Bayesian Persuasion by Stress Test Disclosure

Wolfgang Gick*

Harvard University and Free University of Bozen/Bolzano

and

Thilo Pausch†

Deutsche Bundesbank

Abstract

This paper argues that banking supervisors can design economy-wide stress tests (macro stress tests) to improve welfare. We show in a multi-receiver framework of Bayesian persuasion that a supervisor can create value when disclosing the stress-testing methodology (signal-generating process) together with the stress test result (signal). By optimally choosing the two pieces of information, supervisors can design the disclosure process in a way to generate a higher expected utility for prudent investors when the latter act accordingly. Overall, the disclosure mechanism suggested in this paper reduces uncertainty and leads to better risk-adjusted behavior on the investors' side, and to less financial market volatility.

Keywords: Stress Tests, Supervisory Information, Bayesian Persuasion, Multiple Receivers, Disclosure.

JEL classification: D81, D84, G28.

Paper prepared for presentation to the ESEM 2013,
Gothenburg, Aug 28, 2013

---

*Corresponding author. Center for European Studies, Harvard University, 27 Kirkland St., Cambridge MA 02138, and School of Economics, Free University of Bozen/Bolzano, Universitätsplatz 1, I-39100 Bozen, Italy, email: wolfgang.gick@unibz.it

†Department of Banking and Financial Supervision, Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, D-60431 Frankfurt am Main, Germany, e-mail: thilo.pausch@bundesbank.de. The views expressed in this paper represent the authors' personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

We thank seminar participants at the Deutsche Bundesbank, the Austrian National Bank, TU Dresden, the Max Planck Institute for Research on Collective Goods in Bonn, furthermore session participants at the EARIE 2012, the German Finance Association 2012 meeting, and the workshop "Banken und Finanzmärkte" at the University of Magdeburg. Special thanks to Monika Bucher, Günther Franke, Mark Le Quement, Dominik Grafenhofer, Christian Leuz, Monika Piazzesi, Jennifer Reinganum, Christoph Schmidhammer, Eva Schliephake, Mihkel Tombak and Alex Weissensteiner for helpful comments. All remaining errors are ours.

1 Introduction

What is the use of having an institutional supervisor disclose information about stress tests that cover the entire banking sector (macro stress tests)? And if there is a use to it, how should banking supervisors optimally design stress tests and the related disclosure mechanisms? This paper focuses on these questions because they deserve particular attention as one of the lessons learned from the financial crisis of 2007-2009 is that intransparency in the banking sector prevents welfare creation.

It is common knowledge that banks improve the allocation of capital in an economy and in so doing, they increase social welfare. Since Diamond and Dybvig (1983), scholars have systematically researched how such processes work. The literature found that transformation processes – in particular maturity transformation, by which banks convert securities with short maturities into long-term assets like bank loans – provide basis for welfare creation.

It lies in the very nature of such transformation processes that they make banking a risky business. Moreover, riskiness of banking affects decisions of potential investors whether or not to provide funds to banks and may, hence, cause financial fragility as the banking system may become unstable when investors do not have the necessary knowledge about banks' risk exposure.[1] This instability may spread, moreover, throughout the whole banking system.[2] Transparency, then, is considered valuable as it may improve the information accessible to investors, and this increased transparency is expected to reduce uncertainty, leading to better risk-adjusted behavior on the investors' side and to less financial market volatility over the business cycle.

In this context, it is easy to understand why Jaime Caruana, the General Manager of the Bank for International Settlements, has emphasized in particular that

> *"[. . .] strengthened, transparent disclosure is good for markets, because it helps investors make more informed decisions."*[3]

Stress test disclosure represents a very prominent as well as a very special instrument of institutional supervisors to enhance transparency in the financial system. The European Banking Authority, the FSA in the UK, the Securities and Exchange Commission (SEC) in the U.S. as well as other national supervisory agencies have designed and performed a multitude of new stress tests over the past few years. Additionally, the number of countries *publishing* bank stress tests has vastly increased from 0 to 40 over the past decade.[4]

[1] See e.g. Freixas and Rochet (2008), Diamond and Rajan (2001).

[2] See Allen and Gale (2000).

[3] Caruana, J. (2011), p. 2.

[4] See Horvath and Vasko (2012).

The frequency by which such macroeconomic stress tests are being performed is clearly rising. One may safely assume that large and more centralized supervisory authorities will prefer to make more use of such tests rather than less, as they provide information about the resilience of the banking sector as a whole. In fact, due to their mandate, institutional banking supervisors appear to be the only entities to generate this kind of valuable information.

We show that an optimal stress-test disclosure mechanism improves investors' assessments of whether the banking sector as a whole is unstable. In our model with heterogenous prior beliefs, investors gain as increased stress test disclosure will, to some extent, adjust distorted prior beliefs, ensuring that investors make correct decisions more often. In addition, optimal informative disclosure increases the banking supervisor's expected utility – compared to a situation with no stress test disclosure. This is notable as we show that informative disclosure exposes the supervisor to risk regarding investors' final behavior. A negative welfare effect, then, may appear when (nearly) all investors behave the same way (herding) as a response to supervisory information disclosure. Our equilibrium analysis, however, shows that this will not happen under an optimally designed disclosure mechanism.

Indeed, many recent contributions confirm that more public disclosure would actually lead to more transparency, which seems to support our argument that disclosing stress-test results enhances transparency and thus financial stability.

Goldstein and Sapra (2012), who discuss the impact of information disclosure on ex-ante incentives as well as on ex-post actions of market participants, do actually favor public disclosure of stress-test results, and argue that

> *[...] at least from a financial stability perspective – the benefits of disclosing stress test results are undeniable."*[5]

Horvath and Vasko (2012) while offering a transparency index for central banks show that the degree of central bank transparency has significantly increased since 2000. More interestingly, they find that greater transparency is beneficial during typical financial periods when financial stress is low. Yet, they also detect that more transparency may be detrimental to stability in times of financial stress.[6] We address this particular point in our paper in a borderline case where the banking sector is completely vulnerable. Our framework suggests for this special case that non-informative disclosure is optimal.

Petrella and Resti (2011), based on the stress test performed by the European Banking Authority in 2011, find that stress tests do provide relevant information to investors and so reduce bank opaqueness. While they find that the market is

[5] Goldstein and Sapra (2012), p. 2.

[6] We also thank Günther Franke to make a similar point.

not able to anticipate the stress test results, they also conclude on the existence of significant positive market reactions as a consequence of stress tests, which indicates that stress tests do help investors make better decisions. Lastly, also Peristiani et al. (2010) conclude in their study that stress tests will generally *reduce* the inherent degree of opacity in the banking sector - the latter being regarded as the epitome of an "opaque industry."[7]

1.1 Theory background

Policy papers as well as empirical studies when referring to terms like "opacity" or "opaqueness" generally mean that transparency is a quality that should be understood against the background of disclosure processes in second-best environments. The most influential theory that has shaped this understanding is work by Morris and Shin (2000, 2002, and 2006), who have studied the disclosure of *one* piece of information (e.g. a macroeconomic fundamental) by a public actor such as a central bank. In their setup, this triggers two distinct actions by market participants: first they decide based on the economic fundamentals, *and*, second, they form expectations on what other market participants will do. As a result, greater transparency (that is, greater precision of public information) is not always desirable. This view borrows from Keynes' (1936) idea of a beauty contest where market participants wish to best respond to public information by conjecturing what the majority perception of the best action will be.

While quoting Morris and Shin (2002), many authors have interpreted their theory as one in which investors wish to act like each other: when a central bank discloses the value of a fundamental, investors will typically put too much weight on this information, and Keynes' beauty contest hits. As Goldstein and Sapra (2012) put it, there can be over-reaction to public information. This leads to a result where all investors will *reduce* the weight of their private information, which is detrimental as this reduction is no longer indicated by the disclosure itself, leading to a situation where individuals, *as long as the central bank's goal is solely encompassing individual action that follows the fundamentals disclosed by the central bank* - more transparency is bad as it crowds out the use of other pertinent information. In the words of Goldstein and Sapra,

"[...] instead of providing market discipline, if not properly designed, disclosure of these stress test results may actually create more panic, therefore lowering confidence in the banking sector."[8]

We show in this paper that this situation can be mitigated as there exist ways to properly design stress test disclosure, and that such design options come naturally

[7] Morgan (2000), Haggard and Howe (2007).

[8] Goldstein and Sapra, (2012), p. 2.

with a fundamental property of the stress test disclosure itself as it involves more than one piece of information. Typically, publicly appointed supervisors do not only use their disclosure *to make sure that investors follow solely the fundamentals* as a simple interpreting of Morris and Shin (2002) would suggest. In fact, bank supervisors do rely on a very peculiar form of disclosure. As a matter of fact, stress test disclosure comes with two pieces of information: the stress test result, *and* the stress test mechanism. In other words, stress test disclosure encompasses a very specific form of signaling, distinct from the known way used by central banks to convey information about fundamentals.

Two papers are in line with our findings. First, Angeletos and Pavan (2004) revert the specific beauty contest argument used in Morris and Shin (2002), and show that as long as investors use public information to a *suboptimal degree* concerning the internalization of the positive externality of their own investment on the return on others, more transparency will necessarily increase welfare. Secondly, Svensson (2006) argues that conditions under which greater transparency would in fact *reduce* welfare will, however, hinge on assumptions that are quite unlikely to occur. Specifically, as Svensson shows, Morris and Shin involve situations with disclosure that actually imply a *lower* welfare situation than no disclosure at all. In turn, the minimum level of disclosure would be above the level from which onward welfare increases with transparency. Thus, Svensson finds that a realistic interpretation of Morris and Shin will remain in favor of transparency.

We argue that the way macroeconomic stress tests are perceived by financial markets calls for a new theory.[9] What we know is that a supervisory agency both discloses the stress test *design, and the result.* It is this quality that sets the stage for a new approach, which this paper is set out to deliver.

The theoretical framework that we offer belongs to the larger strand of literature on cheap talk games or *strategic information transmission,* a strand that developed after the seminal contribution of Crawford and Sobel (1982) (CS hereafter). Indeed, to use strategic information scenarios to explain central bank disclosure practices is not new. For example, Stein (1989) has argued that a central bank (sender) when setting a target exchange rate at a point in time, should take into account that the market (receiver) may hold beliefs that could include a policy reversal for a later date. Consequently, disclosure should optimally remain coarse and limited as the market would show less of a reaction compared to what the sender would like to see.

Sender-receiver models of cheap talk are generally plagued by a high degree of strategic complexity. Theorists have thus looked for game forms that are more applicable to real-world setups. Bayesian persuasion games differ in an important aspect from the original CS game. In a persuasion game, the sender can, to some

[9] See also Borio et al. (2011).

extent, *control the communication environment of the receiver(s).*[10] The sender now relies on more than just one signal, which now permits him to make use of the signal-generating process as well as of the stress-testing procedure. This permits to deliver *additional* information to the receiver, and to trigger specific actions for different types of receivers in our model. The novelty of the setup is that the new game displays properties that are much different from those found under the standard CS framework: as the sender knows that receivers will update their prior beliefs by using Bayes' rule, he can, *without actually being better informed* than the receiver(s), design stress-test disclosure mechanisms in a way that permits him to influence the receivers' actions. The very feature to operate with two forms of disclosure now eliminates a well-known drawback of cheap talk games, *namely that messages are arbitrary for a wide range of beliefs.* Thus, the problem of coordinating on a common language disappears.[11] In persuasion games, this problem is now mitigated by the property that the sender creates a "meaning" of messages based on the availability of two pieces of information.

By providing an analysis based on Bayesian persuasion, our paper no longer follows opacity arguments. Instead, it opens the path for a new understanding of stress tests disclosure *based on the option that the supervisor will structure disclosure itself to persuade investors as he has more than one piece of information available.* What we show is that a public banking authority can generally "persuade" investors over a wide range of (prior) beliefs to take actions that lead to a socially optimal trade-off between individual risk bearing and the provision of liquidity to the banking sector. By making optimal use of two informational components (the result and the signal-generating process), stress tests will influence a continuum of prudent investors, in a way that permits to increase welfare.

The specific literature on persuasion games has been laid out in a seminal article by Kamenica and Gentzkow (2011) (KG hereafter), a setting with one sender and one receiver. In what follows, we provide a continuous multi-receiver model with properties that differ from KG.

The paper is organized as follows. Section 2 introduces the model, defines the timing of the game, introduces investors' preferences and the supervisor's utility function as well as the concept of Bayesian Plausibility. Section 3 defines the equilibrium and derives conditions for optimal disclosure, adding a benchmark of uninformative disclosure, and expands on welfare implications. Section 4 concludes. The proofs to all lemmas appear in the appendix.

[10]See Sobel (2013).

[11]See Sobel (2013).

2 Model

2.1 Primitives

Consider a setting with two types of agents: a single banking supervisor (sender) and a (infinitely) large number of investors (receivers). The supervisor (S) owns a technology (a stress testing mechanism) that provides him with reliable information about the true status of the banking sector under his control. While information acquisition is costly, these costs are fixed and do not depend on the information revealed about the banking sector.

We assume a binary state space of the supervisor's information: either the banking sector is firm / sound (F) or vulnerable (V), meaning that some adverse situation or crisis hurts the banking sector either marginally or heavily. We formally summarize this binary state space by $\Theta = \{F, V\}$ where θ denotes a realization of a certain stress-testing exercise. The prior (objective) probability distribution over Θ is $Pr(V) = p$ and $Pr(F) = 1 - p$.

The investors (I) do not know the state of the banking sector, but they hold prior beliefs (b) about the probability that the banking sector is vulnerable. These priors may be considered to arise from experiences over past periods, or they may be based on the evaluation of other institutions such as rating agencies. Prior beliefs are, hence, considered to be heterogenous among investors.

We assume that the total number of investors is infinitely large and can be normalized to one. Further, let $g(b)$ denote the continuous function which represents the distribution of prior beliefs over all investors. The corresponding continuous function $G(b)$ denotes the cumulative distribution function of prior beliefs as well as the number of investors having prior beliefs of, at most, b.

Based on their individual beliefs, investors make their investment decisions and choose an action in the action space. They decide what their behavior in the banking sector will be. The action space of investors is assumed to be binary as well: investors may either act prudently (P) or riskily (R). Prudent behavior means that the investor under consideration believes that the banking sector is vulnerable and is, therefore, willing to provide funding to banks (e.g. by depositing money with banks or by buying bank bonds) only to the minimum amount that is backed by some deposit insurance mechanism or through State guarantees. All investors whose individual beliefs are *beyond* some threshold probability b_T behave this way. Risky behavior, in contrast, refers to decisions that result in investors providing much more funds to the banking sector than the benchmark would minimally suggest. This latter behavior occurs when investors choose individual beliefs below the threshold probability, b_T. The binary action space will be denoted by $A = \{P, R\}$ in what follows, and a denotes a certain realization. The threshold probability b_T

then defines when investors prefer to switch from risky to prudent behavior because they believe that the probability that the banking sector is vulnerable is too high.

What essentially follows from these assumptions is that they imply an overall distribution of beliefs, which now determines the total number of investors who behave prudently or riskily, respectively. Specifically, $G(b_T) = Pr(b \leq b_T)$ is the total number of investors who prefer a risky strategy whereas $1-G(b_T) = Pr(b > b_T)$ is the total number of prudent investors given their prior beliefs.

2.2 Timing

The timing of the game is as follows. First, nature determines the true state of the banking system $\theta \in \Theta$, not observed by any player. Second, the supervisor chooses a stress-test disclosure mechanism π. This mechanism consists of two elements: a signal from a binary realization space of the supervisor's stress tests results $D \in \{f, v\}$, plus a related family of conditional distributions $\{\pi|\cdot\}_{\theta\in\Theta}$ over D. Note that the supervisor does not have any superior knowledge concerning the true state of the banking system at the time when designing the stress test mechanism. Rather, $\pi(\cdot)$ defines the reliability of stress test results $d \in D$.

The conditional distributions follow the design and accuracy of macroeconomic stress tests as they are commonly observable. In mathematical terms, the design of a stress test determines the probabilities of the result $d \in D$, given that the true state of the banking sector is $\theta \in \Theta$. In other words, the supervisor chooses his degree of informativeness before sending the two messages.

Let $\pi(v|V)$ $(\pi(f|V))$ and $\pi(f|F)$ $(\pi(v|F))$ denote the probabilities that the stress test generates correct (incorrect) results. The stress-test design determines the following probabilities:

$$\begin{array}{c} \pi(v|V) \\ \pi(f|V) = 1 - \pi(v|V) \end{array} \quad \text{and} \quad \begin{array}{c} \pi(f|F) \\ \pi(v|F) = 1 - \pi(f|F) \end{array}.$$

In the third step, the supervisor carries out a stress-testing exercise, observes (privately) stress-test result $d \in D$ and reports this result together with the full information about the stress test's design, i.e. $\pi(\cdot)$, to investors.

Forth, investors observe both the supervisor's choice of the stress test mechanism and the stress test realizations and chooses between a prudent or a risky action.

As already mentioned in the introduction, information disclosure in our model as well as in KG does not only mean to send out a single message as in standard models of cheap talk. Instead, disclosure now includes information about the signal-generating process *and* about the obtained signal as the supervisor in our model reveals information to the public about stress-test design (including underlying assumption, information regarding the data analysis, and so on) *and* about the outcome of the stress test.

2.3 Preferences

Each investor considers combinations of investor actions a and states of the banking sector θ to decide on optimal investment behavior. For simplicity, an investor's perceived value of a specific action depends on him making correct investment decisions and takes values of 1 or 0. An investor's perceived value is 1 if he makes the correct decision: $U_I(P,V) = U_I(R,F)$. Otherwise, his value remains $U_I(P,F) = U_I(R,V) = 0$.[12]

The general aim of banking supervisors is the promotion of the stability of the banking sector.[13] Supervisory disclosure, in particular, intends to improve transparency and market discipline in the banking sector.[14] The standard argument in this regard is that transparency disciplines banking as investors would refrain from investing in banks which are obviously vulnerable.[15] The drawback with this line of argument, however, is that herding behavior may be a (extreme) consequence of supervisory disclosure. That is, a very high level of transparency may drive investors to either refrain form investing in the banking sector in case of bad information or provide excessive funds to banks in case of good information. As a result, transparency and market discipline as a consequence of supervisory disclosure may come at the cost of high volatility in investor behavior, which may translate into a higher instability of the banking sector. Herding behavior on the part of investors - either too many investors acting prudently or too many investors acting riskily - may negatively impact the banking sector as a whole. With too many prudent investors banks may face severe funding problems which may in the end make the system even more vulnerable. In turn, with too many investors acting riskily, banks may find themselves in a situation of excess liquidity which may trigger excessive risk taking by banks in order to profitably invest available funds. However, excessive

[12]Although each single investor makes his own decision we do not label investors. This is only for notational convenience and does not affect results because investors do not differ regarding utility functions. Moreover, our viewpoint is compatible with the class of single-peaked, quadratic-loss functions in the tradition of Crawford and Sobel (1982). Without loss of generality, we have omitted stating them mathematically in the same way as we use a multi-receiver model with the supervisor aggregating individual decisions in a way, that is different from the literature.

[13]See BCBS (2011), paragraph 1.

[14]See BCBS (2011), paragraph 91.

[15]Note that Jaime Caruana's (2011) statement we quoted in the Introduction perfectly supports this argument.

risk taking will also aggravate the vulnerability of the banking sector.[16] In other words, banking supervisors face a tradeoff between transparency and the stability impact of investor behavior when designing disclosure mechanisms.

A supervisor's utility function should therefore reflect this tradeoff. With p denoting the given objective probability of the banking sector being vulnerable, we assume that there exists a certain number of prudent investors $|P|^{max}$ that maximizes the supervisor's utility. That is, for a given p, $|P|^{max}$ denotes the number of prudent investors which best addresses the before mentioned tradeoff. The more the number of prudent investors deviates from $|P|^{max}$ the higher the danger of ending up in a situation when all investors either act prudently or riskily. A supervisor's utility should then be lower than with $|P|^{max}$. Moreover, given the timing of the game, the final outcome of the stress test and therefore the investors' responses to information disclosure are uncertain from the supervisor's perspective at the time when he decides on the disclosure mechanism – a fact that crucially determines the equilibrium construction offered in the next section. The supervisor can only *form expectations about investor behavior*. Therefore, and in accordance with the equilibrium construction, the supervisor's utility function is invariant to the specific stress-test outcome. In sum, the utility function appears as hill-shaped, reaching a maximum at $|P|^{max}$ and zero at the extremes when either all investors act prudently or all investors act riskily.

Let $U_S(|P|)$ denote the supervisor's utility as a function of the number $|P|$ of prudent investors, i.e. investors who choose $a = P$.[17] The supervisor's utility becomes zero when either all investors act riskily or all investors act prudently. Moreover, there exists a number $|P|^{max} \in (0; 1)$ where the supervisor's utility reaches a maximum U_S^{max}.[18] In sum, we assume the following continuously differentiable

[16] This argument would call for an endogenous determination of p as informative disclosure affects investor behavior, which in turn affects the stability of the banking sector and, therefore, p. But as the goal of the paper is on optimal disclosure, we leave the endogenous determination of p for future research.

[17] Due to our assumptions, $|P|$ is found by inserting the threshold probability b_T into the cumulative distribution function of investor beliefs in a given situation. For instance, in the case of a non-informative disclosure mechanism we have $|P| = 1 - G(b_T)$. In the case of effective information disclosure we have $|P| = 1 - \hat{G}_v(b_T)$ and $|P| = 1 - \hat{G}_f(b_T)$ when the supervisor discloses signals $d = v$ and $d = f$, respectively.

[18] We do not explicitly consider p to be an argument of $U_S(\cdot)$ as we earlier assumed that p is given exogenously. That is, given the (objective) probability p there exists a socially-optimal number of prudent investors that maximizes the supervisor's utility.

hill-shaped curve representing the supervisor's utility:[19]

$$U_S = U_S(|P|) \qquad \text{with } U'_S(|P|) > (<)0 \;\forall\; |P| < (>)|P|^{max},\; U'_S(|P|^{max}) = 0 \quad (1)$$
$$\text{and } U'_S(0) \to \infty,\; U'_S(1) \to -\infty.$$

2.4 Investor beliefs

2.4.1 Bayesian updating

By selecting the appropriate stress test design, the supervisors define the framework that shows investors how to update prior beliefs using Bayes' rule when the stress test shows a specific outcome. Related to KG and Wang (2012), the supervisor is able to influence investor behavior in a specific way.

Formally, this reads as follows. First and as in the literature, let $\mu_b(\theta|d)$ denote the posterior belief of an investor with individual prior realization b that the true state of the banking sector is θ when the supervisor discloses d and applies stress-test design $\{\pi|\cdot\}_{\theta\in\Theta}$. In particular, the posteriors for any prior belief b are:

$$\begin{array}{lcl} \mu_b(V|v) = \frac{\pi(v|V)\cdot b}{\pi(v|V)b+\pi(v|F)(1-b)} & & \mu_b(F|f) = \frac{\pi(f|F)\cdot(1-b)}{\pi(f|V)b+\pi(f|F)(1-b)} \\ & \text{and} & \\ \mu_b(F|v) = \frac{\pi(v|F)\cdot(1-b)}{\pi(v|V)b+\pi(v|F)(1-b)} & & \mu_b(V|f) = \frac{\pi(f|V)\cdot b}{\pi(f|V)b+\pi(f|F)(1-b)} \end{array} \quad (2)$$

when the supervisor discloses $d = v$ and $d = f$, respectively. As a consequence, Bayesian updating will affect the cumulative distribution function of investors' beliefs. The following subsection describes how the stress test design $\{\pi|\cdot\}_{\theta\in\Theta}$ and the stress test outcome d jointly affect investor beliefs.

2.4.2 Distribution of investors' posterior beliefs and uncertainty

Since stress-test design $\{\pi|\cdot\}_{\theta\in\Theta}$ is public information and does *not* depend on the prior belief of any single investor, the outcome of the stress test will affect beliefs of *all* investors in the same direction. Consider the borderline case in which a stress test is completely uninformative. That is, it must be true that $\mu_b(V|v) = \mu_b(V|f)$ and $\mu_b(F|v) = \mu_b(F|f)$ for any b and $\pi(v|V) = \pi(f|V) = \pi(v|F) = \pi(f|F) = \frac{1}{2}$.[20]

---

[19] The second line of the definition adapts the idea of Inada conditions to more standard utility functions.

[20] Note that $\mu(V|v) = \mu(V|f)$ and $\mu(F|v) = \mu(F|f)$ both require

$$\frac{\pi(v|F)}{\pi(v|V)} = \frac{\pi(f|F)}{\pi(f|V)}.$$

Due to $\pi(v|V) = 1 - \pi(f|V)$ and $\pi(v|F) = 1 - \pi(f|F)$ the former condition holds if and only if $\pi(v|V) = \pi(f|V) = \pi(v|F) = \pi(f|F) = \frac{1}{2}$.

Now, making the stress test informative - i.e. setting $\pi(v|V) > \frac{1}{2}$ and/or $\pi(f|F) > \frac{1}{2}$ - works as follows:[21]

$$\pi(v|V) > \tfrac{1}{2} \;\Rightarrow\; \mu_b(V|v) > \mu_b(V|f) \text{ and } \mu_b(F|f) > \mu_b(F|v) \;\forall\; b,$$

where the last part follows from $\pi(f|V) = 1 - \pi(v|V)$ and the definition of posterior beliefs (2) above. Moreover, perfectly analogously and using the same reasoning, we now observe the following:

$$\pi(f|F) > \tfrac{1}{2} \;\Rightarrow\; \mu_b(F|f) > \mu_b(F|v) \text{ and } \mu_b(V|v) > \mu_b(V|f) \;\forall\; b.$$

In a next step, let $x \in [0, \frac{1}{2}]$ and $y \in [0, \frac{1}{2}]$ denote the level of precision – ie the amount by which $\pi(v|V)$ and $\pi(f|F)$ exceed $\frac{1}{2}$ – of the signal $d = v$ and $d = f$, respectively.

These observations already lead to two results. First, although the supervisor can decide about the precision of a certain signal, i.e. $d = v$ and $d = f$, basically separately, there is an interaction between both. This interaction will make signals reinforce each other in terms of their precision. As a consequence, posterior beliefs are driven by the precision of both signals as well as by their interaction. The second result is that the relations hold for any arbitrary prior investor belief b.

This gives the model the following interesting twist: when the supervisor's disclosure mechanism is informative, the public signal $d \in D$ will crucially affect the cumulative distribution function of posterior beliefs. Therefore, let $\hat{G}_v(\cdot)$ and $\hat{G}_f(\cdot)$ denote the cumulative distribution functions of posterior beliefs in the case of signals $d = v$ and $d = f$, respectively. Further, for any b, let us denote $\hat{b}_v(b) = \mu_b(V|v)$ and $\hat{b}_f(b) = \mu_b(V|f)$ as the respective posterior beliefs that the banking sector is vulnerable when signals $d = v$ and $d = f$ are observable. Then, for any b, the above arguments imply:

$$\pi(v|V) > \tfrac{1}{2} \quad\Rightarrow\quad \hat{b}_v(b) \geq b \text{ and } \hat{b}_f(b) < b \;\forall\; b \tag{3}$$
$$\Rightarrow\quad \hat{G}_v(b) \leq G(b) \text{ and } \hat{G}_f(b) > G(b) \;\forall\; b. \tag{4}$$

and

$$\pi(f|F) > \tfrac{1}{2} \quad\Rightarrow\quad \hat{b}_f(b) \leq b \text{ and } \hat{b}_v(b) > b \;\forall\; b \tag{5}$$
$$\Rightarrow\quad \hat{G}_f(b) \geq G(b) \text{ and } \hat{G}_v(b) < G(b) \;\forall\; b. \tag{6}$$

[21] While $\pi(v|V) < \frac{1}{2}$ may occur, we limit our attention to the case with $\pi(v|V) > \frac{1}{2}$ in the case of informative disclosure. The reason is twofold: first, $\pi(v|V) > \frac{1}{2}$ refers to a situation of truthful disclosure, which directly translates into investor utility. Second, $\pi(v|V) < \frac{1}{2}$ simply means that the supervisor is likely to send a signal that is exactly the opposite of the true state. Since investors will know the disclosure mechanism, this supervisory strategy will cause investors to decide inversely.

In words: an informative disclosure mechanism increases or decreases the beliefs of any single investor *relative to his priors*, depending on the disclosed signal – this is what (3) and (5) reveal. As a result, we have for the aggregate setting, depending on the supervisor's signal, that information disclosure now shifts mass to the left tail or to the right tail of the cumulative distribution function – see (4) and (6). We summarize these findings as follows:

Lemma 1 *Informative disclosure by the supervisor shifts the cumulative distribution function of investors' beliefs in the sense of first-order stochastic dominance (FSD). The signal v (f) deteriorates (improves) the cumulative distribution function of investor beliefs in the sense of first-order stochastic dominance (FSD) compared to the prior distribution.*

Proof: See Appendix A. ■

Note that the FSD shift of the cumulative distribution functions may be formally written as

$$\frac{\partial \hat{G}_v(b|x,y)}{\partial x} \leq 0 \text{ and } \frac{\partial \hat{G}_v(b|x,y)}{\partial y} < 0 \ \forall b \tag{7}$$

$$\frac{\partial \hat{G}_f(b|x,y)}{\partial x} > 0 \text{ and } \frac{\partial \hat{G}_f(b|x,y)}{\partial y} \geq 0 \ \forall b. \tag{8}$$

The FSD shift and the finding that the direction of the FSD shift depends on the stress-test outcome d leads to an important implication regarding the supervisor's decision environment: Implementing an informative stress-testing mechanism *replaces the certain decision-making situation of the supervisor with an uncertain one.* Recall that the supervisor does not know the true state of the banking sector at the time when he decides on the optimal stress testing mechanism. Therefore, the supervisor faces uncertainty about the signal to be sent to investors in a later stage of the game. In the case of an uninformative stress-testing mechanism this fact, however, is irrelevant for the supervisor. In this latter case, investors will decide based on their prior beliefs, which implies that the supervisor knows the number of prudent investors *with certainty.* In contrast, in a situation of informative disclosure the stress test outcome d determines the direction the investors' beliefs shift and, as a result, the number of prudent investors. As the stress test outcome d is unknown at the time when the stress test mechanism is designed, the supervisor creates an environment of *uncertainty* regarding the effective number of prudent investors.

To summarize: when designing an informative disclosure mechanism, the supervisor actually replaces a certain decision-making situation with an uncertain decision-making situation *(disclosure lottery).* Whether this is valuable depends on both the supervisor's utility function and on the level of expected utility generated by the disclosure lottery, relative to the utility of the certain situation with uninformative disclosure. In other words, the disclosure lottery may have a negative welfare impact. Yet, in what follows, we show that this will not be the case in the optimum.

2.4.3 Investor decisions

Investors, in general, have to choose an action out of their action space $A = \{P, R\}$. Let, as a general representation, $Pr_i(V)$ and $Pr_i(F)$ denote the individual probabilities (beliefs) of an arbitrary investor i that the true state of the banking system is vulnerable (V) or firm (F). Then investor i's expected value of choosing prudent behavior ($a = P$) or risky behavior ($a = R$) is

$$\mathrm{E}\left(U_I(a=P)\right) = Pr_i(V)U_I(P,V) + Pr_i(F)U_I(P,F) = Pr_i(V) \quad (9)$$

or

$$\mathrm{E}\left(U_I(a=R)\right) = Pr_i(V)U_I(R,V) + Pr_i(F)U_I(R,F) = Pr_i(F), \quad (10)$$

respectively, because of $U_I(P,V) = U_I(R,F) = 1$ and $U_I(P,F) = U_I(R,V) = 0$.

From expected values (9) and (10) it is easily verified that an arbitrary investor i is indifferent between $a = P$ and $a = R$ if and only if $Pr_i(V) = Pr_i(F) = \frac{1}{2}$. As a consequence, any investor who believes that $Pr_i(V) \leq \frac{1}{2}$ will choose $a = R$, and any investor who thinks that $Pr_i(V) > \frac{1}{2}$ will choose $a = P$. In other words: the threshold probability b_T that we mentioned earlier, which defines when investors switch from a risky to a prudent strategy is unambiguously

$$b_T = \tfrac{1}{2}.$$

Specifically, it is irrelevant in this context whether investors revert to their prior beliefs for decision making or whether they update and build posterior beliefs. The mode of decision making is unaffected by setting $Pr_i(V) = b$, $Pr_i(V) = \hat{b}_v(b)$, or $Pr_i(V) = \hat{b}_f(b)$. Therefore, it must be true that $b_T = \frac{1}{2}$ is constant, regardless of the information disclosed by the supervisor.

Supervisory information disclosure, however, may affect the *level* of the expected value of a specific investor action. In the previous subsection (Lemma 1) we found that an informative disclosure mechanism causes $\hat{b}_v(b) \geq b \; \forall \; b$ when the supervisor sends a signal $d = v$, and $\hat{b}_f(b) \leq b \; \forall \; b$ when the supervisor's signal is $d = f$. That is, in the case where $d = v$, any investor's expected value from deciding prudently increases, compared to the situation without supervisory information disclosure. Instead, when a supervisor sends $d = f$, any investor's expected value of prudent behavior will *decrease*. Therefore, on one hand, investors may prefer to adjust their investment decisions according to the signal received. The supervisor, on the other hand, is able to affect the investors' expected value and their choice of actions in a specific way by applying an optimally designed disclosure mechanism.

2.4.4 Bayesian Plausibility and the value of informative disclosure

The standard persuasion literature, which considers games with a single sender and a single receiver, argues that Bayesian Plausibility (BP) is the only restriction imposed on a sender's mechanism.[22] This implies that the expected posteriors must be equal to the *objective probability* of a specific situation.

The present model now adds a special twist to the basic setup: with an infinitely large number of investors and with heterogenous prior beliefs, we are now able to generalize BP to find an equilibrium condition. This becomes evident by taking a closer look at the formal representation of BP for the single investor (receiver) case:

$$\mu_b(V|v)Pr(v) + \mu_b(V|f)Pr(f) = p \tag{11}$$

with $Pr(v) = \pi(v|V) \cdot p + \pi(v|F) \cdot (1-p)$ and $Pr(f) = \pi(f|V) \cdot p + \pi(f|F) \cdot (1-p)$.

Although (11) is written for an arbitrary realization b of investors' prior beliefs, it can be shown that (11) does not hold for all (heterogenous) b simultaneously. Calculating the posterior beliefs and taking into account the previous definition of precision of the supervisor's disclosure mechanism, we can rewrite (11)[23]

$$\left(\tfrac{1}{2}+x\right) b \frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))} + \left(\tfrac{1}{2}-x\right) b \frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))} \quad = \quad p. \tag{12}$$

Recall that in the basic persuasion model of KG, Bayesian Plausibility (11) requires that posterior beliefs need to be unbiased *in expected terms.* That is, for a persuasion mechanism to work properly, the sender needs to make sure that beliefs are not distorted in a way such that receivers would suffer a welfare loss from acting according to information disclosed. Note that the left-hand side of (11) - ie $\mu_b(V|v)Pr(v) + \mu_b(V|f)Pr(f)$ - represents a (single) investor's expected value of prudent behavior.[24] If persuasion increases (decreases) this term beyond (below) the objective probability p, prudent behavior becomes more (less) attractive for a

[22] See Sobel (2010), p. 20, and Kamenica and Gentzkow (2011), p. 2596.

[23] Due to previous definitions we have:

$$\begin{aligned}
\mu_b(V|v) &= \frac{\left(\frac{1}{2}+x\right) b}{\left(\frac{1}{2}+x\right) b + \left(\frac{1}{2}-y\right)(1-b)} = \frac{\left(\frac{1}{2}+x\right) b}{\frac{1}{2}-(y-b(x+y))} \\
\mu_b(V|f) &= \frac{\left(\frac{1}{2}-x\right) b}{\left(\frac{1}{2}-x\right) b + \left(\frac{1}{2}+y\right)(1-b)} = \frac{\left(\frac{1}{2}-x\right) b}{\frac{1}{2}+(y-b(x+y))} \\
Pr(v) &= \left(\tfrac{1}{2}+x\right) p + \left(\tfrac{1}{2}-y\right)(1-p) = \tfrac{1}{2}-(y-p(x+y)) \\
Pr(f) &= \left(\tfrac{1}{2}-x\right) p + \left(\tfrac{1}{2}+y\right)(1-p) = \tfrac{1}{2}+(y-p(x+y)).
\end{aligned}$$

[24] In section 2.4.3 it was shown that the value of a specific action a to an investor is given by the (posterior) belief regarding the status of the banking sector.

(single) investor. However, due to the distortion, an investor will find himself *ex post* too often in a situation where his initial decision proves to be incorrect. The investor then realizes $U_I(P,F) = 0$ or $U_I(R,V) = 0$ instead of $U_I(P,V) = 1$ or $U_I(R,F) = 1$ – compared to decision-making based on prior beliefs when ignoring the supervisor's information.

However - and here we differ from the literature such as Wang's (2011) voting model - *this cannot hold for every single investor when there are many investors with heterogenous prior beliefs and when the supervisor's disclosure mechanism is publicly known*, i.e. (x, y) stays the same for all investors. In this situation (11) - or (12) for our particular situation - can only hold for an investor whose prior belief b is exactly identical to the objective probability p of a vulnerable banking sector. That is, except for this latter investor, individual prior beliefs of all other investors *will appear more or less distorted* compared to the objective probability p.

For instance, all investors with $b < (>)p$ believe that the probability of a vulnerable banking sector is lower (higher) than p. For all these "distorted" investors, an informative disclosure mechanism may correct their individual distortions to *some* degree. Put differently: an informative disclosure mechanism of the supervisor generates expected posterior beliefs that are higher (lower) than the individual priors of investors in the case of $b < p$ $(b > p)$. We so state and prove

Lemma 2 *Bayesian Plausibility in our model with multiple investors (receivers) and heterogenous investor prior beliefs requires*

$$\mu_b(V|v)Pr(v) + \mu_b(V|f)Pr(f) = p \ \Leftrightarrow \ b = p.$$

Proof: See Appendix B. ■

From the investors' point of view information disclosure suggests that their individual priors understate or overstate the probability of a vulnerable banking sector in the case of $b < p$ or $b > p$, respectively:

$$\left(\tfrac{1}{2}+x\right) b \frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))} + \left(\tfrac{1}{2}-x\right) b \frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))} \; > \; b \qquad \forall\, b < p \quad (13)$$

or

$$\left(\tfrac{1}{2}+x\right) b \frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))} + \left(\tfrac{1}{2}-x\right) b \frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))} \; < \; b \qquad \forall\, b > p \quad (14)$$

For these "distorted" investors, the supervisor's disclosure mechanism now corrects (part of) this distortion of prior beliefs *in a way that investor beliefs move toward the true probability of a vulnerable banking sector.* As a consequence, information disclosure helps investors to make correct investment decisions. Against this background, the value of an informative disclosure mechanism for investors becomes evident:

Corollary 3 *A disclosure mechanism that is Bayesian plausible according to Lemma 2 raises the expected value of investors whose prior beliefs deviate from the objective probability of a vulnerable banking sector.*

Proof: Note that the left-hand side of (12) represents an investor's expected value from behaving prudently, acting according to the supervisor's information disclosure, and having prior belief b. Under a Bayesian plausible disclosure mechanism, then, from the proof of Lemma 2 it is immediately clear that informative disclosure by the supervisor increases the expected value of prudent behavior for all investors who understate the probability of the banking sector being vulnerable ($b < p$). That is, informative disclosure better aligns the evaluations of latter investors with the true state of the banking system. Moreover, informative disclosure causes investors with $\mu_b(V|v) \geq \frac{1}{2} > b$ to switch from a risky to a prudent investment strategy.

The opposite effect appears to be the case with investors whose priors overstate the true vulnerability of the banking sector. Their expected value from prudent behavior decreases under a Bayesian plausible disclosure mechanism (see proof of Lemma 2). As a result, investors with $\mu_b(V|v) < \frac{1}{2} \leq b$ switch from a prudent to a risky investment strategy. For both situations, the proof of Lemma 2 (see Appendix B) shows that under a Bayesian-plausible disclosure mechanism the expected value of those investors who change their investment strategy in response to supervisory disclosure, will increase. ■

In other words, under a Bayesian plausible disclosure mechanism, "distorted" investors' expected value will increase when they base their decisions on supervisory information instead of prior beliefs. This result is novel to the literature on the disclosure of supervisory information in the banking sector.

3 Equilibrium and Optimality Analysis

3.1 Equilibrium

We are now able to define the equilibrium of the game. Given that the ideal point is $|P|^{max}$ for the supervisor, now let $\{\hat{G}_v(b), \hat{G}_f(b)\}_\pi$ be the posterior distribution of receiver beliefs under the mechanism π. Given threshold $b_T = \frac{1}{2}$ that is independent of the mechanism, the aggregate receiver behavior is summarized by $\{|P|_v, |P|_f\}_\pi = \{1 - \hat{G}_v(\frac{1}{2}), 1 - \hat{G}_f(\frac{1}{2})\}_\pi$. Supervisor's (realization of uncertain) utility under mechanism π is $\{U_S(|P|_v), U_S(|P|_f)\}_\pi$.

Any mechanism π^* that maximizes the expected utility $\mathrm{E}(\{U_S(|P|_v), U_S(|P|_f)\}_\pi)$ of the supervisor, subject to Bayesian Plausibility provides a SPE:

Definition. A SPE is a choice of R or P by each investor and a choice of mechanism π^* by the supervisor such that:

• An investor with Bayesian-updated beliefs $\{\hat{b}_v(b), \hat{b}_f(b)\}_{\pi^*}$ chooses his action with maximum expected value under beliefs $\{\hat{b}_v(b), \hat{b}_f(b)\}_{\pi^*}$; resulting in $\{|P|_v, |P|_f\}_{\pi^*} = \{1 - \hat{G}_v(\frac{1}{2}), 1 - \hat{G}_f(\frac{1}{2})\}_{\pi^*}$, and

• π^* maximizes the supervisor's expected utility $\mathrm{E}(\{U_S(|P|_v), U_S(|P|_f)\}_\pi)$, s.t. Bayesian Plausiblility.

3.2 Optimal disclosure

To answer whether or not the supervisor should optimally implement an informative stress testing (disclosure) mechanism, we now state the supervisor's optimization problem, followed by the existence of an informative disclosure mechanism and then derive some conclusions on welfare.

3.2.1 The supervisor's problem

The supervisor's goal is to maximize his utility while taking into account all the factors that were analyzed in the previous sections. Note that the supervisor's utility function has a unique maximum when the number of prudent investors is exactly $|P|^{max}$. Thus, when the supervisor finds himself in a situation where the investors' prior beliefs result in[25]

$$|P|^{max} = 1 - G(\tfrac{1}{2}),$$

the supervisor will refrain from implementing any informative disclosure mechanism, as this cannot increase his utility.

In what follows, we consider situations where prior beliefs generate numbers of prudent investors that deviate from $|P|^{max}$. In this context there are two possibilities: either we have $1 - G(\frac{1}{2}) < |P|^{max}$ (case a)) or we have $1 - G(\frac{1}{2}) > |P|^{max}$ (case b)). In this context, the supervisor's objective is to design an informative disclosure mechanism – by choosing x and y – such that the distance between the supervisor's highest possible utility and the expected utility realized by informative disclosure is

[25] Note that in section 2.4.3 we argued that the unique threshold when investors switch from a risky to a prudent investment strategy is $b_T = \frac{1}{2}$.

minimized:[26]

$$
\begin{aligned}
\min_{x,y}\ \Delta U_S \ \equiv\ & Pr(v)\left(U_S^{max}-U_S(1-\hat{G}_v(\tfrac{1}{2}))\right)+Pr(f)\left(U_S^{max}-U_S(1-\hat{G}_f(\tfrac{1}{2}))\right)\\
\text{with}\quad & Pr(v)=(\tfrac{1}{2}+x)\,p+(\tfrac{1}{2}-y)\,(1-p)\\
& Pr(f)=(\tfrac{1}{2}-x)\,p+(\tfrac{1}{2}+y)\,(1-p) \qquad (15)\\
& x\leq\tfrac{1}{2} \qquad (\lambda_x)\\
& y\leq\tfrac{1}{2} \qquad (\lambda_y)\\
& x,y\geq 0.
\end{aligned}
$$

Recall that in optimization problem (15) U_S^{max} denotes the supervisor's hightest possible utility, which is achieved when $1-G(\frac{1}{2})=|P|^{max}$. Furthermore $1-\hat{G}_v(\frac{1}{2})$ and $1-\hat{G}_f(\frac{1}{2})$ denote the number of prudent investors when the supervisor sends signals $d=v$ and $d=f$, respectively. Bayesian Plausibility is implicitly considered in the cumulative distribution functions of the investors' posterior beliefs $\hat{G}_v(\cdot)$ and $\hat{G}_f(\cdot)$.[27]

3.2.2 Optimality of informative disclosure

The analysis of optimal information disclosure begins with the derivation of the first-order necessary conditions of optimization problem (15) using the Kuhn-Tucker Theorem:

$$
\begin{aligned}
\frac{\partial\mathcal{L}}{\partial x} =\ & p\left(U_S^{max}-U_S(1-\hat{G}_v(\tfrac{1}{2}))\right)-p\left(U_S^{max}-U_S(1-\hat{G}_f(\tfrac{1}{2}))\right)+\\
& +Pr(v)U_S'(1-\hat{G}_v(\tfrac{1}{2}))\frac{\partial\hat{G}_v(\frac{1}{2})}{\partial x}+Pr(f)U_S'(1-\hat{G}_f(\tfrac{1}{2}))\frac{\partial\hat{G}_f(\frac{1}{2})}{\partial x}+\lambda_x\geq 0\\
& x\geq 0\ ;\ \frac{\partial\mathcal{L}}{\partial x}x=0 \qquad (16)\\
\frac{\partial\mathcal{L}}{\partial y} =\ & -(1-p)\left(U_S^{max}-U_S(1-\hat{G}_v(\tfrac{1}{2}))\right)+(1-p)\left(U_S^{max}-U_S(1-\hat{G}_f(\tfrac{1}{2}))\right)+\\
& +Pr(v)U_S'(1-\hat{G}_v(\tfrac{1}{2}))\frac{\partial\hat{G}_v(\frac{1}{2})}{\partial y}+Pr(f)U_S'(1-\hat{G}_f(\tfrac{1}{2}))\frac{\partial\hat{G}_f(\frac{1}{2})}{\partial y}+\lambda_y\geq 0\\
& y\geq 0\ ;\ \frac{\partial\mathcal{L}}{\partial y}y=0 \qquad (17)\\
\frac{\partial\mathcal{L}}{\partial\lambda_x} =\ & x-\tfrac{1}{2}\leq 0\ ;\ \lambda_x\geq 0\ ;\ \frac{\partial\mathcal{L}}{\partial\lambda_x}\lambda_x=0 \qquad (18)\\
\frac{\partial\mathcal{L}}{\partial\lambda_y} =\ & y-\tfrac{1}{2}\leq 0\ ;\ \lambda_y\geq 0\ ;\ \frac{\partial\mathcal{L}}{\partial\lambda_y}\lambda_y=0 \qquad (19)
\end{aligned}
$$

[26] λ_x and λ_y denote the Lagrange multipliers for the constraints $x\leq\frac{1}{2}$ and $y\leq\frac{1}{2}$, respectively, in the supervisor's optimization problem.

[27] See Appendix C for a formal proof.

where $\mathcal{L}$ denotes the Lagrangean of problem (15) which can be found in Appendix C in explicit form.

The following considerations, which are essential to our analysis, build on a number of insights concerning the cumulative distribution functions $\hat{G}_v$ and $\hat{G}_f$ to which we refer in great detail in section 2.4.2 as well as in Appendix A. First note that the numbers of prudent investors are determined by the cumulative distribution functions of prior and posterior beliefs and that – due to the properties of these functions – the following relation always holds:

$$1 - \hat{G}_f(\tfrac{1}{2}) \leq 1 - G(\tfrac{1}{2}) \leq 1 - \hat{G}_v(\tfrac{1}{2}).$$

That is, starting from the benchmark of uninformative disclosure, in which investors act according to their prior beliefs and the number of prudent investors amounts to $1 - G(\frac{1}{2})$, informative disclosure reduces the number of prudent investors to $1 - \hat{G}_f(\frac{1}{2})$ when the stress-test outcome indicates that the banking sector is firm ($d = f$). Otherwise, when the stress-test outcome points to a vulnerable banking sector ($d = v$) the number of prudent investors increases to $1 - \hat{G}_v(\frac{1}{2})$.

The supervisor, instead, when designing the optimal stress-test mechanism is still uninformed about which one of the latter two cases will materialize. An informative stress test will generate $d = f$ ($d = v$) with some strictly positive probability $Pr(f) \in (0, 1)$ ($Pr(v) \in (0, 1)$). As a result, by creating an informative stress test the supervisor replaces a situation of certainty, i.e. a situation when investors only decide according to their certain prior beliefs, by a situation of uncertainty, i.e. a situation when investor behavior *depends on an ex-ante uncertain stress test outcome.*[28] To take this into account, we now analyze whether the banking supervisor is able to design the disclosure lottery in a way that the resulting expected utility exceeds the certain utility a banking supervisor would earn in the benchmark situation of an uninformative stress test.

Second, setting the partial derivatives of the cumulative distribution functions of investor beliefs to $\gamma = \frac{1}{2}$ (see Appendix A) we have:

$$\begin{array}{lcl} \frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} = -g\left(\frac{\frac{1}{2}-y}{1+x-y}\right)\frac{\frac{1}{2}-y}{[1+x-y]^2} \leq 0 & & \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} = g\left(\frac{\frac{1}{2}+y}{1-x+y}\right)\frac{\frac{1}{2}+y}{[1-x+y]^2} > 0 \\ & \text{and} & \\ \frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y} = -g\left(\frac{\frac{1}{2}-y}{1+x-y}\right)\frac{\frac{1}{2}+x}{[1+x-y]^2} < 0 & & \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} = g\left(\frac{\frac{1}{2}+y}{1-x+y}\right)\frac{\frac{1}{2}-x}{[1-x+y]^2} \geq 0 \end{array} \tag{20}$$

with

$$\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y} < \frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} \leq 0 \quad \text{and} \quad \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} > \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} \geq 0 \ . \tag{21}$$

This permits us to derive the following lemma:

Lemma 4 *It is never optimal for the supervisor to implement an informative disclosure mechanism that either shifts $1 - \hat{G}_v(\frac{1}{2})$ beyond $|P|^{max}$ in case a), i.e. in the*

[28] We will refer to this latter situation of uncertainty as the *disclosure lottery* in what follows.

case when $1 - G(\frac{1}{2}) < |P|^{max}$*, or that shifts* $1 - \hat{G}_f(\frac{1}{2})$ *below* $|P|^{max}$ *in case b), i.e. in the case when* $1 - G(\frac{1}{2}) > |P|^{max}$*.*

In words: $1 - \hat{G}_f(\frac{1}{2})$ and $1 - \hat{G}_v(\frac{1}{2})$ can both lie either to the left or to the right of $|P|^{max}$. It cannot be optimal if only one of the two values moves from one side of $|P|^{max}$ to the other.

Proof: See Appendix D. ■

The intuition behind Lemma 4 is straightforward: starting with case a) or b), the stress test lottery affects the supervisor's utility. In case a) with $1 - G(\frac{1}{2}) < |P|^{max}$, the stress-test outcome $d = f$ *reduces* the number of prudent investors as well as the supervisor's utility compared to the benchmark ($U_S(1-\hat{G}_f(\frac{1}{2})) < U_S(1-G(\frac{1}{2}))$) since $U'_S(\cdot) > 0 \; \forall \; |P| < |P|^{max}$. A stress test result $d = v$ in the same case a) will instead *increase* the number of prudent investors.

More generally, the supervisor's utility will increase if he reports $d = v$ in this case ($U_S(1 - \hat{G}_v(\frac{1}{2})) > U_S(1 - G(\frac{1}{2}))$), but there is a maximum possible utility level for the supervisor in this situation. Note that $U_S(1 - \hat{G}_v(\frac{1}{2}))$ reaches U_S^{max} when $1-\hat{G}_v(\frac{1}{2})$ approaches $|P|^{max}$, and increasing $1-\hat{G}_v(\frac{1}{2})$ beyond $|P|^{max}$ results in lower utility levels $U_S(1-\hat{G}_v(\frac{1}{2})) < U_S^{max}$ as the supervisor's utility function is decreasing beyond $|P|^{max}$. In section 2.4.2 we have shown that the precision parameters x and y will reinforce each other in the sense that each parameter affects both $1 - \hat{G}_f(\frac{1}{2})$ as well as $1 - \hat{G}_v(\frac{1}{2})$. Consequently, the lower utility level $U_S(1 - \hat{G}_v(\frac{1}{2})) < U_S^{max}$ – when $1-\hat{G}_v(\frac{1}{2}) > |P|^{max}$ is accompanied by an even lower utility level $U_S(1-\hat{G}_f(\frac{1}{2}))$ – when the stress test finds $d = f$. In other words: it cannot be optimal to move $1 - \hat{G}_v(\frac{1}{2})$ beyond $|P|^{max}$ in case a) because this would reduce the supervisor's expected utility from informative disclosure compared to all other situations where disclosure is informative and $1 - \hat{G}_v(\frac{1}{2}) \leq |P|^{max}$. An analogous argument holds for case b): with $1-G(\frac{1}{2}) > |P|^{max}$ it cannot be optimal for the supervisor to set x and y such that $1 - \hat{G}_f(\frac{1}{2}) < |P|^{max}$.

Building on the above lemma, we now state our main result in the following proposition:

Proposition 1 *In either case (case a) as well as case b)), there exists an an informative but not fully revealing disclosure mechanism that minimizes the distance between the supervisor's maximum utility and the expected utility arising from informative disclosure.*

Proof: See Appendix E. ■

Proof of Proposition 1 formally shows that it is always beneficial for a supervisor to design an informative stress test mechanism, i.e. a mechanism that provides investors with useful information about the true status of the banking sector. As a result, correct investor decisions become more likely, *yet they are not certain.* Evidently, a fully revealing mechanism, i.e. a mechanism which eliminates all uncertainty of investors about the true state of the banking sector, induces *extreme* investor behavior in the sense that *all* investors act either prudently or riskily. With a fully revealing stress test mechanism, any investor could infer that the banking sector is really vulnerable (firm) if the supervisor reports $d = v$ ($d = f$). Through such a disclosure, a supervisor would induce extreme volatility concerning investor behavior in the banking system – all investors would act prudently (riskily) when they observe $d = v$ ($d = f$): the utility of the supervisor would be zero.

An optimal stress-testing mechanism, although improving investors' information about the true status of the banking sector, will also leave investors with *some* amount of remaining uncertainty. Moreover, the proof of Proposition 1 also illustrates that the supervisor can choose the precision parameters x and y such that the expected utility resulting from his disclosure lottery exceeds the utility resulting from the benchmark situation of uninformative disclosure. Note that the size of x affects the magnitude of the impact of y and vice versa (see relations (20)). Thus, the supervisor is able to limit utility losses and to exploit potential utility gains. Note further that the values of p do not affect the result.

In the previous analysis we have excluded corner solutions, limiting p to values of less then 1. These former results remain valid as long as there is no crisis going on in the banking sector. It is easy to show under conditions when the banking sector is hurt by a systemic crisis that the supervisor will not apply an informative disclosure mechanism. To see this, let $p = 1$. In words, when the banking sector is hit by a crisis, the *objective probability* of a vulnerable banking sector will approach unity. This implies $Pr(v) = \pi(v|V)$ and $Pr(f) = \pi(f|F)$. Moreover, using equation (2) we find that $\mu_b(V|v) = 1$ and $\mu_b(V|f) = 1$ for investors' posterior beliefs when $b = p = 1$. Applying Bayesian Plausibility (equation (11) and *Lemma* 2) now requires $Pr(v) = \pi(v|V) = \frac{1}{2}$ and $Pr(f) = \pi(f|F) = \frac{1}{2}$. This implies that $Pr(v) = \pi(f|V) = \frac{1}{2}$ and $Pr(f) = \pi(v|F) = \frac{1}{2}$. The outcome in the case of an ongoing banking crisis shows that the supervisor's disclosure will optimally need to remain uninformative – a result perfectly in line with a number of recent observations made during the subprime crisis of 2007-2009, as well as during the sovereign crisis that has gone on since 2010.[29]

[29] See Horvath and Vasko (2012).

3.3 Welfare

Our welfare implications follow immediately from Proposition 1 and Corollary 3. Recall that in Section 2 we argued based on Bayesian Plausibility and the underlying decision-making process that investors will either *gain* from information disclosure or realize *at least the same value* compared to a situation without informative disclosure. Moreover, the supervisor will always gain, as Proposition 1 has shown. There exists an optimal *informative* disclosure mechanism that minimizes the distance between the supervisor's maximum possible utility (for a given objective probability p) and the expected utility arising from informative disclosure. In sum, total welfare increases as a consequence of the supervisor's optimal information-disclosure mechanism. Proposition 1 also reveals that, although informative disclosure may have a negative welfare impact because of replacing a certain decision-making situation by an uncertain one, equilibrium welfare is always improved.

4 Conclusion

The goal of this paper was to explain why macro stress tests, based on their property of carrying two pieces of information, can be designed in a way to increase welfare. Stress tests, as already mentioned, have been implemented with a frequency that should lead to some serious discussions about what they can achieve when they are optimally designed. We have delivered a multi-receiver model of Bayesian persuasion that explains some of the options available to supervisors.

In the wider sense, the paper sheds new light on transparency and financial stability, which in our model is achieved through the use of a very peculiar form of disclosure. More generally, several issues that are central to the current debate on public signals and the value of transparency should be reconsidered by practitioners in the light of our analysis. Supervisors, as we have shown, may contribute to improve investor decisions, given the underlying trade-off that economies face between market discipline and financial stability.

In a more narrow sense, we have shown in a multi-receiver persuasion game that senders (supervisors) can optimally design a disclosure mechanism with two pieces of information, namely the signal-generating process together with the resulting test signal. With an eye on transparency, we have revealed a new but important aspect that institutional supervisors (or banking authorities) may make use of, namely the fact that disclosure processes will influence Bayesian receivers to act in an overall welfare-enhancing way.

The disclosure mechanism that we have suggested has a series of attractive properties. It shows a unique interior optimum, is generally robust, and permits the implementation of better disclosure practices with respect to social welfare. In this way, our paper supports Jaime Caruana's (2011) call in favor of more transparency. In addition, we have shown that for the borderline case of a systemic crisis, disclosure should optimally remain uninformative.

Our paper leads to several extensions. A first one would be to determine the objective probability p endogenously, as we have stressed already in Section 2.3. The merit of such an extension would consist in providing greater generalization of the supervisor's utility function.

A second extension could include banks in the set of players, differentiating them along a new (type-) dimension. Extending the model in such a way would make the supervisor a middleman, thus permitting additional disclosure options to be included in the setup. A possible advantage of a more complex treatment could provide the option to segment investment into the matching of bank types to investor types. This, in turn, could make the analysis of differentiated disclosure processes a new field of study, in the light of risk-adjusted behavior. Given the already high degree of complexity that we have reached in this model, this aspect is left for future research.

Appendix

A. Proof of Lemma 1: FSD shift of investor beliefs distribution

Consider the posteriors of any investor with any prior belief b

$$\begin{aligned}
\hat{b}_v(b) &= \mu_b(V|v) = \frac{\left(\frac{1}{2}+x\right)b}{\left(\frac{1}{2}+x\right)b+\left(\frac{1}{2}-y\right)(1-b)} \\
\hat{b}_f(b) &= \mu_b(V|f) = \frac{\left(\frac{1}{2}-x\right)b}{\left(\frac{1}{2}-x\right)b+\left(\frac{1}{2}+y\right)(1-b)}
\end{aligned}$$

where we used the notion of signal precision as defined in section 2.4.2. Changing the precision parameters $x \in [0, \frac{1}{2}]$ and $y \in [0, \frac{1}{2}]$, the mechanism has the following general effect on posterior beliefs:

$$\begin{aligned}
\frac{\partial \hat{b}_v(b)}{\partial x} &= \frac{\left(\frac{1}{2}-y\right)(1-b)b}{\left[\left(\frac{1}{2}+x\right)b+\left(\frac{1}{2}-y\right)(1-b)\right]^2} \geq 0 \\
\frac{\partial \hat{b}_f(b)}{\partial x} &= -\frac{\left(\frac{1}{2}+y\right)(1-b)b}{\left[\left(\frac{1}{2}-x\right)b+\left(\frac{1}{2}+y\right)(1-b)\right]^2} < 0 \\
\frac{\partial \hat{b}_v(b)}{\partial y} &= \frac{\left(\frac{1}{2}+x\right)(1-b)b}{\left[\left(\frac{1}{2}+x\right)b+\left(\frac{1}{2}-y\right)(1-b)\right]^2} > 0 \\
\frac{\partial \hat{b}_f(b)}{\partial y} &= -\frac{\left(\frac{1}{2}-x\right)(1-b)b}{\left[\left(\frac{1}{2}-x\right)+\left(\frac{1}{2}+y\right)(1-b)\right]^2} \leq 0
\end{aligned}$$

where the first and the last line become equal to zero when $y = \frac{1}{2}$ and $x = \frac{1}{2}$, respectively. As a result, a higher level of x implies, *ceteris paribus*, $\hat{b}_v(b) \geq b$ and $\hat{b}_f(b) < b$ for any b whereas a higher level of y implies, *ceteris paribus*, $\hat{b}_v(b) > b$ and $\hat{b}_f(b) \leq b$ for any b. That is, from a formal perspective the disclosure mechanism $(D, \{\pi|\cdot\}_{\theta\in\Theta})$ is a monotonic transformation of investor beliefs.[30]

Let us now denote $\hat{g}_v(b)$ and $\hat{g}_f(b)$ the distribution functions of investors' posterior beliefs when the supervisor sends $d = v$ and $d = f$, respectively. Given the impact of the signaling mechanism on investor beliefs above, the distribution functions of posteriors can be determined to be:

$$\hat{g}_v(b) \; : \; \hat{b}_v(b) \longmapsto g(b) \; \forall \; b \tag{22}$$

$$\hat{g}_f(b) \; : \; \hat{b}_f(b) \longmapsto g(b) \; \forall \; b. \tag{23}$$

The corresponding cumulative distribution functions are, by definition,

$$\begin{aligned} \hat{G}_v(\gamma) &= \int_0^\gamma \hat{g}_v(\hat{b}_v(b))d\hat{b}_v(b) \\ \hat{G}_f(\gamma) &= \int_0^\gamma \hat{g}_f(\hat{b}_f(b))d\hat{b}_f(b). \end{aligned}$$

Applying the definitions of $\hat{b}_v(b)$ and $\hat{b}_f(b)$ above allows for the calculation of these cumulative distribution functions based on the distribution of prior beliefs:

$$\hat{G}_v(\gamma) = \int_0^{\frac{\left(\frac{1}{2}-y\right)\gamma}{\left(\frac{1}{2}+x\right)(1-\gamma)+\left(\frac{1}{2}-y\right)\gamma}} g(b)db \tag{24}$$

$$\hat{G}_f(\gamma) = \int_0^{\frac{\left(\frac{1}{2}+y\right)\gamma}{\left(\frac{1}{2}-x\right)(1-\gamma)+\left(\frac{1}{2}+y\right)\gamma}} g(b)db. \tag{25}$$

Equations (20) and (21) show first that $x \in [0, \frac{1}{2}]$ and $y \in [0, \frac{1}{2}]$ represent parameters which determine the upper limit of the integrals. Therefore the cumulative distributions may be considered to be conditional on x and y, denoted

$$\hat{G}_v(\gamma) \equiv \hat{G}_v(\gamma|x, y) \text{ and } \hat{G}_f(\gamma) \equiv \hat{G}_f(\gamma|x, y) \; \forall \; \gamma.$$

Second, the impact of x and y on the cumulative distributions of posteriors at any

[30]Note that x and y reinforce each other regarding the impact on investor posteriors when a certain signal is received.

$\gamma \in [0,1]$ can be determined by calculating the partial derivatives:

$$\begin{aligned}
\frac{\partial \hat{G}_v(\gamma|x,y)}{\partial x} &= -g\left(\frac{(\frac{1}{2}-y)\,\gamma}{(\frac{1}{2}+x)\,(1-\gamma)+(\frac{1}{2}-y)\,\gamma}\right)\frac{(\frac{1}{2}-y)\,(1-\gamma)\gamma}{[(\frac{1}{2}+x)\,(1-\gamma)+(\frac{1}{2}-y)\,\gamma]^2} \leq 0\\
\frac{\partial \hat{G}_v(\gamma|x,y)}{\partial y} &= -g\left(\frac{(\frac{1}{2}-y)\,\gamma}{(\frac{1}{2}+x)\,(1-\gamma)+(\frac{1}{2}-y)\,\gamma}\right)\frac{(\frac{1}{2}+x)\,(1-\gamma)\gamma}{[(\frac{1}{2}+x)\,(1-\gamma)+(\frac{1}{2}-y)\,\gamma]^2} < 0\\
\frac{\partial \hat{G}_f(\gamma|x,y)}{\partial x} &= g\left(\frac{(\frac{1}{2}+y)\,\gamma}{(\frac{1}{2}-x)\,(1-\gamma)+(\frac{1}{2}+y)\,\gamma}\right)\frac{(\frac{1}{2}+y)\,(1-\gamma)\gamma}{[(\frac{1}{2}-x)\,(1-\gamma)+(\frac{1}{2}+y)\,\gamma]^2} > 0\\
\frac{\partial \hat{G}_f(\gamma|x,y)}{\partial y} &= g\left(\frac{(\frac{1}{2}+y)\,\gamma}{(\frac{1}{2}-x)\,(1-\gamma)+(\frac{1}{2}+y)\,\gamma}\right)\frac{(\frac{1}{2}-x)\,(1-\gamma)\gamma}{[(\frac{1}{2}-x)\,(1-\gamma)+(\frac{1}{2}+y)\,\gamma]^2} \geq 0
\end{aligned}$$

where die inequalities follow from $g(\cdot) > 0$, $x, y \in [0, \frac{1}{2}]$, and $\gamma \in [0,1]$.

B. Proof of Lemma 2: Bayesian Plausibility

Let b_{BP} denote the prior belief of the investor for which (12) holds.

Consider the situation $b < b_{BP}$ first. For a given decision (x,y) of the supervisor we observe $y - b(x+y) > y - b_{BP}(x+y)$ which implies $\frac{1}{2} - (y - b(x+y)) < \frac{1}{2} - (y - b_{BP}(x+y))$ and $\frac{1}{2} + (y - b(x+y)) > \frac{1}{2} + (y - b_{BP}(x+y))$. For the fraction terms in (12) we therefore find

$$\begin{aligned}
\frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))} &> \frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b_{BP}(x+y))}\\
&\text{and}\\
\frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))} &< \frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b_{BP}(x+y))}.
\end{aligned}$$

Since $(\frac{1}{2}+x)\,b \geq (\frac{1}{2}-x)\,b$ – and note that $b \in [0,1]$ and $x \in [0,\frac{1}{2}]$ – we have

$$(\tfrac{1}{2}+x)\,b\frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))} + (\tfrac{1}{2}-x)\,b\frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))} > b \qquad \forall\ b < b_{BP}.$$

Regarding the situation $b > b_{BP}$ the arguments are analogous but the previous relations turn in the opposite direction. That is, for a given decision (x,y) we observe $y - b(x+y) < y - b_{BP}(x+y)$, $\frac{1}{2} - (y - b(x+y)) > \frac{1}{2} - (y - b_{BP}(x+y))$ and $\frac{1}{2} + (y - b(x+y)) < \frac{1}{2} + (y - b_{BP}(x+y))$. As a consequence the relations between the fraction terms in (12) are:

$$\begin{aligned}
\frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))} &< \frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b_{BP}(x+y))}\\
&\text{and}\\
\frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))} &> \frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b_{BP}(x+y))}.
\end{aligned}$$

With $(\frac{1}{2}+x)\,b \geq (\frac{1}{2}-x)\,b$ we finally have in the current situation:

$$(\tfrac{1}{2}+x)\,b\frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))}+(\tfrac{1}{2}-x)\,b\frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))}<b \qquad \forall\ b>b_{BP}.$$

Moreover, the above arguments actually prove that $b_{BP}=p$ is the only feasible opportunity to make persuasion work: Note that for $(x,y)=(0,0)$ Bayesian Plausibility (12) holds for any possible b. This is trivial because $(x,y)=(0,0)$ means that the disclosure mechanism is completely non-informative and investors' posteriors are equivalent to their prior beliefs.

Conversely, in the case where $(x,y)\neq(0,0)$ it is easily verified that Bayesian Plausibility (to reach a high degree of transparency) holds if and only if

$$\frac{\frac{1}{2}-(y-p(x+y))}{\frac{1}{2}-(y-b(x+y))}=\frac{\frac{1}{2}+(y-p(x+y))}{\frac{1}{2}+(y-b(x+y))}$$

which requires $b=p$ to hold. In words: in the current context Bayesian Plausibility needs to be met only for an investor whose prior belief b equals the *objective* probability for a vulnerable banking sector p.

C. Bayesian Plausibility in the supervisor's optimization problem

Consider the supervisor's problem in an explicit form, i.e. including the Bayesian Plausibility constraint (BP):

$$\begin{aligned}
\min_{x,y}\ \Delta U_S \quad &\equiv \quad Pr(v)\left(U_S^{max}-U_S(1-\hat{G}_v(\tfrac{1}{2}))\right)+Pr(f)\left(U_S^{max}-U_S(1-\hat{G}_f(\tfrac{1}{2}))\right)\\
\text{with} \quad & \quad Pr(v)\hat{b}_v(p)+Pr(f)\hat{b}_f(p)=p \quad (BP)\\
& \quad Pr(v)=(\tfrac{1}{2}+x)\,p+(\tfrac{1}{2}-y)\,(1-p)\\
& \quad Pr(f)=(\tfrac{1}{2}-x)\,p+(\tfrac{1}{2}+y)\,(1-p)\\
& \quad x\leq\tfrac{1}{2} \qquad\qquad (\lambda_x)\\
& \quad y\leq\tfrac{1}{2} \qquad\qquad (\lambda_y)\\
& \quad x,y\geq 0.
\end{aligned} \tag{26}$$

Starting from the corresponding Lagrangian

$$\begin{aligned}
\mathcal{L} \ = \ & Pr(v)\left(U_S^{max}-U_S(1-\hat{G}_v(\tfrac{1}{2}))\right)+Pr(f)\left(U_S^{max}-U_S(1-\hat{G}_f(\tfrac{1}{2}))\right)+\\
& +\lambda\left[Pr(v)\hat{b}_v(p)+Pr(f)\hat{b}_f(p)-p\right]+\lambda_x\left[x-\tfrac{1}{2}\right]+\lambda_y\left[y-\tfrac{1}{2}\right]
\end{aligned}$$

and using the Kuhn-Tucker Theorem yields the following first-order necessary conditions:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x} &= p\left(U_S^{max} - U_S(1-\hat{G}_v(\tfrac{1}{2}))\right) - p\left(U_S^{max} - U_S(1-\hat{G}_f(\tfrac{1}{2})\right) + \\
&+ Pr(v)U_S'(1-\hat{G}_v(\tfrac{1}{2}))\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} + Pr(f)U_S'(1-\hat{G}_f(\tfrac{1}{2}))\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} + \\
&+ \lambda\left[p\hat{b}_v(p) + Pr(v)\frac{\partial \hat{b}_v(p)}{\partial x} - p\hat{b}_f(p) + Pr(f)\frac{\partial \hat{b}_f(p)}{\partial x}\right] \geq 0; \\
& x \geq 0\ ;\ \frac{\partial \mathcal{L}}{\partial x}x = 0 \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial y} &= -(1-p)\left(U_S^{max} - U_S(1-\hat{G}_v(\tfrac{1}{2}))\right) + (1-p)\left(U_S^{max} - U_S(1-\hat{G}_f(\tfrac{1}{2})\right) + \\
&+ Pr(v)U_S'(1-\hat{G}_v(\tfrac{1}{2}))\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y} + Pr(f)U_S'(1-\hat{G}_f(\tfrac{1}{2}))\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} + \\
&+ \lambda\left[-(1-p)\hat{b}_v(p) + Pr(v)\frac{\partial \hat{b}_v(p)}{\partial y} + (1-p)\hat{b}_f(p) + Pr(f)\frac{\partial \hat{b}_f(p)}{\partial y}\right] \geq 0; \\
& y \geq 0\ ;\ \frac{\partial \mathcal{L}}{\partial y}y = 0. \qquad (28) \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= Pr(v)\hat{b}_v(p) + Pr(f)\hat{b}_f(p) - p = 0 \qquad (29) \\
\frac{\partial \mathcal{L}}{\partial \lambda_x} &= x - \tfrac{1}{2} \leq 0\ ;\ \lambda_x \geq 0\ ;\ \frac{\partial \mathcal{L}}{\partial \lambda_x}\lambda_x = 0 \qquad (30) \\
\frac{\partial \mathcal{L}}{\partial \lambda_y} &= y - \tfrac{1}{2} \leq 0\ ;\ \lambda_y \geq 0\ ;\ \frac{\partial \mathcal{L}}{\partial \lambda_y}\lambda_y = 0. \qquad (31)
\end{aligned}
$$

Inspection of terms in square brackets in (27) and (28), which are the derivatives of (BP), shows that they both equal zero: using the explicit formulations of $Pr(v)$,

$Pr(f)$, $\hat{b}_v(p)$, $\hat{b}_f(p)$, $\frac{\partial \hat{b}_v(p)}{\partial x}$, $\frac{\partial \hat{b}_f(p)}{\partial x}$, $\frac{\partial \hat{b}_v(p)}{\partial y}$, and $\frac{\partial \hat{b}_f(p)}{\partial y}$ (see the proof of Lemma 1) yields

$$\begin{aligned}
& \left[p\hat{b}_v(p) + Pr(v)\frac{\partial \hat{b}_v(p)}{\partial x} - p\hat{b}_f(p) + Pr(f)\frac{\partial \hat{b}_f(p)}{\partial x}\right] = \\
= \; & p\left[\frac{(\frac{1}{2}+x)p + (\frac{1}{2}-y)(1-p)}{(\frac{1}{2}+x)p + (\frac{1}{2}-y)(1-p)} - \frac{(\frac{1}{2}-x)p + (\frac{1}{2}+y)(1-p)}{(\frac{1}{2}-x)p + (\frac{1}{2}+y)(1-p)}\right] = \\
= \; & p\,[1-1] = 0 \\
& \text{and} \\
& \left[-(1-p)\hat{b}_v(p) + Pr(v)\frac{\partial \hat{b}_v(p)}{\partial y} + (1-p)\hat{b}_f(p) + Pr(f)\frac{\partial \hat{b}_f(p)}{\partial y}\right] = \\
= \; & -(1-p)\left[\frac{(\frac{1}{2}+x)p - (\frac{1}{2}+x)p}{(\frac{1}{2}+x)p + (\frac{1}{2}-y)(1-p)} - \frac{(\frac{1}{2}-x)p - (\frac{1}{2}-x)p}{(\frac{1}{2}-x)p + (\frac{1}{2}+y)(1-p)}\right] = \\
= \; & -(1-p)\,[0-0] = 0
\end{aligned}$$

due to $p \in (0, 1)$. Including Bayesian Plausibility (BP) in the supervisor's optimization problem, hence, does not affect the relevant first-order necessary conditions for the optimum. Rather, calculations show that the probability distributions already comprise the crucial features of (BP).

D. Proof of Lemma 4

Consider case a) with $0 < 1 - G(\frac{1}{2}) < |P|^{max}$ first. Assume that the supervisor's decision on x and y implies $0 < 1 - \hat{G}_f(\frac{1}{2}) < 1 - G(\frac{1}{2}) < |P|^{max} < 1 - \hat{G}_v(\frac{1}{2}) < 1$. Let ΔU_S denote the supervisor's valuation of this situation with

$$\begin{aligned}
\Delta U_S \;&=\; Pr(v)\left(U_S^{max} - U_S(1 - \hat{G}_v(\tfrac{1}{2}))\right) + Pr(f)\left(U_S^{max} - U_S(1 - \hat{G}_f(\tfrac{1}{2}))\right) \\
\text{with} \quad & Pr(v) = \tfrac{1}{2} + xp - y(1-p) \\
& Pr(f) = \tfrac{1}{2} - xp + y(1-p).
\end{aligned}$$

It can be easily shown that this cannot be optimal because there exists a solution $x^* < x$ and $y^* < y$ to the supervisor's optimization problem for which the corresponding valuation ΔU_S^* is smaller than ΔU_S above.[31] In this regard note first totally differentiating equations for $Pr(v)$ and $Pr(f)$ above shows that $Pr(v)$ and $Pr(f)$ can be held constant as long as a one-unit increase of y is accompanied by a $\frac{1-p}{p}$ increase of x:

$$\begin{aligned}
dPr(v) \;&=\; pdx - (1-p)dy \overset{!}{=} 0 \;\Rightarrow\; \frac{dx}{dy} = \frac{1-p}{p} \\
dPr(f) \;&=\; -pdx + (1-p)dy \overset{!}{=} 0 \;\Rightarrow\; \frac{dx}{dy} = \frac{1-p}{p}.
\end{aligned}$$

[31] Note, the supervisor aims to minimize ΔU_S.

Second, let $\left[1-\hat{G}_f(\frac{1}{2})\right]^*$ and $\left[1-\hat{G}_v(\frac{1}{2})\right]^*$ denote the numbers of prudent investors if the supervisor reports $d=f$ and $d=v$, respectively, and sets x^* and y^*. Then starting from $0<1-\hat{G}_f(\frac{1}{2})<1-G(\frac{1}{2})<|P|^{max}<1-\hat{G}_v(\frac{1}{2})<1$ reducing x and y in the afore derived relation increases $1-\hat{G}_f(\frac{1}{2})$ and simultaneously reduces $1-\hat{G}_v(\frac{1}{2})$ such that

$$0<1-\hat{G}_f(\tfrac{1}{2})<\left[1-\hat{G}_f(\tfrac{1}{2})\right]^*<1-G(\tfrac{1}{2})<|P|^{max}<\left[1-\hat{G}_v(\tfrac{1}{2})\right]^*<1-\hat{G}_v(\tfrac{1}{2})<1$$

due to $\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x}\leq 0$, $\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y}<0$, $\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x}>0$, $\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y}\leq 0$. This, however, implies $U_S^*\left(\left[1-\hat{G}_v(\frac{1}{2})\right]^*\right)>U_S\left(1-\hat{G}_v(\frac{1}{2})\right)$ and $U_S^*\left(\left[1-\hat{G}_f(\frac{1}{2})\right]^*\right)>U_S\left(1-\hat{G}_f(\frac{1}{2})\right)$ due to $U_S'\left(1-\hat{G}_v(\frac{1}{2})\right)<0$ and $U_S'\left(1-\hat{G}_f(\frac{1}{2})\right)>0$ while $Pr(v)$ and $Pr(f)$ have been held constant. As a result we have

$$\Delta U_S^*<\Delta U_S$$

when the supervisor reduces x and y to $x^*<x$ and $y^*<y$ in the present situation.

Because the previous arguments hold as long as $1-\hat{G}_v(\frac{1}{2})>|P|^{max}$ an optimum in the case of $1-G(\frac{1}{2})<|P|^{max}$ requires $1-\hat{G}_v(\frac{1}{2})\leq|P|^{max}$.

Consider now case b) with $|P|^{max}<1-G(\frac{1}{2})<1$. Assume that the supervisor's decision on x and y implies $0<1-\hat{G}_f(\frac{1}{2})<|P|^{max}<1-G(\frac{1}{2})<1-\hat{G}_v(\frac{1}{2})<1$ with a corresponding valuation ΔU_S. Then the same arguments as have been explained in the previous case can be applied to show that there exists a pair $x^*<x$ and $y^*<y$ for which the corresponding valuation ΔU_S^* is smaller than ΔU_S, i.e. $\Delta U_S^*<\Delta U_S$. Therefore, an optimum in the case of $|P|^{max}<1-G(\frac{1}{2})<1$ requires $1-\hat{G}_f(\frac{1}{2})\geq|P|^{max}$.

E. Proof of Proposition 1

a. Non-optimality of a fully-revealing disclosure mechanism

A fully revealing mechanism is characterized by $x=y=\frac{1}{2}$.

Kuhn-Tucker conditions (18) and (19) then imply $\lambda_x>0$ and $\lambda_y>0$, respectively. Furthermore Kuhn-Tucker conditions (16) and (17) require

$$\frac{\partial\mathcal{L}}{\partial x}=0 \text{ due to } x=\tfrac{1}{2}>0$$

and

$$\frac{\partial\mathcal{L}}{\partial y}=0 \text{ due to } y=\tfrac{1}{2}>0.$$

However, relations (20) and the definitions of cumulative probability distribution functions $\hat{G}_v(\cdot)$ and $\hat{G}_f(\cdot)$ (see (24) and (25)) imply

$$\hat{G}_v(\tfrac{1}{2}) = 0 \ , \ \hat{G}_f(\tfrac{1}{2}) = 1$$

and

$$\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} = \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} = \frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y} = \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} = 0,$$

respectively, which causes $U_S(1 - \hat{G}_v(\frac{1}{2})) = U_S(1 - \hat{G}_f(\frac{1}{2})) = 0$ and therefore yields

$$\frac{\partial \mathcal{L}}{\partial x} = \lambda_x > 0 \text{ and } \frac{\partial \mathcal{L}}{\partial y} = \lambda_y > 0.$$

But this conflicts with the above requirement for optimality of $x = y = \frac{1}{2}$. As a result, in the optimum, at least one of both parameters x and y must be strictly less than $\frac{1}{2}$, and the optimal disclosure mechanism cannot be fully revealing.

b. Non-optimality of corner solutions with either $x = \frac{1}{2}$ or $y = \frac{1}{2}$

Consider the case of $x = \frac{1}{2}$ and $y \in (0, 1)$ first. Then equations (25), (1) and (20) imply $\hat{G}_f(\frac{1}{2}) = 1$, $U_S(1 - \hat{G}_f(\frac{1}{2})) = 0$ with $U_S'(1 - \hat{G}_f(\frac{1}{2})) \to \infty$ and $\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} = 0$, respectively. In addition Kuhn-Tucker condition (18) implies $\lambda_x > 0$, and optimality of $x = \frac{1}{2}$ requires

$$\frac{\partial \mathcal{L}}{\partial x} = 0$$

due to Kuhn-Tucker condition (16).

However, in the current case Kuhn-Tucker condition (16) reduces to

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= -pU_S\left(1 - \hat{G}_v(\tfrac{1}{2})\right) + \\ & +Pr(v)U_S'\left(1 - \hat{G}_v(\tfrac{1}{2})\right)\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} + Pr(f)U_S'(0)\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} + \lambda_x > 0. \end{aligned}$$

The inequality is a result of $p, Pr(v), Pr(f), U_S(1-\hat{G}_v(\frac{1}{2})), U_S'(1-\hat{G}_v(\frac{1}{2})), \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} > 0$ and the fact that although $\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} < 0$ our earlier result that $x = y = \frac{1}{2}$ cannot be optimal; $1 - \hat{G}_v(\frac{1}{2}) < 1$ and hence $U_S'(1 - \hat{G}_v(\frac{1}{2})) << \infty$ must be true. Thus $U_S'(0) \to \infty$ dominates, yielding the positive sign of $\frac{\partial \mathcal{L}}{\partial x}$ which, in turn, conflicts with the optimality requirement above.

The argument regarding the second case – i.e. $y = \frac{1}{2}$ and $x \in (0, 1)$ – is analogous: equations (24), (1) and (20) imply $\hat{G}_v(\frac{1}{2}) = 0$, $U_S(1 - \hat{G}_v(\frac{1}{2}) = 0$ with $U_S'(1) \to -\infty$

and $\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x} = 0$, respectively. In addition the Kuhn-Tucker condition (19) implies $\lambda_y > 0$, and optimality of $y = \frac{1}{2}$ requires

$$\frac{\partial \mathcal{L}}{\partial y} = 0$$

due to Kuhn-Tucker condition (17).

However, in the current case the Kuhn-Tucker condition (17) reduces to

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial y} &= -(1-p)U_S\left(1-\hat{G}_f(\tfrac{1}{2})\right)+ \\ &+Pr(v)U_S'(1)\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y}+Pr(f)U_S'\left(1-\hat{G}_f(\tfrac{1}{2})\right)\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y}+\lambda_y > 0.\end{aligned}$$

The inequality is a result of $(1-p), Pr(v), Pr(f), U_S(1-\hat{G}_f(\frac{1}{2})), U_S'(1-\hat{G}_f(\frac{1}{2})), \frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} > 0$ and the fact that the effect of $U_S'(1) \to -\infty$ is turned positive by factor $\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y} < 0$. Furthermore this latter effect is dominant because the non-optimality of $x = y = \frac{1}{2}$ (see above) implies $1 - \hat{G}_f(\frac{1}{2}) > 0$ and $U_S'(1 - \hat{G}_f(\frac{1}{2})) << \infty$. Thus $y = \frac{1}{2}$ and $x \in (0, 1)$ cannot be optimal.

In sum both parts of this proof imply that optimality of the disclosure mechanism requires $x < \frac{1}{2}$ *as well as* $y < \frac{1}{2}$. Corner solutions with either $y = \frac{1}{2}$ or $x = \frac{1}{2}$ are not optimal.

Non-optimality of an uninformative disclosure mechanism

A (completely) uninformative disclosure mechanism is characterized by $x = y = 0$.

Kuhn-Tucker conditions (18) and (19) in this situation imply $\lambda_x = \lambda_y = 0$. In addition Kuhn-Tucker conditions (16) and (17) say that $x = y = 0$ requires

$$\frac{\partial \mathcal{L}}{\partial x} > 0 \text{ as well as } \frac{\partial \mathcal{L}}{\partial y} > 0.$$

However, from the definitions of the cumulative probability distribution functions $\hat{G}_v(\cdot)$ (24) and $\hat{G}_f(\cdot)$ (25), the definitions of $Pr(v)$ and $Pr(f)$ and relations (20) one observes $\hat{G}_f(\frac{1}{2}) = \hat{G}_v(\frac{1}{2}) = G(\frac{1}{2})$ with $U_S(1-\hat{G}_v(\frac{1}{2})) = U_S(1-\hat{G}_f(\frac{1}{2})) = U_S(1-G(\frac{1}{2}))$ as well as $U_S'(1-\hat{G}_v(\frac{1}{2})) = U_S'(1-\hat{G}_f(\frac{1}{2})) = U_S'(1-G(\frac{1}{2}))$, $Pr(v) = Pr(f) = \frac{1}{2}$ and $\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial x} = -\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial x}$ as well as $\frac{\partial \hat{G}_f(\frac{1}{2})}{\partial y} = -\frac{\partial \hat{G}_v(\frac{1}{2})}{\partial y}$. Inserting this into Kuhn-Tucker conditions (16) and (17) yields

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \text{ as well as } \frac{\partial \mathcal{L}}{\partial y} = 0$$

which conflicts with the earlier stated requirement that $x = y = 0$ be optimal.

As a result an optimal disclosure mechanism cannot be (completely) uninformative. In the optimum at least one of the precision parameters x and y must be strictly positive, but both must be less than $\frac{1}{2}$.[32]

References

Allen, Franklin and Gale, Douglas (2000), "Financial Contagion." *Journal of Political Economy 108*, 1-33.

Angeletos, George M., and Pavan, Alessandro (2004) "Transparency of information and coordination in economies with investment complementarities." *American Economic Review, Papers and Proceedings 94*, pp. 91 – 98.

BCBS, Basel Committee on Banking Supervision (2011),"Basel III: A Global Regulatory Framework For More Resilient Banks and Banking Systems." *Bank for International Settlements, Basel.*

Borio, Claudio, Drehmann, Mathias and Tsatsaronis, Kostas (2012),"Stress-testing Macro Stress Tests: Does It Live Up To Expectations?" *Bank for International Settlement*, BIS Working Paper 269.

Caruana, Jaime (2011), "Financial Stability and Risk Disclosure." *Keynote Speech to the FSB Roundtable on Risk Disclosure, Basel*, December 9, 2011.

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

Diamond, Douglas (2007), "Liquidity, Banks, and Markets." *Journal of Political Economy 105*, 928-956.

Diamond, Douglas and Dybvig, Philip (1983), "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy 91*, 401-419.

[32]Note that this result allows for an optimum with only one of both parameters being zero as long as the other one is strictly positive. In such a case the disclosure mechanism is also informative because the analysis in section 2.4.2 and Appendix A above showed that even a single strictly positive precision parameter is able to generate a disclosure lottery with $1 - \hat{G}_f(\frac{1}{2}) < 1 - G(\frac{1}{2}) < 1 - \hat{G}_v(\frac{1}{2})$ and $Pr(f), Pr(v) \neq \frac{1}{2}$.

Diamond, Douglas and Rajan, Raghuram (2001), "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking." *Journal of Political Economy 109*, 287-327.

Freixas, Xavier and Rochet, Jean-Jacques (2008), *Microeconomics of Banking,* 2^{nd} edition, Cambridge, MA: MIT Press.

Gick, Wolfgang, and Pausch, Thilo (2012), "Persuasion By Stress Testing - Optimal Disclosure of Supervisory Information in the Banking Sector." Economic Research Centre of the Deutsche Bundesbank, Discussion paper 32/2012, ISBN 978-3-86558-867-8.

Goldstein, Itay, and Sapra, Haresh (2012), "Should Bank's Stress Test Results Be Disclosed? An Analysis of the Costs and Benefits," Wharton School, University of Pennsylvania, mimeo, April 2012.

Haggard, K. Stephen, and Howe, John S. (2007), "Are Banks Opaque?" Department of Economics, Finance and International Business University of Southern Mississippi, mimeo.

Horvath, Roman, and Vasko, Dan (2012), "Central Bank Transparency and Financial Stability: Measurements, Determinants, and Effects," Institute of Economic Studies, Charles University, Prague, mimeo, September 2012.

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

Kamenica, Emir, and Gentzkow, Matthew (2010), "Bayesian Persuasion - Web Appendix," avaliable at
http://faculty.chicagobooth.edu/matthew.gentzkow/research/BayesianPersuasion_appendix.pdf

Keynes, John Maynard (1936). *The General Theory of Employment, Interest, and Money.* London: Macmillan.

Morgan, Donald P. (2000), "Rating Banks: Risk and Uncertainty in an Opaque Industry." mimeo, Federal Reserve Bank of New York.

Morris, Stephen and Shin, Hyung Song (2000),"Welfare Effects of Public Information." Economic Research Centre of the Deutsche Bundesbank, Discussion paper 7/00.

Morris, Stephen and Shin, Hyung Song (2002),"Social Value of Public Information." *American Economic Review 92*(5), 1521-1534.

Morris, Stephen, Shin, Hyung Song, and Tong, Hui (2006),"Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con: Reply." *American Economic Review 96*(1), 453-455.

Peristiani, Stavros, Morgan, Donald P., and Savino, Vanessa (2010), "The Information Value of the Stress Test and Bank Opacity," Federal Reserve Bank of New York.

Petrella, Giovanni, and Resti, Andrea (2011), "Supervisors as Information Producers: Do Stress Tests Reduce Bank Opaqueness?" available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1968681

Sobel, Joel (2013), "Giving and Receiving Advice," in: *Advances in Economics and Econometrics*, D. Acemoglu, M. Arellano, and E. Dekel (eds.), forthcoming.

Stein, Jeremy (1989), "Cheap Talk and the Fed: A Theory of Imprecise Policy Announcements," *American Economic Review 79*(1), 32 - 42.

Svensson, Lars E. O. (2006), "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro Transparency, Not Con. *American Economic Review 96*(1), 448-452.

Wang, Yun (2012), "Bayesian Persuasion with Multiple Receivers." Job Market Paper, Department of Economics, University of Pittsburgh, mimeo, November 2012.