Time Inconsistency in Stress Test Design

Markus Parlasca *

December 8, 2021

Abstract

We show that central banks face a time inconsistency problem when publishing bank stress test results. Before a stress test, they want to *appear* tough as the threat of letting banks fail the stress test incentivizes prudent behaviour. After the stress test, they want to *act* soft by releasing only partial information in order to reassure financial markets about bank health. We characterize an institutional design solution to this commitment problem: a social planner sets the framework within which the central bank communicates. We find that a hurdle rate framework, where all banks are judged to pass or fail relative to a common threshold, is optimal in many settings as it generates intermediate levels of both incentives and reassurance. With a hurdle rate framework, stress tests become an informational contagion channel, as changes in the health of one bank affect beliefs about the health of other banks. Thus, informational contagion can be a feature of a socially optimal institutional design in the presence of a time inconsistency problem.

Keywords: *Bank Stress Tests, Strategic Communication, Contagion, Central Bank Design, Financial Stability*

JEL Classification: D83, E58, G21

*WU Vienna and Vienna Graduate School of Finance, markus.parlasca@wu.ac.at
I thank Franklin Allen, John Cochrane, Vincent Crawford, Stephen Hansen, Ole Jann, Emir Kamenica, Christian Laux, Mikko Leppämäki, Gyöngyi Loranth, Margaret Meyer, Alan Morrison, Daniel Quigley, Chris Roth, John Vickers, Sebastian Vogel, Ben Waltmann, Josef Zechner, and various seminar and conference audiences for helpful comments and suggestions. Remaining errors are mine. I gratefully acknowledge financial support from the Economic and Social Research Council and Merton College, Oxford.

1 Introduction

"*Had we been fully open and fully transparent about what was going on during the financial crisis, it would, let me tell you, have been a lot, lot worse. That would have been shouting 'fire' in the theatre.*"

— Andy Haldane, Chief Economist, Bank of England[1]

"*The disclosure of stress test results [...] promotes market discipline.*"

— Ben Bernanke, Chairman of the Board of Governors of the Federal Reserve System[2]

"*The stress test has been a catalyst for pressure to raise capital.*"

— Andrea Enria, Chair of the ECB Supervisory Board[3]

Following the recent financial crisis, the frequent publication of bank stress test results has become the centrepiece of central bank communication about bank stability. While central banks view the communication of bank stress test results as serving two purposes – incentivize banks to take prudent actions and reassure financial markets about bank stability – the academic literature has largely focused on the latter (Bouvard et al., 2015; Goldstein and Leitner, 2018; Inostroza and Pavan, 2018).

This paper studies stress tests as serving both purposes. We find that a trade-off between these two purposes exists and that they create a time inconsistency problem which results in suboptimal combinations of incentives and reassurance. The trade-off arises because incentives for banks to improve their health are created by the central bank's threat to reveal bank weakness, which to financial markets is the opposite of a reassuring message. Additionally, these two roles create a time-inconsistency problem: Before the stress test, when banks decide whether to take prudent actions, the central bank wants to *appear* tough, i.e. create a threat of revealing bank weakness. After the test, when banks cannot alter their health anymore, the central bank wants to *act* soft, i.e. not reveal bank weakness. Rational agents anticipate this. Banks know that the threat is not credible and therefore do not take prudent actions.

We propose a central bank design solution to this time inconsistency problem. This approach reflects existing institutional arrangements. In practice, via the Dodd-Frank Act, Congress asks the Fed to run stress tests annually and disclose their results, but leaves it to the Fed to choose the appropriate adverse scenario in a given year.[4] In our model, a social planner (Congress)

[1] Quoted in *The Times*, 6 October 2017. Andy Haldane was the Bank of England's Executive Director for Financial Stability during the financial crisis.

[2] Speech at the "Maintaining financial stability: holding a tiger by the tail" financial markets conference, sponsored by the Federal Reserve Bank of Atlanta, Stone Mountain, Georgia, 8 April 2013.

[3] Quoted in *Financial Times*, 15 July 2011. Andrea Enria was Chairperson of the European Banking Authority from 2011 to 2018.

[4] For example, in the 2021 stress test, the Fed included strong stress on commercial real estate with prices having dropped by 40% after three years. The Fed could have chosen more or less stress on commercial real estate prices without violating the Dodd-Frank Act.

chooses a design, and the central bank (Fed) runs stress tests in the constraints given by the design.[5] While the central bank design is chosen for the long term and is not adjusted to changing economic conditions, the stress severity is adjusted for every test. Thus, while the choice of stress severity is subject to a time inconsistency problem, central bank design choices offer a limited form of commitment.

Our definition of central bank design has parallels in the monetary economics literature. A central bank design consists of a communication framework and the central bank's (CB's) objective function (henceforth: mandate). A framework is either full disclosure, zero disclosure, or a hurdle rate framework which means that all banks are judged to pass or fail relative to a common level of stress.[6] Thus, while the planner cannot commit to a state contingent communication rule, the design affects the verifiable disclosure game the CB plays and thus determines the strength of incentives and reassurance. This approach echoes Rogoff's (1985) work on time inconsistency in monetary policy where the central banker's preferences affect the interest rate setting game the central banker plays and thus determines the size of the inflation bias.

We find that a hurdle rate framework is optimal for a large set of parameters as it provides an intermediate combination of both incentives and reassurance. A full disclosure framework generates strong incentives but no reassurance. The reverse is true for a zero disclosure framework.

We show that, in a hurdle rate framework, *strategic delegation* provides additional benefits. This means the planner maximises ex-ante welfare by setting the CB's mandate to differ from ex-post welfare maximisation. The optimal mandate can differ in either direction. If taking prudent actions has high costs, the planner optimally specifies the CB's mandate to be "tougher" than welfare maximisation, so that the CB is less concerned about supporting weak banks than society. This makes it credible that the CB will reveal bank weakness by letting weak banks fail the stress test and thus generates strong incentives. If the costs of prudent actions are low, the optimal mandate is "softer" than welfare maximisation, so that the CB is very concerned about weak banks. This achieves high reassurance while still generating sufficient incentives to induce the prudent action.

Stress tests become an informational contagion channel when a hurdle rate framework is used. This means that changes in the health of one bank affect beliefs about the health of other banks. Contagion arises because the CB optimally responds to a deterioration in one bank's health by adjusting the severity of stress. In some cases, it is optimal for the CB to lower the severity of stress. All banks pass, but as they pass a weaker test, beliefs about all banks are lower. More surprisingly, there are also cases where a deterioration in one bank's health leads to higher beliefs about other banks. This is the case when the CB increases the severity of stress. As a result, the weak bank fails the test and beliefs about its health deteriorate. However, beliefs about other banks improve as they pass a tougher test.

[5]Similar provisions exist in other jurisdictions. In the European Union, the EBA Act mandates the EBA to run stress tests and in the UK, the Bank of England sets out a medium term framework within which it runs stress tests annually.

[6]A full disclosure framework is currently used by the Fed whereas a hurdle rate framework is used by the Bank of England.

Informational contagion can thus be a feature of an optimal design when a time-inconsistency problem exists. Contagion need not reduce welfare. The mechanism that turns stress tests into an informational contagion channel - judging all banks against the same severity of stress - is the same mechanism that makes a hurdle rate framework optimal for many parameters. For the CB, using one severity of stress is a constraint. Optimal behaviour given the constraint implies financial contagion. From the planner's perspective, this constraint is valuable, as it results in some but less than full revelation of bank health and thus achieves intermediate levels of both incentives and reassurance. Without the constraint, there would be no contagion but also no resolution to the time inconsistency problem.

Our examination of stress tests has broader implications for economic theory. We contribute to the strategic communication literature by characterizing the equilibrium in a verifiable disclosure communication game with a cross-message constraint. In these games, the sender communicates about multiple banks, truthfully releases pass or fail results for each of them, but is constrained by having to judge all banks relative to a common level of stress, rather than using bank specific stress levels. The sender therefore faces a trade-off: While higher levels of stress result in more favourable messages for those banks that pass, they have the cost of more banks failing the test which is a very unfavourable message.

Our paper also contributes to the literature on incentives generate by communication, which has commonly been framed in the context of teachers grading exams (Boleslavsky and Cotton, 2015; Dubey and Geanakoplos, 2010). Our model can be reinterpreted in that setting: A class teacher (central bank) wants to incentivize students (banks) to study for an exam (take a prudent action) by threatening to let bad students fail the test. However, once the exam has been taken, the teacher wants to give out good marks to help students get good jobs. A head teacher (planner) tries to solve this time inconsistency problem by specifying the school's grading policy (central bank design).[7]

The remainder of this paper is structured as follows: Section 2 reviews the related literature. Section 3 outlines the model and discusses key assumptions. Section 4 focuses on reassurance. First, it takes an ex-post perspective and characterises how much reassurance a central bank achieves for given health outcomes in equilibrium in a given design. For a hurdle rate framework, this involves solving the verifiable disclosure game with a cross-message constraint. Then, it takes the planner's ex-ante perspective and characterises how much reassurance different designs achieve in expectation, i.e. across possible health outcomes. Section 5 compares designs in terms of incentives. Section 6 combines incentive and reassurance concerns to solve for the optimal institutional design. Section 7 shows that a hurdle rate framework implies that stress tests become an informational contagion channel. Section 8 concludes.

[7] This paper differs from Boleslavsky and Cotton (2015) in not assuming commitment of the teacher (central bank) and in focusing on effort of a third party (students, banks) rather than the sender. It differs from Dubey and Geanakoplos (2010) in focusing on effort prior to and thus in anticipation of communication rather than afterwards.

2 Literature Review

This paper is related to three different strands of the literature. Our idea of studying institutional design solutions to a time inconsistency problem is related to the monetary economics and particularly the "central bank design" literature (Reis, 2013). Rather than focusing on a monetary policy maker who is tempted to boost output by creating surprise inflation[8] we focus on a bank regulator tempted to shield weak banks from market pressure by not releasing information. Our modelling framework for this communication game builds on and contributes to the theoretical literature on strategic communication. Our topic of bank stress tests is rooted in the central bank communication literature.

Central Bank design and strategic delegation

CB design solutions to the time inconsistency problem in monetary policy range from strategic delegation to performance contracts. Rogoff (1985) argues that delegating policy to a "conservative central banker" who places more weight than society on inflation stabilisation relative to output stabilisation increases welfare. Walsh (1995) uses insights from principal agent theory to argue that the time inconsistency problem can be reduced by a performance contract which rewards the central banker for low inflation.

Both institutional design solutions have the common feature that the objective function of the policy maker which maximises ex-ante welfare differs from ex-post welfare maximisation.[9] This is an application of strategic delegation as outlined by Vickers (1985) who shows that in a strategic context a principal achieves the highest pay-off by delegating to an agent who maximises an objective function other than the principal's pay-off.[10]

This paper contributes to the CB design literature in three ways. We characterise a new problem by showing that central banks face a time inconsistency when publishing bank stress test results. To solve this problem, we take CB design and strategic delegation ideas from monetary policy to the new setting of financial stability and from a focus on policy makers' actions to a focus on communication. Our institutional design solution shows that benefits of strategic delegation exist also in this new setting.

Strategic communication theory

Our communication game builds on models of verifiable disclosure (Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981; Guttman et al., 2014; DeMarzo et al., 2019). That means the

[8]As agents with rational expectations anticipate this temptation, the result is "inflation bias": an equilibrium with inflation above target and no output gains (Lucas, 1972, 1975; Kydland and Prescott, 1977; Barro and Gordon, 1983a,b).

[9]In Rogoff (1985), the objective function comes from intrinsic types. In Walsh (1995), it results from a performance contract.

[10]Vickers (1985) studies Cournot competition. A firm's profit increases when decisions are delegated to a manager who places weight on market share, not just on profits, as this commits the firm to aggressive behaviour which in turn leads rivals to cut output. Strategic delegation allows the principal to become a Stackelberg leader.

sender (central bank) is privately informed and can only make true statements.[11] Readers more familiar with the macroeconomic literature may equally view our model as "communication with discretion", i.e. the CB is not able to commit to a communication rule. The opposite case of "communication with commitment" corresponds to Bayesian persuasion.

We extend verifiable disclosure models to capture communication about multiple items (banks). In the single item literature, receivers always learn bank health perfectly (Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981) unless communication is costly (Jovanovic, 1982; Verrecchia, 1983) or there is uncertainty on whether the sender is informed (Dye, 1985; Jung and Kwon, 1988).[12] We show that a new departure from the stark result of full revelation can arise when communication is about multiple items. In our stress test setting, the sender judges multiple banks to pass or fail relative to the same stress scenarios. While for senders using as many (or more) stress scenarios as there are banks, there is full revelation in equilibrium, this is no longer the case when there are fewer stress scenarios than banks. We formalise the resulting novel verifiable disclosure game where the choice of the stress scenario affects the information revealed about all banks, i.e. there is a constraint across messages, characterise the equilibria, and show that all involve partial disclosure.

Central bank communication

The growing importance of central bank communication as a policy tool is reflected in a vast literature spanning both monetary policy and financial stability.[13] The disclosure of stress test results is a prominent topic in the latter field. Closely related to our paper are models without commitment, i.e. where the central bank decides on disclosure after becoming privately informed. Models with commitment, i.e. where the CB commits to a communication strategy before becoming privately informed, are a useful comparison.[14]

Our key contributions to the hitherto small literature on the case without commitment are that we consider the incentive motive of stress test disclosure in addition to the reassurance motive and that we add the possibility of partial disclosure (hurdle rate framework) to the previously studied extremes of full and zero disclosure. Bouvard et al. (2015) consider a regulator who communicates only because of the reassurance motive and faces a binary choice between full and zero disclosure. Intermediate levels of informativeness such as a hurdle rate framework are not considered. Moreover, Bouvard et al.'s (2015) regulator makes the choice between full and zero disclosure once he has private information on bank health. In our model, the planner chooses a design (full, zero, hurdle rate) without having private information, but the regulator's choice in a design (e.g. severity of stress) is made based on private information. In Shapiro and

[11] Alternative classes of communication games are models of cheap talk (Crawford and Sobel, 1982) where the sender is privately informed but can lie, or Bayesian persuasion (Kamenica and Gentzkow, 2011) where an uninformed sender commits to a messaging rule before becoming privately informed. A recent survey of the literature based on commitment is Bergemann and Morris (2019).

[12] Recent work which builds on these verifiable disclosure models includes Guttman et al. (2014) who study the timing of disclosure in a multi-period model and DeMarzo et al. (2019) where the sender is initially uninformed, decides what test to undertake to become informed, and decides on the disclosure of results after observing them.

[13] Surveys on monetary policy communication include Geraats (2002) and Blinder et al. (2008).

[14] A broad survey of arguments for and against the disclosure of stress test results is provided by Goldstein and Sapra (2013).

Zeng (2018) a regulator tests one bank repeatedly.

Models with commitment are applications of Bayesian persuasion and include Goldstein and Leitner (2018) where disclosure destroys risk sharing among banks (Hirshleifer (1971) effect) and Inostroza and Pavan (2018) who add an explicit model of how receivers coordinate by introducing additional heterogeneous private information. Faria-e-castro et al. (2016) study how a government's fiscal position and disclosure policy interact. Parlatore (2015) views stress tests as generating private signals in a Diamond and Dybvig (1983) model of bank runs. Orlov et al. (2017) endogenise the response of asset prices. Williams (2017) considers an extension in which banks can adjust to the communication rule. He finds that banks reduce their liquidity buffers when stress tests are conducted since pass results make runs less likely and thus act as substitute for liquidity buffers. While in Williams (2017) banks respond to a CB which is committed to a disclosure rule, in our paper banks respond to the anticipated strategic communication of a CB which cannot commit. Huang (2021) studies persuasion on a network.[15]

The conflict our central bank faces between creating incentives and providing reassurance is new in the stress test and central bank communication literature, but also arises in the context of bail-outs. Bail-outs distort banks' incentives, but can prevent runs. More formally, in Shapiro and Skeie (2015), a regulator's decision whether to bail out a bank signals the regulator's type to depositors, who may run, and to banks, who may adjust their future risk-taking behaviour. Keister's (2016) regulator faces a trade-off and a commitment problem. The anticipation of the regulator's action (bail-out) affects the incentives banks face and the depositor's decision whether to run. Keister then compares a policy of committing to not bailing out banks, i.e. a ban, to discretion and finds that a ban may reduce welfare.

Our results on the interaction of disclosure and contagion complement those by Alvarez and Barlevy (2021). While they focus on a setting with connected banks and study whether banks' private disclosure decisions are socially optimal, we study a setting with independent banks and show that disclosure by a regulator can be a source of contagion.

3 Model

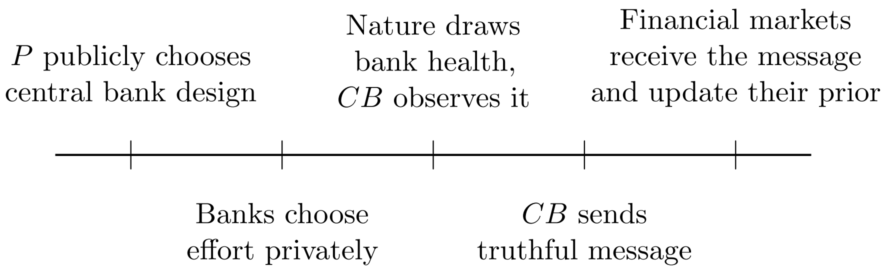
Our model of stress tests distinguishes between a social planner's (Congress) choice of a central bank design (Should stress test results be disclosed? How granular should disclosure be?) and the central bank's (Fed) choice when running a stress test (Should commercial real estate prices be subjected to a 20% or a 30% drop in the stress scenario?). While the central bank design is chosen for the long term and is not adjusted to changing economic conditions, the stress severity is adjusted for every test. Thus, while the choice of stress severity is subject to a time inconsistency problem, central bank design choices offer a limited form of commitment.

Formally, a welfare maximising social planner P chooses a central bank design. There are two

[15]While this literature review focused on the disclosure of stress test results, other aspects of stress tests have also been studied. For example, Parlatore and Philippon (2020) model stress tests as generating information for an uninformed regulator, not as transmitting information to financial markets. Leitner and Williams (2020) study the disclosure of stress test models rather than of results. Quigley and Walther (2020) study the effect on banks' incentives to disclose additional information.

ex-ante identical banks which both observe the design and decide whether to take a prudent action which makes high bank health realisations more likely.[16] Bank health is realised ($H_i \in [0,1]$) and observed only by the central bank (CB). Then, the CB runs a stress test in the constraints specified by the design and publishes results truthfully. Financial markets receive these messages and update their belief on bank health (μ_i) to $\mu_i = \mathbb{E}[H_i \mid \mathcal{I}]$ where $\mathcal{I}$ is all publicly available information. The sequence of moves is summarised in Figure 1 below:

Figure 1: Order of Moves



3.1 Description of the Model

Planner At the first stage, planner P chooses a central bank design $\mathcal{D} = \{\mathcal{C}, \mathcal{M}\}$ which consists of a communication framework $\mathcal{C}$ and a mandate for the central bank $\mathcal{M}$. We consider frameworks that have been used in practice: full disclosure (US), zero disclosure (EU in 2009), a hurdle rate framework (UK, EU in 2011). A hurdle rate framework specifies that the CB must judge all banks to pass or fail relative to a common stress scenario, but does not specify its severity.[17] A mandate ($\mathcal{M}$) is a mapping from a pair of beliefs about bank health $(\mu_1, \mu_2) \in [0,1]^2$ to $\mathbb{R}^1$, i.e. $\mathcal{M} : [0,1]^2 \to \mathbb{R}^1$. We denote the value which a mandate assigns to (μ_1, μ_2) as $V(\mu_1, \mu_2)$, i.e. $\mathcal{M} : (\mu_1, \mu_2) \mapsto V(\mu_1, \mu_2)$. The planner maximises welfare ($\mathcal{W}$). Formally, $\mathcal{W}$ is a mapping from a pair of beliefs $(\mu_1, \mu_2) \in [0,1]^2$ to $\mathbb{R}^1$, i.e. $\mathcal{W} : [0,1]^2 \to \mathbb{R}^1$. We denote the value which welfare assigns to (μ_1, μ_2) as $W(\mu_1, \mu_2)$, i.e. $\mathcal{W} : (\mu_1, \mu_2) \mapsto W(\mu_1, \mu_2)$.[18] $\mathcal{D}$ becomes public knowledge.

Bank moral hazard Banks simultaneously choose an unobservable effort level $e_i \in \{\underline{e}, \bar{e}\}$ where $\underline{e} < \bar{e}$. High effort makes higher health realisations more likely in a monotone likelihood ratio sense (MLRP),[19] but is privately costly for the bank ($C(\bar{e}) = c > 0$) while low effort costs $C(\underline{e}) = 0$. Thus, high effort is efficient.[20] The bank's effort can reflect improved risk management, diligence in evaluating loan applications (Aghion et al., 1999), project choice where c is the

[16] A model with two banks is the simplest possible model which allows us to study stress tests as revealing information on multiple banks simultaneously.

[17] In models with one dimensional bank health, the following two ways of thinking about stress tests are equivalent: A bank passes a stress test of severity $s = 0.5$ if stressed health $\tilde{H}_i = H_i - s$ satisfies $\tilde{H}_i > 0$. A bank passes a stress test if its health is sufficiently high $H_i > s$, i.e. clears the hurdle set by s.

[18] The reduced form approach of specifying pay-offs in terms of beliefs is standard in the literature (Guttman et al., 2014; DeMarzo et al., 2019).

[19] The MLRP assumption, introduced by Milgrom (1981), is common in the moral hazard literature (Holmstrom, 1982; Lambert, 1983), in the communication literature (Ottaviani and Sorensen, 2006; DeMarzo et al., 2019), and in finance applications (Dewatripont and Tirole, 1993).

[20] This formalization of bank moral hazard is from Dewatripont and Tirole (1993, 1994). Aghion et al. (1999) develop a similar approach.

bank's private benefit of the bad project (Dewatripont and Tirole, 1993), or choosing a funding mix which is privately costly but socially desirable (e.g. due to externalities from bank failure). The bank maximises expected profit $\mathbb{E}[\pi] = \mathbb{E}[\mu_i] - C(e_i)$. A bank's strategy is thus: $\mathcal{B}_i : \mathcal{D} \rightarrow e_i$.

Nature Bank health realisations H_i are drawn by nature from the distributions implied by e_i and observed only by the central bank, not by financial markets. Bank health is independent across banks.

Central Bank The central bank observes bank health $H_i \, \forall \, i$ and, within the framework $\mathcal{C}$, conducts a stress test to maximise its mandate $\mathcal{M}$. In ZDF (FDF) financial markets learn nothing (bank health perfectly) for any mandate $\mathcal{M}$. In HRF, the central bank chooses a level of stress s strategically to maximise its mandate $\mathcal{M}$ and must report truthfully whether each bank passed or failed relative to the common stress s, which is also reported. The resulting message is a triplet $\{s, o_1, o_2\}$ where $s \in [0, 1]$ and $o_i \in \{p, f\}$ is the outcome for bank B_i. For example $\{0.2, p, f\}$ corresponds to "*Stress was* 0.2 *and* B_1 *passed while* B_2 *failed.*". This does not only reveal the number of pass and fail marks but also reveals which bank passed or failed.

In the hurdle rate communication game, where we treat $\mathcal{D}$ and e_1, e_2 as exogenously fixed, a central bank's messaging rule S is a mapping from any health state pair $(H_1, H_2) \in [0, 1]^2$ to a message $\{s, o_1, o_2\} \in [0, 1] \times \{p, f\}^2$ subject to the truth-telling constraint that $o_i = p$ if and only if $H_i \geq s$, otherwise $o_i = f$, i.e. $S : (H_1, H_2) \rightarrow \{s, o_1, o_2\}$. The CB cannot lie, but it could obfuscate by sending a universally true message such as $\{0, p, p\}$. In the overall institutional design game, the central bank's strategy $\mathcal{S}$ is a profile of such messaging rules, i.e. $\mathcal{S} : \mathcal{D} \rightarrow S$.

Financial markets Financial markets do not observe bank health H_i directly but observe the CB's message $\{s, o_1, o_2\}$ and update their belief accordingly to $\mu_i = \mathbb{E}[H_i \mid \mathcal{I}]$.[21] Financial market participants are interested in bank health because they trade bank shares and credit default swaps or because they assess the risk of interbank loans. In the pure communication game, we say that financial markets follow an interpretation rule $R : \{s, o_1, o_2\} \rightarrow (\mu_1, \mu_2)$. In the overall institutional design game, financial market's strategy $\mathcal{R}$ specifies an R for every design $\mathcal{D}$, i.e. $\mathcal{R} : \mathcal{D} \rightarrow R$. Based on the beliefs (μ_1, μ_2), pay-offs are realised.

3.2 Equilibrium Concept

We focus on perfect Bayesian equilibria (PBE) as solution concept. This means that we restrict our attention to those weak perfect Bayesian equilibria of the overall central bank design game which are weak perfect Bayesian equilibria in every subgame. We first define weak PBE in an important subgame, the hurdle rate communication game, and then define PBE in the overall central bank design game.

Definition 1 *An equilibrium of the hurdle rate communication game is a collection* $\{S, R\}$ *that satisfies:*

[21] Summarizing beliefs by the expected value is in line with the communication literature (Guttman et al., 2014; DeMarzo et al., 2019) and the stress test literature (Goldstein and Leitner, 2018). It can be microfounded by financial market participants being risk neutral and trading an asset in a competitive market, or by participants solving a quadratic loss problem.

(i) Given R; S maximises $V(\mu_1, \mu_2)$ *at every health state pair* (H_1, H_2).

(ii) $\mu_i = \mathbb{E}[H_i \mid \{s, o_1, o_2\}, S] \; \forall \; i$

Condition (ii) states that R uses all available information $\{s, o_1, o_2\}$ and interprets it based on the equilibrium messaging strategy S.

Definition 2 *An equilibrium of the overall central bank design game is a collection of strategy profiles* $\{\mathcal{D}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{S}, \mathcal{R}\}$ *that satisfy*

(i) Given $\mathcal{B}_1, \mathcal{B}_2, \mathcal{S}, \mathcal{R}$*;* $\mathcal{D}$ *maximises* $\mathbb{E}[W(\mu_1, \mu_2)]$.

(ii) For every $\mathcal{D}$ *and given* $\mathcal{B}_{-i}$*,* $\mathcal{S}$*,* $\mathcal{R}$*;* $\mathcal{B}_i$ *maximises* $\mathbb{E}[\pi_i]$.

(iii) For every $\mathcal{D}$ *and given* $\mathcal{B}_1, \mathcal{B}_2, \mathcal{R}$*;* $\mathcal{S}$ *maximises* $V(\mu_1, \mu_2)$ *at every health state pair* (H_1, H_2).

(iv) $\mu_i = \mathbb{E}[H_i \mid \{s, o_1, o_2\}, \mathcal{D}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{S}] \; \forall \; i$

Condition (iv) states that $\mathcal{R}$ uses all available information $\{s, o_1, o_2\}$ and interprets it based on all player's equilibrium strategies.

3.3 Assumptions

Welfare We assume that welfare is increasing in beliefs about bank health ($\frac{\partial W}{\partial \mu_i} > 0 \forall i$) and is particularly concerned with the weak bank as the risks of belief driven phenomena such as runs, higher funding costs, or problems to roll over short term debt are particularly acute for weak banks. We model this by assuming that $\frac{\partial W}{\partial \mu_i} \geq \frac{\partial W}{\partial \mu_j} \; \forall \; \mu_i \leq \mu_j$. Therefore, key feature of welfare is that the expected welfare maximising social planner is risk averse. In our model, it is convenient to use $W(\mu_1, \mu_2) = \lambda \; \min(\mu_1, \mu_2) + (1 - \lambda) \; \frac{\mu_1 + \mu_2}{2}$ where $\lambda \in [0, 1)$ indexes the concern for the weak bank. Results are qualitatively unchanged and in some cases exactly identical when we instead assume that $W(\mu_1, \mu_2) = u(\mu_1) + u(\mu_2)$ where $u'(\mu_i) > 0$ and $u''(\mu_i) < 0$ (see Appendix D.1 and D.2)[22] or when we adopt a more explicit model of bank runs as in Goldstein and Leitner (2018) (see Appendix D.3 and D.4).[23] The common feature of these approaches is risk aversion. Thus, for a given bank health distribution, the planner prefers zero disclosure to full disclosure (the Hirshleifer (1971) effect, i.e. there is full risk sharing among banks; the reassurance motive), but for a given level of disclosure, he prefers better health distributions (incentive motive).[24]

Communication framework The planner chooses one framework from ZDF, FDF, HRF. We consider this choice set for three reasons. First, all three frameworks have been used in practice

[22]This functional form is close to Hirshleifer (1971) who shows that disclosure can destroy risk sharing among agents in a non-bank setting where agents have concave utility functions.

[23]Our results do not rely on continuity or differentiability of $W(\mu_1, \mu_2)$. Appendix D.3 and D.4 show that our results are qualitatively unchanged when we define welfare as in Goldstein and Leitner (2018). Their welfare function has a discontinuity to capture the cost of bank runs, but satisfies the assumption that welfare is increasing in beliefs when restated as $W(\mu_1', \mu_2) - W(\mu_1, \mu_2) > 0 \; \forall \; \mu_1' > \mu_1$.

[24]While health (H_i) does not enter welfare directly, it enters indirectly via beliefs (μ_i). An equilibrium property of Bayesian beliefs is that $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$ (Bayes plausibility). Thus, equilibria with higher $\mathbb{E}[H_i]$ also have higher $\mathbb{E}[\mu_i]$ and thus higher welfare.

and thus are feasible. Second, in introducing HRF we extend the existing literature which mainly compares ZDF and FDF (e.g. Bouvard et al. (2015)). Third, these frameworks can be microfounded. When the planner can choose the number of stress scenarios and the central bank chooses their severity, then zero scenarios results in the same information being transmitted in equilibrium as ZDF, one scenario as HRF, and two or more scenarios as FDF. Thus, framework choice can be reinterpreted as choosing the number of stress scenarios.[25]

Mandate The mandate is endogenously chosen by the planner and does not need to be identical to ex-post welfare maximisation. While we characterize the equilibrium of the communication game played by a CB with $\frac{\partial V}{\partial \mu_i} > 0 \, \forall \, i$ in general, for our central bank design results it is convenient to assume that $V(\mu_1, \mu_2) = \omega \; \min(\mu_1, \mu_2) + (1 - \omega) \; \frac{\mu_1 + \mu_2}{2}$ where $\omega \in [0, 1)$ as this results in closed form expressions. For $\omega > 0$ the CB values increases in beliefs more strongly when they refer to a weak bank. Thus, we refer to the CB as being "tougher" than welfare if $\omega < \lambda$. Results are qualitatively unchanged and in some cases even identical when $V(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ where $v'(\mu_i) > 0$ and $v''(\mu_i) \leq 0$ (see Appendix D.1) or when $V(\mu_1, \mu_2)$ includes a more explicit model of bank runs as in Goldstein and Leitner (2018) (see Appendix D.3). The key common feature of these approaches is that the mandate maximising CB prefers higher beliefs.

MLRP assumption In the context of our model, it is useful to note that the MLRP assumption is equivalent to assuming that for every subinterval $[a, b]$ with $0 \leq a \leq b \leq 1$ the conditional distribution can be ranked according to first-order stochastic dominance (Shaked and Shanthikumar, 2007) and therefore implies that:

$$\mathbb{E}_G[H \mid a \leq H \leq b] \leq \mathbb{E}_F[H \mid a \leq H \leq b] \quad \forall \quad 0 \leq a \leq b \leq 1 \tag{1}$$

where G (F) is the CDF of bank health when effort is low (high).[26] Thus, when financial markets learn that $H_i \in [a, b]$, the resulting belief μ_i is higher when financial markets conjecture that the bank exerted high effort. To get a meaningful incentive problem, we restrict our attention to $c \leq \mathbb{E}_F[H] - \mathbb{E}_G[H]$.[27]

Central bank: no commitment, information We assume that the Central Bank chooses the level of stress strategically after observing bank health and reports results truthfully. The assumption that the CB cannot commit is in line with CBs referring to their choice of stress scenarios as judgements which indicates discretion,[28] with frequent debates on how meaningful a test is,[29] and with the extreme experience of the 2011 European stress test.[30] Our assumption that the CB knows bank health when deciding on the message is motivated by the CB's access

[25] A formal treatment of the game where the planner chooses the number of stress scenarios, together with proofs, is provided in Appendix A. For the remainder of the paper, we adopt the language of framework choice.

[26] An example is $G(H) = H$, $F(H) = H^a$ where $a \geq 1$ ensures an MLRP ordering.

[27] If $c > \mathbb{E}_F[H] - \mathbb{E}_G[H]$, then no disclosure strategy, not even FDF, can induce high effort. Thus, the planner only considers the reassurance motive and therefore ZDF is the optimal design.

[28] "The FPC and PRC judge the stress scenario to be appropriate [...]." Bank of England, Key elements of the 2019 stress test, 5 March 2019.

[29] "Bank stress tests fail to tackle deflation spectre." Financial Times, 27 October 2014.

[30] "The tests were long ago branded as flawed, so the results [most banks passed] were never going to serve their market-soothing purpose. The exams refused to countenance a sovereign default, even as such an event appears imminent. Investors understood that the tests had been overtaken by events, and that a 'pass' or 'fail' was largely meaningless. Instead, they dumped bank shares [. . .]" Financial Times, EU bank stress tests, 18 July 2011.

to private information in their role as bank supervisor.[31]

4 Reassurance

We show that it can be socially beneficial to set the CB a mandate which is softer than welfare maximisation as this increases reassurance. This benefit of a softer mandate, or equally of delegating to a softer CB, arises because the CB's mandate affects how financial markets interpret a given message. As a softer CB chooses a low stress severity in many cases while tougher CBs opt for it only if all banks are very unhealthy, a stress test with a low severity is interpreted less sceptically when it comes from a softer CB. Hence, softer CBs achieve more reassurance.

To derive this, we formally define reassurance and apply the definition to compare reassurance in ZDF and FDF. Then, we turn to HRF. First, we characterise the equilibrium of the novel verifiable disclosure game played in a HRF. Then, we use the equilibrium to characterise reassurance in a HRF and compare it to reassurance in ZDF, FDF, and in HRF with alternative mandates.

Definition 3 *Central bank design $\mathcal{D}$ is said to provide more reassurance than $\mathcal{D}'$ if, for given bank effort, $\mathbb{E}[W(\mu_1, \mu_2)]$ is higher in $\mathcal{D}$ than in $\mathcal{D}'$.*

By comparing $\mathbb{E}[W(\mu_1, \mu_2)]$ across designs while holding effort constant, this definition focuses purely on reassurance. Incentives are considered in section 5, i.e. the possibility that a design might induce higher effort.

Intuitively, this definition adopts the planner's view of reassurance which differs slightly from the CB's view. Recall that previously we introduced the reassurance motive as a desire to increase the beliefs financial markets hold about bank health, potentially with added concern for the weaker bank. For the mandate maximising CB, whose messages affect μ_i, this is captured by $\frac{\partial V}{\partial \mu_i} > 0 \, \forall \, i$. However, the expected welfare maximising planner's design choice affects only the dispersion of beliefs, but not the expected belief since, when effort is held fixed, by Bayes plausibility $\mathbb{E}[\mu_i] = \mathbb{E}[H_i] \, \forall \, \mathcal{D}$.[32] Thus, for the planner, maximising reassurance corresponds to minimising dispersion of beliefs. An alternative interpretation is that the planner maximises reassurance by allowing risk sharing among banks to function, i.e. the Hirshleifer (1971) effect.

For $W(\mu_1, \mu_2) = \lambda \, \min(\mu_1, \mu_2) + (1-\lambda) \, \frac{\mu_1+\mu_2}{2}$, reassurance provided by a design is captured by $\mathbb{E}[\min(\mu_1, \mu_2)]$. This arises because holding effort fixed ensures that Bayes plausibility holds, i.e. $\mathbb{E}[\mu_i] = \mathbb{E}[H_i] \, \forall \, \mathcal{D}$, and therefore we have that $\mathbb{E}[W(\mu_1, \mu_2)] = \lambda \, \mathbb{E}[\min(\mu_1, \mu_2)] + (1 - \lambda) \, \frac{\mathbb{E}[H_1]+\mathbb{E}[H_2]}{2}$. Thus, comparisons of $\mathbb{E}[W(\mu_1, \mu_2)]$ reduce to comparisons of $\mathbb{E}[\min(\mu_1, \mu_2)]$.[33]

[31] Related papers also assume that the CB has private information on bank health, e.g. Shapiro and Skeie (2015). Moreover, stress tests as a diagnostic tool for CBs to learn about bank health existed already prior to the financial crisis. The novelty of the post-crisis regime is that results are published in a specific frequency unaffected by economic events.

[32] Bayes plausibility refers to the equilibrium property of Bayesian beliefs that $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$. Intuitively, any message which leads to $\mu_i > \mathbb{E}[H_i]$ must mean that another message leads to $\mu_i < \mathbb{E}[H_i]$.

[33] The results are qualitatively unchanged for alternative functional forms and alternative measures of dispersion. See Appendix D.2 and D.4.

Proposition 1 *Central bank designs $\mathcal{D}$ can be ranked by how much reassurance they provide:*

(i) ZDF provides more reassurance than HRF which provides more than FDF.

(ii) Within HRF, reassurance is weakly increasing in the weight the mandate puts on the weak bank.

Proof: See Appendix C.1.

ZDF provides more reassurance than FDF. In ZDF, the CB releases no information and thus creates no dispersion of beliefs ($\mu_i = \mathbb{E}[H_i]$).[34] This corresponds to high reassurance. In FDF, the CB fully reveals bank health ($\mu_i = H_i$) which is a deterioration in beliefs if $H_i < \mathbb{E}[H_i]$.[35] This corresponds to low reassurance.

The characterization of reassurance in a HRF builds on the equilibrium of the hurdle rate communication game, which is played if the planner chooses a HRF. The next subsection characterises this equilibrium.

4.1 Hurdle Rate: Equilibrium Communication

We show that an equilibrium exists in the hurdle rate communication game and prove that it is unique when we restrict our attention to equilibria that satisfy a notion of monotonicity. In this equilibrium, the CB chooses a level of stress such that both banks pass if their health is similar and choose a level such that one bank passes and one fails if their health differs strongly. Thus, financial markets interpret stress tests where both banks pass as meaning that both banks' health is at the level of stress or slightly above, but that no bank's health is far above as then the CB would have wanted to choose a higher level of stress and let the weaker bank fail the stress tests. Stress tests where only one bank passes and one fails are interpreted as meaning that the passing bank's health is exactly at the level of stress and the failing bank's health is substantially lower.

In a hurdle rate communication game, the CB faces a trade-off as it judges multiple banks against the same stress scenario. Formally, the CB chooses one strategic variable (s) and truthfully communicates whether a bank's health exceeds s (pass) or not (fail). Thus, higher levels of stress have the benefit of increasing beliefs about those banks which continue to pass but have the cost of resulting in low beliefs about those banks who fail the stress test.

Characterising the equilibrium is challenging because financial markets know that messages are not just true, but also chosen strategically. Thus, the interpretation of a message takes into account that the CB preferred this message to all other feasible messages. Therefore, the interpretation depends on the CB's mandate. The interpretation in turn affects the CB's trade-off. Theorem 1 characterises the equilibrium mathematically and Figure 2 depicts it graphically. Denoting the CDF of H_1 as F and of H_2 as G.

[34]Formally, since in ZDF $\mu_i = \mathbb{E}[H_i]$, reassurance is $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(\mathbb{E}[H_i], \mathbb{E}[H_i])] = \mathbb{E}[H_i]$.

[35]Formally, since in FDF $\mu_i = H_i$, reassurance is $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_i, H_i)] < \mathbb{E}[H_i]$.

Theorem 1 *In the Hurdle Rate Framework communication game there exists a Perfect Bayesian Equilibrium (PBE). This equilibrium is characterised by two indifference frontiers* $x(H_2)$ *and* $y(H_2)$*, visualised in Figure 2a, such that the central bank (sender) plays:*

$$
\begin{array}{llll}
s = & H_1 & \text{iff} \quad H_1 > x(H_2) & \rightarrow \quad \{H_1, p, f\} \\
s = & \min(H_1, H_2) & \text{iff} \quad x(H_2) \geq H_1 \geq y(H_2) & \rightarrow \quad \{\min(H_1, H_2), p, p\} \\
s = & H_2 & \text{iff} \quad y(H_2) > H_1 & \rightarrow \quad \{H_2, f, p\}
\end{array}
$$

where $x(0) = y(0) = 0$*;* $x(H_2) \geq y(H_2)\ \forall\ H_2$*,* $x(H_2) \leq 1\ \forall\ H_2$*; and* $x(H_2)$ *and* $y(H_2)$ *are both continuous and monotonically increasing* ($\frac{dx(H_2)}{dH_2} > 0$, $\frac{dy(H_2)}{dH_2} > 0$)*.*

Financial markets (receivers) form beliefs accordingly, visualised in Figure 2b.

$$
\begin{array}{lll}
\{s, p, f\} & \text{is interpreted as} & \mu_1 = s; \quad \mu_2 = \mathbb{E}_G[H \mid H < x^{-1}(s)] \\
\{s, f, p\} & \text{is interpreted as} & \mu_2 = s; \quad \mu_1 = \mathbb{E}_F[H \mid H < y(s)] \\
\{s, p, p\} & \text{is interpreted as} & \mu_1 = \alpha\ s + (1-\alpha)\ \mathbb{E}_F[H \mid s \leq H \leq x(s)] \\
 & & \mu_2 = (1-\alpha)\ s + \alpha\ \mathbb{E}_G[H \mid s \leq H \leq y^{-1}(s)] \\
\{s, f, f\} & \text{is interpreted as} & \mu_1 = \mu_2 = 0\ \forall\ s > 0
\end{array}
$$

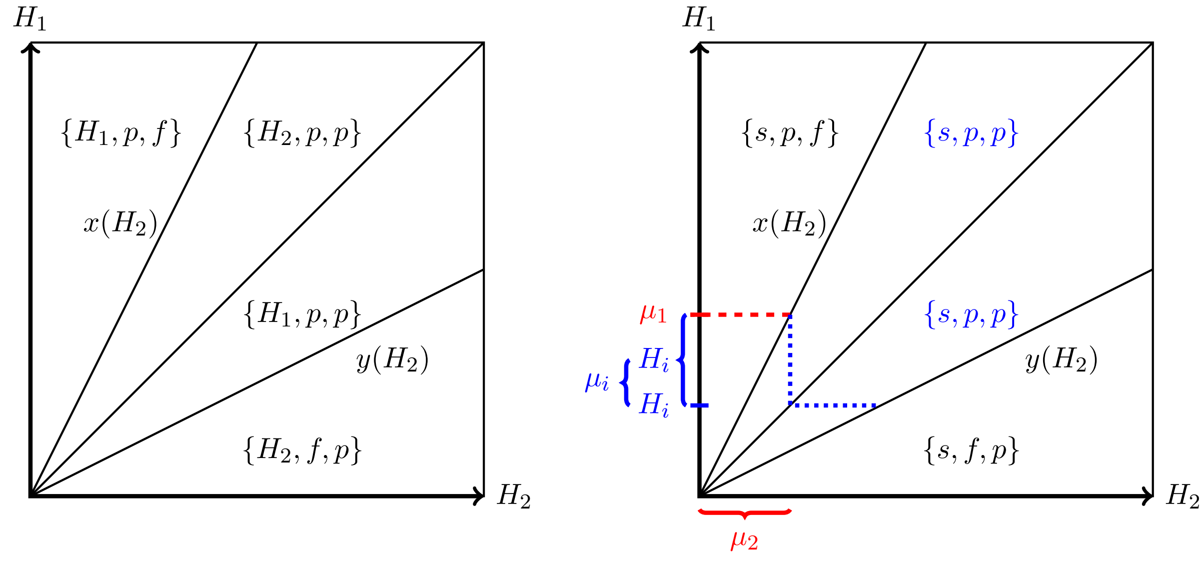
where

$$\alpha = \frac{f(s)\ [G(y^{-1}(s)) - G(s)]}{f(s)\ [G(y^{-1}(s)) - G(s)]\ +\ g(s)\ [F(x(s)) - F(s)]} \tag{2}$$

Proof: See Appendix B.1.

Figure 2: Equilibrium in a Hurdle Rate Framework

(a) Central Bank's Communication Strategy (b) Financial Markets' Beliefs



Note: Figure 2 depicts an example: $H_i \sim U[0,1]\ \forall\ i = 1, 2$; $V(\mu_1, \mu_2) = \frac{\mu_1 + \mu_2}{2}$. Then $x(H_2) = 2H_2$; $y(H_2) = \frac{1}{2}H_2$.

The two indifference frontiers $x(H_2)$ and $y(H_2)$ are defined by the following set of indifference equations where $\mu_i \mid \{s, o_1, o_2\}$ denotes the posterior formed upon receiving $\{s, o_1, o_2\}$.

On $x(H_2)$:

$$V\Big(\mu_1 \mid \{H_1, p, f\},\ \mu_2 \mid \{H_1, p, f\}\Big) = V\Big(\mu_1 \mid \{H_2, p, p\},\ \mu_2 \mid \{H_2, p, p\}\Big) \tag{3}$$

On $y(H_2)$:

$$V\Big(\mu_1 \mid \{H_1, p, p\},\ \mu_2 \mid \{H_1, p, p\}\Big) = V\Big(\mu_1 \mid \{H_2, f, p\},\ \mu_2 \mid \{H_2, f, p\}\Big) \tag{4}$$

When bank health is identically distributed, the indifference frontiers are symmetric, i.e. $y(H_2) = x^{-1}(H_2)$ and $\alpha = 0.5$.

Example: Let $H_i \sim U[0, 1]\ \forall\ i = 1, 2$ and let $V(\mu_1, \mu_2) = \omega \min(\mu_1, \mu_2) + (1 - \omega)\ \frac{\mu_1 + \mu_2}{2}$. Then the equilibrium is characterised by:

$$x(H_2) = \frac{2 - \omega}{1 - 2\omega} H_2 \quad ; \quad y(H_2) = x^{-1}(H_2) \tag{5}$$

This holds for $\omega < 0.5$. For $\omega \geq 0.5$ the CB sends $\{\min(H_1, H_2), p, p\}$ in almost all cases. The only exception arises when $\min(H_1, H_2) = 0$ as then the CB sets $s = \max(H_1, H_2)$ which results in one bank passing and one bank failing the stress test.

To discuss uniqueness of this equilibrium, consider the following notion of monotonicity:

Definition 4 *An equilibrium is said to satisfy monotonicity if beliefs satisfy the following condition:*

$$\mu_i \mid \{s, o_1, o_2\} > \mu_i \mid \{s', o_1, o_2\} \quad \forall\ s > s',\ \forall i = 1, 2,\ \forall o_1,\ \forall o_2 \tag{6}$$

This means that beliefs are increasing in the level of stress for every given stress test outcome, i.e. for passing and for failing the test. This seems close to commentary of stress tests where passing weak tests is seen as a relatively weak signal while passing tough tests is viewed as a strong signal. While this condition imposes a theoretical restriction, it seems plausible in the context of real world bank stress tests.

Theorem 2 *The equilibrium described in Theorem 1 is not a unique weak PBE, but it is the only weak PBE which satisfies monotonicity.*

The proof and a discussion of alternative equilibria which violate monotonicity is provided in Appendix B.2. For the remainder of the paper, we focus on the equilibrium as described in Theorem 1 when referring to a HRF.

To build an understanding of how delegation affects the equilibrium, the following proposition studies a comparative static in the CB's mandate.

Proposition 2 *The more weight the mandate places on the weak bank, the larger the set of states where* $\{\min(H_1, H_2), p, p\}$ *is sent and the smaller the set of states where* $s = \max(H_1, H_2)$ *is set.*

Proof: See Appendix C.2.

The intuition behind proposition 2 builds on the differences in dispersion of beliefs created by setting $s = \min(H_1, H_2)$ or $s = \max(H_1, H_2)$. Recall from Theorem 1 that for any (H_1, H_2) on the $x(H_2)$-indifference frontier, $\{H_2, p, p\}$ results in $\mu_1 = \mu_2$ while $\{H_1, p, f\}$ results in dispersed posteriors $\mu_1' \neq \mu_2'$. Moreover, $\mu_2' < \mu_1 = \mu_2 < \mu_1'$. Thus, while the CB with mandate $\mathcal{M}$ is indifferent between $\{H_2, p, p\}$ and $\{H_1, p, f\}$ on its indifference frontier, a CB with mandate $\mathcal{M}'$ that places more weight on reassurance than $\mathcal{M}$ strictly prefers $\{H_2, p, p\}$ as this achieves a higher $\min(\mu_1, \mu_2)$ than $\{H_1, p, f\}$.

Messages where all banks pass occur in equilibrium even if the mandate places no extra weight on the weak bank.[36] This arises because messages with high levels of stress which both banks pass are only feasible if both H_1 and H_2 are high. Thus, any possible interpretation must result in high $\frac{\mu_1+\mu_2}{2}$ (e.g. $\{0.9, p, p\}$). However, messages with a high level of stress which only one bank passes are feasible also when one bank is extremely weak. E.g. $\{0.95, p, f\}$ is feasible both at $(H_1 = 0.95, H_2 = 0)$ and at $(H_1 = 0.95, H_2 = 0.9)$. Thus, in equilibrium $\{0.95, p, f\}$ cannot be a strong message about H_2 and hence at $(H_1 = 0.95, H_2 = 0.9)$ it is optimal to send $\{0.9, p, p\}$.

Moreover, low levels of stress which all banks pass and which result in financial markets understanding that all banks are weak can arise in equilibrium even if the mandate places no extra weight on the weak bank. We return to this observation in section 8 where we argue that the European stress test of 2011 (most banks passed, financial markets continued to believe that bank health was low) is no evidence of regulatory incompetence or softness, but rather that such an event arises when the stress tester has committed to using a HRF and learns that banks are weak.

Messages where one bank fails occur in equilibrium only if the mandate places no weight, or at most a small weight, on the weak bank. Mandates with medium and large weights result in an equilibrium where the CB almost always sends $\{\min(H_1, H_2), p, p\}$ and sets $s = \max(H_1, H_2)$ iff $\min(H_1, H_2) = 0$. For example, when $H_i \sim U[0, 1]\ \forall\, i = 1, 2$ and $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1 - \omega)\ \frac{\mu_1+\mu_2}{2}$, the CB almost always sends $\{\min(H_1, H_2), p, p\}$ for all $\omega \geq 0.5$.

The comparative statics show that additional concern for the weak bank is not necessary for the equilibrium to take the form depicted in Figure 2 and that large degrees of concern for the weak bank are not necessary to result in an equilibrium where banks fail the stress test only if they are extremely weak.

4.2 Hurdle Rate: Reassurance

This subsection shows that it can be socially beneficial to set the CB a mandate which is softer than welfare maximisation as this increases reassurance. To establish this, we characterise reassurance provided in a HRF and show that reassurance is higher for softer mandates. We also show that reassurance in a HRF lies between that of ZDF and FDF.

[36] This can be seen in the example used to construct Figure 2 where a CB which places no extra weight on the weak bank sends $\{\min(H_1, H_2), p, p\}$ at every second (H_1, H_2).

Characterisations of reassurance in a HRF face a mathematical difficulty stemming from the numerous thresholds in the equilibrium of the communication game, and the conceptual difficulty that changes in the mandate have competing, partially offsetting effects on reassurance. On the one hand, a softer CB sends $\{\min(H_1, H_2), p, p\}$ more often which increases reassurance since $\{\min(H_1, H_2), p, p\}$ creates less dispersion of beliefs than messages resulting from $s = \max(H_1, H_2)$. On the other hand, the interpretation of messages differs across CB mandates. While $\{\min(H_1, H_2), p, p\}$ messages are more reassuring when sent by a softer CB, $\{H_1, p, f\}$ messages are less reassuring when sent by a softer CB.

Intuitively, a softer CB chooses $\{\min(H_1, H_2), p, p\}$ messages in many cases while tougher CBs opt for it only if both banks are very unhealthy. Thus, $\{\min(H_1, H_2), p, p\}$ is interpreted more sceptically if it is sent from a tough CB. However, this also means that messages which one bank fails, e.g. $\{H_1, p, f\}$, result in higher beliefs about the weak bank if the CB is tougher.

To overcome these challenges, note that mathematically reassurance in a HRF is:

$$\mathbb{E}[\min(\mu_1, \mu_2)] = \int_0^1 \int_0^1 \min(\mu_1, \mu_2)\ f(H_1)\ dH_1\ g(H_2)\ dH_2 \tag{7}$$

As beliefs build on the CB's strategy, which is characterised by thresholds, this becomes:

$$\begin{aligned}
\mathbb{E}[\min(\mu_1, \mu_2)] = \int_0^{x^{-1}(1)} \Bigg[&\int_0^{y(H_2)} \mathbb{E}\Big[H \mid H < y(H_2)\Big]\, f(H_1)\, dH_1 \\
&+ \int_{y(H_2)}^{H_2} \frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}\Big[H \mid H_1 < H < x(H_1)\Big]\, f(H_1)\, dH_1 \\
&+ \int_{H_2}^{x(H_2)} \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}\Big[H \mid H_2 < H < x(H_2)\Big]\, f(H_1)\, dH_1 \\
&+ \int_{x(H_2)}^{1} \mathbb{E}\Big[H \mid H < x^{-1}(H_1)\Big]\, f(H_1)\, dH_1 \Bigg]\, g(H_2)\, dH_2 \\
+ \int_{x^{-1}(1)}^{1} \Bigg[&\int_0^{y(H_2)} \mathbb{E}\Big[H \mid H < y(H_2)\Big]\, f(H_1)\, dH_1 \\
&+ \int_{y(H_2)}^{H_2} \frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}\Big[H \mid H_1 < H < 1\Big]\, f(H_1)\, dH_1 \\
&+ \int_{H_2}^{1} \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}\Big[H \mid H_2 < H < 1\Big]\, f(H_1)\, dH_1 \Bigg]\, g(H_2)\, dH_2
\end{aligned} \tag{8}$$

We find that reassurance is weakly increasing in the weight the mandate puts on the weak bank (Proposition 1). This means that when stress tests are conducted by a softer CB, the increase in reassurance stemming from $\{\min(H_1, H_2), p, p\}$ messages becoming more likely and from these messages being interpreted more favourably outweighs the decrease in reassurance stemming from $\{H_1, p, f\}$ being interpreted more sceptically. We also find that reassurance in a HRF lies between ZDF and FDF for all possible mandates.

Delegating stress tests to a CB with a mandate that is softer than welfare maximisation increases reassurance and can thus be socially beneficial. To see the benefit of delegation formally, recall that effort is taken as exogenous. Then, for any planner who places some strictly positive extra weight on the weak bank, welfare maximisation becomes reassurance maximisation. Supposing that ZDF is not available, the planner's optimal choice therefore is to choose a mandate which is so soft that the CB implements the softest possible communication rule, i.e. almost always sends $\{\min(H_1, H_2), p, p\}$.[37]

The benefit of delegation arises because the CB's mandate affects how financial markets interpret messages. For the planner, choosing a soft mandate is effectively a commitment to sending $\{\min(H_1, H_2), p, p\}$ messages often and setting $s = \max(H_1, H_2)$ only in the rare case when one bank is extremely weak. This results in the interpretation of $\{\min(H_1, H_2), p, p\}$ messages being less sceptical and thus increases reassurance.[38] The alternative - setting the CB's mandate to equal welfare maximisation and announcing that $\{\min(H_1, H_2), p, p\}$ will be sent often - would not affect the interpretation of messages as financial markets realise that it is optimal for the CB to deviate from the stated communication rule.

While this section showed that delegation to a softer CB can be beneficial as it increases reassurance, the next section shows that delegation to a tougher CB can be beneficial as it increases incentives.

5 Incentives

We show that it can be socially beneficial to set the CB a mandate which is tougher than welfare maximisation as this increases incentives. This benefit of a tougher mandate, or equally of delegating to a tougher CB, arises because the CB's mandate affects what messages banks expect the CB to send. As banks know that a tougher CB lets banks fail the stress test at more health states, delegating to a tougher CB makes the threat of revealing bank weakness credible at more health states. The credible threat incentivises banks to take prudent actions, which improve bank health. We show that the increase in incentives caused by delegation can result in a shift from an equilibrium with low prudent actions to an equilibrium with high prudent actions.

To derive this, we formally define incentives and show that FDF implies stronger incentives than ZDF. Then, we characterise incentives in a HRF and show that they are higher for tougher mandates and lie between those of ZDF and FDF. Drawing on these results, we characterise equilibria in the game with endogenous bank effort and show that, when incentives are stronger, high effort occurs in equilibrium for a larger set of parameters.

[37] This is not necessarily the same as delegating to the softest possible CB. As argued in the context of proposition 2, there exists a range of mandates which result in almost always sending $\{\min(H_1, H_2), p, p\}$. In our example, $\omega \in [0.5, 1]$. The planner is indifferent between these mandates as they imply the same level of reassurance.

[38] A planner can achieve even more reassurance by choosing ZDF. The result here, which compares different mandates in a HRF, is important when there is a reason to disclsoe some information. In our full model, this reason is the need to create incentives. More generally, CBs may view some degree of communication as necessary for public accountability reasons.

5.1 Defining Incentives

A bank, which maximises $\mathbb{E}[\pi] = \mathbb{E}[\mu_i] - C(e_i)$, will choose high effort if the benefit in terms of expected beliefs $\mathbb{E}[\mu_i]$ more than compensates for the certain cost c. While effort increases $\mathbb{E}[H_i]$ by assumption, the design determines whether effort also increases $\mathbb{E}[\mu_i]$. We thus say that a design provides incentives if $\mathbb{E}[\mu_i]$ is larger for high effort than for low effort.

The key reason why the design determines the presence and strength of incentives is that effort is unobserved. If financial markets can observe effort, they know the distribution of bank health and use it when interpreting the CB's messages. For example, upon learning that $H_i \in [a, b]$, financial markets form the belief that $\mu_i = \mathbb{E}[H \mid a < H < b]$ where the expectation is taken across the bank health distribution implied by effort. Moreover, across all possible health states $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$. The design is thus irrelevant for the strength of incentives.

However, when effort is not observed, as in our model, the design affects the presence and strength of incentives. If financial markets do not observe effort, they do not know the distribution of bank health and thus instead use their conjectured effort level and the implied conjectured bank health distribution when interpreting the CB's messages. Upon learning that $H_i \in [a, b]$, financial markets form the belief that $\mu_i = \mathbb{E}[H \mid a < H < b]$ and calculate the expectation based on the conjectured bank health distribution. We denote conjectured effort as $\varepsilon \in \{\underline{e}, \bar{e}\}$. While in equilibrium conjectured effort is correct ($\varepsilon_i = e_i \forall i$) this need not be the case outside equilibrium. Hence, we do not have that across all possible health states $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$ and thus the design does affect the strength of incentives.

To formalize this, let $\mu_i(\underline{\boldsymbol{e}})$ denote the belief financial markets have when they conjecture that $\varepsilon_1 = \underline{e}$, $\varepsilon_2 = \underline{e}$. Let $\underline{e}$ lead to H_i following CDF $G(H)$, and $\bar{e}$ to CDF $F(H)$. Then, upon learning that $H_i \in [a, b]$, financial markets could have $\mu_i(\underline{e}) = \mathbb{E}_G[H \mid a < H < b] = \int_a^b Hg(H)dH$ or $\mu_i(\bar{\boldsymbol{e}}) = \mathbb{E}_F[H \mid a < H < b] = \int_a^b Hf(H)dH$.

While the interpretation of a message depends purely on conjecture effort ε_i, which banks treat as fixed, the probability of reaching a given health state and thus triggering a given message depends on a bank's effort e_i. This means that while a bank knows that it cannot affect the interpretation resulting from message $H_i \in [a, b]$, it knows that it can affect the probability of H_i being in interval $[a, b]$ and thus of triggering the message. Thus, expected beliefs as calculated across all health states depend on both conjectured effort ε_i and effort e_i. We use $\mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2\big]$ to denote the expected belief when financial markets interpret messages based on the conjectures $\varepsilon_1 = \varepsilon_2 = \underline{e}$ and when the probability of a health state occuring is determined by $e_1 = \bar{e}$ and $e_2 = \underline{e}$.

Definition 5 *Central bank design $\mathcal{D}$ generates effort incentives if:*

$$\mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2\big] - \mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \underline{e}_1, \underline{e}_2\big] > 0 \tag{9}$$

where by Bayes plausibility of beliefs $\mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \underline{e}_1, \underline{e}_2\big] = \mathbb{E}[H_1 \mid \underline{e}_1] = \mathbb{E}_G[H]$.

Intuitively, incentives capture a bank's expected benefit from choosing high effort. We define incentives as the change in expected beliefs when a bank deviates from an equilibrium,[39] i.e. the other bank's effort and conjectured effort are fixed.

For example, in FDF markets always learn bank health perfectly ($\mu_i = H_i$). Thus, high effort results in $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}_F[H]$. Hence, FDF generates strong incentives $\mathbb{E}_F[H] - \mathbb{E}_G[H]$. In contrast, in ZDF markets never learn bank health. Thus, $\mu_i = \mathbb{E}[H_i]$ where the expectation is calculated based on financial market's conjectured effort. Hence, high effort results in $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}_G[H]$ and therefore ZDF generates zero incentives. We now rank designs by the strength of incentives they generate.

5.2 Comparing Designs in Terms of Incentives

This subsection shows that it can be socially beneficial to set the CB a mandate which is tougher than welfare maximisation as this increases incentives. To establish this, we characterize incentives for all central bank designs and show that incentives in a HRF are higher for tougher mandates and lie between those of ZDF and FDF.

Proposition 3 *Central bank designs $\mathcal{D}$ can be ranked by the strength of incentives generated:*

(i) FDF provides more incentives than HRF which provides more than ZDF.

(ii) Within HRF, the strength of incentives is weakly decreasing in the weight the mandate places on the weak bank.

Proof: See Appendix C.3.

Calculating incentives in a HRF is complicated because banks choose effort but take conjectured effort as given.[40] While in some cases the constant conjectured effort makes expected beliefs more responsive to effort than expected fundamentals, the reverse is true in other cases. For example, when effort leads to an improvement in H_1 from $H_1 < y(H_2)$ to H_1' where $H_1 < H_1' < y(H_2)$, then beliefs are unaffected since the CB sends $\{H_2, f, p\}$ in both cases. On the other hand, if $H_1 < y(H_2) < H_1'$, then beliefs increase more strongly than H_1 as the CB switches from sending $\{H_2, f, p\}$ to $\{H_1', p, p\}$.[41] To calculate incentives, we need to find the expected effect across all possible states.

[39] Definition 5 considers deviations from a low effort equilibrium. An analogous definition in terms of deviations from a high effort equilibrium is $\mathbb{E}_F[H] - \mathbb{E}[\mu(\bar{e}) \mid \underline{e}_1, \bar{e}_2] > 0$. The ordering of communication frameworks in terms of incentives is identical in both cases.

[40] Incentives in ZDF and FDF are more straightforward to calculate and were thus used to explain Definition 5.

[41] As a third case, beliefs can change in line with fundamentals. When effort leads to an improvement from $H_1 > x(H_2)$ to $H_1' > H_1 > x(H_2)$, the CB changes from sending $\{H_1, p, f\}$ to $\{H_1', p, f\}$.

To overcome this challenge, note that incentives in a HRF are:

$$\begin{aligned}\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \int_0^{x^{-1}(1)} & \Bigg[\int_0^{y(H_2)} \mathbb{E}_G\Big[H \mid H < y(H_2)\Big] f(H_1)dH_1 \\ & + \int_{y(H_2)}^{H_2} \frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}_G\Big[H \mid H_1 < H < x(H_1)\Big] f(H_1)dH_1 \\ & + \int_{H_2}^{x(H_2)} \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}_G\Big[H \mid H_2 < H < x(H_2)\Big] f(H_1)dH_1 \\ & + \int_{x(H_2)}^{1} H_1 f(H_1)dH_1\Bigg] g(H_2)dH_2 \\ + \int_{x^{-1}(1)}^{1} & \Bigg[\int_0^{y(H_2)} \mathbb{E}_G\Big[H \mid H < y(H_2)\Big] f(H_1)dH_1 \\ & + \int_{y(H_2)}^{H_2} \frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}_G\Big[H \mid H_1 < H < 1\Big] f(H_1)dH_1 \\ & + \int_{H_2}^{1} \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}_G\Big[H \mid H_2 < H < 1\Big] f(H_1)dH_1\Bigg] g(H_2)dH_2\end{aligned} \tag{10}$$

We find that incentives in a HRF are weakly decreasing in the weight the mandate places on the weak bank and lie between the incentives provided by FDF and ZDF. Moreover, in a HRF even the softest possible mandate creates strictly stronger incentives than ZDF. This holds because in a HRF the level of stress provides some information on bank weakness for any mandate.[42] Additionally, in a HRF even a mandate which is concerned only with average bank health creates strictly weaker incentives than FDF.

Delegating stress tests to a CB which is tougher than welfare can be socially beneficial as this increases incentives. The increase in incentives arises because delegation affects what messages banks expect the CB to send. For the planner, delegating to a tougher CB is thus a form of committing to reveal bank weakness at more health states. As banks know that delegating has made the threat of revealing weakness credible, they are incentivised to choose high effort. While this section defined incentives in terms of a deviation from equilibrium, the next section solves for the equilibrium.

5.3 Incentives and Equilibrium Bank Effort

This subsection shows that a communication framework which generates strong incentives results in banks taking high prudent actions in equilibrium. Thus, the increase in incentives caused by delegating to a CB which is tougher than welfare can cause a shift from an equilibrium with low prudent actions to an equilibrium with high prudent actions. This can increase welfare.

Proposition 4

[42] Example: For any mandate, it is optimal to send $\{0.1, p, p\}$ only if at least one bank has health exactly at 0.1.

(i) In ZDF, there always exists a unique pure strategy equilibrium. In this equilibrium both banks choose low effort.

(ii) In FDF, there always exists a unique pure strategy equilibrium. In this equilibrium both banks choose high effort.

(iii) In HRF, there exist two thresholds $\underline{t}(\mathcal{M})$ and $\bar{t}(\mathcal{M})$ with $\underline{t}(\mathcal{M}) < \bar{t}(\mathcal{M})$ such that:

(a) For all $c \leq \underline{t}(\mathcal{M})$, there exists a unique pure strategy equilibrium. In this equilibrium, both banks choose high effort.

(b) For all $c \geq \bar{t}(\mathcal{M})$, there exists a unique pure strategy equilibrium. In this equilibrium, both banks choose low effort.

(c) For all $\underline{t}(\mathcal{M}) < c < \bar{t}(\mathcal{M})$, there does not exist a pure strategy equilibrium.

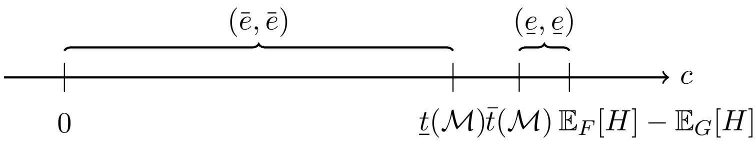
where $\bar{t}(\mathcal{M}) = \mathbb{E}[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2] - \mathbb{E}_G[H]$ and $\underline{t}(\mathcal{M}) = \mathbb{E}_F[H] - \mathbb{E}[\mu_1(\bar{\boldsymbol{e}}) \mid \underline{e}_1, \bar{e}_2]$.

Proof: See Appendix C.4.

Proposition 4 means that stronger incentives ensure that high effort occurs in equilibrium also when effort is more costly. To see how incentives determine equilibrium effort, consider ZDF and FDF. ZDF provides no incentives. This means that expected beliefs are unaffected by effort. Therefore, banks have no benefit of exerting effort. As this reasoning is independent of conjectured effort or the other bank's effort, the unique equilibrium exhibits low effort. FDF provides strong incentives ($\mu_i = H_i$). Thus, effort has the same effect on expected beliefs and expected health, independent of conjectured effort or the other bank's effort. Since $\mathbb{E}_F[H] - \mathbb{E}_G[H] > c$, this ensures that the equilibrium exhibits high effort.

The role of incentives in shaping equilibrium effort applies also in a HRF, though the formal treatment is complicated by the distinction between effort and conjectured effort. Fig 3 illustrates the result of Proposition 4 (iii). The thresholds $\underline{t}(\mathcal{M})$ and $\bar{t}(\mathcal{M})$ capture the strength of incentives and are higher if the mandate is tougher. Thus, a tougher mandate ensures that the high effort equilibrium occurs even at larger c. The thresholds diverge ($\underline{t}(\mathcal{M}) \neq \bar{t}(\mathcal{M})$). While $\underline{t}(\mathcal{M})$ and $\bar{t}(\mathcal{M})$ both capture incentives, the calculation of incentives is based on a different equilibrium and thus on different conjectured effort and other bank's effort. Whereas $\bar{t}(\mathcal{M})$ captures incentives from a low effort equilibrium, $\underline{t}(\mathcal{M})$ refers to the high effort equilibrium.

Figure 3: Equilibrium Effort in a Hurdle Rate Framework



Note: Thresholds depicted for the case where $G(H) = H, F(H) = H^2, V(\mu_i, \mu_2) = \frac{\mu_1+\mu_2}{2}$.

There exist costs of effort where no $\{HRF, \mathcal{M}\}$ can generate incentives which are sufficiently strong to induce high effort. This arises because even the toughest possible mandate in a

HRF provides weaker incentives than FDF. Henceforth, we refer to "high costs of effort" if no $\{HRF, \mathcal{M}\}$ can induce high effort and otherwise refer to "low costs of effort".

Setting the CB with a mandate that is tougher than welfare maximisation, or equally delegating to a tougher CB, can be socially beneficial as this increases incentives. Increased incentives, in turn, can increase equilibrium effort. Since $\mathbb{E}_F[H] - \mathbb{E}_G[H] > 0$ and c is a private cost, the switch from an equilibrium with low effort to one with high effort is socially beneficial, ceteris paribus. However, as this switch is achieved by delegating to a tougher CB, reassurance has decreased. Thus, the incentive and the reassurance motive of stress tests are conflicting. While, so far, we studied reassurance and incentives separately, the next section studies the planner's problem when he trades off incentives and reassurance.

6 Central Bank Design

We characterise the optimal central bank design and find that a hurdle rate framework with the softest possible mandate that induces high effort is the the optimal design when the cost of effort is low and the effect of effort and bank health is large relative to the planner's concern for the weak bank. Intuitively, HRF is optimal in many settings as it provides an intermediate combination of incentives and reassurance. FDF generates strong incentives but no reassurance and is therefore only optimal if no HRF can induce high effort, effort has a large effect on bank health, and society places only a small weight on reassurance. ZDF generates strong reassurance but no incentives and is therefore only optimal if effort has small effects on bank health and society places a large weight on reassurance.

Interestingly, the optimal mandate can differ in either direction from welfare maximisation. Formally, we establish that for any weight the planner places on the weak bank, there exist parameters such that the optimal design is a HRF with the softest possible mandate that induces high effort. This means that in some cases it is optimal for the planner to set the CB a mandate which is tougher than welfare maximisation while in other cases it is optimal to set a softer mandate.

To derive this, we first show that the planner's problem reduces to a choice among three alternatives: ZDF, FDF, or HRF with the softest possible mandate that induces high effort.[43] Then, we solve for the optimal design. We interpret these results in terms of strategic delegation.

6.1 The Optimal Mandate

As a first step in solving for the optimal design, this subsection shows that the design choice problem reduces to a choice among three alternatives: ZDF, FDF, or HRF with the softest possible mandate that induces high effort. All HRF with mandates which do not induce high effort are dominated by ZDF since ZDF provides more reassurance and neither design induces high effort. Similarly, all HRF with mandates that induce high effort are dominated by the

[43] Since in ZDF all mandates result in the same equilibrium, we economize on notation and refer to $\{ZDF, \mathcal{M}\}$ as ZDF. Similarly, $\{FDF, \mathcal{M}\}$ becomes FDF. In HRF, $\mathcal{M}$ affects the equilibrium. Thus, we maintain $\{HRF, \mathcal{M}\}$.

softest possible mandate which induces high effort as the extra incentives generate no benefit in terms of effort while they have the cost of reducing reassurance.

Proposition 5 *The design choice problem reduces to a choice among three alternatives: FDF, ZDF, $\{HRF, \tilde{\mathcal{M}}\}$ where $\tilde{\mathcal{M}}$ is the softest possible mandate which incentivises high effort. This means $\tilde{\mathcal{M}}$ is defined by $\underline{t}(\tilde{\mathcal{M}}) = c$ and thus is tougher for larger c.*

Proof: See Appendix C.5.

This result arises because reassurance is socially beneficial per se but incentives are beneficial only if they cause an increase in effort. Thus, all $\{HRF, \mathcal{M}\}$ which do not induce high effort result in lower welfare than ZDF as neither achieves high effort while ZDF provides more reassurance. Similarly, all $\{HRF, \mathcal{M}\}$ which induce high effort result in lower welfare than $\{HRF, \tilde{\mathcal{M}}\}$ as all these designs achieve high effort while $\{HRF, \tilde{\mathcal{M}}\}$ provides more reassurance.

When the cost of effort is low, the optimal design is either ZDF or $\{HRF, \tilde{\mathcal{M}}\}$. Recall that we defined low costs to ensure that $\{HRF, \tilde{\mathcal{M}}\}$ exists. Thus, $\{HRF, \tilde{\mathcal{M}}\}$ achieves higher welfare than FDF as both incentivise high effort while $\{HRF, \tilde{\mathcal{M}}\}$ achieves more reassurance.

When the cost of effort is high, the optimal design is either ZDF or FDF. High costs of effort mean that all $\{HRF, \mathcal{M}\}$ result in low effort. Thus, all $\{HRF, \mathcal{M}\}$ result in lower welfare than ZDF. However, FDF induces high effort and can, depending on parameters, result in higher welfare than ZDF. We now characterize the optimal design for every parameter combination.

6.2 Optimal Central Bank Design

We show that, for a large set of parameters, the optimal design is a hurdle rate framework with the softest possible mandate that induces high effort ($\{HRF, \tilde{\mathcal{M}}\}$). A HRF is optimal because it provides an intermediate combination of incentives and reassurance whereas FDF generates strong incentives but no reassurance. The reverse is true for ZDF.

Design $\{HRF, \tilde{\mathcal{M}}\}$ is the optimal design if cost of effort are low and simultaneously the effect of effort on bank health is large relative to the planner's concern for the weak bank. When costs of effort are so high that no HRF induces high effort, but high effort has a strong effect on bank health relative to the planner's concern for the weak bank, FDF is the optimal design. ZDF is only optimal when effort has a small effect on bank health and simultaneously the planner places a large weight on the weak bank.

To formalise these optimal central bank design results, we draw on the familiar functional forms $W(\mu_1, \mu_2) = \lambda \ \min(\mu_1, \mu_2) + (1 - \lambda) \ \frac{\mu_1 + \mu_2}{2}$ and $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1 - \omega) \ \frac{\mu_1 + \mu_2}{2}$. Additionally, let a capture the strength with which effort improves the bank health distribution. Formally, let low effort lead to $G(H) = H$ and high effort to $F(H) = H^a$ where $a > 1$ ensures an MLRP ordering.

Proposition 6 *Optimal Central Bank Design*

(i) When costs of effort are low, i.e. $c \leq \underline{t}(\mathcal{M}(0))$, the optimal central bank design is:

(a) $\{HRF, \tilde{\mathcal{M}}\}$ *iff* $a > \lambda'(a, C)$

(b) ZDF otherwise

where $\lambda'(a, C) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1, \mu_2)]}$ *and thus* $\lambda'(a, C)$ *is increasing in the cost of effort because* $\tilde{\omega}$ *is decreasing in the cost of effort.*

(ii) When costs of effort are high, i.e. $c > \underline{t}(\mathcal{M}(0))$*, the optimal central bank design is:*

(a) FDF iff $a > \lambda^\star(a)$

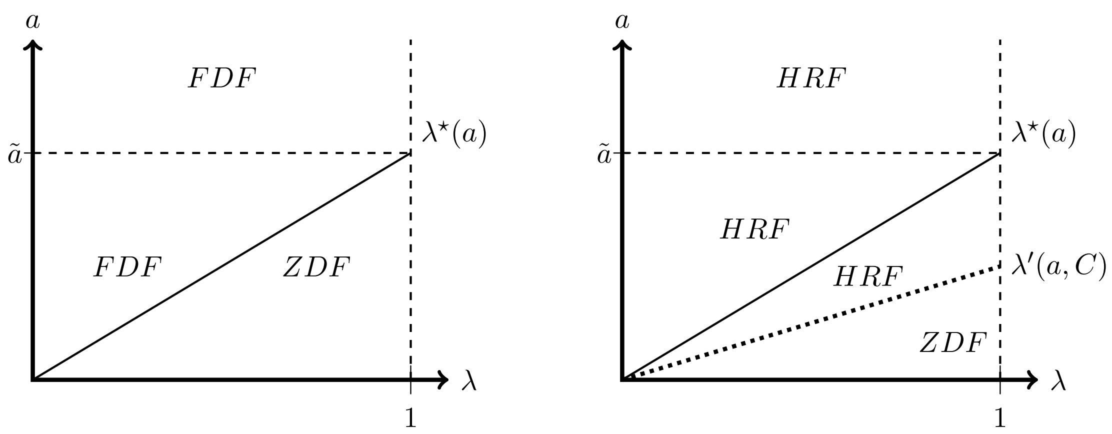
(b) ZDF otherwise

where $\lambda^\star(a) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)]}$ *and thus* $\lambda^\star(a) > \lambda'(a, C)$ *always holds.*

Proof: See Appendix C.6.

Figure 4: Optimal Central Bank Design

(a) High cost of effort (b) Low cost of effort



Note: $G(H) = H$, $F(H) = H^a$ with $a > 1$; HRF refers to $\{\text{HRF}, \tilde{\mathcal{M}}\}$; $\tilde{a} = \mathbb{E}_F[\min(H_1, H_2)] - \mathbb{E}_G[H]$.

For high costs of effort, the optimal design is FDF if and only if the improvement in bank health outweighs the reduction in reassurance. Otherwise ZDF is optimal. For a given weight on reassurance (λ), there exists a threshold benefit of bank effort (a) at which the planner is indifferent between ZDF and FDF. For all larger a, it is optimal to choose FDF. The more weight the planner places on reassurance, the larger the threshold benefit of effort becomes. This is depicted in Figure 4a where the partitioning frontier $\lambda^*(a)$ is increasing.

For low costs of effort, the optimal design is $\{HRF, \tilde{\mathcal{M}}\}$ if and only if the improvement in bank health outweighs the reduction in reassurance. Otherwise ZDF is optimal. Analogous to the case with high costs of effort, this gives rise to a threshold benefit of effort at which the planner is indifferent between $\{HRF, \tilde{\mathcal{M}}\}$ and ZDF. When the reassurance motive is larger, the planner requires a stronger effect of effort to choose $\{HRF, \tilde{\mathcal{M}}\}$ over ZDF. The resulting partitioning frontier $\lambda'(a, C)$, which is increasing, is depicted in Figure 4b. Since the reduction in reassurance is lower for $\{HRF, \tilde{\mathcal{M}}\}$ than for FDF, $\lambda'(a, C)$ lies strictly below $\lambda^*(a)$.

A hurdle rate framework is the optimal design for a large set of parameters because it provides an intermediate combination of incentives and reassurance. Specifically, when the cost of effort is low and the benefit high, $\{HRF, \tilde{\mathcal{M}}\}$ improves on FDF, which would otherwise have been chosen, because $\{HRF, \tilde{\mathcal{M}}\}$ provides more reassurance while both achieve high effort. Moreover, $\{HRF, \tilde{\mathcal{M}}\}$ improves on ZDF for parameters in the area between $\lambda^*(a)$ and $\lambda'(a, C)$. While for these parameters ZDF is preferred to FDF, which induces effort at a high cost in terms of foregone reassurance, $\{HRF, \tilde{\mathcal{M}}\}$ induces effort at a lower cost and achieves even higher welfare than ZDF.

6.3 Optimal Strategic Delegation

We show that the optimal mandate can differ in either direction from welfare maximisation. This means that in some cases it is optimal for the planner to delegate stress tests to a tougher CB while in others, perhaps surprisingly, it is optimal to delegate to a softer CB. This arises because the increase in reassurance caused by delegating to a softer CB is socially valuable per se while the corresponding reduction in incentives constitutes a social cost only if it results in lower effort. Moreover, even when the change in incentives shifts effort, but this shift only has a small effect on bank health, it is possible that the increase in reassurance outweighs the reduced incentives. Formally, the strategic delegation result is a Corollary of Proposition 6.

Corollary 1

(i) The optimal central bank mandate generically differs from welfare maximisation.

(ii) There exist cases where the optimal mandate is tougher than welfare while in other cases it is softer.

Formally, for any strength of society's reassurance motive (λ), there exist combinations of a and c such that $\{HRF, \tilde{\mathcal{M}}\}$ is the optimal design. Recall that $\tilde{\mathcal{M}}$ is independent of λ. For high λ, $\tilde{\mathcal{M}}$ is a tougher mandate, while for low λ, $\tilde{\mathcal{M}}$ is a softer mandate.

Delegating to a tougher CB can be optimal when a mandate to maximise welfare generates incentives which are not strong enough to induce high effort. Provided costs of effort are low, there exist tougher mandates that induce high effort. When effort has large effects on bank health, the optimal design is $\{HRF, \tilde{\mathcal{M}}\}$ where $\tilde{\mathcal{M}}$ is tougher than welfare. In this case, the benefit of delegating to a tougher CB (increased incentives due to banks anticipating that their weakness will be revealed) outweighs the cost (reduced reassurance due to financial markets interpreting messages more sceptically when stress is low and all banks pass).

More surprisingly, it is possible that delegating to a softer CB is optimal. This is the case when a mandate to maximise welfare generates stronger incentives than needed to induce high effort. Then, there exist mandates which are softer than welfare maximisation and also achieve high effort, but provide more reassurance. Such softer mandates increase welfare provided welfare puts some, potentially small, weight on reassurance.

The result that a HRF can be the optimal design leads to the delegation interpretation discussed above and also implies that stress tests become an informational contagion channel. We discuss this in the next section.

7 Informational Contagion

We show that stress tests become an informational contagion channel when a HRF is used. This means that changes in the health of one bank affect beliefs about the health of other banks. This arises because the CB optimally responds to a change in one bank's health by adjusting the severity of stress. As all banks are judged against the same severity, the messages, and thus the beliefs, about all banks are affected.

Informational contagion can take different forms. There are cases where a deterioration in the health of one bank results in lower beliefs about the other bank. More surprisingly, there are also cases where a deterioration in the health of one bank increases the belief about the other bank.

As HRF can be the optimal design, these results show that informational contagion can be a feature of an optimal design when a time inconsistency problem exists. Contagion need not reduce welfare. We first outline the different forms of contagion. Then, we show that the mechanism that turns stress tests into an informational contagion channel - judging all banks against the same severity of stress - is the same mechanism that makes HRF optimal for many parameters.

7.1 Contagion Patterns

A deterioration in the health of one bank can decrease the belief about the health of the other bank. This arises when the CB responds to the deterioration by lowering the severity of stress. Compare (H_1, H_2) where $x(H_2) > H_1 > H_2$ and thus $\{H_2, p, p\}$ is sent, to (H_1, H_2') where $H_2' < H_2$ and $x(H_2') > H_1$. At (H_1, H_2'), the CB sends $\{H_2', p, p\}$. Thus, the CB has lowered the severity of stress. While both banks continue to pass, they pass a weaker test, which is a less favourable message. As a result, beliefs about both banks are lower.

A deterioration in the health of one bank can increase the belief about the health of the other bank. This arises when the CB responds to the deterioration by increasing the severity of stress. Consider (H_1, H_2) where $x(H_2) > H_1 > H_2$ and a deterioration to (H_1, H_2') where $H_2' < H_2$ and $H_1 > x(H_2')$. In response to this deterioration, the CB changes its message from $\{H_2, p, p\}$ to $\{H_1, p, f\}$. Thus, the CB has increased the severity of stress. While this amplifies the effect on the weak bank - μ_2 falls more strongly than H_2 - it increases beliefs about the strong bank as it passes a tougher test which is a more favourable message.

Neither pattern requires assumptions on the mandate. Instead, contagion arises in a HRF for every mandate. Intuitively, the CB faces a choice between lowering the severity of stress, which results in sending a slightly weaker message about both banks, and raising the severity of stress, which results in one bank failing the test, a much weaker message, and in a slightly stronger message about the bank that passes. The former can be preferred to the latter even when the mandate places no additional weight on the weak bank.[44]

7.2 Contagion in the Optimal Design

While contagion has been found to be socially harmful (e.g. Allen et al. (2012)), our model shows that contagion can be a feature of the socially optimal design. Formally, this holds because contagion is a feature of any HRF and HRF is the optimal design for a large set of parameters.

Intuitively, the mechanism that turns stress tests into an informational contagion channel - judging all banks against the same severity of stress - is the same mechanism that makes HRF optimal for many parameters as it results in an intermediate combination of incentives and reassurance. Without the constraint, i.e. in FDF or ZDF, there would be not contagion and either incentives or reassurance, but not a combination of both.[45]

The result that contagion is a feature of the socially optimal design arises because of three components of our model: stress tests are used to create incentives for banks, to reassure financial markets, and commitment is not possible. Without the need for incentives, ZDF is optimal. Without the need for reassurance, FDF is optimal. With the ability to commit, choosing e.g. $s = 0.5$ for all health outcomes achieves an intermediate combination of incentives and reassurance. None of these alternatives implies contagion. When incentives, reassurance, and no commitment are present, the optimal design can imply contagion.

8 Conclusion

We showed that central banks face a time inconsistency problem when publishing bank stress test results. Before a stress test, they want to *appear* tough as the threat of letting banks fail the stress test incentivizes prudent behaviour. After the stress test, they want to *act* soft by releasing only partial information in order to reassure financial markets about bank health.

We characterised an institutional design solution to this commitment problem: a social planner specifies the central bank's mandate and sets the framework within which the central bank communicates. We find that a hurdle rate framework, where all banks are judged to pass or fail relative to a common level of stress, is optimal in many settings as it generates intermediate levels of both incentives and reassurance.

In a hurdle rate framework, *strategic delegation* provides additional benefits. This means the planner maximises ex-ante welfare by setting the central bank's mandate to differ from ex-post welfare maximisation. The optimal mandate can differ in either direction. If taking prudent actions has high costs, the planner optimally specifies the CB's mandate to be "tougher" than welfare maximisation, so that the CB is less concerned about supporting weak banks than society. This makes it credible that the CB will reveal bank weakness by letting weak banks fail the stress

[44] Formally, contagion arises in any equilibrium in which $\{\min(H_1, H_2), p, p\}$ and $\{H_1, p, f\}$ messages are sent. Theorem 1 shows that this is the case for all mandates.

[45] Recall that the case with as many stress scenarios as banks equals FDF in equilibrium. Thus, communicating results as pass or fail relative to thresholds need not result in contagion. Contagion arises only when there are fewer thresholds than banks.

test and thus generates strong incentives. If the costs of prudent actions are low, the optimal mandate is "softer" than welfare maximisation, so that the CB is very concerned about weak banks. This achieves more reassurance while still generating sufficient incentives to induce the prudent action.

Informational contagion can be a feature of the socially optimal design when a time inconsistency problem exists. Thus, contagion need not reduce welfare. Contagion arises in a hurdle rate framework because the central bank optimally responds to a deterioration in one bank's health by adjusting the severity of stress. In some cases, it is optimal for the central bank to lower the severity of stress. All banks pass, but as they pass a weaker test, beliefs about all banks are lower. More surprisingly, there are also cases where a deterioration in one bank's health leads to higher beliefs about other banks. This is the case when the central bank increases the severity of stress.

These results lead to a reinterpretation of past stress tests. Stress tests by the EBA or its predecessor (the CEBS) were repeatedly criticised for being too soft, because "many European banks passed the annual exams in July yet still had their shares trashed by investors."[46] This led to debates about the regulator's competence. Our model shows that such an episode can arise from optimal behaviour by the regulator. In our model, a regulator who faces unhealthy banks and has committed to judging them against a common stress severity, optimally chooses a low severity of stress, which all banks pass. Financial markets understand that the low stress severity was not an accident but was chosen strategically because banks are weak. As a result, financial markets believe that banks are weak. Our model shows that this occurs in equilibrium for every mandate, i.e. even for "tough" mandates which place no additional weight on the weak bank. Moreover, committing to use a hurdle rate framework can itself be an optimal choice as it results in an intermediate combination of incentives and reassurance.

This paper adds a new dimension to the debate on stress test communication. We highlight that welfare depends not just on the communication framework, but also on the central bank's mandate. Current central bank designs do not specify an objective which a central bank should pursue in its communication. Absent an explicit mandate, central banks are likely to attempt to maximise ex-post welfare. Thus, our paper suggests that unexploited welfare gains exist. Giving central banks an explicit mandate which differs from ex-post welfare maximisation generates gains from strategic delegation.

[46] "Which part of 'stress test' do the eurozone's policy makers not understand? That so many European banks passed the annual exams in July yet still had their shares trashed by investors says it all: the pass mark was too low and the questions were too narrow." Financial Times, Lex, European stress tests: a grim backdrop, 6 October 2011.

References

Aghion, Philippe, Patrick Bolton, and Steven Fries, "Optimal Design of Bank Bailouts: The Case of Transition Economies," *Journal of Institutional and Theoretical Economics*, March 1999, *155* (1), 51–70.

Allen, Franklin, Ana Babus, and Elena Carletti, "Asset commonality, debt maturity and systemic risk," *Journal of Financial Economics*, 2012, *104* (3), 519–534.

Alvarez, Fernando and Gadi Barlevy, "Mandatory disclosure and financial contagion," *Journal of Economic Theory*, 2021, *194*.

Barro, Robert J. and David B. Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy*, 1983, *91* (4), 589–610.

_ and _ , "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, 1983, *12* (1), 101–121.

Bergemann, Dirk and Stephen Morris, "Information Design: A Unified Perspective," *Journal of Economic Literature*, 2019, *57* (1), 1–57.

Blinder, Alan S, Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen, "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence," *Journal of Economic Literature*, 2008, *46* (4), 910–945.

Boleslavsky, Raphael and Christopher Cotton, "Grading Standards and Education Quality," *American Economic Journal: Microeconomics*, May 2015, *7* (2), 248–279.

Bouvard, Matthieu, Pierre Chaigneau, and Adolfo De Motta, "Transparency in the Financial System: Rollover Risk and Crises," *The Journal of Finance*, 2015, *70* (4), 1805–1837.

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

DeMarzo, Peter M., Ilan Kremer, and Andrzej Skrzypacz, "Test Design and Minimum Standards," *American Economic Review*, 2019, *109* (6), 2173–2207.

Dewatripont, Mathias and Jean Tirole, "Efficient governance structure: implications for banking regulation," in Colin Mayer and Xavier Vives, eds., *Capital Markets and Financial Intermediation*, Cambridge University Press, 1993, chapter 2, pp. 12–35.

_ and _ , "A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence," *The Quarterly Journal of Economics*, November 1994, *109* (4), 1027–1054.

Diamond, Douglas W. and Philip H. Dybvig, "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 1983, *91* (3), 401.

Dubey, Pradeep and John Geanakoplos, "Grading exams: 100, 99, 98, ... or A, B, C?," *Games and Economic Behavior*, 2010, *69*, 72–94.

Dye, Ronald A., "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 1985, *23* (1), 123–145.

Faria-e-castro, Miguel, Joseba Martinez, and Thomas Philippon, "Runs versus Lemons: Information Disclosure and Fiscal Capacity," *Review of Economic Studies*, 2016, pp. 1–25.

Geraats, Petra M, "Central Bank Transparency," *The Economic Journal*, 2002, *112*, F532–F565.

Goldstein, Itay and Haresh Sapra, "Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits," *Foundations and Trends in Finance*, 2013, *8* (1), 1–54.

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

Grossman, Sanford J, "The Informational Role of Warranties and Private Disclosure about Product Quality," *The Journal of Law and Economics*, 1981, *24* (3), 461–483.

Grossman, S.J. and O.D. Hart, "Disclosure Laws and Takeover Bids," *Journal of Finance*, 1980, *35* (2), 323–334.

Guttman, Ilan, Ilan Kremer, and Andrzej Skrzypacz, "Not Only What but Also When: A Theory of Dynamic Voluntary Disclosure," *American Economic Review*, August 2014, *104* (8), 2400–2420.

Hirshleifer, Jack, "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, September 1971, *61* (4), 561–574.

Holmstrom, Bengt, "Moral Hazard in Teams," *The Bell Journal of Economics*, 1982, *13* (2), 324–340.

Huang, Jing, "Optimal Stress Tests in Financial Networks," *Working Paper*, 2021.

Inostroza, Nicolas and Alessandro Pavan, "Persuasion in Global Games with Application to Stress Testing," *Working Paper*, January 2018.

Jovanovic, Boyan, "Truthful Disclosure of Information," *The Bell Journal of Economics*, 1982, *13* (1), 36–44.

Jung, Woon-Oh and Young K. Kwon, "Disclosure When the Market is Unsure of Information Endowment of Managers," *Journal of Accounting Research*, 1988, *26* (1), 146–153.

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

Keister, Todd, "Bailouts and Financial Fragility," *Review of Economic Studies*, 2016, *83*, 704–736.

Kydland, Finn E. and Edward C. Prescott, "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, June 1977, *85* (3), 473–492.

Lambert, Richard A., "Long-Term Contracts and Moral Hazard," *The Bell Journal of Economics*, 1983, *14* (2), 441–452.

Leitner, Yaron and Basil Williams, "Model Secrecy and Stress Tests," *Working Paper*, 2020.

Lucas, Robert E., "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 1972, *4*, 103–124.

_ , "An Equilibrium Model of the Business Cycle," *Journal of Political Economy*, December 1975, *83* (6), 1113–1144.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 1981, *12* (2), 380.

Orlov, Dmitry, Pavel Zryumov, and Andrzej Skrzypacz, "Design of Macro-prudential Stress Tests," *Working Paper*, May 2017.

Ottaviani, Marco and Peter Norman Sorensen, "Professional advice," *Journal of Economic Theory*, 2006, *126*, 120–142.

Parlatore, Cecilia, "Transparency and Bank Runs," *Working Paper*, 2015.

_ **and Thomas Philippon**, "Designing Stress Scenarios," *Working Paper*, 2020.

Quigley, Daniel and Ansgar Walther, "Inside and Outside Information: Reverse Unraveling and Stress Test Design," *Working Paper*, 2020.

Reis, Ricardo, "Central Bank Design," *The Journal of Economic Perspectives*, 2013, *27* (4), 17–43.

Rogoff, Kenneth, "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, November 1985, *100* (4), 1169–1189.

Shaked, Moshe and George J. Shanthikumar, *Stochastic Orders* Springer Series in Statistics, Springer, 2007.

Shapiro, Joel and David Skeie, "Information Management in Banking Crisis," *The Review of Financial Studies*, 2015, *28* (8), 2322–2363.

_ **and Jing Zeng**, "Stress Testing and Bank Lending," *Working Paper*, July 2018.

Verrecchia, Robert E., "Discretionary Disclosure," *Journal of Accounting Research*, 1983.

Vickers, John, "Delegation and the Theory of the Firm," *The Economic Journal*, 1985, *95*, 138–147.

Walsh, Carl E., "Optimal Contracts for Central Bankers," *The American Economic Review*, March 1995, *85* (1), 150–167.

Williams, Basil, "Stress Tests and Bank Portfolio Choice," *Working Paper*, 2017.

A Full Revelation with Multiple Items

Proposition 7 *When the sender (central bank) communicates about multiple items (banks) simultaneously, sets at least as many hurdles as there are banks, and truthfully communicates pass or fail for each bank relative to every hurdle, then every equilibrium is fully revealing, i.e. receivers (financial markets) learn each bank's health perfectly in every equilibrium of this communication game.*

This proposition serves two roles. It relates our results about communication on multiple items to the single item setting with its seminal results of full revelation (Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981). It also justifies the name of Full Disclosure Framework (FDF) that we give to a framework with at least as many hurdles as banks.

Proof: We focus on the case with two banks and two hurdles. Then, the sender's messages can be expressed as $\{s^a, o_1^a, o_2^a; s^b, o_1^b, o_2^b\}$ where s^j denotes a level of stress and $o_i^j \in \{p, f\}$ is the outcome (pass or fail) for bank i relative to stress level j. We establish three Lemmata which together prove proposition 7.

Lemma A.1 *There cannot exist a fully pooling equilibrium.*

Proof: To sustain a fully pooling equilibrium, i.e. an equilibrium where no information is released, the sender would have to send the same message at every possible health state pair (H_1, H_2) or randomise over the same messages. Thus, financial markets would believe that $\mu_1 = E(H_1)$, $\mu_2 = E(H_2)$. The only message which can implement full pooling, i.e. that can be sent from all health state pairs, is $\{0, p, p; 0, p, p\}$, i.e. both hurdles are set at zero and all banks pass relative to both hurdles.

This candidate equilibrium never is an equilibrium because the sender has a strictly beneficial deviation at some health state pairs. This deviation is beneficial for any interpretation rule which receivers could use. For example, at $(H_1 = 1, H_2 = 1)$, the sender can truthfully send $\{1, p, p; 1, p, p\}$. As this message is true if and only if $H_1 = 1, H_2 = 1$, receivers must interpret it as $\mu_1 = 1, \mu_2 = 1$. The sender with objectives $V(\mu_1, \mu_2)$ always benefits from this deviation since $\frac{\partial V}{\partial \mu_i} > 0 \; \forall \; i$ and $1 > E(H_i) \; \forall \; i$. *Q.E.D.*

Lemma A.2 *There cannot exist a partially pooling equilibrium.*

Proof: In any partially pooling equilibrium[47] the central bank has to send the same message for all health state pairs which are part of the pool or randomise over the same messages. Let Ω denote the set of health state pairs (H_1, H_2) which form the pool. First, we outline how such a pool could be implemented. Then, we show that for every pool and every implementation the sender has a strictly beneficial deviation.

[47] A partially pooling equilibrium is an equilibrium where receivers do not learn (H_1, H_2) perfectly, but learn that it is in a subset of the state space.

How could a pool be implemented? Define $\bar{H}$ to be the highest health level in the pool Ω, that is:

$$\bar{H} = \max\big[\sup_{H_1 \in \Omega} H_1, \sup_{H_2 \in \Omega} H_2\big] \tag{11}$$

and $\underline{H}$ to be lowest health level in the pool:

$$\underline{H} = \min\big[\inf_{H_1 \in \Omega} H_1, \inf_{H_2 \in \Omega} H_2\big] \tag{12}$$

Such a pool can be implemented by choosing hurdles s^1, s^2 which are not in the interval $[\underline{H}, \bar{H}]$. If at least one hurdle was in the interval, this would reveal information and destroy the pool.

Having outlined all possible implementations of a pool, we proceed to show that for every pool and every implementation, it is always possible for the central bank to deviate and benefit strictly. Hence, no partially pooling equilibrium can exist.

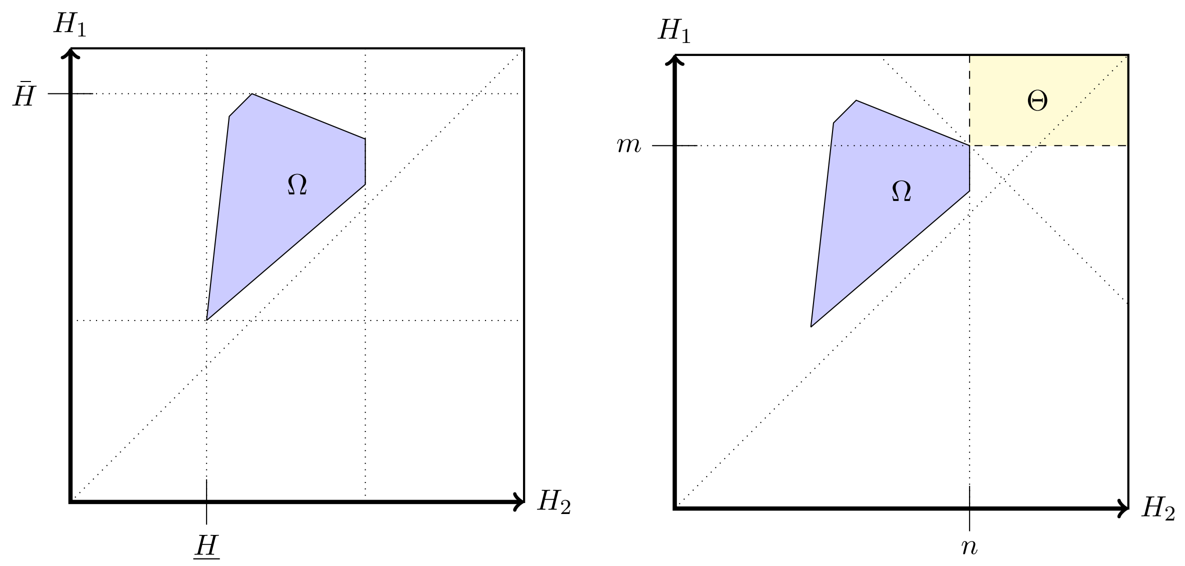
We introduce the term "truthful pay-off" and define it as the pay-off the central bank would achieve if health state pairs were perfectly revealed. That is, the truthful pay-off at (H_1, H_2) is $V(H_1, H_2)$.

Sublemma A.1 *A pool across different truthful pay-offs never exists.*

In every pool Ω, there exists a maximum truthful pay-off set M, which is the set of health state pairs that achieve the highest truthful pay-off. Denote one such health state pair (m, n). The central bank strictly benefits from deviating from the pool at (m, n) and can feasibly do so. Suppose without loss of generality that $m > n$. The central bank can deviate by sending $\{s^1 = m, p, f; s^2 = n, p, p\}$. This must be a deviation as it cannot be the implementation of pool Ω. Financial markets know that the health state pair must be to the top-right of (m, n), i.e. that $H_1 \geq m$ and $H_2 \geq n$. For any belief financial markets hold over the states supporting this message, the central bank is strictly better off than when sending the pooling message and pooling in Ω. This is depicted in Figure 5b.

Figure 5: A Beneficial Deviation from Partial Pooling Always Exists

(a) A Candidate Partially Pooling Equilibrium (b) Beneficial Deviation



Note: Ω is the set of health state pairs which constitute the pool. Only $s^i \notin]\underline{H}, \bar{H}]$ implement the pool. $(m, n) \in M$. Θ is the set of health state pairs at which it is feasible to send message $\{s^1 = m, p, f; s^2 = n, p, p\}$.

Importantly, even for the worst possible interpretation about where the deviation to $\{s^1 = m, p, f; s^2 = n, p, p\}$ comes from, the deviation is still a strict improvement for the central bank. Hence, no partially pooling equilibrium across different truthful pay-off pairs can exist. *Q.E.D.*

Sublemma A.2 *A pool of different health state pairs with the same truthful pay-off never exists.*

It is not possible that health state pairs with the same truthful pay-off form a pool because there does not exist a message which is *i)* feasible at health state pairs with the same truthful pay-off and which is *ii)* not feasibly at health state pairs with a lower truthful pay-off.

Let (H_1^a, H_2^a) and (H_1^b, H_2^b) have the same truthful pay-off. Any message with $s^j \in]\min(H_i^a, H_i^b), \max(H_i^a, H_i^b)]$ allows receivers to distinguish between state H_i^a and state H_i^b. Thus, the only feasible messages which implement the pool have $s^j \leq \min(H_1^a, H_1^b, H_2^a, H_2^b)$ or $s^j > \max(H_1^a, H_1^b, H_2^a, H_2^b)$.[48]

However, all messages which implement the pool are also feasible at any (H_1^c, H_2^c) where $H_i^c \in [\min(H_1^a, H_1^b, H_2^a, H_2^b), \max(H_1^a, H_1^b, H_2^a, H_2^b)]\ \forall i = 1, 2$. Some (H_1^c, H_2^c) must have lower truthful pay-offs than (H_1^a, H_2^a). To see this, recall that $V(\mu_1, \mu_2)$ is concave. Thus, if $H_1^a > H_1^b$ and (H_1^a, H_2^a) has the same truthful pay-off as (H_1^b, H_2^b), we must have $H_2^a < H_2^b$. Thus $H_1^c = H_1^b$ and $H_2^c = H_2^a$ must have a lower truthful pay-off than (H_1^a, H_2^a), but at (H_1^c, H_2^c) it is feasible to send any message that could implement the pool of (H_1^a, H_2^a) and (H_1^b, H_2^b).

Thus, any message which can implement a pool across health state pairs with the same truthful pay-off is also feasible at some health state pairs with a lower truthful pay-off. Hence,

[48]It is also possible that s^1 satisfies one inequality and s^2 the other.

there cannot exist an equilibrium where health state pairs with the same truthful pay-off pool as there will always exist health state pairs with a lower truthful pay-off at which the CB can feasibly deviate to the pooling message. Moreover, any such deviation is beneficial since no alternative way to achieve a pay-off higher than the truthful pay-off can exist (see Sublemma A.1 above). *Q.E.D.*

Lemma A.3 *A fully revealing equilibrium exists.*

Proof: When the sender uses as many hurdles as there are banks, a fully revealing equilibrium is feasible. In the two bank case, the sender can choose $s_1 = H_1$, $s_2 = H_2$ and receivers interpret it as $\mu_1 = H_1$, $\mu_2 = H_2$, since every message is sent from a unique health state pair. No beneficial deviation exists for the sender since messages which would be interpreted more favourably (higher levels of stress, banks still passing) are not true. *Q.E.D.*

Combining the three lemmata above proves proposition 7. *Q.E.D.*

B Equilibrium in the Hurdle Rate Communication Game

B.1 Proof of Theorem 1: Existence

The collection $\{S, R\}$ specified in Theorem 1 is an equilibrium of the Hurdle Rate communication game if and only if conditions (i) and (ii) defined in Definition 1 both hold. We prove these in turn, using the logic of revealed preference. As in Theorem 1, we allow bank health distributions to be any distributions and to differ across banks. We denote the CDFs as $F(H_1)$ and $G(H_2)$ respectively.

Condition (i): The interpretation rule R means that, for a given outcome pair $\{o_1, o_2\}$, higher levels of stress result in higher beliefs μ_i for both banks. Since $\mathcal{M}$ is strictly increasing in $\mu_i \, \forall \, i$, we have:

Lemma B.1 *Given the interpretation rule R, the CB always prefers the highest level of stress possible for a given outcome pair to any alternative level of stress for that outcome pair.*

Formally, this means that the CB strictly prefers $\{\min(H_1, H_2), p, p\}$ to all $\{s', p, p\}$ where $s' < \min(H_1, H_2)$; the CB strictly prefers $\{H_1, p, f\}$ to all $\{s'', p, f\}$ where $s'' < H_1$; the CB strictly prefers $\{H_2, f, p\}$ to all $\{s''', f, p\}$ where $s''' < H_2$.

Proof: Since $\mathcal{M}$ is strictly increasing in μ_i and R specifies that μ_i are increasing in s for a given outcome pair, this Lemma holds. *Q.E.D.*

Lemma B.1 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$ in response to interpretation rule R. We proceed by showing that the sets of health state pairs (H_1, H_2) where a given $\{s, o_1, o_2\}$ is sent can be expressed via indifference frontiers $x(H_2)$ and $y(H_2)$ as in Theorem 1.

Lemma B.2 *If, at a health state pair (H_1, H_2) with $H_2 > 0$, sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1, H_2') with $H_2' < H_2$ it is mandate maximising to send $\{H_1, p, f\}$.*

Proof: At (H_1, H_2), $\{H_1, p, f\}$ was revealed preferred to $\{H_2, p, p\}$, i.e. $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Given the interpretation rule R, $\{H_2, p, p\}$ results in higher beliefs μ_i about both banks than $\{H_2', p, p\}$ and thus $V(\{H_2, p, p\}) > V(\{H_2', p, p\})$. Hence, $\{H_1, p, f\}$ must be preferred to $\{H_2', p, p\}$, i.e. $V(\{H_1, p, f\}) > V(\{H_2', p, p\})$. *Q.E.D.*

Lemma B.3 *If, at a health state pair (H_1, H_2), sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1', H_2) with $H_1' > H_1$ it is mandate maximising to send $\{H_1', p, f\}$.*

Proof: (Analogous to the proof of Lemma B.2) Revealed: $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Given R, $\{H_1', p, f\}$ results in higher μ_i for both banks than $\{H_1, p, f\}$. Thus, $V(\{H_1', p, f\}) > V(\{H_1, p, f\})$. Hence, we must have $V(\{H_1', p, f\}) > V(\{H_2, p, p\})$. *Q.E.D.*

Lemma B.4 *There never exists a perfectly vertical part of the $x(H_2)$ indifference frontier.*[49]

Proof by contradiction: Suppose a perfectly vertical part of the $x(H_2)$ indifference frontier existed. Then, on this part, there would be at least two health state pairs, (H_1^b, H_2) and (H_1^a, H_2) with $H_1^a > H_1^b$, at which the sender is indifferent between $\{H_2, p, p\}$ and respectively $\{H_1^a, p, f\}$ or $\{H_1^b, p, f\}$. Thus $V(\{H_1^b, p, f\}) = V(\{H_2, p, p\})$ and $V(\{H_1^a, p, f\}) = V(\{H_2, p, p\})$ must both hold. Therefore, we must have $V(\{H_1^a, p, f\}) = V(\{H_1^b, p, f\})$. But this conflicts with the interpretation rule R. Specifically, given R, $\{s, p, f\}$ results in higher μ_i than $\{s', p, f\}$ $\forall\ s > s'$ and $\forall \mu_i$ (monotonicity property). Thus, since $\frac{\partial V(\mu_1, \mu_2)}{\partial \mu_i} > 0 \forall i$, R ensures that $V(\{H_1^a, p, f\}) > V(\{H_1^b, p, f\})$ and therefore no perfectly vertical part of the $x(H_2)$ indifference frontier can exist. *Q.E.D.*

Proof of condition (i): Lemma B.1 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$. Lemma B.2 - B.3 showed that the health state pairs at which the CB sends $\{H_1, p, f\}$ form a connected set in the top left part of the state space $(H_1, H_2) \in [0, 1]^2$. Lemma B.4 showed that this connected set does not have a perfectly vertical boundary. The boundary also cannot be decreasing. Thus, given R, the CB's optimal strategy is the strategy described in Theorem 1, i.e. sending sending $\{H_1, p, f\}$ for all (H_1, H_2) above an $x(H_2)$ indifference frontier which is monotonically increasing. *Q.E.D.*

Condition (ii) : Given the strategy S, the interpretation rule is correct in the sense of using all available information. Naive interpretations of $\{s, p, p\}$ would result in $\mu_1 = \mathbb{E}[H \mid s \leq H \leq 1]$. Partially sophisticated interpretations would realise that $s = \min(H_1, H_2)$ and result in $\mu_1 = \alpha\ s + (1 - \alpha)\ \mathbb{E}_F[H \mid s \leq H \leq 1]$. Fully sophisticated beliefs realise that $s = \min(H_1, H_2)$

[49] The same logic applies to horizontal parts of $y(H_2)$.

and additionally that the sender preferred $s = \min(H_1, H_2)$ to $s = \max(H_1, H_2)$. Thus, fully sophisticated beliefs result in $\mu_1 = \alpha\, s + (1-\alpha)\, \mathbb{E}_F[H \mid s \leq H \leq x(s)]$ which corresponds to the beliefs in Theorem 1. *Q.E.D.*

B.2 Proof of Theorem 2: Uniqueness and Monotonicity

To prove Theorem 2, we first characterise an alternative weak PBE which exists but violates monotonicity. Then, we prove that the equilibrium described in Theorem 1 is the unique weak PBE which satisfies monotonicity.

An Alternative Equilibrium which Violates Monotonicity

Lemma B.5 *The following actions and beliefs can constitute an equilibrium of the hurdle rate communication game.*

$$
\begin{array}{llllll}
s = & H_1 & \textit{iff} & H_1 > x(H_2) & \rightarrow & \{H_1, p, f\} \\
s = & H_2 & \textit{iff} & y(H_2) > H_1 & \rightarrow & \{H_2, f, p\} \\
s = & \min(H_1, H_2) & \textit{iff} & x(H_2) \geq H_1 \geq y(H_2) \ \textit{and} & & \\
 & & & \textit{either} \min(H_1, H_2) < \kappa_1 \ \textit{or} \min(H_1, H_2) \geq \kappa_2 & \rightarrow & \{\min(H_1, H_2), p, p\} \\
s = & \kappa_1 & \textit{iff} & x(H_2) \geq H_1 \geq y(H_2) \ \textit{and} & & \\
 & & & \kappa_1 \leq \min(H_1, H_2) < \kappa_2 & \rightarrow & \{\kappa_1, p, p\}
\end{array}
$$

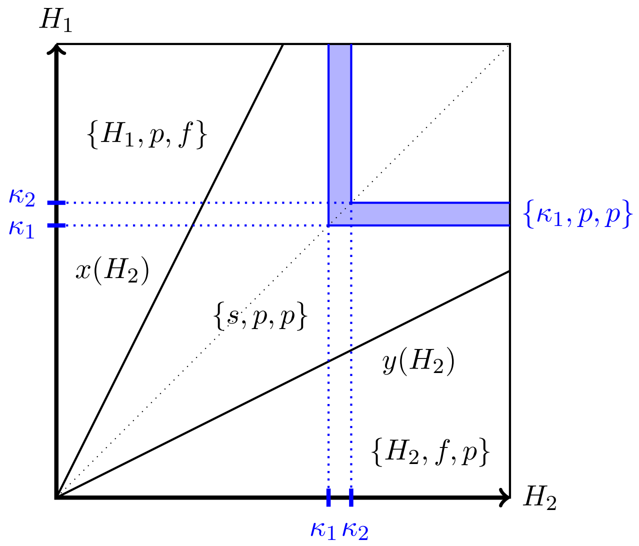
where $x(H_2)$ and $y(H_2)$ have the same properties as in the equilibrium described in Theorem 1.

Recipients form beliefs accordingly. Denoting the set of health state pairs from which $\{\kappa_1, p, p\}$ is sent as Ω, and assuming that bank health distributions are identical, beliefs are given by:

$$
\begin{array}{llll}
\{s, p, f\} & & \textit{is interpreted as} & \mu_1 = s; \quad \mu_2 = \mathbb{E}[H \mid H < x^{-1}(s)] \\
\{s, f, p\} & & \textit{is interpreted as} & \mu_2 = s; \quad \mu_1 = \mathbb{E}[H \mid H < y(s)] \\
\{s, p, p\} & \textit{if } s < \kappa_1 \textit{ or } s > \kappa_2 & \textit{is interpreted as} & \mu_1 = \mu_2 = \frac{1}{2}\, s + \frac{1}{2}\, \mathbb{E}[H \mid s \leq H \leq x(s)] \\
 & \textit{if } \kappa_1 < s < \kappa_2 & \textit{is interpreted as} & \mu_1 = \mu_2 = s \\
 & \textit{if } s = \kappa_1 & \textit{is interpreted as} & \mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] \\
\{s, f, f\} & & \textit{is interpreted as} & \mu_1 = \mu_2 = 0 \ \forall\, s > 0
\end{array}
$$

The messages $\{s, p, p\}$ for $\kappa_1 < s < \kappa_2$ are never sent in equilibrium. A PBE therefore implies no restrictions on how these messages are interpreted other than that beliefs are formed over health state pairs from which the message could have been sent. In this alternative equilibrium, these messages are interpreted in the most sceptical way possible. Diagrammatically:

Figure 6: Alternative Equilibrium: Communication Strategy



Proof: We prove the possibility that such an equilibrium can exist by example. Let $\kappa_1 = 0.6$, $\kappa_2 = 0.65$ and $F(H) = H$, $G(H) = H$. For these values we have $\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] = \frac{431}{600} \approx 0.718$.

Condition (i): No beneficial deviation exists for the CB: Focus on the region where $\min(H_1, H_2) \in [\kappa_1, \kappa_2[$. For other regions, e.g. the indifference frontiers $x(H_2)$, $y(H_2)$, the proof of Theorem 1 applies. For $\min(H_1, H_2) < \kappa_1$ and health state pairs where the CB sets $s = \max(H_1, H_2)$, it is technologically not possible to deviate to $\{\kappa_1, p, p\}$. For $\min(H_1, H_2) \geq \kappa_2$ it is possible but not optimal to deviate from $\{\min(H_1, H_2), p, p\}$ to $\{\kappa_1, p, p\}$. To see this, consider $\min(H_1, H_2) = \kappa_2$. Then $\{\kappa_2, p, p\}$ results in $\mu_1 = \mu_2 = \frac{1}{2}0.65 + \frac{1}{2}\mathbb{E}[H \mid H > 0.65] = 0.7375$. As $\{\kappa_1, p, p\}$ results in $\mu_1 = \mu_2 = 0.718$, the deviation to $\{\kappa_1, p, p\}$ is suboptimal. The same logic applies to all $\min(H_1, H_2) > \kappa_2$.

For $\min(H_1, H_2) \in [\kappa_1, \kappa_2[$ it is optimal to send $\{\kappa_1, p, p\}$. Alternative messages $\{s, p, p\}$ where $s < \kappa_1$ result in lower sender pay-off $V(\mu_1, \mu_2)$ as both μ_1 and μ_2 are lower; where $s > \kappa_2$ these messages are technologically available. When deviations to $\{s, p, p\}$ with $s \in [\kappa_1, \kappa_2[$ are available, these are also not optimal as the resulting beliefs $\mu_1 = \mu_2 = s$ are always strictly below the beliefs resulting from $\{\kappa_1, p, p\}$, i.e. $\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] = \frac{431}{600} \approx 0.718$. Messages of the form $\{s, p, f\}$ are available, but have been revealed to result in lower sender pay-offs than $\{s, p, p\}$ with s below but close to κ_1.

Condition (ii): is also satisfied. This concludes the proof of Lemma B.5. *Q.E.D.*

Lemma B.6 *The equilibrium characterised in Lemma B.5 violates monotonicity.*

Proof: In this equilibrium, $\{0.6, p, p\}$ is interpreted as $\mu_1 = \mu_2 = \mathbb{E}[H \mid (H_1, H_2) \in \Omega] = \frac{431}{600} \approx 0.718$. The off-equilibrium path message $\{0.61, p, p\}$ is interpreted as $\mu_1 = \mu_2 = 0.61$. This violates monotonicity which requires that passing a tougher stress test results in higher beliefs μ_i. *Q.E.D.*

Uniqueness Given Monotonicity

To prove that the equilibrium characterised in Theorem 1 is the unique weak PBE which satisfies monotonicity, we show that, for any interpretation rule R which results in beliefs that satisfy monotonicity, the CB's optimal strategy is unique and always identical to the strategy specified in Theorem 1. The proof is completed by the argument that the beliefs specified in Theorem 1 are the unique set of beliefs consistent with the CB's strategy.

Lemma B.7 *For any interpretation rule R which results in beliefs that satisfy monotonicity, the CB always prefers the highest level of stress possible for a given outcome pair to any alternative level of stress for that outcome pair.*

Formally, this means that the CB strictly prefers $\{\min(H_1, H_2), p, p\}$ to all $\{s', p, p\}$ where $s' < \min(H_1, H_2)$, the CB strictly prefers $\{H_1, p, f\}$ to all $\{s'', p, f\}$ where $s'' < H_1$, the CB strictly prefers $\{H_2, f, p\}$ to all $\{s''', f, p\}$ where $s''' < H_2$.

Proof: Since $\mathcal{M}$ is strictly increasing in μ_i and by monotonicity μ_i is increasing in s for a given outcome pair, this Lemma holds. *Q.E.D.*

Lemma B.7 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$. We proceed by showing that the sets of health state pairs (H_1, H_2) where a given $\{s, o_1, o_2\}$ is sent can be expressed via indifference frontiers $x(H_2)$ and $y(H_2)$ as in Theorem 1.

Lemma B.8 *If, at a health state pair (H_1, H_2) with $H_2 > 0$, sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1, H_2') with $H_2' < H_2$ it is mandate maximising to send $\{H_1, p, f\}$.*

Proof: At (H_1, H_2), $\{H_1, p, f\}$ was revealed preferred to $\{H_2, p, p\}$, i.e. $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. Given beliefs satisfy monotonicity, $\{H_2, p, p\}$ results in higher beliefs μ_i about both banks than $\{H_2', p, p\}$ and thus $V(\{H_2, p, p\}) > V(\{H_2', p, p\})$. Hence, $\{H_1, p, f\}$ must be preferred to $\{H_2', p, p\}$, i.e. $V(\{H_1, p, f\}) > V(\{H_2', p, p\})$. *Q.E.D.*

Lemma B.9 *If at a health state pair (H_1, H_2) sending $\{H_1, p, f\}$ is the message which maximises the CB's mandate $\mathcal{M}$, then for all (H_1', H_2) with $H_1' > H_1$ it is mandate maximising to send $\{H_1', p, f\}$.*

Proof: (Analogous to the proof of Lemma B.8) Revealed: $V(\{H_1, p, f\}) > V(\{H_2, p, p\})$. For beliefs which satisfy monotonicity, $\{H_1', p, f\}$ results in higher μ_i for both banks than $\{H_1, p, f\}$. Thus, $V(\{H_1', p, f\}) > V(\{H_1, p, f\})$. Hence, $V(\{H_1', p, f\}) > V(\{H_2, p, p\})$. *Q.E.D.*

Lemma B.10 *There never exists a perfectly vertical part of the $x(H_2)$ indifference frontier.*[50]

[50] The same logic applies to horizontal parts of $y(H_2)$.

Proof by contradiction: Suppose a perfectly vertical part of the $x(H_2)$ indifference frontier existed. Then, on this part, there would be at least two health state pairs, (H_1^b, H_2) and (H_1^a, H_2) with $H_1^a > H_1^b$, at which the sender is indifferent between $\{H_2, p, p\}$ respectively $\{H_1^a, p, f\}$ or $\{H_1^b, p, f\}$. Thus $V(\{H_1^b, p, f\}) = V(\{H_2, p, p\})$ and $V(\{H_1^a, p, f\}) = V(\{H_2, p, p\})$ must both hold. Therefore, we must have $V(\{H_1^a, p, f\}) = V(\{H_1^b, p, f\})$. But this conflicts with beliefs satisfying monotonicity. For beliefs which satisfy monotonicity $\{s, p, f\}$ results in higher μ_i than $\{s', p, f\}$ $\forall s > s'$ and $\forall \mu_i$. Thus, since $\frac{\partial V(\mu_1, \mu_2)}{\partial \mu_i} > 0 \forall i$, monotonicity of beliefs ensures that $V(\{H_1^a, p, f\}) > V(\{H_1^b, p, f\})$ and therefore no perfectly vertical part of the $x(H_2)$ indifference frontier can exist. *Q.E.D.*

Lemma B.7 means that the CB chooses either $s = \max(H_1, H_2)$ or $s = \min(H_1, H_2)$. Lemma B.8 - B.9 showed that the health state pairs at which the CB sends $\{H_1, p, f\}$ form a connected set in the top left part of the state space $(H_1, H_2) \in [0, 1]^2$. Lemma B.10 showed that this connected set does not have a perfectly vertical boundary. The boundary also cannot be decreasing. Thus, when beliefs satisfy monotonicity, the CB's optimal strategy is unique and always identical to the strategy specified in Theorem 1. *Q.E.D.*

C Proof of Propositions

C.1 Proof of Proposition 1: Ordering Designs by Reassurance

Lemma C.1 *The Zero Disclosure Framework generates* $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[H]$.

Proof: Since in ZDF never any information is released, recipients can never update their beliefs from their prior $\mathbb{E}[H_i]$. Thus, beliefs are always $\mu_i = \mathbb{E}[H_i]$ and therefore $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(\mathbb{E}(H), \mathbb{E}(H))] = \mathbb{E}[H]$. *Q.E.D.*

Lemma C.2 *The Full Disclosure Framework generates* $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)]$.

Proof: Since in FDF health realisations H_i are always fully revealed, recipients update their belief to the true fundamental: $\mu_i = H_i$. Thus, the entire distribution of fundamentals and beliefs coincide, not just the mean (Bayes plausibility). Hence, $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)]$. *Q.E.D.*

ZDF always generates more reassurance than FDF as $\mathbb{E}[H] \geq \mathbb{E}[\min(H_1, H_2)]$ holds for all distributions.[51]

Lemma C.3

(i) For all mandates $\mathcal{M}$, the Hurdle Rate Framework generates a level of reassurance that lies between the levels of FDF and ZDF.

[51] When bank health is drawn from a uniform distribution, $\mathbb{E}[H] = \frac{1}{2} > \frac{1}{3} = \mathbb{E}[\min(H_1, H_2)]$.

(ii) *The level of reassurance in a HRF is monotonically increasing in the weight the mandate places on the weak bank.*

Proof: We know from (5) that the $x(H_2)$ indifference frontier is $x(H_2) = \frac{2-\omega}{1-2\omega} H_2$ when $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1-\omega) \ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0,1)$ and $H_i \sim U[0,1] \ \forall \ i$. For ease of exposition, we define $m \equiv \frac{2-\omega}{1-2\omega}$. Since $\omega \geq 0$, our parameter space is $m \geq 2$. There is a one-for-one mapping between m and ω, $\omega = \frac{m-2}{2m-1}$. Using the resulting $x(H_2) = mH_2$ in (8), reassurance in a HRF is:

$$\begin{aligned}\mathbb{E}[\min(\mu_1, \mu_2)] = &\int_0^{\frac{1}{m}} \Big[\frac{H_2^2}{2m^2} + \frac{3+m}{8}(\frac{m^2-1}{m^2})H_2^2 + \frac{3+m}{4}(m-1)H_2^2 + \frac{1}{4m} - \frac{m}{4}H_2^2\Big] g(H_2)dH_2 \\ &+ \int_{\frac{1}{m}}^1 \Big[\frac{H_2^2}{2m^2} + \frac{3}{8}(\frac{m^2-1}{m^2})H_2^2 + \frac{1}{4}(\frac{m-1}{m})H_2 + \frac{1}{4} - \frac{3}{4}H_2^2 + \frac{1}{2}H_2\Big] g(H_2)dH_2\end{aligned} \tag{13}$$

which simplifies to:

$$\mathbb{E}[\min(\mu_1, \mu_2)] = \frac{1}{2} + \frac{-7m^3 + m^2 + 3m - 1}{24m^4} \tag{14}$$

From (14), we establish three sublemmata which together prove Lemma C.3.

Sublemma C.1 *Even when the sender's mandate is concerned only with average bank health, a hurdle rate framework implies more reassurance than the full disclosure framework.*

Proof: Using that the mandate is concerned only with average bank health ($\omega = 0$ or equally $m = 2$) in (14), we obtain $\mathbb{E}[\min(\mu_1, \mu_2)] = \frac{145}{384} \approx 0.378$. In FDF, using Lemma C.2 and the distributional assumption, $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)] = \frac{1}{3}$, which is below reassurance in a HRF. *Q.E.D.*

Sublemma C.2 *The stronger the weight the mandate places on the weak bank, the more reassurance is provided in equilibrium.*

Proof: We aim to show that $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{d\omega} \geq 0 \ \forall \ \omega \in [0,1)$. This corresponds to showing that $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{dm} \geq 0 \ \forall \ m \geq 2$. From (14) we obtain:

$$\frac{d\mathbb{E}[\min(\mu_1, \mu_2)]}{dm} = \frac{7m^3 - 2m^2 - 9m + 4}{24m^5} \tag{15}$$

and

$$\frac{d^2\mathbb{E}[\min(\mu_1, \mu_2)]}{dm^2} = \frac{-7m^3 + 3m^2 + 18m - 10}{12m^6} \tag{16}$$

Solving $\max_m \mathbb{E}[\min(\mu_1, \mu_2)]$ yields the FOC $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{dm} \stackrel{!}{=} 0$ which is satisfied for $m = 1$ and for two lower values of m. The SOC confirms that $m = 1$ is indeed a minimum. Thus, for all $m > 1$ we have $\frac{d\mathbb{E}[\min(\mu_1,\mu_2)]}{dm} > 0$. Hence, reassurance is monotonically increasing in m for all $m \geq 2$. *Q.E.D.*

Sublemma C.3 *As the weight the mandate places on the weak bank increases, the amount of reassurance provided by a HRF approaches, but never exceeds, that of ZDF.*

Proof: Higher ω result in higher m. In the extreme case ($m \to \infty$), the CB almost always sends $\{\min(H_1, H_2), p, p\}$. Only if $\min(H_1, H_2) = 0$, does the CB let the weak bank fail the test. To characterise reassurance in the limit case, apply limit theorems to (14):

$$\lim_{m\to\infty} \mathbb{E}[\min(\mu_1, \mu_2)] = \lim_{m\to\infty}(\frac{1}{2}) + \lim_{m\to\infty}(\frac{-7m^3 + m^2 + 3m - 1}{24m^4}) = \frac{1}{2} \tag{17}$$

From Lemma C.1 we know that for ZDF $\mathbb{E}[\min(\mu_1, \mu_2)] = \mathbb{E}[\min(H_1, H_2)] = \frac{1}{2}$. Thus, $\mathbb{E}[\min(\mu_1, \mu_2) \mid HRF, m \to \infty] = \mathbb{E}[\min(\mu_1, \mu_2) \mid ZDF]$. *Q.E.D.*

Note that, given our definition of reassurance, this limit result focuses on the first moment of the distribution of $\min(\mu_1, \mu_2)$ and finds that the first moment in the extreme HRF equals that in ZDF. However, higher moments differ. In ZDF, minimum beliefs have a degenerate distribution ($\min(\mu_1, \mu_2) = \mathbb{E}[H_i]$), while in the extreme HRF the distribution of $\min(\mu_1, \mu_2)$ has full support on the unit interval. Thus, under a broader definition of reassurance which includes higher moments we expect that the extreme HRF provides strictly less reassurance than ZDF. As both the ranking of frameworks in terms of reassurance and the ranking of mandates within HRF remain unaffected, our results on optimal institution design would remain valid.

Proof of Lemma C.3: Combining Sublemmata C.1, C.2, C.3 proves Lemma C.3. *Q.E.D.*

Proof of Proposition 1: Combining Lemmata C.1, C.2, C.3 proves Proposition 1. *Q.E.D.*

C.2 Proof of Proposition 2: Comparative Static on the Sender's Mandate

Let there be two mandates $V(\mu_1, \mu_2)$ and $U(\mu_1, \mu_2)$ where $U(\mu_1, \mu_2)$ is more concerned with the weak bank. Let the equilibrium in the communication game for $V(\mu_1, \mu_2)$ be characterised by indifference frontier $x(H_2)$ and the interpretation rule R while the equilibrium for $U(\mu_1, \mu_2)$ is characterised by $\tilde{x}(H_2)$ and $\tilde{R}$. On $x(H_2)$ we have that $V(\{H_1, p, f\}) = V(\{H_2, p, p\})$. Given the interpretation rule R, $\{H_2, p, p\}$ results in $\mu_1 = \mu_2$ while $\{H_1, p, f\}$ results in dispersed posteriors $\mu_1' \neq \mu_2'$. Moreover, $\mu_2' < \mu_1 = \mu_2 < \mu_1'$. Therefore, a CB with mandate $U(\mu_1, \mu_2)$, which places more weight on the weak bank than $V(\mu_1, \mu_2)$, strictly prefers $\{H_2, p, p\}$ on $x(H_2)$ for interpretation R. For interpretation rule R, $U(\mu_1, \mu_2)$ strictly prefers $\{H_2, p, p\}$ to $\{H_1, p, f\}$ for all (H_1, H_2) between the 45-degree line and the $x(H_2)$-frontier. Thus, the hypothetical frontier $\bar{x}(H_2)$, at which a CB with $U(\mu_1, \mu_2)$ is indifferent between $\{H_1, p, f\}$ and $\{H_2, p, p\}$ given R, must lie strictly above $x(H_2)$. However, $\bar{x}(H_2)$ does not characterise an equilibrium as it is based on R rather than the appropriate $\tilde{R}$. Allowing the interpretation rule to adjust to $\tilde{R}$ results in $U(\mu_1, \mu_2)$ strictly preferring $\{H_2, p, p\}$ also on $\bar{x}(H_2)$. Thus, the resulting equilibrium indifference frontier $\tilde{x}(H_2)$ must lie above $\bar{x}(H_2)$. Since $\tilde{x}(H_2) > \bar{x}(H_2) \ \forall \ H_2$ and $\bar{x}(H_2) > x(H_2) \ \forall \ H_2$, we have that $\tilde{x}(H_2) > x(H_2) \ \forall \ H_2$. *Q.E.D.*

C.3 Proof of Proposition 3: Ordering Designs by Incentives

Recall that low effort leads to bank health being drawn from CDF $G(H)$ and high effort from $F(H)$.

Lemma C.4 *The Zero Disclosure Framework generates no incentives.*

Proof: Since in ZDF never any information is released, recipients can never update their prior on bank health. This means that the expected belief is fully determined by conjectured effort and unaffected by actual effort. Formally $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}\big[\mu_1(\underline{e}) \mid \underline{e}_1, \underline{e}_2\big] = \mathbb{E}_G[H]$. *Q.E.D.*

Lemma C.5 *The Full Disclosure Framework generates full incentives.*

Proof: Since in FDF health realisations H_i are always fully revealed, recipients update their belief to the true fundamental: $\mu_i = H_i$. Hence, $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}_F[H]$. Thus, incentives are $\mathbb{E}_F[H] - \mathbb{E}_G[H]$ which is strictly positive by the MLRP ordering. *Q.E.D.*

Lemma C.6

(i) For all mandates $\mathcal{M}$, the Hurdle Rate Framework generates incentives. These incentives have a strength between that of ZDF and HRF.

(ii) The strength of incentives in a HRF is monotonically decreasing in the weight the mandate places on the weak bank.

Proof: We know from (5) that the $x(H_2)$ indifference frontier is $x(H_2) = \frac{2-\omega}{1-2\omega} H_2$ when $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1-\omega) \ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0, 1)$, $G(H) = H$, $F(H) = H^2$, and the CB conjectures that banks choose low effort. Analogous to Appendix C.1, using the resulting $x(H_2) = mH_2$ in (10), we have that for an HRF:

$$\begin{aligned}
\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = &\int_0^{\frac{1}{m}} \Bigg[\frac{H_2}{2m} F(\frac{H_2}{m}) + \frac{3+m}{4} \int_{\frac{H_2}{m}}^{H_2} H_1 f(H_1) dH_1 \\
&+ \frac{3+m}{4} H_2 [F(mH_2) - F(H_2)] + \int_{x(H_2)}^{1} H_1 f(H_1) dH_1 \Bigg] g(H_2) dH_2 \\
+ &\int_{\frac{1}{m}}^{1} \Bigg[\frac{H_2}{2m} F(\frac{H_2}{m}) + \int_{\frac{H_2}{m}}^{H_2} \frac{1}{4} + \frac{3}{4} H_1 f(H_1) dH_1 \\
&+ (\frac{1}{4} + \frac{3}{4} H_2)\big(1 - F(H_2)\big) \Bigg] g(H_2) dH_2
\end{aligned} \tag{18}$$

and

$$\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \frac{1}{48m^6} \quad \big(27m^6 + 15m^5 - 13m^4 - m^3 + 4m - 2\big) \tag{19}$$

From (19), we establish three sublemmata which together prove Lemma C.6.

Sublemma C.4 *When the mandate is concerned only with average bank health, a HRF generates incentives with a strength strictly between that of ZDF and FDF.*

Proof: Using that $\mathcal{M}$ is only concerned with average bank health ($m = 2$) in (19), we obtain $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \frac{333}{512}$. In FDF, using Lemma C.5 and the distributional assumptions, $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}_F[H] = \frac{2}{3}$ and in ZDF, using Lemma C.4, $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}_G[H] = \frac{1}{2}$. Hence, $\mathbb{E}_G[H] < \mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2, HRF\big] < \mathbb{E}_F[H]$ and thus effort incentives in a HRF are strictly positive (more than in ZDF) but less than in FDF. *Q.E.D.*

Sublemma C.5 *Incentives are monotonically decreasing in the weight the mandate places on the weak bank.*

Proof: We aim to show that $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] - \mathbb{E}_G[H]$ is decreasing in $\omega \in [0, 1)$. Since $\mathbb{E}_G[H]$ is independent of $\mathcal{M}$, this is equivalent to showing that $\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big]$ is decreasing in m for all $m \geq 2$. From (19) we obtain:

$$\frac{d\mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big]}{dm} = \frac{-15m^5 + 26m^4 + 3m^3 - 20m + 12}{48m^7} \tag{20}$$

For our entire parameter space ($m \geq 2$), this expression is negative. *Q.E.D.*

Sublemma C.6 *For all $\mathcal{M}$, a Hurdle Rate Framework creates strictly positive incentives. Even when the CB is purely concerned with the weak bank, incentives remain strictly positive.*

Proof: If the mandate is purely concerned with the weak bank ($\omega = 1$), then $m \to \infty$. The limit of (19) becomes $\lim_{m\to\infty} \mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] = \frac{9}{16}$. Thus, $\lim_{m\to\infty} \mathbb{E}\big[\mu_1(\underline{e}) \mid \bar{e}_1, \underline{e}_2\big] - E_G[H] > 0$ which means that incentives are strictly positive. Since by Sublemma C.5 incentives are monotonically decreasing in ω, all $\omega \in [0, 1)$ must have even higher and thus strictly positive incentives. *Q.E.D.*

Recall that when the CB is purely concerned with the weak bank, then in equilibrium the CB almost always sends $\{\min(H_1, H_2), p, p\}$. The sole exception is when $\min(H_1, H_2) = 0$ as then $s = \max(H_1, H_2)$ is chosen. This messaging strategy differs from zero disclosure as it reveals some information ($\min(H_1, H_2)$). Thus, the messaging strategy generates strictly positive incentives. Bank i anticipates that the CB chooses $s = \min(H_1, H_2)$ and knows that in some states $\min(H_1, H_2) = H_i$. Higher effort by bank i makes higher H_i more likely which makes more favourable messages more likely and thus increases expected beliefs. Thus, strictly positive incentives exist.

Proof of Lemma C.6: Combining Sublemma C.4, C.5, C.6 proves Lemma C.6. *Q.E.D.*

Proof of Proposition 3: Combining Lemma C.4, C.5, C.6 proves Proposition 3. *Q.E.D.*

C.4 Proof of Proposition 4: Equilibrium in the Endogenous Effort Game

Proof of Part (i): Since in ZDF never any information is released, receivers can never update their prior on bank health. This means that the expected belief is fully determined by conjectured effort and unaffected by actual effort, i.e. $\mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2\big] = \mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \underline{e}_1, \underline{e}_2\big] = E_G[H]$. When banks choose effort e_i to maximise $\mathbb{E}[\pi_i] = \mathbb{E}[\mu_i] - C(e_i)$, they anticipate no benefit of high effort in terms of $\mathbb{E}[\mu_i]$ and know that there is a strictly positive cost. Thus, banks choose low effort independent of the receivers' conjecture. Thus, there exists a unique equilibrium. In this equilibrium banks choose low effort and receivers conjecture low effort. *Q.E.D.*

Proof of Part (ii): In FDF, health realisations H_i are always fully revealed and recipients thus update their belief to the true fundamental: $\mu_i = H_i$. Therefore, expected beliefs $\mathbb{E}[\mu_i]$ and expected fundamental health $\mathbb{E}[H_i \mid e_i]$ coincide for any effort choice and are unaffected by conjectured effort. When banks choose effort e_i to maximise $\mathbb{E}[\pi_i] = \mathbb{E}[\mu_i] - C(e_i)$, they anticipate a benefit of high effort in terms of $\mathbb{E}[\mu_i]$ that equals the effect of effort on fundamental health ($\mathbb{E}_F[H] - \mathbb{E}_G[H]$) and weigh it against the strictly positive cost (c). Thus, banks exert high effort iff $\mathbb{E}_F[H] - \mathbb{E}_G[H] > c$, which holds by the assumption that effort is socially beneficial. Thus, banks choose high effort independent of the receivers' conjecture. Hence, there exists a unique equilibrium. In this equilibrium banks choose high effort and receivers conjecture high effort. *Q.E.D.*

Proof of Part (iii):

Lemma C.7 *A low effort equilibrium exists if and only if* $c \geq \bar{t}(\mathcal{M})$

Proof: In a HRF, an equilibrium where both banks choose low effort exists only if a bank does not benefit from unilaterally deviating to high effort, i.e. $\mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2\big] - \mathbb{E}_G[H] \leq c$. We know from Proposition 3 that the strength of incentives generated in a HRF depends on the CB's mandate. Defining $\mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2\big] - \mathbb{E}_G[H] = \bar{t}(\mathcal{M})$, the condition above can be restated as $c \geq \bar{t}(\mathcal{M})$. All other equilibrium conditions are also satisfied. *Q.E.D.*

Lemma C.8 *A high effort equilibrium exists if and only if* $c \leq \underline{t}(\mathcal{M})$.

Proof: In a HRF, an equilibrium where both banks choose high effort exists only if a bank does not benefit from unilaterally deviating to low effort, i.e. $\mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\bar{\boldsymbol{e}}) \mid \underline{e}_1, \bar{e}_2\big] \geq c$. The strength of incentives generated in a HRF depends on the CB's mandate. Defining $\underline{t}(\mathcal{M}) = \mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\bar{\boldsymbol{e}}) \mid \underline{e}_1, \bar{e}_2\big]$, the condition can be restated as $c \leq \underline{t}(\mathcal{M})$. All other equilibrium conditions are also satisfied. *Q.E.D.*

Lemma C.9 $\underline{t}(\mathcal{M}) < \bar{t}(\mathcal{M})$

When both banks play high effort which results in $F(H) = H^2$, then the hurdle rate communication game for $V(\mu_1, \mu_2) = \frac{\mu_1 + \mu_2}{2}$ results in $x(H_2) = \sqrt{3}\ H_2$. In the endogenous effort game, there exists an equilibrium where both banks play high effort if and only if $c \leq \mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\bar{\boldsymbol{e}}) \mid \underline{e}_1, \bar{e}_2\big] \approx 0.1285$. Correspondingly, both banks playing low effort results in $x(H_2) = 2H_2$ and is an equilibrium if and only if: $c \geq \mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2\big] - \mathbb{E}_G[H] = \frac{77}{512} \approx 0.1504$. Since $0.1285 < 0.1504$, Lemma C.9 holds. *Q.E.D.*

Proof of Part (iii): Since the only candidate symmetric pure strategy equilibria have $(\underline{e}_1, \underline{e}_2)$ or $(\bar{e}_1, \bar{e}_2)$, combining Lemma C.7, C.8, C.9 proves Part (iii). Lemma C.7 and C.8 show regions in which $(\underline{e}_1, \underline{e}_2)$ or $(\bar{e}_1, \bar{e}_2)$ equilibria exist and Lemma C.9 shows that the regions of the cost space in which these equilibria exist are disjoint, leading to a region where no pure stretegy symmetric equilibrium exists between $\underline{t}(\mathcal{M})$ and $\bar{t}(\mathcal{M})$. *Q.E.D.*

C.5 Proof of Proposition 5: The Optimal Mandate

Lemma C.10 *If there exists a $\{HRF, \mathcal{M}\}$ which incentivises high effort, then $\{HRF, \tilde{\mathcal{M}}\}$, where $\tilde{\mathcal{M}}$ is the softest possible mandate which just incentivises high effort, achieves higher welfare than all tougher HRF mandates and than FDF.*

Formally, let $\tilde{\omega}$ be the softest possible mandate which in a HRF results in high effort, i.e. $\tilde{\omega}$ is defined by $c = \mathbb{E}_F[H] - \mathbb{E}\big[\mu_1(\bar{\boldsymbol{e}}) \mid \underline{e}_1, \bar{e}_2, \tilde{\omega}\big]$. Then all tougher mandates $\omega < \tilde{\omega}$ also result in high effort. Thus, Lemma C.10 states that

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1, \mu_2)] > \mathbb{E}_{HRF,\omega}[W(\mu_1, \mu_2)] \; \forall \; \omega < \tilde{\omega} \tag{21}$$

and

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1, \mu_2)] > \mathbb{E}_{FDF}[W(\mu_1, \mu_2)] \tag{22}$$

Proof: FDF, {HRF, $\tilde{\omega}$}, and all {HRF, ω} with $\omega < \tilde{\omega}$ result in high effort. Therefore, for $W(\mu_1, \mu_2) = \lambda \; \min(\mu_1, \mu_2) + (1 - \lambda) \; \frac{\mu_1 + \mu_2}{2}$ with any $\lambda > 0$ the welfare ranking of these designs reduces to a ranking in terms of reassurance, $\mathbb{E}[\min(\mu_1, \mu_2)]$. From Proposition 1 we know that FDF provides the lowest reassurance and that in a HRF, reassurance is monotonically increasing in the softness of the mandate (ω). Thus, equations (21) and (22) hold. *Q.E.D.*

Lemma C.11 *ZDF achieves higher welfare than all hurdle rate frameworks which result in low effort.*

Formally, let $\bar{\omega}$ be the toughest possible mandate which in a HRF results in low effort, i.e. $\bar{\omega}$ is defined by $c = \mathbb{E}\big[\mu_1(\underline{\boldsymbol{e}}) \mid \bar{e}_1, \underline{e}_2, \bar{\omega}\big] - \mathbb{E}_G[H]$. Then, all softer mandates $\omega > \bar{\omega}$ also result in low effort. Thus, Lemma C.11 states that

$$\mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] > \mathbb{E}_{HRF,\omega}[W(\mu_1, \mu_2)] \; \forall \; \omega > \bar{\omega} \tag{23}$$

Proof: ZDF and all $\{HRF, \omega\}$ with $\omega > \bar{\omega}$ result in low effort. Therefore, for $W(\mu_1, \mu_2) = \lambda \; \min(\mu_1, \mu_2) + (1 - \lambda) \; \frac{\mu_1 + \mu_2}{2}$ with any $\lambda > 0$ the welfare ranking of these designs reduces to a ranking in terms of reassurance, $\mathbb{E}[\min(\mu_1, \mu_2)]$. From Proposition 1 we know that ZDF is the design which provides the highest reassurance among all designs. Thus, equation (23) holds.

Q.E.D.

Proof of Proposition 5: Combining Lemma C.10 and Lemma C.11 proves Proposition 5. *Q.E.D.*

C.6 Proof of Proposition 6: Central Bank Design

Proof of Part (ii): From Proposition 5 we have that, in this case of high effort costs, $W(\{\text{ZDF}\}) > W(\{\text{HRF}, \mathcal{M}\}) \; \forall \; \mathcal{M}$ and thus the planner's choice reduces to a choice among ZDF and FDF. We also established that ZDF results in a low effort equilibrium. Thus, if the planner chooses ZDF:

$$\mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] = \mathbb{E}_G[H] \; \forall \; \lambda \tag{24}$$

We also established that FDF results in a high effort equilibrium. Thus, if the planner chooses FDF:

$$\mathbb{E}_{FDF}[W(\mu_1, \mu_2)] = \lambda \; \mathbb{E}_F[\min(H_1, H_2)] - (1 - \lambda) \; \mathbb{E}_F[H] \tag{25}$$

The planner chooses FDF if and only if

$$\mathbb{E}_{FDF}[W(\mu_1, \mu_2)] > \mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] \tag{26}$$

Let λ^* denote the value of λ at which (26) holds with equality. Then, the equality can be written as

$$\lambda^* \; \mathbb{E}_F[\min(H_1, H_2)] - (1 - \lambda^*) \; \mathbb{E}_F[H] = \mathbb{E}_G[H] \tag{27}$$

Which yields:

$$\lambda^\star(a) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)]} \tag{28}$$

Thus, for a given benefit of effort a, at $\lambda < \lambda^\star(a)$ the planner chooses FDF and otherwise ZDF.

For the full characterisation of the optimal central bank design as depicted in Figure 4a note that since $\mathbb{E}_F[H] - \mathbb{E}_G[H] > 0$ and $\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1, H_2)] > 0$ always hold, it must be that $\lambda^\star(a) > 0$. Moreover, if the benefit of effort is so large that $\mathbb{E}_F[\min(H_1, H_2)] > \mathbb{E}_G[H]$, then $W(\{FDF\}) > W(\{ZDF\}) \; \forall \; \lambda$ which gives rise to the top area in Figure 4a where FDF is chosen for all values of λ. For the functional forms $G(H) = H$, $F(H) = H^a$, $\mathbb{E}_F[\min(H_1, H_2)] > \mathbb{E}_G[H]$ holds for $a > 1.78$. Thus, we consider $\lambda^\star(a)$ only for $a \in [1, 1.78]$ where $a > 1$ is needed to ensure an MLRP ordering of distributions. Over this range, $\lambda^\star(a)$ is positive and monotonically increasing as $\lambda^\star(a) = -\frac{1}{2} - \frac{1}{2a} + a$. Thus, the optimal institutional design in the high effort case is characterised by Figure 4a. *Q.E.D.*

Proof of Part (i): From Proposition 5 we know that, in this case of low effort costs, $W(\{\text{HRF}, \tilde{\omega}\}) > W(\{\text{HRF}, \omega\}) \; \forall \; \omega < \tilde{\omega}$ and $W(\{\text{HRF}, \tilde{\omega}\}) > W(\{\text{FDF}\})$. We also know that $W(\{\text{ZDF}\}) > W(\{\text{HRF}, \omega\}) \; \forall \; \omega > \tilde{\omega}$. Thus, the planner's choice reduces to a choice among ZDF and $\{HRF, \tilde{\omega}\}$. As before, $\mathbb{E}_{ZDF}[W(\mu_1, \mu_2)] = \mathbb{E}_G[H] \; \forall \; \lambda$. By construction of $\tilde{\omega}$ as the softest possible mandate which induces high effort:

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1,\mu_2)] = \lambda\, \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1,\mu_2)] + (1-\lambda)\, \mathbb{E}_F[H] \tag{29}$$

The planner chooses $\{HRF, \tilde{\omega}\}$ if and only if:

$$\mathbb{E}_{HRF,\tilde{\omega}}[W(\mu_1,\mu_2)] > \mathbb{E}_{ZDF}[W(\mu_1,\mu_2)] \tag{30}$$

Let λ' denote the value of λ at which (30) holds with equality. Then, the equality can be written as

$$\lambda'\, \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1,\mu_2)] + (1-\lambda')\, \mathbb{E}_F[H] = \mathbb{E}_G[H] \tag{31}$$

Which yields:

$$\lambda'(a,C) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1,\mu_2)]} \tag{32}$$

Thus, for a given benefit of effort a, at $\lambda < \lambda'(a,C)$ the planner chooses $\{\text{HRF}, \tilde{\omega}\}$ and otherwise ZDF.

For the full characterisation of the optimal central bank design as depicted in Figure 4b note that $\lambda'(a,C) > 0$ always holds. Moreover, comparing $\lambda'(a,C)$ and $\lambda^*(a)$ we always have that $\lambda'(a,C) > \lambda^*(a)$, since this equals:

$$\frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1,\mu_2)]} > \lambda^\star(a) = \frac{\mathbb{E}_F[H] - \mathbb{E}_G[H]}{\mathbb{E}_F[H] - \mathbb{E}_F[\min(H_1,H_2)]} \tag{33}$$

which simplifies to:

$$\mathbb{E}_{HRF,\tilde{\omega}}[\min(\mu_1,\mu_2)] > \mathbb{E}_F[\min(H_1,H_2)] \tag{34}$$

which holds by the reassurance ranking. Thus, the optimal institutional design in the low effort case is characterised by Figure 4b. *Q.E.D.*

The value of $\lambda'(a,C)$ depends indirectly on the cost of effort. This arises because $\lambda'(a,C)$ depends directly on mandate $\tilde{\omega}$, that is the softest possible mandate which just induces high effort, which in turn depends on the cost of effort. If costs of effort are higher, a tougher mandate is needed to incentivise effort. This corresponds to a smaller value of $\tilde{\omega}$. Then, $\{\text{HRF}, \tilde{\omega}\}$ releases more information. $\lambda'(a,C)$ thus becomes steeper, i.e. rotates towards the $\lambda^*(a)$ line. When costs of effort become so high that they switch from the low cost of effort to the high cost of effort regime, then $\lambda'(a,C)$ becomes $\lambda^*(a)$ and the design chosen to induce effort switches from being $\{\text{HRF}, \tilde{\omega}\}$ to FDF. *Q.E.D.*

D Extensions and Robustness

This Appendix shows that our results are qualitatively unchanged for alternative functions forms of welfare and the CB's mandate.

D.1 Concave utility and CB communication

In the paper, we characterized the equilibrium of the verifiable disclosure game played in a HRF (Theorem 1) assuming that the CB's mandate $V(\mu_1, \mu_2)$ is strictly increasing in beliefs. We considered extensions where the CB may be particularly concerned with the weak bank as the risks of belief driven phenomena such as runs, higher funding costs, or problems to roll over short term debt are particularly acute for weak banks. While the main paper formalised this via concern for the weaker bank, i.e. via the ranking of banks, mathematically $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0, 1)$, this Appendix shows that results are qualitatively unchanged, in some cases even identical, when concern is about the level of banks, i. e. where $V(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ where $v'(\mu_i) > 0$ and $v''(\mu_i) \leq 0$. The concavity of $v(\mu_i)$ makes the CB risk averse and relates closely to the seminal paper on risk sharing by Hirshleifer (1971). Thus, our results do not depend on the functional form $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ but generalize to concave utility functions.

Let the CB mandate by $\tilde{V}(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ where $v'(\mu_i) > 0$ and $v''(\mu_i) \leq 0$. Then, the equilibrium takes the form described in Theorem 1 and depicted in Figure 2. The proof in Appendix B.1 also applies to $\tilde{V}(\mu_1, \mu_2)$, since the proof applies to any mandate $V(\mu_1, \mu_2)$ that is increasing in beliefs ($\frac{\partial V}{\partial \mu_i} > 0\, \forall\, i$) which is the case for $\tilde{V}(\mu_1, \mu_2)$.

Having established that the equilibrium resulting from $\tilde{V}(\mu_1, \mu_2)$ has the same form as the equilibrium resulting from $V(\mu_1, \mu_2)$, we now establish the stronger result that the equilibrium resulting from mandate $\tilde{V}(\mu_1, \mu_2)$ can be identical to the equilibrium resulting from mandate $V(\mu_1, \mu_2)$. This means that there exists a mandate $\tilde{V}(\mu_1, \mu_2)$ which results in exactly the same equilibrium as a mandate $V(\mu_1, \mu_2)$.

Proposition 8 *There exist cases where a CB with a mandate of the form*

$$\tilde{V}(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2) \text{ where } v'(\mu_i) > 0 \text{ and } v''(\mu_i) \leq 0$$

follows the same strategy as a CB with a mandate of the form

$$V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1-\omega)\ \tfrac{\mu_1+\mu_2}{2} \text{ where } \omega \in [0, 1).$$

Thus, the resulting equilibrium is identical.

Proof: Example 1: Let $\tilde{V}(\mu_1, \mu_2) = \sqrt{\mu_1} + \sqrt{\mu_2}$ and $H_i \sim U[0, 1]\, \forall\, i$. Conjecture that there exists an equilibrium of the form described in Theorem 1 and depicted in Figure 2. Then, on the $x(H_2)$-indifference frontier, the CB must be indifferent between sending $\{H_1, p, f\}$ and $\{H_2, p, p\}$. $\{H_1, p, f\}$ results in $\mu_1 \mid \{H_1, p, f\} = H_1$, and $\mu_2 \mid \{H_1, p, f\} = \mathbb{E}[H \mid H < x^{-1}(H_1)] = \frac{x^{-1}(H_1)}{2}$, while $\{H_2, p, p\}$ results in $\mu_1 \mid \{H_2, p, p\} = \mu_2 \mid \{H_2, p, p\} = \frac{1}{2}H_2 + \frac{1}{2}\,\mathbb{E}[H \mid H_2 < H < x(H_2)] = \frac{3}{4}H_2 + \frac{1}{4}x(H_2)$. On the indifference frontier we have $x(H_2) = H_1$ and respectively $x^{-1}(H_2) = H_2$. Thus, on the $x(H_2)$ indifference frontier we have that $\{H_1, p, f\}$ results in $\mu_1 \mid \{H_1, p, f\} = x(H_2)$ and $\mu_2 \mid \{H_1, p, f\} = \frac{H_2}{2}$.

For the point (H_1, H_2) to actually be on the $x(H_2)$ indifference frontier, the CB needs to be indifferent between sending $\{H_1, p, f\}$ and $\{H_2, p, p\}$, i.e.

$$\tilde{V}(\{H_1, p, f\}) = \tilde{V}(\{H_2, p, p\}) \tag{35}$$

which equals

$$v\Big(x(H_2)\Big) + v\Big(\frac{H_2}{2}\Big) = v\Big(\frac{3}{4}H_2 + \frac{1}{4}x(H_2)\Big) + v\Big(\frac{3}{4}H_2 + \frac{1}{4}x(H_2)\Big) \tag{36}$$

Using that $v(\mu_i) = \sqrt{\mu_i}$, (36) becomes

$$\sqrt{x(H_2)} + \sqrt{\frac{H_2}{2}} = 2\sqrt{\frac{3}{4}H_2 + \frac{1}{4}x(H_2)} \tag{37}$$

which solves to

$$x(H_2) = \frac{25}{8}\, H_2. \tag{38}$$

Thus, when $\tilde{V}(\mu_1, \mu_2) = \sqrt{\mu_1} + \sqrt{\mu_2}$ and $H_i \sim U[0,1]\,\forall\, i$, the CB's strategy is characterised by the $x(H_2)$ indifference frontier $x(H_2) = \frac{25}{8}\, H_2$.

Recall from equation (5) that when $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ and $H_i \sim U[0,1]\,\forall\, i$, then the $x(H_2)$ indifference frontier is

$$x(H_2) = \frac{2-\omega}{1-2\omega}H_2.$$

Since for $\omega = \frac{3}{14}$ we have that $\frac{2-\omega}{1-2\omega} = \frac{25}{8}$, the equilibrium resulting from $V(\mu_1, \mu_2) = \frac{3}{14}\ \min(\mu_1, \mu_2) + \frac{11}{14}\ \frac{\mu_1+\mu_2}{2}$ is identical to the equilibrium resulting from $\tilde{V}(\mu_1, \mu_2) = \sqrt{\mu_1} + \sqrt{\mu_2}$. *Q.E.D.*

Example 2: $\tilde{V}(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ where $v'(\mu_i) > 0$ and $v''(\mu_i) = 0$, or equally let $\tilde{V}(\mu_1, \mu_2) = \kappa(\mu_1 + \mu_2)$ where $\kappa > 0$. Then the preference ranking over belief pairs (μ_1, μ_2) described by $\tilde{V}(\mu_1, \mu_2)$ is the same for any $\kappa > 0$. Mathematically, $\tilde{V}(\mu_1, \mu_2)$ describes a preference ranking such that $(\mu_1, \mu_2) \succ (\mu_1', \mu_2')$, where $\succ$ denotes strict preference, is equivalent to $\kappa(\mu_1 + \mu_2) > \kappa(\mu_1' + \mu_2')$. This can be rearranged to

$$\kappa[(\mu_1 + \mu_2) - (\mu_1' + \mu_2')] > 0 \tag{39}$$

If (39) holds for any $\kappa > 0$, then it must hold for all other $\kappa > 0$.

Since for $\omega = 0$, $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ becomes $V(\mu_1, \mu_2) = \frac{\mu_1+\mu_2}{2}$ which in turn corresponds to $\tilde{V}(\mu_1, \mu_2) = \kappa(\mu_1 + \mu_2)$ with $\kappa = \frac{1}{2}$, we have that the preference ranking implies by $V(\mu_1, \mu_2)$ with $\omega = 0$ and the ranking implied by $\tilde{V}(\mu_1, \mu_2)$ with $v''(\mu_i) = 0$ are identical. Thus, the resulting equilibrium is also identical. *Q.E.D.*

The comparative static result of Proposition 2 also generalises, i.e. continues to hold when the CB mandate takes the form $\tilde{V}(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ where $v'(\mu_i) > 0$ and $v''(\mu_i) \leq 0$ rather than $V(\mu_1, \mu_2) = \omega\ \min(\mu_1, \mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0, 1)$.

Proposition 2 states that when the CB's mandate is more concerned with the weak bank, the set of health states where the CB sets $s = \min(\mu_1, \mu_2)$ is larger and the set where it sets $s = \max(\mu_1, \mu_2)$ is smaller. This corresponds to the $x(H_2)$ indifference frontier shifting to the left.

Comparing the two cases used to prove Proposition 8 proves that Proposition 2 continues to hold. When $\tilde{V}_a(\mu_1, \mu_2) = \sqrt{\mu_1} + \sqrt{\mu_2}$ the CB mandate is more concerned about the weak bank than when $\tilde{V}_b(\mu_1, \mu_2) = \mu_1 + \mu_2$. Formally, a mean preserving spread of beliefs does not reduce the CB's pay-off as specified by $\tilde{V}_b$, but it does reduce the pay-off as specified by $\tilde{V}_a$. Thus, while neither $\tilde{V}$ expression includes the min-operator familiar from V expressions, the curvature of $\tilde{V}$ and thus the presence of risk aversion creates a concern for the weak bank in the sense that a given increase in beliefs is more valuable in terms of $\tilde{V}$ if it increases the belief about the weaker bank rather than about the stronger bank.

Recall that for $H_i \sim U[0,1] \, \forall \, i$, $\tilde{V}_a(\mu_1, \mu_2) = \sqrt{\mu_1} + \sqrt{\mu_2}$ results in $x(H_2)$ indifference frontier $x(H_2) = \frac{25}{8} H_2$ and $\tilde{V}_b(\mu_1, \mu_2) = \mu_1 + \mu_2$ in $x(H_2) = 2\, H_2$.[52] Since $\frac{25}{8} > 2$, the $x(H_2)$ frontier in $\tilde{V}_a$ lies to the left of that in $\tilde{V}_b$. Thus, the set of health states where the CB sets $s = \min(\mu_1, \mu_2)$ is larger in $\tilde{V}_a$ than in $\tilde{V}_b$. Thus, Proposition 2 holds. *Q.E.D.*

D.2 Concave utility and central bank design

In the paper, we established that HRF provides more reassurance than FDF but less than ZDF. This Appendix shows that results are unchanged in a model analogous to the seminar work on risk sharing by Hirshleifer (1971), where the benefits of risk sharing arise from the concavity of utility functions. Thus, this Appendix highlights that risk aversion is the key feature of welfare which drives our results, not a particular functional form.

Let welfare be

$$W(\mu_1, \mu_2) = u(\mu_1) + u(\mu_2) \quad \textit{where} \quad u'(\mu_i) > 0 \textit{ and } u''(\mu_i) < 0 \tag{40}$$

The expected welfare maximising planner thus maximises

$$\mathbb{E}[W(\mu_1, \mu_2)] = \mathbb{E}[u(\mu_1)] + \mathbb{E}[u(\mu_2)] \tag{41}$$

As in section 4, we compare designs in terms of reassurance by assuming that effort is exogenously fixed and comparing designs in terms of $\mathbb{E}[W(\mu_1, \mu_2)]$. However, while in section 4 we assumed that $W(\mu_1, \mu_2) = \lambda \, \min(\mu_1, \mu_2) + (1-\lambda) \, \frac{\mu_1+\mu_2}{2}$ and thus reduced comparisons in terms of $\mathbb{E}[W(\mu_1, \mu_2)]$ to comparisons in terms of $\mathbb{E}[\min(\mu_1, \mu_2)]$, (40) implies that $\mathbb{E}[W(\mu_1, \mu_2)]$ reduces to the following:

Definition 6 *When welfare is risk averse as modelled by concave utility as described in equation (40), central bank design $\mathcal{D}$ is said to provide more reassurance than $\mathcal{D}'$ if, for given bank effort, $\mathbb{E}[u(\mu_1)] + \mathbb{E}[u(\mu_2)]$ is larger in $\mathcal{D}$ than in $\mathcal{D}'$.*

[52] The case of $\omega = 0$ was derived in the main paper and the equivalence of the equilibrium with $\omega = 0$ to the equilibrium resulting from $\tilde{V}_b$ was established above.

Applying this definition shows that HRF provides more reassurance than FDF but less than ZDF. Consider the analogous case to Appendix D.1, i.e. let $W(\mu_1, \mu_2) = \sqrt{\mu_1} + \sqrt{\mu_2}$ and $H_i \sim U[0,1] \forall i = 1,2$.

ZDF: Since in ZDF no information is disclosed, $\mu_i = \mathbb{E}[H_i]$ for every health outcome H_i. Thus $\mathbb{E}[W(\mu_1,\mu_2)] = \mathbb{E}[\sqrt{\mathbb{E}[H_i]}] + \mathbb{E}[\sqrt{\mathbb{E}[H_i]}] = 2\sqrt{\mathbb{E}[H_i]}$ and using that $H_i \sim U[0,1] \forall i$, we have that $\mathbb{E}[W(\mu_1,\mu_2)] = 2\,\frac{1}{\sqrt{2}} = \sqrt{2}$.

FDF: In FDF bank health is fully revealed ($\mu_i = H_i$) for every health outcome H_i. Thus, $\mathbb{E}[W(\mu_1,\mu_2)] = \mathbb{E}[\sqrt{H_i}] + \mathbb{E}[\sqrt{H_i}] = 2\ \mathbb{E}[\sqrt{H_i}]$ and using that $H_i \sim U[0,1] \forall i$, we have that $\mathbb{E}[W(\mu_1,\mu_2)] = \frac{4}{3}$

HRF: Assume that the CB's mandate is $V(\mu_1,\mu_2) = \mu_1 + \mu_2$. Then, the equilibrium is characterised by the $x(H_2)$ indifference frontier $x(H_2) = 2H_2$. As the equilibrium is characterised by thresholds and as the resulting beliefs are not continuous or differentiable functions of health H_i, $\mathbb{E}[\sqrt{\mu_i}] = \int_0^1 \int_0^1 \sqrt{\mu_1}\, f(H_1)\, dH_1\, g(H_2)\, dH_2$ is

$$
\begin{aligned}
\mathbb{E}[\sqrt{\mu_1}] = & \int_0^{x^{-1}(1)} \Bigg[\int_0^{y(H_2)} \sqrt{\mathbb{E}\Big[H \mid H < y(H_2)\Big]} f(H_1) dH_1 \\
& + \int_{y(H_2)}^{H_2} \sqrt{\frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}\Big[H \mid H_1 < H < x(H_1)\Big]} f(H_1) dH_1 \\
& + \int_{H_2}^{x(H_2)} \sqrt{\frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}\Big[H \mid H_2 < H < x(H_2)\Big]} f(H_1) dH_1 \\
& + \int_{x(H_2)}^{1} \sqrt{H_1} f(H_1) dH_1 \Bigg] g(H_2) dH_2 \\
+ & \int_{x^{-1}(1)}^{1} \Bigg[\int_0^{y(H_2)} \sqrt{\mathbb{E}\Big[H \mid H < y(H_2)\Big]} f(H_1) dH_1 \\
& + \int_{y(H_2)}^{H_2} \sqrt{\frac{1}{2}H_1 + \frac{1}{2}\mathbb{E}\Big[H \mid H_1 < H < 1\Big]} f(H_1) dH_1 \\
& + \int_{H_2}^{1} \sqrt{\frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}\Big[H \mid H_2 < H < 1\Big]} f(H_1) dH_1 \Bigg] g(H_2) dH_2
\end{aligned}
\tag{42}
$$

which simplifies to: $\mathbb{E}[\sqrt{\mu_i}] = \frac{337}{270} - \frac{23}{108}\sqrt{10} - \frac{1}{120}\sqrt{5} + \frac{49}{1080}\sqrt{7} \approx 0.676105$. Thus, $\mathbb{E}[W(\mu_1,\mu_2)] = \mathbb{E}[\sqrt{\mu_1}] + \mathbb{E}[\sqrt{\mu_2}] = 2 \times 0.676105 = 1.35221$.

Since $\sqrt{2} > 1.35221 > \frac{4}{3}$, we have that ZDF results in higher $\mathbb{E}[W(\mu_1,\mu_2)]$ and thus higher reassurance than HRF which in turn results in higher reassurance than FDF. *Q.E.D.*

Repeating the analysis for the more general CB mandate $V(\mu_1,\mu_2) = \omega\ \min(\mu_1,\mu_2) + (1-\omega)\ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0,1)$ and the resulting $x(H_2)$ indifference frontier shows that, for any mandate, reassurance provided by a HRF lies between reassurance in ZDF and FDF. Moreover, tougher mandates provide less reassurance. Thus, our CB design results are driven by welfare being risk averse and generalise to different functional forms of risk aversion.

An alternative generalisation is to consider $Var(\mu_i)$ as measure of reassurance. We find that

FDF results in higher $Var(\mu_i)$ than HRF which results in larger $Var(\mu_i)$ than ZDF. This further supports the statement that risk aversion of welfare is the key driver of our results.

The CB design results on incentives also generalise. In the paper, we assumed $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1 - \omega) \ \frac{\mu_1+\mu_2}{2}$ where $\omega \in [0, 1)$ and established that HRF generates stronger incentives than ZDF but weaker incentives than FDF. Appendix D.1 established that $\tilde{V}(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ where $v'(\mu_i) > 0$ and $v''(\mu_i) \leq 0$ results in an equilibrium which is qualitatively unchanged and can even be identical to that derived for $V(\mu_1, \mu_2)$. Recalling that the strength of incentives generated by a mandate is determined purely by the equilibrium of the communication game, it follows that the incentives generated by mandate $\tilde{V}(\mu_1, \mu_2)$ can be identical to those in the main paper generated by $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1 - \omega) \ \frac{\mu_1+\mu_2}{2}$. Hence, the central bank design results on reassurance and incentives both generalise.

Detailed discussion of functional forms: This Appendix highlighted that risk aversion is the key feature of welfare which drives our results and that results do not rely on a particular functional form. The main paper uses $W(\mu_1, \mu_2) = \lambda \ \min(\mu_1, \mu_2) + (1 - \lambda) \ \frac{\mu_1+\mu_2}{2}$ to ensure comparability with the mandate where the form $V(\mu_1, \mu_2) = \omega \ \min(\mu_1, \mu_2) + (1 - \omega) \ \frac{\mu_1+\mu_2}{2}$ is used to get closed form expressions with a parameter to vary the degree of concern for the weak bank. $W(\mu_1, \mu_2) = \lambda \ \min(\mu_1, \mu_2) + (1 - \lambda) \ \frac{\mu_1+\mu_2}{2}$ gives rise to comparisons of reassurance reducing to comparisons of $\mathbb{E}[\min(\mu_1, \mu_2)]$.

In this Appendix, we assume $\tilde{W}(\mu_1, \mu_2) = u(\mu_1) + u(\mu_2)$ where $u'(\mu_i) > 0$ and $u''(\mu_i) < 0$ to show that our CB design results do not rely on specific functional forms, e.g. the min operator, but hold also for other more general representations of risk aversion. This emphasizes that the key feature of welfare is that the expected welfare maximising planner is risk averse.

The functional forms W and $\tilde{W}$ have in common that welfare is strictly increasing in μ_i, but increases at a decreasing rate. Thus, the expected welfare maximising planner is risk-averse in both cases. With $\tilde{W}$, the increase is decreasing because of the level of μ_i. With W the increase is decreasing because of the ranking of μ_i, i.e. holding μ_j constant, an increase in μ_i has a larger effect on welfare when $\mu_i < \mu_j$ but a smaller effect when $\mu_i > \mu_j$. Mathematically, $\frac{\partial W}{\partial \mu_i} \geq \frac{\partial W}{\partial \mu_j} \ \forall \ \mu_i \leq \mu_j$.

D.3 Bank runs and CB communication

In the paper, we characterized the equilibrium of the verifiable disclosure game played in a HRF (Theorem 1) assuming that the CB's mandate $V(\mu_1, \mu_2)$ is strictly increasing in beliefs. This was motivated by referring to higher beliefs as reducing the risk of bank runs, mitigating problems to roll over short term debt, and reducing funding costs. This Appendix shows that results are qualitatively unchanged in a more explicit model of bank runs as in Goldstein and Leitner (2018). As a byproduct, this Appendix shows that our results only require $V(\mu_1, \mu_2)$ to be strictly increasing in beliefs and hold for both the case where $V(\mu_1, \mu_2)$ is continuous (in the paper) and where there are discontinuities as in Goldstein and Leitner (2018) (this Appendix).

Let the CB mandate be $V(\mu_1, \mu_2) = v(\mu_1) + v(\mu_2)$ and define $v(\mu_i)$ analogous to Goldstein and Leitner (2018) as:

$$v(\mu_i) = \begin{cases} \mu_i & \text{if} \quad \mu_i > \underline{\mu} \\ \mu_i - \Omega & \text{if} \quad \mu_i \leq \underline{\mu} \end{cases} \tag{43}$$

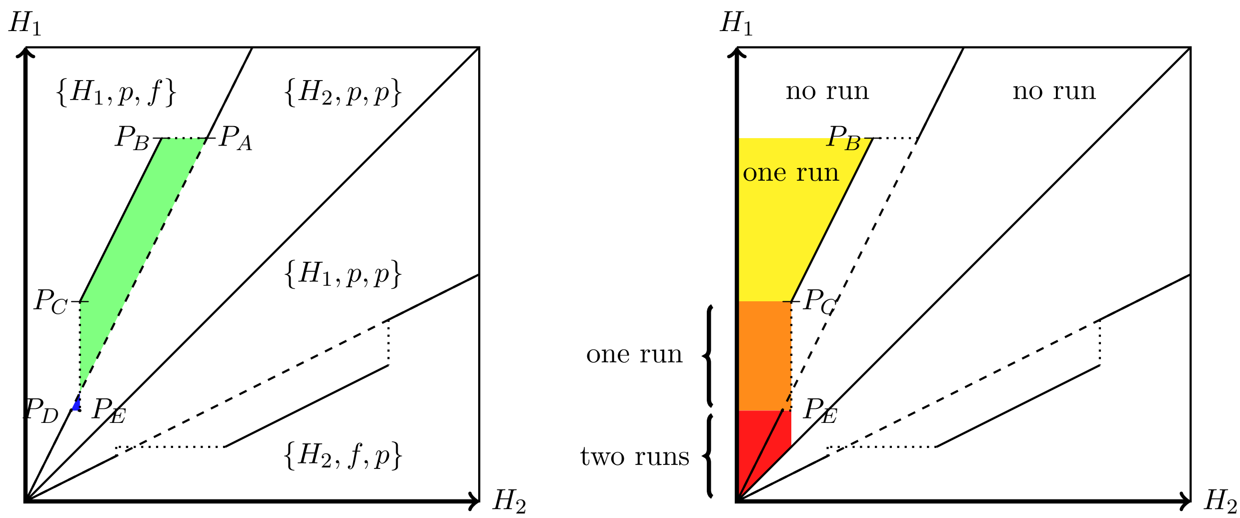
where $0 < \underline{\mu} < 1$ is the belief below which a bank run occurs and $\Omega > 0$ is the cost associated with a bank run.

Let $H_i \sim U[0,1] \forall i = 1, 2$, $\underline{\mu} = 0.2$, $\Omega = 0.1$. Then the CB's equilibrium strategy is as depicted in Figure 7a.

Figure 7: Equilibrium in a HRF when costs of bank runs exist

(a) Central Bank's Communication Strategy

(b) Bank runs



Proof: The proof in Appendix B.1 still applies also for the case here with bank run costs as the proof only uses that $V(\mu_1, \mu_2)$ is increasing in beliefs. The mandate $V(\mu_1, \mu_2)$ used here satisfies this assumption. The shape depicted in Figure 7 is derived in a numerical example further below.

Intuition: Intuitively, the equilibrium in Figure 7 is best understood in comparison to the equilibrium without bank runs. Following the messaging rule of the main paper, i.e. the rule derived without bank run costs (Ω), would, in the green area, result in one bank run, but the changed strategy prevents that. In this area, preventing the run makes $\{s, p, p\}$ more attractive and thus expands the set of health states where $\{s, p, p\}$ is sent. In the blue area, the messaging rule derived without considering bank run costs explicitly would result in two bank runs, but the changed strategy reduces this to just one bank run. In this blue area, reducing the number of runs makes $\{s, p, f\}$ more attractive and thus expands the set of health states where $\{s, p, f\}$ is sent. While the CB adjusts it behaviour to reduce the number of runs, it is not possible to prevent runs completely. The states where runs occur are depicted in Figure 7b.[53]

Comparison to main paper: The equilibrium depicted in Figure 7 is qualitatively unchanged from the equilibrium described in Theorem 1 in the sense that (i) the CB chooses either

[53]For a more detailed discussion of runs in a HRF and the reassurance implied by a HRF when there are runs, see Appendix D.4.

$s = \min(H_1, H_2)$ or $s = \max(H_1, H_2)$, (ii) that $s = \min(H_1, H_2)$ is chosen when H_1 and H_2 are similar and $s = \max(H_1, H_2)$ when they differ strongly, (iii) the set of health state pairs where a message is sent is a connected set, i.e. if at (H_1, H_2) the CB sends $\{H_1, p, f\}$, then at all (H_1, H_2') with $H_2' < H_2$ the CB also sends $\{H_1, p, f\}$.

Thus, the equilibrium is also unchanged in the sense that a HRF results in partial disclosure and that a HRF can be a source of information contagion.[54] Appendix D.4 shows that central bank design results are also qualitatively unchanged, i.e. a HRF provides intermediate levels of reassurance.

While the equilibrium depicted in Figure 7 partitions the state space into four areas where the CB sends $\{H_1, p, f\}$, $\{H_2, p, p\}$, $\{H_1, p, p\}$, or $\{H_2, f, p\}$ respectively, there is no representation via indifference frontiers as bank run costs introduce the possibility of jumps from one strict preference to the other. As a result, the frontier partitioning these four areas is not continuous. No closed form expressions as in the example in equation (5) exist. Moreover, while the CB in Theorem 1 was indifferent between messages on the frontier, the bank run costs (Ω) mean that the CB's preference can switch from strict preferences of e.g. $\{H_1, p, f\}$ to strict preferences of $\{H_2, p, p\}$ without being indifferent in between. Thus, in Figure 7, the frontier is partitioning, but the CB is indifferent only on solid parts while it has strict preferences on the dotted lines.

Numerical example

The frontier partitioning the areas where $\{s, p, f\}$ and $\{s, p, p\}$ are sent is the line composed of the lines connecting points $(H_2 = \frac{1}{2}, H_1 = 1)$ to P_A, P_A to P_B, P_B to P_C, P_C to P_E, P_E to P_D, and P_D to $(0, 0)$.

Define $\bar{x}(H_2)$ as the indifference frontier resulting from a CB who does not face runs, i.e. $V(\mu_1, \mu_2) = \mu_1 + \mu_2$.

Then, point P_A is the highest point on the $\bar{x}(H_2)$ frontier where a message would result in a run. Since on $\bar{x}(H_2)$ both $\{s, p, f\}$ and $\{s, p, p\}$ result in the same $\mu_1 + \mu_2$, but $\{s, p, f\}$ results in $\mu_1 \neq \mu_2$ while $\{s, p, p\}$ results in $\mu_1 = \mu_2$, the message which results in a run at P_A must be $\{H_1, p, f\}$. Setting $\mu_2 \mid \{H_1, p, f\} = \underline{\mu}$ yields: $P_A = (0.8, 0.4)$. Henceforth, we denote the H_1-coordinate of P_A as P_A^1 and the H_2 coordinate as P_A^2. Here $P_A^1 = 0.8$, $P_A^2 = 0.4$. While a CB with $V(\mu_1, \mu_2) = \mu_1 + \mu_2$ is indifferent between $\{H_1, p, f\}$ and $\{H_2, p, p\}$ at point P_A, the CB which faces runs has a strict preference. Since at P_A the CB could send $\{H_2, p, p\}$, which does not lead to a run, or $\{H_1, p, f\}$, which leads to the same $\mu_1 + \mu_2$ but triggers a run and thus causes loss Ω, the CB strictly prefers $\{H_2, p, p\}$ at point P_A.

At the points between P_B and P_C the CB is indifferent between $\{H_2, p, p\}$, which results in no run, and $\{H_1, p, f\}$, which results in one run (on bank B_2). Among these points, P_B has the highest H_2 and P_C the lowest H_2. Intuitively, P_B has $P_B^2 > 0$ since while $\{H_2, p, p\}$ avoids a run, it does so at the cost of lower beliefs $\mu_1 + \mu_2$. Thus, for P_A^1 there exists a H_2 at which the higher sum of beliefs $\mu_1 + \mu_2$ resulting from $\{H_1, p, f\}$ compensates the CB for the cost of a run (Ω). For $H_2 < P_B^2$, the CB strictly prefers $\{H_1, p, f\}$ to $\{H_2, p, p\}$.

Formally, $P_B = (P_B^1, P_B^2)$ is the point where the CB is indifferent between $\{P_B^2, p, p\}$, when it

[54] Information contagion arises for the same comparative statics outlined in section 7.

results in no bank run, and $\{P_A^1, p, f\}$ when it results in one bank run. Using that here $P_A^1 = 0.8$, we have: $\mu_1 \mid \{0.8, p, f\} = 0.8$, and $\mu_2 \mid \{0.8, p, f\} = \frac{P_B^2}{2}$. Thus, $V(\{0.8, p, f\}) = 0.8 + \frac{P_B^2}{2} - \Omega = 0.7 + \frac{P_B^2}{2}$. For $\{P_B^2, p, p\}$, $\mu_1 \mid \{P_B^2, p, p\} = \mu_2 \mid \{P_B^2, p, p\} = \frac{1}{2}P_B^2 + \frac{1}{2}\mathbb{E}[H \mid P_B^2 < H < 0.8] = \frac{3}{4}P_B^2 + 0.2$. Thus, $V(\{P_B^2, p, p\}) = \frac{3}{2}P_B^2 + 0.4$. Setting $V(\{0.8, p, f\}) = V(\{P_B^2, p, p\})$ yields $P_B^1 = 0.8$, $P_B^2 = 0.3$.

Formally, $P_C = (P_C^1, P_C^2)$ is the point with the lowest H_2 that still satisfies that the CB is indifferent between $\{P_C^1, p, f\}$ when it results in one bank run, and $\{P_C^2, p, p\}$ when it results in no run. Thus, $\mu_1 \mid \{P_C^1, p, f\} = P_C^1$ and $\mu_2 \mid \{P_C^1, p, f\} = \frac{P_C^2}{2}$. Thus $V(\{P_C^1, p, f\}) = \mu_1 + \mu_2 - \Omega = P_C^1 + \frac{P_C^2}{2} - 0.1$. And

$$\mu_1 \mid \{P_C^2, p, p\} = \mu_2 \mid \{P_C^2, p, p\} = \frac{1}{2}P_C^2 + \frac{1}{2}\frac{P_C^1 + P_C^2}{2} = \frac{3}{4}P_C^2 + \frac{1}{4}P_C^1 \tag{44}$$

Thus $V(\{P_C^2, p, p\}) = \mu_1 + \mu_2 = \frac{3}{2}P_C^2 + \frac{1}{2}P_C^1$. Setting $V(\{P_C^1, p, f\}) = V(\{P_C^2, p, p\})$ yields

$$P_C^1 = P_C^2 + 2\Omega \tag{45}$$

which characterises the indifference frontier between P_C and P_B.

P_C is the point which satisfies (45) and has the lowest possible H_2 such that $\{H_2, p, p\}$ does not result in a run on both banks. Thus, $\mu_i \mid \{P_C^2, p, p\} = \underline{\mu}$. Using (44) and (45) results in $\frac{3}{4}P_C^2 + \frac{1}{4}(2P_C^2 + 2\Omega) = \underline{\mu}$, which yields that P_C is $P_C^1 = \frac{11}{25}$, $P_C^2 = \frac{3}{25}$.

For health state pairs with low H_1 and low H_2, both $\{H_1, p, f\}$ and $\{H_2, p, p\}$ result in two runs. Since both messages result in the same number of runs, the indifference frontier is as in the case without runs. Mathematically, $\mu_1 \mid \{H_1, p, f\} = H_1$; $\mu_2 \mid \{H_1, p, f\} = \frac{H_2}{2}$. Thus, $V(\{H_1, p, f\}) = H_1 + \frac{H_2}{2} - 2\Omega$. $\mu_1 \mid \{H_2, p, p\} = \mu_2 \mid \{H_2, p, p\} = \frac{3}{4}H_2 + \frac{1}{4}H_1$. Thus $V(\{H_2, p, p\}) = \frac{3}{2}H_2 + \frac{1}{2}H_1 - 2\Omega$. On the indifference frontier $V(\{H_1, p, f\}) = V(\{H_2, p, p\})$, which reduces to $H_1 = 2H_2$ which coincides with the $\bar{x}(H_2)$ frontier.

Point P_D is the highest point on the $\bar{x}(H_2)$ frontier where both $\{H_1, p, f\}$ and $\{H_2, p, p\}$ result in a run on both banks. By arguments analogous to those made when characterizing the point P_A, P_D is the point where $\{H_2, p, p\}$ results in a run on both banks but where $\{H_1, p, f\}$ results in a run only on the weaker bank with the stronger bank just avoiding the run. Formally, at P_D we must have $\mu_1 \mid \{P_D^1, p, f\} = \underline{\mu}$. Using that $\mu_1 \mid \{H_1, p, f\} = H_1$, this yields that $P_D^1 = \underline{\mu} = 0.2$. Since P_D is on the $\bar{x}(H_2)$ frontier: $P_D^2 = \frac{1}{2}P_D^1$ and thus $P_D^2 = 0.1$. Hence, we have that $P_D^1 = 0.2$; $P_D^2 = 0.1$.

Point P_E is defined by $P_E^1 = P_D^1$ and $P_E^2 = P_C^2$. Thus, for these numbers P_E is $P_E^1 = 0.2$; $P_E^2 = \frac{3}{25}$. The frontier between P_C and P_E and between P_E and P_D partitions the state space into sets of health state pairs where $\{s, p, f\}$ is sent and where $\{s, p, p\}$ is sent. However, on the frontier the CB has a strict preference for one type of message (hence dotted lines are used to denote the difference from indifference frontiers which are denoted by solid lines).

On the line connecting P_C and P_E, the CB strictly prefers $\{H_1, p, f\}$ to $\{H_2, p, p\}$. To see this, note that $\mu_1 \mid \{H_1, p, f\} = H_1$ and for the health states on this line $H_1 \in [P_D^1, P_C^1]$. $\mu_2 \mid \{H_1, p, f\} = \frac{P_C^2}{2}$. Thus, $V(\{H_1, p, f\}) = H_1 + \frac{P_C^2}{2} - \Omega = H_1 - \frac{1}{25}$. $\mu_1 \mid \{H_2, p, p\} = \mu_2 \mid$

$\{H_2,p,p\} = \frac{1}{2}P_C^2 + \frac{1}{2}\mathbb{E}[H \mid P_C^2 < H < P_D^1] = \frac{3}{4}P_C^2 + \frac{1}{4}P_D^1 = \frac{7}{50}$. Thus, $V(\{H_2,p,p\}) = 2(\frac{3}{4}P_C^2 + \frac{1}{4}P_D^1) - 2\Omega = \frac{2}{25}$. The CB strictly prefers $\{H_1,p,f\}$ to $\{H_2,p,p\}$ iff $V(\{H_1,p,f\}) > V(\{H_2,p,p\})$ which is equivalent to $H_1 - \frac{1}{25} > \frac{2}{25} \forall H_1 \in [P_D^1, P_C^1]$. This equals $H_1 > \frac{3}{25}$ which holds since $P_D^1 = 0.2$ and $P_C^1 = \frac{11}{25}$.

On the line connecting P_D and P_E, the CB strictly prefers $\{H_2,p,p\}$ to $\{H_1,p,f\}$. To see this, note that $\mu_1 \mid \{H_1,p,f\} = P_D^1$ and $\mu_2 \mid \{H_1,p,f\} = \frac{P_D^2}{2}$. Thus $V(\{H_1,p,f\}) = P_D^1 + \frac{P_D^2}{2} - 2\Omega = \frac{1}{20}$. And $\mu_1 \mid \{H_2,p,p\} = \mu_2 \mid \{H_2,p,p\} = \frac{1}{2}H_2 + \frac{1}{2}\mathbb{E}[H \mid H_2 < H < P_D^1] = \frac{3}{4}H_2 + \frac{1}{4}P_D^1$ and for the health states on this line $H_2 \in]P_D^2, P_C^2]$. Thus, $V(\{H_2,p,p\}) = 2(\frac{3}{4}H_2 + \frac{1}{4}P_D^1) - 2\Omega = \frac{3}{2}H_2 - \frac{1}{10}$. The CB strictly prefers $\{H_2,p,p\}$ to $\{H_1,p,f\}$ iff $V(\{H_2,p,p\}) > V(\{H_1,p,f\})$ which is equivalent to $\frac{3}{2}H_2 - \frac{1}{10} > \frac{1}{20} \forall H_2 \in]P_D^2, P_C^2]$. This reduces to $H_2 > \frac{1}{10}$ which holds since $P_D^2 = \frac{1}{10}$ and $P_C^2 = \frac{3}{25}$.

Thus, the CB sets either $s = \min(H_1, H_2)$ or $s = \max(H_1, H_2)$. However, with the cost of bank runs Ω the CB's strategy can no longer be expressed via an indifference frontier with a closed form expression. Instead, the CB's strategy is characterized by a partitioning frontier (on which it is not always indifferent) which is the set of lines connecting the points $(H_2 = \frac{1}{2}, H_1 = 1)$ to P_A, P_A to P_B, P_B to P_C, P_C to P_E, P_E to P_D, and P_D to $(0,0)$. *Q.E.D.*

D.4 Bank runs and central bank design

In the paper, we established that HRF provides more reassurance than FDF but less than ZDF. This Appendix shows that results are unchanged in a more explicit model of bank runs as in Goldstein and Leitner (2018).

Let welfare be $W(\mu_1, \mu_2) = u(\mu_1) + u(\mu_2)$ and, as in Goldstein and Leitner (2018), let

$$u(\mu_i) = \begin{cases} \mu_i & \text{if} \quad \mu_i > \underline{\mu} \\ \mu_i - \Omega & \text{if} \quad \mu_i \leq \underline{\mu} \end{cases} \tag{46}$$

As in Appendix D.3, $0 < \underline{\mu} < 1$ is the belief below which a bank run occurs and $\Omega > 0$ is the cost of a run. Equation (46) is equivalent to writing welfare as:

$$W(\mu_1, \mu_2) = \mu_1 + \mu_2 - \Omega\, \mathbb{1}(\mu_1 \leq \underline{\mu}) - \Omega\, \mathbb{1}(\mu_2 \leq \underline{\mu}) \tag{47}$$

where $\mathbb{1}(\mu_i \leq \underline{\mu})$ is an indicator function which takes value 1 iff $\mu_i \leq \underline{\mu}$. Otherwise it takes value zero. The expected welfare maximising planner thus maximises:

$$\mathbb{E}[W(\mu_1, \mu_2)] = \mathbb{E}[\mu_1] + \mathbb{E}[\mu_2] - \Omega\, \mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] - \Omega\, \mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})] \tag{48}$$

As in section 4, we compare designs in terms of reassurance by assuming that effort is exogenously fixed and comparing designs in terms of $\mathbb{E}[W(\mu_1, \mu_2)]$. However, while in section 4 we assumed that $W(\mu_1, \mu_2) = \lambda \, \min(\mu_1, \mu_2) + (1 - \lambda) \, \frac{\mu_1 + \mu_2}{2}$ and thus reduced comparisons in terms of $\mathbb{E}[W(\mu_1, \mu_2)]$ to comparisons in terms of $\mathbb{E}[\min(\mu_1, \mu_2)]$, (48) implies that $\mathbb{E}[W(\mu_1, \mu_2)]$ reduces to the following:

Definition 7 *When welfare includes costs of bank runs as described in equation (47), central bank design $\mathcal{D}$ is said to provide more reassurance than $\mathcal{D}'$ if, for given bank effort,* $\Omega\,\mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] + \Omega\,\mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})]$ *is lower in $\mathcal{D}$ than in $\mathcal{D}'$.*

The definition follows naturally from (48). As our focus on reassurance means that we assume that effort is fixed, we have that by Bayes plausibility $\mathbb{E}[\mu_i] = \mathbb{E}[H_i]$ for any design. Thus, designs do not affect $\mathbb{E}[\mu_1] + \mathbb{E}[\mu_2]$ and only affect $\Omega\,\mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] + \Omega\,\mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})]$. Thus, maximising expected welfare is equivalent to minimising the expected cost of bank runs $(\Omega\,\mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] + \Omega\,\mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})])$ or equally to minimising the probability of a bank run, i.e. $\mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] + \mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})]$.

Applying this definition shows that HRF provides more reassurance than FDF but less than ZDF. Consider the same case as in Appendix D.3, i.e. let $H_i \sim U[0,1] \forall i = 1,2$, $\underline{\mu} = 0.2$, $\Omega = 0.1$.

ZDF Since in ZDF no information is disclosed, $\mu_i = \mathbb{E}[H_i] = 0.5 > \underline{\mu} = 0.2$ and thus there never is a run. Hence, $\mathbb{E}[W(\mu_1, \mu_2)] = \mathbb{E}[\mu_1] + \mathbb{E}[\mu_2]$. No design can achieve more reassurance than ZDF as ZDF has minimised the probability of a run to zero.

FDF Since in FDF bank health is always fully revealed ($\mu_i = H_i$), there is a run iff $H_i \leq \underline{\mu_i} = 0.2$. Thus, $\mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] = \mathbb{E}[\mathbb{1}(H_1 \leq 0.2)] = 0.2$. Hence, $\mathbb{E}[W(\mu_1, \mu_2)] = \mathbb{E}[\mu_1] + \mathbb{E}[\mu_2] - 0.4\,\Omega = \mathbb{E}[\mu_1] + \mathbb{E}[\mu_2] - 0.04$.

HRF Assume that the CB's mandate is to maximise ex-post welfare, i.e. $V(\mu_1, \mu_2) = W(\mu_1, \mu_2)$. Then, the equilibrium is as depicted in Figure 7. In this equilibrium, one bank run occurs in the yellow area and in the orange area, and two runs occur in the red area. By symmetry, similar areas exist below the 45 degree line.

To calculate the expected cost of bank runs in a HRF, we proceed in three steps. First, for each area in Figure 7b, we calculate the probability of health state pairs being in that area. Second, we weigh these probabilities with the corresponding number of runs to get the expected cost of runs in the parameter space where $H_1 > H_2$. Third, we multiply these costs by two to account for the other half of the parameter space which follows by symmetry.

The probability of a health state pair being in the yellow region, denoted $Prob(A)$, is

$$Prob(A) = \int_0^{p_C^2} \int_{p_C^1}^{p_B^1} 1\, f(H_1)\, dH_1\, f(H_2)\, dH_2 + \int_{p_C^2}^{P_B^2} \int_{2H_2+2\Omega}^{p_B^1} 1\, f(H_1)\, dH_1\, f(H_2)\, dH_2 \tag{49}$$

The probability of a health state pair being in the orange region, denoted $Prob(B)$, is

$$Prob(B) = \int_0^{P_C^2} \int_{P_D^1}^{P_C^1} 1\, f(H_1)\, dH_1\, f(H_2)\, dH_2 \tag{50}$$

The probability of a health state pair being in the red region, denoted $Prob(C)$, is

$$Prob(C) = \int_0^{P_C^2} \int_{H_2}^{P_D^1} 1\, f(H_1)\, dH_1\, f(H_2)\, dH_2 \tag{51}$$

For the numbers used in this example, we have $Prob(A) = \frac{189}{2500}$, $Prob(B) = \frac{18}{625}$, and

$Prob(C) = \frac{21}{1250}$.

The expected cost of bank runs, given $H_1 > H_2$, is $\Omega Prob(A) + \Omega Prob(B) + 2\ \Omega Prob(C)$. For this example, we have that $\Omega Prob(A) + \Omega Prob(B) + 2\ \Omega Prob(C) = \frac{69}{5000}$.

Multiplying the expected cost of bank runs, given $H_1 > H_2$, by two accounts for the lower half of the parameter space. Thus, we have that:

$$\Omega\, \mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] + \Omega\, \mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})] = 2\Big(\Omega Prob(A) + \Omega Prob(B) + 2\ \Omega Prob(C)\Big) \tag{52}$$

For this example we thus have that $\Omega\, \mathbb{E}[\mathbb{1}(\mu_1 \leq \underline{\mu})] + \Omega\, \mathbb{E}[\mathbb{1}(\mu_2 \leq \underline{\mu})] = \frac{69}{2500} = 0.0276$.

Thus, expected welfare in a HRF is

$$\mathbb{E}[W(\mu_1, \mu_2)] = \mathbb{E}[\mu_1] + \mathbb{E}[\mu_2] - 0.0276 \tag{53}$$

Therefore, the ranking of designs in terms of reassurance derived for this model with bank runs is identical to the ranking in the main paper. Here, ZDF results in maximal reassurance, i.e. a bank run never occurs. In FDF, bank runs can occur and the expected cost of bank runs is 0.04. In HRF bank runs can occur, but the expected cost is 0.0276 and thus lies between the expected costs of ZDF and FDF. *Q.E.D.*