Limits of stress-test based bank regulation*

Isha Agarwal† and Tirupam Goel‡

January 14, 2021

Abstract

Stress-tests provide complementary information about banks' risk exposures. Recent empirical evidence, however, has uncovered potential inaccuracies in stress-test based assessments. We investigate the regulatory implications of these inaccuracies. Without stress-tests, the regulator cannot observe bank's type, and sets the same requirement across banks. Stress-testing provides a noisy signal about the banks' types, and enables bank specific surcharges, which can improve welfare. Yet, when stress-tests are less accurate, they distort banks' ex-ante incentives to improve their risk-return profiles. As a result, higher capital surcharges can lead banks to be more risky. The optimal surcharge depends on the accuracy of stress-tests, and can be zero.

JEL Codes: G21, G28, C61

Keywords: Capital regulation; Stress-tests; Information asymmetry; Adverse incentives; Moral hazard; Covid-19.

*The authors thank Elena Carletti, Jean-Edouard Colliard, Ingo Fender, Neil Esho, Eswar S. Prasad, Nikola Tarashev, Alberto Teguia, Kostas Tsatsaronis, Egon Zakrajšek, our referees, and seminar participants at the 2021 ASSA meetings, the BIS, and Cornell University for useful comments. Matlab programs to reproduce the results in this paper are available at https://sites.google.com/site/tirupam/. The views expressed in this paper are those of the authors and not necessarily of the Bank for International Settlements.

†Affiliation: Sauder School of Business, University of British Columbia. Email: Isha.Agarwal@sauder.ubc.ca. Declarations of interest: none.

‡Corresponding author. Affiliation: Bank for International Settlements, Centralbahnplatz 2, 4051 Basel, Switzerland. Email: Tirupam.Goel@bis.org. Phone: +41 61 280 8433. Declarations of interest: none.

1 Introduction

Stress-tests have become an important policy tool for regulators globally after the Global Financial Crisis (GFC) [Baudino et al., 2018]. They complement financial reporting and disclosures in revealing private information about banks' risk exposures to regulators [Morgan et al., 2014].[1] This enables regulators to better align baseline capital requirements to individual banks' risk profiles. In both the U.S. and the Euro Area, for instance, stress-test results are used to determine bank-specific quantitative and/or qualitative regulatory requirements.[2]

There are, however, limits to the accuracy of stress-tests, which means that regulation based on test results can be misdirected. For one, models used in the stress-test may not fully capture banks' risk exposures or their response to the crisis, and may lead to assessments that turn out to be inaccurate ex-post [Acharya et al., 2014; Philippon et al., 2017]. Relatedly, reliance on historical data to define stress scenarios may miss unprecedented events [Breuer et al., 2018]. In addition, bank-level inputs to stress-test models may be noisy or not comparable across banks [Ong et al., 2010]. Moreover, technical and computational glitches can lead to faulty results.[3] As such, stress-tests may exhibit Type-I and Type-II errors, ie a less (more) risky bank may fail (pass) the test and face excessive (insufficient) regulation.

Despite empirical evidence of potential inaccuracies in stress-testing, there is a lack of theoretical studies on the implications of these inaccuracies for optimal capital requirements. Our goal in this paper is to start filling in this gap. We develop a tractable

[1] Indeed, the GFC underscored that banks, especially the large and more complex ones, can be very opaque [Gorton, 2008], and that this can give rise to information asymmetries *viz-a-viz* the regulators.

[2] In the U.S., capital surcharges (among other requirements) are determined on the basis of stress-test results. See https://www.govinfo.gov/content/pkg/FR-2020-03-18/pdf/2020-04838.pdf for more details. In the Euro Area, stress-tests conducted by the European Banking Authority (EBA) are a crucial input into the Supervisory Review and Evaluation Process (SREP) which entails capital planning, reporting, and governance requirements tailored to individual banks. See https://eba.europa.eu/eba-launches-2020-eu-wide-stress-test-exercise for more details.

[3] For example, in September 2020, the U.S. Federal Reserve Bank published corrections to its previously issued stress-test results [Fed, 2020].

framework to study the welfare and policy implications of capital requirements based on stress-tests that are not perfect. We use the framework to also investigate the trade-offs a regulator faces in making stress-tests more accurate.

The key players in our model are a banker and a regulator. The banker runs a bank that takes deposits from the household and invests in a risky project. The return on the project can be high or low, depending on the bank's *type*, which in turn depends on the effort it exerts ex-ante. On the back of a mis-priced deposit insurance combined with limited liability, the bank over-borrows relative to the social optimal.[4] Over-borrowing increases the probability of failure (which can be costly) and poses an inefficiency from the social planner's point of view. This inefficiency rationalises a minimum capital-ratio requirement in our model, and allows us to study the welfare implications of counterfactual policies.[5]

We assume that the regulator cannot observe the bank's type, which means it cannot impose bank-specific requirements. We then consider stress-tests as a regulatory (supervisory) tool that provides a potentially inaccurate signal about the bank's type, based on which the regulator can impose a capital surcharge on top of the baseline requirement. In doing so, the regulator faces a trade-off. Stress-tests help overcome (some) information frictions and align regulation to individual banks' risk profiles. This improves welfare. Yet, inaccuracies can lead to inefficiently low or high requirements for some banks, and thus distort banks' ex-ante incentives.[6] This lowers welfare. We use the model to assess this trade-off formally.

Our main contribution is to show that under information frictions higher capital requirements can induce banks to become more risky, and to derive the attendant relationship

[4] Typical reasons for a mis-priced deposit insurance include the inability of the insurer to observe banks' risk profiles or impose risk-sensitive premium payments. See Flannery et al. [2017] for elaboration.

[5] A large literature provides several rationales for capital-ratio requirements, such as fire-sale externalities [Kara and Ozsoy, 2016], moral hazard issues [Christiano and Ikeda, 2016], implicit guarantees [Nguyen, 2015], and household preference for safe and liquid assets [Begenau, 2019]. The approach in this paper is related to that of Kareken and Wallace [1978], Santos [2001], and Van den Heuvel [2008] who show that over-borrowing, led by mis-priced deposit insurance or otherwise, justifies capital regulation.

[6] In a similar vein, Prescott [2004] shows that poorly executed supervisory audits can create adverse incentives ex-ante.

between optimal capital surcharge and stress-test accuracy. When test accuracy is below a threshold, a high-type bank can fail the test more often and face an inefficiently high requirement. This not only reduces welfare *ceteris paribus*, but because the opportunity cost of a tighter constraint is greater for a high-type bank, ex-ante incentives of the bank to exert effort towards becoming a high-type are also diminished. We show that the optimal surcharge in this case is zero. For intermediate levels of accuracy, we show that the optimal surcharge increases with accuracy, but is still smaller than what the full information benchmark would imply. In case of a perfectly accurate stress-test, any surcharge has a strong disciplining effect in terms of eliciting greater ex-ante effort from banks, and accordingly the optimal surcharge is the highest.[7]

Next we examine the problem of a regulator who jointly chooses the level of test accuracy and the optimal surcharge. We consider two cases. In the first case, there is no trade-off between reducing the false positive and false negative rates, i.e. the regulator can alter the test design and reduce at least one of the Type-I or Type-II errors. Yet it incurs a social cost related to re-designing the test and improving its accuracy (e.g. due to higher supervisory burden on both regulators and banks). We show that as the cost of improving test accuracy becomes larger, the optimal accuracy of the test as well as the surcharge for banks that fail the test must be adjusted downwards.

In the second case, we assume that it is not possible to improve one error rate without worsening the other error rate. This is typically the case when the regulator cannot improve the design of the test (say because it is prohibitively costly to do so) and can only adjust the signal-cutoff below which a bank is considered to have failed the test. We show that in this case, while a lower cost of accuracy may induce the regulator to improve the test accuracy, say by reducing the rate of false positive at the expense of a higher false negative rate, the

[7] Our paper formalises the intuition James Bullard (President of the Federal Reserve Bank of St. Louis) had in the context of quantitative easing: *while state-contingent policies are generally desirable, they work well when the states on which the policy is contingent are known.* See this article for a coverage of his remark. Relatedly, our paper supports the remarks made by Mark Zelmer (Deputy Superintendent, OSFI Canada) in 2013 in the context of risk-sensitivity of capital requirements.

optimal surcharge may be lower – thus creating a dichotomy between test accuracy and the capital surcharges for failing banks.

The regulatory trade-offs discussed above are aggravated when bank failures are more costly, such as in the case of too-big-to-fail banks. We show that in this case, not only is the optimal baseline capital requirement stricter, the optimal surcharge for a given level of accuracy is also higher.

To illustrate our analytical results, we calibrate the parameters of the model using data on U.S. banks. Numerical computations confirm our qualitative insights, and allow us to fully characterise the phase shift in the relationship between accuracy and optimal surcharge. Indeed, the surcharge is zero if the accuracy is below a threshold, and increases non-linearly with accuracy otherwise.

We conclude our discussion by alluding to potential inaccuracies in the 2020 Dodd-Frank Act Stress Test of large bank holding companies in the U.S. that was conducted right before the Covid-19 crisis. We find no correlation between banks' performance in the stress-test and their actual performance in the crisis. This observation raises the question of whether the capital surcharges imposed on banks based on the results of the stress test were too high or low for some banks.

Our paper belongs to a growing literature on bank stress-tests in the post-GFC period. Most studies in this literature have focused on the trade-offs associated with transparency and disclosure policy in stress-testing. For instance, greater disclosure can help enhance market discipline but also hamper ex-ante risk-sharing [Goldstein and Sapra, 2013; Goldstein and Leitner, 2018]. Relatedly, secrecy of stress-test models can prevent gaming but may discourage productive investment[Leitner and Williams, 2020].[8] These studies typically assume that stress-tests reveal the true risk profiles of banks.

[8] Other studies in this literature include Corona et al. [2017] who assess how bailout regime and disclosure policy interact, Orlov et al. [2018] who characterise the optimal disclosure policy for high- and low-risk banks, and Bouvard et al. [2015] who show that the optimal disclosure policy must vary along the business cycle.

A smaller strand of the stress-testing literature provides evidence of potential errors in stress-testing. Acharya et al. [2014] find that the stress-test based assessments of banks' capital adequacy are not in line either with market-data based assessments, nor with banks' actual performance during the European Sovereign Debt crisis in 2011. For the 2014 stress-test conducted by the European Banking Authority, Philippon et al. [2017] find that while model-based losses are good predictors of realized losses around announcements of macroeconomic news, banks headquartered in countries with weak banking system have higher realized losses compared with their losses predicted by the 2014 stress test. Frame et al. [2015] show that stress-tests conducted by the U.S. Office of Federal Housing Enterprise Oversight in the pre-GFC period failed to detect risks on the balance sheets of Fannie Mae and Freddie Mac. Covas et al. [2014] show that stress-test assessments can be more informative and less prone to gaming if they are based on density (instead of linear-model based point) forecasts.

Despite evidence on potential limitations of stress-tests, the literature has not formally assessed the attendant welfare and policy implications. Our paper develops a tractable framework to study these implications.[9] We derive analytical characterisations of the optimal capital surcharge when stress-tests are given to be less accurate, and when the regulator can choose stress-test accuracy. In addition, this paper discusses how the Covid-19 crisis may help shed light on the ex-post validation of stress-testing.

More broadly, our paper contributes to the literature on bank-specific regulation. For instance, Marshall and Prescott [2001] show that state-contingent fines on banks can increase welfare, but assume that the states are observable. Lohmann [1992] shows that when future states are not fully known, it is sub-optimal to commit to a state-contingent policy. By contrast, we consider a setup where the policy maker can choose the degree of

[9]Parlatore and Philippon [2018] also model stress-tests, but focus on the optimal design of stress scenarios that enable efficient information acquisition by the regulator. More generally, Morrison and White [2005] study the the effectiveness of capital regulation in avoiding crisis as public confidence in the regulator's ability to screen banks varies. Our paper, in contrast, focuses on how capital regulation must be optimally adjusted as stress-test accuracy changes.

state-contingency, and characterise the optimal degree. More recently, Ahnert et al. [2020] show that sensitivity of regulation to banks' types must depend on the precision of the attendant signal, like in our paper. Yet, while they show that starting from high precision, lower precision implies greater sensitivity of regulation to risk, we show that such a strategy can decrease welfare by creating adverse incentives ex-ante. This difference stems from the fact that we allow banks to affect the probability that they face a capital surcharge, due to which regulation can affect ex-ante incentives.

2 Model

Our goal is to analyse the welfare and policy implications of stress-test based capital requirements when stress-tests are potentially inaccurate. To this end, we develop a model with the following main elements. First is a general equilibrium setup that enables us to capture the welfare effect of regulation on the (representative) household's utility. Second is a dynamic setup that allows us to assess the effect of future stress-test and regulation on banks' ex-ante behavior. Third is a rationale for capital-regulation – specifically, a social inefficiency that warrants regulatory intervention. Fourth is information frictions – i.e., the unobservability of a bank's type by the regulator – that justify the use of stress-tests. Accordingly, we consider an economy that lasts three periods (0, 1, and 2), and consists of a representative household, a banker whose decisions are socially inefficient and whose type is stochastic, a regulator that cannot (fully) observe the bank's type, and a government that runs a deposit insurance program.

Household The household is representative, and receives an unconditional income endowment $\bar{Y}$ on dates 1 and 2. On date-1, it decides how much to consume, c_1, and how much to deposit, d, in the bank.[10] Deposits are risk-free, and pay a gross return of R on

[10] A time subscript is used only for those quantities that are relevant on multiple dates. For instance, since d is only chosen once, on date-1, a time subscript is omitted.

date-2.

Banker The banker has a capital endowment of k on date-1. It runs a bank that issues deposits d to invest $k + d$ in a risky project that pays $\psi g(k + d)$ on date-2. $g(.)$ is a decreasing returns to scale (DRS) return function. ψ is an investment shock whose density f_s depends on the banker's type s on date-1, which can be high (H) or low (L). Specifically, we assume that while both types face the same standard deviation of ψ, namely σ, the high-type bank has a higher expected return, $\mu_H > \mu_L$, so that a high-type bank has a higher *risk-adjusted return.*The probability p with which the bank is of high-type depends on the effort e the banker exerts on date-0. The cost of exerting effort is $\zeta(e)$.

The bank's deposit liabilities on date-2 equal Rd, and thus the net cash-flow n equals $\psi g(k+d) - Rd$. When ψ is sufficiently high and the bank is solvent, the entire cash-flow is paid as dividends to the banker. However, when ψ is low enough so that the cash-flow is negative, the bank fails and banker receives null. We assume that the banker only consumes on date-2, and that it has limited liability, so that it cannot be asked for additional capital to rescue a failing bank. Instead, the government takes the bank into receivership.

Government The government runs the deposit insurance scheme and ensures that depositors are fully protected against bank failure. When a bank fails, the government takes its assets into custody, liquidates the same, and covers any shortfall in the failed bank's liabilities. To fund the scheme, the government imposes a tax lumpsum T on the household. We assume that the insurance scheme is mis-priced – ie insensitive to the risks banks take – which, as we prove later, leads to a social inefficiency.[11] The government runs a balanced budget.

[11]The reason for introducing an inefficiency in our model is to rationalise capital requirements. A mis-priced deposit insurance is not the only way to do so, but it is a relatively simple method that helps keep our model tractable. Another paper to have taken this route is Van den Heuvel [2008]. A moral hazard between banks and its creditors [Gertler and Kiyotaki, 2010], or implicit government guarantees [Nguyen, 2015] are among the several other ways in which capital requirements can be justified.

Regulator The regulator is benevolent, i.e. it strives to maximise the joint welfare of the household and the banker. On date-0, it announces the minimum capital-ratio requirement χ that the bank must satisfy on date-1. However, we assume that the regulator cannot *observe* the bank's type on date-1.[12] In the baseline economy, as such, it must announce a requirement that does not depend on banks' types, i.e. applies universally to both types of banks on date-1. In the economy with stress-tests, the regulator is able to obtain a noisy signal about the bank, and classify it as a high- or low-type depending on whether it passes or fails the test. The regulator then announces a surcharge x for failing the stress-test, effectively imposing a bank-type specific requirement $\chi_s, s \in \{H, L\}$.

Recursive formulation We now formally setup the problem statements of the agents in the economy. The household chooses d on date-1 to maximize its expected utility over dates 1 and 2:

$$U = \max_{d} \quad c_1 + \beta \mathbb{E} c_2 \quad s.t. \quad c_1 = \bar{Y} - d \quad and \quad c_2 = \bar{Y} + Rd - T. \tag{1}$$

The banker chooses e on date-0 which determines the probability of being an H-type on date-1:

$$[Date - 0]: \quad \max_{e} \quad -\zeta(e) + \beta\Big(p(e)V_H(\chi_H) + (1 - p(e))V_L(\chi_L)\Big). \tag{2}$$

where $V_s(\chi_s)$ is defined in equation (3). The bank of type $s \in \{H, L\}$ chooses d on date-1 to maximize the expected dividend it pays on date-2:

$$[Date - 1]: \quad V_s(\chi_s) = \max_{d} \quad \beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\underbrace{\psi g(k + d) - Rd}_{n}\Big) f_s(\psi) d\psi \quad s.t. \quad \frac{k}{\chi_s} \geq d. \tag{3}$$

[12] In reality, regulators do have some knowledge about banks' characteristics (such as via regulatory filings). We assume that the observable characteristics are embedded in the return function $g(.)$ of the bank while *type* simply summarizes the unobservable characteristics. Furthermore, we assume that the bank cannot credibly communicate its type to the regulator, except via its performance in a stress-test.

The lower limit on the integral is the ψ cut-off – call it ψ_c – below which the bank fails (and dividends n equal zero). χ_s are the bank-type-specific minimum capital-ratio requirements (although the requirements will be the same in the absence of stress tests). The government's budget constraint is as follows:

$$T = \begin{cases} Rd - \psi g(k+d) & \text{If the bank fails i.e. } \psi \leq \frac{Rd}{g(k+d)} \\ 0 & \text{Otherwise} \end{cases} \tag{4}$$

3 Qualitative Analysis

We begin by assessing the equilibrium conditions in the baseline economy. We then characterise – as a benchmark – the optimal regulation in the absence of stress tests. Finally, we analyse the optimal capital surcharge based on stress-test results, including when bank failure is socially costly.

3.1 The competitive equilibrium

The first-order condition (FOC) of the bank's problem on date-0 shows that the effort the bank exerts depends on the *wedge*, say ω, between the value of being a high- as opposed to low-type on date-1:

$$-\zeta'(e) + \beta p'(e) \underbrace{\Big(V_H(\chi_H) - V_L(\chi_L)\Big)}_{\omega} = 0 \tag{5}$$

To see how the effort changes as the wedge increases, we take the total derivative of Equation 5 with respect to ω, from where it is straightforward to note Lemma 1:

$$-\zeta''(e)\frac{de}{d\omega} + \beta p''(e)\omega\frac{de}{d\omega} + \beta p'(e) = 0 \tag{6}$$

Lemma 1. *If $\zeta(.)$ is increasing and convex, and $p(.)$ is increasing and concave, then the* *bank exerts more effort when the difference in the value of being a high type compared to a low type increases, i.e.* $de/d\omega > 0$.

It's intuitive to see from equation 5 why effort would increase with the wedge ω. The optimal choice of effort trades off the marginal benefit of effort with its marginal cost. As the relative value of being a high-type bank increases, the marginal benefit of effort increases while the marginal cost is unaffected by the relative difference between the value of the high- and low-type bank. Hence, the bank optimally chooses to increase effort.

The pre-conditions for Lemma 1 to hold are sufficient but not necessary. For instance, the result still holds if $\zeta(.)$ is increasing and linear. Nonetheless, that becoming a high-type bank is increasingly difficult is a realistic assumption to have.

Lemma 1 underscores an important insight. The minimum requirements (χ_H, χ_L) announced on date-0 affect the wedge ω by impacting the value of the bank on date-1. As such, the minimum requirements are a key factor in bank's effort choice on date-0, and will shape the regulator's choice of optimal ex ante capital requirement as we show later in Section 3.3.[13]

As regards the date-1 FOCs, we have the following:

$$\text{Bank:} \qquad \beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi - \Lambda_s = 0 \tag{7}$$

$$\text{Household:} \qquad R = 1/\beta \tag{8}$$

Note in the bank's FOC that Λ_s is the Lagrange multiplier on the regulatory constraint, and that two of the three terms which arise from a routine application of the Leibniz rule are equal to zero. The system of FOCs (5), (7), (8) and the government's budget constraint (4) together characterise the competitive equilibrium of the model economy for a given set of minimum capital-ratio requirements (χ_H, χ_L).

[13]Lemma 1 is related to a similar result proven in Christiano and Ikeda [2016], but the channel through which regulation has an impact on the banker's effort is different in the two papers.

3.2 Optimal ex-post regulation

We now assess the efficiency of the competitive equilibrium, and discuss the role that regulation could play in improving welfare. In this section, we focus on the date-1 economy, and turn to the date-0 economy (and the discussion of stress tests) in the next subsection.

Inefficiency of the competitive equilibrium We compare allocations in an unregulated date-1 economy with a benevolent social planner's allocations. Without loss of generality, we focus on the case of an s-type bank, where s could be high or low. As such, for now we ignore the effect banker's effort on date-0 has on overall welfare (and return to this consideration later).

We consider a constrained social planner that maximizes the date-1 and date-2 equally weighted welfare of the household and the banker by choosing the level of deposit-based funding on behalf of the banker, taking as given the household's first order condition:

$$\max_{d} \quad c_1 + \beta \mathbb{E}(c_2 + n) \quad s.t. \quad R = 1/\beta; \quad c_1 = \bar{Y} - d; \quad c_2 = \bar{Y} + Rd - T$$

Recall that the banker does not consume on date-1, and note that $c_2 + n$ denotes the combined consumption of the household and the banker on date-2. Since the planner internalises the effect of choosing d on n and T, we can solve for $c_2 + n$ using expressions for n and T from equations (3) and (4) respectively:

$$c_2 + n = \bar{Y} + \psi g(k + d) \tag{9}$$

Next, we rewrite the planner's objective after plugging in the expressions for c_1, c_2, n, rearranging terms using the household's FOC, and segregating the expectation (i.e. the integral on c_2) at the ψ cutoff for failure of the bank:

$$\max_d \quad (1+\beta)\bar{Y} + \underbrace{\beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g(k+d) - Rd\Big) f_s(\psi) d\psi}_{\text{Banker's date-1 objective}} + \beta \int_0^{\frac{Rd}{g(k+d)}} \Big(\psi g(k+d) - Rd\Big) f_s(\psi) d\psi. \tag{10}$$

By segregating the integral into two parts, the first part matches the bank's objective function, and thus facilitates a comparison of bank's and planner's FOCs, as shown below:

$$\beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi + \underbrace{\beta \int_0^{\frac{Rd}{g(k+d)}} \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi}_{\text{Bank-failure inefficiency}} = 0. \tag{11}$$

Equation (11) uncovers a wedge between the planner's FOC and the bank's FOC in the unregulated economy (i.e. equation (7) with $\Lambda_s = 0$). This wedge stems from limited liability and a mis-priced deposit insurance. Because of limited liability, the bank does not internalise the losses corresponding to the left tail of the distribution of ψ – the part that corresponds to bank failure. And because of deposit insurance, the depositors do not charge a premium for risk of non-repayment of deposit principle plus interest in full post bank failure. The bank thus over borrows. The planner, on the contrary, chooses the level of deposits taking into account the entire distribution of ψ. We refer to this wedge as the bank-failure inefficiency, which the following lemma characterises.

Lemma 2. *The bank's capital ratio, defined as k/d, is smaller in the competitive equilibrium as compared to that in the constrained planner's problem, i.e. the second best.*

Proof. Assume that the inefficiency term is positive. Then, $\beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi$ must also be positive. But this is a contradiction since the overall expression for the planner's FOC must equal zero. As such, the inefficiency term must be negative. In turn, this implies that $\beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi > 0$. We know that d^{CE} (the level of deposits in the competitive equilibrium) satisfies $\beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi = 0$. But since $g(.)$ is concave, it must be that $d^{CE} > d^*$ where d^* solves the constrained planner's problem.[14] ■

Implementability of the constrained efficient allocation That the competitive equilibrium exhibits an inefficiency implies that $W^{CE} \leq W^*$ where W^{CE} is the welfare in the competitive equilibrium and W^* is the second-best welfare. The question that follows is whether a regulatory intervention can help implement or approach the second best.

To this end, we consider a benevolent regulator that sets a minimum capital-ratio requirement $k/d \geq \chi_s$ on the bank in order to maximize welfare. In choosing χ_s, the regulator faces the following trade-off. A higher χ_s forces the bank to reduce deposit-based funding and accordingly its failure probability, which has a welfare improving effect due to a smaller bank-failure inefficiency. Yet, a higher χ_s depresses expected output, which has a welfare reducing effect.

In effect, the regulator's decision problem is very similar to that of a constrained planner. This is because choosing deposits on behalf of the bank to maximise welfare is equivalent to imposing a minimum capital-ratio requirement with the same objective when capital is fixed and the requirement is binding. This is formally seen by comparing equations (7) and (11). Indeed, the first terms are identical. And to the extent the Lagrange multiplier Λ_s on (i.e. the shadow cost of) the regulatory constraint in (7) is equal to the absolute value of the bank-failure inefficiency term in (11), the solution to the two equations must be identical. We note this result in the lemma below, and denote the optimal regulation for an s-type bank by χ_s^o.

Lemma 3. *The solution to the constrained planner's problem can be implemented via a minimum capital-ratio requirement.*[15]

[14] The finding that the bank takes more leverage than what is socially optimal is not unique to this paper, nor is it our main contribution. Several other studies have related findings, such as Van den Heuvel [2008] and Christiano and Ikeda [2016], for instance. Our approach is to develop a relatively parsimonious model that has the mechanisms needed to study the welfare effects of stress-test based capital requirements.

[15] A capital-ratio requirement is not the only regulatory tool that can implement the second best. A

Before turning to the date-0 problem, we document a result that will be useful later. It compares the optimal date-1 regulation for high- and low-type banks.

Lemma 4. *The regulator optimally sets a higher ex-post requirement on the low-type bank as compared to a high-type bank, i.e.* $\chi^o_L > \chi^o_H$.

Proof. Consider the non dis-aggregated version of the planner's date-1 FOC – i.e. equation (11) – for both high- and low-type banks. This characterises the optimal level of deposits in each case.

$$0 \;=\; \int_0^\infty \Big(\psi g'(k+d) - R\Big) f_s(\psi) d\psi \;=\; \mu_s g'(k+d) - R \quad s \in \{H, L\} \tag{12}$$

The total derivative of d with respect μ_s implies:

$$g'(k+d) + \mu_s g''(k+d)\frac{\partial d}{\partial \mu_s} = 0 \implies \frac{\partial d}{\partial \mu_s} > 0 \quad s \in \{H, L\} \tag{13}$$

This immediately implies that the optimal d is higher, or equivalently, the optimal χ^o is lower for a high-type bank. ■

Intuitively, for a given level of deposits, a low-type bank not only generates lower expected output, but is also more likely to fail. This underpins the stricter regulation for the low-type bank.

3.3 Optimal ex-ante regulation

The bank forms expectations and chooses its date-0 decisions based on date-1 requirements announced by the regulator on date-0.[16] However, because the bank's type on date-1 is its private information, the regulator cannot announce a type-specific requirement (such

tax (or a deposit insurance premium) that is a function of the balance sheet choice of the bank may also achieve the same objective.

[16]We abstract away from time-inconsistency issues, and assume that regulatory announcements are credible.

as $\mathcal{X}_L^o$ and $\mathcal{X}_H^o$ for low- and high-type banks respectively). As a result, the regulator must adopt a bank-type independent capital requirement – say $\mathcal{X}$ – which is applicable on date-1 irrespective of the bank's type. To characterize the optimal $\mathcal{X}$, we begin with the following result.

Lemma 5. *Assume that regulation $\mathcal{X}$ binds for both bank types on date-1. Then the effort the bank exerts on date-0 decreases as $\mathcal{X}$ rises.*

Proof. As shown in Lemma 1, the bank's date-0 effort e depends on $\omega = V_H(\mathcal{X}) - V_L(\mathcal{X})$, i.e. the wedge between the value of being a high- versus low-type on date-1. The key then to proving this lemma is to characterise how regulation impacts ω.

$$\omega = \beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g(k+d) - Rd\Big) f_H(\psi) d\psi - \beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g(k+d) - Rd\Big) f_L(\psi) d\psi$$

where $d = k/\mathcal{X}$. The derivative of ω with respect to $\mathcal{X}$ gives:

$$\frac{\partial \omega}{\partial \mathcal{X}} = -\frac{k\beta}{\mathcal{X}^2} \left(\underbrace{\int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_H(\psi) d\psi}_{\Lambda_H} - \underbrace{\int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f_L(\psi) d\psi}_{\Lambda_L} \right) \tag{14}$$

where Λ_s is the Lagrange multiplier on the regulatory constraint in the bank's problem.

To sign this expression, we proceed as follows. First note that since the regulatory requirement is the same for both types of banks, their deposit choices and thus the failure cutoffs ψ_c are also the same. Then let $\hat{F}_H$ and $\hat{F}_L$ be the distribution functions of ψ for high- and low-type banks, truncated below at ψ_c. Since $\mu_H > \mu_L$ (while the variances are the same), $\hat{F}_H$ *FOSD* $\hat{F}_L$, that is $\hat{F}_H(\psi) \leq \hat{F}_L(\psi) \ \forall \ \psi$. Finally, since $(\psi g'(k+d) - R)$ is an increasing function of ψ, it follows that:

$$\int \Big(\psi g'(k+d) - R\Big) d\hat{F}_H(\psi) - \int \Big(\psi g'(k+d) - R\Big) d\hat{F}_L(\psi) = \Lambda_H - \Lambda_L > 0.$$[17]

In turn, this implies that $\frac{\partial \omega}{\partial \chi} < 0$. Then from Lemma 1 we know that $\frac{\partial e}{\partial \omega} > 0$, which completes the proof since:

$$\frac{\partial e}{\partial \chi} = \frac{\partial e}{\partial \omega}\frac{\partial \omega}{\partial \chi} < 0.$$

■

Lemma 5 captures a key insight of this paper. Because capital requirement is more costly for a high-type bank because of higher loss of output following restrictions on raising deposits, an increase from a given level of requirement leads to a greater decline in the expected value of the high-type bank than a low type bank. This, in turn, lowers the returns to exerting more effort. In contrast to the conventional wisdom that more skin-in-the-game via higher capital requirement can induce banks to become safer, our finding is that under information frictions banks might respond to stricter regulation by becoming more risky.

This insight thus points to an important trade-off the regulator faces while setting χ. Compared to no regulation ($\chi = 0$), a higher χ can improve welfare *ex-post* by mitigating some of the inefficiency associated with the bank's choices, especially in case of a low-type bank. Yet, a higher χ can reduce welfare due to its adverse impact on effort exerted *ex-ante*.

Before characterising the optimal ex-ante regulation, we note that the assumption that regulation binds for both bank types is not critical for Lemma 5. The case where regulation

[17] To prove this formally, consider continuous distribution functions G and H such that $\forall x, H(x) \leq G(x)$, and define $y(x) = H^{-1}(G(x))$. Then for any increasing function $w(x)$, $\int w(y(x))dH(y(x)) = \int w(y(x))dG(x)$. Next, note that $y(x) = H^{-1}(G(x)) \implies y(x) \geq x$ since $\forall x, H(x) \leq G(x)$. In turn, since $w(.)$ is an increasing function, $w(y(x)) \geq w(x)$. Thus, $\int w(y(x))dG(x) \geq \int w(x)dG(x)$. Indeed, intuitively, the *shadow cost* of the minimum capital-ratio constraint should be greater for a bank whose assets are *ceteris paribus* more profitable.

binds for only one bank – which has to be the high type bank since it chooses higher leverage in the unregulated economy and also since $\mathcal{X}^o_L > \mathcal{X}^o_H$ – leads to the same result because in that case $\Lambda_L = 0$. The case where regulation does not bind for any bank is not relevant nor interesting because we already showed that an inefficiency rationalises *some* regulation.

Proposition 1. *The optimal ex-ante requirement $\mathcal{X}^o$ in the case where the regulator cannot observe the bank's type, is saddled by the optimal ex-post requirement for low- and high-type banks, ie $\mathcal{X}^o_L \geq \mathcal{X}^o \geq \mathcal{X}^o_H$.*

Proof. The problem of a benevolent regulator on date-0 when it cannot impose bank-specific requirements, is as follows:

$$\max_{\mathcal{X}} \quad \beta p(e) U_H(\mathcal{X}) + \beta(1 - p(e)) U_L(\mathcal{X}) - \zeta(e)$$

Here U_s is the household's and banker's combined expected lifetime consumption utilities when the banker turns out to be of type s, while $\zeta(e)$ accounts for the banker's effort on date-0. We will prove the proposition via the method of contradiction. Let $\mathcal{X}^o$ solve the above problem. Then, if $\mathcal{X}^o > \mathcal{X}^o_L > \mathcal{X}^o_H$, it means that the requirement is more strict than the optimal requirement for both bank types, and thus a lower $\mathcal{X}^o$ would improve welfare in case of each bank type, as well as the total expected welfare. Similarly, if $\mathcal{X}^o_L > \mathcal{X}^o_H > \mathcal{X}^o$, it means that the requirement is more liberal than the optimal requirement for both bank types, and thus a higher $\mathcal{X}^o$ would improve total welfare. ■

Intuitively, this proposition shows that when there is information asymmetry, the regulator chooses a *middle-ground* relative to the optimal bank-type specific requirements.

3.4 Mitigating information frictions via stress-tests

Stress test allows the regulator to gather information about banks' types. It thus helps mitigate some information frictions and allows capital requirements to be better aligned to the banks' types. This is desirable as it can improve welfare.

We incorporate stress tests in our model as follows (see Figure 1 for the timeline). We assume that the test delivers a noisy signal η to the regulator about the bank's type on date-1. The signal distribution Q_H of high-type banks dominates (in the first order stochastic (FOSD) sense) the signal distribution Q_L of low-type banks. Depending on its preferences for true- and false- positive and negative rates, the supervisor uses a signal cutoff η^c above (below) which the bank is considered pass (fail) and is deemed to be of the high- (low-) type. Thus the probability that a high-type bank passes the test is given as $q_H = 1 - Q_H(\eta^c)$, and the same for a low-type bank is given as $q_L = 1 - Q_L(\eta^c)$. Moreover, $Q_H \succcurlyeq_{FOSD} Q_L \implies q_H > q_L$.[18]

The accuracy of the stress-test is fully captured by the tuple (q_H, q_L). Any test can thus be represented by a point in the set $[0,1] \times [0,1]$, as shown in Figure 2. In this format, $(1 - q_H)$ denotes the 'false positive' or Type-I error rate (high-type bank fails the test), while q_L is the 'false negative' or Type-II error rate (low-type bank passes the test). A convenient benchmark, which is equivalent to the full-information case, is when $q_H = 1$ and $q_L = 0$, i.e. a *perfect* stress-test that exactly identifies the type of the bank. In all other cases, we refer to the test as *imperfect* because an H-type bank can fail the test ($q_H < 1$) or an L-type bank can pass the test ($q_L > 0$).

The regulator uses the outcome of the stress-test to adjust the baseline capital requirement χ^o. We assume that a bank that passes the stress test is deemed high-type and is allowed to operate at χ^o. A failed bank is deemed to be of the low-type, and the regulator strives to align the capital-ratio requirement towards $\chi_L^o \geq \chi^o$ by imposing a surcharge $x \geq 0$.

[18] We do not model, and instead take as a given, the signal distribution and the regulator's preferences for Type-I and Type-II error rates that jointly map into the signal cutoff η^c and thus the pass probabilities. This mapping is based on the receiver operating characteristics (ROC) curve of the signal, and is standard in the literature. In section 3.5 we assess the implications of a change in the signal cutoff for the optimal policy, and also the case where the regulator can incur a cost and improve the overall accuracy of the stress-test – i.e. increase the area under the ROC curve.

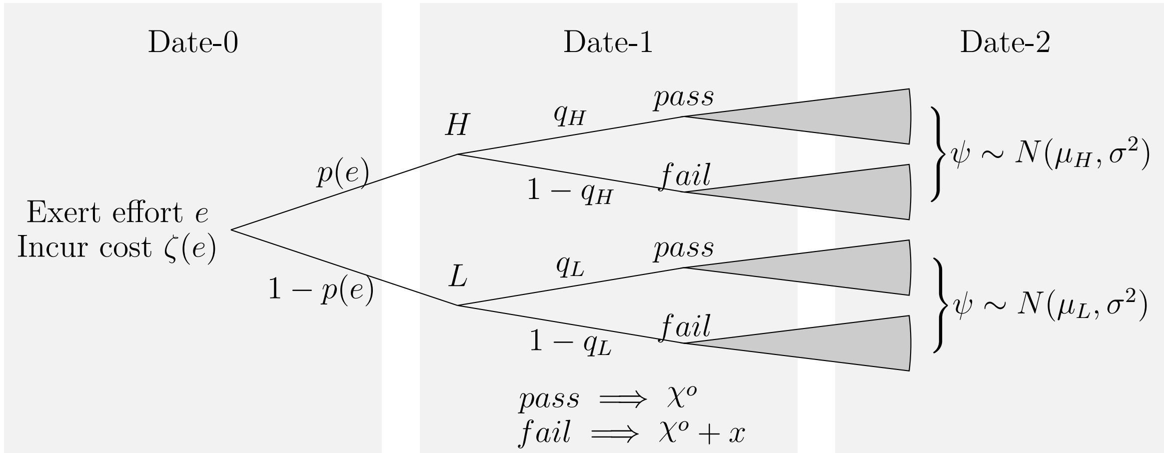

Figure 1: The timeline of events when there is information asymmetry about the bank's type, and stress tests serves as a tool to (partially) mitigate information frictions.

Our capital surcharge specification follows the spirit of stress-testing in practice. In the U.S., for instance, banks that perform poorly in the test (in terms of their capital ratios under stress) must satisfy a higher capital surcharge, while those that do well are not given any relief and must continue to satisfy the baseline requirements. Consistent with this observation, we assume that the surcharge x cannot be negative.[19]

The core question of interest then is as follows: what is the welfare maximising level of surcharge x that the regulator must announce on date-0. The choice of x is non-trivial, and encompasses a three-way trade-off.

1. In case of the low-type bank, the surcharge (upon failing the test) **increases** welfare *ceteris paribus* as long as $x \leq \chi_L^o - \chi^o$. This is because the surcharge brings the ex-ante requirement $(\chi^o + x)$ closer to the ex-post optimal (χ_L^o).

2. In case of the high-type bank, the surcharge (upon failing the test) **decreases** welfare *ceteris paribus*. This is because $\chi^o + x > \chi^o \geq \chi_H^o$, as a result of which the surcharge takes the ex-ante requirement away from the ex-post optimal.

[19]Note that in practice, most jurisdictions no longer assign an official pass or fail grade to the banks. Yet, whether a bank has done well or not in absolute (i.e. pass or fail) terms can still be inferred from its capital ratio under stress and the minimum capital-ratio requirement in the economy.

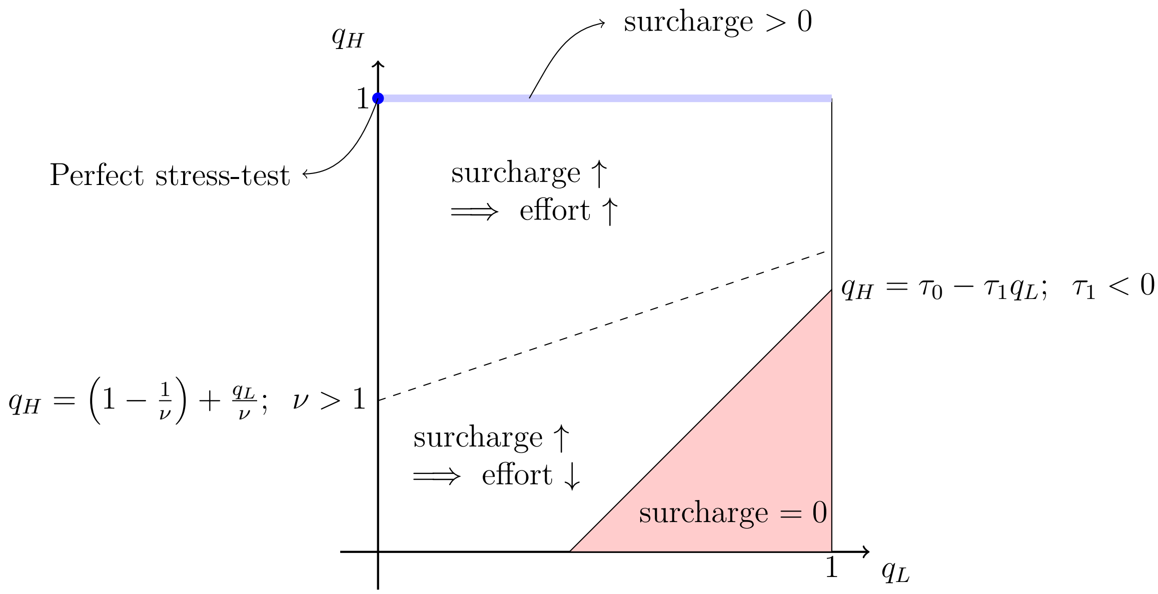

Figure 2: **Stress-test accuracy, effect on ex-ante effort, and optimal penalties:** Each point on the unit square characterises the accuracy of a stress-test. A higher q_H and a lower q_L indicate a more accurate test. The optimal surcharge is zero if accuracy is below the solid diagonal line (i.e. in the shaded area), and positive if $q_H = 1$. Effort increases with the surcharge above the dotted diagonal line, and decreases otherwise.

3. The surcharge affects the wedge between the expected value of being high- versus low-type on date-1, and thus impacts the bank's behaviour on date-0. Depending on the accuracy of the stress test, this can lead to an increase or decrease in the bank's effort. We prove this result in Lemma 6 below. Accordingly, *ceteris paribus*, a higher surcharge can **increase or decrease** welfare through its effect on effort.

Lemma 6. *The bank's effort may increase or decrease with a surcharge, depending on the accuracy of the stress test.*

Proof. The date-0 problem of the bank is:

$$\max_{e} \quad -\zeta(e) + \beta p(e) \underbrace{\left(q_H V_H(\mathcal{X}^o) + (1-q_H) V_H(\mathcal{X}^o + x)\right)}_{\mathbb{E}V_H} + \beta(1-p(e)) \underbrace{\left(q_L V_L(\mathcal{X}^o) + (1-q_L) V_L(\mathcal{X}^o + x)\right)}_{\mathbb{E}V_L} \tag{15}$$

We begin by noting that similar to the case without stress testing, the effort the bank exerts increases with the *expected* value function wedge $\omega = \mathbb{E}V_H - \mathbb{E}V_L$. Taking the derivative of ω with respect to x at $x = 0$ gives:

$$\left.\frac{\partial \omega}{\partial x}\right|_{x=0} = (1 - q_H)V_H'(\mathcal{X}^o) - (1 - q_L)V_L'(\mathcal{X}^o)$$

where V' indicates the derivative of the value function. To determine the sign of this expression, divide everything by $V_L'(\mathcal{X}^o)$:[20]

$$sgn\left(\left.\frac{\partial \omega}{\partial x}\right|_{x=0}\right) = -sgn\left((1 - q_H)\underbrace{\frac{V_H'(\mathcal{X}^o)}{V_L'(\mathcal{X}^o)}}_{\nu} - (1 - q_L)\right)$$

Next, recall from the proof of Lemma (5) that $V_H'(\mathcal{X}^o) - V_L'(\mathcal{X}^o) < 0$, which implies that $\nu > 1$ since $V_H'(\mathcal{X}^o) < 0$ and $V_L'(\mathcal{X}^o) < 0$. Thus, the effect of surcharge on the bank's effort choice depends on the accuracy of the test as follows:

$$(1 - q_L) - (1 - q_H)\nu \begin{cases} > 0 & \implies \text{efforts increases with surcharge} \\ = 0 & \implies \text{efforts does not change with surcharge} \\ < 0 & \implies \text{effort decreases with surcharge} \end{cases}$$

■

Intuitively, ν captures the relative shadow cost of tightening regulation for the high- and low-type banks. *Ceteris paribus*, a higher ν makes imposing a surcharge less desirable by making it more likely that the bank reduces effort. Similarly, for a given ν, a higher Type-I (i.e. lower q_H) or Type-II error rate (higher q_L) would make $(1 - q_L) - (1 - q_H)\nu$ more negative and cause the bank to reduce effort following a higher surcharge. Indeed,

[20] Since the value of a more regulated bank is lower, $V_L'(\mathcal{X}^o) < 0$. As such, we add a minus sign to the RHS expression.

if a high-type bank is sufficiently likely to fail the stress-test and the low-type bank is sufficiently likely to pass, then the high-type bank will often face a surcharge while the low-type bank will not, thereby reducing the relative benefit of being a high-type bank, which the wedge measures. This will induce the bank to exert less effort in the first place. Relatedly, it is clear from Lemma 6 that with a perfect stress test, i.e. when $(q_H = 1, q_L = 0)$, effort increases with surcharge. And that with an imperfect stress test, such as when $q_H = q_L = 0.5$, effort decreases with surcharge. We indicate these insights qualitatively (i.e., not to scale) in Figure 2.

Next we assess the relationship between accuracy of the stress-test and the optimal surcharge.

Proposition 2. *No surcharge must be imposed if the accuracy of stress testing as measured by a (well-defined) linear combination of the Type-1 and Type-II error rates is higher than a cutoff.*

Proof. Welfare as a function of the surcharge x can be written based on the regulator's problem as follows (note that e also depends on x in this expression):

$$\max_{x} \quad W(x) = \beta p(e)\Big(q_H U_H(\mathcal{X}^o) + (1-q_H)U_H(\mathcal{X}^o + x)\Big) +$$

$$\beta(1-p(e))\Big(q_L U_L(\mathcal{X}^o) + (1-q_L)U_L(\mathcal{X}^o + x)\Big) - \zeta(e)$$

Our goal is to identify 'a' non-trivial set of (q_H, q_L) where $W(0) > W(x)\ \forall\ x > 0$, i.e. a zero surcharge is optimal.[21] A sufficient condition for this to be the case is $W'(x) < 0\ \forall\ x > 0$. To this end, we consider the first-order condition of the regulator's problem:

$$\frac{dW}{dx} = p'(e)e'(x)\Big(q_H U_H(\mathcal{X}^o) + (1-q_H)U_H(\mathcal{X}^o + x)\Big) + p(e)(1-q_H)U_H'(\mathcal{X}^o + x) -$$

[21]Our goal is to not fully characterise the set of (q_H, q_L) for which the optimal surcharge is zero. We only wish to show that with low-enough accuracy, imposing a surcharge is sub-optimal.

$$p'(e)e'(x)\Big(q_L U_L(\mathcal{X}^o) + (1-q_L)U_L(\mathcal{X}^o + x)\Big) + (1-p(e))(1-q_L)U'_L(\mathcal{X}^o + x) - \zeta'(e)e'(x)$$

To characterise the sign of this expression, we make a few assumptions, again with the goal to find *sufficient* conditions under which the optimal surcharge is zero.

– First we assume that $x \in [0, \mathcal{X}^o_L - \mathcal{X}^o]$. The upper bound corresponds to a surcharge amount that results in a requirement for the low-type banks that is equal to the ex-post optimal requirement $\mathcal{X}^o_L$. In principle, the optimal surcharge could be higher (due to its effect on improving ex-ante effort), but that would entail a welfare decreasing effect in case of both high- and low-type banks.

– Second, we assume that (q_H, q_L) are such that the effort exerted by the bank decreases as surcharge increases (as per Lemma 6).

Next, since $U_s(\mathcal{X}^o + x), s \in \{L, H\}$ is a concave function of x, $\mathcal{X}^o_L \geq \mathcal{X}^o \geq \mathcal{X}^o_H$ implies the following: (i) $U_H(\mathcal{X}^o) \geq U_H(\mathcal{X}^o + x)$; (ii) $U'_H(\mathcal{X}^o + x) \leq 0$; (iii) $U_L(\mathcal{X}^o) \leq U_L(\mathcal{X}^o + x)$; and (iv) $U'_L(\mathcal{X}^o + x) \geq 0$; $\forall\, x \in [0, \mathcal{X}^o_L - \mathcal{X}^o]$. It then follows that:

$$\frac{dW}{dx} \leq p'(e)e'(x)U_H(\mathcal{X}^o) + p(e)(1-q_H)U'_H(\mathcal{X}^o) - p'(e)e'(x)U_L(\mathcal{X}^o) +$$

$$(1-p(e))(1-q_L)U'_L(\mathcal{X}^o) - \zeta'(e)e'(x)$$

Finally, we re-arrange and set the right-hand-side expression to zero:

$$p(e)U'_H(\mathcal{X}^o) + (1-p(e))U'_L(\mathcal{X}^o) - p(e)q_H U'_H(\mathcal{X}^o) - (1-p(e))q_L U'_L(\mathcal{X}^o) +$$

$$\underbrace{p'(e)e'(x)\Big(U_H(\mathcal{X}^o) - U_L(\mathcal{X}^o)\Big)}_{A<0} - \zeta'(e)e'(x) = 0$$

$$\implies \underbrace{\frac{A}{p(e)U'_H(\mathcal{X}^o)} + 1 + \frac{(1-p(e))U'_L(\mathcal{X}^o)}{p(e)U'_H(\mathcal{X}^o)} - \zeta'(e)e'(x)}_{\tau_0 <> 0} - q_L \underbrace{\frac{(1-p(e))U'_L(\mathcal{X}^o)}{p(e)U'_H(\mathcal{X}^o)}}_{\tau_1 < 0} = q_H$$

$$\implies q_H = \tau_0 - \tau_1 q_L \tag{16}$$

In equation (16), while the slope is positive, the intercept can be positive or negative, depending on the underlying parameters. Also, when $q_L = 1$, $q_H < 1$. The equation implies that when $q_H < \tau_0 - \tau_1 q_L$ the surcharge should be zero, as also indicated in Figure 2. ■

Intuitively, the proposition shows that when q_H is low and/or q_L is high – both of which reflect a relatively less accurate stress-test – the surcharge must be zero. Next we explore conditions under which the optimal surcharge can be strictly positive.

Consider a stress-test that is accurate in identifying high-type banks i.e. $q_H = 1$, but is possibly inaccurate in identifying low-type banks i.e. $1 > q_L \geq 0$. In this case a higher x does not affect $\mathbb{E}V_H$, but decreases $\mathbb{E}V_L$ (recall equation (15)). As a result, the banker increases effort as surcharge increases. Second, consider the regulator's problem:

$$\max_x \quad \beta p(e) U_H(\chi^o) + \beta(1 - p(e))\Big(q_L U_L(\chi^o) + (1 - q_L) U_L(\chi^o + x)\Big) - \zeta(e)$$

A higher x does not affect welfare when the bank passes the test, but increases welfare when it fails the test as long as $x \leq \chi_L^o - \chi^o$ (recall from Proposition 4 that beyond this threshold, the effective requirement on the low-type bank is higher than the optimal requirement χ_L^o.)

Combining the effect of a surcharge on effort e and $U_L(\chi^o + x)$, both of which increase as x increases, and given that $U_H(\chi^o) > U_L(\chi^o)$, it is clear that welfare, ignoring the effect of the surcharge on the cost of effort, must increase as x rises above zero. Thus, if the cost of effort is sufficiently small, the optimal surcharge must be strictly positive. Together with proposition 2, this insight points to a phase shift in the relation between optimal surcharge and stress-test accuracy, with the optimal surcharge being zero (positive) if the level of accuracy of the stress tests is sufficiently low (high).

In what follows, we formalise this insight using a simpler version of the model where the probability that a bank is of a given type is fixed.[22]

Proposition 3. *The optimal surcharge increases with stress-test accuracy.*

Proof. The regulator's problem in this case is as follows:

$$\max_{x} \quad \beta p\Big(q_H U_H(\mathcal{X}^o) + (1-q_H)U_H(\mathcal{X}^o + x)\Big) + \beta(1-p)\Big(q_L U_L(\mathcal{X}^o) + (1-q_L)U_L(\mathcal{X}^o + x)\Big)$$

The first order condition is:

$$[x] \quad 0 = p(1-q_H)U_H'(\mathcal{X}^o + x) + (1-p)(1-q_L)U_L'(\mathcal{X}^o + x)$$

Next consider an increase in accuracy via a higher q_H (the proof in case of a lower q_L is similar):

$$0 = -pU_H'(\mathcal{X}^o + x) + p(1-q_H)U_H''(\mathcal{X}^o + x)\frac{\partial x}{\partial q_H} + (1-p)(1-q_L)U_L''(\mathcal{X}^o + x)\frac{\partial x}{\partial q_H}$$

Since U is concave, and $U_H'(\mathcal{X}^o + x)$ is negative (because $\mathcal{X}^o$ is higher than the optimal requirement for the high-type bank), $\frac{\partial x}{\partial q_H} > 0$. ∎

3.5 Endogenous accuracy

Thus far, we consider the accuracy of the stress test – as summarised by (q_H, q_L) – to be given exogenously. In reality, regulators may be able to influence or even choose the level of accuracy, and may prefer to set it at a high level given the welfare gains it entails. Yet, there may be constraints in choosing a high degree of accuracy.[23]

For one, by subjecting banks to a harder test – one that entails a more severe crisis

[22] A similar result cannot be proven analytically in the fully specified model. Although numerical simulations show that the result also holds in the fully specified model.

[23] See Parlatore and Philippon [2018] for the trade-offs associated with the design of stress tests.

scenario for example – the regulator may be able to lower the false negative rate, and yet, the false positive rate may surge. Then, in order to reduce the false positive rate, the test may have to become more comprehensive and intrusive, which is likely to be more costly not just for the regulator, but also for the banks. Indeed, a more extensive review of the banks' balance sheet and model would not only require additional supervisory force, but also more bank employees dedicated to satisfying regulation and attending to supervision. Moreover, there may be fundamental constraints to designing a more accurate stress-test – indeed, predicting bank performance in a hypothetical scenario rests on a number of assumptions, and is an inherently hard endeavor.

To examine the trade-offs in this case, we consider the problem of a regulator who jointly chooses surcharge x and a test-design parameter $y \geq 0$ that maps to the pass probability of the high-type bank: $q_H(y) \uparrow 1$ as $y \to \infty$, while keeping q_L fixed (the other case where q_L is adjusted can be handled similarly). We assume that greater accuracy by adjusting the design of the test y entails a social cost $C(y) = \gamma_c y$. This setup leads to the following result.

Proposition 4. *The regulator increases stress-test accuracy q_H as well as the surcharge for failing banks as the cost of accuracy decreases.*[24]

Proof. The regulator's problem in this case is as follows:

$$\max_{x,y} \quad \beta p\Big(q_H(y)U_H(\mathcal{X}^o)+(1-q_H(y))U_H(\mathcal{X}^o+x)\Big)+\beta(1-p)\Big(q_LU_L(\mathcal{X}^o)+(1-q_L)U_L(\mathcal{X}^o+x)\Big)-\gamma_c y$$

The first order conditions are:

$$[x] \quad 0 = p(1-q_H(y))U'_H(\mathcal{X}^o+x) + (1-p)(1-q_L)U'_L(\mathcal{X}^o+x)$$

$$[y] \quad 0 = \beta p q'_H(y)(U_H(\mathcal{X}^o) - U_H(\mathcal{X}^o+x)) - \gamma_c$$

[24] A similar result holds in case where q_L can be chosen while q_H is fixed.

Next consider an increase in the cost of accuracy γ_c. A total derivative of the FOCs leads to:

$$[x]: \quad 0 = p\Big((1-q_H(y))U''_H(\mathcal{X}^o+x)\dot{x} - q'_H(y)\dot{y}U'_H(\mathcal{X}^o+x)\Big) +$$
$$(1-p)\Big((1-q_L)U''_L(\mathcal{X}^o+x)\dot{x}\Big)$$
$$[y]: \quad 0 = \beta p\Big(q''_H(y)\dot{y}(U_H(\mathcal{X}^o) - U_H(\mathcal{X}^o+x)) - q'_H(y)U'_H(\mathcal{X}^o+x)\dot{x}\Big) - 1$$

where $\dot{y} = \frac{\partial y}{\partial \gamma_c}$ and $\dot{x} = \frac{\partial x}{\partial \gamma_c}$. The first total derivative implies that $\dot{x}$ and $\dot{y}$ are of the same sign since U is concave, $U'_H < 0$, and $q'_H > 0$. This means that accuracy and surcharge go hand in hand. To assess the direction of change, consider the second total derivative, and replace $\dot{x}$ using the first total derivative. This results in the following expression:

$$\dot{y}\Big(\beta p q''_H(y)(U_H(\mathcal{X}^o) - U_H(\mathcal{X}^o+x)) - \frac{\beta(pq'_H(y)U'_H(\mathcal{X}^o+x))^2}{p(1-q_H(y))U''_H(\mathcal{X}^o+x) + (1-p)(1-q_L)U''_L(\mathcal{X}^o+x)}\Big) = 1$$

The coefficient on $\dot{y}$ is negative following the second-order sufficiency optimality condition (i.e. negative-definite Hessian matrix). This implies that $\dot{y} < 0$, and from the above discussion, that $\dot{x} < 0$. ■

Intuitively, higher accuracy along one or both dimensions of the stress-test reduces the likelihood that a high-type bank is penalised, and this in turn mitigates the adverse incentives that a capital surcharge can generate. As such, a higher surcharge is optimal.

Yet, in practice, it may not be possible to improve accuracy along one dimension (say q_H) while keeping the other fixed (i.e. q_L) or not lower it. For instance, one way to increase q_H, i.e. reduce false positives, would be to set the failure cutoff less conservatively – this would make a good bank less likely to fail. But given that there are limits to how comprehensive supervision can be (due to technical, financial, and logistical reasons), a less conservative failure cutoff may lead to, at the same time, a higher false negative rate – i.e. more bad banks passing the test. Relatedly, making the test harder to reduce the false negative rate might lead more good banks to fail and increase the false positive rate.

To obtain the regulatory implications in this more practical case, we assume that in order to increase q_H (reduce false positive rate), not only does the regulator (and potentially the bank) incur a higher cost, but also that q_L increases i.e. false negative rate increases (or vice versa). In this case, we find that a lower cost of increasing q_H does induce the regulator to reduce the false positive rate (while increasing the false negative rate), but it does not necessarily lead the bank to choose a higher x, creating a dichotomy between accuracy and surcharge (see proof in the Appendix).

All in all, our analysis suggests that stress-test design and the subsequent capital surcharge decisions are intricately linked, and must inform each other. This is especially since higher accuracy along one dimension does not necessarily imply the room to impose a higher surcharge on banks.

3.6 Failure costs

Failure of a bank can impose a social cost. This cost can stem from, for instance, forced sale of a failed bank's assets, as well as due to resolution related expenses. It can be a major cost in the case of large banks (due to contagion/knock-on effects), when the resolution framework is not well functioning, or during a crisis when many banks are in insolvency at the same time.

Failure costs exacerbate the trade-off for regulators. A higher surcharge (compared to the case without failure costs) may be justified on the grounds that it lowers the expected failure rate and attendant social costs. Yet, to the extent the stress test is not sufficiently accurate, a higher surcharge would not only lower welfare in the case of a high type bank, but also would lower the ex-ante effort exerted by the bank. As such, it is not obvious as to whether the surcharge must be adjusted upwards or downwards in the presence of failure costs.

To formally assess the effect of failure cost on optimal regulation, we adapt the model as follows. We assume that once a bank fails, the recovery value of its assets is less that a hundred percent. This cost – denoted Δ – is borne by the deposit insurance program and is funded via taxes:

$$T(\psi) = \begin{cases} Rd - \psi g(k+d)(1-\Delta) & \text{If the bank fails i.e. } \psi \leq \frac{Rd}{g(k+d)} \\ 0 & \text{Otherwise} \end{cases}$$

In what follows, we prove that the failure cost exacerbates the inefficiency banks pose, and rationalises a higher ex-post requirement χ^o and also a higher ex-ante surcharge x associated with failing the stress test.

We begin by assessing the ex-post requirement, while abstracting away from bank-type as before. Household and banker consumption on date-2 in this case is given as:

$$c_2 + n = \bar{Y} + \psi g(k+d) - \Delta \psi g(k+d) \mathbb{1}\left(\psi \leq \frac{Rd}{g(k+d)}\right).$$

Accordingly, the planner's problem is:

$$\max_d \quad (1+\beta)\bar{Y} + \beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g(k+d) - Rd\Big) df(\psi) + \beta \int_0^{\frac{Rd}{g(k+d)}} \Big(\psi g(k+d) - Rd - \Delta \psi g(k+d)\Big) df(\psi),$$

while the attendant first-order-condition is:

$$\begin{aligned} 0 \;=\; & \beta \int_{\frac{Rd}{g(k+d)}}^{\infty} \Big(\psi g'(k+d) - R\Big) f(\psi) d\psi \;+ \\ & \underbrace{\beta \int_0^{\frac{Rd}{g(k+d)}} \Big(\psi g'(k+d)(1-\Delta) - R\Big) f(\psi) d\psi - \beta \Delta \psi g(k+d) \frac{\partial \frac{Rd}{g(k+d)}}{\partial d} f\left(\frac{Rd}{g(k+d)}\right)}_{\text{Bank-failure inefficiency}} \end{aligned} \tag{17}$$

We know from the discussion of equation (11) that the inefficiency term in that equation, namely $\beta \int_0^{\frac{Rd}{g(k+d)}} \Big(\psi g'(k+d) - R\Big) f(\psi) d\psi$, is negative. This means that the left term in the second row of equation (17) is also negative, and even lower in value. At the same time, since $g(.)$ is concave:

$$\frac{\partial \frac{Rd}{g(k+d)}}{\partial d} = \frac{R\Big(g(k+d) - dg'(k+d)\Big)}{g(k+d)^2} > 0.$$

As such, the inefficiency term in equation (17) is negative and larger in magnitude relative to the inefficiency term in equation (11). Thus failure cost amplifies the bank-failure inefficiency. In turn, as shown in Lemma 3, greater inefficiency rationalises a higher minimum capital-ratio requirement. We note this result in Lemma 7.

Lemma 7. *The regulator must optimally impose a higher ex-post minimum capital-ratio requirement on a bank that, all else equal, exhibits a higher failure cost.*

Next we examine how the optimal surcharge must change as failure cost increases. Unfortunately, it is not possible to characterise the change generally. For analytical tractability, we continue to assume that the probability of being a high-type (or equivalently low-type) bank is given and that there is no effort choice involved. The regulator's problem is as follows:

$$\max_x \quad W(x) = \beta p\Big(q_H U_H(\mathcal{X}^o, \Delta) + (1-q_H)U_H(\mathcal{X}^o + x, \Delta)\Big)$$
$$\beta(1-p)\Big(q_L U_L(\mathcal{X}^o, \Delta) + (1-q_L)U_L(\mathcal{X}^o + x, \Delta)\Big)$$

Here Δ in the utility function formally expresses the dependence of welfare on failure costs. The attendant first-order condition is as follows, where the D_i operator indicates the derivative with respect to the i^{th} argument of U:

$$p(1-q_H)D_1U_H(\mathcal{X}^o + x, \Delta) + (1-p)(1-q_L)D_1U_L(\mathcal{X}^o + x, \Delta) = 0$$

Next, we take the total derivative of this expression with respect to Δ:

$$p(1-q_H)\Big(D_{11}U_H(\mathcal{X}^o + x, \Delta)\frac{dx}{d\Delta} + D_{12}U_H(\mathcal{X}^o + x, \Delta)\Big) +$$

$$(1-p)(1-q_L)\Big(D_{11}U_L(\chi^o+x,\Delta)\frac{dx}{d\Delta}+D_{12}U_L(\chi^o+x,\Delta)\Big)=0$$

$$\implies -\underbrace{\Big(p(1-q_H)D_{11}U_H(\chi^o+x,\Delta)+(1-p)(1-q_L)D_{11}U_L(\chi^o+x,\Delta)\Big)}_{A}\frac{dx}{d\Delta}=$$

$$p(1-q_H)D_{12}U_H(\chi^o+x,\Delta)+(1-p)(1-q_L)D_{12}U_L(\chi^o+x,\Delta)$$

Since both U_H and U_L are concave functions of x, $A<0$. To sign the RHS, consider $U_s, s=\{H,L\}$:

$$U_s(\chi^o+x,\Delta)=\bar{Y}-d+\beta g(k+d)(\mu_s-\Delta\int_0^{\frac{Rd}{g(k+d)}}\psi f_s(\psi)d\psi)\quad where\quad d=\frac{k}{\chi^o+x}$$

$$\implies D_2U_s(\chi^o+x,\Delta)=-\beta g(k+d)\int_0^{\frac{Rd}{g(k+d)}}\psi f_s(\psi)d\psi)$$

$$\implies D_{21}U_s(\chi^o+x,\Delta)=-\beta\Bigg(g'(k+d)\frac{dd}{dx}\int_0^{\frac{Rd}{g(k+d)}}\psi f_s(\psi)d\psi+$$

$$g(k+d)\frac{d}{dd}\left[\frac{Rd}{g(k+d)}\right]\frac{Rd}{g(k+d)}f_s\left(\frac{Rd}{g(k+d)}\right)\frac{dd}{dx}\Bigg)$$

As x increases, d decreases i.e. $\frac{dd}{dx}<0$. Also, as d increases, the upper limit on the integral is increases (recall $g(.)$ is concave), which means that by application of Leibniz rule, $D_{21}U_s(\chi^o+x,\Delta)>0$. Since U is a continuous function in both its arguments, $D_{21}U_s(\chi^o+x,\Delta)=D_{12}U_s(\chi^o+x,\Delta)>0$ for both $s=H,L$. This immediately leads to the following Proposition.

Proposition 5. *Assuming $p(e)\equiv p$, the optimal surcharge must increase as Δ increases.*

Relaxing the assumption that $p(e)=p$ does not lead to a general result, that is, $\frac{dx}{d\Delta}$ cannot be signed unless the specific values of the parameters of the model are known. As such, we pursue this more general case in the quantitative analysis. Nonetheless, the above proposition suggests that if the stress test is sufficiently accurate so that effort e and thus the probability of being a high-type bank increase as the surcharge increases, then it is

Parameter	Description	Value	Target moments	Value
α	Payoff exponent: $(\text{k+d})^{\alpha}$	0.914	Gross Return on risk-adjusted assets	10.19%
μ	Mean of ψ	1.336	Equity capital to assets ratio	10.38%
σ	Standard-deviation of ψ	0.102	Value-at-risk threshold	1%
$\bar{Y}$	Household income	117.8	Household savings rate	7.32%
β	Discount factor	0.99	Deposit interest-rate	1%
Δ	Failure cost	0.22	US bank failure losses	22%

Table 1: Parameter values and target moments. Bank micro-data are sourced from Fitch, US household savings rate from FRED, and bank failure losses from FDIC. Note that the last two parameters and target moments have a one-to-one mapping (i.e. they need not be estimated jointly), and that without loss of generality k is normalised to unity. The value of the moments in data are exactly match with those implied by the mode.

likely that the surcharge must be optimally adjusted upwards as the failure cost increases.

4 Numerical illustration

We now calibrate the model parameters using data on U.S. commercial banks. Our goal is not to pursue a quantitative analysis of the model or draw empirical predictions, but provide a relevant numerical illustration of our analytical results. To this end, we set the parameters such that model generated moments are equal to the corresponding data moments (see Table 1). We focus on the post-GFC to pre-Covid period – i.e. 2010-2019 – to abstract away from any crisis led large movements in the data.

We consider the following moments as targets. First is the pooled mean of return on risk-weighted assets, while taking into account interest as well as non-interest income. Dividing by risk-weighted assets (instead of just assets) helps align the moment condition with the interpretation of high- and low-type bank in our model (recall that high- and low-type banks have the same standard deviation of return on assets, and vary only in terms of the mean return on assets). Second is the pooled mean of equity capital to assets ratio. Third is a typical regulatory or bank-management imposed value-at-risk threshold of 1%. Fourth is the household savings rate, defined as the average savings of US households out of their personal disposable income during 2010-2019. Next, we set the interest rate to

1% – a standard value in the literature. Finally, Δ is set in line with the losses associated with bank failures in the US during 2010-2019. According to the Federal Deposit Insurance Commission (FDIC), there have been 367 bank failures during this period, and the median estimated loss is about 21% of the failed bank's assets, while the attendant inter-quartile range is 13% to 30%. Our target moment is the mean, which is 22%.

As regards the functional forms, we assume the cost of exerting effort by the bank on date-0 as $\zeta(e) = \gamma_e e^2$, and the attendant probability of the bank becoming a high-type on date-1 as $p(e) = 1 - 1/(1 + e)$. γ_e is chosen so that the probability that the bank is of high-type is close to fifty percent. The exact functional forms do not matter for our qualitative results as long as $\zeta(.)$ is (weakly) convex and $p(.)$ is concave. As regards μ_H and μ_L, we assume a symmetric perturbation of 50 basis points around μ. Finally, we treat q_H and q_L as free parameters that we conduct comparative statics with respect to.

Optimal ex-post regulation We begin by analyzing the impact of a minimum capital-ratio requirement on the bank's behavior and overall welfare on date-1. Without loss of generality, we consider a high-type bank. Starting from the unregulated economy, a higher minimum capital-ratio requirement forces the bank to deleverage (first panel in Figure 3). This reduces the failure probability (second panel), but also lowers expected output (third panel). The overall effect – one that weighs welfare gains from lower bank failure against the welfare loss from lower expected output – is an inverted U-shaped welfare profile as a function of χ. This finding is consistent with Lemmas 2 and 3 where we showed that the unregulated equilibrium is sub-optimal and that a minimum capital-ratio requirement can improve welfare, and also with the broader literature (e.g. Begenau [2019], Christiano and Ikeda [2016]).

Relatedly, as bank failure costs increases, not only is the optimal requirement higher (as proven in Lemma 7), the welfare gain from regulation is also higher (see left-hand panel in Figure 4).

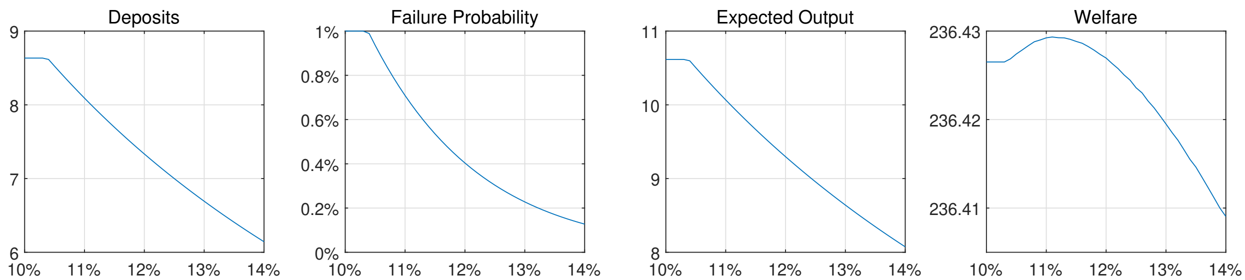

Figure 3: The effect of minimum capital-ratio requirement (x-axis) on the high-type bank and on overall welfare.

Finally, we compare the optimal ex-post requirement for low- and high-type banks. Consistent with Lemma 4, we find that the requirement is higher for the low-type bank (see right-hand panel of Figure 4, dotted lines).

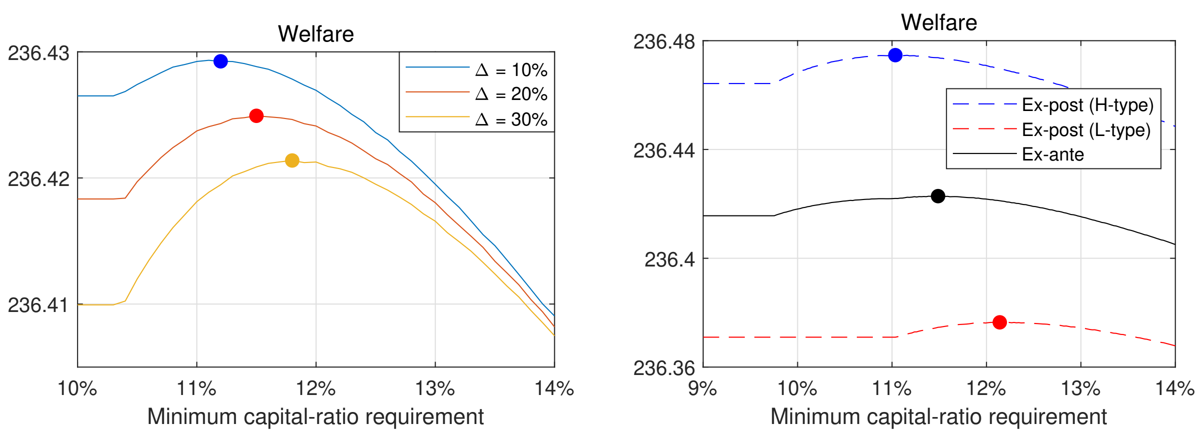

Figure 4: *Left-hand panel*: The welfare maximizing regulation for varying levels of bank failure costs. *Right-hand panel*: Optimal ex-post requirement depending on bank type, and the optimal ex-ante requirement in the absence of stress tests.

Optimal ex-ante regulation When the regulator cannot observe banks' types ex-post, the optimal ex-ante requirement announced on date-0 cannot be bank-type specific. Consistent with Proposition 1, we find that it is saddled by the ex-post optimal requirements (see solid line in the right-hand panel of Figure 4).

Next we assess how a stress-test led surcharge affects bank's behavior. A higher surcharge decreases the value of both high- and low-type banks (left-hand panel of Figure 5). The decrease is starker for a high-type bank – indeed the opportunity cost of not being able to use its balance sheet capacity is higher for a bank whose assets have a higher return. And as long as the stress test is not fully perfect, both $\mathbb{E}V_H$ and $\mathbb{E}V_L$ decrease as x increases.

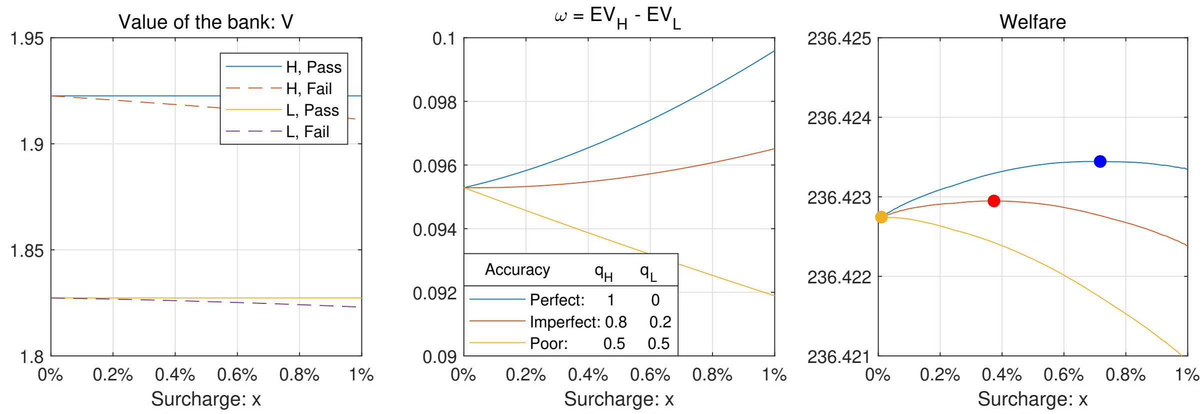

Figure 5: *Left-hand panel:* The value V of the bank in various cases as a function of the surcharge. *Centre panel:* The expected value wedge changes in response to the surcharge for different levels of accuracy of the stress test. *Right-hand panel:* Optimal surcharge.

The difference between $\mathbb{E}V_H$ and $\mathbb{E}V_L$, namely ω – as we showed in Lemma 6 – can increase or decrease depending on the accuracy of the test (see centre panel of Figure 5). This immediately means that the effort banks exert can also increase or decrease as the surcharge is raised (recall that e depends on ω; see the proof of Lemma 5). This is a key insight of the paper – a higher surcharge may not necessarily act as a disciplining device if the basis on which the surcharge is imposed is not sufficiently accurate.

Overall, the optimal surcharge depends on the following trade-off. Penalizing banks that fail the stress tests can improve welfare to the extent a low-type bank is penalised. As such, a sufficiently inaccurate test may not increase expected welfare. Moreover, in this case, banks may reduce the effort they exert. We confirm this insight quantitatively. For very low level of accuracy, consistent with proposition 2, the optimal surcharge is zero (right-hand panel of Figure 5). For higher levels of accuracy, including the case of a perfect stress test, the optimal surcharge is higher (recall Proposition 3).

We illustrate the optimal surcharge for each accuracy level of the stress-test in the left-

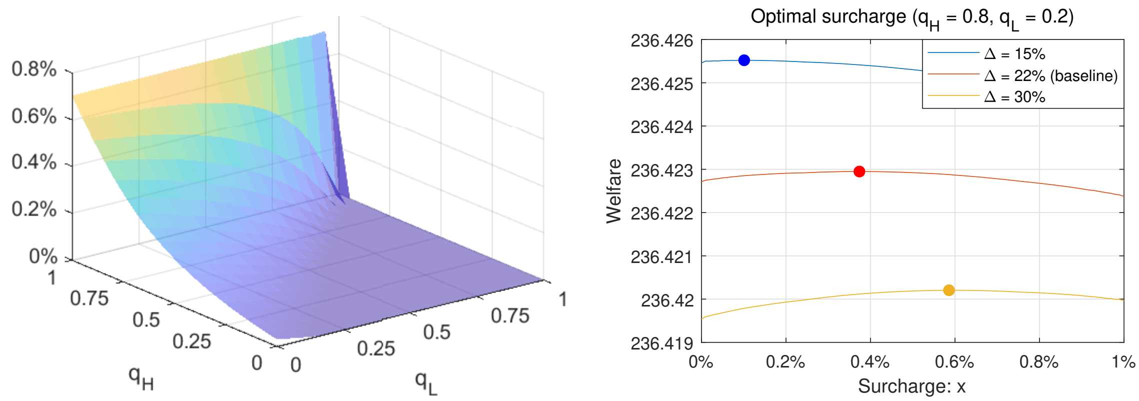

Figure 6: *Left-hand panel*: Optimal surcharge as a function of the accuracy of the stress-test. *Right-hand panel*: Change in optimal surcharge as the cost of failure increases.

hand panel of Figure 6, thus confirming the broad indications sketched in Figure 2. Indeed, a phase shift is evident: for sufficiently low levels of accuracy, the optimal surcharge is zero. Moving closer to a perfect stress test ($q_H = 1, q_L = 0$) increases the size of the optimal surcharge.

Next we elaborate upon the result proven in Proposition 5, and show that as the failure cost increases, the optimal surcharge also increases (see the right-hand panel of Figure 6).

Finally we illustrate the implications of endogensing accuracy (recall the setup in subsection 3.5). We assume that $C(y) = \gamma_c y$, $q_H(y) = 1 - \gamma_q/(1+y)$, and $q_L(y) = 1 - q_H(y)$ (see left-hand panel of Figure 7 for an example). Consistent with the analytical result in proposition 4, we find that as the cost of accuracy decreases ($\gamma_c \downarrow$), the regulator must optimally work with more accurate stress tests ($y \uparrow$), and at the same time, revise upwards the surcharge it imposes on failings banks ($x \uparrow$) (see right-hand panel of Figure 7).

5 The Covid-19 crisis: A test of stress-tests?

In this section, we review supervisory stress-tests in the U.S., and explore whether the Covid-19 crisis can provide insights about potential inaccuracies in stress-testing.

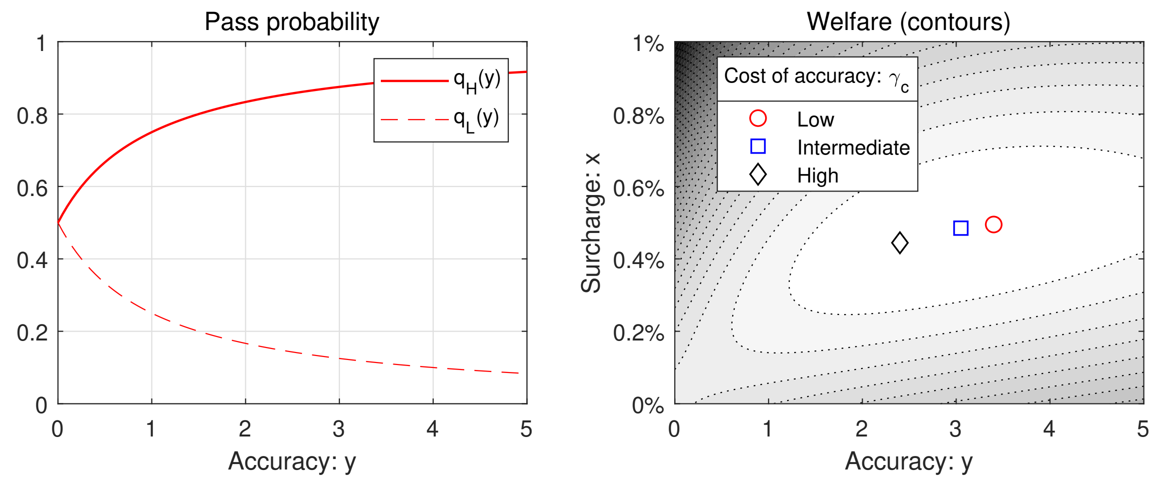

Figure 7: *Left-hand panel:* Pass probabilities for high and low-type banks as a function of accuracy. *Right-hand panel:* The jointly optimal accuracy and surcharge for different levels of cost of accuracy γ_c. Welfare contours correspond to the intermediate level of γ_c.

5.1 Institutional Background

The Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act) was enacted in response to the Global Financial Crisis (GFC). It requires the U.S. Federal Reserve Bank (Fed) to conduct an annual stress-test – known as the Dodd Frank Act Stress Test (DFAST) – of large bank holding companies (BHCs).[25] The goal is to evaluate whether the tested entities have sufficient capital to absorb losses resulting from adverse economic conditions. The DFASTs have evolved quite a bit over time.

The DFAST considers a hypothetical severely adverse scenario – one in which the U.S. economy experiences a significant recession and financial market stress while other major economies also experience contraction in economic activity – and projects the revenues, expenses, losses, and, crucially, the capital ratios of the participating banks. The projections are generated using inputs provided by the tested banks and forecasting models developed or selected by the Fed. The projections use a standard set of capital action assumptions that entail zero common stock dividend distribution, and no issuance or repurchase of

[25]Non-bank financial companies designated by the Financial Stability Oversight Council (FSOC) for Fed supervision are also included in the exercise.

common or preferred stock.[26]

The Federal Reserve imposes a capital surcharge on banks based on their performance in the test, as measured by the projected decline in their Common Equity Tier 1 (CET1) capital ratios in the severely adverse scenario.[27] Banks that perform poorly face a higher Stressed Capital Buffer (SCB) – a surcharge on top of the baseline capital requirements and any other surcharges (such as the G-SIB surcharge) – among other qualitative and quantitative requirements.

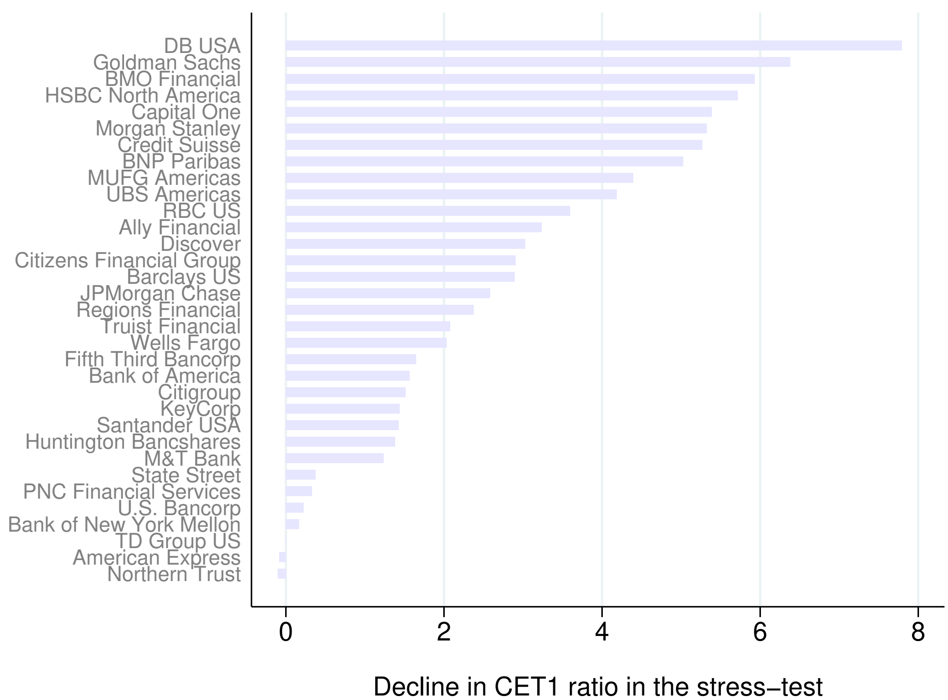

Figure 8: Decline in the CET1 ratio in the 2020 DFAST. Unit of the x-axis is percentage points.

[26]Scheduled dividend, interest, or principal payments that qualify as additional tier 1 capital or tier 2 capital are assumed to be paid, but repurchases of these instruments is assumed to be zero.

[27]The results also contain other capital ratios, namely the tier 1 and total capital ratios, and the tier 1 and supplementary leverage ratios. We focus our attention on the CET1 ratio since it is a core measure of capital adequacy, and also because capital surcharge is expressed in these terms.

5.2 The 2020 Stress-Test

The severely adverse scenario in the 2020 DFAST comprised of a peak unemployment rate of 10 percent, a decline in real GDP of 8.5 percent, and a drop in equity prices of 50 percent through the end of 2020, among other macroeconomic developments.[28] Thirty-three entities participated in the test and the results were published in June 2020, which included projections for the decline in the CET1 ratio relative to the end-2019 value (see Figure 8). The average decline was 2.7 percentage points, with Deutsche Bank USA and the Goldman Sachs Group being the worst performers, and Northern Trust and American Express being the best performers.

Performance in the stress-test and the attendant capital surcharge have a strong and positive relationship.[29] Beyond the minimum SCB of 2.5%, the two go hand-in-hand to a large extent (see Figure 9). In fact, for some banks (eg DB USA, MUFG, HSBC, RBC) the SCB is equal to decline in CET1 ratio in the stress-test. This observation suggests that a bank's performance – especially poor performance – in the test is tightly linked to the capital requirement it faces.

5.3 Bank performance during the Covid-19 crisis

In this subsection, we compare banks' performance in the 2020 DFAST with their actual performance in the Covid-19 crisis and discuss if we can learn something about potential inaccuracies in stress testing. Specifically, we compare the decline in CET1 ratios of banks in the test with that observed during the first half of 2020.[30] Several factors make the

[28]The severely adverse scenario was designed in late 2019 and was published in February 2020. While the 2020 DFAST did not adapt the severely stress scenario to incorporate the Covid-19 crisis, it disclosed additional information about predicted aggregate losses in the banking sector based on a sensitivity analysis viz-a-viz the Covid-19 crisis. Bank-level results from this exercise were not disclosed.

[29]This is despite other considerations that go into calibrating the SCB.

[30]While the pandemic was already underway at the end of Q1 2020, Q1 CET1 ratios are unlikely to reflect the effect of the shock. As such, we use Q2 CET1 ratios as a better indicator of the effect of the shock. Basing the analysis on Q3 CET1 ratios does not affect our qualitative conclusions. Moreover, an issue with Q3 data is that it is likely to nest even more (compared to Q2 data) the effect of the various liquidity and regulatory relief packages central banks adopted, whereas our aim is to identify the raw effect

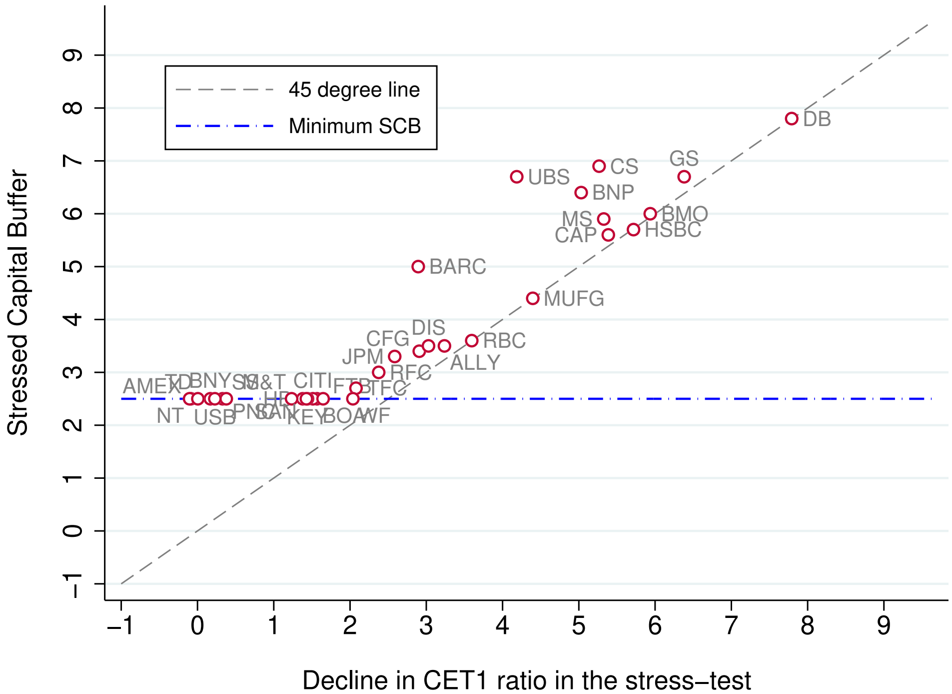

Figure 9: A comparison of the decline in CET1 ratio in the 2020 DFAST and the Stressed Capital Buffer (SCB) imposed on banks. Unit of both axes is percentage points.

Covid-19 crisis a useful natural experiment to appraise the 2020 US stress-test, but there are caveats too.

For one, while several of the key macroeconomic indicators (such as GDP, employment, and stock prices) line up well in the stress scenario and during Covid-19[31] some indicators, like the house price index, move in opposite directions. The similarities facilitate the comparison of the decline in a bank's capital ratio in the test and during the crisis, but the differences render the comparison less meaningful. Indeed a bank may be better positioned to handle specific aspects of the Covid-19 shock (house prices in this case), so that its *actual* and *test* performances may not be comparable. Nonetheless, a general concordance in banks' relative performances in the test and in a crisis is to be expected as long as there are no systematic differences in the the stress scenario and the crisis (which would be

of the shock.

[31]The U.S. economy contracted by close to 30% (YoY) in Q2 2020; the peak unemployment rate was 15%; and the Dow Jones Index plunged by close to 30% in March 2020.

undesirable either from a stress-test design perspective). Therefore, the pandemic offers a feasible, if not ideal, test of stress-testing.

Second, the Covid-19 shock was completely unexpected, like in the case of stress-tests where the hypothetical scenarios are not known to banks in advance. The test results were announced on 25th June, which means that it is less likely that banks were able to anticipate the attendant capital surcharges (i.e. the SCBs) and adjust/raise capital in time for their second-quarter (i.e., end-June) earnings reports. These factors imply that endogeneity issues associated with the Q2 2020 CET1 ratios can be ruled out.[32]

Third, that risk-weighted assets and loan loss provisions (LLPs) are forward looking, and that banks front-loaded their response to the crisis by increasing LLPs substantially in Q2 2020, means that the Q2 capital ratios likely reflect how banks would eventually perform in the crisis. This help address, at least partially, claims that the CET1 ratio of banks may not have bottomed out yet.[33] Although concerns that extraordinary support was provided to banks and to the broader economy have dented the impact of the crisis on banks (so far) may still apply. Despite this caveat, we consider that focusing on the relative performance of banks is a potential inroad to appraising stress-tests.

With this view, a comparison of the test-implied and actual CET1 ratios shows that they do not line-up (see Figure 10).[34] In fact, while the ratios declined for almost all banks in the test, it rose for many in reality. In the case of Deutsche Bank USA, for instance, the CET1 ratio declined by 8 percentage points (pp) in the test, while during H1 2020, the same ratio rose by 5 pp. A similar although less stark narrative applies to most other banks. Even in terms of *relative* (cross-sectional) performance of banks in the test

[32]Potential endogeneity issues also suggest that the 2019 DFAST or the 2018 European Banking Authority (EBA) stress-tests cannot be appraised against banks' actual performances in the Covid-19 crisis as banks are likely to act on the test results and evolve materially in the meantime. Note that the 2020 EBA stress-test was postponed to give banks operational relief.

[33]The DFAST results only report the minimum and the end-of period ratios, and not the complete trajectory.

[34]Note that for most banks, the primary driver of change in the CET1 ratio is risk-weighted assets as opposed to the level of CET1 capital. This is consistent with the capital action assumption in DFAST which does not allow for new issuance of CET1 capital.

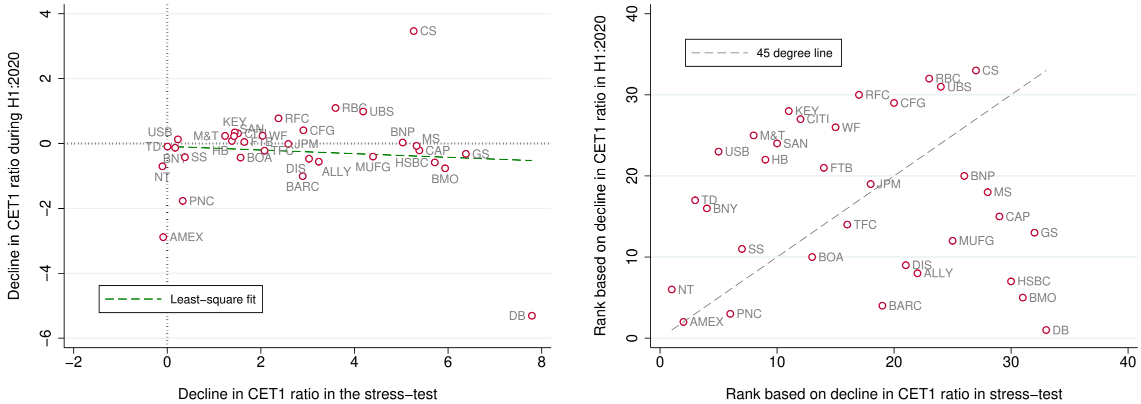

Figure 10: (Left-hand panel: A comparison of the decline in CET1 ratio in the stress-test and the actual decline observed between Q4:2019 and Q2:2020. Unit of both axes is percentage points. Right-hand panel: A comparison of the rank based on decline in CET1 ratio in the stress-test and rank based on the actual decline observed between Q4:2019 and Q2:2020. A lower rank (number) indicates a smaller decline in the ratio.

as compared to reality, we notice a high degree of discordance (see Figure 10).

Comparing test performance with market indicators of banks resilience also points to a qualitatively similar conclusion. Specifically, we consider how banks' CDS spreads evolved during the course of 2020. To gauge the immediate impact of the Covid-19 shock on banks, which also enables us to mitigate any effect of the various support programs, we evaluate the change in CDS spreads between January and the average in March and April viz-a-viz the banks' performances in the test (left hand panel of Figure 11). This comparison shows a somewhat positive correlation between test performance and market assessment of banks' risk profiles at impact. Yet, given that the movement in CDS spreads on impact may be prone to herding or other market imperfections, we also look at how spreads have evolved until end of 2020 (right-hand panel). In this case too we find limited evidence of concordance in the banks' test and real performances.

The observed discordances may not be taken as definitive evidence of Type-I/II errors in stress-testing. Yet, they raise the question whether a differently designed stress test could have generated a lower discordance between banks' performance in the stress test and their actual performance in the crisis. As we show in this paper, the discordance matters as in the medium to long term, it may impact how banks view and respond to stress-tests, and more fundamentally, banks' incentives to improve their risk-return profile.

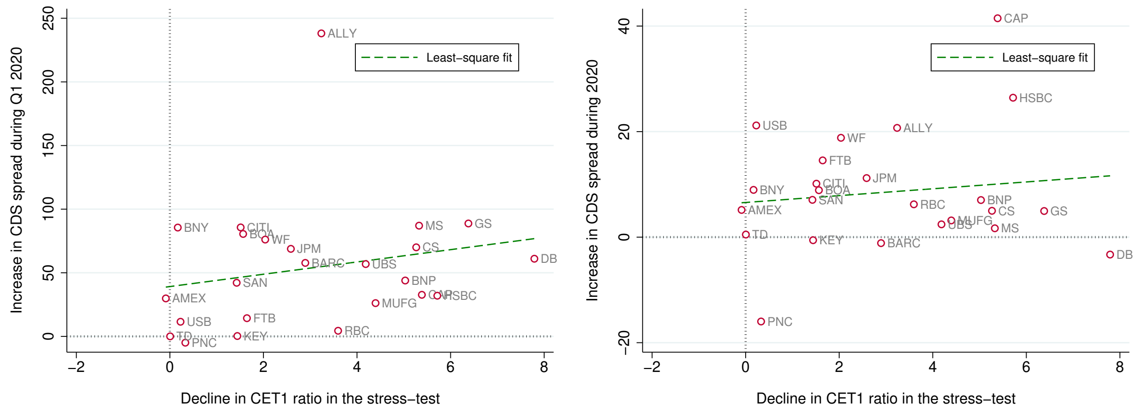

Figure 11: A comparison of the decline in CET1 ratio in the stress-test and rank based on the increase in CDS spreads from January to March (left hand panel) and to December (right-hand panel). CDS data is not available for all banks in the sample.

6 Conclusion

Stress tests have become an important tool for regulators in the post-GFC era. They have helped regulators in gauging banks' riskiness and in bolstering financial stability. They continue to evolve and improve based on lessons learnt over the years. Despite these enhancements, stress-tests may not be perfect yet, not least due to fundamental difficulties inherent in prediction of bank resilience in a hypothetical scenario. Given that stress-test results underpin banks' capital requirements, inaccurate results can have a large impact on banks' capital costs, on their operational incentives, and on overall economic welfare.

To assess the implications, we build a model of stress-testing, and show that inaccuracies may not only reduce welfare directly, but also by creating adverse ex-ante incentives. Going against the conventional wisdom, we show that in the presence of information frictions, higher capital requirements may lead to more risky banks. As such, stress-test based regulatory actions must be less strenuous when accuracy is lower, or when the regulator can choose the level of accuracy but it is fundamentally challenging or prohibitively costly to do so.

The parsimony and tractability of our model makes it amenable to extensions of interest. For instance, some regulators have discussed maintaining a surprise element in stress-tests on the grounds that it can help avoid pre-positioning or complacency by banks.[35] The welfare effects of surprise or extraneous noise in stress tests is not obvious because while it can limit the scope for gaming by banks, higher regulatory uncertainty can weaken the link between the effort banks exert and their performance in the stress-test. This can make banks exert less effort towards improving their risk-return profile. Our model can be extended to study this trade-off by adding some uncertainty around the type-specific capital requirements. Relatedly, it is possible to study the impact disclosures have by allowing depositors (or other creditors) to react to stress-test results in our model.

References

Acharya, V., R. Engle, and D. Pierret (2014). Testing macroprudential stress tests: The risk of regulatory risk weights. *Journal of Monetary Economics 65*, 36–53.

Ahnert, T., J. Chapman, and C. Wilkins (2020). Should bank capital regulation be risk sensitive? *Journal of Financial Intermediation*, 100870.

Baudino, P., R. Goetschmann, J. Henry, K. Taniguchi, and W. Zhu (2018). Stress-testing banks: A comparative analysis. *FSI Insights on policy implementation* (12).

Begenau, J. (2019). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics*.

[35]See, for instance, the remarks by Mr Jerome H Powell, Chair of US Federal Reserve System, at the research conference titled "Stress Testing: A Discussion and Review" on 9 July 2019. In fact, the continuous evolution of the stress-test regime may be motivated by this pursuit.

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

Breuer, T., M. Jandacka, K. Rheinberger, and M. Summer (2018). How to find plausible, severe, and useful stress scenarios. *Eighteenth issue (September 2009) of the International Journal of Central Banking*.

Christiano, L. and D. Ikeda (2016). Bank leverage and social welfare. *American Economic Review 106*(5), 560–64.

Corona, C., L. Nan, and G. Zhang (2017). The coordination role of stress test disclosure in bank risk taking. Technical report, Working paper.

Covas, F. B., B. Rump, and E. Zakrajšek (2014). Stress-testing us bank holding companies: A dynamic panel quantile regression approach. *International Journal of Forecasting 30*(3), 691–713.

Fed (2020). Press release: Federal reserve board releases corrected stress test results stemming from an error in projected trading losses and as a result, revised the capital requirements for two banks (september).

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

Frame, W. S., K. Gerardi, and P. Willen (2015). The failure of supervisory stress testing: Fannie mae, freddie mac, and the of federal housing enterprise oversight.

Gertler, M. and N. Kiyotaki (2010). Financial intermediation and credit policy in business cycle analysis. *Handbook of Monetary Economics 3*, 547.

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

Goldstein, I. and H. Sapra (2013). Should banks' stress test results be disclosed? an analysis of the costs and benefits. *Foundations and Trends in Finance, forthcoming*.

Gorton, G. B. (2008). The panic of 2007. Technical report, National Bureau of Economic Research.

Kara, G. and S. M. Ozsoy (2016). Bank regulation under fire sale externalities.

Kareken, J. H. and N. Wallace (1978). Deposit insurance and bank regulation: A partial-equilibrium exposition. *Journal of business*, 413–438.

Leitner, Y. and B. Williams (2020). Model secrecy and stress tests. *Available at SSRN 3606654*.

Lohmann, S. (1992). Optimal commitment in monetary policy: credibility versus flexibility. *The American Economic Review 82*(1), 273–286.

Marshall, D. A. and E. S. Prescott (2001). Bank capital regulation with and without state-contingent penalties. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 54, pp. 139–184. Elsevier.

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

Morrison, A. D. and L. White (2005). Crises and capital requirements in banking. *American Economic Review 95*(5), 1548–1572.

Nguyen, T. T. (2015). Bank capital requirements: A quantitative analysis. *Charles A. Dice Center Working Paper* (2015-14).

Ong, M. L. L., M. R. Maino, and N. Duma (2010). *Into the great unknown: Stress testing with weak data.* Number 10-282. International Monetary Fund.

Orlov, D., P. Zryumov, and A. Skrzypacz (2018). Design of macro-prudential stress tests.

Parlatore, C. and T. Philippon (2018). Designing stress scenarios. In *2018 Meeting Papers*, Number 1090. Society for Economic Dynamics.

Philippon, T., P. Pessarossi, and B. Camara (2017). Backtesting european stress tests. Technical report, National Bureau of Economic Research.

Prescott, E. S. (2004). Auditing and bank capital regulation. *FRB Richmond Economic Quarterly 90*(4), 47–63.

Santos, J. A. (2001). Bank capital regulation in contemporary banking theory: A review of the literature. *Financial Markets, Institutions & Instruments 10*(2), 41–84.

Van den Heuvel, S. J. (2008). The welfare cost of bank capital requirements. *Journal of Monetary Economics 55*(2), 298–320.

Appendix

The regulator's problem in this case is as follows, where both q_H and q_L increase as y, the underlying accuracy of the test, increases:

$$\max_{x,y} \quad \beta p\Big(q_H(y)U_H(\mathcal{X}^o) + (1 - q_H(y))U_H(\mathcal{X}^o + x)\Big) +$$

$$\beta(1-p)\Big(q_L(y)U_L(\mathcal{X}^o) + (1 - q_L(y))U_L(\mathcal{X}^o + x)\Big) - \gamma_c y$$

The first order conditions are:

$$[x] \quad 0 = p(1 - q_H(y))U'_H(\mathcal{X}^o + x) + (1-p)(1 - q_L(y))U'_L(\mathcal{X}^o + x)$$

$$[y] \quad 0 = \beta p q'_H(y)(U_H(\mathcal{X}^o) - U_H(\mathcal{X}^o + x)) + \beta(1-p)q'_L(y)(U_L(\mathcal{X}^o) - U_L(\mathcal{X}^o + x)) - \gamma_c$$

Next consider an increase in the cost of accuracy γ_c. A total derivative of the FOCs leads to:

$$[x]: \quad 0 = p\Big((1-q_H(y))U_H''(\mathcal{X}^o+x)\dot{x} - q_H'(y)\dot{y}U_H'(\mathcal{X}^o+x)\Big)+$$

$$(1-p)\Big((1-q_L(y))U_L''(\mathcal{X}^o+x)\dot{x} - q_L'(y)\dot{y}U_L'(\mathcal{X}^o+x)\Big)$$

$$[y]: \quad 0 = \beta p\Big(q_H''(y)\dot{y}(U_H(\mathcal{X}^o) - U_H(\mathcal{X}^o+x)) - q_H'(y)U_H'(\mathcal{X}^o+x)\dot{x}\Big)+$$

$$\beta(1-p)\Big(q_L''(y)\dot{y}(U_L(\mathcal{X}^o) - U_L(\mathcal{X}^o+x)) - q_L'(y)U_L'(\mathcal{X}^o+x)\dot{x}\Big) - 1$$

where $\dot{y} = \frac{\partial y}{\partial \gamma_c}$ and $\dot{x} = \frac{\partial x}{\partial \gamma_c}$. The first total derivative no longer (compare to Proposition 4) implies that $\dot{x}$ and $\dot{y}$ are of the same sign given that U is concave, $U_H' < 0$, $q_H' > 0$, $U_L' > 0$, and $q_L' > 0$. This means that accuracy and surcharge do not go hand in hand. To assess the direction of change in y, consider the second total derivative, and replace $\dot{x}$ using the first total derivative. Like before, the second-order sufficiency condition for optimality (ie negative-definite Hessian matrix) implies that $\dot{y} < 0$.