Kosse, Anneke; Lu, Zhentong

Working Paper

Transmission of cyber risk through the Canadian wholesale payments system

Bank of Canada Staff Working Paper, No. 2022-23

Provided in Cooperation with:
Bank of Canada, Ottawa

Suggested Citation: Kosse, Anneke; Lu, Zhentong (2022) : Transmission of cyber risk through the Canadian wholesale payments system, Bank of Canada Staff Working Paper, No. 2022-23, Bank of Canada, Ottawa, https://doi.org/10.34989/swp-2022-23

This Version is available at:
https://hdl.handle.net/10419/265217

BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel—2022-23

Last updated: May 16, 2022

Transmission of Cyber Risk Through the Canadian Wholesale Payments System

by Anneke Kosse[1] and Zhentong Lu[2]

[1]Banking and Payments Department
[2]Financial Stability Department
Bank of Canada
ZLu@bankofcanada.ca

DOI: https://doi.org/10.34989/swp-2022-23 | **ISSN** 1701-9397

Acknowledgements

The authors would like to thank Nellie Zhang for developing the simulation modules used for this study, Jacob Sharples for excellent research assistance, Ajit Desai and Narayan Bulusu for their detailed comments, as well as participants of the Bank of Canada and Payments Canada Quarterly Research Workshop and the RWPS (Resilience of Wholesale Payments Systems) Industry Workshop for their useful suggestions and comments. This work is part of the RWPS initiative led by the Bank of Canada. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the Bank of Canada, the Committee on Payments and Market Infrastructures or the Bank for International Settlements. As at the date of publication, Anneke works at the Bank for International Settlements, but the majority of her contributions to this paper were made during her employment at the Bank of Canada.

Abstract

In this paper, we study how the impact of a cyber-attack that paralyzes one or multiple banks' ability to send payments would transmit to other banks through the Canadian wholesale payments system. Based on historical payment data, we simulate a wide range of scenarios and evaluate the total payment disruption in the system. We find that depending on the type and number of banks under attack, the time of the attack and the design of the payments system, the attack can quickly become systemic and result in a significant loss of liquidity in the system. For instance, a three-hour attack on one bank can in the worst case impair the payments capacity of seven other banks within less than an hour and eventually disrupt 25% of the daily payments value. We also demonstrate that the system-wide impact of an attack can be significantly reduced by contingency plans that enable attacked banks to still send high-value payments. Given the interconnectedness of banks, we conclude that the cyber-resilience of a wholesale payment system strongly depends on the cyber-resilience of its participants and underline the importance of strong sectoral collaboration and coordination.

Topics: Payment clearing and settlement systems; Financial institutions; Financial stability
JEL codes: C49, E4, E42, E47, G2, G21

1 Introduction

Wholesale payment systems process transactions between financial institutions. These payments are typically large in value and often need to be settled by a particular time. Therefore, the safe and efficient functioning of these payment systems is essential to maintaining and fostering financial stability and economic growth. A wholesale payment system provides a link between its participating financial institutions. As such, the system could become a channel through which shocks to one financial institution are transmitted across the broader domestic or even international financial markets. The payment system itself, when not properly managed, can also become the source of a shock and, through its linkages, have a broader system-wide impact.[1] In recent years, cyber incidents have become a prominent source of such shocks, causing serious operational disruptions to individual financial institutions, as well as to payment systems and other financial market infrastructure. In fact, due to the relatively low risk of prosecution and the widespread availability of easy-to-use attack tools, cyber attacks have been rising globally (Adelmann et al. (2020)). According to Forbes, the financial sector was the most targeted industry in 2019. Doffman (2019) and Khiaonarong et al. (2021) show how cyber incidents targeting financial institutions have been rising over the past decade.[2] Aldasoro et al. (2020b) found that the financial sector incurs a larger number of cyber attacks than other sectors. When looking at Canada, the latest Financial System Survey of the Bank of Canada (Bank of Canada (2020)) highlights cyber-incidents as one of the top two risks to both individual firms and the Canadian financial system as a whole.

In response to growing concerns pertaining to cyber risks and other operational risks,

[1] A recent outage at the U.S. Federal Reserve caused a disruption to its key payment systems of more than three hours, including the automated clearing house system known as FedACH, the Fedwire Funds Service, as well as several other systems comprising the U.S. payments infrastructure. See: `https://www.reuters.com/article/idUSKBN2AO2I1`.

[2] More generally, the recent cyber-attack on the Colonial Pipeline (an American oil pipeline system that originates in Houston, Texas, and carries gasoline and jet fuel mainly to the Southeastern United States) raised general concerns about the vulnerability of critical infrastructures. See: `https://en.wikipedia.org/wiki/Colonial_Pipeline_ransomware_attack`.

such as those related to climate change, international organizations, central banks and private sector entities have taken collaborative actions to increase the operational and data resilience of financial institutions and financial market infrastructure, including payment systems.[3] In Canada, the Bank of Canada has established and leads the Canadian Financial Sector Resiliency Group (CFRG) and the Resilience of Wholesale Payments Systems (RWPS) initiative, which offer a forum for coordinating a national sectoral response to systemic operational incidents, such as cyber-attacks (see Dinis and Bal (2021) for details).

Cyber-attacks can be broadly categorized into confidentiality, availability and integrity attacks (see Eisenbach et al. (2020)). Confidentiality attacks, such as data breaches, compromise the confidentiality of data or systems, whereas availability attacks disrupt their availability, e.g., by shutting down computer systems. Integrity attacks damage the integrity of data or systems, e.g., by impairing account balances. Cyber-attacks on banks that participate in a wholesale payment system may have a significant impact on the payments processed in that system. In particular, an availability or integrity attack on a bank's information system might disrupt its capability to submit payments into the system. As a result, at some point, other banks that rely on these payments as a source of liquidity will a face a liquidity shortage and, in turn, no longer be able to make payments either. Due to these spillovers, an attack on one bank may eventually have a system-wide effect. Moreover, integrity attacks on banks can cause reputational damage to these specific banks or even affect the public's trust in the financial system more broadly.

Availability and integrity attacks differ from other operational outages, such as those due to technical failures, in various ways (see also Eisenbach et al. (2020)):

- Malicious intent: whereas technical outages are expected to happen randomly, cyber-attacks are meant to cause maximum damage and are therefore more likely to happen at times when the impact is greatest. The magnitude of the losses is therefore likely

[3]For example, in 2016, the Committee on Payments and Market Infrastructures (CPMI) and the Board of the International Organization of Securities Commissions (IOSCO) issued guidance for FMIs to enhance their cyber resilience (`https://www.bis.org/cpmi/publ/d146.pdf`), and in March 2021, the Basel Committee issued principles for banks for operational resilience (`https://www.bis.org/bcbs/publ/d516.htm`).

to be higher.

- Complexity: due to technological advancement, cyber-attacks are getting increasingly sophisticated, and thanks to their digital nature, they can easily spread through different systems. Consequently, they can be more difficult to detect, prevent and solve than traditional outages.

- Uncertainty: because of their malicious intent and complexity, cyber-attacks can remain hidden for a considerable time before being detected, and there might be more uncertainty about the causes and potential fixes, and hence about the duration and scale of the impact.

Given this background, the aim of this paper is to assess how a cyber-attack on one or several banks in Canada can be transmitted through the Canadian financial system via the Canadian wholesale payment system. Up until September 2021, wholesale payments in Canada were processed through the Large Value Transfer System (LVTS).[4] Therefore, we use 2019 LVTS transaction data, as well as the in-house built Canadian Payment System Simulator (CPSS), to study how a disruption of an LVTS participant's ability to send payments would affect the overall payment activity in the LVTS. We simulate various scenarios to see how the impact varies with the type and number of attacked banks, the day and time of the attack, and the duration of the resulting disruption. We also examine the scenario in which attacked banks are only partially impaired and can still process a certain share of their payment obligations.

To assess the impact of an attack, we develop an analytical framework to decompose the total impact into initial, reciprocal and spillover impacts. We measure the initial impact as the number/value of payments that the bank under attack would have submitted into the system if it had not been attacked. The reciprocal and spillover impacts are measured as the number/value of payments that non-attacked banks failed to make as a result of not

[4]The LVTS got replaced by Lynx in September 2021. See: `https://www.bankofcanada.ca/2021/09/bank-canada-designates-lynx-systemically-important-payment-system/`.

receiving any funds from the attacked bank, either back to the attacked bank (reciprocal) or to other non-attacked banks (spillovers). The decomposition allows us to understand the liquidity dependence structure and identify vulnerable points in the system. The framework can be used more generally to analyze other types of shocks to financial networks.

Our key conclusions are: 1) An availability attack on one bank can quickly become systemic and result in a significant loss of liquidity in the system. For instance, a three-hour attack on one bank can, in the worst case, impair the payment capacity of seven other banks within less than an hour and eventually disrupt 25% of the daily payment value. 2) A similar system-wide impact can be generated if all smaller banks are attacked simultaneously. 3) Payment system design is an important factor in the transmission of shocks, e.g., bilateral and multilateral risk controls can lead to different transmission patterns. 4) The transmission of a cyber attack can be significantly reduced when a contingency plan is implemented that allows the attacked bank to still submit high-value payments. For example, the number of banks that would face liquidity shortages due to a 3-hour attack on another bank would be reduced by 50% if the bank under attack would still be able to submit payments (e.g., manually) higher than $100M.

Given the demonstrated interconnectedness of banks, we conclude that the cyber-resilience of a wholesale payment system strongly depends on the cyber-resilience of its participants, and we underline the importance of strong sectoral collaboration and coordination in the prevention, as well as detection and resolution, of cyber attacks.

The rest of our paper is structured as follows. Section 2 reviews the relevant literature. Section 3 provides an overview of the Canadian wholesale payment system and explains the transmission mechanism of a cyber attack. Section 4 describes the methodology and data, as well as various simulation scenarios, used in this paper. Section 5 presents the main results, while Sections 6 and 7 conclude and discuss some policy implications.

2 Literature review

This paper is inspired by Eisenbach et al. (2020), who study how a cyber attack on one bank can be transmitted through the U.S. Fedwire Funds Service. Although the underlying objectives are the same, we tailored our simulation approach to the specificities of the Canadian wholesale payment system. Moreover, our paper is unique in the way we distinguish between initial, reciprocal and spillover effects when assessing different attack scenarios.

Our work fits into the growing literature on cyber risks in financial systems. Duffie and Younger (2019) analyze the resilience of large U.S. banks when it comes to severe cyber attacks and point out that, although they seem to be resilient because of the sufficiency of their liquidity buffers, their inter-bank payments may be disrupted because of strategic uncertainty and coordination failure. A series of IMF working papers, including Kopp et al. (2017), Adelmann et al. (2020), Bouveret (2018) and Goh et al. (2020), provide a broad overview of cyber risks in financial systems, present a framework of assessing and monitoring these, and discuss their potential implications for financial stability. Other closely related discussions of cyber risks in the financial sector can be found in Healey et al. (2018), Kaffenberger and Kopp (2019), Collins et al. (2020) and Aldasoro et al. (2020a). The latter, for instance, employ a cross-country data set containing over 700,000 operational loss events from 2002 until end-2017 at 74 large banks and conclude that cyber losses still constitute a small portion of operational losses but can account for a significant share of total operational value-at-risk.

Our paper is also related to the more traditional literature on operational disruptions in payment systems. After all, a cyber incident can be regarded as a special kind of shock causing operational disruption. Examples of work in this area include McAndrews and Potter (2002), who examine the impact of the 9-11 attack on payment behavior in the Fedwire Funds Service, and Bedford et al. (2005), who propose a simulation approach to measure the consequences of different types of operational incidents and apply it to CHAPS, the UK's main wholesale payment system. Similarly, Clarke and Hancock (2012) simulate operational

disruptions at financial institutions in Australia's wholesale payment system and highlight the implications of alternative liquidity saving features for the resilience of the system. More recently, Heijmans and Wendt (2020) developed a risk indicator to evaluate the criticality of participants in a large value payment system network and applied it to the European wholesale payment system TARGET2.

This is not the first paper that examines shocks in the Canadian wholesale payment system. McVanel (2005) empirically tests LVTS participants' robustness to unanticipated defaults and finds that all LVTS participants were able to withstand their loss allocations rather easily. A few years later, Zhang and Hossfeld (2010) studied the impact of an LVTS participant's default on the survivors' additional settlement obligations and their end-of-day collateral positions. A key difference between these two papers and ours is that the former assume the complete default of a bank in the sense that it is no longer able to make nor receive any payments. In our analysis, however, we assume that the banks under attack are still participating in LVTS and still able to receive payments. Another related study is that of Arjani and Heijmans (2020), who develop an algorithm to detect operational outages at LVTS participants that could be used by the LVTS operator to monitor operational risks in the system.

Finally, our paper is related to the theoretical literature on systemic risks posed by attacks on or disruptions of inter-bank networks. To name a few, Bech and Garratt (2012) provide a theoretical framework to analyze the effects of wide-scale disruptions in payment systems. This framework highlights that there are two forces at play in such disruptions, namely, operational problems and changes in participants' behavior. Freixas et al. (2000) develop a model of the inter-bank market and investigate the ability of the banking system to withstand the insolvency of a bank, i.e., whether the closure of a bank generates a chain reaction in the rest of the system. Goyal and Vigier (2014) study the optimal way to design and defend networks (e.g., in terms of network structure design and defence resource allocation) in the face of attacks, under different conditions.

3 The Canadian wholesale payment system

3.1 Overview

Wholesale payments are generally defined as payments between banks and other financial institutions. Contrary to payments originating from consumers or businesses, wholesale payments are mostly large in value and time-critical. In Canada, between 1999 and September 2021, wholesale payments were processed through the Large Value Transfer System (LVTS).[5] Owned and operated by Payments Canada, the LVTS processed payments in real time with the certainty that the system would settle at the end of the day. In dollar terms, the LVTS settled most of the payment value transacted in Canada every day. In 2020, it processed transactions equivalent to Canada's annual GDP every six days, or about $398B across 41,000 payments per business day.[6]

3.2 Transmission mechanism of a cyber attack

If a bank participates in a wholesale payment system, its liquidity position and payment capacity might be impacted when another bank is experiencing a cyber attack, as the impact would spread through the payment system. Putting aside the strategic responses of unattacked banks, e.g., hoarding cash or delaying payments, the design of the payment system's risk controls plays a key role in the transmission of such a shock. In order to better understand what such a transmission would have looked like for the Canadian financial system when the LVTS was still in place, it is essential to understand the LVTS risk control measures.

[5]The LVTS got replaced by Lynx in September 2021. See: `https://www.bankofcanada.ca/2021/09/bank-canada-designates-lynx-systemically-important-payment-system/`.

[6]Based on our calculations using LVTS transaction data provided by Payments Canada.

The LVTS had two tranches that participants could choose from when submitting a payment: Tranche 1 (T1) and Tranche 2 (T2). A participant could send a T1 payment as long as its net owing position as a result of all its T1 payments sent and received did not exceed the collateral it had pledged to the Bank of Canada for T1 payments, i.e., its T1 net debit cap (T1NDC). Under T2, each participant began the day by granting a bilateral credit limit (BCL) to every other participant, which was the largest net exposure it was willing to accept vis-à-vis the other participants on that day. In addition, each participant had a T2 multilateral net debit cap (T2NDC), which was calculated as the sum of all BCLs extended to it, multiplied by a system-wide percentage. During that day, participants could send T2 payments if their net owing bilateral positions did not exceed the BCLs that they had been granted and their total multilateral net position did not exceed their T2 multilateral net debit cap.[7]

A payment submitted into the LVTS was rejected if it did not pass the above-mentioned risk controls within the tranche it was sent. In particular:

- a T1 payment was rejected if it generated a net owing position that was larger than the payer's T1NDC.

- a T2 payment was rejected if it generated a bilateral owing position vis-à-vis the payee that was larger than the BCL that the payee had granted to the payer, or if it resulted in an overall net owing position that exceeded the payer's T2NDC.

T1 and T2 payments larger than CAD 100M, i.e., jumbo payments, that did not pass the risk controls were placed into a queue. As soon as the required liquidity arrived (through incoming payments, through pledging additional collateral (if a T1 payment) or an increase of the BCL (if a T2 payment)), queued jumbo payments were released and processed as so-called delayed payments. If, however, the queuing time exceeded the maximum queuing time of 35 minutes, they would ultimately be rejected too.

[7]For more details about the LVTS, see Arjani and McVanel (2006).

The T1 and T2 risk controls would have been crucial in the transmission of a potential cyber-attack on one of the LVTS participants. If a participant was no longer able to submit T1 payments into the LVTS, the other banks that relied on these payments would start to face liquidity shortages and at some point hit their T1NDC. From that moment onward, all their T1 payments would be delayed and/or rejected and their payees would soon hit their T1NDC too. Similarly, if an LVTS participant was no longer capable of submitting T2 payments into the system, a domino-effect would be created where other banks that intended to receive these funds would start to hit their T2 BCL vis-à-vis the attacked bank and/or their multilateral T2NDC, at which point their payments would be queued or entirely come to a halt too.

Figure 1 illustrates the aforementioned transmission mechanism in the event of a simulated attack on Bank A. In this simulation, Bank A's ability to send T2 payments is simulated to be disrupted between 15:00 and 18:00 on a random day in 2019. The left panel shows the actual (blue line) and simulated (purple line) net liquidity positions of Bank A in T2. The growing divergence of the two lines demonstrates that Bank A became a liquidity sinkhole soon after the simulated attack. After three hours, the bank had absorbed about CAD 6B from the system as it was still receiving payments while not being able to make payments itself. The middle and right panels show how the simulated disruption at Bank A affects Bank B's multilateral and bilateral liquidity positions in T2. Soon after the attack on Bank A, Bank B's net positions decrease due to lost incoming liquidity from Bank A. In fact, the panel in the middle suggests that Bank B hits its T2 liquidity limit (i.e., its T2NDC) within two hours after the disruption after which it can no longer make any T2 payments either.

From a theoretical perspective, the aforementioned transmission mechanism is a special case of "coordination failure," as discussed in the bank run literature (see e.g., Diamond and Dybvig (1983)). In particular, Morris and Shin (2004) discuss how privately set loss limits by traders can mutually reinforce selling behavior and create liquidity shortages in the market. Though in a different context, their key insight resembles the transmission mechanism for

the LVTS: risk control measures designed to control risk exposures can actually accelerate the propagation of risks through the payment network in the event of a severe shock.

4 Methodology

4.1 Simulation approach and data

We use the Canadian Payment System Simulator (CPSS) to simulate the potential financial system-wide impact of a cyber attack on one or more banks. The CPSS fully replicates the functionalities and above-mentioned risk controls of the LVTS to generate the settlement and rejection times of payments that were submitted into the LVTS. Without changing any parameters, the CPSS generates the same settlement and rejection times as in the actual historical LVTS data. See Zhang and Hossfeld (2010) for more information about the simulation tool.

For the purpose of this study, we employ actual 2019 LVTS transaction data and use a tailored module to modify this data by removing all the payments initiated by the banks that are assumed to be under attack during a particular time window. We create multiple adjusted data-sets for different attack scenarios, including different banks, attack times and attack lengths. Subsequently, we run these modified data-sets through the simulator to generate the settled and rejected payments of the other (non-attacked) banks under the various scenarios. We then compare the simulated outcome with that from the non-attack benchmark simulation to measure the impact of the attack.

We do not include in the simulations the pre-settlement of the LVTS that occurs between 18:00 and 18:30 due to the difficulty of making assumptions about banks' behavior in response to their adjusted end-of-day positions. Moreover, we exclude the Bank of Canada when analyzing the results and only focus on the impact on the payments made and received by other financial institutions in the system.

When analysing the simulation results, we measure the payment disruption in the system

by decomposing the overall impact into three categories:

- Initial impact: the number/value of payments that the bank under attack would otherwise have submitted into the system if it had not been attacked.
- Reciprocal impact: the number/value of payments that non-attacked banks would otherwise have sent to the bank under attack if the latter had not been attacked.
- Spillover impact: the number/value of payments that non-attacked banks would otherwise have sent to each other if the attack had not taken place.

4.2 Description of scenarios

4.2.1 Baseline scenario

In the baseline scenario, we assume that one bank is attacked in such a way that it can no longer submit any payments into the LVTS. Moreover, we assume "passive responses" from other banks, i.e., other banks do not actively respond to the attack by adjusting the number, size and/or timing of their payments. This assumption is reasonable since a cyber-attack of one bank is typically private information and hard to detect by others, especially in the first few hours. Also, this assumption allows us to zero in on the transmission mechanism based on the liquidity dependency among participants, by shutting down the channel of behavioral responses that are not easy to model.[8] Furthermore, in this baseline scenario, we assume that the LVTS is still properly functioning, as a result of which the bank under attack is still able to receive payments from others and hence becomes a "liquidity sinkhole." We also assume that the non-attacked banks do not re-submit any of their rejected payments, that queued payments cannot stay in the queue longer than the current maximum time of 35 minutes and that banks do not raise additional liquidity from other sources (e.g., by increasing their LVTS collateral or changing T2 BCLs).

[8]Reinforcement learning techniques could be employed to model participants' behavior in unusual circumstances, see e.g., Castro et al. (2021). This is, however, beyond the scope of our paper.

We simulate this baseline scenario for the ten largest Canadian banks individually, for every hour of the day, every day of the week and for different attack lengths. Specifically, 136,950 scenarios are simulated and examined.[9]

4.2.2 Alternative scenario 1: Multiple banks under attack

In the first alternative scenario, we simulate and assess a cyber-attack on multiple banks. We conduct this exercise for various combinations of banks and again run it for different attack times, days and lengths. We do so using every potential combination of the six largest banks, generating over 780,000 simulations.[10]

The results of this scenario not only provide insight into the direct and system-wide effect of a situation where cyber-criminals impair the payment capabilities of multiple banks, they also demonstrate the potential consequences of an availability attack on a shared service provider. Financial institutions commonly rely on numerous third-party service providers that supply support services that are needed for the operation of their business functions, such as IT or payment processing services. By outsourcing these services, financial institutions are able to focus their resources on their core functions and generate efficiencies. However, this reliance comes with a risk, as cyber or operational events at the third party could have an impact on the financial institutions' abilities to continue operations. This impact could amplify systemic risks if these third parties provide the same services to multiple institutions at the same time.

4.2.3 Alternative scenario 2: Partial impairment of payment capacity

The second alternative scenario assumes that the bank(s) under attack have certain contingency measures in place that allow them to continue submitting some of their payments into

[9]For each hour in the day, we simulated all possible attack lengths until the system's closure at 6PM. Hence, for a 1-hour attack, the lengths vary from 8AM to 5PM, whereas 2-hour attacks are only simulated from 8AM to 4PM. As a result, the total number of simulations = 10 participants * 249 days * $\sum_{i=1}^{10} i$ attack lengths.

[10]The total number of simulations = 57 bank combinations * 249 days * $\sum_{i=1}^{10} i$ attack lengths.

the LVTS, e.g., by manually entering these.

When only being able to submit a portion of their payments, banks will prioritize those with the largest values.[11] Therefore, under this second alternative scenario, we compare three situations:

- Attacked bank can still submit its payments of $500M and over;
- Attacked bank can still submit its payments of $250M and over;
- Attacked bank can still submit its payments of $100M and over.

5 Results

5.1 Baseline scenario: One bank under attack

Figures 2 and 3 provide a high-level summary of the various scenarios simulated. In particular, we focus on the total value of rejected payments in the whole system due to a 3-hour attack on one bank.[12] It is clear that the transmission and impact largely vary by bank, the attack time as well as the day of the attack. Also, the size and composition of the impact strongly depends on the type of system used (T1 or T2). For example, the impact in T1 is greater in the late afternoon than other times of the day, while the impact in T2 remains much the same during the whole day.

Figure 4 shows the average and maximum effects of a 3-hour cyber-attack on one particular bank for a given attack time (as displayed on the x-axis). The averages (dark-colored bars) are calculated by simulating the attack for each day in 2019 and taking the average across these days. The light-colored bars represent the maximum impact across these days. On average, the total value of system-wide foregone payments caused by a 3-hour attack

[11] Based on internal discussions with various LVTS participants.

[12] The 3-hour duration seems a reasonable response time based on our discussions with employees of the participants. However, we acknowledge that there are many uncertainties in reality and that the 3-hour window may be arbitrary. So we use this particular time window mainly to illustrate our general findings. Other lengths of attack time can be easily simulated and analyzed using the same framework.

fluctuates between 5% and 10% of the daily LVTS value processed. However, in the worst case, the value of foregone payments in the system amounts to 25% of the total LVTS.

Looking at the composition of the impact, the majority constitutes the so-called initial impact, i.e., payments that the bank under attack would otherwise have made. The reciprocal impact, i.e., forgone payments that the bank under attack would otherwise have received from other banks, is much smaller, yet significant. The so-called spillover impact, i.e., foregone payments between non-attacked banks, accounts for the smallest portion. This suggests that the system-wide impact of the attack remains limited to those banks that are direct recipients of payments from the bank under attack. Yet, the spillover effects become considerably greater towards the end of the day, especially in the extreme case. This suggests that a cyber incident becomes more "contagious" in the hours prior to the closing of the system.

Figure 5, for instance, shows how the impact of an attack increases with its duration. This is not surprising—the longer the payment capability of the bank is impaired, the greater the system-wide impact. However, more interestingly, Figure 5 also demonstrates that the transmission pattern largely differs by payment tranche used. This simulated multi-hour attack generates relatively more spillover effects (i.e., foregone payments between non-attacked banks) in T1 than in T2. In contrast, a significant portion of all foregone T2 payments concern reciprocal payments. Also, the total of impacted T2 transactions linearly increases with the duration of the attack, whereas the impact in T1 shows a more exponential pattern. This suggests that the design of the system, or better, the design of the underlying risk controls, influences the transmission of shocks through the financial system.

As discussed in Section 1, one way in which cyber-attacks differ from random operational outages is that the latter are often meant to cause maximum damage and hence are more likely to be targeted so that the impact is greatest. Therefore, it makes sense to look at the tail events, i.e., the cyber-attacks that have the greatest system-wide impact. Figure 6 presents, out of all scenarios simulated, the results of the 3-hour attack that generates the largest total impact. The attacked bank is represented by the black dot in the middle, and

the dots around it represent the other banks in the LVTS. Each of the dots are connected to the centre dot by a red line, which means that in this worst-case scenario, all the other banks would be impacted by this 3-hour attack and no longer receive any payments from the attacked bank. On top of this so-called first-round effect, the five blue lines indicate that after three hours, five non-attacked banks are no longer able to make any payments back to the attacked bank. Yet, the system-wide effect does not end here either, as the green lines demonstrate how one non-attacked bank can no longer make payments to six other non-attacked banks. In sum, a 3-hour cyber-attack on one LVTS participant can rapidly spread through the system and affect the payment capacity of others. Figure 7 demonstrates how this worst-case scenario impacts seven other banks in less than an hour and ten banks within 135 minutes. These results underline the importance of speedy detection and resolution mechanisms that minimize the duration of payment outages.

5.2 Alternative scenario 1: Multiple banks under attack

In order to assess the potential system-wide impact of a multiple-bank attack, we first re-run the simulations while assuming that not one, but multiple banks, are no longer able to make any payments. We do so by simulating all potential bank combinations, focusing on the six largest banks.

Figure 8 presents the distribution of the total impact in terms of the percentage share of the daily LVTS payment value, of an end-of-day, 3-hour attack on all potential bank combinations (out of the top six banks). The impact of a 2-bank attack is clearly larger than that of a 1-bank attack, as shown by the distribution of the impact, which shifts to the right. And as expected, the impact increases when more banks are attacked. Interestingly, the total impact does not considerably increase much further when more than three banks are attacked. This suggests that an attack on three or more banks hardly impairs any payments that are not already affected in the case of a 1- or 2-bank attack. However, this also means that simultaneous attacks on three banks have a very serious system-wide impact and cause

disruptions to most payment activities of the payment system.

Figure 9 compares the impact of a multiple-bank attack on all smaller banks (i.e., all non-big-six banks), shown in the last column, with that of an individual attack on each of the big six banks, shown in the first six columns. Overall, such an attack would generate a substantial shock to the system. The total value of system-wide foregone payments would be of a similar magnitude as that of a single attack on one of the big six banks. Moreover, it would generate a higher share of spillover effects (between non-attacked banks) than an attack on one of the big banks. This suggests that, although at an individual level small banks are only responsible for a small share of payments made in the LVTS, their total payment activity constitutes a non-negligible source of liquidity to other banks in the system.

5.3 Alternative scenario 2: Partial impairment of payment capacity

So far, the analyses have assumed that a cyber attack would fully paralyze a bank's payment ability. This would, however, overestimate the total impact if that bank had contingency measures in place such that it could still submit a certain share of its payments in case of an outage. When only being able to submit a portion of their payments, banks are likely to prioritize those with the largest values as these are often most critical. Therefore, we have re-run the single-bank attack, as summarized in Section 5.1, while assuming the attacked banks are still able to submit payments above a certain value. In particular, we considered three different thresholds: \$500M, \$250M and \$100M. The results are displayed in Figure 10.

The top left panel shows the impact of the baseline scenario in which the bank's payment function is fully impaired. The top right panel demonstrates how the total impact on the system decreases if the bank under attack can still submit payments of \$500M and higher: the magnitudes of the initial and reciprocal impacts (the thicknesses of the red and blue lines) are smaller and there are fewer spillovers between non-attacked banks (green lines).

The bottom left and right panels show how the initial and reciprocal impacts further decline and that the spillovers eventually disappear when lowering the threshold to $250M and $100M. These results suggest how bank-level contingency measures may help contain the system-wide impact of a cyber-attack. This conclusion is also supported by Figures 11 and 12, which illustrate that the affected number of banks and share of LVTS value, respectively, decrease as the payment capacity of the attacked bank increases (i.e., the threshold value is reduced).

6 Conclusion

In this paper, we assess how a potential cyber-attack on one or multiple banks can impact the broader financial system due to banks' interconnectedness through the wholesale payment system. We do so by simulating the total system-wide number and value of disrupted payments in case the payment function of one or more banks is paralysed. We use 2019 LVTS transaction data and study various scenarios to see how the impact varies with the type and number of attacked banks, the day and time of the attack, and the duration of the outage. We also examine the scenario in which a cyber-attack only partially impairs a bank's payment capability.

Four key conclusions can be taken away from the million-plus scenarios simulated. First, we find that a cyber attack disabling the payment capability of one LVTS participant can spread through the entire wholesale payment system in less than a few hours, effecting the payment capacity of others.

Second, we show that the total system-wide impact of a cyber-attack does not considerably increment much further when more than two large banks are attacked. This suggests that a joint outage of the six biggest banks in Canada would hardly impair payments that are not already affected if only one or two large banks were attacked. This suggests that each of the largest banks plays a crucial role in the functioning of the wholesale payment

system. Hence, the resiliency of the entire wholesale payment system strongly depends on the resiliency of its participants. In fact, we show that this not only holds for the largest banks. Simultaneous disruptions of all small banks are likely to generate a shock to the system similar to that of an outage of one of the large banks. Hence, the total payment activity of small banks constitutes a non-negligible source of liquidity for others in the system.

Third, we demonstrate how the total system-wide impact of an outage decreases if banks under attack can still submit a certain share of their payments, e.g., through the existence of contingency arrangements such as manual payment processes.

Finally, we demonstrate that a disruption of T1 payments generates relatively more spillover and less reciprocal effects than a disruption of T2 payments. This suggests that the design of a wholesale payment system influences the transmission of shocks through the financial system. Our analysis shows that bilateral, as opposed to multilateral, credit limits have the potential to encapsulate system-wide contagion such that payments between non-attacked banks are less affected. Based on this, it might be worthwhile to redo the analysis in a few years when there is a sufficiently large data set available from the new wholesale payment system Lynx that replaced the LVTS in September 2021. As opposed to LVTS, Lynx is a full real-time gross settlement (RTGS) system in which the liquidity needed for making payments is generated from incoming payments and collateral pledged with the Bank of Canada. This means that banks will no longer make use of bilaterally granted credit limits.[13] Therefore, a potential payment disruption at a Lynx participant might have an impact similar to the results we found for our LVTS T1 simulations. Future research, however, is needed to better understand the system-wide effect of cyber-attacks in the Lynx era.

[13]For more information about Lynx, see Kosse et al. (2021).

7 Policy implications and future research

Wholesale payment systems are critical for the functioning of a country's economy and the stability of its financial system. Therefore, wholesale payment systems are subject to oversight from the respective central banks. This oversight is generally based on the Principles for Financial Market Infrastructures (PFMIs)[14] and requires payment systems to use appropriate cyber security tools and practices. Yet, our paper demonstrates that the cyber-resiliency of a country's financial system as a whole depends on more than the cyber-resiliency of the payment system. By contrast, we show that banks, being participants of these systems, have an important role to play as well.

Concretely, the results of our simulation exercises underline the importance of banks having speedy detection and resolution mechanisms that minimize the duration of payment outages caused by cyber-attacks. The use of advanced analytical tools such as machine learning, might be useful here — not only at the system level but also at the individual-bank level. In fact, we suggest that all banks, no matter their size, should have proper cyber security measures in place to prevent such attacks from happening and to strengthen their response and recovery functions. Moreover, given the interconnectedness of banks, strong collaboration and coordination might be required to avoid cyber-attacks becoming financial stability events. The development of contingency measures might be such an area for collaboration and coordination, as we demonstrate how the total system-wide impact of an outage decreases if banks under attack can still submit a certain share of their payments. This could be realized, e.g., through sector-wide agreements on manual payment processing or the use of contingency systems.

To the question what can central banks do, in addition to conducting oversight, to minimize the system-wide impact of a cyber-attack on the availability of a bank's payment func-

[14]The PFMIs are international standards related to the risk management, efficiency and transparency of systemically important financial market infrastructures (FMIs), published in April 2012 by the Committee on Payments and Market Infrastructures (CPMI) of the Bank for International Settlements (BIS) and the Technical Committee of the International Organization of Securities Commissions (IOSCO) (see `https://www.bis.org/cpmi/info_pfmi.htm`).

tion, various answers can be given. For instance, central banks could provide liquidity to the system's participants in the form of emergency liquidity to make up for the forgone payments and minimize the reciprocal and spillover effects from and between "healthy" banks. Our results, however, demonstrate that every second counts and that such emergency lending would have to take place as soon as possible to prevent the attack from having a systemic impact. Hence, ex-ante measures to prevent such attacks from happening in the first place, including sector-wide cooperation and coordination, are likely to be more important and effective.

Our paper provides a first step towards understanding the potential systemic implications of a cyber attack on participants of wholesale payment systems. Our results provide a benchmark for future work, which could consider richer and more realistic scenarios by relaxing some of the assumptions made in this paper. For example, it might be worthwhile to broaden the scope of the analysis by accounting for potential behavioral changes. This would call for the use of other approaches beyond the current simulations, such as game-theoretic methods, agent-based modeling and reinforcement learning. Also, future research might be done to assess the system-wide impact of multiple-day attacks that disconnect certain participants from the system for more than one day. One can argue that the closure of the system at the end of the day acts as a sort of fire-wall that prevents the damage from spreading even further. With many countries currently considering or moving towards faster payment systems that are open 24/7/365, an interesting question is how this impacts the transmission of shocks since there is no longer a natural barrier that stops a shock from disseminating further.

References

Adelmann, F., J. A. Elliott, I. Ergen, T. Gaidosch, N. Jenkinson, T. Khiaonarong, A. Morozova, N. Schwarz, and C. Wilson (2020). Cyber risk and financial stability; it's a small world after all. IMF Staff Discussion Notes 2000/007, International Monetary Fund.

Aldasoro, I., L. Gambacorta, P. Giudici, and T. Leach (2020a, February). Operational and cyber risks in the financial sector. BIS Working Papers 840, Bank for International Settlements.

Aldasoro, I., L. Gambacorta, P. Giudici, and T. Leach (2020b, May). The drivers of cyber risk. BIS Working Papers 865, Bank for International Settlements.

Arjani, N. and R. Heijmans (2020, June). Is there anybody out there? detecting operational outages from large value transfer system transaction data. *Journal of Financial Market Infrastructures 8*(4), 23–41.

Arjani, N. and D. McVanel (2006). A primer on canada's large value transfer system. Technical report, Bank of Canada.

Bank of Canada (2020). Financial system survey highlights – november 2020. Technical report.

Bech, M. L. and R. J. Garratt (2012, August). Illiquidity in the Interbank Payment System Following Wide-Scale Disruptions. *Journal of Money, Credit and Banking 44*(5), 903–929.

Bedford, P., S. Millard, and J. Yang (2005). Analysing the impact of operational incidents in large-value payment systems: A simulation approach. *Liquidity, risks and speed in payment and settlement systems–a simulation approach*, 247–74.

Bouveret, A. (2018). *Cyber risk for the financial sector: A framework for quantitative assessment*. International Monetary Fund.

Castro, P. S., A. Desai, H. Du, R. Garratt, and F. Rivadeneyra (2021, February). Estimating Policy Functions in Payments Systems Using Reinforcement Learning. Staff Working Papers 21-7, Bank of Canada.

Clarke, A. and J. Hancock (2012). Payment system design and participant operational disruptions. *Journal of Financial Market Infrastructures 2*(2), 53–76.

Collins, R., C. O'Connor-Close, and A. Zhang (2020, February). Cyber incident cost estimates and the importance of building resilience. *Reserve Bank of New Zealand Bulletin 84*, 1–17.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Dinis, F. and I. Bal (2021). Collaborating for the greater good: enhancing operational resilience within the canadian financial sector. *The Capco Institute Journal of Financial Transformation*, 08–13.

Doffman, Z. (2019). Cybercrime: 25% of all malware targets financial services, credit card fraud up 200%. Technical report, Forbes.

Duffie, D. and J. Younger (2019). Cyber runs. Technical report, Hutchins Center (Brookings).

Eisenbach, T. M., A. Kovner, and M. J. Lee (2020, January). Cyber Risk and the U.S. Financial System: A Pre-Mortem Analysis. Staff Reports 909, Federal Reserve Bank of New York.

Freixas, X., B. M. Parigi, and J.-C. Rochet (2000). Systemic risk, interbank relations, and liquidity provision by the central bank. *Journal of Money, Credit and Banking 32*(3),

611–638.

Goh, J., H. Kang, Z. X. Koh, J. Way Lim, C. Wei Ng, G. Sher, and C. Yao (2020). Cyber risk surveillance: a case study of singapore.

Goyal, S. and A. Vigier (2014, 07). Attack, Defence, and Contagion in Networks. *The Review of Economic Studies 81*(4), 1518–1542.

Healey, J., P. Mosser, K. Rosen, and A. Tache (2018). The future of financial stability and cyber risk. *The Brookings Institution Cybersecurity Project, October*.

Heijmans, R. and F. Wendt (2020). Measuring the impact of a failing participant in payment systems. *IMF Working Papers 20*.

Kaffenberger, L. and E. Kopp (2019). Cyber risk scenarios, the financial system, and systemic risk assessment. *Carnegie's Cyber Policy Initiative Working Paper Series*.

Khiaonarong, T., H. Leinonen, and R. Rizaldy (2021). Operational resilience in digital payments: Experiences and issues. IMF Working Papers 2021/288, International Monetary Fund.

Kopp, E., L. Kaffenberger, and C. Wilson (2017). Cyber risk, market failures, and financial stability. Working paper, International Monetary Fund.

Kosse, A., Z. Lu, and G. Xerri (2021). Predicting payment migration in canada. *Journal of Financial Market Infrastructures 9*(1).

McAndrews, J. J. and S. M. Potter (2002). Liquidity effects of the events of September 11, 2001. *Economic Policy Review 8*(Nov), 59–79.

McVanel, D. (2005). The Impact of Unanticipated Defaults in Canada's Large Value Transfer System. Working paper, Bank of Canada.

Morris, S. and H. S. Shin (2004). Liquidity black holes. *Review of Finance 8*(1), 1–18.

Zhang, N. and T. Hossfeld (2010). Losses from Simulated Defauts in Canada's Large Value Transfer System. Discussion Papers 2010-14, Bank of Canada.

Figures

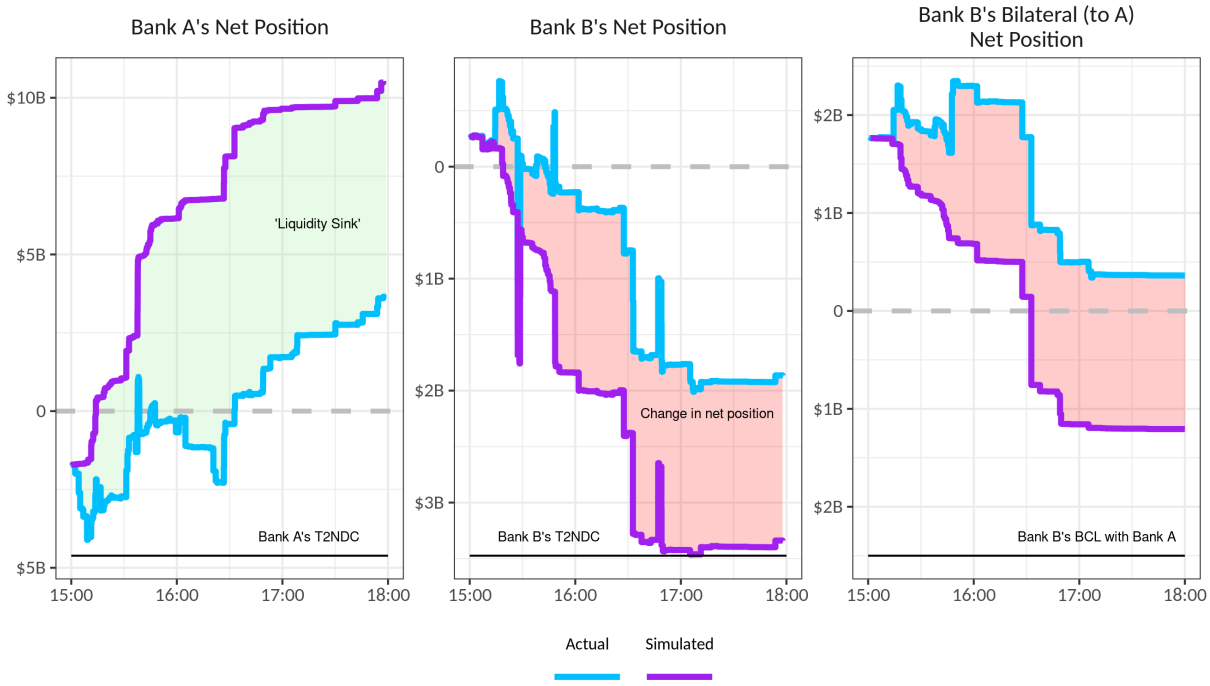

Figure 1: Illustration of Transmission Mechanism

Notes: These graphs illustrate the transmission mechanism of an availability attack on Bank A. Bank A's ability to send T2 payments is simulated to be disrupted between 15:00PM and 18:00PM on a random day in 2019. The left panel shows the actual (blue) and simulated (purple) net liquidity positions of Bank A in T2 during this window. The increasing distance between the two lines demonstrates that Bank A becomes a liquidity sinkhole soon after the simulated attack. At 18:00PM, Bank A has absorbed about CAD 6B from the system. The middle and right panels show how the simulated disruption at Bank A affects Bank B's multilateral and bilateral liquidity positions in T2. Soon after the attack, Bank B's net positions decrease due to lost incoming liquidity from Bank A. The panel in the middle suggests that Bank B hits its T2 liquidity limit (i.e., its T2NDC) within two hours after the disruption after which it can no longer make any T2 payments either.

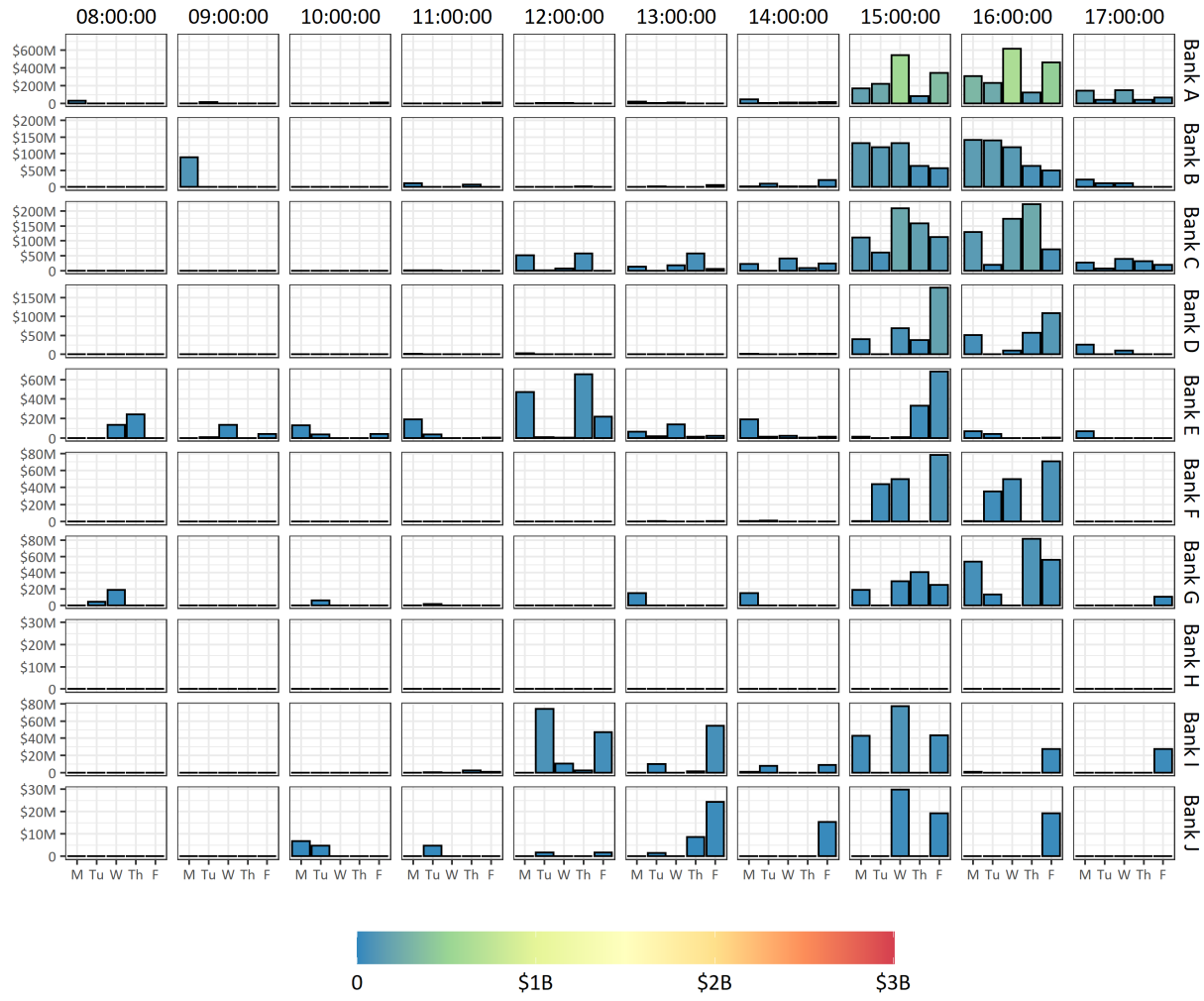

Figure 2: Rejected Payments, by Bank/Time/Day of Attack, in T1

Notes: These graphs provide an overview of the impact of a 3-hour attack, by bank, attack time and weekday, as measured by the total value of rejected T1 payments.

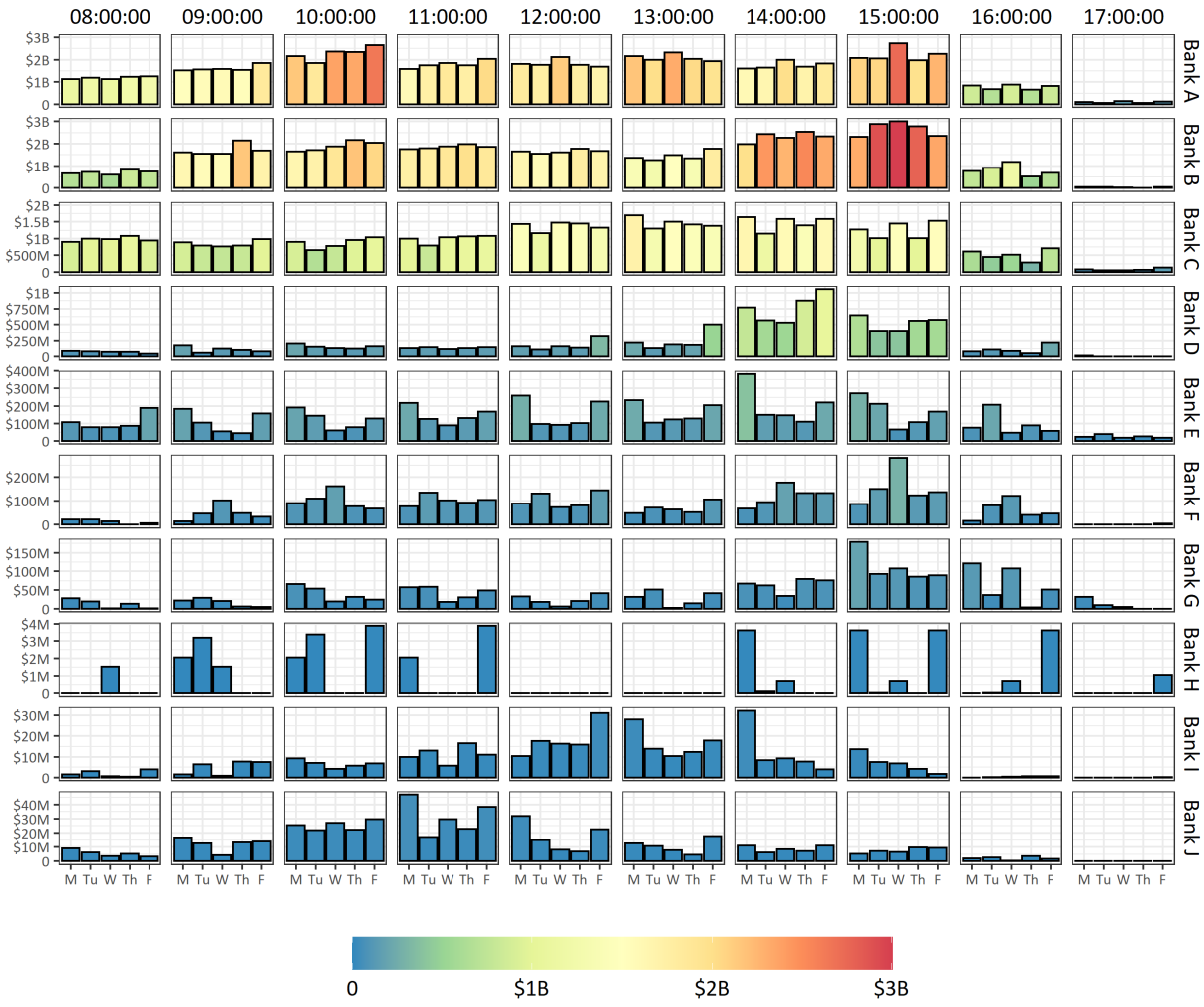

Figure 3: Rejected Payments, by Bank/Time/Day of Attack, in T2

Notes: These graphs provide an overview of the impact of a 3-hour attack, by bank, attack time and weekday, as measured by the the total value of rejected T2 payments.

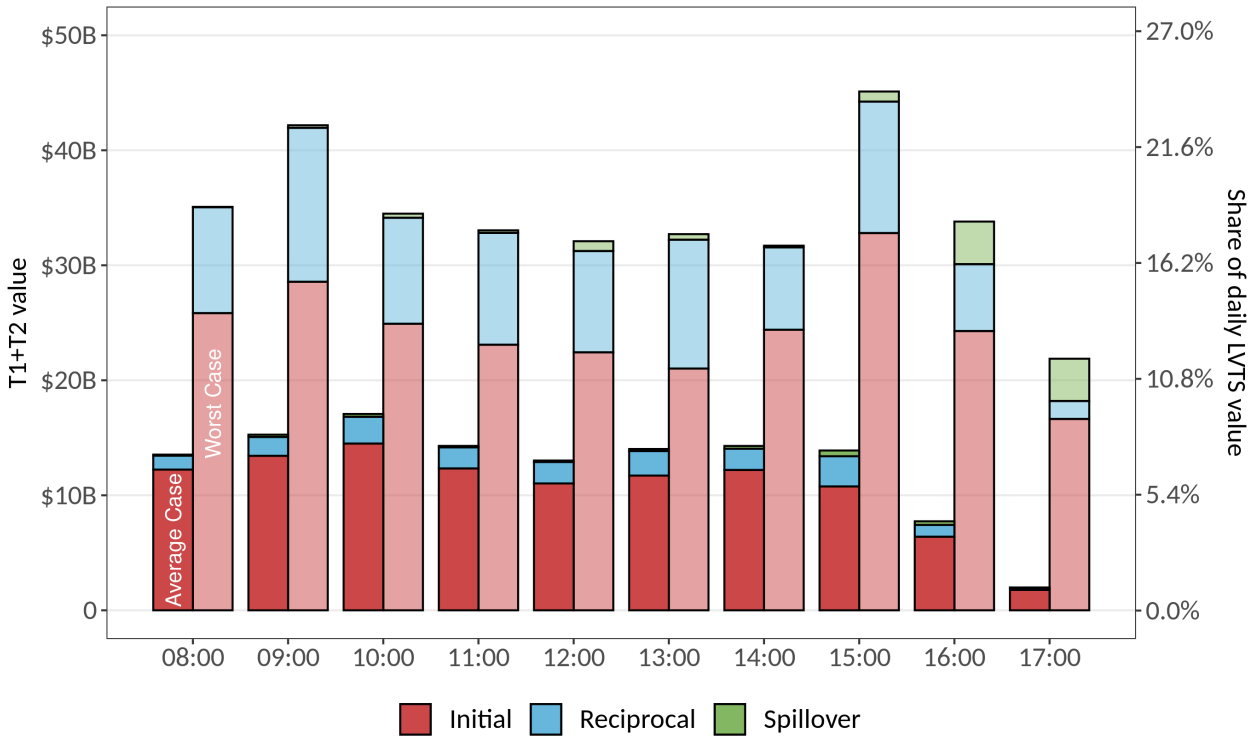

Figure 4: Impact of a Cyber Attack on Bank A

Notes: This graph shows the impact of a 3-hour attack on Bank A. The dark bars show the average impact across days, while the light bars show the maximum (i.e., worst case) impact. The three different colors (each with a dark and a light shade) represent the decomposition of the total system-wide impact into initial, reciprocal and spillover effects.

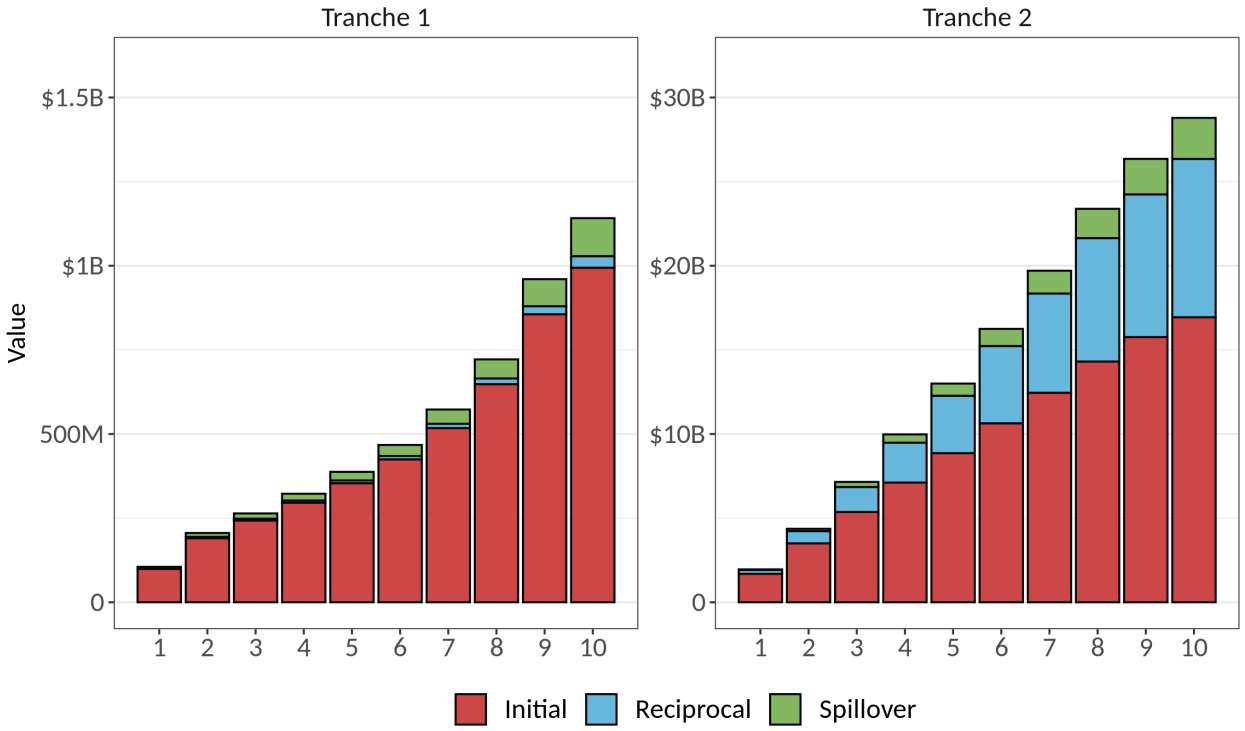

Figure 5: Impact by Attack Length and Tranche

Notes: These graphs show the impact of an attack on Bank A on the value of rejected T1 (left panel) and T2 (right panel) payments for different attack lengths (on the x-axis). The bars represent the average value across all days in 2019. The three colors reflect the decomposition of the total system-wide impact into initial, reciprocal and spillover effects.

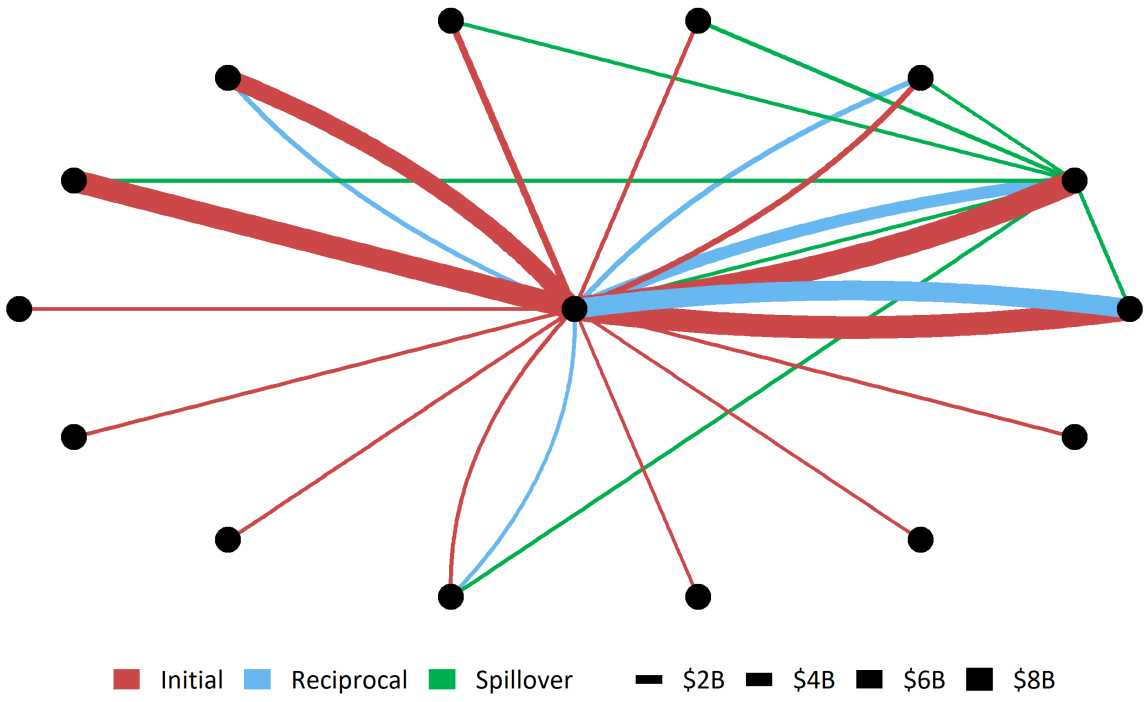

Figure 6: Costliest 3-Hour, Single-Bank Attack in 2019

Notes: This image demonstrates how the costliest 3-hour attack on the bank in the center transmits to the other participants (the peripheral nodes) in the system. The red, blue and green lines represent the initial, reciprocal and spillover effects, respectively. The thickness of the lines reflects the magnitude of the impact.

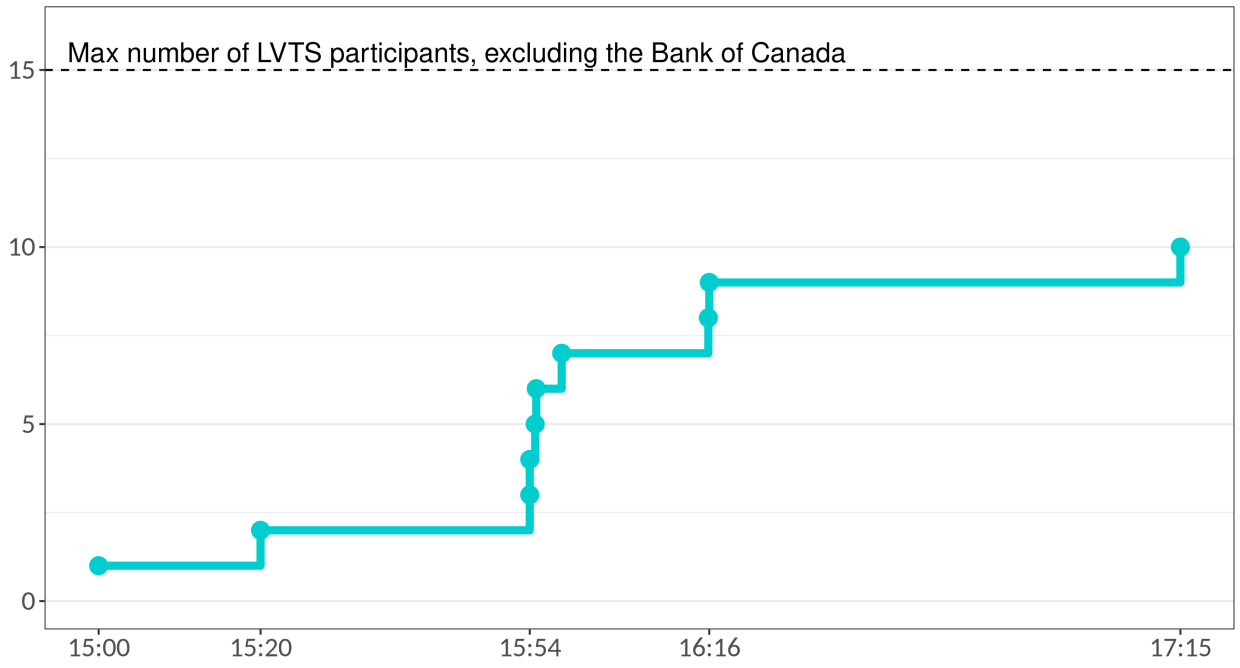

Figure 7: Number of Affected Participants in the Costliest Attack

Notes: This graph shows how the number of impacted banks increases with time in case of the costliest 3-hour, single-bank attack in 2019. Within one hour after the attack at 15:00PM, six other banks start to experience a liquidity shortage, as illustrated in Figure 1.

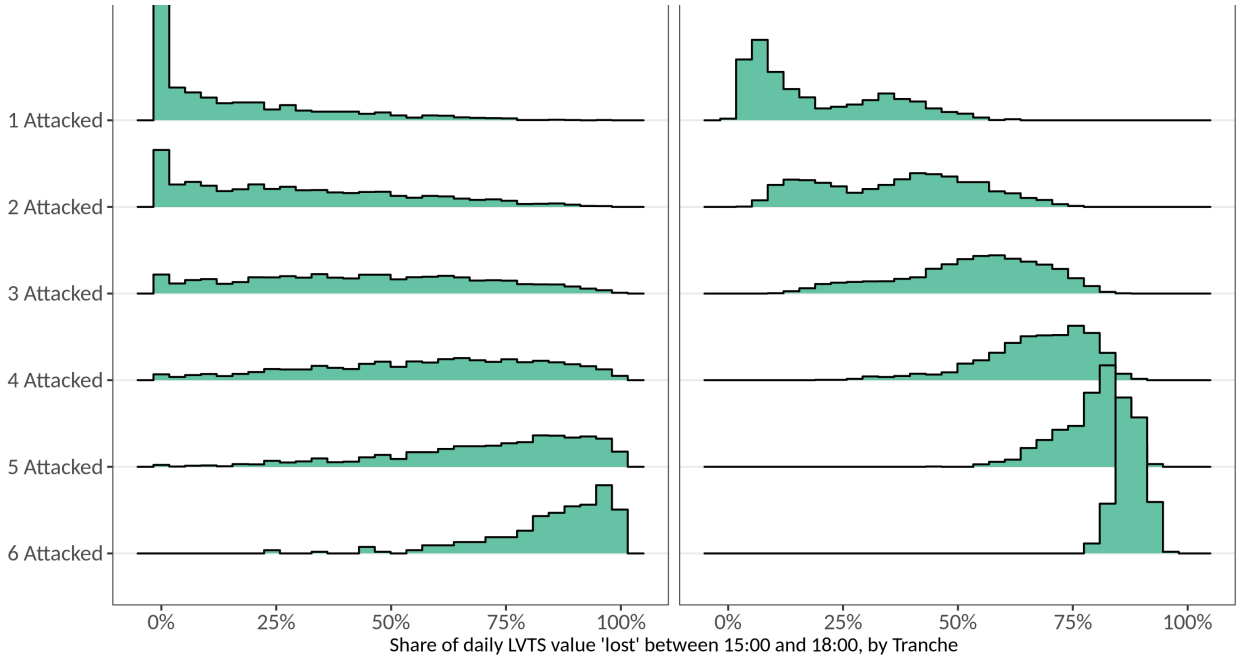

Figure 8: Incremental Impact of Simultaneous Attacks

Notes: These graphs show the distributional impact of simultaneous 3-hour attacks on n banks (out of the "Big Six") across different combinations of attacked banks and business days in 2019.

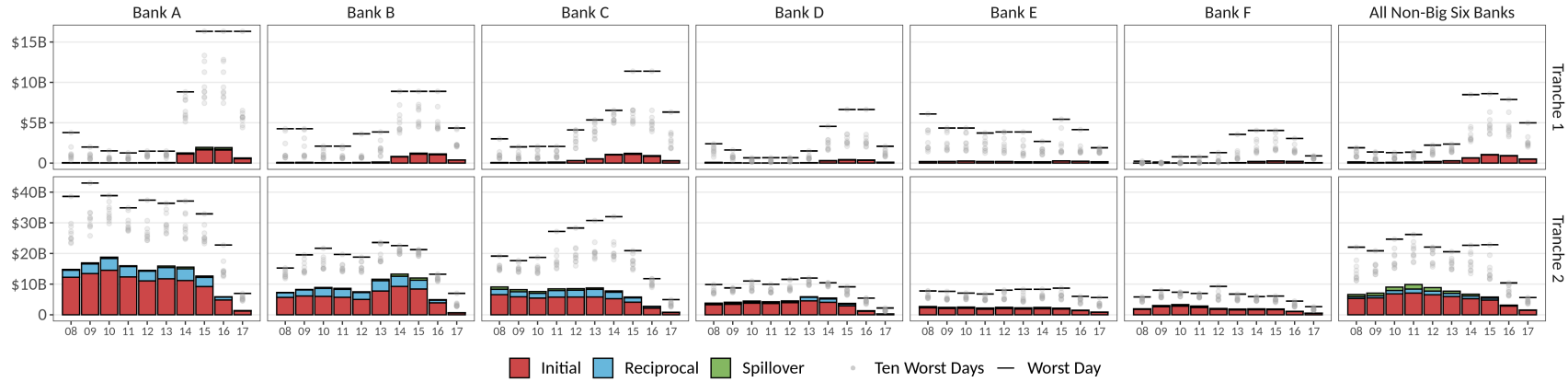

Figure 9: Simultaneous Attacks on Small Banks

Notes: These graphs compare the impact of simultaneous 3-hour attacks on all small banks (last column) with that of a 3-hour single-bank attack on each of the six large banks (columns 1–6), for T1 and T2. The bars show the average impact, and the points and lines represent the maximum effects.

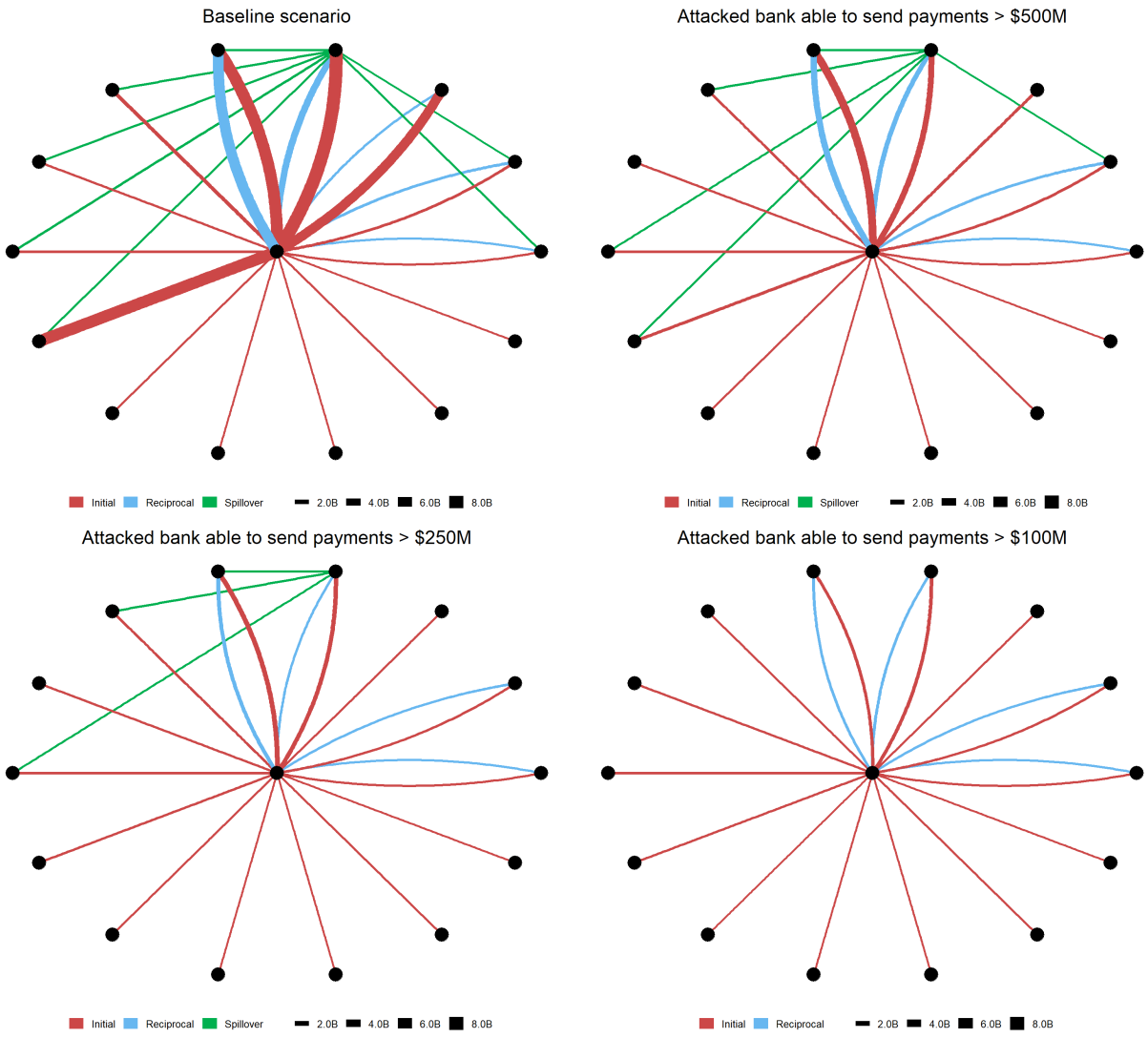

Figure 10: Single-Bank Attacks with Contingency Plans

Notes: These images show how contingency plans that would enable attacked banks to continue sending large payments can mitigate the total system-wide impact of an attack. The top-left scenario represents the baseline scenario from Figure 6. The other three scenarios assume that the bank under attack is still able to submit payments above a certain value threshold.

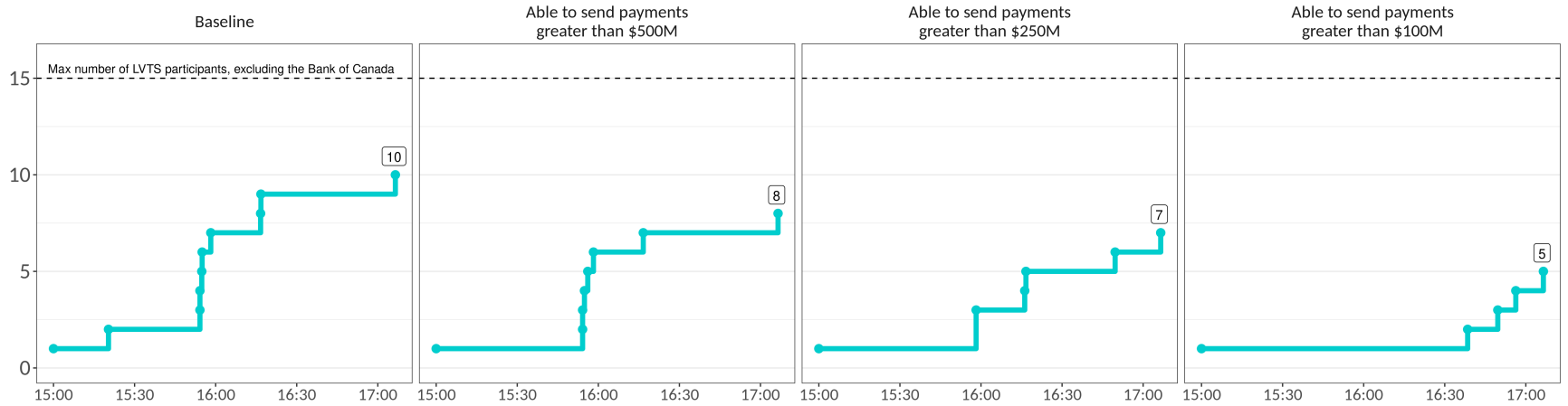

Figure 11: Number of Impacted Banks in the Case of Contingency Plans

Notes: These graphs show that contingency plans that would enable attacked banks to continue sending large payments could reduce the number of impacted banks in the system.

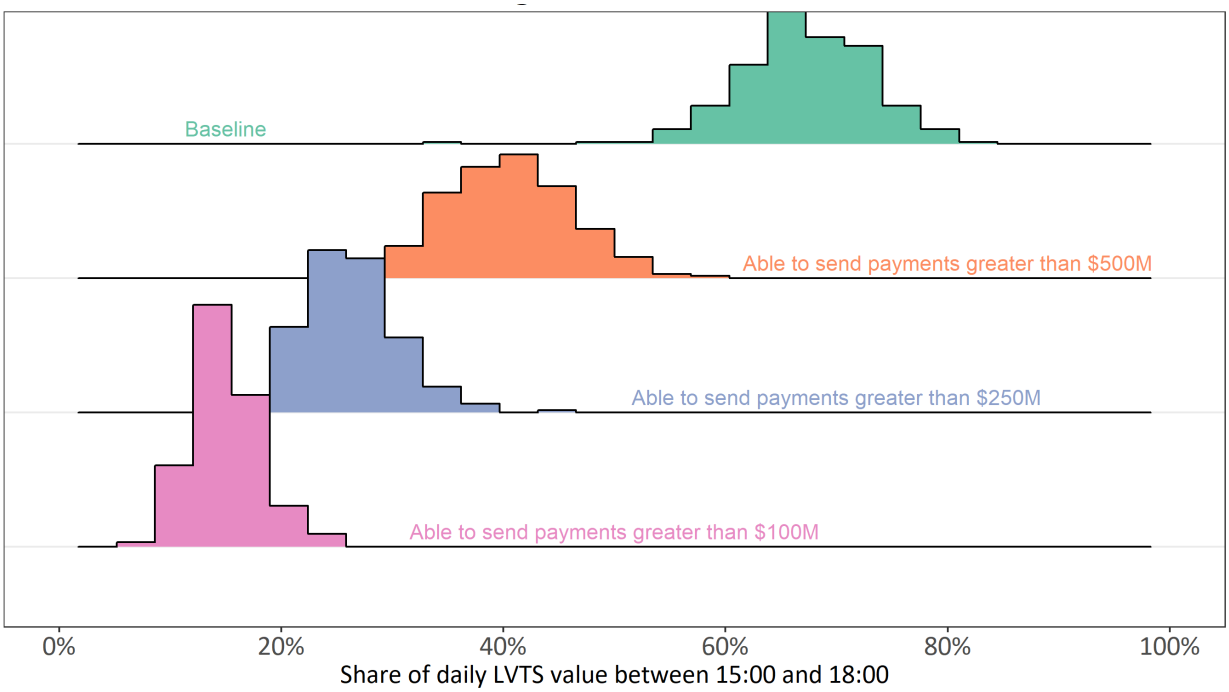

Figure 12: Simultaneous Attacks on Big Banks, in the Case of Contingency Plans

Notes: This graph shows the impact of 3-hour simultaneous attacks on the six largest banks and how the total impact could be mitigated by contingency plans that would enable the attacked banks to continue sending large payments while under attack.