for credible disclosure, our results suggest an advantage to summer testing.

compatibility constraint in (9). The reason for this discrepancy is that our incentive compatibility constraint is driven by the firm's expected utility after learning the true realization θ of its type, which depends on firms' higher-order expectations of investors' beliefs, and is generally different from investor's expectations z. Hence, the distribution of z is no longer sufficient to describe the regulator's problem.

We extend the Dworczak-Martini method to our constrained problem. Central to our approach is another change of variables. Instead of z, we study the distribution of the *expected grade* $x = E_G[z|\theta]$ that type θ achieves, not yet knowing the realization of z. We let the regulator choose the cumulative distribution of x instead of z. This change preserves the basic simplification in Dworczak and Martini (2019): Any feasible test necessarily induces a $H(x)$ which is a mean preserving contraction of the prior F.

For any expected grade x, we define the set $\delta(x)$ as follows:

$$\delta(x) = \{\theta : \bar{u}(\theta) \geq u(x, \theta)\}. \tag{10}$$

Here, $\delta(x)$ stands for the set of types that prefer their outside option, i.e., full disclosure, to obtaining expected grade x. Moreover, we write $F \circ \delta(x)$ for the prior probability mass of these types, which is defined as

$$F \circ \delta(x) = \int 1_{\theta \in \delta(x)} dF(\theta). \tag{11}$$

Key to our solution is the following result, which states a relaxed version of the regulator's problem:

Lemma 3. ***[Relaxed Problem]*** *Consider the following problem, in which the regulator chooses the distribution $H(x)$ of expected grades $x = E[z|\theta]$:*

$$\begin{aligned} \overline{V} = \max_H \quad & \int v(x) dH(x) \ \textit{subject to} \\ & F \textit{ is a mean-preserving spread of } H, \\ & H(x) \leq F \circ \delta(x) \quad , \textit{ for all } x \geq t \end{aligned} \tag{12}$$

where $F \circ \delta(x)$ is defined in Equation (11). The maximized value of this problem is weakly higher than the maximized value in Problem (9), so that $\bar{V} \geq V$.

Lemma 3 describes a new maximization problem, in which the regulator directly chooses a distribution $H(x)$ over expected grades, subject to two sets of constraint. We show in the Appendix that (12) is a relaxed version of (9), in that it always achieves a higher value.

The formulation of problem (12) introduces a second constraint in addition to the requirement that F is a mean-preserving spread of H. To understand the second constraint, recall that $\bar{u}(\theta)$ is the value of deviating to disclosure for type θ of the firm. The constraint compares two measures: First, the mass of firms who prefer their outside option to an expected grade x is $1 - F \circ \delta(x)$. Second, for any candidate distribution $H(x)$ of expected grades, we can write the mass of types who obtain an expected grade above x as $1 - H(x)$. The second constraint in problem (12) replaces the incentive constraints with the requirement that the second quantity must be greater than the first for all $x \geq t$. The key advantage of (12) is that some of the results in Dworczak and Martini (2019) extend to this type of "first order stochastic dominance" constraint (see Appendix A.2.2).

Two additional technical points are are worth considering. First, the objective function in problem (12) focuses on $x = E[z|\theta]$, which is not always a sufficient statistic for expected utility because the firm's objective function in our model is not linear in z when $z < t$. Second, we have imposed the second constraint in problem (12) only on outside options above the threshold t, effectively ignoring incentive constraints for firms with outside options below this threshold. The proof of Lemma 3 shows that there is no loss of generality associated with these shortcuts. Indeed, for any feasible test G in problem (9), there is a test G' which achieves weakly higher value, and for which i) types $\theta \geq t$ get grades $z \geq t$ with probability 1, confining these types to the linear part of the objective function, and ii) types $\theta < t$ do not have binding incentive constraints. We then show that the distribution of expected grades implied by G' also yields value V in problem (12), which implies that this is indeed a relaxed version of the regulator's problem.

We now turn to the solution of the regulator's problem, which leverages Lemma 3 to find optimal tests. Our strategy is to find a solution to the relaxed problem (12) that is also feasible in the regulator's original problem (9), and must therefore be optimal.

3.3 Optimal Tests

In Lemma 2, we stated the regulator's problem of finding an optimal test subject to the incentive compatibility constraint that firms must stay quiet in equilibrium. In addition, Lemma 3 in the previous subsection developed an auxiliary problem that is useful in solving for optimal tests. In Appendix A.3, we show how existing characterizations of optimal tests can be extended to solve this auxiliary problem.

We now use these intermediate results to fully characterize an optimal solution to the regulator's problem. To state our first main result, it is useful to define firms' *disclosure-*

equivalent grades, denoted $\bar{x}(\theta)$:

$$\bar{x}(\theta) = \frac{1}{\phi}\left[\bar{u}(\theta) - (1-\phi)\theta\right]. \tag{13}$$

Intuitively, if type θ of the firm expects to receive a grade of $\bar{x}(\theta)$, then it is indifferent between silence and full disclosure. Notice that, since verifiable disclosures are costly, we have $\bar{x}(\theta) < \theta$ for all θ. Given this definition, our characterization of optimal tests is as follows:

Proposition 1. ***[Characterization of Optimal Tests]*** *The following conditions define an optimal test G in the regulator's problem (9):*

1. *Every type $\theta < \theta^\star$ is fully revealed with probability 1, where $\theta^\star < t$. We refer to $\theta^\star$ as the* failure threshold.

2. *Every type $\theta \geq \theta^\star$ is fully revealed with probability $\alpha(\theta) \in [0,1]$ and receives grade $z = t$ with probability $1 - \alpha(\theta)$, where $\alpha(\theta)$ is the smallest value $\alpha \in [0,1]$ that satisfies*

$$\alpha\theta + (1-\alpha)t \geq \bar{x}(\theta), \tag{14}$$

where the disclosure-equivalent grade $\bar{x}(\theta)$ is defined as in Equation (13).

3. *The failure threshold $\theta^\star$ is smallest value satisfying the inequality:*

$$\int_{\theta^\star}^{\bar{\theta}} (\theta - \max\{t, \overline{x}(\theta)\})dF(\theta) \geq 0. \tag{15}$$

Proposition 1 defines an optimal test for the regulator's problem (9), which is illustrated in Figure 2. A useful point to keep in mind when interpreting our results is that the regulator is effectively trying to minimize the expected losses associated with financial frictions, which are given by $z\lambda(z)$ whenever a firm obtains a grade $z < t$.[16] We refer to any such grade as a "failing" grade, because it implies that the firm faces binding financial frictions after investors learn the test result.

[16]To see this more directly, notice that the objective function of the regulator in problem (9) can be written

$$E_G[v(z)] = E_G[z(1-\lambda(z))] = \mu_0 - E_G[z\lambda(z)],$$

where $\mu_0 = \int \theta dF(\theta)$ is the prior mean of θ, which is independent of the regulator's chosen test G. Hence, the problem can be written equivalently as minimizing $E_G[\theta\lambda(z)]$, subject to the same constraints as in (9).

First, we show that there is always an optimal test that fully reveals all firms with type $\theta < \theta^\star$. These types fail the test with probability 1, which is why we refer to $\theta^\star$ as the *failure threshold* in the rest of the paper. Below, we explain why the threshold property is optimal.

Second, we show that the optimal test assigns a binary lottery to the remaining types $\theta \geq \theta^\star$. Each of these types receives either a passing grade $z = t$, which takes the firm exactly to the threshold that eliminates financial frictions, or a grade that fully reveals the true realization of θ. The intuition is as follows: One can partition the firm into "weak" types $\theta < t$, for whom financial frictions would bind under full disclosure, and "strong" types with $\theta \geq t$. An ideal scenario for the regulator would be a pooling scheme, in which she passes strong types with probability 1, and also assigns a passing grade to as many weak types as possible. However, this pooling scheme would also give some strong types an incentive to deviate towards verifiable disclosure. To avoid this issue, the regulator assigns such types a grade that passes "with flying colors" and reveals their true quality with some probability $\alpha(\theta)$. This probability is chosen as the smallest value that satisfies type θ's incentive compatibility constraint, which in this case reduces to Equation (14). As we elaborate below, this is the most efficient way to satisfy the incentive constraints. Notice that the regulator sets $\alpha(\theta) = 0$ for weak types $\theta \in [\theta^\star, t]$ who pass the test. This is because these types can only benefit from the regulator's pooling scheme, so they need not be given an additional incentive to stay quiet.

The third part of Proposition 1 defines how the failure threshold $\theta^\star$ is chosen. The regulator would like $\theta^\star$ to be as small as possible in order to minimize financial frictions. However, she is constrained by the requirement that the pass grade must credibly induce investors expectations z to lie above the threshold t at which financial constraints start to bind. Using Bayes' rule, investors' posterior probability on type θ upon observing the passing grade can be written (informally) as

$$p(\theta|\text{pass}) = \frac{(1-\alpha(\theta))\, dF(\theta)}{\int_{\theta^\star}^{\bar{\theta}} (1-\alpha(\theta'))\, dF(\theta')}$$

The credibility requirement is that the passing grade actually induces a posterior mean $z = \int \theta p(\theta|\text{pass})\, d\theta \geq t$ or, equivalently,

$$\int_{\theta^\star}^{\bar{\theta}} (\theta - t)(1-\alpha(\theta))\, dF(\theta) \geq 0 \tag{16}$$

Substituting the revelation probabilities $\alpha(\theta)$ implied by Equation (14) now also yields the inequality in Equation (15). The regulator optimally chooses the lowest threshold $\theta^\star$ that is consistent with credibility. From this characterization, it also follows that the regulator can pass all types, choosing $\theta^\star = \underline{\theta}$, if and only if investors' prior beliefs satisfy

$$E[\theta] \geq E[\max\{t, \bar{x}(\theta)\}]. \tag{17}$$

Intuitively, when investors are sufficiently optimistic a priori in the sense of Equation (17), the regulator can design a test which ensures that financial frictions are never binding. Conversely, if investors' prior belief $E[\theta]$ is below the right-hand side of (17), then the regulator is forced to choose $\theta^\star > \underline{\theta}$ and faces binding financial frictions for a positive mass of weak types.

Notice that, whenever financial frictions bind, the regulator optimally chooses a failure threshold that satisfies the credibility constraint in Equation (16) with equality, and therefore induces a passing grade that satisfies $z = t$. Intuitively, it is always wasteful to assign a higher passing grade $z > t$ to weak types. Such a policy only transfers excessive surplus to weak types without alleviating financial frictions. For instance, if a set of types with non-binding incentive constraints enjoys a grade $z > t$ while some other set faces binding financial frictions, then it is obviously always better to pool more of the latter into the passing grades until investors expectations are reduced to exactly $z = t$.

However, when some types' incentive constraints are binding, the test should discourage disclosures by offering those types more surplus; this helps to explain why the regulator satisfies strong types' incentive constraints using the *stochastic* grades of Proposition 1, which randomize between full revelation and the passing grade $z = t$. To see the intuition, suppose a test assigned strong types to higher passing grades $z > t$ with probability one, potentially pooling some weak types into the same grade. This policy is suboptimal, because it again involves an inefficient transfer of surplus to the weak types. A better approach would be to promise those strong types receiving grade z a small probability of being revealed. Doing so increases their utility at the expense of the weak, and so relaxes incentive constraints. With incentives now relaxed, it becomes possible to pool yet more weak types into z and thereby reduce financial frictions. Interestingly, an argument along these lines can be used to show that a deterministic test can *never* be optimal, since it would again involve an inefficient split of surplus between strong and weak types. We prove this result formally in Proposition 2 below.

In Figure 2, panel (a), we illustrate these properties of optimal tests. In the figure, type $\theta_0 < \theta^\star$ is an example of a failing type. The thick arrow from this type to itself denotes that type θ_0 obtains a grade equal to its true value (i.e., is fully revealed) with probability one. Type θ_1 is a passing weak type, who obtains the passing grade $z = t$ with probability one, again represented by a thick arrow. Type θ_2 is a strong type who does not have a binding incentive compatibility constraint, so that the regulator is able to set $\alpha(\theta_2) = 0$ and assign the passing grade with probability one. From Equation (14) and the figure, one can verify that setting $\alpha(\theta_2) = 0$ is feasible because $\bar{x}(\theta_2) < t$. Finally, type θ_3 is a strong type with a binding incentive compatibility constraint. The regulator assigns a stochastic grade, which is represented in the figure by thin arrows that are labeled with the associated probabilities. Type θ_3 is fully revealed with probability $\alpha(\theta_3) > 0$ and obtains the passing grade with the complementary probability.

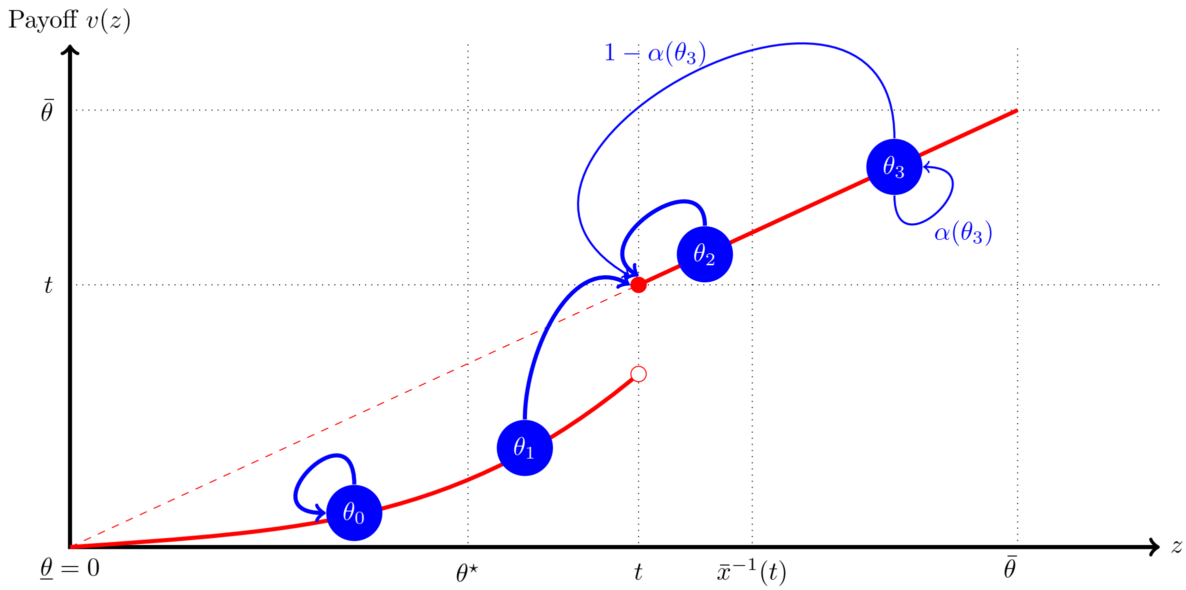
Panel (b) of Figure 2 further illustrates the cumulative distribution $G(z)$ of grades that is induced by an optimal test with these properties. This distribution can be understood by considering three regions. First, all types below the failure threshold $\theta^\star$ are fully revealed, so that the distribution of grades is identical to the prior distribution of types below this threshold. Second, there is bunching (i.e., a point mass) of types that are assigned the passing grade, implying an upward jump in $G(z)$ at the point $z = t$.[17] Third, among types with a binding incentive constraint (i.e., $\theta > \bar{x}^{-1}(t)$), truthful revelation of grades that pass "with flying colors" occurs with probability $\alpha(\theta)$. Hence, the distribution of grades in this region is a "flatter" version of the prior distribution among these types.

As a final note on Proposition 1, we can use the intuition above to explain why a single failure threshold is optimal. Suppose that the regulator considers a marginal increase in the probability that a weak type $\theta < t$ achieves the passing grade t. The expected marginal cost of this change is that the passing grade becomes less credible. Indeed, we can write the marginal cost as $\upsilon(t-\theta)\,dF(\theta)$, where υ stands for the shadow price associated with the credibility constraint (15). The marginal benefit, which arises because type θ obtains larger grades, is given by $(v(t) - v(\theta))\,dF(\theta)$. The policy change increases the regulator's expected utility if marginal benefits exceed marginal costs, that is:

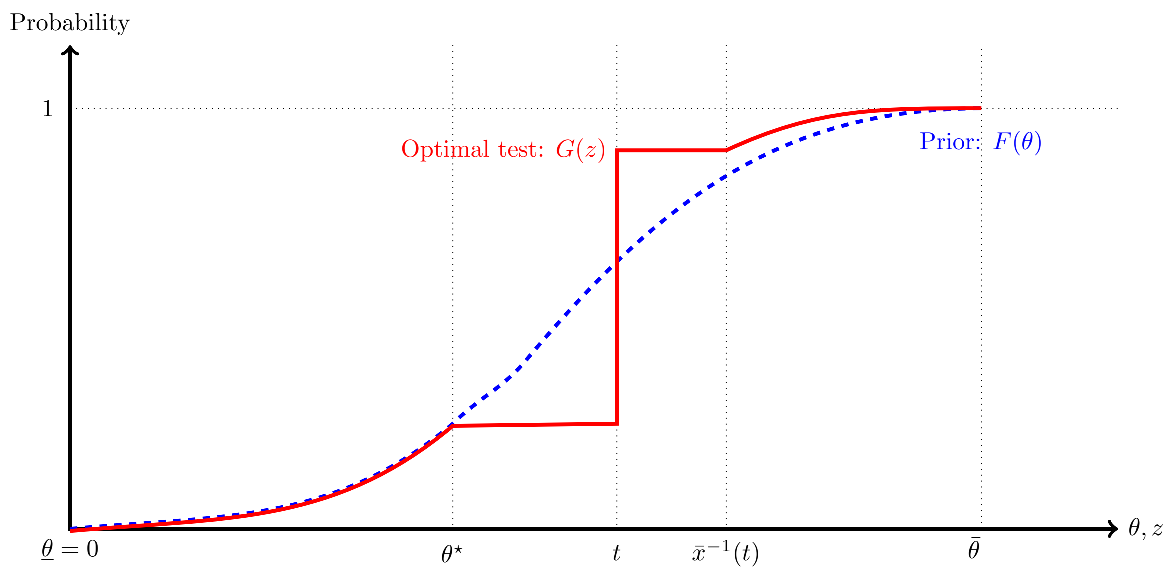
$$\frac{v(t) - v(\theta)}{t-\theta} \geq \upsilon.$$

The left-hand side of this inequality is the average gain obtained by promoting type θ to the passing grade t. Importantly, this is increasing in θ because $v(.)$ is convex on the interval $[0,t]$. Therefore, the regulator always prefers to raise the probability of passing higher types until her credibility budget is used up, resulting in the single failure threshold $\theta^\star$.

[17] Between the failure threshold and the passing grade t, the distribution of grades is flat, meaning that the regulator never reveals grades $z \in (\theta^\star, t)$.



(a) Examples of Optimal Grade Assignment



(b) Distribution of Optimal Grades.

Figure 2: **Illustration of Optimal Tests.** Panel (a) plots the optimal assignment of types θ to grades z under the optimal test described in Proposition 1. Assignments are drawn for four example types $\theta \in \{\theta_0, \theta_1, \theta_2, \theta_3\}$, and the associated expected payoffs $v(z)$ are on the vertical axis. Thick arrows denotes a deterministic grade that is assigned with probability one to a given type. Thin arrows denotes a stochastic grade and are labeled with the associated probabilities. Panel (b) shows the implied distribution of optimal grades. The blue / dashed curve is the cumulative prior distribution of types θ. The red / solid curve is the cumulative distribution of grades $z = E[\theta|s]$ implied by the optimal test described in Proposition 1. The incentive constraint associated with private silence is binding for all types $\theta > \bar{x}^{-1}(t)$.

3.4 Necessary Conditions for Optimal Tests

Proposition 1 fully describes an optimal test, but this implementation is not necessarily unique.[18] In particular, there are potential indeterminacies due to the fact that the regulator's objective function is linear in investors' expectations z on some intervals. For example, different tests can induce the same distribution of surplus across passing types, $\theta \geq \theta^\star$. Since the regulator's objective function $v(z)$ is linear for $z \geq t$, these alternatives are equally good for her — mean-preserving spreads / contractions among passing grades do not affect her expected utility so long as firms' incentives to stay quiet are preserved. Moreover, in applications such as Example 1 (bank runs – see Figure 1, panel (a)), the objective function is also linear for $z < t$, with a jump at $z = t$. In these cases, there are alternative optimal ways to assign failing grades to types $\theta < \theta^\star$, for example, by pooling all firms below the threshold $\theta^\star$ into a single failing grade instead of fully revealing them.

As a result, it becomes important to understand necessary conditions for optimality, so as to extract the economically important features that are shared by *all* optimal tests. We focus below on the interesting case where (17) is violated, so that $\theta^\star > \underline{\theta}$. The following result states two sets of necessary conditions.

Proposition 2. ***[Necessary Conditions for Optimal Tests]*** *All optimal tests G in the regulator's problem (9) satisfy the following conditions:*

1. *Every type $\theta < \theta^\star$ obtains a grade $z < t$ with probability 1, and every type $\theta \geq \theta^\star$ obtains a grade $z \geq t$ with probability 1, where the failure threshold $\theta^\star$ is the same as in Proposition 1.*

2. *For almost every type θ of the firm such that $\bar{x}(\theta) > t$, the incentive compatibility constraint binds, and the firm achieves expected utility $\bar{u}(\theta)$.*

3. *For every type θ with a binding incentive compatibility constraint, the test result z is stochastic conditional on θ.*

The first condition in Proposition 2 shows that all optimal tests share the stark separation around the failure threshold that we demonstrated in Proposition 1. All firm types below the threshold fail with probability one, and all types above pass with probability one, in any optimal test. Moreover, all optimal tests share the same threshold $\theta^\star$, which we characterized by Equation (15). Intuitively, the failure threshold captures the optimal trade-off between passing as many types as possible while also ensuring that the passing grade remains credible.

[18] However, as we show in the appendix, the implied distribution of expected grades x *is* unique.

The trade-off is not affected by any re-design of tests along linear regions of the regulator's objective and, therefore, we can find a unique optimal solution $\theta^\star$.

The second condition considers the incentive compatibility constraint in the regulator's problem, which ensures that no firm has an incentive to make additional disclosures. We show that the set of types θ for which this constraint is binding is the same at all optimal solutions. Indeed, exactly as in the particular solution that we define in Proposition 1, the constraint binds for all types such that $\bar{x}(\theta) > t$, who would have an incentive to deviate to disclosure if they were pooled to the passing grade $z = t$ with probability one.

An interesting third point is that, whenever the incentive constraint binds, the regulator assigns a stochastic grade z conditional on θ. This contrasts other models of information design with a continuum of types, in which optimal tests are often deterministic. The optimality of stochastic tests results from the nature of the regulator's trade-off. The regulator needs to provide strong types with high grades in order to encourage private silence, but also wants to pool them with weak types. As described in section 3.3, the most efficient way to manage this trade-off is to ensure weak types are pooled with stronger ones only on the grade $z = t$ that exactly avoids financial frictions. Randomization allows the regulator to balance these objectives, while deterministic tests do not.

In Online Appendix C.4, we derive an additional necessary condition for optimal tests. Recall that our results so far have relied on Lemma 1, which allows the regulator to focus on tests that induce private silence without loss of optimality. We further show that, even in the broader class of all possible tests, any optimal test *must* induce silence from all types $\theta \geq \theta^\star$ above the failure threshold. In other words, the regulator's optimum cannot be achieved via a combination of a test and subsequent equilibrium disclosures. Private speech by the strong is *strictly* detrimental for the regulator, emphasizing the desirability of the *public speech and private silence* principle.

In this subsection we provided a characterization of optimal tests. Next, we analyze some interesting comparative statics of optimal tests with respect to model primitives.

3.5 Comparative Statics

A useful observation is that the optimal test described in Proposition 1 depends only on two sets of parameters: The expected grades $\bar{x}(\theta)$ that make each type of firm indifferent between silence and disclosure, and the prior distribution of types $F(\theta)$. Knowledge of these two objects is enough to solve Equations (14) and (15), which determine the structure of the optimal test.

Interestingly optimal tests are thus independent of the nature of financial frictions, as measured by the function $\lambda(z)$: holding $\bar{x}(\theta)$ and $F(\theta)$ fixed, the same optimal test arises regardless of whether one studies bank runs in the tradition of Diamond and Dybvig (1983) (as in Example 1), or corporate financing frictions à la Holmstrom and Tirole (1997) (Example 2). Of course, the regulator's maximized utility does depend on frictions, since the total surplus lost to frictions is given by $\int_{\underline{\theta}}^{\theta^\star} \theta\lambda(\theta)\, dF$. However, these differences do not affect optimal policy.

In order to understand the comparative statics of optimal tests, we therefore focus on variations in $\bar{x}(\theta)$ and $F(\theta)$. We begin with an analysis of the firm's disclosure-equivalent grade $\bar{x}(\theta)$:

Proposition 3. ***[Comparative Statics: Disclosure Incentives]*** *If $\bar{x}(\theta)$ increases for all θ, then:*

1. *The failure threshold $\theta^\star$ in any optimal test increases and the regulator's maximized expected utility decreases.*

2. *The optimal test described in Proposition 1 becomes more informative: The distribution of optimal grades $z = E_G[\theta|s]$ undergoes a mean-preserving spread.*

Proposition 3 shows what happens when $\bar{x}(\theta)$ increases for all θ, so that higher expected grades are needed to discourage inside disclosures. The first point in the Proposition is straightforward: The regulator would like to enforce private silence, so that an increase in firms' incentives to make private disclosures is problematic. As a result, the regulator is forced to assign grades strictly above the passing threshold t to strong types with higher probability. In turn, the passing grade becomes less credible, and the regulator must raise the failure threshold $\theta^\star$ and pass fewer weak firms. This change is clearly detrimental for the regulator's expected utility.

The second point in Proposition 3 is more subtle and concerns the informativeness of the optimal test that we described in Proposition 1.[19] Whenever outside options increase, the grade $z = E_G[\theta|s]$ that the regulator releases as a result of her test becomes more variable in the sense of mean-preserving spreads (second order stochastic dominance).[20] In our setting, any decision maker whose utility depends only the expectation of θ would prefer to observe the optimal test corresponding to the higher disclosure payoff function $\bar{x}(\theta)$. This phenomenon is the result of two economic forces. First, when the disclosure-equivalent grade

[19]Due to the potential multiplicity of optimal tests, it is not possible to say that *any* optimal test for high $\bar{x}(\theta)$ is more informative than *any* optimal test for low $\bar{x}(\theta)$.

[20]We note that this ranking is weaker than the Blackwell (1953) criterion, which would require a mean-preserving spread in the entire perceived distribution of $\theta|s$ as opposed to its mean. As the regulator cares only about z, Blackwell's criterion is less useful here: a given distribution over z can often be implemented by more than one test, which need not be Blackwell-ranked.

$\bar{x}(\theta)$ increases, the failure threshold rises and the regulator fails a larger mass of weak types $\theta < \theta^\star$ by fully revealing them. This change provides more information about weak firms to an outside observer. Second, the regulator assigns grades that pass strong types "with flying colors" more frequently so as to incentivize their silence. An outside observer therefore sees more accurate information about both weak and strong types.

Proposition 3 is also useful because it gives us a direct sense of the impact of verifiable disclosures on optimal tests. Indeed, the "unconstrained" case, in which firms cannot make verifiable disclosures at all, is equivalent to $\bar{x}(\theta) \leq \underline{\theta}$ for all θ in our model. Proposition 3 immediately implies that, compared to the constrained solution that we have analyzed, unconstrained-optimal tests are always less informative and fail a smaller mass of firms: the regulator engages in more elaborate public speech so as to maintain private silence.

To obtain further insights, recall from Proposition 1 that the disclosure-equivalent grade $\bar{x}(\theta)$ depends on i) the disclosure technology available to firms, via the the worst-case disclosure payoff $\bar{u}(\theta)$, and on ii) the preference parameter ϕ, which measures the sensitivity of firms' payoffs to investors' beliefs. The following auxiliary result explores the relationship between disclosure incentives and model primitives:

Lemma 4. ***[Sources of Incentives to Disclose]*** *The firm's disclosure-equivalent grade $\bar{x}(\theta)$ increases for all θ and, therefore, optimal tests become more informative, if one of the following applies:*

- *Lower disclosure costs: The cost $c(m)$ of every verifiable message m decreases; or*
- *Firms are more sensitive to investors' beliefs: The parameter ϕ in the firm's utility function in Equation (5) increases.*

Lemma 4 first demonstrates that firm's incentives to disclose become stronger if disclosure becomes less costly, which is intuitive. The third point in the Lemma states that incentives to disclose become stronger when the sensitivity of firms' payoffs to investors' beliefs is large. This property is also intuitive. Verifiable disclosures in our model are driven by firms' desire to pre-emptively supplement outside grades that do not reflect their true quality. The more firms care about investors' beliefs, the stronger their incentive to make verifiable disclosures.

In our extension with more general message spaces, we can further derive comparative statics when the firm's available messages change:

Lemma 5. ***[Sources of Incentives to Disclose: Verifiability of Information]*** *Consider the model with general message spaces, described in Appendix A.1, in which a firm of type θ is able to send messages in a general set $M(\theta)$. Fix the set $M = \cup_\theta M(\theta)$ of all possible messages. The firm's disclosure-equivalent grade $\bar{x}(\theta)$ increases for all θ and, therefore,* *optimal tests become more informative, if the set $M^{-1}(m)$ of types that can send message m becomes becomes smaller (in the set inclusion order) for all m.*

This Lemma shows that incentives to disclose also become stronger if the messages available to the firm are "more verifiable". The sense in which we define verifiability is the size of the set of types θ that have access to each message. For example, in a cheap talk setting, every type can send every message so that $M^{-1}(m)$ is as large as possible. By contrast, a scenario in which $M^{-1}(m)$ becomes smaller captures more verifiability. In this scenario, investors' skeptical beliefs, which are used to derive the worst-case disclosure payoff $\bar{u}(\theta)$ in Equation (19), become more forgiving from the firm's perspective. The firm still expects to be judged as the worst type that can send any given m, but that worst type (weakly) increases when messages become more verifiable. Hence, incentives to disclose become stronger.

Combining these insights with Proposition 3 shows that either a decrease in disclosure costs or an increase in verifiability make the regulator worse off, forcing her to fail more firms and to release more informative tests overall. In the limit as disclosure costs become prohibitive or all messages become cheap talk, the regulator is not constrained by inside disclosures at all. The other extreme, where fully verifiable disclosure is costless, is the least favorable for the regulator.

As a complementary result, we also present comparative statics of optimal tests with respect to investors' prior beliefs $F(\theta)$.

Proposition 4. ***[Comparative Statics: Prior Beliefs]*** *Assume that the disclosure-equivalent grade $\bar{x}(\theta)$ has a single crossing with t. If the prior distribution $F(\theta)$ becomes more optimistic in the sense of the Monotone Likelihood Ratio Property (MLRP), then the failure threshold $\theta^{\star}$ in any optimal test decreases and the regulator's maximized utility increases.*

In information design problems related to ours, it tends to be optimal to release less information when investors' prior beliefs are optimistic (e.g., Bouvard et al., 2015; Goldstein and Leitner, 2018). In models without inside disclosures, this property arises because it becomes easier to make passing grades credible. Proposition 4 confirms this intuition, but the reasoning in our case is more nuanced. In principle, investors' optimism can make the regulator's credibility constraint in Equation (15) either easier or more difficult to fulfill. On one hand, investors place larger weight on strong types a priori, so that credibility is easier to achieve. On the other hand, because the probability weight on strong types increases, it becomes more costly to provide these types with incentives to stay quiet.

Proposition 4 provides conditions under which the former effect dominates, so that the standard intuition prevails. First, we assume that the disclosure-equivalent grade $\bar{x}(\theta)$ has at most one crossing with the threshold t below which financial frictions bind. For example, consider the case where the messages available to the firm are $M(\theta) = \{\theta, \emptyset\}$, so that the firm can either disclose the true θ or stay quiet. If the costs of full disclosure are a constant c, then we have $\bar{x}(\theta) = \theta - \frac{1}{\phi}c$, which is linear, and therefore has the single-crossing property. More generally, single crossing holds in this message space as long as the costs of disclosure do not rise too fast in the evidence disclosed, θ. Second, we consider optimism in the sense of MLRP, which places enough structure on prior beliefs to ensure that the integral defining the credibility constraint in Equation (15) increases with optimism. By contrast, without inside disclosure, a weaker notion of optimism such as first-order stochastic dominance would have been sufficient to make the regulator better off.

3.6 Private Silence as the Only Motivation for Public Speech: A Special Case

We briefly highlight an instructive special case of the characterizations we have developed:

Corollary 1. ***[Private Silence as the Only Motivation for Public Speech]*** *Suppose that the investor's prior beliefs satisfy*

$$t < E[\theta] < E[\max\{t, \bar{x}(\theta)\}]. \tag{18}$$

If firms cannot make verifiable disclosures, then the regulator's optimal test conveys no information to investors. By contrast, if firms can make verifiable disclosures, then the regulator optimally releases an informative test and fails a strictly positive mass of firm types, as described in Proposition 1.

Corollary 1 considers parameter constellations where investors' prior expectation $E[\theta]$ is in an intermediate range, as defined in Equation (18). In this case, a regulator who fully controls the informational environment does not need to provide any information to investors at all. Indeed, since $E[\theta] > t$, financial frictions never bind if investors judge all firms according to their prior expectation. The regulator can do no worse than to preserve this situation by sending a single, uninformative grade $z_0 \equiv E[\theta]$ as her test result. However, if firms are now given the opportunity to make verifiable inside disclosures, the regulator's opaque test "unravels". A set of strong firms have an incentive to deviate to disclosure, which means that they are no longer pooled into the regulator's test result z_0. Hence, investors' posterior belief given z_0 become more pessimistic, providing an incentive for further strong firms to disclose. This process makes the test result gradually less credible, and financial frictions will eventually bind for any firm that does not make a verifiable disclosure.[21] An optimal

[21] When disclosure costs are strictly positive, this process does not necessarily result in *all* types $\theta \geq t$

response to this problem is for the regulator to release some information that encourages private silence. This example makes the key point of our paper quite starkly, since enforcing private silence is now the only motivation for public speech.

Before moving on to the applications of our model, we discuss the robustness of our result to alternative assumptions about the timing of disclosure.

3.7 Alternative Timing Assumptions

In our baseline model, the firm can make only pre-emptive disclosures before the test result s realized. As discussed in Section 2, we view this assumption as realistic for corporate disclosures. Indeed, the available evidence in the accounting literature indicates that firms prefer early disclosures (e.g., Skinner, 1994; Kays, 2022). Moreover, it is reasonable to assume that verifiable disclosures take time to prepare, audit, and circulate. As extensions to this baseline case, we also characterize optimal tests under two sets of alternative assumptions about timing. Both cases are analyzed rigorously in Online Appendix C.1.[22]

Late disclosure option: We consider a game in which the firm has to decide whether to prepare the verifiable message $m = \theta$, and whether to sink the associated cost $c(\theta)$, before it knows the result s of the regulatory rest. However, having prepared its message, the firm has a late disclosure option: It can delay its decision whether to disclose $m = \theta$ to investors after z is realized. This setting can also capture the idea that any information that is prepared at the last minute is likely to be perceived as cheap talk.

We show that the optimal test described in Proposition 1 *remains* optimal in the game with a late disclosure option. The intuition for this equivalence is as follows: Under the optimal test, the only types of firm that might be tempted to pre-prepare a disclosure are strong types $\theta > t$. The optimal test assigns strong types a lottery over grades $z \in \{t, \theta\}$, which are always weakly below their true type. Thus, conditional on pre-preparing the disclosure, releasing it ex post is a weakly dominant strategy, and the payoff from preparing it is equivalent to the payoff from a pre-emptive disclosure. However, by construction, pre-emptive disclosure is not a profitable deviation for the firm since the optimal test induces private silence in the baseline model. Therefore, firms also do not have an incentive to deviate from the optimal test by pre-preparing disclosures.

making a disclosure. However, it remains true that non-disclosing firms will face an effective grade $z < t$ with probability 1.

[22] There, we also provide a stylized analysis of another cost of delay—namely, that late disclosures may be less credible than prompt ones.

Ex post disclosures: We also consider a version of our model in which the firm can prepare and release evidence ex post, i.e., after z is realized. We begin by showing that the principle of *public speech and private silence* continues to apply: The regulator can optimally focus on tests that induce equilibria without any ex post disclosure. Intuitively, given any test that induces the firm to make a disclosure, the regulator can design an equivalent test that induces silence and replicates the firm's disclosures. We again leverage this principle to provide a full characterization of optimal tests, under some weak restrictions on primitives. An optimal test pools weak and strong types *pairwise*, according to a negative assortative matching function. We further show that this matching function is *sufficient* for optimality as long as it is combined with *coarse pooling* of a positive-measure interval of intermediate types around the critical threshold $z = t$.

In our baseline model with pre-emptive disclosures, Proposition 2 shows that optimal tests are necessarily stochastic. Moreover, the test is monotonic in that higher types enjoy a higher surplus. In contrast, optimal tests with ex post disclosures are non-montonic and deterministic, so that every type θ of the firm knows with certainty the grade that it will be assigned. Being deterministic, reactive disclosures necessarily lead to a strict increase in financial frictions. As we argued in Section 2, pre-emptive disclosures are an empirically relevant case, especially if it takes time to prepare verifiable information. In this context, our results indicate a strict benefit of conducting stress tests which ensure strong banks pass, but leave them uncertain about exactly how well they will fare. Historically, this kind of ambiguity has been a feature of the Fed's stress test.[23]

Our extension with ex post disclosures also clarifies the relationship between our work and Goldstein and Leitner (2018) (GL), who study optimal stress tests when banks can decide ex post whether to sell or retain their assets. In their paper, the bank's "outside option" is asset retention while, in our case, it is disclosure. One point of similarity is that negative assortative matching emerges as part of an optimal test in both cases.[24] However, there are also two important differences. First, the nature of the outside option leads to an interesting difference in the nature of optimal tests. GL's regulator cannot always replicate the firm's outside option with her test design and, therefore, some banks may choose to retain assets in equilibrium. As a result, in their setting the relevant constraints may depend in general on both the way firms' reservation prices compare with their type and the regulator's objective

[23]Despite a recent move towards greater transparency, Leitner and Williams (2023) point out that the Fed still does not reveal everything about its tests. They caution that too much transparency can distort investment decisions. We identify a complementary benefit—uncertainty allows the test to better exploit pooling and thereby reduce financial frictions.

[24]Garcia and Tsur (2021) also find negative assortative matching in a finite-types insurance setting. In a model with a continuous state and smooth payoffs, Kolotilin et al. (2022) find conditions for negative assortative matching to be optimal; their setting involves no incentive constraints.

(as these determine the set of types she wishes to sell). By contrast, in our disclosure context the regulator can focus on tests that induce all types of firm/bank to forgo the outside option of disclosure, regardless of the bank's or the regulator's objective. Second, we are able to provide a full characterization of optimal tests. We identify that optimal tests pass banks above a threshold, where passing grades comprise a combination of *fine-tuned*, negatively-assorting grades and a *coarse*, bunching region of intermediate types. This insight is not available in GL, perhaps due to discreteness issues introduced by their finite state space.

4 Applications

4.1 Bank Stress Tests in Times of Financial Crises

In Examples 1 and 3 above, we demonstrated that our framework can be applied to models of banking crises. In this context, our characterization of optimal tests highlights the preferred policy of a regulator who can release credible information to investors, and who is concerned about market expectations dropping below a threshold and triggering bank runs. We view this setting as informative for the problem of real-word financial regulators who have to design stress tests in times of financial crises.

For example, during the financial crisis of 2007-9, the Federal Reserve designed a novel stress testing exercise as part of its Supervisory Capital Assessment Program (SCAP). The Fed's stated rationale for announcing this test was to provide credible information that would resolve uncertainty about potential losses in the US banking system.[25] All large US bank holding companies were required to take part. For each of the 19 participating banks, the Fed calculated projected losses in hypothetical, adverse macroeconomic scenarios. The Fed also released granular breakdowns of these losses for the "severely adverse" scenario in its test to the public.[26] Several studies have confirmed that the SCAP conveyed credible information to investors, which manifested itself in bank stock returns and trading volumes (Morgan et al., 2014; Flannery et al., 2017), CDS spreads and systematic risk (Sahin et al., 2020), and indicators of information asymmetry and uncertainty (Fernandes et al., 2020).

Interestingly, our results also offer a novel perspective on the contrasting performances of stress tests conducted in the US and Europe following the 2007-08 financial crisis. As Schuermann (2014) explains, the Federal Reserve's tests were widely regarded as having restored investor confidence more successfully than those subsequently conducted by the European Banking Authority. Moreover, evidence suggests that the Federal Reserve's stress

[25] See https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20090424a1.pdf for a description of the framework and its rationale, which was published before the results were announced.

[26] For example, see Table 3 in https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20090507a1.pdf.

tests have not been associated with increased voluntary disclosures by banks,[27] while the EBA's tests did lead to voluntary disclosures of banks' sovereign debt exposures (Bischof and Daske, 2013).[28] The principle of *public speech and private silence* in our analysis suggests that the Fed's stress tests may have been more effective, in part, because they did not trigger additional bank disclosures, allowing information to flow more efficiently to market participants.

Since the financial crisis, regulators have also adopted a more regular stress testing regime, which releases information about banks' health in normal times. The stated goal of these tests is different, and focuses more on their ability to reinforce transparency and to mitigate moral hazard in the financial system. While our analysis can shed some light on the optimal design of tests in normal times (recall Corollary 1), we do not formally model moral hazard or market discipline. Hence, we view our results as primarily useful for the design of stress tests *in times of financial distress,* such as the SCAP and its international equivalents, which were driven mostly by the desire to avoid market distress.

Below, we discuss three novel insights for stress test design that arise from our analysis.

Optimal granular stress tests: A novel feature of optimal tests in our model is that the regulator releases three types of grades: Failing grades $z < t$ lead to binding financial frictions, a "marginal" passing grade $z = t$ pools weak and strong banks, and a range of grades $z > t$ that pass strong firms "with flying colors". Thus, optimal tests are quite granular. They feature significant bunching at the "marginal" grade, but they also offer the strong a rare opportunity for distinction in order to discourage disclosures. In addition, if there is is strict convexity in the regulator's objective function $v(z)$ when frictions bind at $z < t$ (e.g., in Example 3), then any optimal test also involves multiple failing grades.

Interestingly, the granularity of the optimal test in our model is somewhat consistent with how the 2009 SCAP was implemented in practice. Indeed, the Fed published detailed results about each bank's losses in adverse scenarios. This included losses among banks such as Goldman Sachs and JP Morgan, who "passed" the test, in the sense that the Fed did not conclude that they had a capital shortfall. Our model provides a possible rationale for these practices, if one of the motives for public disclosure of stress test results is to prevent the

[27]Consistent with greater bank opacity, Flannery et al. (2017) find that analyst earnings forecast errors do not significantly reduce after a bank is stress tested–despite a rise in coverage. Interestingly, they note that while the Fed's test allows banks to voluntarily include additional disclosures from scenarios beyond the "severely adverse" one published by the Fed, most have not taken up this option. Shahhosseini (2016) studies banks' reactions to being included in stress tests. She finds stress-tested banks made fewer loan charge-offs and more frequently changed the classification of loan losses, which one could interpret as increased opacity.

[28]In addition, Bischof and Daske (2013) show that the latter shift was associated with a significant *increase* in average bid-ask spreads across stress-tested banks. Petrella and Resti (2013) also study the EBA's stress tests in 2011.

"unraveling" of information via signals from strong banks.

These granularity results are new to the existing literature on stress test design, which shows that simple pass/fail tests are optimal in many settings. This includes Faria-e-Castro et al. (2016), who study the interaction of stress tests with government bailouts, Inostroza and Pavan (2022), where a regulator can communicate privately with dispersed investors, and Orlov et al. (2021), who advocate for an initial binary test followed by capital injections and ultimately full disclosure. Inostroza (2019) emphasizes the merits of multiple failing grades, as in our Example 3, but in his model there is no rationale for granularity in passing grades. In Goldstein and Leitner (2018), strong firms have the outside option of retaining an asset rather than selling it to investors; we discussed the relationship in section 3.7. Particularly novel features of our analysis are that optimal grades involve strict separation around the failure threshold $\theta^{\star}$, and assign stochastic test results to all types with a binding incentive compatibility constraint. Our work also connects to Philippon and Skreta (2012), who show in a very different environment that regulatory policy in a crisis should take into account its effect on endogenous information and adverse selection in financial markets.

Determinants of Regulatory Trade-offs in Stress Test Design: Our analysis of comparative statics in Section 3.5 sheds light on the institutional details that affect the optimal design of bank stress tests. In common with existing literature, we show that regulators must release more accurate information, usually at the cost of failing more banks, when investors' beliefs become more pessimistic in a financial crisis. In addition, we show that it is not enough to reassure investors in isolation. Instead, regulators need to consider the endogenous response of banks' disclosures, make stress tests informative enough to prevent "unraveling" of information, and encourage a regime with *public speech and private silence.*

Our analysis of banks' incentives and disclosure technologies in Proposition 3 and Lemma 4 highlights additional features that regulatory designs must take into account. First, the required informativeness of stress tests is increasing in banks' ability to make verifiable disclosures. A sizable applied literature in accounting studies the extent to which banks' disclosures have the potential to credibly reassure investors. An emerging consensus is that banks' disclosures about opaque areas of their business model, in particular about derivatives and securitization activities, provide particularly valuable information to investors during crises (e.g., Barth and Landsman, 2010; Iren et al., 2014). By contrast, standard fair value accounting in normal times is considered less important. Sowerbutts et al. (2013) also document that banks made heightened disclosures during the crisis of 2008. Our analysis implies that the more detailed banks' potential verifiable disclosures become, the more regulators should be concerned about "unraveling" effects, and respond accordingly with more informative stress tests and a greater emphasis on multiple passing grades that incentivize private silence.

Moreover, Lemma 4 highlights that banks' incentives to disclose may be driven by the incentives of their managers and, in particular, by the sensitivity of managers' payoffs to investors' expectations. This finding is particularly salient in banking, because senior bankers tend to have bonus-based compensation schemes. These patterns have been documented both by theoretical analyses of optimal compensation in banking (e.g., Thanassoulis, 2012; Acharya et al., 2016) and by the associated empirical evidence (e.g., Bannier et al., 2013; Célérier and Vallée, 2019). All else equal, our results suggest that regulators should be more concerned about inside disclosures when managers' payoffs are more sensitive to current market expectations, for example because managers own their bank's stock / options or because their bonuses are sensitive to annual stock price performance.

Perhaps surprisingly, once these institutional features have been taken into account, the optimal test described in Proposition 1 does not depend on the details of a particular model of financing frictions that regulators have in mind. Indeed, the optimal test has the same structure, and the same failure threshold, when financial frictions take the shape of a binary "run / don't run" outcome (as in our Example 1), or when they are manifested in a smooth, convex function that captures the gradual deterioration of market conditions as investors become more pessimistic (as in Example 3). Therefore, our analysis shows that tests need not be "fine-tuned" to particular financial frictions. Rather, optimal designs depend crucially on market beliefs and insiders' incentives.

A final point worth highlighting is our recommendation for stress testing at the "eve of a crisis", when investors' prior beliefs are relatively pessimistic but lie in the intermediate region characterized in Corollary 1. In these scenarios, there is no need for stress tests at all unless regulators expect inside disclosures by banks. Running a stress test in these situations is still beneficial because it ensures that information does not unravel, and presents information to investors in a way that minimizes the probability of future bank runs.

Credibility and Regulatory Commitments to Stress Test Accuracy: Several studies of stress tests have highlighted potential time inconsistencies in regulator's behavior. For example, Bouvard et al. (2015) show that regulators may have an incentive to withhold private information about bank vulnerability ex post, in an attempt to signal strength to investors, while it would be optimal ex ante to commit to an informative stress test.

In our framework, we can address the question whether a regulator would change the nature of the optimal test after learning the true quality θ of banks' assets. Concretely, notice that the regulators' and the firms' preferences, defined in Equations (2) and (5), are exactly the same when we set the sensitivity parameter ϕ to zero. Hence, analyzing stress tests that induce private silence when $\phi = 0$ is equivalent to analyzing tests that would induce the *regulator* to refrain from further disclosures, and to obey her ex-ante information design, after she learns the realization of θ.

Using our characterization of optimal silence-inducing tests in Proposition 1, we can infer that incentive constraints associated with private silence are never binding in the case where $\phi = 0$. The economic intuition is as follows: Incentive constraints in our model when strong banks have an incentive to signal their quality to investors, whose beliefs drive bank managers' payoffs whenever $\phi > 0$. By contrast, under the regulators' preferences with $\phi = 0$, there is no gain from signaling quality as long as one achieves a grade $z \geq t$ that avoids financial frictions. Interestingly, this allows us to slightly relax our assumptions about regulatory commitment. If the regulator were able to make a full disclosure after learning θ, she would not want to do so under the optimal test that we have characterized.

We stress that other types of commitment issues for the regulator remain relevant in the light of the recent literature. First, as pointed out by Parlasca (2021), if one of the objectives of stress tests in normal times is to alleviate moral hazard, then the regulators have an incentive to threaten tough tests ex ante, but relax these tests ex post if there is a possibility of binding financial frictions. Second, in the model considered by Orlov et al. (2021), it is sometimes optimal to fail strong types of bank in order to make it easier for failing banks to raise capital after the test is released. This type of mechanism may not be time-consistent if the regulator prefers to pass strong banks instead after their strength has been realized.

4.2 Optimal Credit Ratings

In Example 2, we demonstrated that our setting nests a standard model of corporate financing in the spirit of Holmstrom and Tirole (1997). Financial frictions in this model arise when the firm's borrowing capacity becomes insufficient to make all available investments that have positive NPV. Releases of information about asset quality in this setting can be interpreted as credit ratings, whereby favorable ratings allow firms to borrow more and ensure that financial constraints do not bind. The regulator can be interpreted as an idealized ratings agency that maximizes the joint utility of investors and the firm.

This interpretation allows us to contribute insights to a growing literature on the design of credit ratings. Motivated by the crisis of 2008, this literature focuses to a large extent on imperfections in how credit ratings are produced and consumed. These issues include firms' incentives to shop for favorable ratings (Skreta and Veldkamp, 2009; Bolton et al., 2012; Kashyap and Kovrijnykh, 2016), the interaction of ratings with financial regulations (Opp et al., 2013), and the incentives for firms to block the publication of unfavorable ratings Sangiorgi and Spatt (2017). Daley et al. (2020), in perhaps the closest paper to ours in this literature, show how ratings affect firms' incentives to signal their quality via asset retention, with further implication for screening and credit supply. Given the various market failures pointed out by the existing literature, we do not view credit ratings agencies in current reality as social surplus maximizers. Therefore, our findings should be interpreted as an efficient benchmark for credit ratings, rather than a prediction of ratings market equilibrium. Future research investigating the reforms needed to approach this benchmark would be interesting.

All of the studies cited above consider a simple, usually exogenous structure for the informative signals that are conveyed by credit ratings. In particular, the standard approach is to assume that the firm has a binary type and investors observe a binary distributed signal of this type.[29] These assumptions are geared towards conveying intuition about various incentive problems in the clearest possible way.

Our model contributes a different perspective to this literature. Applying our general results to Example 2, we can analyze how credit ratings should be designed *in principle* when there are many possible types of firm, who can also make additional corporate disclosures to increase their borrowing capacity. Once again, the problem of designing an optimal rating boils down to ensuring public speech and private silence. This strategy, which crowds out all voluntary additional disclosures, ensures that the overall impact of financial frictions is minimized, and also economizes on any deadweight costs associated with corporate disclosure.

Unlike existing studies, we show how a wide spectrum of optimal ratings emerges endogenously from the objective to minimize financial frictions. First, there is a range of low "junk" ratings, whose recipients end up financially constrained with probability one. Second, there is significant bunching of firms around an *"investment grade"* rating, at which financing constraints become slack. Finally, there is a range of "premium" ratings, which deliver additional surplus to the firms who receive them. This array of ratings resembles the wide range of grades that the major credit ratings agencies deliver in practice. We would not expect credit ratings in the data to satisfy our optimality conditions exactly, since the ratings industry is subject to the various frictions and agency problems mentioned above. However, we view this result as an interesting and novel benchmark for ratings design.

[29] A notable exception is Manso (2013), who models a stochastic process of more granular ratings but still takes the structure of ratings as given.

5 Conclusion

We have characterized the optimal informational policy of a regulator whose goal is to alleviate financial frictions. Our results provide a characterization of optimal tests that is robust to an informational version of the "Lucas critique": One should not assume that regulators have full control over the informational environment in which financial markets operate, and should expect that any change in informational policy will trigger further changes in the endogenous production of information.

We demonstrate that optimal tests take a markedly different shape once we account for this feature. We show that regulators should aim for public speech and private silence, refraining from private communication of test results to firms, and preserving firms' incentives not to communicate with investors. As a result, optimal tests become more granular, more informative and involve novel types of grades that pass firms "with flying colors" in order to preserve their incentives to stay quiet. Moreover, they should maintain some uncertainty about the likely grade firms will receive.

We further illustrated the applied insights that arise for the design of bank stress tests during financial crises, and for the design of an idealized credit ratings agency that maximizes social surplus. Beyond the applications covered in this paper, we hope that our approach and methodology will spark further work on constrained information design, which could consider additional applications in which the information designer does not have full control of the informational environment.

References

Acharya, V., M. Pagano, and P. Volpin (2016). Seeking alpha: Excess risk taking and competition for managerial talent. *The Review of Financial Studies 29*(10), 2565–2599.

Acharya, V. V., P. DeMarzo, and I. Kremer (2011). Endogenous information flows and the clustering of announcements. *American Economic Review 101*(7), 2955–2979.

Allen, L. and A. Saunders (1992). Bank window dressing: Theory and evidence. *Journal of Banking & Finance 16*(3), 585–623.

Banks, J. S. and J. Sobel (1987). Equilibrium selection in signaling games. *Econometrica 55*(3), 647–661.

Bannier, C. E., E. Feess, and N. Packham (2013). Competition, bonuses, and risk-taking in the banking industry. *Review of Finance 17*(2), 653–690.

Barth, M. E. and W. R. Landsman (2010). How did financial reporting contribute to the financial crisis? *European accounting review 19*(3), 399–423.

Bertomeu, J., E. Cheynel, and D. Cianciaruso (2021). Strategic withholding and imprecision in asset measurement. *Journal of Accounting Research 59*(5), 1523–1571.

Best, J. and D. Quigley (2020). Persuasion for the long run. *Available at SSRN 2908115*.

Billingsley, P. (1995). *Probability and measure* (3. ed ed.). John Wiley & Sons.

Bischof, J. and H. Daske (2013). Mandatory disclosure, voluntary disclosure, and stock market liquidity: Evidence from the eu bank stress tests. *Journal of accounting research 51*(5), 997–1029.

Bischof, J., C. Laux, and C. Leuz (2021). Accounting for financial stability: Bank disclosure and loss recognition in the financial crisis. *Journal of Financial Economics 141*(3), 1188–1217.

Blackwell, D. (1953). Equivalent comparison of experiments. *Annals of Mathematical Statistics 24*(2), 265–272.

Bolton, P., X. Freixas, and J. Shapiro (2012). The credit ratings game. *The Journal of Finance 67*(1), 85–111.

Bond, P. and Y. Zeng (2021). Silence is safest: information disclosure when the audiences preferences are uncertain. *Journal of Financial Economics*.

Bouvard, M., P. Chaigneau, and A. de Motta (2015). Transparency in the financial system: Rollover risk and crises. *Journal of Finance 70*(4), 1805–1837.

Célérier, C. and B. Vallée (2019). Returns to talent and the finance wage premium. *The Review of Financial Studies 32*(10), 4005–4040.

Cho, I.-K. and D. M. Kreps (1987a). Signaling games and stable equilibria. *The Quarterly Journal of Economics 102*(2), 179–221.

Cho, I.-K. and D. M. Kreps (1987b). Signaling games and stable equilibria. *Quarterly Journal of Economics 102*(2), 179–221.

Daley, B. and B. Green (2014). Market signaling with grades. *Journal of Economic Theory 151*, 114–145.

Daley, B., B. Green, and V. Vanasco (2020). Securitization, ratings, and credit supply. *The Journal of Finance 75*(2), 1037–1082.

DeMarzo, P. M., I. Kremer, and A. Skrzypacz (2019, June). Test design and minimum standards. *American Economic Review 109*(6), 2173–2207.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–19.

Dworczak, P. and G. Martini (2019). The simple economics of optimal persuasion. *Journal of Political Economy 127*(5), 1993–2048.

Ederer, F., R. Holden, and M. Meyer (2018). Gaming and strategic opacity in incentive provision. *The RAND Journal of Economics 49*(4), 819–854.

Einhorn, E. (2018). Competing information sources. *The Accounting Review 93*(4), 151–176.

Faria-e-Castro, M., J. Martinez, and T. Philippon (2016). Runs versus lemons: information disclosure and fiscal capacity. *The Review of Economic Studies*.

Feltovich, N., R. Harbaugh, and T. To (2002). Too cool for school? Signalling and countersignalling. *RAND Journal of Economics 33*(4), 630–649.

Fernandes, M., D. Igan, and M. Pinheiro (2020). March madness in wall street:(what) does the market learn from stress tests? *Journal of Banking & Finance 112*, 105250.

Flannery, M., B. Hirtle, and A. Kovner (2017). Evaluating the information in the federal reserve stress tests. *Journal of Financial Intermediation 29*, 1–18.

Frenkel, S., I. Guttman, and I. Kremer (2018). The effect of exogenous information on voluntary disclosure and market quality.

Friedman, H., J. Hughes, and B. Michaeli (2015). Bayesian persuasion in the presence of discretionary disclosure. *Mimeo*.

Garcia, D. and M. Tsur (2021). Information design in competitive insurance markets. *Journal of Economic Theory 191*, 105160.

Gentzkow, M. and E. Kamenica (2017). Competition in persuasion. *The Review of Economic Studies 84*(1), 300–322.

Gigler, F. and T. Hemmer (1998). On the frequency, quality, and informational role of mandatory financial reports. *Journal of Accounting Research* (36), 117–147.

Goldstein, I. and Y. Leitner (2018). Stress tests and information disclosure. *Journal of Economic Theory 177*, 34–69.

Goldstein, I. and A. Pauzner (2005). Demand–deposit contracts and the probability of bank runs. *Journal of Finance 60*(3), 1293–1327.

Goldstein, I. and L. Yang (2017). Information disclosure in financial markets. *Annual Review of Financial Economics 9*, 101–125.

Grossman, S. J. and O. D. Hart (1980). Disclosure laws and takeover bids. *Journal of Finance 35*(2), 323–334.

Guo, Y. and E. Shmaya (2019). The interval structure of optimal disclosure. *Econometrica 87*(2), 653–675.

Hagenbach, J., F. Koessler, and E. Perez-Richet (2014). Certifiable pre-play communication: Full disclosure. *Econometrica 83*(3), 1093–1131.

Holmstrom, B. and J. Tirole (1997). Financial intermediation, loanable funds, and the real sector. *the Quarterly Journal of economics 112*(3), 663–691.

Inostroza, N. (2019). Persuading multiple audiences: An information design approach to banking regulation. *Available at SSRN 3450981*.

Inostroza, N. and A. Pavan (2022). Adversarial coordination and public information design. *Mimeo*.

Iren, P., A. K. Reichert, and D. Gramlich (2014). Information disclosure, bank performance and bank stability. *International Journal of Banking, Accounting and Finance 5*(4), 374–417.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Kashyap, A. K. and N. Kovrijnykh (2016). Who should pay for credit ratings and how? *The Review of Financial Studies 29*(2), 420–456.

Kays, A. (2022). Voluntary disclosure responses to mandated disclosure: Evidence from australian corporate tax transparency. *The Accounting Review 97*(4), 317–344.

Kolotilin, A., R. Corrao, and A. Wolitzky (2022). Persuasion with non-linear preferences.

Leitner, Y. and B. Williams (2023). Model secrecy and stress tests. *The Journal of Finance 78*(2), 1055–1095.

Li, F. and P. Norman (2021). Sequential persuasion. *Theoretical Economics 16*(2), 639–675.

Lyu, C. (2021). Information design for selling search goods and the effect of competition. *Working Paper,*.

Mailath, G. J., M. Okuno-Fujiwara, and A. Postlewaite (1993). Belief-based refinements in signalling games. *Journal of Economic Theory 60*(2), 241–276.

Manso, G. (2013). Feedback effects of credit ratings. *Journal of Financial Economics 109*(2), 535–548.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics 12*(2), 380–391.

Morgan, D. P., S. Peristiani, and V. Savino (2014). The information value of the stress test. *Journal of Money, Credit and Banking 46*(7), 1479–1500.

Morris, S. and H. S. Shin (2000). Rethinking multiple equilibria in macroeconomic modeling. *NBER Macroeconomics Annual 15*, 139–182.

Myerson, R. B. (1982). Optimal coordination mechanisms in generalized principal–agent problems. *Journal of mathematical economics 10*(1), 67–81.

Opp, C. C., M. M. Opp, and M. Harris (2013). Rating agencies in the face of regulation. *Journal of financial Economics 108*(1), 46–61.

Orlov, D., P. Zryumov, and A. Skrzypacz (2021). Design of macro-prudential stress tests.

Parlasca, M. (2021). Time inconsistency in stress test design. *Available at SSRN 3803876*.

Petrella, G. and A. Resti (2013). Supervisors as information producers: Do stress tests reduce bank opaqueness? *Journal of Banking & Finance 37*(12), 5406–5420.

Philippon, T. and V. Skreta (2012). Optimal interventions in markets with adverse selection. *American Economic Review 102*(1), 1–28.

Rothschild, M. and J. E. Stiglitz (1970). Increasing risk: I. A definition. *Journal of Economic Theory 2*(3), 225–243.

Sahin, C., J. de Haan, and E. Neretina (2020). Banking stress test effects on returns and risks. *Journal of Banking & Finance 117*, 105843.

Sangiorgi, F. and C. Spatt (2017). Opacity, credit rating shopping, and bias. *Management Science 63*(12), 4016–4036.

Schuermann, T. (2014). Stress testing banks. *International Journal of Forecasting 30*(3), 717–728.

Shahhosseini, M. (2016). The unintended consequences of bank stress tests. *Mimeo*.

Shin, H. S. (2003). Disclosures and asset returns. *Econometrica 71*(1), 105–133.

Skinner, D. J. (1994). Why firms voluntarily disclose bad news. *Journal of accounting research 32*(1), 38–60.

Skreta, V. and L. Veldkamp (2009). Ratings shopping and asset complexity: A theory of ratings inflation. *Journal of Monetary Economics 56*(5), 678–695.

Sowerbutts, R., P. Zimmerman, and I. Zer (2013). Banks disclosure and financial stability. *Bank of England Quarterly Bulletin*, Q4.

Thanassoulis, J. (2012). The case for intervening in bankers pay. *The Journal of Finance 67*(3), 849–895.

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics 5*, 179–194.

A Proofs of Main Results

A.1 General Message Spaces

In our baseline model, the firm chooses its message of investors from the set $m \in \{\theta, \emptyset\}$, where $m = \theta$ stands for a verifiable disclosure and $m = \emptyset$ stands for silence. All of the proofs in this appendix apply to a more general setting, in which the firm can choose $m \in M(\theta)$, where $M(\theta)$ defines a general set of inside disclosures that the firm can make after learning its type θ. As in our baseline model, we write the firm's utility as $u(z, \theta) - c(m)$, where $c(m)$ denotes the cost of a message $m \in M(\theta)$. We assume the cost depends only on the message: this keeps our analysis in line with most costly disclosure models, and avoids conflating with 'costly signaling' motives. Of course, this places some limitations on the scope of our results (in particular, see Online Appendix C.2). Still, we note that the model allows for some type-dependence in preferences, via u.

We continue to assume that $M(\theta)$ always contains a message denoted $m = \emptyset$, which corresponds to silence and has $c(\emptyset) = 0$. We make the technical assumptions that each $M(\theta)$ is compact and continuous in θ. More rigorously, the space $M = \cup_\theta M(\theta)$ is always endowed with some underlying topology. By continuity, we mean that the correspondence $M(\theta)$ is both upper and lower semicontinuous in the usual sense. We also assume that the sets $M^{-1}(m) = \{\theta : m \in M(\theta)\}$ are continuous in m.

We define the *worst case* disclosure payoff, which generalizes Equation (6), as follows:

$$\bar{u}(\theta) = \max_{m \in M(\theta)} \left\{ \min_{\theta' : m \in M(\theta')} u(\theta', \theta) - c(m) \right\}. \tag{19}$$

This payoff $\bar{u}(\theta)$ is the highest utility that the firm can achieve, net of disclosure costs, if investors respond skeptically, i.e., by assuming that any message m comes from the worst type that is able to send it. Under our assumptions on M, and if the cost function $c(m)$ is continuous, the Theorem of the Maximum applies, so that (19) is well-defined and is moreover an upper semi-continuous function of θ.

Finally, we assume that $\bar{u}(\theta) < u(\theta, \theta)$ for all $\theta > \underline{\theta}$. This assumption is equivalent to our assumption in the baseline model that disclosure is costly: Any message certifying that the type is at least θ must incur strictly positive costs, except (trivially) for the very worst type.

A.2 Preliminaries

In this section we introduce some auxiliary concepts and solution methods used in our main arguments. The proofs of results presented in this section are provided in Appendix B.

With a slight abuse of notation, for any test/signal structure $\{G(s \mid \theta)\}_{s\in S,\theta\in\Theta}$, we write $G(z|\theta)$ for the associated conditional distribution of grades $z = E[\theta|s]$ when the firm does not make inside disclosures. We similarly write $G(z) = \int G(z|\theta)\, dF(\theta)$ for the *unconditional* distribution of grades induced by a test G.

A.2.1 Threshold-separable tests

We introduce a class of tests important to our main analysis, characterize them, and provide a preliminary result on which our main analysis draws.

Definition 1. A test is *threshold-separable* with threshold $\theta' \in [\underline{\theta}, \overline{\theta}]$ if for any pair $\theta_1, \theta_2 \in [\underline{\theta}, \overline{\theta}]$ such that $\theta_1 < \theta' \leq \theta_2$,

$$\text{supp } G(\cdot \mid \theta_1) \cap \text{supp } G(\cdot \mid \theta_2) = \emptyset.$$

We derive an equivalent definition of threshold-separable tests as follows:

Lemma 6. *A test is threshold-separable with threshold θ' if and only if*

$$\int_{\underline{\theta}}^{r} G(z)dz \leq \int_{\underline{\theta}}^{r} F(z)dz \tag{20}$$

for all $r \in [\underline{\theta}, \overline{\theta}]$, with equality at $r = \theta', \overline{\theta}$. Moreover, any test G satisfying (20) must also satisfy $G(\theta') = F(\theta')$.

We can also show that one can focus on threshold-separable tests in the regulator's Problem 9:

Lemma 7. *For any solution $\{G(z \mid \theta)\}$ to Problem (9), there exists an alternative threshold-separable solution $\{G''(z \mid \theta)\}$ with threshold $\theta' \leq t$. Moreover, G'' induces a distribution over grades $z \geq t$ if $\theta \geq \theta'$, and otherwise reveals θ.*

Lemma 7 is not sufficient to pin down optimal tests: it says little about the distribution of z on either side of the threshold. However, in our main analysis, we leverage it to draw yet stronger conclusions about the form of *every* optimal test.

A.2.2 Solution Methods for Relaxed Problem

We now develop solution methods for the regulator's relaxed problem (12), which involves fewer constraints than her main problem (9). As Lemma 3 establishes, the solution to (12) indeed yields a weakly higher value than (9). In this sense, the latter is a relaxed version of the former.

Yet because it involves both first and second order stochastic dominance constraints, Problem (12) itself is not amenable to analysis by the standard tools of Bayesian persuasion. Nonetheless, the next lemma shows that the methods of Dworczak and Martini (2019) can be readily extended to deal with the additional FOSD constraints.[30]

Lemma 8. *If there exist a convex function $\gamma : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$, an decreasing function $\psi : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ and a distribution $H^\star \preceq_{co} F$ such that:*

1. $\gamma(x) + \psi(x) \geq v(x)$ *for all* $x \in [\underline{\theta}, \overline{\theta}]$*, with equality for* $x \in supp\ H^\star$
2. $\int \gamma(x) dH^\star = \int \gamma(x) dF$
3. $\int \psi(x) dH^\star = \int \psi(x) dF \circ \delta$

and $H^\star$ first-order stochastically dominates $F \circ \delta$, then $H^\star$ solves (9). Moreover, any solution H' to (9) must satisfy 1.-3. for the <u>*same pair*</u> (γ, ψ).

Making use of Lemma 8, we now describe the properties of any optimal distribution $H^\star$ over the firm's interim expected grades. We use these findings to prove our main results, Propositions 1 and 2, which translate these findings into implications for optimal tests. The next result establishes that an optimal solution to Problem (12) exists:

Lemma 9. *The distribution*

$$H^\star(x) = \begin{cases} F(x), & x < \theta^\star \\ F(\theta^\star), & \theta^\star \leq x < t \\ F \circ \delta(x), & t \leq x, \end{cases}$$

where $\theta^\star$ is chosen to satisfy:

$$\int_{\underline{\theta}}^{\theta^\star} x dF(x) + t(F(b) - F(\theta^\star)) + \int_{t}^{\overline{\theta}} x dF \circ \delta(x) = \int \theta dF, \tag{21}$$

[30] In a search context without any equivalent of our disclosure problem, Lyu (2021) uses a similar extension. In our setting, the presence of disclosure constraints means that reframing the problem in terms of higher-order beliefs, x, is necessary for identifying this kind of relaxed problem. Indeed, we find that optimal tests are meaningfully stochastic, where most persuasion problems with a continuous state variable find optimal signals which are deterministic.

solves problem (12).

$H^\star$ describes a distribution over the firm's conditional expected payoffs which solves Problem (12). We emphasize that $\theta^\star$ refers to a particular optimal threshold (i.e. that consistent with $H^\star$), where by contrast θ' introduced in Definition 1 denotes an arbitrary (not necessarily optimal) threshold. Of course, $H^\star$ is not a solution to (9). In particular, note that Lemma 9 does not identify whether a test G that induces this distribution of private beliefs exists, nor does it describe what such a G looks like. We establish the existence of such a test in the proof of Proposition 1 (see Appendix A.3).

Nevertheless, Lemma 8 allows us to identify some necessary properties of any solution to Problem (12). These are important to the proof of Proposition 2:

Proposition 5. ***[Properties of any solution** H' **to (12)]**. *Up to a set of measure zero:*

1. *supp* $H' \subset [\underline{\theta}, \theta^\star] \cup [t, \overline{\theta}]$,
2. $H'(x) = F \circ \delta(x)$ *for* $x \geq t$,
3. $x \geq t$ *if and only if* $\theta \geq \theta^\star$. *If* v *is strictly convex on* $[\underline{\theta}, \theta^\star]$ *then* $H'(\theta) = F(\theta)$ *for all* $x \leq \theta^\star$,

where $\theta^\star$ *is unique and identified by Lemma 9.*

In particular, we find that all optimal tests must involve (i) the same set of types enjoying an expected grade greater than t (those above the threshold $\theta^\star$ identified in Lemma 9), and (ii) the FOSD constraint binds everywhere above t.

For purposes of reference, we end this section with a brief recount of the (well-known) solution to the benchmark design problem in which the firm cannot disclose its type:

$$\begin{aligned} V^{FB} = \max_{G} \quad & \int v(z) dG(z) \\ \text{subject to} \quad & G \preceq_{co} F \end{aligned} \tag{22}$$

Corollary 2. *Suppose* $\mathbb{E}[\theta] < t$. *Any solution to (22) is characterized by a threshold* θ^{FB} *in which, with probability* 1, *types* $\theta \geq \theta^{FB}$ *(and only these types) are pooled into a single grade* $z = t$, *where* θ^{FB} *satisfies*

$$\mathbb{E}[\theta \mid \theta \geq \theta^{FB}] = t.$$

For every type $\theta < \theta^{FB}$, $\Pr[\tilde{z} \geq t \mid \theta] = 1 - \lim_{z \uparrow t} G(z \mid \theta) = 0$. *Moreover, full revelation of such types is weakly optimal.*

A.3 Proofs

Proof of Lemmas 1 and 2

Fix a test $\{G(s \mid \theta)\}_{s\in S,\theta\in\Theta}$ and some corresponding equilibrium disclosure strategies $\mu : \Theta \to \Delta M$, where $\mu(M(\theta)^c \mid \theta) = 0$.[31] The pair (G,μ) induces a distribution $\hat{G}$ over posterior means $z = \mathbb{E}[\theta \mid s,m]$, defined by

$$\begin{aligned} \hat{G}(z \mid \theta) &= \int_{\underline{\theta}}^{\overline{\theta}} \mu(M_s(z) \mid \theta) dG(s \mid \theta) \\ \hat{G}(z) &= \int_{\underline{\theta}}^{\overline{\theta}} \hat{G}(z \mid \theta) dF(\theta), \end{aligned} \tag{23}$$

where $M_s(z) := \{m \in M : \mathbb{E}[\theta \mid s,m] \leq z\}$. We note that, as μ is an equilibrium strategy, it satisfies

$$m \in \text{supp } \mu(\cdot \mid \theta) \implies m \in \arg\max_{m\in M(\theta)} \int u(\mathbb{E}[\theta \mid s,m],\theta) dG(s \mid \theta) - c(m) \tag{24}$$

for all $\theta \in [\underline{\theta},\overline{\theta}]$. Let the maximized value in (24) be $U^e(\theta)$.

We show that the test $\{\hat{G}(z \mid \theta)\}_{(z,\theta)\in[\underline{\theta},\overline{\theta}]^2}$ and disclosure strategy $\hat{\mu}(\{m = \emptyset\} \mid \theta) = 1$, for all θ, is an equilibrium, supported by the skeptical off-path beliefs identified in section 2 following messages $m \neq \emptyset$. To that end, denote $\underline{\theta}(m) = \inf\{\theta : m \in M(\theta)\}$. The payoff to $m = \emptyset$ for type θ is

$$\begin{aligned} \int u(z,\theta) d\hat{G}(z \mid \theta) &= \int\int u(\mathbb{E}[\theta \mid s,m],\theta) d\mu(m \mid \theta) dG(s \mid \theta) \\ &\geq U^e(\theta) \end{aligned}$$

where the first line uses $z = \mathbb{E}[\theta \mid s,m]$, (23) and Fubini's theorem, and the inequality follows from (24) and $c(m) \geq 0$. By contrast, for any $m \in M$ we have

$$U^e(\theta) \geq \int u(\mathbb{E}[\theta \mid s,m],\theta) dG(s \mid \theta) - c(m) \geq \int u(\underline{\theta}(m),\theta) dG(s \mid \theta) - c(m)$$

which implies $U^e(\theta) \geq \overline{u}(\theta)$, for all $\theta \in [\underline{\theta},\overline{\theta}]$, and hence $\hat{\mu}$ is an equilibrium strategy. Moreover, it is obvious by definition that $\hat{G}$ induces the same distribution over z as does the pair (G,μ) and hence the regulator's payoff is unchanged.

[31] ΔX denotes the set of probability measures on a set X, and X^c the complement of X.

Since we have just shown that the use of skeptical off-path beliefs is without loss of generality, Lemma 2 follows immediately.

Proof of Lemma 3

Let G be feasible in Problem (9). Using Lemma 7, assume without loss that G is threshold-separable with a failure threshold at some level θ'; moreover, we may assume that (i) $z = \theta$, and so $x = \theta$, for all $\theta < \theta'$, and (ii) $\Pr_G[z \geq t \mid \theta] = 1$ for $\theta \geq \theta'$. Letting H be the distribution of $x(\theta) = \mathbb{E}[z \mid \theta]$ induced by G, the value of the objective in the relaxed problem evaluated at G can be written:

$$\begin{aligned}\int v(x)dH(x) &= \Pr[\theta < \theta'] \, \mathbb{E}[v(\theta)| \, \theta < \theta'] + \Pr[\theta \geq \theta'] \, \mathbb{E}[v(\mathbb{E}[z|\theta])| \, \theta \geq \theta'] \\ &= \Pr[\theta < \theta'] \underbrace{\mathbb{E}[v(z)| \, \theta < \theta']}_{z=\theta, \text{ for } \theta<\theta'} + \Pr[\theta \geq \theta'] \underbrace{\mathbb{E}[\mathbb{E}[z|\theta]| \, \theta \geq \theta']}_{\text{by linearity for } z\geq t} \\ &= \Pr[\theta < \theta'] \, \mathbb{E}[v(z)| \, \theta < \theta'] + \Pr[\theta \geq \theta'] \underbrace{\mathbb{E}[v(z)| \, \theta \geq \theta']}_{\text{by LIE + linearity}} \\ &= \mathbb{E}[v(z)] \equiv V\end{aligned}$$

Now we show that G satisfies the constraints in Problem (12) if it satisfies those in Problem (9). By assumption G satisfies the incentive constraints in (9), and so

$$\begin{aligned}E[u(z,\theta)|\theta] &= E[(\phi z + (1-\phi)\theta)|\theta] \\ &= \phi E[z|\theta] + (1-\phi)\theta \\ &= \phi x(\theta) + (1-\phi)\theta \\ &\geq \overline{u}(\theta), \forall \theta \geq \theta'.\end{aligned}$$

Fix some $x' \geq t$. For any type $\theta \in \delta(x')$,$\bar{u}(\theta) \geq u(x', \theta)$ holds by definition. By contrast, $\bar{x}(\theta)$ satisfies $\bar{u}(\theta) = u(\bar{x}(\theta), \theta)$. Given that utility is increasing in its first argument, this implies $\bar{x}(\theta) \geq x'$ if and only if $\theta \in \delta(x')$. Similarly, the above chain of inequalities implies $x(\theta) \geq \bar{x}(\theta)$. Taken together, this means that for any value x', G must satisfy

$$\begin{aligned}\Pr_G[x(\theta) \geq x'] &\geq \Pr[\overline{x}(\theta) \geq x'] \\ &= \Pr[\theta \in \delta(x')]\end{aligned}$$

or, equivalently,

$$1 - H(x) \geq 1 - F \circ \delta(x)$$

Since x' was arbitrary, this also holds for all $x' \geq t$.

Finally, since x is a conditional expectation of z, it is trivially a mean preserving contraction of z. Hence, $H' \preceq_{co} G'$. As G' is feasible in (9), we have $G' \preceq_{co} F$. The convex order is transitive, so that $H' \preceq_{co} F$.

We briefly remark that $\mathbb{E}[w(z,\theta) \mid z] = z\lambda(z) = \mathbb{E}[u(z,\theta) \mid z]$. Hence, maximizing $\mathbb{E}[v]$ is equivalent to maximizing $\mathbb{E}[u]$. Hence, Problem (12) also constitutes a relaxed problem if the objective in Problem (9) is replaced with the firm's ex ante expected payoff.

Proof of Proposition 1

By Lemma 9 (Appendix A.2.2), we need only verify that the candidate test induces the distribution $H^\star$ over $x(\theta) = \mathbb{E}[z \mid \theta]$. First, for any value $x' \leq \theta^\star$, the test clearly induces $\Pr[x(\theta) \leq x'] = F(x')$. Moreover, the test ensures $E[v(z) \mid \theta] = E[z \mid \theta] = x(\theta) = \overline{x}(\theta)$ for any type $\theta \in \delta(t)$ (where $\overline{x}(\theta) > t$), and $x \leq t$ for $\theta \notin \delta(t)$. This implies $H^\star(x) = F \circ \delta(x)$ for $x > t$. Since probabilities must sum to 1, this pins down the required mass at t.

Proof of Proposition 2

Precisely, we prove the following two statements, which imply Proposition 2.

Lemma 10. ***[Properties of any optimal test.]*** *Up to a set of measure zero:*

1. *$z \geq t$ if and only if $\theta \geq \theta^\star$. If w is strictly convex on $[\overline{\theta}, t]$ then each $\theta < \theta^\star$ is fully revealed,*
2. *Any type $\theta \in \delta^c(t) \cap [\theta^\star, \overline{\theta}]$ receives grade t with probability 1,*
3. *Any type $\theta \in \delta(t)$ receives an expected grade $x(\theta) = \mathbb{E}[z \mid \theta] = \overline{x}(\theta)$,*

where $\theta^\star$ is unique and identified by Lemma 9.

Proof. Property 1 follows from Lemma 7 and Proposition 5.3. If $\theta \in \delta(t)$ then $\theta \geq t$ and $\overline{x}(\theta) > t$. Combining this observation with Proposition 5.2, property 3 follows. Moreover, this also implies $\Pr[x > t \mid \theta \in \delta^c(t) \cap [\theta^\star, \overline{\theta}]] = 0$. Hence, $x(\theta) = \mathbb{E}[z \mid \theta] = t$ for all $\theta \in \delta^c(t) \cap [\theta^\star, \overline{\theta}]$. Since such types cannot receive a grade less than t, this implies they must also have a zero probability of earning a grade strictly in excess of t. □

Lemma 11. *There does not exist a deterministic test which solves (9).*

Proof. We show that no deterministic test can simultaneously satisfy the necessary optimality conditions of Lemma 10 and induce a distribution G over z which constitutes a mean preserving contraction of F. For the sake of a contradiction, suppose such a deterministic mapping $\pi_d : \Theta \to S$ did exist. Then, Lemma 10.3 implies that $\pi_d(\theta) = z(\theta) = \overline{x}(\theta) < \theta$, for all $\theta \in \delta(t)$.[32] By Lemma 10.1-2 we can further conclude $z > t \iff \theta \in \delta(t)$.

Consider the conditional distributions of z and θ, respectively, given $z > t$. Since $z(\theta) = \overline{x}(\theta) < \theta$ for all $\theta \in \delta(t)$ and $z(\theta) = t$ for all $\theta \in \delta^c(t) \cap [\theta^\star, \overline{\theta}]$, we necessarily have $\Pr[z < z' \mid z > t] > \Pr[\theta < z' \mid z > t]$ for all $z' \in (t, \overline{\theta}]$. Hence, $\mathbb{E}[z \mid z > t] < \mathbb{E}[\theta \mid z > t] = \mathbb{E}[\theta \mid \theta \in \delta(t)]$, and:

$$\begin{aligned}
\mathbb{E}[z] &= \Pr[z \leq t]\mathbb{E}[z \mid z \leq t] + \Pr[z > t]\mathbb{E}[z \mid z > t] \\
&= \Pr[\theta \in \delta^c(t)]\mathbb{E}[\mathbb{E}[\theta \mid z] \mid z \leq t] + \Pr[\theta \in \delta(t)]\mathbb{E}[z \mid z > t] \\
&< \Pr[\theta \in \delta^c(t)]\mathbb{E}[\theta \mid \theta \in \delta^c(t)] + \Pr[\theta \in \delta(t)]\mathbb{E}[\theta \mid \theta \in \delta(t)] \\
&= \mathbb{E}[\theta],
\end{aligned}$$

where the second line uses $\Pr[z \leq t] = \Pr[\theta \in \delta^c(t)]$ and $z = \mathbb{E}[\theta \mid z]$, and the inequality follows after applying LIE for $z \leq t$ and $\mathbb{E}[z \mid z > t] < \mathbb{E}[\theta \mid \theta \in \delta(t)]$. We conclude $\mathbb{E}[\theta] = \mathbb{E}[\mathbb{E}[\theta \mid z]] = \mathbb{E}[z] < \mathbb{E}[\theta]$—a contradiction. □

Proof of Proposition 3

Let $G^\star_{\overline{x}}$ denote an optimal test when the outside option is described by the function $\overline{x}$, and let the corresponding equilibrium threshold be $\theta^\star_{\overline{x}}$. Fix two functions $\overline{x}, \overline{x}'$ satisfying $\overline{x}(\theta) \leq \overline{x}'(\theta)$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. To show property 1, recall condition (15) defining the threshold can be written

$$\int_{\theta^\star_{\overline{x}}}^{\overline{\theta}} (\theta - \max\{t, \overline{x}(\theta)\}) dF(\theta) = 0.$$

Direct comparison shows that $\int_k^{\overline{\theta}} (\theta - \max\{t, \overline{x}'(\theta)\}) dF(\theta) \leq 0$ for $k = \theta^\star_{\overline{x}}$. Since $\theta^\star_{\overline{x}} \leq t$, the integral is locally increasing in k and hence $\theta^\star_{\overline{x}'} \geq \theta^\star_{\overline{x}}$. Note that the inequality is strict if and only if $\Pr[\{\theta : \overline{x}'(\theta) > \overline{x}(\theta)\}] > 0$.

To see property 2, note that, by the linearity of v on $[t, \overline{\theta}]$, the regulator's payoff from test $G^\star_{\overline{x}}$ can be written

$$V(\overline{x}) = \int_{\underline{\theta}}^{k} v(\theta) dF(\theta) + \int_{k}^{\overline{\theta}} \theta dF(\theta)$$

for $k = \theta^\star_{\overline{x}}$. As $v(\theta) < \theta$ for all $\theta < t$, it is clear that the right side of this expression strictly

[32] Since the test is deterministic, we can identify the grade $z(\theta)$ as a function of θ.

decreases in k for $k \leq t$. Hence $V(\overline{x}') \leq V(\overline{x})$, with strict inequality if $\theta^\star_{\overline{x}} < \theta^\star_{\overline{x}'}$.

Finally, for comparative statics of informativeness in $\overline{x}$ we focus on the implementation of the optimal test described in Proposition 1. In this case, the result is almost immediate. By property 1, $\theta^\star_{\overline{x}'} > \theta^\star_{\overline{x}}$. Moreover, note that $\delta_{\overline{x}}(t) \subset \delta_{\overline{x}'}(t)$ for all x. The distribution $G^\star_{\overline{x}}$ satisfies

$$G^\star_{\overline{x}}(z) = \begin{cases} F(z), & z < \theta^\star_{\overline{x}} \\ F(\theta^\star) + \int_{\theta \in \delta_{\overline{x}}(t) \cap [\underline{\theta}, z]}^{z} \alpha(\theta) dF(\theta), & \theta^\star_{\overline{x}} \leq z. \end{cases}$$

As they are clearly ordered by single-crossing and have the same mean, $G^\star_{\overline{x}'}$ second-order stochastically dominates $G^\star_{\overline{x}}$ (see, for instance, Rothschild and Stiglitz (1970)).

Proof of Lemma 4

As described in section 2, message space M, costs c, and ϕ induce the function $\bar{u}(\theta)$, which defines the disclosure-equivalent grades $\bar{x}(\theta) = \phi^{-1}(\bar{u}(\theta) - (1-\phi)\theta)$.

Consider comparative statics with respect to $M(\theta)$: fix two message spaces $(M(\theta))_{\theta \in \Theta}$, $(M'(\theta))_{\theta \in \Theta}$ such that $\bigcup_\Theta M(\theta) = M = \bigcup_\Theta M'(\theta)$ and $M^{-1}(m) \subset M'^{,-1}(m)$ for all $m \in M$. Since $M^{-1}(m) \subset M'^{,-1}(m)$, we have $\underline{\theta}'(m) \leq \underline{\theta}(m)$ for all $m \in M$, where $\underline{\theta}(m) = \min\{\theta : m \in M(\theta)\}$ and $\underline{\theta}'(m)$ is similarly defined. Since this holds for all m, it follows that $\overline{u}'(\theta) \leq \overline{u}(\theta)$, where $\overline{u}'(\theta)$ has the obvious definition. From here, $\bar{x}'(\theta) \leq \bar{x}(\theta)$ follows immediately. By similar consideration of (19), it is easy to see that $\bar{x}'(\theta) \geq \bar{x}(\theta)$ if $c'(m) \leq c(m)$ for all $m \in M$.

Finally, we establish comparative statics in ϕ. Note that

$$\begin{aligned} \overline{x}(\theta) &= \phi^{-1}(\max_{m \in M(\theta)} \{u(\underline{\theta}(m), \theta) - c(m)\} - (1-\phi)\theta) \\ &= \max_{m \in M(\theta)} \{\phi^{-1}(u(\underline{\theta}(m), \theta) - c(m) - (1-\phi)\theta)\} \\ &= \max_{m \in M(\theta)} \{\underline{\theta}(m)(1 - \lambda(\underline{\theta}(m))) - \frac{1-\phi}{\phi}\theta\lambda(\underline{\theta}(m)) - \frac{1}{\phi}c(m)\}, \end{aligned}$$

where the third line uses the definition of $u(z, \theta)$. Clearly, the maximand in the last line increases in ϕ for each m, and therefore $\bar{x}(\theta)$ increases in ϕ too.

Proof of Proposition 4

Consider a family $\{F_y\}_{y \in Y}$ of distributions on $[\underline{\theta}, \overline{\theta}]$ indexed by y. In order that we can compare distributions using MLRP, assume that that each F_y has a corresponding density f_y; for simplicity let $f_y(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$.[33] For the purposes of describing some comparative

[33] Our results do not rely on this simplification. It serves to simplify our description of the MLRP order.

statics, assume Y is a totally ordered set. We order distributions by the monotone likelihood ratio (MLR) order, so that if $y < y'$ and $\theta < \theta'$ then

$$\frac{f_y(\theta')}{f_y(\theta)} < \frac{f_{y'}(\theta')}{f_{y'}(\theta)}.$$

We can now prove the result. Consider condition (15)

$$\int_{\theta^\star_y}^{\overline{\theta}} (\theta - \max\{t, \overline{x}(\theta)\}) f_y(\theta) d\theta = 0. \tag{25}$$

Fixing $\theta^\star = \theta^\star_y$, the integrand on the left-hand side again obeys the single-crossing property. Hence, for $y' > y$ the left-hand side is positive. Since $\theta^\star_y < t$, this expression is increasing in $\theta^\star$. Hence, $\theta^\star_{y'} < \theta^\star_y$. To see that the regulator's utility is increasing in y, note that her optimal payoff can be written as a function of y as follows:

$$V(y) = \int_{\underline{\theta}}^{\theta^\star_y} v(\theta) f_y(\theta) d\theta + \int_{\theta^\star_y}^{\overline{\theta}} \theta f_y(\theta) d\theta.$$

As $v(\theta)$ is increasing in θ, the right-hand side is increasing in y for fixed $\theta^\star_y$. Moreover, as $\theta > v(\theta)$ for all $\theta < t$, the right-hand side is also decreasing in $\theta^\star_y$. Putting these together shows that $V(y) < V(y')$ whenever $y < y'$.

B Proofs of Preliminary Results

Proof of Lemma 6

We first prove that, if (20) holds for all r, with equality at $r = \theta', \overline{\theta}$ then $G(\theta') = F(\theta')$. For the sake of a contradiction, suppose first that $G(\theta') > F(\theta')$. By the right-continuity of distribution functions, there exists a $r > \theta'$ such that $F(z) < G(z)$ for all $z \in [\theta', r]$. But then,

$$\int_{\underline{\theta}}^{\theta'} F(z)dz + \int_{\theta'}^{r} F(z)dz \quad < \quad \int_{\underline{\theta}}^{\theta'} G(z)dz + \int_{\theta'}^{\theta'} G(z)dz,$$

which violates (20). On the other hand, if $G(\theta') < F(\theta')$, then—since F is continuous by assumption—there is similarly a $r < \theta'$ such that $1 - G(z) > 1 - F(z)$ on $[r, \theta']$. From here, an almost identical argument, applied to the integrals $\int_r^{\overline{\theta}}(1-G(z))dz$, $\int_r^{\overline{\theta}}(1-F(z))dz$, yields another contradiction.

We now prove the equivalence between threshold-separable signals and the condition in the Lemma. If a test is threshold-separable, then it follows immediately from the mean preserving contraction criterion for signals, applied to $[\underline{\theta}, \theta')$, $[\theta', \overline{\theta}]$, respectively, that (20) holds. Similarly, it follows from Rothschild and Stiglitz (1970) applied to $[\underline{\theta}, \theta')$, $[\theta', \overline{\theta}]$ respectively that for any distribution G satisfying (20)—with equality at $r = \theta', \overline{\theta}$—there is a threshold separable signal which generates it.

The only remaining issue to prove is that no other test could generate G. For a contradiction, suppose there is a test $\{G(\cdot \mid \theta)\}_{\theta\in\Theta}$ with $\text{supp}\, G(\cdot \mid \theta) \cap \text{supp}\, G(\cdot \mid \theta) \neq \emptyset$ which induces a G satisfying (20). Hence, G must also satisfy $G(\theta') = F(\theta')$. Integrating by parts, $G(\theta') = F(\theta')$ and $\int_{\underline{\theta}}^{\theta'} G(z)dz = \int_{\underline{\theta}}^{\theta'} F(z)dz$ together imply that $\mathbb{E}_G[z \mid z \leq \theta'] = \mathbb{E}_F[z \mid z \leq \theta']$. But if G is not threshold-separable, then $\Pr_G[\theta \leq k, s \leq \theta'] \leq F(k)$ for all $k \leq \theta'$, with *strict inequality* for some k. Hence, $G(\theta') = F(\theta')$ also implies the conditional distribution $G(z \mid z \leq \theta')$ first order stochastically dominates $F(z \mid z \leq \theta')$, so that $\mathbb{E}_G[z \mid z \leq \theta'] > \mathbb{E}_F[z \mid z \leq \theta']$—a contradiction.

Proof of Lemma 7

Fix some candidate optimal test inducing a distribution G. If $G(t) = 0$ then G satisfies the hypothesis of the Lemma with $\theta' = \underline{\theta}$. So suppose from now on that $G(t) > 0$. We argue that, if there is no such threshold, then we can find another feasible, incentive compatible structure G'' which strictly increases the objective in Problem (9). We construct G'' in two steps.

Step 1: We construct an intermediate test G' which weakly improves on G, and which ensures that no type $\theta \leq t$ induces a grade $z > t$ with positive probability. Define $T := \{\theta : \theta > t\}$. By the Radon-Nikodym theorem, the random variables $\Pr[T \mid z] = 1 - G(t \mid z)$, $h_z := \mathbb{E}[\theta \mid T, z]$ and $l_z := \mathbb{E}[\theta \mid T^c, z]$ are all measurable functions of z.[34] Notice that $h_z > l_z$ for all z, and $l_z \leq t < h_z$ for all $z > t$. We construct G' by splitting each grade $z > t$ into a pair $\{t, h_z\}$ where h_z is induced with a conditional probability $(1 - G(t \mid z))\alpha_z$, with α_z defined by

$$\alpha_z = \begin{cases} 0, & z \leq t \\ 1 - \frac{G(t|z)(t-l_z)}{(1-G(t|z))(h_z-t)}, & z > t \end{cases}$$

[34]If $G(t \mid z) \in \{0, 1\}$, define $h_z = \overline{\theta}$ or $l_z = \underline{\theta}$ as appropriate.

α_z is continuous in $G(t \mid z)$, h_z and l_z for $z > t$, and so inherits measurability. Moreover, it lives in $[0,1]$ and satisfies, for all $z > t$:

$$\Pr[T \mid z]\alpha_z h_z + (1 - \Pr[T \mid z]\alpha_z)t = z$$

α_z is a conditional probability of inducing grade h_z, given the outcome of test G and $\theta \in T$. Under this new test, the implied distribution of grades is

$$G'(z) = \begin{cases} G(z), & z < t \\ G(t) + \int (1 - \alpha_{z'} \Pr[T \mid z'])dG(z'), & z = t \\ G'(t) + \int_{\{z' : h_{z'} < z\}} \alpha_{z'} \Pr[T \mid z]dG(z'), & z > t. \end{cases}$$

G' corresponds to augmenting G with a commitment to reveal $\{\theta \in T\}$ with conditional probability α_z. Since this is clearly less informative than augmenting the random variable z' induced by G with full revelation of θ, the conditional distributions trivially satisfy $G'(z \mid z') \preceq_{co} F(z \mid z')$ for each z', where $\preceq_{co}$ denotes the convex order.[35] But the convex order is preserved under integration, and so $G' \preceq_{co} F$ too.

Next, notice that

$$\begin{aligned} \int v(z)dG'(z) &= \int (\alpha_{z'} \Pr[T \mid z']v(h_{z'}) + (1 - \alpha_{z'} \Pr[T \mid z'])v(t))dG(z') \\ &= \int v(\alpha_{z'} \Pr[T \mid z']h_{z'} + (1 - \alpha_{z'} \Pr[T \mid z'])t)dG(z') \\ &= \int v(z')dG(z') \end{aligned}$$

since v is linear for $z \geq t$. Hence, G' is a feasible test which weakly improves the objective. We postpone verification of incentive compatibility to the next step.

Step 2: We construct a distribution G'' with the required properties which strictly improves on G'. Our aim is to alter the distribution of those grades $z \leq t$ that would have been realized under G' so that G'' becomes threshold-separable (Appendix A.2.1) with threshold θ'. To that end, consider the conditional distribution $G'(\theta \mid \tilde{z} \leq t)$ of types given a grade no higher than t. Applying Corollary 2 to $G'(\cdot \mid \tilde{z} \leq t)$, there is a threshold-separable distribution $G^t \preceq_{co} G'(\cdot \mid \tilde{z} \leq t)$ with some threshold θ', such that $\int v(z)dG^t(z) \geq \int v(z)dG'(z \mid \tilde{z} \leq t)$. Moreover, by Corollary 2 we may further assume $G^t(z) = G'(z \mid \tilde{z} \leq t)$ for $z \leq \theta'$, so that G^t reveals types below θ'. Now consider the

[35] With a slight abuse, here we briefly adopt the notation $G'(z \mid z')$, $F(z \mid z')$ to emphasize that we are referring to conditional distributions (i.e., given z') over *final grades* from the respective augmented tests.

distribution G'' defined by

$$G''(z) = \begin{cases} G'(t) \cdot G^t(z) & \text{, if } z \leq t \\ G'(z) & \text{, otherwise.} \end{cases}$$

Noting that $G'(z) = G'(t) \cdot G'(z \mid \tilde{z} \leq t)$, G' can be written

$$G'(z) = \begin{cases} G'(t) \cdot G'(z \mid \tilde{z} \leq t) & \text{, if } z \leq t \\ G'(z) & \text{, otherwise.} \end{cases}$$

It is clear by inspection that G'' adjusts G' by replacing the distribution $G'(z \mid \tilde{z} \leq t)$ of grades below t with G^t. Using this characterization, it follows from $G^t \preceq_{co} G'(\cdot \mid \tilde{z} \leq t)$ that $\int_{\underline{\theta}}^{z} G''(s)ds \leq \int_{\underline{\theta}}^{z} G'(s)ds$ for all z, and hence $G'' \preceq_{co} G' \preceq_{co} F$, so that G'' is a feasible distribution of grades. Moreover, by construction of G', every type $\theta \leq t$ send signals $\tilde{z} \leq t$ with probability 1 so that $\Pr_{G'}[\tilde{z} \leq t, \tilde{\theta} \leq \theta] = \Pr_{G'}[\tilde{\theta} \leq \theta] = F(\theta)$. Since $G^t(z) = G'(\cdot \mid \tilde{z} \leq t)$ for $z \leq \theta'$ this implies $G'(t) \cdot G^t(z) = G'(t) \cdot G'(z \mid \tilde{z} \leq t) = F(z)$ for all $z \leq \theta'$. Applying Lemma 6, this proves that G'' is threshold separable with threshold θ'. Moreover, since G^t assigns grade t for $\theta \geq \theta'$, this shows that if $\theta \geq \theta'$ then $G''(\cdot \mid \theta)$ induces $z \geq t$ with probability 1.

To prove incentive compatibility, we first identify a test $\{G''(\cdot \mid \theta)\}_{\theta \in \Theta}$ that induces the marginal distribution G''. In particular, consider the structure defined by

$$G''(z \mid \theta) = \mathbb{1}_{\theta \leq \theta'} \cdot \mathbb{1}_{z \geq \theta} + \mathbb{1}_{\theta > \theta'} \cdot \mathbb{1}_{z \geq t} \cdot G'(z \mid \theta).$$

To see intuitively that this test induces G'', note that for $\theta \leq \theta'$, the signal $z = \theta$ occurs with probability 1 under $G''(\cdot \mid \theta)$. By contrast, for $\theta \geq \theta'$, $G''(\cdot \mid \theta)$ reassigns all the mass under $G'(\cdot \mid \theta)$ from grades $z < t$ to t, and otherwise leaves the conditional distribution unchanged. Hence, this test indeed adjusts G' in the way described in the previous paragraph. To prove this intuition formally, we must show two things: (i) the structure does induce marginal $\Pr_{G''}[\tilde{z} \leq z] = G''(z)$ and (ii) the grades satisfy $\mathbb{E}_{G''}[\theta \mid z] = z$, for all z.[36] Integrating across θ shows that

$$\Pr_{G''}[\tilde{z} \leq z] = \int_{\underline{\theta}}^{\overline{\theta}} G''(z \mid \theta) dF(\theta) = \int_{\underline{\theta}}^{\theta'} \mathbb{1}_{z \geq \theta} dF(\theta) + \int_{\theta'}^{\overline{\theta}} \mathbb{1}_{z \geq t} G'(z \mid \theta) dF(\theta).$$

For $z < t$, the second integral is 0 and the expression reduces to $\Pr_{G''}[\tilde{z} \leq z] = F(\min\{z, \theta'\}) =$

[36]For each z, G'' is the sum of products of measurable functions, and hence measurable.

$G''(z)$. For $z \geq t$, the second integral reduces to $\int_{\theta'}^{\bar{\theta}} G'(z \mid \theta) dF(\theta) \equiv \Pr_{G'}[\tilde{z} \leq z, \tilde{\theta} > \theta']$. Recalling that for all $z \geq t$, $F(\theta') = G'(t) \cdot G'(z \mid \tilde{z} \leq t) = \Pr_{G'}[\tilde{z} \leq t, \theta \leq \theta'] = \Pr_{G'}[\tilde{z} \leq z, \theta \leq \theta']$, the right-hand side reduces to

$$\Pr_{G'}[\tilde{z} \leq z, \theta \leq \theta'] + \Pr_{G'}[\tilde{z} \leq t, \theta > \theta'] = G'(z).$$

Recalling $G'(z) = G''(z)$ for such z establishes that the signal structure induces G''. To show (ii), we note that it suffices to prove $\mathbb{E}_{G''}[\theta \mid t] = t$: indeed, for any $z \neq t$ $G''(\cdot \mid \theta)$ is either fully revealing or identical to $G'(\cdot \mid \theta)$, and so it is not difficult to see that $\mathbb{E}_{G''}[\theta \mid z] = \mathbb{E}_{G'}[\theta \mid z] = z$.[37] If $\Pr_{G''}[\tilde{z} = t] = 0$, then without loss of generality we can simply assign $\mathbb{E}_{G''}[\theta \mid t] = t$.[38] So suppose $\Pr_{G''}[s = t] = 0$. Then, by construction the distribution of grades $z \leq t$ under distribution G'' satisfies

$$\int_{\{\theta < t\}} \theta dF + t \Pr_{G''}[\tilde{z} = t] = G'(t) \cdot \int_{\underline{\theta}}^{\bar{\theta}} \theta dG'(\theta \mid \tilde{z} \leq t). \tag{26}$$

However, by the law of iterated expectations, the distribution of $\mathbb{E}[\theta \mid z]$ induced by test $\{G''(\cdot \mid \theta)\}$ satisfies

$$\int_{\{z<t\}} \int_{\underline{\theta}}^{\bar{\theta}} \theta G''(d\theta, dz) = \int_{\underline{\theta}}^{t} \mathbb{E}[\theta \mid z] dG''(z) = \int_{\{\theta<t\}} \theta dF + \mathbb{E}_{G''}[\theta \mid t] \Pr_{G''}[\tilde{z} = t], \tag{27}$$

where $G''(\theta, z)$ denotes the joint distribution of $(\tilde{\theta}, \tilde{z})$ induced by $\{G''(\cdot \mid \theta)\}$, and the second equality follows from the threshold-separability of G''. Rearranging the right-hand side shows that

$$\int_{\{z<t\}} \int_{\underline{\theta}}^{\bar{\theta}} \theta G''(d\theta, ds) = \int_{\underline{\theta}}^{\bar{\theta}} \theta G''(d\theta, t) = G'(t) \cdot \int_{\underline{\theta}}^{\bar{\theta}} \theta dG^t(\theta)$$

But since G^t is a MPC of $G'(\cdot \mid \tilde{z} \leq t)$, $\int_{\underline{\theta}}^{\bar{\theta}} \theta dG^t(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} \theta dG'(\theta \mid \tilde{z} \leq t)$. Given this, it is clear on inspection of equations (26) and (27) that $t = \mathbb{E}_{G''}[\theta \mid t]$.

Finally, we show that G'' is incentive compatible. For $\theta < \theta'$, G'' reveals θ. As $\overline{u}(\theta) \leq u(\theta, \theta)$ for all θ, incentive compatibility is trivially satisfied for all $\theta \leq \theta'$. For $\theta \in (\theta', t)$,

[37] One can directly verify that the conditional expectation $\mathbb{E}_{G''}[\theta \mid z]$ satisfies the required properties for all z by finding a valid conditional distribution (Radon-Nikodym derivative) $G''(z \mid \theta)$ associated with $G''(z \mid \theta)$. While not difficult, it is tedious and hence we omit it. Details available on request.

[38] See, for example, Billingsley (1995), pg. 445.

$z = t$, which is again trivially incentive compatible. Finally, for $\theta \geq t$ we have

$$\begin{aligned} \int u(z,\theta) dG''(z \mid \theta) &\geq \int u(z,\theta) dG'(z \mid \theta) \\ &= \int (\alpha_{z'} u(h_{z'},\theta) + (1-\alpha_{z'})t) dG(z' \mid \theta) \\ &\geq \int u(z',\theta) dG(z' \mid \theta) \end{aligned}$$

where the first inequality follows as u is increasing in z and $G''(\cdot \mid \theta)$ first-order stochastically dominates $G(\cdot \mid \theta)$, the second uses the definition of G' and the third uses the linearity of $u(z,\theta)$ for $z \geq t$, $h_z > t$ and $\alpha_{z'} \Pr[T \mid z']h_{z'} + (1 - \alpha_{z'} \Pr[T \mid z'])t = z'$ to establish the integrands are ordered pointwise in z'. Hence, G'' is incentive compatible if G is.

Proof of Lemma 8

Suppose such a γ, ψ and $H^\star$ exist. Then, for any feasible distribution H we have

$$\begin{aligned} \int v dH &= \int (v - \gamma - \psi) dH + \int \gamma dH + \int \psi dH \\ &\leq \int \gamma dF + \int \psi dF \circ \delta \\ &= \int v dH^\star. \end{aligned}$$

where the inequality follows by applying condition (1.) to the first term on the right-hand side, noting $H \preceq_{co} F$ for the second, and noting that H first-order stochastically dominates $F \circ \delta$ for the third. The final line follows immediately from application of (1.)-(3.). It is straightforward to see that any optimal H'must meet conditions (1.)-(3.) (and the constraints in Problem (12)) in order to satisfy the above with equality.

Proof of Lemma 9

We identify functions $\gamma^\star, \psi^\star$ for which the verification result of Lemma 8 applies. Consider the following functions:

$$\gamma^\star(x) = \begin{cases} v(x), & x \leq \theta^\star \\ v(\theta^\star) + \gamma^+(x - \theta^\star), & x > \theta^\star \end{cases}$$

where $\gamma^+ = \frac{v(t)-v(\theta^\star)}{t-\theta^\star}$, and

$$\psi^\star(x) = \begin{cases} 0, & x < t \\ v(x) - \gamma(x), & x \geq t \end{cases}$$

Clearly, $\gamma^\star$ is convex, $\psi^\star$ is decreasing and $\gamma^\star + \psi^\star \geq v$ (with equality on the support of $H^\star$). As $H^\star$ is identical to F on $[0, \theta^\star]$ and $\gamma^\star$ linear on $[\theta^\star, \overline{\theta}]$, we have $\int \gamma^\star dH^\star = \int \gamma^\star dF$. By a similar argument, $\int \psi^\star dH^\star = \int \psi^\star dF \circ \delta$. $H^\star$ trivially satisfies the FOSD constraint. Finally, given the definition of $\theta^\star$, and since $F(x) \leq F \circ \delta(x)$ for $x \geq t$, it is easy to verify (by a single-crossing argument) that $\int_{\underline{\theta}}^{x'} H^\star(x)dx \leq \int_{\underline{\theta}}^{x'} F(x)dx$ for all $x' \in [\underline{\theta}, \overline{\theta}]$.

Proof of Proposition 5

Properties (1) and (2) follow immediately by application of conditions (1)-(3) of Lemma 8 to $\gamma^\star$, $\psi^\star$. To prove property (3), we note that $\gamma^\star$ can be written

$$\gamma^\star(x) = \tilde{\gamma}(x) + \lambda \max\{0, x - \theta^\star\}$$

where $\tilde{\gamma}$ is the convex function defined as follows: $\tilde{\gamma}(x) = v(x)$ for $x \leq \theta^\star$, and $\tilde{\gamma}(x) = \gamma^\star(\theta^\star) + \frac{d\gamma^\star}{d\theta^\star}_- \cdot (x - \theta^\star)$ where $\frac{d\gamma^\star}{d\theta^\star}_- := \lim_{x \uparrow \theta^\star} \frac{\gamma^\star(\theta^\star) - \gamma^\star(x)}{\theta^\star - x}$ for $x > \theta^\star$, and $\lambda = \gamma^+ - \frac{d\gamma^\star}{d\theta^\star}_- > 0$. If test H is not threshold-separable, then by Lemma 6 (Appendix A.2.1), $\int_{\underline{\theta}}^{\theta^\star} H(x)dx < \int_{\underline{\theta}}^{\theta^\star} F(x)dx$. This implies

$$\begin{aligned} \int \gamma^\star dH &= \int \tilde{\gamma} dH + \lambda \int \max\{0, x - \theta^\star\} dH \\ &= \int \tilde{\gamma} dH + \lambda \int_{\theta^\star}^{\overline{\theta}} (1 - H(x))dx \\ &< \int \tilde{\gamma} dF + \lambda \int_{\theta^\star}^{\overline{\theta}} (1 - F(x))dx \\ &= \int \gamma^\star dF \end{aligned}$$

where the strict inequality follows by convexity of $\tilde{\gamma}$ and $\int_{\underline{\theta}}^{\theta^\star} H(x)dx < \int_{\underline{\theta}}^{\theta^\star} F(x)dx$. But this contradicts condition 2. of Lemma 8.

C Online Appendix: Extensions

C.1 Alternative Timing Assumptions

C.1.1 Ex Post Disclosures

In this section we solve for optimal tests when the firm chooses its disclosures *after* observing the result s of the regulator's test. This is the only substantive change relative to section 2. To simplify the exposition, we present the case in which *(i)* $\lambda = \beta$ for $z < t$, where $\beta \in [0, 1)$, and *(ii)* the disclosure-equivalent grade $\overline{x}(\theta)$ is increasing in type. Note that *(i)* corresponds to a piecewise linear v; *(ii)* implies that high types require higher grades in order to prevent disclosure. We focus on the interesting case where $\overline{x}^{-1}(t) < \overline{\theta}$, so that some types would reject the threshold grade t.[39] In addition, we make the technical assumptions that $\overline{x}$ is differentiable for $\theta \geq \overline{x}^{-1}(t)$, and that the prior distribution F of types admits a positive, Lipschitz continuous density f. We refer to such distributions as "admissible". Admissibility plays a role in Lemma 13 below. After proving the main result of this section (Proposition 6), we discuss generalizations.

Even with ex post disclosures, the "*public speech, private silence*" principle still applies (see proof of Lemma 15). Accordingly, the regulator's problem can be written simply as:

Lemma 12. *When banks can make ex post disclosures, the designer's problem is*

$$V^{SB} = \max_{G} \quad E_G\left[v\left(z\right)\right] \quad \textit{subject to}$$
$$u\left(z, \theta\right) \quad \geq \bar{u}\left(\theta\right), \ \forall (\theta, z) \in \textit{supp } G. \tag{28}$$

Clearly, problem (28) involves at least as many constraints as problem (9). Hence, we refer to any solution to (28) as a *second-best test.* Before stating the main result of this section, we define a helpful class of tests:

Definition 2. For some thresholds $\theta' < \theta''$ with $\theta' \leq t$, a **threshold negative-assorting pass (T-NAP)** test assigns grades *deterministically* as follows:

$$z(\theta) = \begin{cases} \theta, & \text{for } \theta < \theta' \\ s(\theta), & \text{for } \theta' \leq \theta < \theta'' \\ \mathbb{E}[\theta \mid \theta \in \{\theta'', \overline{x}^{-1}(\max\{t, \theta''\})\}], & \text{for } \theta'' \leq \theta < \overline{x}^{-1}(\max\{t, \theta''\}) \\ \overline{x}(\theta), & \text{for } \overline{x}^{-1}(\max\{t, \theta''\}) \leq \theta \leq \overline{\theta}, \end{cases} \tag{29}$$

[39] If $\overline{x}^{-1}(t) \geq \overline{\theta}$, the solution is trivial: Corollary 2 applies.

where $s(\theta) : [\theta', \theta'') \to (t, \overline{x}(\overline{\theta})]$ is strictly decreasing, differentiable function, and the thresholds are chosen so that $\mathbb{E}[\theta \mid \theta'' < \theta < \overline{x}^{-1}(\max\{t, \theta''\})] \geq t$.

Remark. A heuristic construction of a T-NAP test is helpful; Lemma 13 establishes the formalities. First, a T-NAP identifies a threshold $\theta' \leq t$ below which types are revealed (and hence failed). There are many passing grades, assigned as follows: the lowest passing type θ' gets grade $s(\theta') = \overline{x}(\overline{\theta})$, and is thereby pooled with the highest type $\overline{\theta}$ (who receives the same grade); the grade makes $\overline{\theta}$'s ex post disclosure constraint bind. Thereafter, a similar pairwise matching process—indexed by y, the next lowest type to be matched—continues: if all types in a range $[\theta', y)$ have so far been pooled with a unique 'high' type in $(\overline{x}^{-1}(s(y)), \overline{\theta}]$, then type y gets the grade $s(y) = \overline{x}(\theta_y)$—and is thereby pooled with the highest unmatched type $\theta_y := \overline{x}^{-1}(s(y))$. As Lemma 13's proof makes clear, the function s determines the 'speed' at which low types are matched to high ones, and is chosen specifically to ensure (i) grades are calibrated so that $z = \mathbb{E}[\theta \mid z]$, and (ii) the grade assigned to pair (y, θ_y) makes θ_y's disclosure constraint bind.

This matching continues until $y = \theta''$, where pooling opportunities are exhausted. This may occur in one of two ways: (i) there are no more high types with binding disclosure constraints (i.e., $\theta_{y=\theta''} = \overline{x}^{-1}(t)$), or (ii) remaining low types are sufficiently good that pooling them with the remaining higher types cannot violate any disclosure constraints (i.e., $\theta_{y=\theta''} = \overline{x}^{-1}(\theta'')$). The process is completed by pooling the remaining types $\theta \in [\theta'', \overline{x}^{-1}(\max\{t, \theta''\})]$ into a single grade. To be considered a T-NAP, this pooled grade must exceed t: we show below that this can be ensured, so long as θ' is chosen large enough.

T-NAP tests clearly always satisfy the constraints in problem (28). Moreover:

Proposition 6. *There exists a T-NAP test with threshold $\theta' = \theta^{SB}$ which solves problem (28). θ^{SB} always weakly exceeds $\theta^\star$. If $\theta^{SB} > \underline{\theta}$, then $\theta^{SB} > \theta^\star$ and moreover $\theta'' < t$.*

Proposition 6 is stated in terms of the mapping from types to grades, but problem (28) is expressed in terms of the distribution G. Still, the latter is easily derived from the former. For example, in the most interesting case where financial conditions are weak enough that some types must fail, we have:

Corollary 3. *If $\theta^{SB} > \underline{\theta}$, then the test in Proposition* 1 *induces grade distribution*

$$G^{SB}(z) = \begin{cases} F(z), & \textit{for } \underline{\theta} \leq z < \theta^{SB} \\ F(\theta^{SB}), & \textit{for } \theta^{SB} \leq z < t \\ F(\overline{x}^{-1}(t)) - F(\theta''), & \textit{for } z = t, \\ F(\overline{x}^{-1}(z)) - F(s^{-1}(z)), & \textit{for } t \leq z \leq x(\overline{\theta}). \end{cases}$$

The following two lemmas are integral to the proof of Proposition 6.

Lemma 13. *For any admissible F, the set of T-NAP tests is non-empty. In particular, there exists a T-NAP test with a minimal threshold, $\theta' = \theta^{SB}$.*

Proof. Fix a threshold θ', and consider mappings $z(\cdot)$ of the form (29). We first show that θ' uniquely pins down the function $s(\cdot)$ and threshold θ''. Existence of a T-NAP with threshold θ' then reduces to checking whether or not $\mathbb{E}[\theta \mid \theta'' < \theta < \overline{x}^{-1}(\max\{t, \theta''\})] \geq t$. We show there exists a minimal $\underline{\theta} \leq \theta^{SB} < t$ such that the latter condition holds for all $\theta' \geq \theta^{SB}$.

Fix some $y > \theta'$ and consider the class of decreasing, differentiable functions $s : [\theta', y] \to [s(y), s(\theta')]$ with $s(\theta') = \overline{x}(\overline{\theta})$ and $s(y) > y$. Below, we identify the unique value of y consistent with the T-NAP. It will be helpful to associate with s the unique function $\sigma : [\overline{x}^{-1}(s(y)), \overline{\theta}] \to [\theta', y]$ defined by

$$\sigma(\theta) = s^{-1}(\overline{x}(\theta)).$$

If $z(\cdot)$ assigns grades to types in $[\theta', y]$ according to $s(\cdot)$ (recall (29)), then for any given 'high' type $\theta \in [s(y), \overline{\theta}]$ the function σ identifies the 'low' type paired with θ to receive grade $z(\theta) = \overline{x}(\theta)$; obviously, σ is decreasing.

As F has a density f and s, $\overline{x}$ are differentiable, the conditional expectation $\mathbb{E}[\tilde{\theta} \mid z = \overline{x}] = \mathbb{E}[\tilde{\theta} \mid \tilde{\theta} \in \{s^{-1}(z), \overline{x}^{-1}(z)\}]$ can be directly calculated from the joint density of z and θ. It is easy to see from the change-of-variables formula (see Billingsley (1995), pg. 225) that (locally around θ) the density of $z(\theta) = \overline{x}(\theta)$ is $f(\theta)\overline{x}'(\theta)$. Similarly, the density (local to $\sigma(\theta)$) is $f(\sigma(\theta)) \times |\sigma'(\theta)| \times \overline{x}'(\theta)$. Hence,

$$\mathbb{E}[\tilde{\theta} \mid z = \overline{x}] = q_\theta \cdot \theta + (1 - q_\theta) \cdot \sigma(\theta), \tag{30}$$

where $q_\theta = \frac{f(\theta)}{f(\theta) - f(\sigma(\theta)) \cdot \sigma'(\theta)}$ is the conditional probability of type θ given $\tilde{\theta} \in \{\theta, \sigma(\theta)\}$.[40] Note $\sigma' < 0$ implies $q_\theta \in (0, 1)$. Recall a T-NAP requires $s(\cdot)$ calibrates grades to ensure that:

$$\overline{x}(\theta) = \mathbb{E}[\tilde{\theta} \mid z = \overline{x}] = q_\theta \cdot \theta + (1 - q_\theta) \cdot \sigma(\theta), \text{ for any } \theta \geq \overline{x}^{-1}(s(y)). \tag{31}$$

Upon rearrangement, (31) can be expressed as a differential equation in σ:

$$\sigma'(\theta) = -\frac{f(\theta)}{f(\sigma(\theta))} \cdot \left(\frac{\theta - \overline{x}(\theta)}{\overline{x}(\theta) - \sigma(\theta)} \right). \tag{32}$$

As s is decreasing, $s(y) > y$ is equivalent to $\overline{x}(\theta) > \sigma$ for all $(\theta, \sigma) \in [\overline{x}^{-1}(s(y)), \overline{\theta}] \times [\theta', y]$.

[40] We use $\tilde{\theta}$ here to emphasize the distinction between the random variable and its realization θ.

Together with the admissibility of F, this implies the right hand side of (32) is continuous in (θ, σ) and bounded away from 0 on $[\overline{x}^{-1}(s(y)), \overline{\theta}] \times [\theta', y]$. It is even Lipschitz continuous in σ.[41] As a result, we can apply the Picard-Lindelöf theorem along with the boundary condition $\sigma(\overline{\theta}) = \theta'$ to conclude that there exists a unique solution $\sigma^{SB}(\cdot; \theta')$ to (32) on $[\overline{x}^{-1}(s(y)), \overline{\theta}]$ and hence, obviously, a unique corresponding $s^{SB}(\cdot; \theta')$ on $[\theta', y]$. We parameterize σ^{SB} and s^{SB} to emphasize that the boundary condition at θ' fully determines these functions; moreover, we note that σ^{SB} and s^{SB} are each pointwise continuous and increasing in θ'.[42]

To finish the description of the unique candidate T-NAP with threshold θ', we must pin down the value of y. There are two cases to consider. *Case (i)*: there exists a $k < t$ with $s^{SB}(k; \theta') = t$. In this case, set $y = \theta'' = k$. Note that $s^{SB}(k; \theta') = t$ implies that the range of s^{SB} is $[t, \overline{x}(\overline{\theta})]$, which means $z(\theta) = \overline{x}(\theta)$ for $\theta \geq \overline{x}^{-1}(s^{SB}(k; \theta')) = \overline{x}^{-1}(t)$, as required for a T-NAP with $\theta'' < t$. Because s^{SB} is decreasing, k is the *unique* candidate for θ'' satisfying the requirement $s^{SB}(\theta''; \theta') = \max\{\theta'', t\}$. *Case (ii)*: $s^{SB}(\theta; \theta') > t$ for all $\theta \leq t$. In this case, as s^{SB} is decreasing and continuous in its first argument, there exists a k such that $k = s^{SB}(k; \theta')$. As above, $\theta'' = k$ therefore uniquely satisfies the domain criteria in (29).

We have shown that for *any* θ', there is a unique function satisfying (29) and $\mathbb{E}[\theta \mid z] = z$. We now show that there exists a smallest θ' for which z also satisfies

$$\mathbb{E}[\theta \mid \theta \in \left\{\theta'', \overline{x}^{-1}(\max\{t, \theta''\})\right\}] \geq t. \tag{33}$$

If (33) satisfied for $\theta' = \underline{\theta}$ then we are done. Suppose (33) is violated at $\theta' = \underline{\theta}$. Then, it must be that $\theta'' < t$—otherwise, the left side of ((33)) would trivially exceed the right. Hence, *case (i)* must apply at $\theta' = \underline{\theta}$, so that θ'' is determined by $s^{SB}(\theta''; \theta') = t$. Noting that $s^{SB}(\theta''; \theta')$ is continuously increasing (decreasing) in θ' (resp., θ''), θ'' is therefore continuously increasing in θ'.[43] As F is continuous, this implies the left-side of (33) increases continuously in θ'. As $\theta' \leq \theta''$, there must exist a $\theta' \in (\underline{\theta}, t)$ such that $\theta'' = t$–by which point, the left-side of (33) exceeds the right. Hence, by continuity, there exists a $\theta^{SB} \in (\underline{\theta}, t)$ such that (33) holds with equality. This is precisely the minimal threshold described in the lemma. □

Similar to problem (9), the tools of duality can be used to solve problem (28). As we no longer need to account for the test's influence on the firm's higher order beliefs, (28) is somewhat simpler than (9). Yet, it is not a straightforward Bayesian persuasion problem *à*

[41]The functions $f(h)$ and $\overline{x} - \sigma$ are each bounded and Lipschitz continuous on the compact set $[\overline{x}^{-1}(s(y)), \overline{\theta}] \times [\theta', y]$. Hence, their product is Lipschitz. Moreover, the function $y(x) = 1/x$ is Lipschitz continuous on $[a, \infty]$ for any $a > 0$. As a composition of Lipschitz functions is Lipschitz, so is the RHS of (32); the Lipschitz constant can be uniformly bounded in θ by continuity-compactness arguments.

[42]This follows from contraction mapping arguments *à la* Picard-Lindelöf; details available on request.

[43]As $\overline{x}$ is continuous, s^{SB} inherits these properties from σ^{SB}.

la Dworczak and Martini (2019), either: the distribution of posterior means is insufficient to characterize its solution because it matters for incentives *which* types receive each grade.

To deal with this, we rewrite (28). First, we temporarily index grades according to the disclosure-equivalent grade of largest type $j \in [\underline{\theta}, \overline{\theta}]$ who may be assigned that grade.[44] Formally, j is the *conditional essential supremum* of θ, given the grade. Because we assumed $\overline{x}$ is increasing, this 'highest' type captures the *relevant* incentive constraints.

Second, with a small abuse of notation we write the design problem as a choice of a *joint* distribution $G(\theta, \overline{x}(j))$ over the type θ and the (*indexed*) grade $\overline{x}(j)$. We write $z(j) = \mathbb{E}[\theta \mid s = \overline{x}(j)]$ for the conditional expected type, given signal $\overline{x}(j)$. Then problem (28) becomes:

$$V^{SB} = \max_{G(\cdot,\cdot)} \quad \beta \mathbb{E}_F[\theta] + (1-\beta) \int_{\overline{x}(j) \geq t} z(j) dG(\tilde{\theta}, \overline{x}(j)) \quad \text{subject to} \tag{34}$$

$$\int_{\tilde{\theta} > \theta, j \leq \overline{\theta}} dG(\tilde{\theta}, \overline{x}(j)) = 1 - F(\theta), \quad \forall \theta \in [\underline{\theta}, \overline{\theta}] \tag{35}$$

$$\int_{j \in [\theta_1, \theta_2]} z(j) dG(\overline{x}(j)) = \int_{\theta \in [\underline{\theta}, \overline{\theta}], j \in [\theta_1, \theta_2]} \tilde{\theta} dG(\tilde{\theta}, \overline{x}(j)), \quad \forall \theta_1 \leq \theta_2 \tag{36}$$

$$G(\theta, \overline{x}(j)) = G(j, \overline{x}(j)), \quad \forall \theta \geq j \tag{37}$$

$$\int_{[j_1, j_2]} z(j) dG(\overline{x}(j)) \geq \int_{[j_1, j_2]} \overline{x}(j) dG(\overline{x}(j)), \quad \forall j_1 \leq j_2. \tag{38}$$

Constraint (35) requires the joint distribution of signals and types be consistent with the marginal distribution of types (Kamenica and Gentzkow (2011)); (36) requires the $z(j)$ reflect investors' posterior means; (37) ensures that the essential supremum corresponding to a grade indexed $\overline{x}(j)$ is in fact no greater than j. Finally, the "public speech, private silence" principle is G-a.e.-equivalent to (38).

As v is piecewise linear, the objective function can be written in two parts: a linear component weighted by β, and a $(1-\beta)$-weighted gain for inducing a grade $z \geq t$. Due to (36), the linear component is pinned down by $\mathbb{E}_F[\theta]$. Due to IC, the gain is necessarily earned for every grade indexed $\overline{x}(j) > t$. Due to linearity for $z \geq t$, we can (without loss) assume at most one additional grade—which we index $\max\{t, \theta''\}$, where θ'' is the unique value corresponding to $\theta' = \theta^{SB}$ in Lemma 13—experiences the gain.[45]

Framed this way, problem (28) becomes amenable to a duality analysis. We show how to find the optimal test in the interesting case where $\theta^{SB} > \underline{\theta}$ (the other case is similar, and so omitted). The joint distribution $G^\star$ over types θ and indexed grades $\overline{x}(j)$ associated with

[44] This is without loss of optimality: take any incentive compatible (IC) test, and adjust every grade $z \leq t$ as per Corollary 2 for an IC improvement. Next, pool by j all signals z whose conditional essential supremum $j_z = j$ satisfies $\overline{x}(j) \geq t$. This clearly retains IC, and leaves the objective unchanged (v is linear for all such signals). Finally, pool any grade with $z \geq t$ and $\overline{x}(j_z) < t$ for a similar weak improvement. The resulting test induces at most one signal for each possible j, and hence can be indexed by $\overline{x}(j)$.

[45] See footnote 44. In the optimal test, this index will reflect the appropriate supremal type. The reader might wonder if it unnecessarily tightens the constraint (37). It does not – if we relax the problem by indexing this grade $\overline{\theta}$, the solution would still be the one described in Proposition 6.

the T-NAP described in Lemma 13 is

$$G^\star(\theta, x) = \begin{cases} F(\min\{\theta, \overline{x}^{-1}(x)\}), & \text{for } \overline{x} < t \\ F(\theta^{SB}) + F(\min\{\theta, \overline{x}^{-1}(x)\}) - F(\sigma\left(\overline{x}^{-1}(x)\right)), & \text{otherwise.} \end{cases}$$

$G^\star$ induces a distribution over $z(j)$, and it is easy to see that this distribution is exactly G^{SB} (recall Corollary 3). Note that $G^\star$ trivially satisfies (35) and (37). Moreover, by construction it is chosen so that $z(j)$ satisfies (36) and (38) (recall (31)).

Lemma 14. *Function $G^\star$ solves problem (34)-(38) if there exists a non-increasing function $\Lambda : [t, \overline{x}(\overline{\theta})] \to \mathbb{R}_+$, and a function $M : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ of bounded variation, such that:*

$$\theta - \Lambda(\overline{x}(j)) \cdot (\overline{x}(j) - \theta) - M(\theta) \leq 0, \qquad \forall \theta < j \text{ s.t. } \overline{x}(j) \geq t \tag{39}$$

with strict equality for $(\theta, \overline{x}(j)) \in supp\ G^\star$ with $\overline{x}(j) \geq t$, and

$$\int_t^{\overline{x}(\overline{\theta})} \int_t^{\overline{x}} (\overline{x}(j) - \tilde{\theta}) dG(\tilde{\theta}, \overline{x}(j)) d\Lambda(\overline{x}) \quad \leq 0, \tag{40}$$

for any $G(\cdot,\cdot)$ satisfying (35)-(38), with strict equality at $G^\star$, and $M(\theta) = 0$ for $\theta \leq \theta^{SB}$.

Proof. Adding and subtracting terms, the objective can be written (up to a constant)

$$\begin{aligned} \int_{\overline{x}(j)\geq t} z(j) dG(\tilde{\theta}, \overline{x}(j)) - \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(j)\geq t} \Lambda(\overline{x}(j))(\overline{x}(j) - \tilde{\theta}) dG(\tilde{\theta}, \overline{x}(j)) - \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(\underline{\theta})}^{\overline{x}(\overline{\theta})} M(\theta) dG(\theta, \overline{x}(j)) \\ + \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(j)\geq t} \Lambda(\overline{x}(j))(\overline{x}(j) - \tilde{\theta}) dG(\tilde{\theta}, \overline{x}(j)) + \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(\underline{\theta})}^{\overline{x}(\overline{\theta})} dM(\theta) dG(\theta, \overline{x}(j)), \end{aligned} \tag{41}$$

which can be rewritten after some algebra as

$$\begin{aligned} \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(j)\geq t} (\theta - \Lambda(\overline{x}(j)) \cdot (\overline{x}(j) - \tilde{\theta}) - M(\theta)) dG(\tilde{\theta}, \overline{x}(j)) - \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(j)< t} M(\theta) dG(\tilde{\theta}, \overline{x}(j)) \\ + \int_{\overline{x}\geq t} \int_{\underline{\theta}}^{\overline{\theta}} \int_t^{x} (\overline{x}(j) - \theta) dG(\theta, \overline{x}(j)) d(-\Lambda(x)) + \Lambda(\overline{x}(\overline{\theta})) \cdot \int_{\underline{\theta}}^{\overline{\theta}} \int_{\overline{x}(j)\geq t} (\overline{x}(j) - \theta) dG(\theta, \overline{x}(j)) \\ + \int_{\underline{\theta}}^{\overline{\theta}} (1 - F(\theta)) dM(\theta). \end{aligned} \tag{42}$$

The first line follows on collection of the first three integrals in (41), after using (36) to express the first integral in terms of θ. For the second line, we substitute $\Lambda(\overline{x}(\overline{\theta})) - \Lambda(\overline{x}(j)) = \int_{\overline{x}(j)}^{\overline{x}(\overline{\theta})} d\Lambda(x)$ into the fourth integral in (36), and invoke Fubini's theorem to change the order of integration. The final line uses (35) to simplify the last integral in (41).

Under condition (39), the first term is non-positive for any G satisfying (37). $M \geq 0$ implies the second term in (42) is non-positive too. Finally, for any G satisfying (38), so too are the terms on the second line: as Λ is non-increasing and non-negative, this follows from (38) evaluated at $\theta_1 = t$ and $\theta_2 \in \{x, \overline{x}(\overline{\theta})\}$, after using (36) to replace z. Hence, the objective is bounded above by $\int_{\underline{\theta}}^{\overline{\theta}} (1 - F(\theta))\, dM(\theta)$. By contrast, (39) holds with equality on the support of $G^\star$ by hypothesis. Moreover, the construction in Lemma 13 makes clear that $G^\star$ satisfies $z(j) = \overline{x}(j)$ for all $\overline{x}(j) \geq t$. From these observations, it follows that the first and third terms of (42) are 0 under $G^\star$. Finally, $M(\theta) = 0$ for $\theta \leq \theta^{SB}$ implies that the second integral is 0 under $G^\star$ (the set $\{\theta > \theta^{SB}\} \cap \{\overline{x}(j) < t\}$ obviously has $G^\star$-probability 0). We conclude that $G^\star$ attains the upper bound and hence is optimal. □

With Lemmas 13 and 14 in hand, we are now ready to prove Proposition 6.

Proof of Proposition 6

Proof. We identify $\langle \Lambda, M \rangle$ for which $\langle G^\star, \Lambda, M \rangle$ satisfies the conditions of Lemma 14. Put

$$\begin{aligned} \Lambda(x) &= \left(\frac{\theta^{SB}}{\overline{x}(\overline{\theta}) - \theta^{SB}} \right) e^{\int_x^{\overline{x}(\overline{\theta})} \left(\tilde{x} - s^{-1}(\tilde{x}) \right)^{-1} d\tilde{x}} \\ M(\theta) &= \max \left\{ \max_{x \geq \max\{t, \overline{x}(\theta)\}} \{ y(\theta, x) \}, 0 \right\}, \end{aligned}$$

where $y(\theta, x) = \theta - \Lambda(x) \cdot (x - \theta)$. As $\theta^{SB} < \overline{x}(\overline{\theta})$, note that $\Lambda(x) \geq 0$ on its domain and is non-increasing, because $\tilde{x} - s^{-1}(\tilde{x}) > 0$ for $\tilde{x} \geq t$.[46] $M(\theta)$ is also non-negative. The proposed $\langle \Lambda, M \rangle$ clearly satisfy inequality (39); the incentive constraints (38) guarantee (40) holds; recalling from (31) that the incentive constraints bind for all $j \geq \overline{x}^{-1}(t)$ under $G^\star$, (40) is moreover equal to 0.

Under the proposed $\Lambda(x)$, $y(\theta, x)$ is maximized at $x = z(\theta)$ for each $\theta \geq \theta^{SB}$, where z is defined by (29) and x is subject to the constraint $x \geq \max\{t, \overline{x}(\theta)\}$. To see this, consider first $\theta \in [\theta^{SB}, \theta'')$. Taking the derivative $\frac{\partial y}{\partial x} = \Lambda(x) \cdot \left(\frac{s^{-1}(x) - \theta}{x - s^{-1}(x)} \right)$, it is easy to see that y is single-peaked in x, with a peak at $x = s(\theta)$. Moreover, as $\Lambda(x)$ is non-increasing, it is easy to verify that $\frac{\partial^2 y}{\partial \theta \partial x} = \Lambda'(x) \leq 0$. From this Spence-Mirrlees condition and the optimality of $x = s(\vartheta) = \overline{x}(\theta)$ for $\vartheta = s^{-1}(\overline{x}(\theta)) < \theta$ it follows that for any $\theta > \overline{x}^{-1}(t)$, $\overline{x}(\theta)$ maximizes $y(\theta, x)$ under the constraint $x \geq \overline{x}(\theta)$. A similar argument proves $x = t$ is constrained optimal for $\theta \in [\theta'', \overline{x}^{-1}(t)]$. But this implies $y(\theta, z(\theta)) = M(\theta)$ for all $\theta \geq \theta^{SB}$, as the lemma requires. A direct calculation verifies that $y(\theta^{SB}, \overline{x}(\overline{\theta})) = 0$ and hence $M(\theta^{SB}) = 0$. From the Spence-Mirrlees condition, we conclude $M(\theta) = 0$ for all $\theta \leq \theta^{SB}$.

[46] Recall from (30) that $\sigma(j) < z(j) = \overline{x}(\theta)$ for all $j \geq \overline{x}^{-1}(t)$. Setting $\tilde{x} = \overline{x}(j)$ shows $\tilde{x} - s^{-1}(\tilde{x}) > 0$.

We now verify that $M(\theta)$ has bounded variation. Using the envelope theorem, $M'(\theta) = 1 + \Lambda\left(z(\theta)\right)$ for $\theta < \overline{x}^{-1}(t)$ and $M'(\theta) = 1 + \Lambda\left(z(\tilde{\theta})\right) \cdot \left(\frac{\overline{x}(\theta)-\theta}{\overline{x}(\theta)-\sigma(\theta)}\right)$ for $\theta \geq \overline{x}^{-1}(t)$.[47] As $\overline{x}$ and z^{-1} are continuous for $\theta \geq \theta^{SB}$ and $\overline{x}(\theta)$ is bounded away from $\sigma(\theta)$ for $\theta \geq \overline{x}^{-1}(t)$, M' is continuous on the intervals $[\underline{\theta}, \overline{x}^{-1}(t))$ and $[\overline{x}^{-1}(t), \overline{\theta}]$. Moreover, $\lim_{\theta \uparrow \overline{x}^{-1}(t)} M'(\theta) = 1 + \Lambda(t)$. Thus M' is bounded and continuous on any bounded subinterval of $[\underline{\theta}, \overline{x}^{-1}(t))$ or $[\overline{x}^{-1}(t), \overline{\theta}]$, which implies M has bounded variation. The conditions of Lemma 14 established, we conclude that $G^{\star}$ solves problem (34)-(38).

For the comparison of θ^{SB} with $\theta^{\star}$, note that any T-NAP is *deterministic* and obeys the incentive constraints (9). Therefore, $V^{SB} < V$ (Proposition 2). Directly evaluating:

$$
\begin{aligned}
V &= \beta \mathbb{E}_F[\theta] + (1-\beta) \cdot \Big(t(F(\overline{x}^{-1}(t))) - F(\theta^{\star}) + \int_{\overline{x}^{-1}(t)}^{\overline{\theta}} \overline{x}(\tilde{\theta}) dF(\tilde{\theta}) \Big) \\
V^{SB} &= \beta \mathbb{E}_F[\theta] + (1-\beta) \cdot \Big(t(F(\overline{x}^{-1}(t)) - F(\theta'')) + \int_{\theta^{SB}}^{\theta''} \overline{x}(\sigma^{-1}(\tilde{\theta})) dF(\tilde{\theta}) + \int_{\overline{x}^{-1}(t)}^{\overline{\theta}} \overline{x}(\tilde{\theta}) dF(\tilde{\theta}) \Big)
\end{aligned}
$$

so that

$$
V - V^{SB} = (1-\beta) \cdot \Big(t(F(\theta^{\star}) - F(\theta^{SB})) - \int_{\theta^{SB}}^{\theta''} \Big(\overline{x}(\sigma^{-1}(\tilde{\theta})) - t \Big) \, dF(\tilde{\theta}) \Big).
$$

Since $\overline{x}(\sigma^{-1}(\theta)) \geq t$ for all $\theta \in [\theta^{SB}, \theta'']$, $V > V^{SB}$ requires $\theta^{\star} < \theta^{SB}$.[48] For $\theta'' < t$, see the last paragraph in Lemma 13's proof. □

Since $\theta^{SB} > \theta^{\star}$ and $\theta'' < t$, a comparison of the tests described by Proposition 1 and Proposition 6 yields the following corollary:

Corollary 4. *The optimal test under ex ante disclosures induces the threshold grade t with a strictly larger probability than does the optimal test under ex post disclosures. Moreover, the regulator's payoff is strictly higher in the former case.*

With ex ante disclosures, the test exploits strictly more 'bunching' of types around the margin. There are three reasons for this: first, with ex ante disclosures even the highest types can be pooled (with some probability $\alpha(\theta) > 0$) into the threshold grade; second, low passing types (below $[\theta^{SB}, \theta'']$) receive a grade in excess of t; and third, strictly more weak types pass $\theta^{\star} < \theta^{SB}$.

A comparison of the grade distributions for higher grades is tricky: while higher types expect to receive a grade z in excess of t less frequently under the ex ante timing, they

[47] For $\theta \geq \overline{x}^{-1}(t)$, note that the envelope theorem for constrained optimization applies.

[48] When $\theta^{SB} = \underline{\theta}$, $V = V^{SB} = \mathbb{E}_F[\theta]$, from which $\theta^{\star} = \underline{\theta}$ follows.

also expect a higher grade in the 'best case' (θ, rather than $\overline{x}(\theta)$). In particular, the grade distributions are *not* ordered by second order stochastic dominance.

Remark. As the example below illustrates, negative assortative matching of passing grades optimally economizes on the excessively high grades required to satisfy ex post nondisclosure constraints. When the lowest types are pooled with the highest, the distribution of remaining types (yet to be assigned a grade) is improved, which allows the designer to raise the value of intermediate grades, ceteris paribus. This relaxes disclosure constraints for types lower down the distribution, and allows the designer to pass more bad types.

By contrast, this non-monotonicity is not optimal under ex ante disclosure constraints. We showed that stochastic grades—which offer high types a binary lottery between pooling with other passing types into a marginal grade t, or (ii) separating entirely from the low types—are optimal. This allows the designer to both offer high types a targeted reward for nondisclosure, and still pool high types as much as possible with weaker types. These stochastic tests are simply not feasible with ex post disclosure, because some grades violate ex post disclosure constraints. In Section C.1.2, we discuss some frictions that motivate a focus on ex ante disclosure constraints, even if the firm has multiple disclosure opportunities.

Example. Suppose for simplicity there are five types $\theta \in \{1, 2, 3, 4, 5\}$, with prior probabilities f_θ, and a threshold $t \in (3, 4)$. Consider a test which fails type 1 and involves two passing grades, 2 and 3, where types 2 and 4 are assigned grade l, with conditional mean z_l, while 3 and 5 get h, for a conditional mean z_h. Clearly, this test is not negative-assorting.

Now consider a new test, which 'swaps' the grades allocated to 2 and 3: a small mass p_2 of type 2 is reassigned from grade l to h, and a small mass p_3 of type 3 is reassigned in the opposite direction, where p_2 and p_3 are chosen to ensure the conditional mean z_h remains unchanged. This requires $p_2 < p_3$: otherwise, h would now place a higher weight on type 2 than it previously did type 3, so z_h would fall. As a result, the swap reduces the overall probability that investors observe h. By the Law of Iterated Expectations, this means z_l must *increase*—since l is now more common, it must have a more moderate effect on investors' expectations. But raising z_l in this way relaxes type 4's disclosure constraints, allowing an additional mass of type 1's to be pooled into grade l, and thereby escape financial frictions.

We conclude this section with a brief discussion of generalizations of Proposition 6. First, the T-NAP of Lemma 13 continues to solve (28) for a subclass of the v functions from section 2. Denote the *convexification* of v by vex v, and say a value function v is *T-NAP-compatible* if its convexification satisfies:[49]

$$\text{vex } v(z) = \begin{cases} v(z), & \text{for } z < t \\ v(t_-) + \frac{z-t}{\overline{\theta}-t}\Big(v(\overline{\theta}) - v(t_-)\Big), & \text{for } z \geq t. \end{cases}$$

where $v(t_-) = \lim_{z \uparrow t} v(z)$. Proposition 6 generalizes to *T-NAP-compatible* value functions. The key step is to prove that for such v optimal tests are threshold-separable (Appendix A.2.1). The argument is similar to step 2, proof of Lemma 7.[50] Once this is established, the proof of Proposition 6 applies almost unchanged to the search for the optimal *threshold-separable* test. Conversely, we can also show that if v is not T-NAP-compatible then there exists a prior F for which the test is not a T-NAP (and hence, not threshold-separable).[51]

Second, when $\overline{x}$ is non-monotone, the appropriate generalization of Proposition 6 requires a broader notion of a T-NAP test. In this case, the optimal test involves negative assortative matching between types in $[\theta', \theta'']$ and those with the highest values of $\overline{x}$–which need no longer perfectly align with the highest types.

C.1.2 Late Disclosure Options

Our main analysis of pre-emptive disclosure was motivated by two observations: *(i)* it takes time to prepare and publish credible, accurate evidence, and *(ii)* there are legal and reputational costs associated with delaying the disclosure of evidence 'material' to the firm's fortunes. By contrast, in this and the next subsection, we study the implications of the firm being able to *choose* the timing of its disclosures. To do this, we introduce two time periods: $t = 0$ corresponds to some time before the outcome of the test is realized, and $t = 1$ to a time after the test result has been realized. In this subsection we relax *(ii)* by separating the choices of evidence preparation and disclosure. We assume the firm can *privately* choose whether to prepare evidence at $t = 0$.[52] Preparation costs c, but allows the firm to disclose θ. If the firm has prepared evidence at $t = 0$, then it may decide whether to disclose at $t = 0$, wait until $t = 1$, or disclose nothing at all. We refer to this extension of section 2 as the *disclosure option* game.

One might wonder whether the option to wait undermines the regulator's test. Indeed, for an arbitrary test, the option value associated with delay may increase the value of preparing evidence, and therefore lead to disclosures on the equilibrium path. However, under the

[49] vex v is the largest convex function pointwise dominated by v.

[50] The argument must now be applied to signals $z > t$. For a T-NAP compatible v, one can show that the reasoning of step 2 still goes through.

[51] Details on request.

[52] Owing to *(i)*, preparing at $t = 1$ will be too late to avoid the costs of a negative test result.

optimal test of Proposition 1, this issue does not arise:

Proposition 7. *In disclosure option game, the regulator's optimal payoff is V.*

Proof. Under the test in Proposition 1, every type $\theta \in \delta_+^c(t)$ receives a *deterministic* grade, and hence the option to delay does not influence their preparation decision, relative to model of section 2. Next, consider any type $\theta \in \delta_+(t)$. If type θ prepares evidence, it pays cost c. Clearly, waiting until $t = 1$ is weakly preferred to disclosing at $t = 0$. Under the test in Proposition 1, there are two possibilities: *(i)* $s = t$, *or (ii)* $s = \theta$. In the former case, the firm will choose $m = \theta$. In the latter, it is indifferent between $m = \theta$ and $m = \emptyset$. Hence, under the test of Proposition 1 the firm's payoff from waiting to disclose is equal to the payoff from pre-emptive disclosure, and so it is optimal for the firm not to prepare evidence, just as in the main analysis. As the disclosure option game imposes more constraints on the regulator than problem (9), the result follows.[53] □

C.1.3 Degrading evidence

We now partially relax assumption *(i)* (full relaxation was analyzed in section C.1.1). In particular, we introduce the possibility that the market's perception of evidence quality might depend on the timing of disclosure. It is natural to think that the informativeness of evidence depends not only on the content of the disclosure itself but also on *when* it has been disclosed. On one hand, producing accurate information requires some time. Yet, with too much time, one might become concerned that the firm could find ways to present itself in an overly favorable manner.[54] Our aim is simply to illustrate the consequences of this friction in a stylized way; we do not aim to provide a full-fledged analysis of all the issues that arise in a game involving dynamic evidence.

We refer to the following as the *evidence degradation* game. As above, there are two time periods $t \in \{0, 1\}$. In contrast to the disclosure option game, we now allow the firm more freedom: it only pays the cost c at the time of making a disclosure. However, the evidence structure depends on t. As before, type θ's message space at time $t = 0$ is $M(\theta) = \{\emptyset, \theta\}$. But at time 1, the firm's message space $M'(\theta)$ may differ: with probability $p \in (0, 1)$, $M'(\theta) = M(\theta)$ and with the remaining probability $M'(\theta) = \Theta \cup \{\emptyset\}$. Investors never observe the firm's message space. This (admittedly stylized) set-up captures the idea that, as time

[53]The "public speech, private silence" principle extends to the disclosure option game in the obvious way.

[54]The SEC requires public companies to file audited annual accounts information (Form 10-K) within 60 days of the end of the fiscal year; similarly less detailed, unaudited quarterly filings (Form 10-Q) must be made within a 40-day deadline. This suggests regulators may see some trade-off between preparation time and informativeness in financial reporting. Relatedly, it is well understood that banks have incentives and at least some opportunities to 'window dress' their financial statements (Allen and Saunders (1992)).

progresses, a firm of type $\theta' < \theta$ may be able to find ways to 'window dress' its accounting and present itself as having type θ.

With this change, the regulator must now satisfy two sets of incentive constraints:

$$\begin{aligned} V = \max_G \quad & E_G\left[v\left(z\right)\right] && \text{subject to} \\ & E_G\left[u\left(z,\theta\right)\middle|\,\theta\right] && \geq \bar{u}\left(\theta\right), \qquad \forall\theta, \\ & u\left(z,\theta\right) && \geq \max_{m\in M'(\theta)} u(\underline{\theta}^1(m,z),\theta) - c(m), \\ & && \qquad \forall z \text{ and } \theta \in \text{supp } F(\theta \mid z), \end{aligned} \tag{43}$$

where $\underline{\theta}^1(m,z) = \min\{\theta : m \in M'(\theta),\ \text{and } \theta \in \text{supp } F(\theta \mid z)\}$ is the most skeptical belief at $t = 1$ consistent with the test result z.[55] When evidence degrades over time, market skepticism can enforce pre-emptive disclosures as the relevant outside option for firms:

Proposition 8. *In the evidence degradation game, the regulator's optimal payoff is V.*

Proof. Consider the optimal test $G^\star$ described in Proposition 1 and an equilibrium in which the sender never discloses. We support this equilibrium with the off-path beliefs $\Pr[\tilde{\theta} = \theta \mid m = \theta, t = 0] = \theta$, $\Pr[\tilde{\theta} = \theta \mid m = \theta, z, t = 1] = \underline{\theta}^1(m,z)$. In particular, since the message space is not observable at $t = 1$, notice that for each type and grade z, the skeptical belief at $t = 1$ obviously satisfies $\underline{\theta}^1(m,z) \leq z$ for all $m \in M(\theta)$. Hence, the disclosure constraints at $t = 1$ are non-binding for the test described in Proposition 1. As problem (43) is clearly more constrained than (9), the result follows. □

We briefly note that the equilibrium described in the proof survives standard signaling refinements, such as D1 and the Intuitive Criterion. The argument is similar to those developed in section C.2 and hence omitted. We remark that the simple equilibrium construction we consider does feature a discontinuity at $p = 1$ (in that case, note that problem (43) is equivalent to problem (28), studied in section C.1). We view this discontinuity as an artifact of the highly stylized set-up, and consider developing a richer model of dynamic disclosures as an interesting avenue for future research.

C.2 Equilibrium Selection with General Message Spaces

The proof of the general version of the *public speech, private silence* principle (section A.3) uses an equilibrium supported by 'skeptical beliefs'. Since Milgrom (1981), skeptical beliefs have appeared commonly in the analysis of disclosure games; for instance, they have recently

[55]Notice that the second set of incentive constraints are taken with respect to $M'(\theta)$ only. Since $M(\theta) \subseteq M'(\theta)$, these constraints already subsume the requirement that the firm not disclose at $t = 1$ in the event that the message space is $(M(\theta))_{\theta\in\Theta}$.

been used by Hagenbach et al. (2014) to characterize conditions under which fully revealing equilibria exist in general disclosure games. Nonetheless, a natural question arises regarding whether these beliefs are compatible with signaling refinements such as the Intuitive Criterion or D1 (Cho and Kreps (1987b); Banks and Sobel (1987)), which impose structure on off-path beliefs. We emphasize that this issue does not arise for the "truth-or-nothing" message spaces considered our the main analysis: the only off-path messages are those of the form $m = \theta$, which are pinned down in *any* sequential equilibrium to $\Pr[\tilde{\theta} = \theta \mid m = \theta] = 1$. Given this lack of flexibility to choose off-path beliefs, the *public speech, private silence* equilibrium corresponding to the test in Proposition 1 (hereafter, the *regulator-preferred equilibrium*) trivially survives D1 and the Intuitive Criterion; the same argument applies for the case of ex post disclosures (see section C.1.1).

We address the question of when the regulator-preferred equilibrium survives the Intuitive Criterion and D1 more generally in two scenarios. First, we consider a setting in which the regulator's test can be conditioned on the firm's disclosures (Proposition 9). Notice that this does not allow the regulator to do better than in Proposition 1: on the equilibrium path, the *public speech, private silence* principle still applies. Off-path, any test the regulator imposes following a disclosure cannot do the firm more harm than the skeptical belief. Hence, it is still optimal to adopt the test of Proposition 1 after non-disclosure. However, when off-path beliefs are no longer skeptical, the regulator may benefit from changing the test in response to disclosures. We show that under a natural assumption on $\overline{x}$ the regulator-preferred equilibrium survives standard signaling refinements.

For completeness, we also consider the case where the regulator's test cannot react to disclosures (Proposition 10). Under tighter conditions on $\overline{x}$, the regulator-preferred equilibrium continues to survive.

Proposition 9. *Suppose the disclosure-equivalent grade $\overline{x}$ is non-decreasing in θ. Then the regulator-preferred equilibrium supported by skeptical beliefs survives the Intuitive Criterion and the D1 refinement if the regulator can respond to disclosure with a less informative test.*

Proof. Let $\underline{\theta}(m) = \arg\min_{\theta' \in M^{-1}(m)} u(\theta, \theta')$ be the minimal type who can send message m. Let $\underline{m}(\theta) = \arg\max_{m \in M(\theta)} u(\underline{\theta}(m), \theta) - c(m)$ be type θ's skeptical outside option, and notice that we can write the disclosure-equivalent grade for any type $\theta \geq \overline{x}^{-1}(t)$ as $\overline{x}(\theta) = \frac{1}{\phi}[\overline{u}(\theta) - (1-\phi)\theta] = \underline{\theta}(\underline{m}(\theta)) - \frac{c(\underline{m}(\theta))}{\phi}$. We consider a regulatory test which responds to $m \neq \emptyset$ by revealing the type if $\theta \in (\underline{\theta}(m), \theta^\star)$ and otherwise stays quiet. With this test, we show that the regulator-preferred equilibrium survives D1.[56] To do this, it is sufficient to prove that for any off-path message m the set $D_{\underline{\theta}(m)}(m)$ of beliefs μ_m supported on $M^{-1}(m)$ for

[56]The argument for the Intuitive Criterion is similar, and hence omitted.

which playing m improves $\underline{\theta}(m)$'s payoff (relative to equilibrium) contains the corresponding set $D_\theta(m)$ for any other type $\theta \in M^{-1}(m)$. We proceed in cases:

Case 1: $\theta \geq \theta^\star$: In this case, a necessary condition for type $\tilde{\theta} \in \{\underline{\theta}(m), \theta\}$ to benefit from disclosure is $\mathbb{E}_{\mu_m}[\theta \mid \theta \notin (\underline{\theta}(m), \theta^\star)] = E_{\mu_m} \geq t$, where we note that E_{μ_m} depends only on m.[57] The payoff to choosing such a message is $(\phi E_{\mu_m} + (1-\phi)\tilde{\theta}) - c(m)$. If $\underline{\theta}(m) \geq \theta^\star$, this payoff exceeds type $\tilde{\theta}$'s equilibrium payoff $\overline{u}(\tilde{\theta})$ if and only if

$$\min\{t, \overline{x}(\tilde{\theta})\} \leq \mathbb{E}_{\mu_m}[\theta] - \frac{1}{\phi}c(m).$$

If $\overline{x}$ is non-decreasing, then $\min\{t, \overline{x}(\underline{\theta}(m))\} \leq \min\{t, \overline{x}(\theta)\}$ for all $\theta \in M^{-1}(m)$. This shows $D_\theta(m) \subset D_{\underline{\theta}(m)}(m)$. If $\underline{\theta}(m) < \theta^\star$, then sending m is beneficial for $\underline{\theta}(m)$ iff $\underline{\theta}(m)\lambda(\underline{\theta}(m)) \leq \mathbb{E}_{\mu_m}[\theta] - \frac{1}{\phi}c(m)$, which only strengthens the argument.

Case 2: $\theta < \theta^\star$: This case is dealt with by the regulator's response to $m \neq \emptyset$. If $\theta \neq \underline{\theta}(m)$, the regulator's test fully reveals θ. This can never be a profitable deviation, since disclosures are costly. Hence, $D_\theta(m) \subset D_{\underline{\theta}(m)}(m)$ is trivially true. $\square$

We note that, when $\overline{x}$ is strictly increasing, the above argument shows something stronger: among all equilibria in which the firm chooses $m = \emptyset$, *only* those involving skeptical beliefs for the binding types $\theta \geq \overline{x}^{-1}(t)$ survive D1. In this sense, they are somewhat natural.

As the proof shows, failing types are always more willing to disclose than passing types. In fact, this would continue to be true even if the regulator responded to disclosures with complete silence. Fixing m, a low type $\theta_1 < \theta_2$ gains less than a high type $\theta_2 \geq t$ from the regulator-preferred equilibrium and hence has a greater incentive to try to undermine it with disclosures. On the other hand, when comparing the incentives of two failing types $\theta_1 < \theta_2 < t$ to disclose, things are not so straightforward. When λ is increasing and $\phi < 1$, θ_2 has a higher marginal benefit of increasing z. The regulator's response deals with this issue by making such types' outcomes insensitive to their disclosures.[58]

Further, we note that Proposition 9 is not a knife-edge consequence of our assumption that disclosure costs $c(m)$ are type-independent.[59] For instance, if $\overline{x}$ is strictly increasing, then the proof shows that $D_\theta(m) \subset D_{\underline{\theta}(m)}(m)$ is strict. Hence the conclusion still applies for range of type-dependent cost functions $c(\theta, m)$; on the other hand, we suspect it may not apply for others.

[57]The conditioning arises from the regulator's response, but it plays no role in the argument for case (i) – it would have equally applied if the regulator chose to stay silent after disclosures.

[58]Indeed, if λ were constant for $\theta < \theta^\star$ or $\phi = 1$, Proposition 9 would go through even if the regulator responded to disclosures with silence.

[59]This is common in disclosure models; type-dependent costs are more akin to signaling models.

One might ask whether skeptical beliefs are still robust when the regulator cannot respond to the test. Under stricter primitive conditions, the answer is still 'yes'.

Proposition 10. *Suppose the regulator's test cannot respond to disclosures, but* $\underline{\theta}(m) \geq \theta^\star$ *for all* $m \in \bigcup_{\theta \geq t} M(\theta)$*. If* $\overline{x}$ *increases sufficiently quickly on* $\{\overline{x} \geq t\}$ *so that* $\alpha(\theta)$ *defined in (14) is non-decreasing, then the regulator-preferred equilibrium supported by some (not necessarily skeptical) beliefs survives the Intuitive Criterion and the D1 refinement.*[60]

Proof. As before, we focus on the D1 refinement. The proof is similar to that of Proposition 9, but relies more on the structure of the regulator-preferred test. Recall that the regulator-preferred test induces a lottery over $\{t, \theta\}$ for each $\theta \geq \theta^\star$, where $\alpha(\theta)$ is the probability of grade θ; it perfectly reveals types below (Proposition 1). As those types cannot influence beliefs about themselves via disclosures, they can be ignored from here onwards (similar to case 2, proof of Proposition 9, $D_\theta(m) = \emptyset$ for such types).

We support the regulator-preferred equilibrium with quasi-skeptical beliefs: for $m \in \bigcup_{\theta \geq t} M(\theta)$, let $\theta = \underline{\theta}(m) \geq \theta^\star$ be the corresponding skeptical belief. For $m \in \bigcup_{\theta \in [\theta^\star, t)} M(\theta)$, let $\theta = \theta^q(m)$, where $\theta^q(m) = \min\{\theta \in M^{-1}(m) \cap \{\theta \geq \theta^\star\}\}$. Note that for any $m \in \bigcup_{\theta \in [\theta^\star, t)} M(\theta)$, we have $\theta^q(m) \leq t$ because there must exist $\theta < t$ for whom $m \in M(\theta)$. As beliefs are skeptical for $m \in \bigcup_{\theta \geq t} M(\theta)$, it follows that these quasi-skeptical beliefs do support the regulator-preferred equilibrium.

For any $m \in \bigcup_{\theta \geq t} M(\theta)$ and $\theta \in M^{-1}(m)$, we must show $D_\theta(m) \subset D_{\underline{\theta}(m)}(m)$. Similarly, for $m \in \bigcup_{[\theta^\star, t)} M(\theta)$, we must show $D_\theta(m) \subset D_{\theta^q(m)}(m)$ for all $\theta \in M^{-1}(m)$. To that end, consider some $m \in \bigcup_{\theta \geq \theta^\star} M(\theta)$ and two types $\theta_1 < \theta_2$ with $\theta_1, \theta_2 \in M^{-1}(m)$ and $\theta_1 \geq \theta^\star$; as before, let μ_m be some belief compatible with m. By an argument similar to case 1, proof of Proposition 9, type $\hat{\theta} \in \{\theta_1, \theta_2\}$ benefits from choosing m iff:

$$t + \frac{1}{1 - \alpha(\theta)} c(m) \leq \mathbb{E}_{\mu_m}[\theta \mid s = t].$$

If $\alpha(\theta)$ increases, then the left hand side of this expression is increasing in θ. In other words, $D_{\theta_2}(m) \subset D_{\theta_1}(m)$. Choosing $\theta_1 \in \{\underline{\theta}(m), \theta^q(m)\}$ for the respective cases shows that $D_\theta(m) \subset D_{\underline{\theta}(m)}(m)$ for $m \in \bigcup_{\theta \geq t} M(\theta)$, and $D_\theta(m) \subset D_{\theta^q(m)}(m)$ for $m \in \bigcup_{[\theta^\star, t)} M(\theta)$. □

Undefeated equilibrium

Finally, we conclude with a few words about the *undefeated equilibrium* (Mailath et al. (1993)) refinement. For a disclosure game with general message space M, it is difficult (even

[60]If $\overline{x}$ is increasing and differentiable, then it is easy to verify from (14) that α is increasing if and only if the elasticity of $\overline{x} - t$ with respect to $\theta - t$ exceeds 1 for $\theta \geq \overline{x}^{-1}(t)$. This is satisfied for instance if $\overline{x} = \beta\theta - k$ for $\theta \in [t, \overline{\theta}]$ and some $k > 0, \beta \geq 1$ satisfying $\overline{x}(\overline{\theta}) \leq \overline{\theta}$.

without additional outside information) to fully characterize the set of all Perfect Bayesian Equilibria.[61] For this reason, a full analysis of undefeated equilibrium—which requires a comparison of payoffs across equilibria—is intractable. Yet, as discussed in section 2.1, the regulator's optimal test is trivially undefeated in the baseline "all-or-nothing" message space. As we illustrate below, this is not the only message space for which the regulator-preferred equilibrium is undefeated.

Suppose $M(\theta)$ is strictly increasing (in the set inclusion order), and the cost of any $m \neq \emptyset$ is some constant $c > 0$. To see that the regulator-preferred equilibrium is undefeated, it is sufficient to argue that there is no other equilibrium in which a set Θ^d of types choose some message $m \neq \emptyset$, with every $\theta \in \Theta^d$ enjoying a weakly higher payoff than in the regulator-preferred equilibrium. Notice that playing m obviously yields a lower payoff for the highest type θ' in Θ^d than if he were simply revealed at a cost c. Since that the latter is exactly $\bar{u}(\theta')$ in this setting, it is clearly no greater than his payoff in the regulator-preferred equilibrium. A similar argument can be applied when costs vary by message—so long as they do not vary too much.[62]

C.3 Regulatory Previews

Here we prove an extended version of Lemma 1, in which we also allow the regulator to provide a preview to the firm about the likely outcome of the test. Specifically, we enrich the regulator's design problem from section 2 by allowing her to provide an additional private signal $\rho : \Theta \times S \to \Delta\mathbb{R}$ to the firm.[63] Note that, in principle, ρ is flexible enough to allow the preview to depend both on the firm's type and on the grade; equivalently, one can think of the regulator being able to tailor information about grades to different types of firm. In particular, notice that this extension allows us to capture variations in the timing of test results as special cases of preview strategy: for instance, pre-disclosure tests correspond to the special case $\rho(\theta, s) = s$, while post-disclosure tests (i.e. the timing considered in the main body) can be captured with the constant function $\rho \equiv 0$. With a slight abuse, we use ρ to denote both the preview strategy and its realizations, where convenient and clear from context. We prove:

Lemma 15. ***[Public Speech, Private Silence]*** *For any valid test $\{G, \rho\}$, there exists an alternative test $\left\{\hat{G}, \hat{\rho}\right\}$ such that i) the regulator obtains the same expected utility, ii) the*

[61] Indeed, much of the literature focuses only on when such models admit a fully revealing equilibrium; see for instance Hagenbach et al. (2014).

[62] Beyond this and the baseline case, we simply do not know for which spaces M and functions c the regulator-preferred equilibrium is undefeated; we suspect there may be configurations for which it is not.

[63] More formally, we require ρ to be a measurable function.

firm makes no verifiable disclosures in equilibrium under $\left\{\hat{G}, \hat{\rho}\right\}$*, and iii) the regulator does not send an informative private signal to the firm with* $\hat{\rho} \equiv 0$*.*

Proof. Fixing ρ, an almost identical argument to the proof of Lemma 1 shows that (G, ρ, μ) can be replaced with a pair $(\hat{G}, \rho, \hat{\mu})$ such that *(i)* and *(ii)* hold.[64] The only difference is that the incentive constraints become:

$$\int u(z, \theta) d\hat{G}(z \mid \theta, r) \geq \overline{u}(\theta) \tag{44}$$

for all θ and $r \in \text{supp}\rho(\cdot \mid \theta)$. Integrating across r, (44) implies $\int u(z, \theta) d\hat{G}(\mid \theta) \geq \overline{u}(\theta)$ for all θ. That is, $\hat{\mu}$ is also an equilibrium strategy under $(\hat{G}, \hat{\rho})$, where $\hat{\rho} \equiv 0$. □

C.4 On the strict optimality of private silence

We argue there does not exist a test G and corresponding equilibrium disclosures μ in which *(i)* a positive measure of types $\theta \geq \theta^\star$ disclose with positive probability and *(ii)* the regulator's equilibrium payoff is V.

We first argue that types $\theta \in \delta_+(t) := \{\theta : \bar{u}(\theta) > u(t, \theta)\}$–i.e., the set where $\overline{x}(\theta) > t$– cannot disclose with positive probability in equilibrium. Otherwise, the payoff of such a type would be

$$\mathbb{E}[u(\mathbb{E}[\tilde{\theta} \mid s, \tilde{m}], \theta) \mid \theta] - \mathbb{E}[c(\tilde{m}) \mid \theta] \geq \overline{u}(\theta)$$

where the inequality follows since type θ has the option to disclose $\underline{m}$, which is worth at least $\overline{u}(\theta)$ in equilibrium. If θ discloses with positive probability, then $\mathbb{E}[c(\tilde{m}) \mid \theta] > 0$ and so

$$\mathbb{E}[u(\mathbb{E}[\tilde{\theta} \mid s, \tilde{m}], \theta) \mid \theta] - \mathbb{E}[c(\tilde{m}) \mid \theta] > \overline{u}(\theta).$$

But if this inequality holds for a positive measure of types, then the implied distribution $G^\star$ of $z = \mathbb{E}[\theta \mid s, m]$ violates Proposition 10.3 and hence cannot be optimal. Therefore if the regulator's payoff is V, then types $\theta \in \delta_+(t)$ must not disclose with positive probability.

We now show this extends to all types $\theta \geq \theta^\star$. By Proposition 10.3 any remaining type $\theta \in [\theta^\star, \overline{\theta}] \cap \delta_+^c(t)$ must receive grade t with probability 1. Since disclosures are costly and all types in $[\theta^\star, \overline{\theta}] \cap \delta_+^c(t)$ must earn the same grade, all types in this set must be disclosing (otherwise, non-disclosers expect a grade of t without paying the disclosure cost). Let m' be

[64] In particular, since $\hat{G}$ is a mean-preserving spread of G (see Lemma 1), it is feasible to provide the firm with the same information structure ρ. Indeed, if ρ is a Blackwell-garbling of G, then it can be constructed as a Blackwell-garbling of $\hat{G}$ too.

this disclosure.[65] Moreover, Proposition 10.3 implies there must be some grade s' such that $t = \mathbb{E}[\theta \mid m', s']$ and (m', s') is induced in equilibrium iff $\theta \in [\theta^\star, \overline{\theta}] \cap \delta_+^c(t)$. By the previous argument, any type $\theta \in \delta_+(t)$ must not send m' with positive probability. We argue this implies $\mathbb{E}[\theta \mid m', s'] < t$—a contradiction. To see this, recall that $\theta^\star$ by definition satisfies

$$tq + (1-q)\mathbb{E}[\overline{x}(\theta) \mid \overline{x}(\theta) > t] = \mathbb{E}[\max\{t, \overline{x}(\theta)\} \mid \theta \geq \theta^\star] = \mathbb{E}[\theta \mid \theta \geq \theta^\star],$$

where $q = \Pr[\theta \in \delta_+^c(t) \mid \theta \geq \theta^\star]$. Further recall that $\overline{x}(\theta) < \theta$ for all θ, and hence

$$tq + (1-q)\mathbb{E}[\overline{x}(\theta) \mid \overline{x}(\theta) > t] > \mathbb{E}[\theta \mid \theta \geq \theta^\star]. \tag{45}$$

But, by the Law of Iterated Expectations, we have

$$\mathbb{E}[\theta \mid \theta \in [\theta^\star, \overline{\theta}] \cap \delta_+^c(t)]q + (1-q)\mathbb{E}[\overline{x}(\theta) \mid \overline{x}(\theta) > t] = \mathbb{E}[\theta \mid \theta \geq \theta^\star], \tag{46}$$

where we have used the fact that $\overline{x}(\theta) > t \iff \theta \in \delta_+(t)$. A direct comparison of (45) and (46) shows that $t > \mathbb{E}[\theta \mid \theta \in [\theta^\star, \overline{\theta}] \cap \delta_+^c(t), s'] = \mathbb{E}[\theta \mid m', s']$.

D Online Appendix: Analysis of Bank Runs

We rigorously derive Example 1 in the paper from a model of bank runs, which follows Diamond and Dybvig (1983) and Morris and Shin (2000).

Model Environment: The model has three periods $t = 0, 1, 2$. A bank interacts with a continuum of investors. Each investor is endowed with \$1 at $t = 0$. At date 1, a measure λ of investors finds out that they are either patient, obtaining utility $u(c_1) = \log(c_1)$ from consumption at date 1, and a unit measure of other depositors find out that they are patient, obtaining utility $u(c_1 + c_2) = \log(c_1 + c_2)$ from consumption at either $t = 1$ or $t = 2$. As in Diamond and Dybvig (1983), it is not possible to verify or contract on the realization of investors' preferences.

The bank has access to a riskless storage technology that transfers dollars across time periods, and also to a long-term investment technology that yields $Re^{-\kappa\ell}$ at $t = 2$ for every dollar invested at $t = 0$, where $R > 1$ denotes the return on investment if it is held to maturity, and where ℓ is the (endogenously determined) fraction of resources that are withdrawn at $t = 1$. In this formulation, κ is a parameter measuring the illiquidity of the

[65] It is without loss of generality to consider a single disclosure message: if there are multiple disclosure messages, the same argument applies after taking expectations across such messages.

bank's technology.

We assume that the baseline return on investment satisfies

$$\log R = \theta + \eta,$$

where $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ is the unknown quality of the bank, defined as in the paper as a random variable with cumulative distribution $F(\theta)$, and where $\eta \sim N\left(0, \frac{1}{\alpha}\right)$ is an idiosyncratic component of banks' returns, which is statistically independent of θ.

Optimal Contracts: Diamond and Dybvig (1983) and Morris and Shin (2000) show that an efficient allocation can be implemented by the following contract between the bank and its investors: Each investor deposits \$1 in the bank at $t = 0$. At $t = 1$, each investor has the right to withdraw her claim and obtain a fixed repayment of \$1, or to leave her claim in the bank and obtain a share of the final returns to investment, which are shared equally between all investors that do not withdraw at $t = 2$. The bank invests \$$\lambda$ in storage and the remaining \$1 in the long-term technology. Hence, any withdrawals from impatient investors can be met without (inefficiently) withdrawing resources from the long-term investment.

Inside and Outside Information: At $t = 1$, investors observe a disclosure from the bank m and the regulatory signal s (outside information). These variables contain information about θ but are independent of η, and are otherwise defined as in the paper. We let $z = E[\theta|s, m]$ denote investors' posterior expectation of θ after these disclosures.

In addition, each investor obtains an outside signal $x_i = \eta + \epsilon_i$ of the idiosyncratic component of bank returns, where $\epsilon_i \sim N\left(0, \frac{1}{\beta}\right)$ is independent of η and θ. This signal is introduced only to generate small deviations from common knowledge among investors, as in the literature on global games in banking (e.g., Goldstein and Pauzner, 2005). We will focus below on the "noiseless" limit where the precision of agent's signals β tends to infinity. Therefore, we do not treat the auxiliary outside signal x as part of the information design problem. The regulator's problem is again to design the test s so as to maximize investors' joint utility.

The Bank Run Game: All impatient investors have a dominant strategy to withdraw at date 1, and the bank repays them in cash. If a measure ℓ of patient consumers also withdraw, a patient investor who withdraws at $t = 1$ obtains \$1, for utility $u(1) = 0$. A patient investor who waits until $t = 2$ obtains $Re^{-\kappa\ell}$, for utility $u\left(Re^{-\kappa\ell}\right) = \theta + \eta - \kappa\ell$. Therefore, patient investor i prefers to wait if

$$E[\theta + \eta|s, m, x_i] \geq \kappa\ell,$$

Equivalently, using the independence of θ and η, we can write

$$z + \rho_i \geq \kappa \ell,$$

where

$$\rho_i = E\left[\eta | x_i\right] = \frac{\beta x_i}{\alpha + \beta}.$$

We look for cutoff equilibria such that investor i withdraws if and only if $z + \rho_i \leq \rho^\star$. In such an equilibrium, suppose that i is the an investor who is just indifferent between withdrawing and waiting, i.e., that i receives a signal x_i that ensures the posterior belief $\rho_i = \rho^\star$.

Investor i expects that the mass of investors withdrawing is

$$\begin{aligned} E\left[\ell | x_i\right] &= Pr\left[\rho_j \leq \rho_i | x_i\right] \\ &= Pr\left[\frac{\beta}{\alpha + \beta} x_j \leq \rho_i | x_i\right] \\ &= Pr\left[x_j \leq \rho_i \left(1 + \frac{\alpha}{\beta}\right) | x_i\right] \end{aligned}$$

Notice also that

$$x_j | x_i \sim N\left(\rho_i, \frac{\alpha + 2\beta}{\beta\left(\alpha + \beta\right)}\right)$$

implying that

$$\begin{aligned} E\left[\ell | x_i\right] &= \Phi\left(\sqrt{\frac{\beta\left(\alpha + \beta\right)}{\alpha + 2\beta}} \frac{\alpha}{\beta} \rho_i\right) \\ &= \Phi\left(\sqrt{\gamma} \rho_i\right), \end{aligned}$$

where we have defined

$$\begin{aligned} \gamma &= \frac{\alpha^2}{\beta^2} \frac{\beta\left(\alpha + \beta\right)}{\alpha + 2\beta} \\ &= \frac{\alpha^2\left(\alpha + \beta\right)}{\beta\left(\alpha + 2\beta\right)} \end{aligned}$$

To complete the verification of equilibrium, we require that investor i is indeed indifferent between withdrawing and waiting. This requirement translates to $z + \rho_i = \kappa E\left[\ell | x_i\right]$ or, equivalently,

$$\kappa \Phi\left(\sqrt{\gamma} \rho^\star\right) = z + \rho^\star. \tag{47}$$

We now focus on the noiseless limit of this model, which is defined as follows: Consider a sequence of parameter values (α, β), describing the precisions of the idiosyncratic component η of bank returns and the noise in investors' private signals. Assume that along this sequence we have $\alpha, \beta \to \infty$, so that in the limit, η collapses to a point mass at zero and investors have perfect information about η. Assume also that $\gamma \to 0$, which requires that β diverges to infinity at a faster rate than α.

In the limit as $\gamma \to 0$, (47) implies that $\rho^{\star} \to \frac{\kappa}{2} - z$. Moreover, as $\alpha, \beta \to \infty$, investors have near-perfect information, so that the distribution of each ρ_i collapses to a point mass at zero. Hence, there is a run in the noiseless limit whenever

$$\begin{aligned} \rho_i &\leq \frac{\kappa}{2} - z \\ \Leftrightarrow z &\leq \frac{\kappa}{2} \equiv t \end{aligned}$$

Hence, we have derived the specification of bank runs in Example 1 as the unique equilibrium in the noiseless limit of the bank run game.

Regulator's Objective Function: The regulator wants to maximize the joint utility of depositors. Here, we can ignore the utility of impatient investors, which is equal to $u(c_1) = u(1) = 0$ regardless of whether there is a bank run. The utility of a representative patient investor if there is a run is also $u(1) = 0$. If there is no run, then it is given by $u(R) = \theta + \eta \to \theta$ in the noiseless limit. Hence, the joint utility of all investors is proportional to

$$\theta \cdot Pr\,[\text{bank run}] = \theta \cdot (1 - \lambda(z)),$$

where $\lambda(z) = 1\{z < t\}$. Hence, we have derived the objective function in Equation (2) in the paper, with $\lambda(z)$ defined as in Example 1.