CYBERSECURITY AND FINANCIAL STABILITY

KARTIK ANAND, CHANELLE DULEY, AND PRASANNA GAI

ABSTRACT. Cyber attacks can trigger bank runs when vulnerabilities in banks' common IT systems are exploited to impair their ability to service deposit withdrawals. We show how cyber attacks and the related run risk shape banks' incentives to invest in cybersecurity. Greater run risk encourages banks to privately invest more in cybersecurity protection against attacks. However, banks do not account for how run risk increases system-wide financial instability in the wake of attacks. Increased run risk, thus, widens the gap between private and socially optimal outcomes. We discuss how regulatory measures can help facilitate constrained efficient cybersecurity.

Keywords: Cybersecurity, bank runs, global games, public good provision.

JEL classifications: G01, G21, G28, H41.

We thank Toni Ahnert, Mei Dong, Thomas Eisenbach, Sayantan Ghosal, Tetsuya Hoshino, Charlie Kahn, Philipp J. König, Stephen Morris, Silvio Petriconi, Hyun Shin, Xavier Vives and participants at the Bank for International Settlements Research Seminar Series, CAFRAL Webinar, the Deutsche Bundesbank Research Brown Bag, the European Banking Theory Brown Bag Webinar 2021, IFABS Conference, Oxford, 2021, Ridge Virtual Forum 2021 Workshop on Financial Stability, Montevideo, the 2021 European Winter Meeting of the Econometric Society, Barcelona, RiskLab/BoF/ESRB Conference on Systemic Risk Analytics 2022, Helsinki, ESRB Mini Workshop of Cyber risks 2022, Frankfurt, Workshop on Financial Stability 2022, Halle, and CEMLA 2022 Workshop on Financial Stability, Mexico City, for helpful comments. All remaining errors are our own. This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem. Anand: Deutsche Bundesbank, Research Department, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt, Germany. Email: kartik.anand@bundesbank.de. Duley and Gai: University of Auckland, 12 Grafton Rd, Auckland 1010, New Zealand. Email: chanelle.duley@auckland.ac.nz and p.gai@auckland.ac.nz.

> *"Cyber security is a public good...the social benefit conveyed by a well functioning and resilient financial system...requires a higher level of investment in cyber security than what individual firms would like to do on their own. In addition, many individual firms rely on shared services....an individual firm may rely on others in the shared network to make investments to increase the security of the network, but if every firm thinks this way, there will be underinvestment in security."*
>
> — Loretta J. Mester, Reserve Bank of Cleveland, 21 November 2019

1. Introduction

Motivation and Contribution. From mobile phone-based solutions for customers to virtual data rooms automating due diligence processes, modern banks bear little resemblance to the brick and mortar institutions of yesteryear. But as the digital transformation in banking has gathered pace, so too have cyber risks to financial stability.[1] It has been argued, for example, that cyber attacks can trigger bank runs (Duffie and Younger, 2019) and may have significant spillovers in wholesale funding markets (Eisenbach et al., 2022). In response, public authorities have mandated that banks run cyber stress-tests (European Systemic Risk Board, 2022) and are redoubling efforts to collect data on cyber risks (Federal Deposit Insurance Corporation, 2022). Despite the growing interest, our understanding of how cyber attacks and the related run risk shapes banks' efforts towards protection is limited. Our paper seeks to fill this gap.

A cyber attack occurs when a malicious adversary takes a deliberate action to exploit vulnerabilities in a bank's IT systems to either disrupt critical processes or gain unauthorised access to data. Banks can protect themselves from cyber attacks by investing in cybersecurity, i.e., taking measures aimed at detecting and patching up vulnerabilities and thwarting unauthorised access. But, due to the increased outsourcing of IT services, shared use of open-source digital solutions, and concentration of banks amongst a handful of IT service providers, banks may fail to internalise

[1] Data from the Carnegie Endowment for International Peace indicates that the number of cyber attacks on financial institutions is increasing four-fold, year-on-year (Mauer and Nelson, 2020).

that other banks also also exposed to the same vulnerabilities. This distinguishes cyber risk from other forms of operational risk. As suggested by Mester (2019) in the epigraph, it means that cybersecurity has the hallmarks of a public good.

We show how banks' ex ante contributions to cybersecurity are shaped by ex post bank runs that threaten financial stability. Since banks do not internalise the cybersecurity contribution decisions of others using common IT systems, they have incentives to free-ride on these contributions. While free riding discourages investment in cybersecurity, the prospect of a bank run encourages greater effort. While free riding discourages investment in cybersecurity, the prospect of a run encourages greater effort by individual banks. But, at the system level, under-investment in cybersecurity by banks is magnified as rollover risk increases. This is because rollover risk heightens financial instability for all banks, making protection more valuable. Although banks increase their cybersecurity investment at the margin, their efforts are less than what is socially desirable to make the financial system resilient. While cast as a model of bank runs, our analysis has relevance to other settings where the ex ante provision of a public good takes place in the shadow of coordination failure.

Framework. Our analysis incorporates cyber attacks and cybersecurity into a model of bank runs (Rochet and Vives, 2004). Banks invest in profitable, risk-free, and liquid projects by issuing demandable debt claims to depositors. To manage their business, banks obtain IT systems from a common third-party technology vendor.[2] But the IT systems have vulnerabilities unknown to both the vendor and banks (Perlroth, 2021).[3] An attacker chooses how much effort to expend to discover and exploit the vulnerabilities. Banks, in turn, choose how much to contribute to cybersecurity to uncover and patch the vulnerabilities. Following Cornes (1993), we model cybersecurity with a CES aggregator where banks' contributions to protection are imperfect substitutes. By contributing more towards cybersecurity, a bank reduces its profitable investments.

[2]Gartner (2020) estimates that Amazon, Microsoft, Alibaba, Google and IBM account for 77% of the market for cloud solutions including data warehousing, run-time services and operating systems that facilitate both customer online banking services and banks' back-end operations.

[3]Vulnerabilities refer to coding and hardware flaws or weaknesses that can be exploited to: (i) gain privileged access to victims' devices and applications; (ii) access data on the affected device or application and (iii) disrupt access for legitimate users. Vulnerabilities are ubiquitous in major platforms, from Amazon Web Services to VMware, and other large and small digital services. Examples include the 'Meltdown' and 'Spectre' hardware vulnerabilities present in every Intel processor since 1995, but only uncovered in 2018 (Lipp et al., 2018; Kocher et al., 2019).

An attacker who successfully identifies the vulnerabilities launches a cyber attack by deploying malicious code, which causes a temporary outage to the IT systems that impairs banks' recourse to liquidity. This, in turn, can lead to deposits being withdrawn early and precipitate a bank run. And, even after services resume, banks may suffer permanent losses that further compromise their ability to repay depositors. We pin down a unique equilibrium using global game methods (Morris and Shin, 2003); a run occurs if the outage exceeds a certain threshold which depends on the bank's investment decisions.

The more that banks contribute to cybersecurity, the more likely it becomes that the vulnerabilities in their common IT systems are discovered and patched up in time. While this reduces the likelihood of the attacker successfully disrupting operations at the individual and system level, protection comes at the expense of profitable investments. Conditional on the cyber attack being successful, banks have fewer resources at their disposal to service withdrawals, which increases the risk of failing due to runs.

Main Results. Banks' contributions to cybersecurity are strategic complements, i.e., the more a bank invests in cybersecurity, the greater are the incentives for others to do likewise. Formally, the cybersecurity contribution game among banks is supermodular, with expected equity values displaying increasing differences (Van Zandt and Vives, 2007). The ex ante contributions to cybersecurity are characterised by multiple equilibria. In one equilibrium, all banks find it optimal to contribute nothing towards cybersecurity, and instead invest everything in profitable investments. In the other, there is a positive contribution by all banks which benefits everyone.

Cybersecurity investments are constrained inefficient. The desire to free-ride on the public good contributions of other banks towards collective security means that a bank under-invests in cybersecurity. Each bank equates its individual marginal rate of substitution between cybersecurity and profitable investment with the marginal rate of transformation between its endowment and the public good, in contrast to the Samuelson (1954) rule for optimal public good provision. Moreover, at the system level, rollover risk enlarges the gap between the private and social optimum. Greater rollover risk heightens financial instability for all banks and this makes protection from attacks

more valuable. Banks fail to account for the social benefits of cybersecurity that extend beyond their own balance sheets and so they increase their contributions by less than what a planner would choose.

The sub-optimal system-wide cybersecurity allocations can be corrected by regulatory measures. We consider three schemes: duty of care penalties, cybersecurity subsidies, and cyber stress tests. By establishing a minimum standard of due care, a regulator can impose penalties on financial firms deemed to have deficient cybersecurity practices. The threat of the penalty in the event of an attack disciplines banks to contribute the socially optimal amount ex ante. Regulators can also attain the constrained efficient outcome by subsidising financial market cybersecurity resilience initiatives. Stress tests and cyber hygiene notices, such as those recently undertaken by central banks, also encourage coordination on the social optimum.

Extensions. We extend our baseline analysis to consider how (i) learning about possible disruptions, and (ii) bank heterogeneity affect cybersecurity provision. We relax the assumption that attackers have no prior knowledge about vulnerabilities when exerting effort to disrupt bank activity. Attackers' effort, while not revealing the technical details of vulnerabilities, signals the extent of possible disruptions to banks. A tripartite classification of cyber attacks emerges, featuring low-grade (trivial) attacks, high-grade (critical) attacks, and the possibility of both types driven by self-fulfilling beliefs. The information banks obtain about the vulnerability serves as a focal point to anchor their beliefs. Multiple equilibria arise on the severity of the attack. While multiplicity in the cybersecurity contribution game is driven by beliefs over other banks' investments, this indeterminacy is a result of self-fulfilling beliefs about whether exploitation of the vulnerability would render banks illiquid. Agencies that disseminate information about cyber risks and vulnerabilities can eliminate this multiplicity.

Bank heterogeneity is systematically related to cybersecurity provision. The substitutability of bank contributions, reflecting similarity in banks' IT solutions, shapes how the burden of cybersecurity production is shared between banks. We show that policy targeting only banks that are significant contributors is sufficient to achieve the social optimum, and discuss how transfers

among heterogeneous banks can sometimes lead to Pareto-improving re-allocations of cybersecurity investment.

Related literature. Our paper contributes to the nascent, largely empirical, literature on cybersecurity and financial stability. Duffie and Younger (2019) describe cyber-runs and conduct a stress test to understand the resilience of systemically important US banks to wholesale depositor withdrawals following a cyber attack.[4] Eisenbach et al. (2022) examine how cyber attacks impair a bank's ability to repay withdrawing creditors and discuss how this influences creditors' incentives to run. They suggest that, since cyber attacks impair the ability of the bank to repay early withdrawals, the first-mover advantage of creditors is weakened. The sequential service constraint means that creditors who withdraw face a lower probability of being repaid in full. Our analysis complements this argument. In our model, whenever rollover risk is low, bank failure following a cyber attack is not driven by coordination failure but instead by the deadweight losses. In this case, since the impairment to banks' ability to repay is permanent, there is no scope for an inefficient run. But when rollover risk is large, the impairment suffered by a bank following a cyber attack is more transient, and inefficient runs are a source of bank failure.

Our paper informs the growing policy literature in this area (Kashyap and Wetherilt, 2019; Adelmann et al., 2020; Elestedt et al., 2021; Fell et al., 2022). These papers recognise that cyber attacks can impair financial stability and note more should be done to bolster banks' resilience in the face of the attacks. Fell et al. (2022) argue that policymakers should improve their monitoring of cyber attacks to expand their macroprudential toolkits to cover cyber risks. Elestedt et al. (2021) suggest that better coordination and cooperation between the financial sector and relevant agencies responsible for cybersecurity is vital to financial stability. Our analysis provides a formal basis for such policy concerns.

[4] Jamilov et al. (2021) contribute to the empirical literature on cyber risk and contagion. They construct a text-based measure of cyber risk, which they use to test the link between balance sheet and income statement information from publicly-listed companies and their exposure to cyber risk. They also explore whether cyber risk exposure influences asset pricing, and whether the effects of this exposure propagate to unaffected peer firms. They demonstrate that cyber risk exposure has a negative and significant effect on stock returns of affected firms and find evidence that cyber risk is a source of systematic risk in financial markets due to contagion effects or firm-to-firm networks. See also Kamiya et al. (2021), Woods et al. (2021) and Florakis et al. (2020).

There are also points of contact with the economics of security. Gordon and Loeb (2002) consider a one-period model of a firm choosing how much to invest in IT security, given an exogenous threat probability. They argue that, since the firm's investment depends on the marginal product of security investment, it may be optimal for the firm to invest very little or nothing at all. Varian (2004) and Grossklags et al. (2008) extend this analysis to the case of multiple firms with network externalities where cybersecurity exhibits properties of public goods. Varian (2004) shows that under-provision of cybersecurity at the system level can be rectified using so-called negligence rules. Grossklags et al. (2008) examine how the option to invest in insurance against damages following a cyber attack influence decisions to contribute to cybersecurity. Our contribution to this literature shows how ex post coordination failures in the form of bank runs, and information on the nature of vulnerabilities shape banks' incentives to contribute to the public good.

Our model also relates to attack and defender games in the economics literature. This literature focuses on how the incentives of attackers and defenders depends on the network structure (Bier et al., 2007; Dziubinski and Goyal, 2013; Goyal and Vigner, 2014 and Acemoglu et al., 2016). In contrast, we abstract from the network perspective and highlight the public good aspect of cybersecurity. Finally, we add to the large literature on bank runs and global games (Morris and Shin, 2003; Goldstein and Pauzner, 2005). We specifically build on Rochet and Vives (2004), where unsecured debt holders delegate their rollover decisions to professional managers, so the decisions to rollover are global strategic complements. Our contribution shows how cyber risk management interacts with run risk in such a setting.

2. Model

A single-good economy comprises $N \geq 2$ identical risk-neutral banks, indexed $b = 1, \ldots, N$. Banks operate over three dates, $t = 0, 1, 2$, and are protected by limited liability. At $t = 0$, the representative bank has access to a safe and liquid investment that matures at $t = 2$ and provides return $R > 1$. The bank is funded by a mix of own funds, $E > 0$, and unsecured, demandable debt, $D = 1 - E$, issued to risk-neutral depositors. As in Rochet and Vives (2004), depositors

delegate rollover decisions to professional fund managers. The face value of the debt, denoted F, is exogenous and independent of the withdrawal date.[5]

Common IT solutions. Banks use common IT solutions provided by third-party vendors to manage their investments. These solutions include software, such as liquidity management programs, or hardware solutions to securely store confidential data. Oracle Corporation, for example, sells FLEXCUBE to banks such as Wells Fargo, Citigroup, and HSBC. FLEXCUBE is a suite of software solutions to help streamline and optimise core bank operations including the management of accounts and loans, transferal of funds, and reporting tools.[6]

Vulnerabilities and cybersecurity. IT solutions have vulnerabilities or "bugs" that are unknown even to the vendors. A malicious actor who discovers the vulnerabilities can write code to exploit them for personal gain. It is, thus, in the banks' interests to apply mitigating measures.

We consider a strategic contest between an attacker and banks who have a first-mover advantage (Dixit, 1987). At $t = 0$, a risk-neutral, deep-pocketed attacker invests $A \geq 0$ to discover vulnerabilities at marginal cost, c. This investment may be thought of as the time spent searching through source code to locate vulnerabilities and writing code to exploit them. Anticipating this behaviour by the attacker, each bank, b, allocates S_b towards *cybersecurity* to uncover and patch vulnerabilities. Cybersecurity involves, for example, paying security experts to hack into systems (often referred to as "red-teaming") or paying a "bug bounty"—monetary rewards to members of the public who report previously undisclosed vulnerabilities. [7] Although vendors also have incentives to identify and patch vulnerabilities, we focus on the role of banks.[8]

[5]In Appendix B, we endogenise the face value of debt and show that this does not qualitatively change our main results.

[6]Other prominent technology vendors include Finestra and Amazon Web Services (AWS).

[7]Kashyap and Wetherilt (2019) report that, in 2019 banks spent up to $124 billion on cybersecurity. In 2022, ethical hackers earned more than $230 million by reporting vulnerabilities to HackerOne, a bug bounty platform (HackerOne, 2022).

[8]Many vendors develop products using open-source solutions including Apache and Linux which are used by banks such as JPMorgan Chase, ING, and Raffeisen. But, in doing so, vendors effectively outsource the problem of finding vulnerabilities to the open-source community, which is rife with free-riding (Johnson, 2002). Our assumption that vendors do not contribute to cybersecurity is an extreme depiction of this environment.

Cybersecurity also involves the sharing of information once weak points have been detected. In practice, this means disclosing information on vulnerabilities to communities such as the National Vulnerability Database (NVD) or the Common Vulnerabilities and Exposures (CVE) system.[9] Thus, by eliciting and sharing information, a bank's expenditure on cybersecurity benefits all banks, making cybersecurity a *public good* (Mester, 2019).

We follow Cornes (1993) and model cybersecurity using a symmetric constant elasticity of substitution function,

$$X(\vec{S}) = \left[\frac{1}{N}\sum_{b=1}^{N} S_b^{\nu}\right]^{1/\nu}, \tag{1}$$

which aggregates the contributions of the individual banks, $\vec{S} = (S_1, \ldots, S_N)$. The parameter ν characterises the extent to which banks' contributions are substitutes. If $\nu = 1$ then banks' contributions to cybersecurity are perfect substitutes. Such a situation may arise where, for example, all banks use similar IT systems without distinct proprietary features. In the limit $\nu \to -\infty$, they become perfect complements and $X(\vec{S})$ converges to a weakest-link public good (Hirshleifer, 1983). While for $\nu \to \infty$, the game is transformed into a best shot public good (Bliss and Nalebuff, 1984). In what follows, we assume $\nu < 0$, reflecting IT solutions that are differentiated across banks, but share enough similarities to make contributions imperfect complements. This assumption has two key implications. First, each bank's marginal contribution to cybersecurity is positive but decreasing in the level of its contribution, i.e., $\partial X/\partial S_b > 0$ and $\partial^2 X/\partial S_b^2 < 0$. This allows us to derive interior solutions for banks' contributions to cybersecurity. And second, there are positive spillovers across banks, i.e., $\partial^2 X/(\partial S_b \partial S_{b'}) > 0$ for all $b \neq b'$ that reflects the gains from sharing information on vulnerabilities.

Finally, by committing S_b towards cybersecurity, bank b reduces its investment to $I_b \equiv 1 - S_b$. Thus, the private opportunity cost to a bank from contributing to cybersecurity is the foregone revenue from investing.

[9]While database administrators admit that public vulnerability disclosure can assist attackers, it has been argued that the benefits of information sharing outweigh the costs (Johnson et al., 2016). Tony Sager, Senior Vice President of the Center for Internet Security, is a proponent of this approach, encouraging "one organization's detection to become another's prevention".

Cyber attacks and bank failure conditions. At $t = 1$, banks are successful in uncovering vulnerabilities before the attacker and applying mitigating measures with probability

$$p(A, X(\vec{S})) = \frac{X(\vec{S})}{A + X(\vec{S})}\,. \tag{2}$$

The attacker, consequently, loses the contest and receives nothing. With converse probability $1 - p(A, X(\vec{S}))$, the attacker identifies the vulnerabilities first and launches a *cyber attack* by deploying malicious code to exploit the vulnerabilities. These exploits cause temporary glitches in the common IT solution and impair banks' recourse to liquidity. The cyber attack on the New Zealand stock exchange in December 2020, which prevented the posting of market announcements and led to trading being suspended for several days is an exemplar (Tarabay, 2021). Specifically, banks are subject to a common impairment shock, $\alpha \in [0, 1]$, which is a uniformly distributed random variable drawn at $t = 1$. When $\alpha = 0$, the impairment does not impact on banks' recourse to liquidity, while $\alpha = 1$ represents a complete denial to liquidity. The impairment shock takes effect at $t = 1$ and is resolved by $t = 2$. Finally, the attacker receives a prize, $V(\alpha)$, at $t = 2$ that is increasing in the impairment shock.

By impairing a bank's recourse to liquidity, the shock can precipitate bank runs (Duffie and Younger, 2019). If a fraction $\ell_b \in [0, 1]$ of debt is withdrawn at $t = 1$, the bank fails due to *illiquidity* whenever

$$(1 - \alpha)RI_b - \ell FD < 0\,, \tag{3}$$

i.e., the value of available assets is insufficient to service withdrawals. So whenever $\alpha > \alpha_b^{IL}(\ell_b) \equiv 1 - \frac{\ell_b FD}{RI_b}$, the bank fails and its equity value is wiped out.[10]

Cyber attacks can also have longer lasting repercussions. These include, for example, the loss of secret information pivotal to the bank's role as a financial intermediary (Dang et al., 2017), losses incurred from paying ransom demands, and even physical damage to critical systems. The credit

[10] In many standard bank run models (e.g., Rochet and Vives, 2004; Ahnert et al., 2019), the cost of runs at $t = 1$ stems from the fire-sale or costly liquidation of assets. We abstract from fire sales and costly liquidation to highlight the core tension between ex ante public good provision and ex post coordination failures even in absence of these features. Instead, the cost of the run at $t = 1$ is that if the bank is unable to service even a single withdrawal, this triggers a bankruptcy that completely erodes the value of its assets. We further exploit this feature in Appendix B when we endogenise the face value of debt.

downgrading of the Maltese bank, Valletta PLC, following a cyber attack and concerns over the bank's operational risk management illustrate how cyber attacks can threaten bank solvency (S&P Global Market Intelligence, 2019).

We capture such possibilities by assuming that each bank is subject to a deadweight loss proportional to the shock. In particular, after a successful attack subsides and banks regain access to their operations at $t = 2$, bank b's assets yield $(1 - \delta\alpha)RI_b$, where $\delta < 1$ reflects the deadweight loss incurred through the cyber attack. The greater the deadweight loss from a cyber attack, the larger are bank losses and the more strain is put on bank solvency.

Bank b fails due to *insolvency* at $t = 2$ whenever

$$(1 - \delta\alpha)RI_b - \ell_b FD < (1 - \ell_b)FD\,, \tag{4}$$

i.e., the gross returns from the asset are insufficient to repay the total debt claims against the bank. So the bank fails whenever $\alpha > \alpha_b^{IN} \equiv \frac{1}{\delta}\left(1 - \frac{FD}{RI_b}\right)$, which is independent of the fraction of withdrawals at the interim date.

Rollover decisions. Depositors delegate rollover decisions to professional fund managers who are rewarded for making the right decision—if the bank does not fail, a fund manager's payoff difference between withdrawing and rolling over is $-c < 0$; if the bank fails, the differential payoff is $r - c > 0$. The *conservatism ratio*, $\gamma \equiv \frac{r-c}{r}$, summarises these payoffs.[11] More conservative managers (i.e., higher γ) are less inclined to rollover since the cost of withdrawing is low. When $\gamma > 0$, fund managers' actions are strategic complements and the bank is subject to rollover risk.

In many cases, outsiders are not perfectly informed about the details of a cyber attack because banks are reluctant to publicly broadcast details due to fear of further attacks and reputational concerns (Biener et al., 2015; Pretty, 2018). To this end, we suppose that the continuum of fund managers have incomplete information about the impairment shock on which they base their rollover

[11] The probability of bank failure, P, needed to render a fund manager just indifferent between withdrawing and rolling over is given by

$$P(-c) + (1 - P)(r - c) = 0 \Leftrightarrow P = \gamma \equiv \frac{r - c}{r}\,. \tag{5}$$

decisions. For each bank, b, a fund manager, indexed $k \in [0, 1]$, receives a noisy signal

$$x_{bk} = \alpha + \epsilon_{bk}\,, \tag{6}$$

where ϵ_{bk} is a zero-mean noise term that is independent of the shock α, and is independently and identically distributed across fund managers according to a continuous distribution H with support $[-\epsilon, \epsilon]$, where $\epsilon > 0$. There is no overlap in sets of fund managers across the different banks.

Table 1 illustrates the timing of events in the model.

$t = 0$	$t = 1$	$t = 2$
1. Banks (first-movers) invest in assets and cybersecurity to uncover vulnerabilities	1. Attacker succeeds with probability p and launches cyber attack	1. Banks' assets mature
2. Attacker invests in discovering and exploiting vulnerabilities	2. Fund managers receive private signals on shocks and roll over or withdraw	2. Attacker receives the prize if successful
		3. Banks and depositors consume

TABLE 1. Timeline of events.

3. ANALYSIS

The symmetric, pure-strategy, perfect Bayesian equilibrium comprises, for the attacker, investment A^* in discovering vulnerabilities and launching attacks, and for each bank, $b = \{1, \ldots, N\}$, critical thresholds, α_b^* and x_b^*, contribution to cybersecurity, S_b^*, and investment, I_b^*, such that

(1) at $t = 1$, fund managers' rollover decisions, x_b^*, are optimal and the run threshold causes bank b to fail whenever $\alpha > \alpha_b^*$, given A^*, S_b^* and I_b^*;

(2) at $t = 0$, the attacker's investment, A^*, maximises the expected attack prize given banks' choices, S_b^* and I_b^*, for all $b = 1, \ldots, N$;

(3) at $t = 0$, the banks' choices, S_b^* and I_b^*, maximise expected equity value given the thresholds, x_b^* and α_b^*, and the attacker's investment, $A^* \equiv A^*(\vec{S})$.

We construct the equilibrium backwards, solving for the optimal rollover decision of fund managers before establishing the optimal investments by the attacker and banks.

Rollover risk. The contributions to cybersecurity, together with the size of the impairment shock, shape the dynamics of rollover risk. For a given mass of early withdrawals, ℓ_b, bank b does not fail provided the outage shock, α, is sufficiently small. But the criteria determining the largest shock that bank b can withstand depend on whether failure is driven by illiquidity or insolvency.

Lemma 1. *There exists a unique threshold,*

$$\widehat{\gamma} \equiv \frac{1}{\delta} - \left(\frac{1}{\delta} - 1\right) \frac{RI_b}{FD}, \tag{7}$$

such that bank b fails due to illiquidity if and only if the mass of withdrawals is large, i.e., $\ell_b > \widehat{\gamma}$.

Figure 1 illustrates Lemma 1 by plotting the illiquidity and insolvency conditions together with their "envelope"—the red line encapsulating the region where the bank does not fail. In region I, where the fraction of withdrawals is small, $\ell_b < \widehat{\gamma}$, the bank is able to service them following the cyber attack. But if $\alpha > \alpha_b^{IN}$, due to the subsequent losses suffered by the bank, it has too few resources to repay claims that are rolled over. So although the bank can meet interim liquidity needs, the cyber attack renders it insolvent at $t = 2$.

In region II, the fraction of withdrawals is so large, given the outage shock, that the bank is unable to service them at $t = 1$. This is despite the bank having sufficient resources at $t = 2$, allowing for losses from the cyber attack, to satisfy all its claims. In this case, the bank fails due to illiquidity even though it is solvent. Since bank equity value is wiped out when the bank fails, region II illustrates the cost of bank runs.

Assumption 1. $(1 - \delta)RI_b < FD; RI_b > FD.$

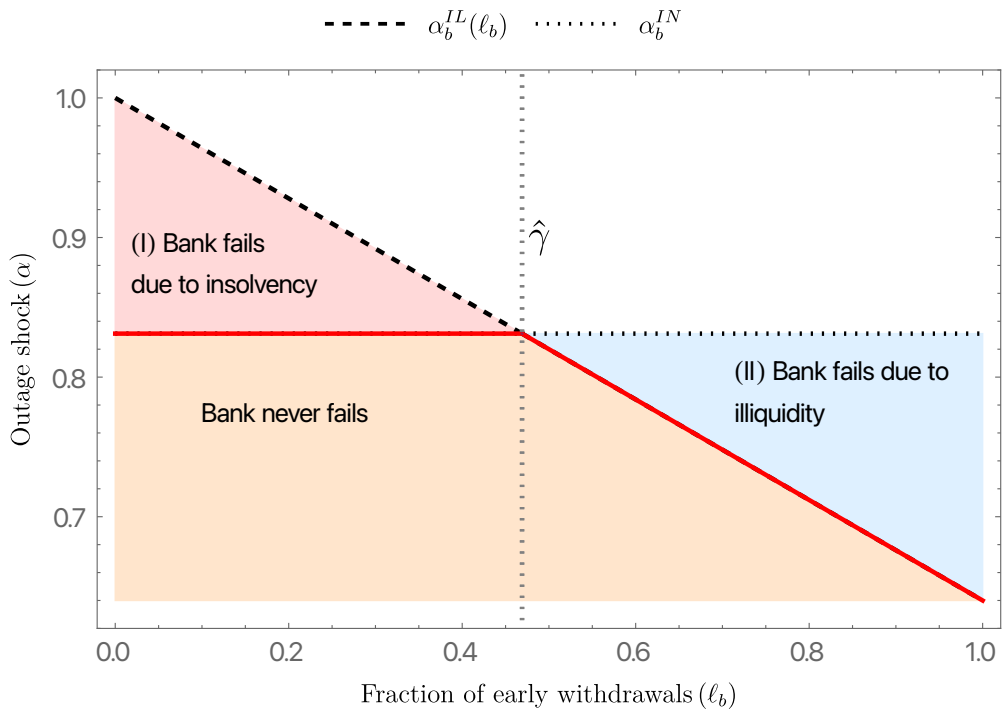

FIGURE 1. Failure conditions for a bank following a successful cyber attack.

Assumption 1 ensures a strict tripartite classification over values of the impairment shock (Morris and Shin, 1998). The first part of the assumption implies that if in the worst case, $\alpha = 1$, all debt claims are rolled over, $\ell_b = 0$, the losses that the bank incurs are so large that the bank cannot service its debts at $t = 2$ and fails due to insolvency. A sufficiently large deadweight loss, δ, is required to ensure that this region – where it is a dominant action for fund managers to withdraw – is well defined. The second part of the assumption ensures that if there is no impairment, $\alpha = 0$, and all fund managers withdraw early, $\ell_b = 1$, then the bank has sufficient resources to repay creditors in full at $t = 1$. Since $I_b = 1 - S_b$, Assumption 1 also places an upper bound on investment so that $S_b < \bar{S} \equiv 1 - \frac{FD}{R}$.

In the limit of vanishing private noise, $\epsilon \to 0$, the unique run threshold converges to the failure threshold, $x_b^* \to \alpha_b^*$, and fund managers face only strategic uncertainty about the behavior of other managers.

Proposition 1. *There exists a unique failure threshold,*

$$\alpha_b^* = \begin{cases} \alpha_b^{IN} \equiv \frac{1}{\delta}\left(1 - \frac{FD}{RI_b}\right) & \text{if } \gamma < \widehat{\gamma} \\ \alpha_b^{IL}(\gamma) \equiv 1 - \frac{\gamma FD}{RI_b} & \text{if } \gamma \geq \widehat{\gamma}\,, \end{cases} \tag{8}$$

such that bank b fails whenever $\alpha > \alpha_b^*$.

Bank failure is driven by illiquidity only when rollover risk is sufficiently large, i.e., $\gamma \geq \widehat{\gamma}$. Following an impairment shock, the bank is only able to service debt claims if few fund managers withdraw. But if a sufficient proportion of fund managers withdraw, it is in each fund manager's best interest to also withdraw because the bank will fail as a result of the run. Under vanishing private noise, there is a unique impairment shock, $\alpha_b^{IL}(\gamma)$, such that the bank fails due to a self-fulfilling bank run whenever $\alpha > \alpha_b^{IL}(\gamma)$. By contrast, when $\gamma < \widehat{\gamma}$, bank failure is driven by insolvency concerns and rollover risk plays no role. If the impairment shock is sufficiently large, $\alpha > \alpha_b^{IN}$, each fund manager has a strictly dominant strategy to withdraw at $t = 1$, since the bank is sure to fail at $t = 2$. In what follows, we define *bank fragility* as the risk of a bank failing due to rollover risk.

Corollary 1. *Conditional on a cyber attack, greater ex ante investment in assets reduces bank fragility, i.e., $\frac{\partial \alpha_b^*}{\partial I_b} > 0$, for all banks $b = 1, \ldots, N$.*

In the event a vulnerability is successfully exploited by an attacker, the more the bank has invested in assets, the greater is its recourse to liquidity to service depositor withdrawals. As a result, bank fragility is reduced. However, the result in Corollary 1 is conditional on a cyber attack taking place and does not account for the trade-off faced by banks between lowering fragility and preventing an attack entirely. We next endogenise banks' ex ante investment allocation and examine this trade-off. To this end, we focus on on the case where bank failures are driven by illiquidity, i.e., $\gamma \geq \widehat{\gamma}$.

Attacker behaviour and optimal cybersecurity. At $t = 0$, the attacker decides how much to invest to discover vulnerabilities, A, to maximise its expected prize, taking as given the level of

cybersecurity. Importantly, at the time of choosing A, the attacker does not know what vulnerabilities it will find and how disruptive their exploitation will be for the banks.[12] We thus express the attacker's problem as

$$A^* \equiv \arg\max_A \left[1 - p\left(A, X(\vec{S})\right)\right] \int_0^1 V(\alpha)d\alpha - cA\,. \tag{9}$$

In what follows, let $\overline{V}$ denote the expected prize to the attacker.

Lemma 2. *The attacker invests*

$$A^*(\vec{S}) = \begin{cases} \sqrt{\frac{X(\vec{S})\overline{V}}{c}} - X(\vec{S}) & \text{if } X(\vec{S}) < \frac{\overline{V}}{c} \\ 0 & \text{otherwise}\,, \end{cases} \tag{10}$$

at $t = 0$ to discover and exploit vulnerabilities. The investment is increasing in the prize but decreasing in the marginal cost.

Lemma 2 provides intuitive results. First, the attacker invests a strictly positive amount as long as the expected prize is sufficiently large. Second, equilibrium investment is larger the higher the expected prize, $\overline{V}$. And finally, as the marginal cost, c, increases, the potential gains from investing are lower and so the attacker invests less.

Turning to banks' investment decisions, we suppose that banks pre-commit their investment choices and, as such, have a first-mover advantage (Dixit, 1987). In practice, this pre-commitment could capture banks announcing their IT budgets at the start of every financial year. Assuming $A^* > 0$, from the bank's perspective, the probability banks will identify vulnerabilities and put sufficient mitigating measures in place is

$$p(\vec{S}) \equiv p\left(A^*(\vec{S}), X(\vec{S})\right) = \sqrt{cX(\vec{S})\Big/\overline{V}}\,, \tag{11}$$

which satisfies the Inada conditions over the unit interval.

The expected profit for bank b is the sum of its profit in the event that it successfully patches its vulnerabilities, and its residual profit if the attacker is successful in exploiting vulnerabilities and

[12] As such, the vulnerabilities may be thought of as "known unknowns" (Rumsfeld, 2002).

launching a cyber attack, i.e.,

$$\pi_b(I_b, S_b) = p(S_b, \vec{S}_{-b})\,(RI_b - FD) + \left(1 - p(S_b, \vec{S}_{-b})\right) \int_0^{\alpha_b^*(I_b)} EV_b(\alpha) d\alpha\,, \tag{12}$$

where $\vec{S}_{-b}$ are the allocations by all other $N-1$ banks, $EV_b(\alpha) \equiv (1-\delta\alpha)RI_b - FD$ is the bank's equity value in the event of a cyber attack, and $I_b + S_b = 1$ is the initial balance sheet condition.

The bank balances the marginal benefit and cost of cybersecurity when making its allocation. A higher contribution to cybersecurity improves the chances of uncovering vulnerabilities first, lowering the probability of a cyber attack. But this comes at the cost of reduced investments. This has two implications. First, irrespective of who wins the contest, the value of the bank's assets are lower which reduces its equity value. Second, in the event that the attacker successfully launches a cyber attack, the bank is more fragile. This gives the bank an incentive to free-ride on the cybersecurity contributions of other banks. Therefore, banks face a *trade-off* between heightened protection afforded ex ante and the shadow of a self-fulfilling run due to fragility ex post.

When banks choose the scale of their cybersecurity contributions and investments, taking as given the choices of all other banks, they equate their marginal rate of transformation with their own marginal rate of substitution.

Proposition 2. *Bank b's contribution to cybersecurity, S_b^*, is given by the solution to*

$$\frac{\partial \pi_b / \partial p}{\partial \pi_b / \partial I_b} = \left(\frac{p}{2X}\frac{\partial X}{\partial S_b}\right)^{-1}\,, \tag{13}$$

where

$$\frac{\partial \pi_b}{\partial p} = R(1 - S_b) - FD - \int_0^{\alpha_b^*(1-S_b)} EV_b(\alpha) d\alpha\,, \tag{14}$$

and

$$\frac{\partial \pi_b}{\partial I_b} = pR + (1-p)\left(\int_0^{\alpha_b^*(I_b)} \frac{\partial EV_b}{\partial I_b} d\alpha + EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial I_b}\right)\,. \tag{15}$$

The equilibrium satisfies $\gamma > \widehat{\gamma}(S_b^*)$, ensuring that bank failure is driven by illiquidity. Finally, the contribution is a maximum if and only if the degree of substitutability is low, $\nu \leq \widehat{\nu}(S_b^*)$.

The left-hand side of Equation (13) is a bank's marginal rate of substitution. It is the ratio of the marginal impact on bank b's profits from an increase in protection against an attack, p, to the marginal impact on bank b's profits from an increase in its investment in assets, I_b. This ratio captures the bank's willingness to forego a unit of investment for a unit increase in the probability an attack is avoided. Each bank chooses an allocation such that the quantity of investment the bank is willing to forego for a unit of extra protection is equal to the quantity of investment needed to produce the additional unit of protection, given by the marginal rate of transformation in the right-hand side of Equation (13).

For S_b^* to be an interior maximum, we require that the degree of substitutability between banks' contributions to cybersecurity to be low. If, however, the degree of substitutability is high, each bank is aware that its contribution to cybersecurity could easily be replaced by that of another bank. Such a situation may arise where, for example, all banks use similar IT systems without distinct proprietary features. Knowing this, banks face strong incentives to free ride on the contributions of others. When, however, the degree of substitutability is low, banks' IT systems are more differentiated. Each bank's contribution to cybersecurity has a lower benefit for other banks. As such, the incentives to free ride are lower leading to banks choosing interior solutions for their contributions to cybersecurity.

Cybersecurity at the system level. To establish a joint equilibrium for all bank decisions and the system-wide consequences for cybersecurity, we first show how each bank, b, responds to an increase in cybersecurity contribution by another bank.

Lemma 3. *An increase in the contribution to cybersecurity by bank b elicits an increase in the contribution by bank $b' \neq b$, i.e., $\partial S_{b'}^*/\partial S_b > 0$ for all $b' = 1, \ldots, N$.*

Contributions to cybersecurity are strategic complements across banks. As each bank increases its contribution, the benefits from having more cybersecurity also rise for all other banks. Increases in cybersecurity contributions by bank b' from a low base are met by substantial increases in other banks' contributions. As $S_{b'}$ becomes sufficiently large, the net benefit to bank b from contributing

more to cybersecurity is still positive but smaller in magnitude. This reflects the increased opportunity cost from contributing to cybersecurity, which is the foregone revenue from scaling up investment in assets.

Lemma 3 implies that the cybersecurity contributions game among banks is supermodular with increasing differences in banks' expected profits (Van Zandt and Vives, 2007). This ensures that banks' best response functions for cybersecurity contribution, given the allocations of other banks, intersect to yield greatest and least Nash equilibria.

Proposition 3. *There exist two Nash equilibria. In the first, all banks contribute nothing to cybersecurity, i.e.,* $S_b^* = 0$ *for all* $b = 1, \ldots, N$. *In the second, all banks,* $b = 1, \ldots, N$, *contribute* $S_b^* = S^* < 1$ *to cybersecurity and invest* $I_b^* = I^* \equiv 1 - S^*$ *in assets, with an equilibrium level of cybersecurity* $X^* = S^*$.

Figure 2 illustrates Proposition 3 for the two-bank case. If bank b anticipates that others will scale up their investments, then it expects the level of cybersecurity to be low. The attacker is more likely to succeed in launching a cyber attack and so it is a best response for bank b to also scale up its investment, thereby shoring up its resilience against impairment shocks. Since all banks reason in this way, there is no investment in this "bad" equilibrium.

In the other equilibrium, each bank contributes $S^* < 1$ to cybersecurity. If bank b expects others to contribute a small amount to cybersecurity, then it is in bank b's interests to do so, too. Once the level of contribution by other banks is sufficiently high, system-wide security is relatively strong. Beyond this point, there are negative returns to bank b from contributing further to cybersecurity and it refrains from doing so. The "good" equilibrium corresponds to the point of diminishing returns for all banks.

Comparative statics. Changes in cyber risk and rollover risk influence the level of cybersecurity and bank fragility. Since banks are identical, it follows that in the "good" equilibrium, $X(\vec{S}^*) = S^*$. Therefore, it suffices to present the comparative statics for S^*, i.e., each individual bank's contribution to cybersecurity.

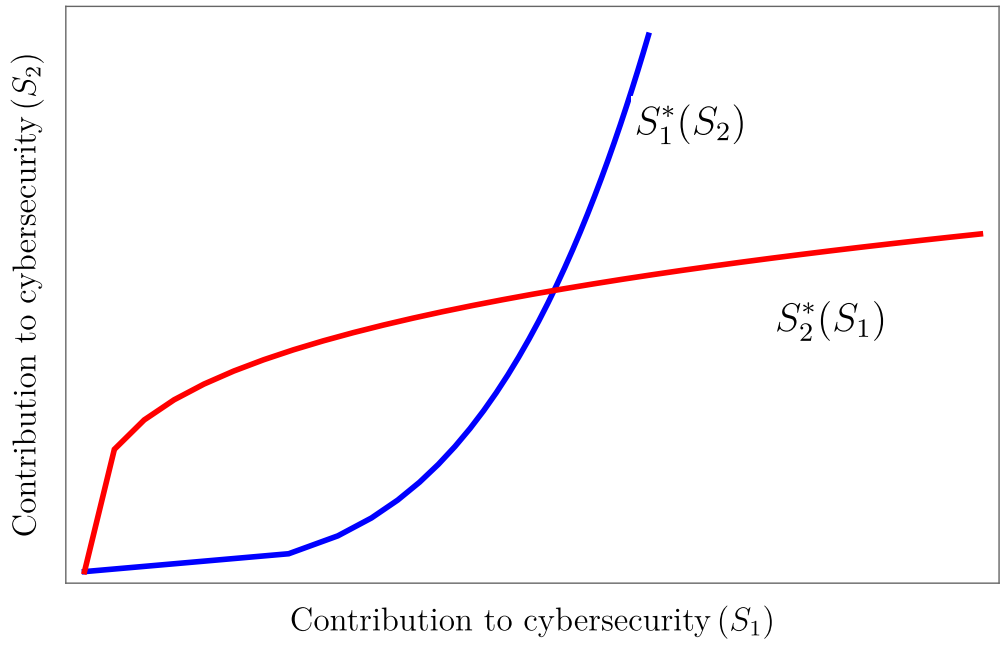

FIGURE 2. The best response of each bank to the cybersecurity contribution of another bank. The best response curves overlap at a least and greatest Nash equilibrium in which each bank contributes nothing, or a positive amount, to cybersecurity respectively.

Proposition 4. *Each bank's contribution to cybersecurity is:*

(1) increasing in the attacker's marginal investment cost, $\frac{\partial S^*}{\partial c} > 0$*;*

(2) increasing in the deadweight loss of cyber attacks, $\frac{\partial S^*}{\partial \delta} > 0$*;*

(3) increasing in rollover risk, $\frac{\partial S^*}{\partial \gamma} > 0$*; and*

(4) decreasing in bank equity, $\frac{\partial S^*}{\partial E} < 0$*.*

Bank fragility following a cyber attack is increasing in c*,* δ *and* γ*, and decreasing in* E*.*

An increase in c implies that it is more costly for the attacker to invest in finding vulnerabilities to exploit. Lemma 2 shows that this causes the attacker to invest less, which improves banks' chances of uncovering vulnerabilities first and taking sufficient mitigating measures. This incentivises banks to contribute more to cybersecurity which also increases fragility in the event a successful cyber attack is launched by the attacker.

An increase in the deadweight loss, δ, has two effects. First, the difference in expected equity value earned by the bank between an unsuccessful and successful cyber attack is larger. This

increases the marginal rate of substitution between cybersecurity and recourse to liquidity. Second, the benefits from scaling up investments in assets are lower for any level of impairment, α. The bank is better off trying to avoid becoming exposed to impairment shocks altogether. In sum, the two effects reinforce the benefits to protection via greater cybersecurity and, as a result, fragility also increases.

Given the fixed balance sheet assumption, an increase in bank equity mechanically leads to a decrease in debt issuance. But since the bank now has more "skin in the game", there is more to lose in the event that the bank fails. Since bank failure is driven by cyber attacks, conditional on a cyber attack being launched, the marginal returns to scaling up investment in assets are greater. This causes banks to reduce their contributions to cybersecurity.

Following an increase in rollover risk, each fund manager, k, is less likely to roll over claims for any given signal, x_{bk}. Ex ante, this means that bank b is more likely to fail for any realisation of the impairment shock. Therefore, the marginal benefit to contributing an additional unit to cybersecurity is high relative to the opportunity cost of scaling up investment in assets. This is because the bank is more likely to fail as fund manager conservatism increases. The bank is better off bolstering cybersecurity to ward off cyber attacks entirely, even though this leads to greater fragility in the event of a successful attack.

4. Regulatory implications

To clarify the interaction between ex ante free riding in cybersecurity by banks and ex post fragility due to coordination failure, we consider a benchmark allocation by a social planner who chooses how much each bank should contribute towards cybersecurity. The allocation problem for a social planner is

$$\max_{\{I_b,S_b\}} \quad \sum_{b=1}^{N} \left\{ p\big(R\,I_b - F\,D\big) + (1-p) \int_0^{\alpha_b^*(I_b)} EV_b(\alpha)\, d\alpha \right\}$$

$$\text{subject to} \quad 1 = I_b + S_b \quad \forall\, b = 1, ..., N$$

$$p = \sqrt{\frac{cX(\vec{S})}{\overline{V}}}\,.$$

The social planner sets bank allocations to maximise the sum of expected profits for all banks, taking into account their balance sheet conditions and how banks' contributions to cybersecurity influence the level of public good provision.

The solution to the planner's problem, $(S_1^P, S_2^P, \ldots, S_N^P)$, is given by the following system of equations:

$$\frac{\partial \pi_b/\partial p}{\partial \pi_b/\partial I_b} + \sum_{k \neq b}^{N} \frac{\partial \pi_k/\partial p}{\partial \pi_b/\partial I_b} = \left(\frac{p}{2X(\vec{S})} \frac{\partial X}{\partial S_b} \right)^{-1} \qquad \forall\ b = 1, \ldots, N\,. \tag{16}$$

Under symmetry, $S_b^P = S^P$ for all banks, $b = 1, \ldots, N$.

The solution presented in Equation (16) is a version of the Samuelson rule (Samuelson, 1954). The left-hand side of (16) is the planner's marginal rate of substitution. It is the sum of ratios of the marginal impact for each bank from an increase in protection against cyber attacks, p, to the marginal impact on bank b's profits from increasing its investments, I_b. This ratio captures the planner's willingness to forego a unit of bank b's private good for a unit increase in the level of the public good. The planner chooses an allocation such that the level of investment that is given by b in favour of contributing more to cybersecurity is equal to the contribution needed to produce the additional unit of cybersecurity i.e., the right-hand side of Equation (16). At low levels of contribution to cybersecurity by bank b, the marginal benefit from greater protection against cyber attacks is high – the right-hand side of Equation (16) is small – and so the planner requires a lower level of foregone returns to produce additional cybersecurity. But b's contribution to cybersecurity

has diminishing returns. This, coupled with bank b's increasing fragility due to substitution away from its investment, tempers the planner's incentives to continue investing in cybersecurity.

Proposition 5. *Banks contribute too little to cybersecurity, $S^* < S^P$. Moreover, the scale of the shortfall, $S^P - S^*$, is increasing in rollover risk, γ.*

The inadequate contribution to cybersecurity in the laissez-faire equilibria arises from the free-riding incentives posed by the public good. Proposition 2 shows that the marginal rate of substitution is given by the ratio of a bank's marginal returns from an increase in cybersecurity to its marginal returns from an increase in the scale of its investments in assets. On the other hand, in the planner's allocation, the marginal rate of substitution includes the *sum of all bank's marginal returns from heightened cybersecurity*. Each bank's failure to internalise the benefits to other banks from greater cybersecurity leads banks to contribute too little individually.

When rollover risk is low, so too is the risk of bank failure due to a run. Ex ante, this dampens banks' incentives to contribute to cybersecurity. Compared with the benchmark allocation in Equation (16), free-riding exerts a negative influence on banks' incentives to contribute to cybersecurity and there is some under-provision of the public good. As rollover risk rises, all banks value ex ante protection more because they all become more susceptible to runs in the event of an attack. But each bank only internalises its own heightened risk of failure due to a run. Accordingly, it increases its equilibrium contribution to cybersecurity by a lesser degree than a planner, who accounts for the social implications of heightened rollover risk, would. So although an increase in rollover risk encourages higher contributions by each bank, there is more under-provision of the public good. In this way, the ex ante free-riding externality is reinforced by ex post coordination frictions.

Proposition 5 clarifies why some policy-makers have emphasised the importance of regulatory safeguards to mitigate cyber risks in financial systems. In what follows, we consider three schemes that a regulator can introduce at $t = 0$ in order to influence banks' contributions to cybersecurity. These include (i) duty of care penalties; (ii) cybersecurity subsidies, and (iii) cyber stress tests.

Duty of care. On August 30, 2021, the Securities and Exchange Commission (SEC) in the United States sanctioned eight firms for failures in their cybersecurity policies and procedures that resulted in the unauthorised disclosure of personal information of thousands of customers and clients of the firms.[13] The financial penalties issued to the firms ranged from $200,000 – $300,000. In issuing the fines, the SEC stated that the firms had violated so-called Safeguards Rule that are designed to protect confidential customer information. In the official press release, Kristina Littman, Chief of the SEC Enforcement Division's Cyber Unit stated that, "Investment advisers and broker dealers must fulfill their obligations concerning the protection of customer information."

In enforcing the fines, the SEC, arguably, sought to enforce a minimum level of *due care*, i.e., that firms exert a minimum standard of care when entrusted with their customers' data. Such negligence rules have previously been studied as measures that can reduce the likelihood of accidents (Brown, 1973; Shavell, 2009). In the context of cybersecurity, we derive the following result.

Proposition 6. *The regulator can achieve the constrained efficient outcome by introducing a negligence rule at $t = 2$ with a penalty κ^*, in the event of a successful attack, that is proportional to banks' contributions to cybersecurity when they are less than S^P.*

The negligence rule elicits the planner's solution as a minimum level of due care for each bank. If a cyber attack is successful and banks suffer losses, then there is no further liability as long as all banks exercise the due care standard, i.e., as long as each bank contributes at least S^P to cybersecurity. Otherwise, banks that exert insufficient care are penalised proportionally to the level of their under-investment with a penalty of κ.

Imposing a penalty increases the marginal returns to each unit invested in cybersecurity because the penalty erodes residual profits in the event of a cyber attack. So banks are better off substituting away from investments and towards heightened cybersecurity. When contributing at the constrained efficient level, banks pay no penalty since their investment meets the standard of due care.

[13] See: https://www.sec.gov/news/press-release/2021-169.

The penalty reflects the distance between the regulator's optimum and what the bank chooses in the absence of any intervention. Although the bank does not directly internalise how its investment impacts other banks, the penalty ensures that the constrained efficient allocation provides the best private outcome for the bank. And since the degree of under-investment is proportional to the level of rollover risk, γ, we have that $\frac{\partial \kappa^*}{\partial \gamma} > 0$. The penalty required to enforce the negligence rule is larger when rollover risk concerns aggravate the free-rider problem caused by the public good.

Subsidising cybersecurity. Since 2019, the Deutsche Bundesbank has offered banks, insurers and financial market infrastructures voluntary services to test their resilience to cyber attacks. These tests are intended to identify vulnerabilities in the participants' systems, which can then be patched and mitigated. The Bundesbank's service is part of a wider initiative introduced by the European Central Bank in 2018 on standardising ethical hacking for enterprises to test their resilience to cyber attacks.[14]

The service offered by the Bundesbank, which takes inspiration from a similar initiative previously introduced by the Dutch Central Bank in 2016, may be viewed as subsidising firms' contributions to cybersecurity. As we show in Proposition 7, revenue-neutral Pigovian subsidies that are funded through lump-sum taxation can also elicit the constrained efficient outcome.

Proposition 7. *The regulator can achieve the constrained efficient outcome using a Pigouvian subsidy, σ^*, at $t = 0$. The subsidy is funded by taxes, $T = N\sigma^*$, imposed as a lump-sum, $\tau = \frac{T}{N}$, on each bank $b = 1, \ldots, N$.*

The subsidy scheme incentivises banks to allocate more resources towards cybersecurity. The scheme is structured such that each bank is rewarded for contributing additional units towards cybersecurity while offsetting the private costs that would be incurred from heightened fragility in absence of the scheme. This makes banks more willing to forego investment in assets in return for a higher probability that cyber breaches are avoided.

[14]Similarly, the Federal Reserve has collaborated with other U.S. regulatory agencies to develop an operational monitoring and assessment partnership with large financial institutions to help enhance their preparedness against cyber attacks. See https://www.federalreserve.gov/publications/2021-november-supervision-and-regulation-report-supervisory-developments.htm.

The subsidy rate that elicits the constrained efficient outcome is equal to the sum of marginal returns to banks other than b from a unit increase in cybersecurity contribution by bank b when other banks contribute at their constrained efficient levels. Under this design, the subsidy is decreasing in S_b: the marginal returns to other banks from additional contributions by b are diminishing in S_b and, when coupled with the increasing fragility from reducing its recourse to liquidity, result in bank b contributing no more than the constrained efficient amount.

The optimal subsidy also takes advantage of complementarities in bank contributions. Because additional contributions by other banks provide positive marginal benefits to bank b, the subsidy that achieves the social optimum is smaller in size than it would be in absence of these spillovers. Moreover, due to the symmetry of the equilibrium, the scheme features no transfers between banks (i.e., $\sigma^* = \tau$); instead, the subsidy is a coordinating device that facilitates the social optimum.

The optimal subsidy is proportional to the degree of rollover risk, $\frac{\partial \sigma^*}{\partial \gamma} > 0$. The subsidy provides additional impetus to contribute to cybersecurity over and above banks' private incentives to increase their contributions.

Cyber hygiene notices and stress tests. The regulator can also achieve the constrained efficient outcome by imposing that each bank contributes S^P at $t = 0$. Clearly, constraining banks to invest at least the level set by the regulator will ensure that there is optimal provision of the public good. But the choices are sub-optimal for banks since they would choose to contribute lower levels. An example is the approach taken by the Monetary Authority of Singapore (MAS). The MAS sets minimum regulatory guidelines in the form of a Cyber Hygiene Notice that obliges banks to implement a set of cybersecurity measures, including network perimeter defenses, malware protection, and baseline configuration standards. Compliance with these regulatory requirements and expectations are verified and enforced by the MAS (Goh et al., 2020).

The use of cyber stress tests, such as those implemented by the Bank of England, is another form of such a regulatory approach. The Financial Policy Committee of the Bank tests the resilience of the UK financial system to cyber attacks by requiring financial firms to meet a system-wide tolerance threshold set by the regulator (Kashyap and Wetherilt, 2019).

5. EXTENSION: LEARNING ABOUT VULNERABILITIES

In our baseline model, both the attacker and banks have no prior information about the bug when making their ex ante choices. In what follows, we relax this assumption by supposing that at $t = 0$, the attacker has some prior knowledge, while the banks do not. This prior knowledge consists of two parts: (i) technical details regarding the nature of the bug and (ii) an accurate estimate of the disruption, α, to banks from successfully exploiting the bug. The attacker subsequently chooses $A^* \equiv A^*(\alpha)$ given this prior knowledge. Banks, given their first-mover advantage, choose how much to invest in cybersecurity given $A^*(\alpha)$ and, thus, learn about the potential disruption that may occur if the bug is exploited. However, they do not obtain any technical knowledge about the bug from knowing $A^*(\alpha)$. Instead, banks ascertain such technical knowledge by investing in cybersecurity.

Following the lines of reasoning previously established, we can express the attacker's problem as

$$A^* \equiv \arg\max_{A} \quad \left[1 - p\left(A, X(\vec{S})\right)\right] V(\alpha) - cA\,,$$

yielding as its solution

$$A^*(\alpha, \vec{S}) = \begin{cases} \sqrt{\frac{X(\vec{S})V(\alpha)}{c}} - X(\vec{S}) & \text{if} \quad X(\vec{S}) < \frac{V(\alpha)}{c} \\ 0 & \text{else} \end{cases}$$

The ex ante probability that banks are successful in finding and patching the bug is now a function of the disruption, i.e., $p \equiv p(\alpha, \vec{S})$. As such, for each bank, b, there are two distinct cases to consider: (i) *high-grade attacks*, whereby the bank fails following a cyber attack, i.e., $\alpha > \alpha_b^*$, and (ii) *low-grade attacks* that do not lead to the bank failing, i.e., $\alpha \leq \alpha_b^*$. We can succinctly express bank b's expected profits as

$$\pi_b = p(\alpha, S_b, \vec{S}_{-b})\Big[R(1 - S_b) - DF\Big] + \Big(1 - p(\alpha, S_b, \vec{S}_{-b})\Big) EV_b(\alpha)\, \mathbb{I}_{\alpha \leq \alpha_b^*}\,,$$

where $\mathbb{I}_{\alpha \leq \alpha_b^*}$ is an indicator function denoting whether the cyber attack is low-grade or not. For tractability, we focus on the special limiting case $\nu \to 0^+$, which implies that cybersecurity is a *weaker-link* public good (Cornes, 1993). Correspondingly, we have $X(\vec{S}) = \left(\prod_{j=1}^{N} S_j\right)^{1/N}$.

Proposition 8. *Suppose* $\frac{1}{2N+1} < \gamma < \frac{1}{\delta(2N+1)}$. *There exist two bounds,* $\widehat{\alpha}^*$ *and* $\widehat{\alpha}^{**}$, *such that* $\widehat{\alpha}^* < \widehat{\alpha}^{**}$ *where*

- *If* $\alpha < \widehat{\alpha}^*$, *the equilibrium is characterised by low-grade cyber attacks;*
- *if* $\alpha > \widehat{\alpha}^{**}$, *only high-grade cyber attacks occur;*
- *for* $\alpha \in [\widehat{\alpha}^*, \widehat{\alpha}^{**}]$, *both high-grade and low-grade cyber attacks can emerge in equilibria driven by self-fulling beliefs.*

Proposition 8 demonstrates that there is a tripartite classification for cyber attacks. First, when $\alpha > \widehat{\alpha}^{**}$, only high-grade cyber attacks occur in equilibrium. In this case, banks believe that the disruptions caused by the cyber attack are sufficiently large for them to fail. They, thus, respond by allocating more towards cybersecurity. This, in turn, implies that if the cyber attack is successful, banks have even fewer resources to defend themselves against runs, which validates the initial held belief that cyber attacks are high-grade. By a similar logic, for $\alpha < \widehat{\alpha}^*$, only low-grade cyber attacks occur in equilibrium. By believing that they can withstand the cyber attack, banks are more willing to invest more in assets, which in turn, enables them to withstand the cyber attack. Finally, for $\alpha \in [\widehat{\alpha}^*, \widehat{\alpha}^{**}]$ there are multiple equilibria. Specifically, a bug that is characterised by the disruption α can manifest as either a high-grade cyber attack or a low-grade one. The equilibrium that is selected depends on banks' initially held beliefs on the severity of the disruption. Moreover, shocks to these beliefs can shift the equilibrium within the region of multiplicity.

Importantly, high-grade cyber attacks can occur in equilibrium when banks are subject to rollover risk. In the presence of rollover risk, disruptions to banks' operations can engender en-masse withdrawals leading them to fail due to illiquidity. Such failure can manifest even though banks remain fundamentally solvent following the cyber attack. Or, conversely, in the absence of rollover risk, when concerns over illiquidity risk are immaterial, banks never fail following cyber attacks. And so all cyber attacks are low-grade.

The information ascertained by banks about the disruption serves as a focal point to anchor their beliefs. Armed with the knowledge that a successful cyber attack would cause a certain level of disruption, each bank can better assess the value of having additional cybersecurity for itself. For $\alpha < \widehat{\alpha}^{*}$ and $\alpha > \widehat{\alpha}^{**}$, banks' higher-order beliefs are also well anchored, leading to the unique equilibrium outcomes. While in the intermediate range, $\alpha \in [\widehat{\alpha}^{*}, \widehat{\alpha}^{**}]$, multiple equilibria are supported since whether or not the disruption will render banks illiquid depends on their contributions to cybersecurity. This, in turn, depends on banks' higher-order beliefs on the contributions of other banks.

Multiple equilibria may be eliminated if banks' higher-order beliefs could be coordinated. This can be achieved, for example, through institutions that both coordinate the cybersecurity activities of banks and also disseminate information on risks. For example, in the United Kingdom, the National Cyber Security Centre supported the creation of the Financial Sector Cyber Collaboration Centre (FSCCC) with the aim of identifying and coordinating responses to incidents that can impact the financial sector. The FSCCC works with both financial authorities and financial firms to combine and distribute information it gathers from across the sector. Many other countries have similar institutional arrangements as well.[15]

6. Extension: Bank heterogeneity

So far, we have assumed that banks are identical. Our baseline model sheds light on the underprovision of cybersecurity even in absence of differences in budgets or balance sheet characteristics across banks. In this extension, we describe how the degree of underprovision is systematically related to the degree of bank heterogeneity, and suggest how targeted policy differs depending on the degree of substitutability of IT solutions, ν, and distribution of bank characteristics. A complete formal analysis of the equilibrium is complex and beyond the scope of the paper. However, we can apply well-established results (Bergstrom et al., 1986; Cornes, 1993; Varian, 2004) to clarify the relationship between bank heterogeneity and IT substitutability in the cyber risk setting.

[15]In the United States, this role is facilitated by the Cybersecurity and Information Security Agency. Likewise, both Australia and New Zealand also have national cybersecurity agencies.

For ease of exposition, we focus on two banks, i and j, who differ in marginal rates of substitution. For example, let $E_i < E_j$ for $i \neq j$. By Proposition 4, this implies that $\text{MRS}_i > \text{MRS}_j$ and so i is a higher contributor than j with $S_i^* > S_j^*$. In this case, the degree of substitutability shapes how the burden of cybersecurity production is shared between banks.

Remark 1. Average public good production. *As $\nu \to 1$, aggregate cybersecurity provision is increasingly borne by bank i. Bank j free rides on bank i.*

When cybersecurity production is given by the simple average contribution across banks, the marginal product of cybersecurity production is identical. Given the discrepancy in marginal rates of substitution, the burden of public good production begins to rely disproportionately on the bank with the *highest* marginal rate of substitution.

Remark 2. Best shot public good. *As $\nu \to \infty$, there are multiple equilibria. In one outcome, bank i provides all contributions and j free rides. In another outcome, bank i free rides and bank j provides all contributions.*

As $\nu \to \infty$, cybersecurity production becomes $X(\vec{S}) = \max[S_i, S_j]$. There will always be a Nash equilibrium in which the bank with the *highest* marginal rate of substitution (bank i) provides all the contributions. But given a contribution of zero by bank i, bank j's best response is to contribute positively to cybersecurity, reinforcing i's decision to free ride. And so there is also a Nash equilibrium in which the bank with the *lowest* marginal rate of substitution (bank j) contributes everything. Since aggregated profits are concave in contributions, the first Nash equilibrium Pareto-dominates the second.

Remark 3. Weaker link public good. *As $\nu \to 0^+$, aggregate cybersecurity provision is shaped by bank j. There are two equilibria: in one outcome, both banks contribute zero to cybersecurity; in another outcome, both banks contribute positively.*

As $\nu \to 0$, cybersecurity production becomes $X(\vec{S}) = \left(\prod_{b=1}^{N} S_b\right)^{1/N}$. Therefore, given a zero contribution by one bank, it is always a best response of the other bank to also contribute nothing given the marginal product of cybersecurity is zero. In the positive-contribution equilibrium,

as bank contributions become increasingly complementary, the marginal product of cybersecurity contribution varies both across banks and in the level of contribution by each bank. Bank i's returns from contributing to cybersecurity are diluted by the low contributions from j and so bank i's marginal rate of substitution is brought down towards j's. Aggregate protection is thus disproportionately affected by the bank with the *lowest* marginal rate of substitution. The positive-contribution outcome Pareto-dominates the zero-contribution outcome.

The weakest-link formulation, $\nu \to -\infty$, is an extreme outcome featuring perfect complementarity. No bank finds it optimal to contribute any more than the bank with the lowest marginal rate of substitution. Therefore, contributions are identical and there is a range of Nash equilibria. The greatest and least equilibrium are equivalent to the identical-bank equilibria in the baseline model.

Remark 4. Policy implications. *In the average cybersecurity production function, transfers between banks do not result in a Pareto improving allocation. In the best shot production function, a transfer from i to j is only Pareto improving in the equilibrium where i free rides; the same transfer lowers aggregate profits in the other equilibrium. Similarly, in the weaker and weakest link cases, a transfer from i to j results in a higher level of cybersecurity and is Pareto improving in the positive-contribution equilibrium.*

When cybersecurity is given by the average of bank contributions, a transfer of resources from i to j produces a mechanical decrease in the contribution of bank i which lowers i's expected profits. However, this is exactly offset by the increased contribution by bank j and, given the identical marginal product of cybersecurity, there is no change in i's expected profits. Bank j's profits remain similarly unchanged. Therefore, transfers are ineffective in the average public good scenario.

On the other hand, transfers play a significant role when cybersecurity has best shot properties. In both equilibria, Pareto improvements occur when the free rider transfers income to the maximum contributor until the social optimum is achieved. In a given equilibrium, this suggests that penalties and subsidies targeted only at the maximum contributor suffice in eliciting the social optimum. In

practice, however, it may be difficult for regulators to identify these maximum contributors owing to the multiplicity of equilibria.

When weaker links determine aggregate protection, income transfers also affect aggregate profits if and only if banks coordinate on the positive-contribution outcome. A transfer from i to j is Pareto-improving, while a transfer from j to i results in a net reduction in aggregate expected profits. This is due to the heterogeneity in the marginal product of cybersecurity protection. When cybersecurity has these properties and banks are positive contributors, policy tools can be targeted towards specific banks in a way that elicits the social optimum.

Remark 5. Subsidies for weaker links. *The revenue-neutral subsidy scheme redistributes bank budgets.*

Bank i's subsidy is smaller than its tax payment, but this cost is offset by the higher protection offered by j's increased contribution. Bank j's subsidy is greater than the lump-sum tax and it finds it optimal to contribute the additional income towards cybersecurity. Overall, the subsidy acts as a transfer from i to j and guarantees Pareto improvements for banks.

The negligence rule can be similarly tailored towards banks with different marginal rates of substitution. By imposing relatively larger penalties on those banks whose contributions matter most in bolstering protection, the costs to affected banks can be offset by coordinated increases in contributions resulting in greater protection for all.[16]

7. Conclusion

We provide an analytical framework to show how cyber attacks might morph into bank runs, and which takes seriously the notion that cybersecurity is a public good (Mester, 2019). In our model, banks trade off bolstering their recourse to liquidity to ward off private run risk against taking measures that benefit the security of the system as a whole. System-wide investment in cybersecurity is suboptimal as a result. We describe the trade-off between fragility and heightened

[16]Redistribution does not take place under the negligence rule. As long as banks internalise the conditional cost of paying the penalty in the event of a successful attack and adjust their contributions accordingly, no bank has to actually pay the penalty. Instead, the penalty acts as a coordinating device that nudges banks towards the social optimum.

protection, and also discuss how negligence rules, subsidies, and cyber hygiene notices facilitate constrained efficient outcomes.

Financial regulators are increasingly focused on cyber risks to financial stability. For example, the European Central Bank has introduced a Threat Intelligence-Based Ethical Red Team (TIBER-EU) framework for EU-based financial entities (Panetta, 2020). And the Monetary Authority of Singapore has examined how banks' capital and liquidity buffers might cope in the face of a 24-hour system outage triggered by a cyber event (Goh et al., 2020). Our work provides a formal basis for such regulatory emphasis. In highlighting the important role of shared IT services in generating cyber risk dependencies across banks, our results emphasise operational resilience, the public disclosure of vulnerabilities, and research into the network of linkages between banks and digital services. Such data might usefully inform "top-down" macroprudential cyber stress testing in much the same way as stress tests on interbank networks (Gai et al., 2011; Glasserman and Young, 2016).

Future work might consider how industry initiatives such as Sheltered Harbor and cyber insurance markets shape the trade-offs identified and mitigate the under provision of cybersecurity. Deeper analysis of the drivers of the deadweight losses from cyber attacks is also warranted. Arguably, cyber attacks compromise the ability of a bank to both make and keep secret information and it is the loss of such information that is, ultimately, most devastating for the integrity of the financial system.

APPENDIX A. PROOFS

A.1. **Proof of Lemma 1.** Bank b fails due to insolvency whenever

$$\alpha > \alpha_b^{IN} \equiv \frac{1}{\delta}\left(1 - \frac{FD}{RI_b}\right), \tag{17}$$

and it fails due to illiquidity whenever

$$\alpha > \alpha_b^{IL}(\ell_b) \equiv 1 - \ell_b \frac{FD}{RI_b}. \tag{18}$$

While α_b^{IN} is invariant to the proportion of withdrawals, the threshold $\alpha_b^{IL}(\ell_b)$ is decreasing in ℓ_b. The proportion of withdrawals, $\widehat{\gamma}_b$, for which the two failure conditions intersect is given by $\alpha_b^{IL}(\hat{\gamma}) = \alpha_b^{IN}$, i.e.,

$$\frac{1}{\delta}\left(1 - \frac{FD}{RI_b}\right) = 1 - \widehat{\gamma}_b \frac{FD}{RI_b}. \tag{19}$$

A.2. **Proof of Proposition 1.** The proof is in three steps. First, we show that the dominance regions at $t = 1$ are well defined. If all fund managers withdraw early, $\ell_b = 1$, then the illiquidity failure threshold is given by $\alpha_b^{IL}(1) = 1 - \frac{FD}{RI_b}$, where $\alpha_b^{IL}(1) < \alpha_b^{IN}$. If, however, no fund manager withdraws at $t = 1$, then $\ell_b = 0$. In this case, the bank never fails due to illiquidity since $\alpha_b^{IL}(0) > 1$. But the bank can, nevertheless, fail at $t = 2$ due to insolvency whenever $\alpha > \alpha_b^{IN}$, since, under Assumption 1, $\alpha_b^{IN} < 1$.

It also follows from Assumption 1 that $\underline{\alpha_b} \equiv \alpha_b^{IL}(1) > 0$ is the largest shock that bank b can withstand even if all fund managers withdraw early. When $\alpha \in [0, \underline{\alpha_b}]$, fund managers have a dominant strategy to roll over their claims. Next, let $\overline{\alpha}_b \equiv \alpha_b^{IN} < 1$ denote the upper dominance bound, beyond which bank b fails regardless of the number of fund managers who withdraw early. When $\alpha \in [\overline{\alpha}_b, 1]$, fund managers have a dominant strategy to withdraw early. Finally, for $\alpha \in (\underline{\alpha_b}, \overline{\alpha}_b)$, if the shock were common knowledge, the run dynamics of fund managers would be characterised by multiple, self-fulfilling equilibria, as shown in Figure 3.

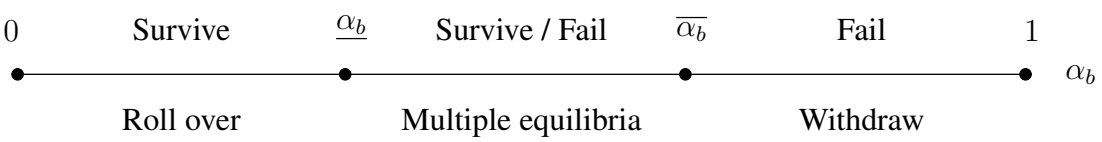

FIGURE 3. Tripartite classification of the accessibility shock.

Second, we define a threshold strategy which, for well defined dominance bounds and sufficiently precise private information, survives iterated deletion of strictly dominated strategies (Morris and Shin, 2003; Frankel et al., 2003). Denote this threshold point x_b^*, with corresponding strategy

$$s(x_{bk}) = \begin{cases} \text{withdraw} & \text{if } x_{bk} > x_b^*, \\ \text{roll over} & \text{if } x_{bk} \leq x_b^*. \end{cases} \tag{20}$$

Finally, we characterise this equilibrium. With switching point x_b^*, the proportion of fund managers who withdraw at $t = 1$, given some realisation of the shock α_b is

$$\ell_b(\alpha, x_b^*) = \Pr(x_{bk} > x_b^* | \alpha) = 1 - H\big(x_b^* - \alpha\big), \tag{21}$$

For $\gamma > \widehat{\gamma}$, the failure threshold, α_b^*, solves

$$\alpha_b^* = 1 - \ell_b^*(\alpha_b^*, x_b^*)\frac{FD}{RI_b}\,. \tag{22}$$

For a given x_b^*, the left-hand side of Equation 22 is strictly increasing in α_b^* and is unbounded over the entire unit interval, while the right-hand side is decreasing in α_b^* and is bounded between 1 and $1 - \frac{FD}{RI_b}$. Hence, there exists a unique failure threshold, α_b^*.

The posterior distribution of the shock conditional on the private signal can be derived using Bayes' rule. At the threshold signal x_b^*, fund managers are indifferent between withdrawing and rolling over, so that

$$\gamma = \Pr(\alpha \leq \alpha_b^* | x_{bk} = x_b^*)\,. \tag{23}$$

For small ϵ, this can be written as $\gamma = 1 - H\Big(x_b^* - \alpha_b^*\Big)$. The indifference condition therefore implies $x_b^* - \alpha_b^* = H^{-1}\big(1 - \gamma\big)$. Inserting this into $\ell\big(\alpha_b^*, x_b^*\big)$, the withdrawal proportion at the

threshold α_b^* becomes $\ell\Big(\alpha_b^*, x_b^*\Big) = 1 - H\Big(x_b^* - \alpha_b^*\Big) = 1 - H\Big(H^{-1}\Big(1-\gamma\Big)\Big) = \gamma$. Thus, for $\gamma \geq \widehat{\gamma}$, we have that $\alpha_b^* = \alpha_b^{IL}(\gamma)$. While for $\gamma < \widehat{\gamma}$, the bank fails only due to insolvency and so $\alpha_b^* = \alpha_b^{IN}$.

A.3. **Proof of Corollary 1.** The failure threshold for bank b is given by $\alpha_b^* = \alpha_b^{IL}(\gamma) = 1 - \frac{\gamma FD}{RI_b}$. This threshold is increasing in investment in assets,

$$\frac{\partial \alpha_b^*}{\partial I_b} = \frac{\gamma FD}{RI_b^2} > 0\,. \tag{24}$$

Since the bank fails whenever $\alpha > \alpha_b^*$, this reduces the event space where the bank fails and lowers its fragility.

A.4. **Proof of Proposition 2.** Our proofs proceeds in three steps. First, we suppose that $\gamma > \widehat{\gamma}$ such that bank failure is driven by illiquidity. Second, we derive each bank's optimal S_b^* and argue that it is a maximum. And finally, we return to our original supposition and argue that the equilibrium threshold $\widehat{\gamma}(S_b^*)$ is well defined.

Given our supposition $\gamma > \widehat{\gamma}$, each bank, $b = 1, \ldots, N$, chooses its cybersecurity contribution, S_b, and investment, I_b, to maximise its expected equity value, where the default threshold is $\alpha_b^* = 1 - \gamma\frac{FD}{RI_b}$ and taking as given the contributions to cybersecurity by all other banks, $\vec{S}_{-b}$. Substituting in the balance sheet constraint, $1 = I_b + S_b$, the first-order condition for bank b is given by

$$\begin{aligned}\frac{\partial \pi_b}{\partial S_b} &= \frac{\partial p}{\partial S_b}\left[R(1-S_b) - FD - \int_0^{\alpha_b^*(S_b)} EV_b(\alpha)d\alpha\right] \\ &- R\left[p + (1-p)\int_0^{\alpha_b^*(S_b)}(1-\delta\alpha)d\alpha\right] + (1-p)EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial S_b}\,.\end{aligned} \tag{25}$$

Since $p(\vec{S})$ satisfies the Inada conditions, it follows that $\lim_{S_b \to 0} \frac{\partial \pi_b}{\partial S_b} = +\infty > 0$ and $\pi_b(S_b = 0) > 0$. Moreover, it follows from economic reasoning that the bank would never choose $S_b \geq \bar{S}$ in equilibrium. If, by contradiction, we suppose that the bank choose $S_b = \bar{S}$, then even if the vulnerabilities are uncovered and patched, the bank immediately defaults since it has generated zero asset returns. And whenever the attacker is successful, the bank always defaults, too. Thus,

for $S_b = \bar{S}$, we have that $\pi_b(S_b = \bar{S}) < 0$, implying that this level of contribution to cybersecurity is not feasible. And so by Rolle's theorem, at least one optimum exists within the unit interval.

Next, we argue that the equilibrium is unique. To see this, note that evaluating the second-order derivative at the first-order condition yields

$$\begin{aligned}
\left.\frac{\partial^2 \pi_b}{\partial S_b^2}\right|_{S_b=S_b^*} &= \left.\frac{\partial^2 p/\partial S_b^2}{\partial p/\partial S_b}\right|_{S_b=S_b^*} R\left(p(S_b^*) + \Big(1-p(S_b^*)\Big)\int_0^{\alpha_b^*(S_b^*)}(1-\delta\alpha)d\alpha\right) \\
&- 2\left.\frac{\partial p}{\partial S_b}\right|_{S_b=S_b^*} R\left(1-\int_0^{\alpha_b^*(S_b^*)}(1-\delta\alpha)d\alpha\right) \\
&- \Big(1-p(S_b^*)\Big)\left.\frac{\partial \alpha_b^*}{\partial S_b}\right|_{S_b=S_b^*}\left[2(1-\delta\alpha_b^*(S_b^*))R + \left.\frac{\partial^2 p/\partial S_b^2}{\partial p/\partial S_b}\right|_{S_b=S_b^*} EV_b(\alpha_b^*(S_b^*))\right] \\
&+ 2EV_b(\alpha_b^*(S_b^*))\left.\frac{\partial \alpha_b^*}{\partial S_b}\right|_{S_b=S_b^*}\left(\frac{1-p(S_b^*)}{1-S_b^*} - \left.\frac{\partial p}{\partial S_b}\right|_{S_b=S_b^*}\right).
\end{aligned}$$

Under a symmetric equilibrium, $S_b^* = S^*$, for all b, we obtain $p(S^*) = \sqrt{\frac{cS^*}{\overline{V}}}$, $\left.\frac{\partial p}{\partial S_b}\right|_{S_b=S^*} = \frac{1}{2N}\sqrt{\frac{c}{\overline{V}S^*}}$, and $\left.\frac{\partial^2 p}{\partial S_b^2}\right|_{S_b=S^*} = \frac{1}{2NS^*}\sqrt{\frac{c}{\overline{V}S^*}}\left[\frac{1}{N}\left(\frac{1}{2}-\nu\right)-(1-\nu)\right]$. We can, thus, re-write the above as

$$\begin{aligned}
\left.\frac{\partial^2 \pi_b}{\partial S_b^2}\right|_{S=S^*} &= \frac{1}{S^*}\left[\frac{1}{N}\left(\frac{1}{2}-\nu\right)-(1-\nu)\right] R\left(p(S^*) + \Big(1-p(S^*)\Big)\int_0^{\alpha_b^*(S^*)}(1-\delta\alpha)d\alpha\right) \\
&- \frac{1}{N}\sqrt{\frac{c}{\overline{V}S^*}} R\left(1-\int_0^{\alpha_b^*(S^*)}(1-\delta\alpha)d\alpha\right) \\
&- \Big(1-p(S^*)\Big)\left.\frac{\partial \alpha_b^*}{\partial S_b}\right|_{S_b=S^*}\left(2(1-\delta\alpha_b^*(S^*))R + \frac{1}{S^*}\left[\frac{1}{N}\left(\frac{1}{2}-\nu\right)-(1-\nu)\right] EV_b(\alpha_b^*(S^*))\right) \\
&+ 2EV_b(\alpha_b^*(S^*))\left.\frac{\partial \alpha_b^*}{\partial S_b}\right|_{S_b=S^*}\left(\frac{1-p(S^*)}{1-S^*} - \left.\frac{\partial p}{\partial S_b}\right|_{S_b=S^*}\right).
\end{aligned}$$

Thus, given S^*, the above expression is strictly increasing in the degree of substitutability, ν. Consequently, whenever $\nu < \widehat{\nu}$, where $\widehat{\nu}$ is given by the solution to $\left.\frac{\partial^2 \pi_b}{\partial S_b^2}\right|_{S=S^*} = 0$, the contribution, S^* is a local maximum. Thus, any symmetric solution to the first-order condition is a local maximum, implying that the solution to the first-order condition is unique.

Finally, we return to our original supposition that $\gamma > \widehat{\gamma}$. The critical threshold is implicitly defined as

$$\widehat{\gamma} = \frac{1}{\delta} - \left(\frac{1}{\delta} - 1\right) \frac{R\Big(1 - S^*(\widehat{\gamma})\Big)}{FD}\,. \tag{26}$$

As we show in Proposition 4, a marginal increase in rollover risk increases banks' contributions to cybersecurity. And so the right-hand side of Equation (26) is increasing in $\widehat{\gamma}$. At $\widehat{\gamma} = 0$, the left-hand side of Equation (26) is smaller than the right-hand side. And at $\widehat{\gamma} = 1$, the right-hand side is less than one. To see this, note that $S^* < \overline{S}$. Thus, the highest possible value that the right-hand side can achieve is $\frac{1}{\delta} - \left(\frac{1}{\delta} - 1\right) \frac{R(1-\overline{S})}{FD}$. To see that this is (weakly) less than one, suppose by way of contradiction that it is strictly greater than one, i.e.,

$$\frac{1}{\delta} - \left(\frac{1}{\delta} - 1\right) \frac{R\Big(1 - \overline{S}\Big)}{FD} > 1 \quad \leftrightarrow \quad (1-\delta)\left(1 - \frac{R(1-\overline{S})}{FD}\right) > 0\,.$$

But, given the definition of $\overline{S}$, the above expression is zero. Thus, the right-hand side is never larger than one. Thus, by the intermediate value theorem, we have a well defined threshold, $\widehat{\gamma}$.

A.5. **Proof of Lemma 3.** The cross derivative of bank b's expected equity value with respect to S_b and the contribution $S_{b'}$ by bank b' is given by

$$\begin{aligned}\frac{\partial^2 \pi_b}{\partial S_b \partial S_{b'}} &= \frac{\partial^2 p}{\partial S_b \partial S_{b'}}\left[R(1-S_b) - FD - \int_0^{\alpha_b^*(S_b)} EV_b(\alpha)d\alpha\right] \\ &- R\frac{\partial p}{\partial S_{b'}}\left(1 - \int_0^{\alpha_b^*(S_b)} (1-\delta\alpha)d\alpha\right) - \frac{\partial p}{\partial S_{b'}} EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial S_b}\,.\end{aligned}$$

Evaluating this at the solution to the first-order condition, S_b^*, and using that

$$\frac{\partial^2 p}{\partial S_b \partial S_{b'}} = \frac{\partial p}{\partial S_{b'}}\left(\frac{\partial X/\partial S_b}{X(\vec{S})}\right)\left(\frac{1}{2} - \nu\right), \tag{27}$$

we get that

$$\begin{aligned}\left.\frac{\partial^2 \pi_b}{\partial S_b \partial S_{b'}}\right|_{S_b=S_b^*} &= \frac{\partial p}{\partial S_{b'}}\left\{\frac{1}{p(S_b^*)}(1-2\nu)R\int_0^{\alpha_b^*(S_b^*)}(1-\alpha\delta)d\alpha\right. \\ &- 2\nu R\left(1-\int_0^{\alpha_b^*(S_b^*)}(1-\alpha\delta)d\alpha\right) \\ &- \left.(1-p(S_b^*))EV_b(\alpha_b^*(S_b^*))\frac{\partial \alpha_b^*}{\partial S_b}\Big|_{S_b=S_b^*}\left(1+\frac{1}{p(S_b^*)}(1-2\nu)\right)\right\}.\end{aligned}$$

The above expression is strictly positive by our assumption that $\nu < 0$. By the implicit function theorem, if follows that $\frac{\partial^2 \pi_b}{\partial S_b \partial S_{b'}} > 0$ in equilibrium.

A.6. **Proof of Proposition 3.** We use the results of Van Zandt and Vives (2007) to show that monotone supermodularity of the ex ante investment decision is sufficient to establish the existence of a greatest and least Nash equilibrium.

Define bank b's best response correspondence, $BR_b(\vec{S}_{-b}) : \vec{S}_{-b} \to [0,1]$, as follows:

$$BR_b(\vec{S}_{-b}) \equiv \arg\max_{S_b \in [0,1]} \pi_b(S_b|\vec{S}_{-b})\,. \tag{28}$$

To prove that the investment decision is supermodular, it is sufficient to establish:

(1) The profit function, $\pi_b(S_b|\vec{S}_{-b})$, is supermodular:

$$\pi_b(S_b|\vec{S}_{-b}) + \pi_b(S_b'|\vec{S}_{-b}) \leq \pi_b(S_b \wedge S_b'|\vec{S}_{-b}) + \pi_b(S_b \vee S_b'|\vec{S}_{-b}),$$

where $S_b, S_b' \neq S_b \in [0,1]$ and $\wedge$ and $\vee$ denote the meet and join of investments S_b and S_b' respectively. Since the investment space is a lattice, it follows that the meet and join of S_b and S_b' are in the investment space. Supermodularity of the profit function follows from the concavity of the profit function.

(2) The action profile has increasing first differences:

$$\frac{\partial^2 \pi_b}{\partial S_b S_{b'}} \geq 0,$$

for all banks $b = 1, \ldots, N$ and $b' \neq b$. This holds from the Proof of Lemma 3.

Together, this is sufficient to establish that the investment decision is monotone supermodular. By the results of Van Zandt and Vives (2007), the best response mapping $BR_b(\vec{S}_{-b})$ must, therefore, contain a well-defined greatest ($\overline{BR}_b(\vec{S}_{-b})$) and least element ($\underline{BR}_b(\vec{S}_{-b})$) and the set of all greatest (least) best responses for each $b = 1, \ldots, N$ form a greatest and a least Nash equilibrium.

Next, we characterise these equilibria. It is straightforward to show that the zero-investment outcome is a (least) Nash equilibrium. If a bank, b', contributes nothing to cybersecurity, then $p = 0$, irrespective of what all other banks contribute. Therefore, it is never optimal for the other banks to contribute to cybersecurity either. This is a Nash equilibrium since

$$\pi_b(0|\vec{S}_{-b}) > \pi_b(S'_b|\vec{S}_{-b})\,, \quad \text{for} \quad S'_b > 0\,,$$

where $S_j = 0$ for some $j \neq b$. That is, given that some other bank, j, has invested 0, the profits accruing to b for investing any amount other than 0 are lower than those received by investing 0.

We can also show that the greatest Nash equilibrium is different from zero (i.e., $\underline{BR}_b(\vec{S}_{-b}) \neq \overline{BR}_b(\vec{S}_{-b})$) by a contradiction. Suppose that $S^*_{b'} > 0$ and that $BR_b(\vec{S}_{-b})|_{S_{b'}>0} = 0$. By the definition of $BR_b(\vec{S}_{-b})$, profits are maximised whenever S_b forms a best response. But since p satisfies the Inada conditions, it follows that $\partial\pi_b/\partial S_b|_{S_b\to 0} > 0$, which implies $BR_b(\vec{S}_{-b})|_{S_{b'}>0} > 0$, a contradiction. Finally, by symmetry, it follows that $S^*_b = S^*_{b'} = S^*$ for all $b \neq b'$.

A.7. **Proof of Proposition 4.** In deriving the comparative statics for the joint equilibrium, we first look at how each bank's best-response correspondence shifts following a marginal change in each exogenous variable. Since all banks' best-response correspondences respond in the same way, it follows from the supermodular structure of the cybersecurity investment game that the effects get compounded at the system level. Thus, the qualitative nature of the comparative static for each individual bank is preserved at the system level. As such, in what follows, we concentrate on deriving the comparative statics for each individual bank.

Attacker's effort cost. The cross derivative of bank b's expected profits with respect to its contribution to cybersecurity and the attacker's cost of effort is

$$\frac{\partial^2 \pi_b}{\partial S_b \partial c} = \frac{\partial p}{\partial c}\left\{\frac{\Delta EV_b}{2NS_b} - R\left(1 - \int_0^{\alpha_b^*(S_b)}(1-\delta\alpha)d\alpha\right) - EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial S_b}\right\},$$

where $\Delta EV_b \equiv \left[R(1-S_b) - FD - \int_0^{\alpha_b^*(S_b)} EV_b(\alpha)d\alpha\right]$. Evaluating the above expression at S_b^*, we get

$$\begin{aligned}\left.\frac{\partial^2 \pi_b}{\partial S_b \partial c}\right|_{S_b = S_b^*} &= \frac{\partial p(S_b^*)/\partial c}{p(S_b^*)}\left\{R\left(1 - \int_0^{\alpha_b^*(S_b^*)}(1-\delta\alpha)d\alpha\right) - EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial S_b}\right\} \\ &= \frac{c}{2}\left\{R\left(1 - \int_0^{\alpha_b^*(S_b^*)}(1-\delta\alpha)d\alpha\right) - EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial S_b}\right\} > 0\,.\end{aligned}$$

Thus, by the implicit function theorem, we have $\frac{\partial S_b^*}{\partial c} > 0$. Since the failure threshold does not depend directly on c, we get that

$$\frac{d\alpha_b^*}{dc} = \frac{\partial \alpha_b^*}{\partial S_b^*}\frac{\partial S_b^*}{\partial c} < 0\,.$$

Deadweight loss. The cross derivative of π_b with respect to S_b and δ is

$$\frac{\partial^2 \pi_b}{\partial S_b \partial \delta} = \frac{\partial p}{\partial S_b}\int_0^{\alpha_b^*(S_b)} \alpha R(1-S_b)d\alpha + (1-p)\int_0^{\alpha_b^*(S_b)} \alpha R d\alpha + (1-p)\frac{\partial EV_b(\alpha_b^*)}{\partial \delta}\frac{\partial \alpha_b^*}{\partial S_b},$$

where

$$\frac{\partial EV_b(\alpha_b^*)}{\partial \delta} = -R(1-S_b)(1-\alpha_b^*) < 0\,.$$

Since $\frac{\partial \alpha_b^*}{\partial S_b} < 0$, it follows that $\frac{\partial^2 \pi_b}{\partial S_b \partial \delta} < 0$ and so by the implicit function theorem, $\frac{\partial S_b^*}{\partial \delta} > 0$.

Since $\alpha_b^* = \alpha_b^{IL}(\gamma)$ is independent of δ, fragility is affected only indirectly by the deadweight loss:

$$\frac{d\alpha_b^*}{d\delta} = \frac{\partial \alpha_b^*}{\partial S_b}\frac{\partial S_b^*}{\partial \delta} < 0\,. \tag{29}$$

Rollover risk. The cross derivative of π_b with respect to S_b and γ is given by

$$\frac{\partial^2 \pi_b}{\partial S_b \partial \gamma} = -\frac{\partial p}{\partial S_b}EV_b(\alpha_b^*)\frac{\partial \alpha_b^*}{\partial \gamma} + (1-p)\left\{-R(1-\delta\alpha_b^*)\frac{\partial \alpha_b^*}{\partial \gamma} + \frac{\partial EV_b(\alpha_b^*)}{\partial \gamma}\frac{\partial \alpha_b^*}{\partial S_b} + EV_b(\alpha_b^*)\frac{\partial^2 \alpha_b^*}{\partial S_b \partial \gamma}\right\}.$$

Since $\partial\alpha_b^*/\partial\gamma < 0$, it follows that the first term is strictly positive. Next, we can re-write the terms that multiply into $1-p$ as

$$\begin{aligned} & R\left(1-\delta\left(1-\frac{\gamma FD}{R(1-S_b)}\right)\right)\frac{FD}{R(1-S_b)} - \frac{(\gamma FD)^2}{R(1-S_b)^2} - EV_b(\alpha_b^*)\frac{FD}{R(1-S_b)^2} \\ = \; & \frac{(1-\delta)FD}{R(1-S_b)} + \frac{\delta\gamma(FD)^2}{R(1-S_b)^2} - \frac{\delta\gamma(FD)^2}{R(1-S_b)^2} - \frac{(1-\delta)FD}{R(1-S_b)} + \frac{(1-\gamma\delta)(FD)^2}{R(1-S_b)^2} > 0\,. \end{aligned}$$

Thus, the cross derivative is strictly positive and so by the implicit function theorem, $\frac{\partial S_b^*}{\partial\gamma} > 0$.

An increase in rollover risk introduces a direct and indirect effect on bank fragility.

$$\frac{d\alpha_b^*}{d\gamma} = \frac{\partial\alpha_b^*}{\partial\gamma} + \frac{\partial\alpha_b^*}{\partial S_b}\frac{\partial S_b^*}{\partial\gamma} < 0\,. \tag{30}$$

For each outage shock, the likelihood of bank failure increases in γ, which represents the equilibrium proportion of fund manager withdrawals in Proposition 1, and so $\partial\alpha_b^*/\partial\gamma < 0$. Furthermore, banks reallocate resources towards cybersecurity in an effort to avoid an outage entirely which, by Corollary 1, increases bank fragility.

Bank equity. The cross derivative of π_b with respect to S_b and D is given by

$$\begin{aligned} \frac{\partial^2\pi_b}{\partial S_b\partial D} &= \frac{\partial p}{\partial S_b}\frac{\partial\Delta EV_b}{\partial D} - (1-p)R(1-\delta\alpha_b^*)\frac{\partial\alpha_b^*}{\partial D} + (1-p)\left[\frac{\partial EV_b(\alpha_b^*)}{\partial D}\frac{\partial\alpha_b^*}{\partial S} + EV_b(\alpha_b^*)\frac{\partial^2\alpha_b^*}{\partial S\partial D}\right] \\ &= \frac{p}{2NS_b}\left[-F(1-\alpha_b^*) + \frac{EV_b(\alpha_b^*)\gamma F}{R(1-S_b)}\right] + (1-p)\frac{\gamma F}{1-S_b}\left[1-\delta\alpha_b^* + \frac{(1-\gamma\delta)FD}{R(1-S_b)}\right. \\ &- \left.\frac{EV_b(\alpha_b^*)}{R(1-S_b)}\right]. \end{aligned}$$

The term multiplying into $\partial p/\partial S_b$ may be expressed as

$$\begin{aligned} & -F + F\left(1-\frac{\gamma FD}{R(1-S_b)}\right) + (1-\delta)R\gamma F - \frac{(1-\gamma\delta)\gamma F^2 D}{R(1-S_b)} \\ = \; & -\frac{\gamma F^2 D}{R(1-S_b)} + (1-\delta)R\gamma F - \frac{(1-\gamma\delta)\gamma F^2 D}{R(1-S_b)} \\ = \; & (1-\delta)R\gamma F + \frac{\delta\gamma^2 F^2 D}{R(1-S_b)} > 0\,. \end{aligned}$$

Next, the term that multiplies into $(1-p)\frac{\gamma F}{1-S_b}$ can be written as

$$\begin{aligned} & 1-\delta+\frac{\delta\gamma FD}{R(1-S_b)}+\frac{1}{R(1-S_b)}-\frac{\delta\gamma FD}{R(1-S_b)}+\frac{1}{R(1-S_b)}-\frac{EV_b(\alpha_b^*)}{R(1-S_b)} \\ = \; & 1-\delta+\frac{1}{R(1-S_b)}-(1-\delta)+\frac{(1-\gamma\delta)FD}{R(1-S_b)} \\ = \; & \frac{1}{R(1-S_b)}+\frac{(1-\gamma\delta)FD}{R(1-S_b)}>0\,. \end{aligned}$$

Thus, $\frac{\partial^2\pi_b}{\partial S_b\partial D}>0$ and by the implicit function theorem, $\partial S_b^*/\partial D>0$. Finally, since $E=1-D$, it follows that $\partial S_b^*/\partial E<0$.

An increase bank equity introduces a direct and indirect effect on bank fragility, which re-enforce each other:

$$\frac{d\alpha_b^*}{dE}=\frac{\partial\alpha_b^*}{\partial E}+\frac{\partial\alpha_b^*}{\partial S_b}\frac{\partial S_b^*}{\partial E}>0\,. \tag{31}$$

A.8. **Proof of Proposition 5.** We first set out the optimisation problem for the planner and derive the benchmark S^P. The planner optimises the sum of banks' profits,

$$\Pi=\sum_{b=1}^{N}\left[\sqrt{\frac{cX}{\overline{V}}}(RI_b-F\,D)+\left(1-\sqrt{\frac{cX}{\overline{V}}}\right)\int_0^{\alpha_b^*(I_b)}EV_b(\alpha)\,d\alpha\right], \tag{32}$$

subject to the level of cybersecurity,

$$X=\left[\left(\frac{1}{N}\right)\sum_{j=1}^{N}S_j^{\nu}\right]^{1/\nu},$$

and the banks' budget constraints, $1=S_b+I_b$ for all $b=1,\ldots,N$. The planner's Lagrangian is, thus,

$$\mathcal{L}=\Pi+\sum_{b=1}^{N}\phi_b\,(1-I_b-S_b)+\lambda\left(\left[\left(\frac{1}{N}\right)\sum_{j=1}^{N}S_j^{\nu}\right]^{1/\nu}-X\right). \tag{33}$$

For all $b = 1, \ldots, N$, the necessary and sufficient Kuhn-Tucker conditions are given by

$$(1) \quad \frac{\partial \mathcal{L}}{\partial I_b} = 0 : \sqrt{\frac{c\,X}{\overline{V}}} R + \left(1 - \sqrt{\frac{c\,X}{\overline{V}}}\right) \left\{ \int_0^{\alpha_b^*(I_b)} \frac{\partial EV_b}{\partial I_b} d\alpha + EV_b(\alpha_b^*) \frac{\partial \alpha_b^*}{\partial I_b} \right\} = \phi_b;$$

$$(2) \quad \frac{\partial \mathcal{L}}{\partial X} = 0 : \frac{1}{2} \sqrt{\frac{c}{\overline{V} X}} \sum_{b'=0}^{N} \left\{ RI_{b'} - F\,D - \int_0^{\alpha_{b'}^*(I_{b'})} EV_b(\alpha) d\alpha \right\} = \lambda;$$

$$(3) \quad \frac{\partial \mathcal{L}}{\partial S_b} = 0 : \lambda \frac{\left[\left(\frac{1}{N}\right) \sum_{j=1}^{N} S_j^{\nu} \right]^{1/\nu} S_b^{\nu-1}}{\sum_{j=1}^{N} S_j^{\nu}} = \phi_b \,.$$

From the above conditions, it is clear that

$$\frac{\frac{1}{2}\sqrt{\frac{c}{\overline{V}X}} \sum_{b'=0}^{N} \left\{ RI_{b'} - F\,D - \int_0^{\alpha_{b'}^*(I_{b'})} EV_b(\alpha) d\alpha \right\}}{\phi_b} = \frac{1}{\left(\left[\left(\frac{1}{N}\right) \sum_{j=1}^{N} S_j^{\nu} \right]^{1/\nu} S_b^{\nu-1} \right) / \sum_{j=1}^{N} S_j^{\nu}},$$

for all $b = 1, \ldots, N$. Substituting for ϕ_b from condition (1), we have

$$\sum_{b'=1}^{N} \frac{\frac{1}{2}\sqrt{\frac{c}{\overline{V}X}} \left\{ RI_{b'} - F\,D - \int_0^{\alpha_{b'}^*(I_{b'})} EV_b(\alpha) d\alpha \right\}}{\sqrt{\frac{c\,X}{\overline{V}}} R + \left(1 - \sqrt{\frac{c\,X}{\overline{V}}}\right) \left\{ \int_0^{\alpha_b^*(I_b)} \frac{\partial EV_b}{\partial I_b} d\alpha + EV_b(\alpha_b^*) \frac{\partial \alpha_b^*}{\partial I_b} \right\}} = \frac{1}{\partial X / \partial S_b}, \tag{34}$$

which is equivalent to

$$\frac{\partial \pi_b / \partial p}{\partial \pi_b / \partial I_b} + \sum_{k \neq b}^{N} \frac{\partial \pi_k / \partial p}{\partial \pi_b / \partial I_b} = \left(\frac{p}{2X(\vec{S})} \frac{\partial X}{\partial S_b} \right)^{-1} \quad \forall b = 1, \ldots, N \,. \tag{35}$$

By symmetry, it follows that the solutions satisfy $S_b^P = S^P$ for all $b = 1, \ldots, N$.

An increase in rollover risk causes the numerator terms to increase, since $\partial\alpha_b^*/\partial\gamma < 0$ for all banks. With regards to the terms in the denominator, note that

$$\begin{aligned}
\frac{\partial^2\pi_b}{\partial I_b\partial\gamma} &\propto (1-\delta\alpha_b^*)R\frac{\partial\alpha_b^*}{\partial\gamma} + \frac{\partial EV_b(\alpha_b^*)}{\partial\gamma}\frac{\partial\alpha_b^*}{\partial I_b} + EV_b(\alpha_b^*)\frac{\partial^2\alpha_b^*}{\partial I_b\partial\gamma} \\
&= -\left(1-\delta\left(1-\gamma\frac{FD}{RI_b}\right)\right)\frac{FD}{I_b} + \gamma\delta\frac{(FD)^2}{RI_b^2} + EV_b(\alpha_b^*)\frac{\partial^2\alpha_b^*}{\partial I_b\partial\gamma} \\
&= -(1-\delta)\frac{FD}{I_b} - \gamma\delta\frac{(FD)^2}{RI_b^2} + \gamma\delta\frac{(FD)^2}{RI_b^2} + (1-\delta)\frac{FD}{I_b} - (1-\gamma\delta)\frac{(FD)^2}{RI_b^2} \\
&= -(1-\gamma\delta)\frac{(FD)^2}{RI_b^2} < 0\,.
\end{aligned}$$

Thus, following a marginal increase in rollover risk, the planner allocates even more towards cybersecurity than the bank would privately choose.

A.9. **Proof of Proposition 6.** With a negligence rule in place, expected profits include a penalty, $\kappa_b(I_b)$, that is implemented conditional on a successful cyber attack

$$\begin{aligned}
\pi_b(S_b, I_b) &= p(S_b, \vec{S}_{-b})\,(R\,I_b - F\,D) \\
&+ \left(1-p(S_b,\vec{S}_{-b})\right)\int_0^{\alpha_b^*(I_b)}\Big[R\,I_b(1-\alpha\,\delta) - F\,D - \kappa_b(I_b)\Big]\,d\alpha\,.
\end{aligned}$$

The introduction of a negligence rule lowers expected profits in all events where a cyber attack is successful and the bank survives, increasing the relative benefits from investing more in cybersecurity. For the penalty to be effective in eliciting the social optimum, it must satisfy two conditions:

$$\frac{R\,I_b - F\,D}{R\,I_b\delta} - \frac{\kappa_b(I_b)}{R\,I_b\delta} > \alpha_b^*(I_b) \tag{36}$$

i.e., the penalty is not so large that it leads the bank to fail for any successful attack, and

$$\pi_b(S_b^P, I_b^P) \geq \pi_b(S_b, I_b)\,,$$

for all $S_b \leq S_b^R$ and $I_b \geq I_b^R$. That is, the penalty should be large enough that profits subject to a negligence rule under the social optimum are preferable to those in the laissez faire optimum.

Suppose the penalty is structured in the following way

$$\kappa_b(I_b;\gamma) = \begin{cases} \kappa_b \times I_b & \text{if } I_b > I_b^P(\gamma) \\ 0 & \text{otherwise}\,. \end{cases} \tag{37}$$

Then the penalties, $\{\kappa_j^*\}_{j=1}^N$, that deliver the social optimum are each given by

$$\frac{R(1-S_j) - F\,D - \int_0^{\alpha_j^*(1-S_j)} EV_j(\alpha) - \kappa_j^*\, I_j d\alpha}{pR + (1-p)\left[\int_0^{\alpha_j^*(1-S_j)} \left(\frac{\partial EV_j}{\partial I_j} - \kappa_j^*\right) d\alpha + (EV_j(\alpha_j^*) - \kappa_j^*\, I_j)\frac{\partial \alpha_j^*}{\partial I_j}\right]} = \tag{38}$$

$$\sum_{b=1}^{N} \frac{R(1-S_b) - F\,D - \int_0^{\alpha_b^*(1-S_b)} EV_b(\alpha) d\alpha}{pR + (1-p)\left[\int_0^{\alpha_j^*(1-S_j)} \frac{\partial EV_j}{\partial I_j} d\alpha + EV_j(\alpha_j^*)\frac{\partial \alpha_j^*}{\partial I_j}\right]}\,.$$

The left-hand side of (38) is the marginal rate of substitution between the public good and investment in assets for bank j, and the right-hand side is the sum of the ratio of marginal returns to each bank from an increase in cybersecurity to the marginal returns to bank j from an increase in investment in assets. Taking the partial derivative of the left-hand side with respect to the penalty, we have

$$\frac{I_j \alpha_j^* \xi(I_j) + \left(R - F\,D - \int_0^{\alpha_j^*} \left[EV_j(\alpha) - \kappa_j^* I_j\right] d\alpha\right)(1-p)\left[\alpha_j^* + I_j \frac{\partial \alpha_j^*}{\partial I_j}\right]}{\left[\xi(I_j)\right]^2} > 0, \tag{39}$$

where $\xi(I_b) = pR + (1-p)\left[\int_0^{\alpha_j^*(1-S_j)} \frac{\partial EV_j}{\partial I_j} d\alpha + EV_j(\alpha_j^*)\frac{\partial \alpha_j^*}{\partial I_j}\right]$. With the left-hand side of (38) increasing in κ_j^*, an optimal penalty exists by the intermediate value theorem.

The size of the optimal penalty is affected by the degree of rollover risk. As we have just shown, the marginal rate of substitution for bank j is increasing in κ_j^*. Further, as we show in the proof of Proposition 5, the degree of under-investment is also increasing in γ (so the marginal effect of an increase in γ on the private rate of substitution is smaller than on the regulator's rate of substitution). Therefore, by the implicit function theorem, it must hold that $\partial \kappa_j^* / \partial \gamma \geq 0$.

A.10. **Proof of Proposition 7.** With a subsidy scheme in place, bank b's expected profits include subsidy, $\sigma_b(S_b)$, and lump-sum tax, τ,

$$\pi_b(S_b, I_b) = p(S_b, \vec{S}_{-b})(R\, I_b - F\, D) + \left(1 - p(S_b, \vec{S}_{-b})\right) \int_0^{\alpha_b^*(I_b)} EV_b(\alpha) d\alpha + \sigma_b(S_b) - \tau\,. \quad (40)$$

For the subsidy to elicit the social optimum, the individual bank's marginal rate of substitution should equal the planner's. Suppose the subsidy is structured as follows

$$\sigma_b(S_b) = \sum_{j \neq b}^{N} \left(p(S_b, \vec{S}^P_{-b}) \left[R\, I_j^P - F\, D - \int_0^{\alpha_j^*(I_j^P)} EV_j(\alpha) d\alpha \right] \right)\,, \quad (41)$$

for each bank $b = 1, \ldots, N$. To fund these subsidies, suppose a uniform lump-sum tax is set such that

$$\tau = \frac{T}{N} = \frac{1}{N} \sum_{j=1}^{N} \sigma_j\,. \quad (42)$$

Then the subsidies, $\{\sigma_b^*\}_{b=1}^N$, that deliver the social optimum are each given by

$$\frac{R(1 - S_b) - F\, D - \int_0^{\alpha_b^*(1-S_b)} EV_b(\alpha) d\alpha + \frac{\partial \sigma_b^*}{\partial p}}{pR + (1-p)\left[\int_0^{\alpha_b^*(I_b)} \frac{\partial EV_b}{\partial I_b} d\alpha + EV_b(\alpha_b^*) \frac{\partial \alpha_b^*}{\partial I_b}\right]} = \left(\frac{p(\vec{S})}{2X(\vec{S})} \frac{\partial X}{\partial S_b} \right)^{-1} \quad \forall b = 1, \ldots, N\,. \quad (43)$$

The system in (43) is equivalent to that of the planner. The left-hand side is the marginal rate of substitution between increased protection from cyber attacks and investment in assets for bank b. With subsidy σ_b^* in place, each bank's marginal rate of substitution increases, resulting in a higher level of protection $X(\vec{S}^P) = S^P$.

The subsidy size is impacted by rollover risk. As we show in the proof of Proposition 5, the degree of under-investment is increasing in γ. The individual marginal rate of substitution in (43) is increasing in the size of the subsidy σ_b^*. Therefore, by the implicit function theorem, it must hold that $\partial \sigma_b^* / \partial \gamma \geq 0$.

Under a symmetric equilibrium, $\sigma_b^* = \sigma^*$ for all $b = 1, \ldots, N$ and the lump sum tax is given by $\tau = \sigma^*$.

A.11. **Proof of Proposition 8.** We present the proof in three parts. First, we consider the case where the disruption is so large that a bank always fails due to illiquidity. From this, we obtain an endogenous upper bound for the shock. Second, we consider the converse case where the shock is sufficiently small that the bank never fails, which yields a second lower bound. Finally, we compare the two bounds. In the limit $\nu \to 0^+$, we can express cybersecurity as

$$X(\vec{S}) = \left(\prod_{j=1}^{N} S_j \right)^{1/N} .$$

Suppose that $\alpha > \alpha_b^*$. Bank b's expected profits are given by $\pi_b = p(\alpha, S_b, \vec{S}_{-b}) \Big[R(1 - S_b) - FD \Big]$. The first order condition, $\frac{\partial \pi_b}{\partial S_b} = 0$ yields

$$\begin{aligned} & \frac{\partial p}{\partial S_b} \Big[R(1 - S_b) - FD] - pR = 0 \\ \implies \quad & \frac{p}{2NS_b} \Big[R(1 - S_b) - FD] - pR = 0 \\ \implies \quad & p \left\{ \frac{R(1 - S_b) - FD - 2NRS_b}{2NS_b} \right\} = 0 \,, \end{aligned}$$

and so

$$S_b^* = \frac{R - FD}{R(2N + 1)} \,.$$

Under symmetry, $S_b^* = S_{b'}^*$, for all $b' \neq b$. We can thus drop the bank index subscript. Returning to our supposition that $\alpha > \alpha^*$, we obtain that the equilibrium failure threshold for high-grade cyber attacks is given by

$$\widehat{\alpha}^* \equiv \alpha^*(S^*) = \frac{R\tilde{N} + \left(\frac{1}{2N+1} - \gamma \right) FD}{R\tilde{N} + \frac{FD}{2N+1}} \,,$$

where $\tilde{N} \equiv \frac{2N}{2N+1}$.

Next, suppose $\alpha < \alpha_b^*$. In this case, the bank does not fail following a cyber attack and the bank's expected profits are given by $\pi_b = p \Big[R(1 - S_b) - FD \Big] + (1 - p) \Big[(1 - \delta\alpha) R(1 - S_b) - FD \Big]$.

From the first-order condition, we obtain

$$\begin{aligned} & \frac{\partial p}{\partial S_b}\delta\alpha R(1-S_b) - pR - (1-p)(1-\delta\alpha)R = 0 \\ \implies \quad & \frac{p}{2NS_b}\delta\alpha R(1-S_b) - pR - (1-p)(1-\delta\alpha)R = 0\,, \end{aligned}$$

which implies

$$S_b^* = \frac{p\delta\alpha(1-S_b^*)}{2N\Big(1-(1-p)\delta\alpha\Big)}\,.$$

Under symmetry, we have that $S_b^* = S_{b'}^* = S^*$ and $p = S^*$. We thus obtain

$$S^* = 1 - \frac{\tilde{N}}{\delta\alpha}\,.$$

Returning to our initial supposition that $\alpha < \alpha^*$, evaulating the failure threshold at S^*, we get

$$\begin{aligned} & \alpha < \alpha^*(S^*) \equiv 1 - \frac{\delta\alpha\gamma FD}{R\tilde{N}} \\ \implies \quad & \alpha < \widehat{\alpha}^{**} \equiv \frac{R\tilde{N}}{R\tilde{N} + \delta\gamma FD}\,. \end{aligned}$$

Finally, comparing the two failure thresholds, we obtain that as long as $\gamma > \frac{1}{2N+1}$ and $\gamma < \frac{1}{\delta(2N+1)}$, we have that $\widehat{\alpha}^{**} > \widehat{\alpha}^*$.

APPENDIX B. ENDOGENISING THE FACE VALUE OF DEBT

In this section, we endogenise the face value and show that the main trade-offs and insights are unaffected in this more generalised model.

The value of a debt claim issued by bank b, which we denote by $V(F, \vec{S}^e)$, depends on creditors' expectations over how much banks invest in cybersecurity, $\vec{S}^e$. Under the simplifying assumption that creditors receive nothing in the event of a bank failure, we obtain

$$V_b(F_b, \vec{S}^e) \equiv p(S_b, \vec{S}_{-b})F_b + \Big(1 - p(S_b, \vec{S}_{-b})\Big)\alpha_b^*(S_b^e, F_b)F_b\ .$$

So if the cyber attack is unsuccessful, creditors are repaid in full for sure. While, if the attacker is successful and the bank suffers an outage, then creditors are only repaid if the bank does not fail due to a run. To keep notation succinct in what follows, we suppress the dependency of bank b's failure threshold on its contribution to cybersecurity and face value of debt.

If creditors have access to a safe outside investment option that yields $r^* > 0$, the face value of debt under perfect competition is given by the solution to

$$V(F_b, \vec{S}^e) = r^* \,. \tag{44}$$

Lemma 4. *If the dead-weight loss is sufficiently high, $\delta > 1 - 2\gamma$, the equilibrium face value, $F_b^*(S_b^e)$, is increasing in cybersecurity investment, $\frac{\partial F_b^*}{\partial S_b^e} > 0$.*

Proof. An increase in face value, increases the value function

$$\begin{aligned} \frac{\partial V}{\partial F_b} &= p + \Big(1-p\Big)\left[\alpha_b^* + F_b \frac{\partial \alpha_b^*}{\partial F_b}\right] = p + \Big(1-p\Big)\left[\alpha_b^* - \frac{\gamma F_b D}{R(1-S_b)}\right] \\ &= p + (1-p)(2\alpha_b^* - 1)\,. \end{aligned}$$

A sufficient condition for $\partial V/\partial F_b > 0$ is to require that $2\alpha_b^* - 1 > 0$. Suppose we define $\tilde{S}_b$ to be the level of contribution to cybersecurity at which $\alpha_b^* = \bar{\alpha}_b$. Then, if $2\alpha_b^*(\tilde{S}_b) - 1 > 0$, then for all $S_b < \tilde{S}_b$, this condition will also hold. This, in turn, requires that $\delta > 1 - 2\gamma$.

Next, to see that the value function is decreasing in S_b, note that

$$\begin{aligned} \frac{\partial V}{\partial S_b} &= \frac{\partial p}{\partial S_b} F_b(1-\alpha_b^*) + \Big(1-p\Big)\frac{\partial \alpha_b^*}{\partial S_b} \\ &= \frac{\gamma F_b^2 D}{R(1-S_b)}\left[\frac{p}{2NS_b} - \frac{1-p}{1-S_b}\right] < 0\,. \end{aligned}$$

Thus, we have by the implicit function that

$$\frac{\partial F_b^*}{\partial S_b} = \frac{-1}{\partial V/\partial F_b}\left(\frac{\partial V}{\partial S_b}\right) > 0. \tag{45}$$

□

Creditors demand compensation for additional cybersecurity investment to offset the heightened fragility to which the bank is exposed in the event of an attack. The marginal from returns investing in assets are large relative to contributing to cybersecurity. So each creditor lends to the bank at a rate that depends positively on expectations of the bank's investment in cybersecurity. The bank, for its part, takes face value as given when setting its optimal allocation.

Proposition 9. *There is a unique equilibrium* (S_b^*, F_b^*) *such that creditors' participation constraints are satisfied and the bank optimally chooses its cybersecurity investment.*

Proof. First, we show that bank b's contribution to cybersecurity is increasing in F_b. This, in turn, requires us to sign the cross derivative of bank b's expected profits with respect to S_b and F_b, which is given by

$$\begin{aligned}\frac{\partial^2\pi}{\partial S_b\partial F_b} &= \frac{\partial p}{\partial S_b}\left\{-D(1-\alpha_b^*) - EV_b(\alpha_b^*)\frac{\partial\alpha_b^*}{\partial F_b}\right\} \\ &+ (1-p)\left\{\frac{\partial EV_b(\alpha_b^*)}{\partial F_b}\frac{\partial\alpha_b^*}{\partial S_b} + EV_b(\alpha_b^*)\frac{\partial^2\alpha_b^*}{\partial S_b\partial F_b} - R(1-\delta\alpha_b^*)\frac{\partial\alpha_b^*}{\partial F_b}\right\}.\end{aligned}$$

The terms that multiply into $1-p$ on the second line may be expressed as

$$\begin{aligned}\frac{(1-\gamma\delta)\gamma D^2F_b}{R(1-S_b)^2} &- \Big[(1-\delta\alpha_b^*)R(1-S_b) - DF_b\Big]\frac{\gamma D}{R(1-S_b)^2} + (1-\delta\alpha_b^*)R\frac{\gamma D}{R(1-S_b)} \\ &= \frac{(1-\gamma\delta)\gamma D^2F_b}{R(1-S_b)^2} + \frac{\gamma D^2F_b}{R(1-S_b)^2} > 0\,.\end{aligned}$$

As for the terms that multiply into $\partial p/\partial S_b$ on the first line, we can express them as $\frac{\gamma D}{R(1-S_b)}\Big[EV_b(\alpha_b^*) - DF_b\Big]$. And so, combining everything together, we obtain

$$\begin{aligned}\frac{\partial^2\pi}{\partial S_b\partial F_b} &= \frac{\partial p}{\partial S_b}\frac{\gamma D}{R(1-S_b)}\Big[EV_b(\alpha_b^*) - DF_b\Big] + \frac{(1-p)(2-\gamma\delta)\gamma D^2F_b}{R(1-S_b)^2} \\ &= \frac{\gamma D}{R(1-S_b)}\left\{\frac{p}{NS_b}(1-\delta)R(1-S_b) - (2-\gamma\delta)DF_b\left(\frac{p}{NS_b} - \frac{1-p}{1-S_b}\right)\right\} > 0\,.\end{aligned}$$

Thus, by the implicit function theorem, it follows that $\frac{\partial S_b^*}{\partial F_b} > 0$.

Next, for a unique intersection between the curves $S_b^*(F_b)$ and $F_b^*(S_b)$, we require that (i) the value $\tilde{F}_b$ such that $S_b^*(\tilde{F}_b) = 1$ satisfies $\tilde{F}_b > F_b^*(1)$; and (ii) the value $\underset{\sim}{F}_b$ such that $S_b^*(\underset{\sim}{F}_b) = 0$ satisfies $\underset{\sim}{F}_b < r^*$. □

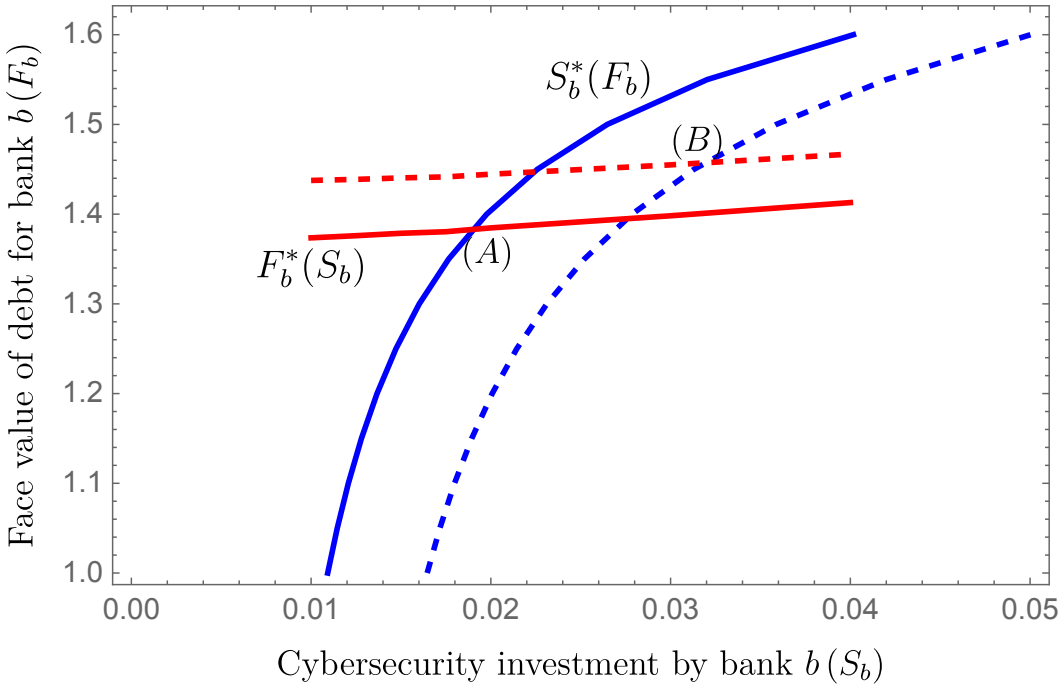

FIGURE 4. The face value of debt demanded by creditors, F_b^*, and level of cybersecurity investment initiated by each bank, S_b^*, are uniquely determined by the intersection (A) of bank b's reaction function (blue line) and creditors' reaction function (red line). Following an increase in rollover risk, the curves shift and the new equilibrium is given by intersection (B).

Figure 4 illustrates the equilibrium. The laissez-faire outcome that arises with an endogenous face value retains the link between rollover risk (ex post coordination failure) and cybersecurity investment (ex ante free riding on the public good). As the optimal level of cybersecurity increases, creditors demand higher face value to choose bank debt over their safe outside option. With a higher face value, bank b is made more fragile for every proportion of early withdrawals in the event that a cyber attack is successful, and so it is better off allocating its working capital towards cybersecurity. Since each bank only internalises its own marginal rate of substitution and face value, the free riding problem remains. The unique equilibrium (S_b^*, F_b^*) satisfies creditors' participation constraints and equates the bank's private marginal rate of substitution with its marginal rate of transformation.

Proposition 10. *Bank b's contribution to cybersecurity is increasing in rollover risk, $\frac{dS_b^*}{d\gamma} > 0$.*

Proof. From a simply graphical analysis of the problem, we note that the total effect of rollover risk may be decomposed into the effects of the bank's best-response, $S_b^*(F_b)$, and that of the investors, $F_b^*(S_b)$. From our earlier analysis in the main text, we know that the direct effect of an increase in γ amounts to an outward-shift in the best-response schedule. To understand how the investors' best-response curve shifts, note that

$$\frac{\partial F_b^*}{\partial \gamma} = \frac{-F_b(1-p)\frac{\partial \alpha_b^*}{\partial \gamma}}{\partial V/\partial F_b} > 0. \tag{46}$$

And so, since a marginal increase in γ leads to an upward shift $F_b^*(S_b)$, it follows that the total effect of an increase in rollover risk is to increase the bank's contribution to cybersecurity. □

Following an increase in γ, each fund manager k is less likely to rollover claims at $t = 1$ for any given signal x_{bk}. Ex ante, this means that bank b is more likely to fail for any realisation of the shock, α, and a given F_b. So the marginal benefit of an additional unit of cybersecurity is high relative to the marginal cost of further investment in assets. The bank is therefore better off shoring up cybersecurity, even though this means an increase in fragility in the event of a successful attack. At the same time, with high rollover risk, creditors demand higher face value to choose bank debt over their safe outside option.

The effects on cybersecurity and face value are, thus, self-reinforcing. This, in turn, implies that endogenising the face value of debt does not qualitatively alter the key results in the main paper with an exogenous face value of debt.

REFERENCES

Acemoglu, D., A. Malekian, and A. Ozdaglar (2016). Network security and contagion. Journal of Economic Theory 116, 536–585.

Adelmann, F., I. Ergen, T. Gaidosch, N. Jenkinson, A. Morozova, N. Schwarz, and C. Wilson (2020). Cyber risk and financial stability: It's a small world after all. Staff Discussion Notes (007), International Monetary Fund, Washington, DC.

Ahnert, T., K. Anand, J. Chapman, and P. Gai (2019). Asset Encumbrance, Bank Funding and Fragility. Review of Financial Studies 32(6), 2422–2455.

Bergstrom, T., L. Blume, and H. Varian (1986). On the private provision of public goods. Journal of Public Economics 29(1), 25–49.

Biener, C., M. Eling, and J. H. Wirfs (2015). Insurability of cyber risk: An empirical analysis. The Geneva Papers on Risk and Insurance-Issues and Practice 40(1), 131–158.

Bier, V., O. Santiago, and L. Samuelson (2007). Choosing what to protect: Strategic defensive allocation against an unknown attacker. Journal of Public Economic Theory 9, 563–587.

Bliss, C. and B. Nalebuff (1984). Dragon-slaying and ballroom dancing: The private supply of a public good. Journal of Public Economics 25(1-2), 1–12.

Brown, J. P. (1973). Toward an economic theory of liability. The Journal of Legal Studies 2(2), 323–349.

Cornes, R. (1993). Dyke maintenance and other stories: Some neglected types of public goods. The Quarterly Journal of Economics 108(1), 259–271.

Dang, T. V., G. Gorton, B. Holmström, and G. Ordonez (2017). Banks as secret keepers. American Economic Review 107(4), 1005–29.

Dixit, A. (1987). Strategic behavior in contests. American Economic Review 77(5), 891–898.

Duffie, D. and J. Younger (2019). Cyber runs. Hutchins Center Working Paper 51, Brookings Institution.

Dziubinski, M. and S. Goyal (2013). Network design and defence. Games and Economic Behavior 79, 30–43.

Eisenbach, T., A. Kovner, and M. J. Lee (2022). Cyber risk and the U.S. financial system: A pre-mortem analysis. Journal of Financial Economics, 145, 802–826.

Elestedt, L., U. Nilsson, and C.-J. Rosenvinge (2021). A cyber attack can affect financial stability. Economic Commentary No. 8, Sveriges Riksbank, Stockholm.

European Systemic Risk Board (2022). Mitigating systemic cyber risk. Technical Report, Frankfurt.

Federal Deposit Insurance Corporation (2022). Report on cybersecurity and resilience. Technical Report, Washington, DC.

Fell, J., N. de Vette, S. Gardó, B. Klaus, and W. Wendelborn (2022). Towards a framework for assessing systemic cyber risk. Financial Stability Review, European Central Bank, Frankfurt.

Florakis, C., C. Louca, R. Michaely, and M. Weber (2020). Cybersecurity risk. NBER Working Paper 28196.

Frankel, D., S. Morris, and A. Pauzner (2003). Equilibrium selection in global games with strategic complementarities. Journal of Economic Theory 108(1), 1–44.

Gai, P., A. Haldane, and S. Kapadia (2011). Complexity, concentration and contagion. Journal of Monetary Economics 58(5), 453–470.

Gartner (2020). Forecast: Enterprise IT Spending for the Banking and Securities Market, Worldwide, 2018-2024, 2Q20 Update. Technical report, Gartner.

Glasserman, P. and H. P. Young (2016). Contagion in financial networks. Journal of Economic Literature 54(3), 779–831.

Goh, J., H. Kang, Z. X. Koh, J. W. Lim, C. W. Ng, G. Sher, and C. Yao (2020). Cyber risk surveillance: A case study of Singapore. MAS Staff Paper No. 57.

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

Gordon, L. and M. Loeb (2002). The economics of information security investment. ACM Transactions on Information and System Security 5(4), 438–457.

Goyal, S. and A. Vigner (2014). Attack, defence, and contagion in networks. The Review of Economic Studies 81, 1518–1542.

Grossklags, J., N. Christin, and J. Chuang (2008). Secure or insure? A game-theoretic analysis of information security games. In WWW '08: Proceedings of the 17th international conference on World Wide Web, pp. 209–218.

HackerOne (2022). 6th annual hacker powered security report. https://www.hackerone.com/resources/reporting/6th-annual-hacker-powered-security-report.

Hirshleifer, J. (1983). From weakest-link to best-shot: The voluntary provision of public goods. Public Choice 41(3), 371–386.

Jamilov, R., H. Rey, and A. Tahoun (2021). The anatomy of cyber risk. NBER Working Paper No. 28906.

Johnson, C., L. Badger, D. Waltermire, J. Snyder, and C. Skorupka (2016). Guide to cyber threat information sharing. Special publication 800-150, National Institute of Standards and Technology.

Johnson, J. (2002). Open source software: Private provision of a public good. Journal of Economics & Management Strategy 11, 637–662.

Kamiya, S., J.-K. Kang, J. Kim, A. Milidonis, and R. M. Stulz (2021). Risk management, firm reputation, and the impact of successful cyberattacks on target firms. Journal of Financial Economics 139(3), 719–749.

Kashyap, A. K. and A. Wetherilt (2019). Some principles for regulating cyber risk. AEA Papers and Proceedings 109, 482–87.

Kocher, P., J. Horn, A. Fogh, D. Genkin, D. Gruss, W. Haas, M. Hamburg, M. Lipp, S. Mangard, T. Prescher, et al. (2019). Spectre attacks: Exploiting speculative execution. In 2019 IEEE Symposium on Security and Privacy (SP), pp. 1–19. IEEE.

Lipp, M., M. Schwarz, D. Gruss, T. Prescher, W. Haas, A. Fogh, J. Horn, S. Mangard, P. Kocher, D. Genkin, Y. Yarom, and M. Hamburg (2018). Meltdown: Reading kernel memory from user space. In 27th USENIX Security Symposium (USENIX Security 18), Baltimore, MD, pp. 973–990. USENIX Association. August.

Mauer, T. and A. Nelson (2020). International strategy to better protect the financial system against cyber threats. Technical report, Carnegie Endowment for International Peace.

Mester, L. J. (2019). Cybersecurity and financial stability. Speech at the Federal Reserve Bank of Cleveland, Cleveland, Ohio. 21 November.

Morris, S. and H. Shin (2003). Global games: Theory and applications. In M. Dewatripont, L. Hansen, and S. Turnovsky (Eds.), Advances in Economics and Econometrics (Proceedings of the 8th World Congress of the Econometric Society). Cambridge University Press.

Morris, S. and H. S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. American Economic Review 88(3), 587–597.

Panetta, F. (2020). Keeping cyber risk at bay: our individual and joint responsibility. Introductory remarks at the fifth meeting of the Euro Cyber Resilience Board for pan-European Financial Infrastructures. Frankfurt, 16 December.

Perlroth, N. (2021). This is how they tell me the world ends: The cyberweapons arms race. Bloomsbury Publishing.

Pretty, D. (2018). Reputation risk in the cyber age: The impact on shareholder value. Technical report, Aon and Pentland Analytics.

Rochet, J.-C. and X. Vives (2004). Coordination failures and the lender of last resort: was Bagehot right after all? Journal of the European Economic Association 2(6), 1116–47.

Rumsfeld, D. (2002). Defense Department Briefing – Secretary Donald Rumsfeld and General Richard Myers. https://www.c-span.org/video/?168646-1/defense-department-briefing.

Samuelson, P. (1954). The pure theory of public expenditure. The Review of Economics and Statistics 36(4), 387–389.

Shavell, S. (2009). Economic analysis of accident law. Harvard University Press.

S&P Global Market Intelligence (2019). S&P downgrades Malta-based Bank of Valletta. https://www.spglobal.com/marketintelligence/en/news-insights/trending/5mvfiykwlxliliri78qd-q2.

Tarabay, J. (2021). How a dated cyber-attack brought a stock exchange to its knees. Bloomberg Businessweek.

Van Zandt, T. and X. Vives (2007). Monotone equilibria in Bayesian games of strategic complementarities. Journal of Economic Theory 134(1), 339–360.

Varian, H. (2004). System reliability and free riding. In L. J. Camp and S. Lewis (Eds.), Economics of information security, pp. 1–15. Springer.

Woods, D. W., T. Moore, and A. C. Simpson (2021). The county fair cyber loss distribution: Drawing inferences from insurance prices. Digital Threats: Research and Practice 2(2), 1–21.