# Settlement Speed and Financial Stability *

Agostino Capponi[†] Jin-Wook Chang[‡]

October 3, 2025

**Abstract**

This paper investigates how settlement speed affects financial stability in payment networks, taking into account netting benefits, liquidity costs, and counterparty risks. Our analysis reveals that faster settlements have ambiguous effects on systemic risk and social welfare. The optimal settlement speed is determined by the network structure and the trade-off between netting efficiency and liquidity costs on one hand, and the probability of counterparty defaults on the other. We identify conditions where faster settlement can increase systemic risk by amplifying crisis severity, despite reducing crisis likelihood. Our insights have important policy implications, arguing against a one-size-fits-all approach to settlement speed design.

**Keywords:** settlement, payment systems, financial network, financial stability, systemic risk

**JEL Classification Numbers:** D49, D53, G01, G21, G33

*We thank Francesca Carapella, Adam Copeland, Michael Gordy, David Rappoport, and Alexandros Vardoulakis for helpful comments and suggestions. We also thank seminar participants at the Federal Reserve Board. We thank Siddhartha Lewis-Hayre for excellent research assistance. This article represents the view of the authors and should not be interpreted as reflecting the views of the Federal Reserve System or its members.

†Department of Industrial Engineering and Operations Research and Columbia Business School, Columbia University, New York, NY, USA. *Email:* ac3827@columbia.edu.

‡Board of Governors of the Federal Reserve System, *Email:* jin-wook.chang@frb.gov.

# 1. Introduction

The Securities and Exchange Commission (SEC) has recently implemented the conversion of the U.S. securities market to a T+1 (one business day after the trade date) settlement cycle, starting May 28, 2024. Following the SEC's implementation of the new settlement speed, there has not been a noticeable change in settlement fails other than trading halts, which were due to technical malfunctions, in the New York Stock Exchange on June 4, 2024. However, the question of "whether shorter settlement lags would increase the likelihood and severity of systemic failures in payment systems or not" remains open.[1] This question is particularly timely, given the recent expansion of real-time payment systems such as FedWire, FedNow, and Real-Time Payments (RTP), as well as the ongoing trend toward faster deferred net settlement systems like the Clearing House Interbank Payments System (CHIPS).

Despite extensive study, research on the systemic risk implications of altering the settlement cycle remains scarce. In particular, payment systems can amplify initial shocks because of their critical role in the financial system and their complex network structure, which can propagate shocks rapidly through cascades of defaults.[2] Therefore, understanding how the interconnectedness of payment systems affects financial stability, and how this relationship changes with settlement speed, is crucial for both policymakers and researchers.

We develop a network model to study the role of settlement in payment systems. Our model incorporates four core features of payment systems: (i) networked payment flows, (ii) netting of obligations, (iii) liquidity costs of funding payments, and (iv) idiosyncratic liquidity shocks, including counterparty defaults.

We contribute to the financial networks literature by developing a model that incorporates an understudied dimension, *time*. Moreover, we consider ex ante social welfare, accounting

[1] Throughout this paper, we use the term "payment system" to refer broadly to payment, clearing, and settlement systems for a variety of financial transactions, including traditional payments, securities, repurchase agreements, and derivatives.

[2] See, for a recent example, Kotidis and Schreft (forthcoming).

for both the probability of shocks and the severity of any resulting contagion. By contrast, most of the financial networks literature focuses on ex post welfare, because standard models hold shock probabilities fixed and thus cannot accommodate changes in their likelihood. Within our framework, we uncover a trade-off between lowering the likelihood of shocks and increasing the severity of contagion conditional on a shock.

We consider a network of $n$ agents, each holding a cash buffer and having senior debt obligations. These agents are interconnected through a payment network, represented by a matrix $D$, where each entry $d_{ij}$ denotes the payment promised by agent $j$ to agent $i$. The key parameter of our model is $\tau$, which represents the settlement time for all transactions in the network. As $\tau$ decreases, the settlement speed increases, with $\tau = 0$ implying instant settlement.

The model has two crucial functions that vary with the settlement time $\tau$: the liquidity cost function $L(\tau)$ and the netting function $\alpha(\tau)$. The liquidity cost function, which is decreasing in $\tau$, captures the costs associated with raising liquidity or posting collateral to finance transactions. The netting function, which is also decreasing in $\tau$, represents the degree of liabilities that are not netted in the network, with $(1 - \alpha(\tau))$ indicating the proportion of transactions netted. Institutional background and empirical evidence supporting $L'(\tau) \leq 0$ and $\alpha'(\tau) \leq 0$ are discussed in Section 1.1.

The model also incorporates counterparty risk through a random liquidity shock that can hit some of the agents in the network before settlement occurs. The probability of this shock arriving before settlement is given by $F(\tau)$, which is increasing in $\tau$. When the shock arrives, it reduces the available cash of affected agents, potentially leading to defaults, which can propagate and cause further defaults through the network. Such defaults incur deadweight losses to the economy that are proportional to the amount of payment shortfalls.

The model implies a clear trade-off: *faster settlement* (smaller $\tau$) reduces the likelihood of default contagion $F(\tau)$, but increases liquidity costs $L(\tau)$ and leaves a larger share of liabilities unnetted—i.e., raises $\alpha(\tau)$ (equivalently, reduces $1 - \alpha(\tau)$). The ex-post welfare implications when a shock arrives is clear, as we show in Lemma 1, because a smaller $\tau$ reduces available cash and amplifies contagion via higher gross (less netted) exposures, thereby lowering ex-post welfare. By contrast, the *ex ante* welfare effect is generally ambiguous: while faster settlement may reduce the probability of stress events, it increases the severity of contagion when they occur.

We obtain key theoretical results that shed light on the complex relationship between settlement speed, network structure, and financial stability.

First, we establish the existence of a discontinuous contagion pattern. Specifically, changes in settlement time can cause social welfare to increase or decrease sharply, depending on the underlying network structure. This effect is especially pronounced when comparing the complete and ring networks in Figure 2. In the complete network, if settlement is sufficiently slow, only a single agent defaults when shocks occur. By contrast, under the same conditions, the ring network exhibits multiple defaults, with the number of defaulting agents rising as settlement speed increases. When settlement becomes fast enough, however, all agents in both networks default. This abrupt shift in contagion dynamics—a phase transition illustrated in Figure 3—is consistent with findings in the financial networks literature and has important implications for the ex ante evaluation of social welfare.

Second, we introduce the concept of default threshold points, defined as settlement times at which the number of defaulting agents in the network changes discontinuously. These threshold points are shown to be critical in determining the ex-ante welfare implications of settlement speed. This insight extends beyond the comparison of complete and ring networks, highlighting that the welfare effects of changes in settlement speed can differ dramatically depending on whether the system operates at or above these threshold points. Furthermore, we show that identifying default threshold points is solvable in polynomial time.

Third, we derive a novel result regarding the differential impact of settlement speed on different network structures. We prove that faster settlement time can improve ex ante social welfare of the ring network but worsen ex ante social welfare of the complete network. This heterogeneity in welfare effects underscores the importance of considering network structure when designing settlement systems or implementing changes to settlement speeds.

Fourth, we show that the node depth centrality is a key measure of an agent's systemic importance. This measure captures how the initial impact of a default by one agent is amplified through the network of defaulting agents, providing a useful tool for identifying systemically important institutions in payment networks.

Finally, we derive conditions under which faster settlement creates a trade-off between reducing the likelihood of systemic events and increasing their severity when they occur.

Altogether, our results provide new insights into the design of optimal settlement systems and emphasize the importance of nuanced interactions between various factors and interconnectedness for policymakers. Our findings have direct policy implications. We show that a one-size-fits-all approach to settlement speed may be suboptimal, given the heterogeneity in network structures. Moreover, our results provide a theoretical underpinning for targeted liquidity support to central entities of payment systems during times of stress. Finally, we show that faster settlements can decrease the likelihood of crisis events but increase their severity, potentially leading to more frequent government interventions due to stronger incentives to intervene.

## 1.1. Related Literature

Our paper contributes to the literature on financial networks and their implications for systemic stability starting from Allen and Gale (2000). The foundational work in this area is Eisenberg and Noe (2001), which has been extended in subsequent studies including Rogers and Veraart (2013), Elliott, Golub, and Jackson (2014), Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015), Glasserman and Young (2015), Bernard, Capponi, and Stiglitz (2022), Capponi, Corell, and Stiglitz (2022), Donaldson, Piacentino, and Yu (2022), Jackson and Pernoud (2024), and Chang and Chuan (2024). Our main contribution is to incorporate the time dimension by examining how shorter settlement horizons affect the tradeoff between counterparty risk and netting efficiency. Specifically, we evaluate ex ante expected social welfare, taking into account both the probability of shocks and the extent of contagion should they materialize. Chang and Chuan (2025) also examine the ex ante expected social welfare in financial networks and find that an increase in collateral reuse decreases the likelihood of crisis but increases the severity of crisis. Their result resembles our finding that an increase in settlement speed decreases the likelihood of stress events but increases their severity. However, in their model the uncertainty arises from the selection among multiple equilibria, whereas in ours it stems from the probability of shock arrival—an exogenous source that is more fundamental and directly tied to netting efficiency and liquidity costs.

A closely related to ours is Khapko and Zoican (2020), who study the effect of settlement speed on the tradeoff between the impatience of traders and costly borrowing. We extend their framework by explicitly incorporating networks and netting, thereby allowing us to analyze the systemic risk implications of settlement speed.

This paper is also related to the literature on portfolio compression or netting of liabilities (e.g. D'Errico and Roukny (2021), Veraart (2022), and Chang (2021)). Our time dimension allows incorporating the tradeoff between a greater degree of netting $\alpha(\tau)$ and greater likelihood of shock arrival $F(\tau)$.

A large body of empirical and institutional research supports our key assumptions regarding the liquidity cost function $L(\tau)$ and the netting efficiency function $\alpha(\tau)$. For example, Rochet and Tirole (1996) emphasize the systemic risks in payment systems, including widespread liquidity shortages and the propagation of failures. They show that net settlement systems, such as the Clearing House Interbank Payment System (CHIPS), can alleviate liquidity pressures through netting, whereas real-time gross settlement systems, such as Fedwire, help prevent failure propagation but are more vulnerable to liquidity constraints. VanHoose and Sellon Jr (1989) find an increase in daylight overdraft amounts commensurate with the increase in payment amounts, Hancock and Wilcox (1996) find that the participants in Fedwire react to the fees of daylight overdraft and holds more reserves in response to policy changes, and McAndrews and Rajan (2000) find that the cost of overdraft fees and from other sources of liquidity drive payments in Fedwire to be concentrated in late afternoon. Kahn, McAndrews, and Roberds (2003) show that net settlement systems can prevent gridlock situations and economize collateral requirements, and avoid trading delays. The timing decisions are further formally analyzed by Mills Jr and Nesmith (2008), who discuss the effects of Real Time Gross Settlement (RTGS) as well as delivery-versus-payment (DVP) and Delayed Net Settlement (DNS). Mills Jr and Nesmith (2008) emphasize that RTGS and DVP systems settle payments and securities one at a time, so liquidity is needed to complete each transaction, incentivizing participants to concentrate trades later in the day. Afonso and Shin (2011) also emphasize that the Fedwire system requires large amounts of liquidity for participating banks that can trigger precautionary demand in liquidity, which could result in a systemic event.

Copeland and Garratt (2019) show that real-time gross settlement systems such as Fedwire require large amounts of cash to process relatively small net transactions, whereas delayed settlement systems such as CHIPS can clear large gross obligations with a much smaller net cash outlay. This difference arises from the fact that CHIPS provides a netting service, which acts as a liquidity-saving mechanism. Furthermore, dozens of large banks (along with some smaller ones) conduct substantial volumes of transactions through CHIPS, whereas thousands of banks rely on Fedwire to settle smaller net amounts.

The study by Ding, Gonzalez, Ma, and Zeng (2025) examines the effect of instant payment systems by using administrative banking data and transaction-level payment data from Brazil's Pix, which is managed by the Central Bank of Brazil (BCB). Ding, Gonzalez, Ma, and Zeng (2025) find that banks increased their liquid asset holdings and lent out more subprime and defaulting loans after the adoption of instant payment systems. This occurred because the convenience of instant payments to consumers comes at the cost of banks' ability to delay and net payments. Together, the findings of Copeland and Garratt (2019) and Ding, Gonzalez, Ma, and Zeng (2025) provide strong empirical support for the central tradeoff in our model, between liquidity costs and netting on one side, and counterparty risk on the other.

Concerns about liquidity costs can even affect trading activity itself. Brunnermeier and Pedersen (2009) uncover a self-reinforcing feedback loop exists between a trader's funding liquidity and an asset's market liquidity. Andolfatto (2020) show that during periods of stress, market participants tend to hoard liquidity, leading to lost trading opportunities. Acharya and Merrouche (2013) provide empirical evidence that large settlement banks hoarded liquidity during the 2007-2008 subprime crisis, because of increased uncertainty about counterparty risk, leading to a freeze of the interbank lending market.

The role of netting efficiency has been studied in the context of clearinghouses (see Duffie and Zhu (2011) for a theoretical analysis and Duffie, Scheicher, and Vuillemey (2015) for an empirical study). Carapella and Mills (2011) formalize the role of central counterparties (CCPs) in payment systems, showing that delaying transactions can help mitigate informational frictions. When CCPs provide netting and mutualize losses (e.g. Capponi, Wang, and Zhang (2022)), netting reduces the cost of trading with informed counterparties, while mutualization lowers potential losses, thereby enhancing incentives to trade. Hence, longer settlement horizons can improve liquidity conditions and netting efficiency by alleviating information asymmetries, ultimately supporting more mutually beneficial trading opportunities and reinforcing the central mechanism of our model.

# 2. Model

We develop a general model that can be applied to both payment systems and securities settlement systems. The following is an overview of the model, which is designed to capture a slice of the continuous transactions in a parsimonious way. For every slice or snapshot of time, there is a set of transactions that are scheduled to be settled sometime later. Some amount of liquidity is needed for each slice of transactions to settle them. Over time, some transactions can be netted. However, over time, some counterparties may receive shocks and default. Although there will be new transactions in the following periods, not being able to settle the initial slice of transactions would lead to default. Thus, our model captures a dynamic tradeoff for given transactions at a particular point in time.

## 2.1. Agents and Payment Network

There are $n$ different agents, and the set of all agents is $N = \{1, 2, \ldots, n\}$ with $n > 2$. Time is continuous and denoted as $t \in [0, \bar{T}]$, where $\bar{T}$ is finite. Agents are risk neutral, do not discount future up to $\bar{T}$,[3] and their utility is determined by how much net payments they receive less liquidity cost (described below) at $t = \bar{T}$. Each agent $i$ holds $e_i > 0$ amount of cash buffer (or operating net cash flow as in Eisenberg and Noe (2001)), and has senior debt (such as deposits) to pay in the amount of $s_i > 0$. Each agent is connected to each other through transaction agreements, which are exogenously given at $t = 0$.[4] The amount of $j$'s payment promised (liability) to $i$ is denoted by $d_{ij}$. The payment network, which is defined as the network of payments promised, is denoted by the matrix $D \equiv [d_{ij}]_{i,j,\in N}$. Denote the sum of payment obligations or the column sums of $D$ as $d_j \equiv \sum_{i \in N} d_{ij}$. Define $n \times 1$ vectors $d \equiv (d_i)_{i \in N}$, $e \equiv (e_i)_{i \in N}$, and $s \equiv (s_i)_{i \in N}$. The payment matrix can be represented as $D = Q \circ d^T$, where $\circ$ is the Hadamard (Schur) product operator, and $Q$ is the matrix of weights with its $(i,j)$-element as

$$q_{ij} = \frac{d_{ij}}{d_j}. \tag{1}$$

[3] Although this is a simplifying assumption, we can consider $\bar{T}$ as a relatively short horizon, which is typically the case, so the effective discount rate is negligible.

[4] For example, banks do not have much control over how payments are made across their own depositors and depositors of other banks. Another example is that large dealers tend to trade with counterparties with whom they have long-term relationships, which are costly to adjust even when there are clear arbitrage opportunities (Chang, Klee, and Yankov, 2025).

## 2.2. Settlement Speed, Netting, and Liquidity Cost

**Settlement speed.** The time to settle all transactions in $D$ is denoted by $\tau$, which is the key parameter of the model. By construction, $\tau \in [0, \bar{T}]$. As $\tau$ becomes smaller (larger), the settlement speed increases (decreases), and $\tau = 0$ implies instantaneous settlement.

**Liquidity cost.** Agents need to raise liquidity (or post collateral) to finance their transactions. Such liquidity needs may decrease over time when settlement happens in later time. For example, if a transaction is the purchase of a security, an agent (dealer) may find a trader (outside of the network) who is willing to purchase the security from the agent over time. However, if an agent (dealer) needs to settle the transaction right away, the agent has to use its own funds to purchase or borrow the security. These effects are captured by a reduced-form *liquidity cost function* denoted as $L(\tau)$, which is differentiable and strictly decreasing in $\tau$ and $L(\tau) \leq e_i - s_i$, $\forall i \in N$ and $\forall \tau \in [0, \bar{T}]$.[5] Further, we assume that the liquidity cost is senior to payments in the network. In sum, $L(\tau)$ represents the amount of the opportunity cost of cash and the external netting of payments.

**Netting.** The matrix of transactions may have cycles or simply a path, which could be compressed through netting of transaction liabilities. Define the *full netting matrix* $\underline{D}$ as the matrix of transactions when all possible netting is conducted for the original network $D$, and $\underline{D} \leq D$. The exact details of what type of netting is conducted are discussed in Appendix B.

However, the actual netting may not reach the full netting matrix for many reasons, such as limited information and staggered maturities or covenants (Donaldson, Piacentino, and Yu, 2022; Jackson and Pernoud, 2024).[6] The degree of netting can be greater when

[5] Khapko and Zoican (2020) provide a microfoundation for this liquidity cost. For example, broker-dealers have to prefund their trades and face inventory risk (or opportunity cost of funds) or borrow securities that they need to settle, and costs associated with both increase as settlement speed increases. In other cases, agents in a payment system may have to rely on external short-term borrowing because mobilizing internal liquidity takes time as settlement speed increases. Moreover, shorter payment delay means that agents cannot use expected incoming funds to cover outgoing payments. In this paper, we are agnostic about the source of the cost and take the decreasing liquidity cost as a reduced-form functional property.

[6] For example, D'Errico and Roukny (2021) find that even the most conservative compression scenario

the settlement occurs later, as agents and the payment system can identify more netting opportunities or simply because some of the transactions in $D$ are actually formed later (Kahn and Roberds, 1998). In other words, as the settlement speed decreases ($\tau$ increases), the total required transactions to be settled decrease. We represent this degree of netting by the *netting function* $\alpha(\tau)$, which is differentiable and decreasing in $\tau$, and $\alpha(\tau) \in [0,1]$ for any $\tau \in [0, \bar{T}]$. We assume the actual required transactions are a convex combination of the full netting matrix and the original transaction matrix as

$$\hat{D} \equiv (1 - \alpha(\tau))\underline{D} + \alpha(\tau)D, \tag{2}$$

which we call as the *partial netting matrix* of $D$ for given $\tau$.[7] We use corresponding notations for the full netting and partial netting matrices: $\underline{d}_{ij} \equiv \underline{D}_{ij}$, $\underline{d}_j \equiv \sum_{i\in N} \underline{d}_{ij}$, $\underline{d} \equiv (\underline{d}_i)_{i\in N}$, $\hat{d}_{ij} \equiv \hat{D}_{ij}$, $\hat{d}_j \equiv \sum_{i\in N} \hat{d}_{ij}$, and $\hat{d} \equiv \left(\hat{d}_i\right)_{i\in N}$. Therefore, $(1 - \alpha(\tau))$ represents the degree of internal netting of payments within the payment network.

## 2.3. Counterparty Risk and Default Spillovers

At any point in time, there is a small probability that some agents receive liquidity shocks and default. The arrival time of liquidity shocks $t_l$ is a random variable with full support $[0, \bar{T}]$. Therefore, the probability of the arrival time being less than or equal to $\tau$ (i.e. the probability of liquidity shocks arriving before the settlement of payments) is $F(\tau) \equiv \Pr(t_l \leq \tau)$, which is differentiable and increasing in $\tau$. In other words, the likelihood of a liquidity shock disrupting the payment system is increasing in $\tau$. Once liquidity shocks arrive, each agent $i$ experiences an immediate reduction in available cash in the amount of $\epsilon_i \geq 0$. Denote the $n \times 1$ vector of liquidity shocks as $\epsilon = (\epsilon_i)_{i\in N}$. For exposition purposes, we slightly abuse notation and denote $\epsilon_i = 0$ when agent $i$ does not receive a liquidity shock.

eliminates on average, more than 85% of the CDS positions in the European Union and all their counterparties.

[7] This general representation of netting encompasses a reduced-form representation of netting in Donaldson, Piacentino, and Yu (2022).

The vector of liquidity shocks $\epsilon$ is a random variable with the probability mass function $G(\epsilon)$ for any $\epsilon \in \mathbb{R}^n_+$.

Due to the liquidity cost and liquidity shocks, some agents may default, as their obligations may exceed their available liquidity. Therefore, agent $j$ can pay the full promised amount (after netting) $\hat{d}_j$ when $j$ has enough liquidity, or $j$ would default when the full promised amount exceeds the total net liquidity—the sum of payments from other agents and cash buffer less of senior debt and liquidity cost paid.[8] Denote the actual total payment made by agent $j$ as $x_j(\tau, D, \epsilon)$. The arguments $\tau$, $D$ and $\epsilon$ are often omitted in the following for expositional simplicity. Agent $j$ will pay the total amount of $x_j < \hat{d}_j$, if agent $j$ is defaulting, or $x_j = \hat{d}_j$, if agent $j$ is solvent. Following the literature, we assume the pro rata rule for distributing the remaining wealth of a defaulting agent to other agents; hence the payment to agent $i$ from agent $j$ is $q_{ij}x_j$.[9]

Finally, we assume that a default incurs deadweight loss in terms of real cost as in Rogers and Veraart (2013), Elliott, Golub, and Jackson (2014), and Glasserman and Young (2015) among many others. The deadweight loss can be interpreted as legal costs or costs of delays and inefficient allocation of resources during the bankruptcy procedure. In particular, following Bernard, Capponi, and Stiglitz (2022) and Capponi, Corell, and Stiglitz (2022), the default of agent $j$ inflicts a deadweight loss of $\min\{\beta\xi_j(\tau, D, \epsilon), A_j(\tau, D, \epsilon)\}$, where $A_j(\tau, D, \epsilon) \equiv e_j + \sum_{k\in N} q_{jk}x_k(\tau, D, \epsilon)$ is the value of agent $j$'s assets (liquidity buffer plus payments received from other agents), $\xi_j(\tau, D, \epsilon) \equiv \left[\hat{d}_j + s_j + \epsilon_j + L(\tau) - A_j(\tau, D, \epsilon)\right]^+$ is agent $j$'s shortfall (total payment obligations + senior debt + liquidity shock + liquidity cost − assets), and $\beta > 0$ is a scaling parameter. The arguments $\tau$, $D$, and $\epsilon$ are often omitted in the following. Note that the deadweight loss is bounded above by the total value of the agent's assets. The deadweight loss can cause either a decline in the defaulting agents'

[8]Here we are simply treating transactions in $\hat{D}$ as being made in terms of cash. This is without loss of generality, as for any transactions with securities, we can just re-normalize everything with the price $p$.

[9]This is a standard assumption in the literature, such as in Eisenberg and Noe (2001) and Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015). Csóka and Herings (2021) provide an axiomatization of the proportional rule in financial networks.

remaining asset value (hence less payment to other agents) or purely a societal cost to the un-modeled external agents. Denote the modeled agents' share of the deadweight loss as $\gamma$.

Mathematically, the total payments made by agent $j$ is

$$x_j(\tau, D, \epsilon) \equiv \begin{cases} \hat{d}_j, & \text{if } A_j(\tau, D, \epsilon) \geq \hat{d}_j + L(\tau) + s_j + \epsilon_j \\ \left[A_j(\tau, D, \epsilon) - L(\tau) - s_j - \epsilon_j - \gamma \min\left\{\beta \xi_j(\tau, D, \epsilon), A_j(\tau, D, \epsilon)\right\}\right]^+, & \text{otherwise,} \end{cases} \tag{3}$$

where $[\cdot]^+ \equiv \max\{\cdot, 0\}$ denotes the positive part. Thus, in a matrix-vector notation, the payment that clear the payment network is

$$x = \left[\min\left\{\hat{d}, Qx + e - L(\tau)\mathbf{1} - s - \epsilon - \gamma \min\left\{\beta\xi, A\right\}\right\}\right]^+, \tag{4}$$

where $x \equiv (x_i(\tau, D, \epsilon))_{i \in N}$, $\xi \equiv (\xi_i(\tau, D, \epsilon))_{i \in N}$, $A \equiv (A_i(\tau, D, \epsilon))_{i \in N}$, and $\mathbf{1}$ is a vector of ones for the appropriate dimension.

The (ex post) equilibrium concept we use is *payment equilibrium*, which satisfies (4) for the given realization of liquidity shocks, liquidity costs, and netting. This equilibrium concept is consistent with that of Eisenberg and Noe (2001) and many other papers in the financial networks literature. The existence of multiple Pareto-ranked payment equilibrium follows directly from Proposition 3 in Glasserman and Young (2015). There could be multiple equilibria due to the fact that defaults can generate discontinuous drops in the asset values, resulting in self-fulfilling defaults.[10] However, all equilibria can be Pareto-ranked, so we focus on the unique Pareto dominant payment equilibrium, i.e. the equilibrium with the lowest total payment shortfalls and social welfare loss.

[10] See Chang and Chuan (2025) for a detailed discussion of this property of financial network models and an analysis of equilibrium selection.

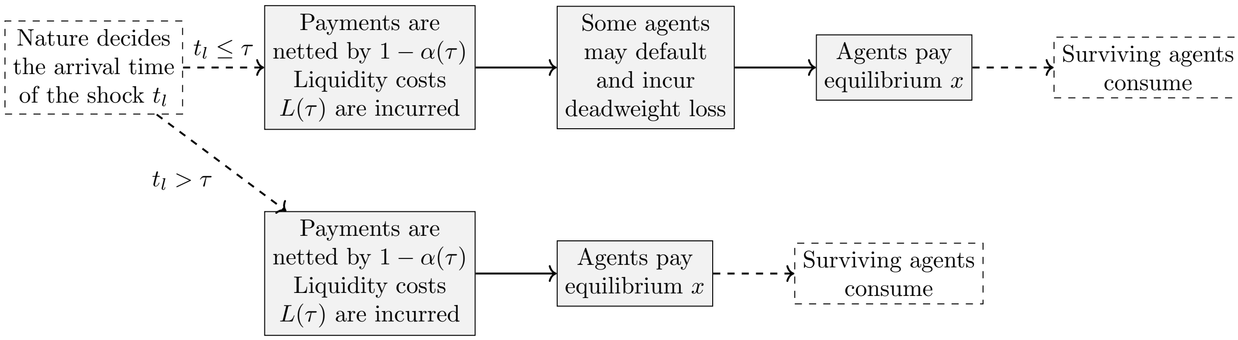



Figure 1: Timeline of the Model

## 2.4. Timeline

Figure 1 summarizes the timeline of the model. The nature decides the arrival time of the shock $t_l$. If the settlement time $\tau$ is less than $t_l$, then there will be no defaults, payments are netted by $1 - \alpha(\tau)$, reducing the actual payment network as $\hat{D} = (1 - \alpha(\tau))\underline{D} + \alpha(\tau)D$, and liquidity costs $L(\tau)$ are paid by agents. Agents will settle all the payments, and surviving agents will consume their resulting wealth at $t = \bar{T}$. If the settlement time $\tau$ is greater than or equal $t_l$, then some agents may default (possibly from spillovers), depending on the equilibrium payments $x$. Surviving agents consume but defaults also incur deadweight loss of $\beta$ multiplied by the amount of payment shortfalls.

## 2.5. Ex Ante Social Welfare Loss

Agents' utility is determined by the net payment they receive less the liquidity cost. Hence, the expected utility depends on the resulting equilibrium vector of payments that varies by whether there is a liquidity shock or not. Note that the amounts of cash buffer, senior debt, liquidity cost, and payment obligations remain the same with or without the arrival of a liquidity shock. Without loss of generality, we assume that all obligations in $D$ can be made in full if no agent defaults.[11] Then, by construction, all payments are made in

[11]This assumption helps us to focus on additional social welfare losses due to changes in the counterparty default probabilities resulting from changes in settlement time and network structure. Our results can be extended to the case in which some agents default on their obligations even if they are paid in full.

full if no liquidity shock has arrived. With probability $F(\tau)$, a liquidity shock arrives before the settlement, and some agents may default and incur deadweight losses. Therefore, the ex ante expected utility of agent $i$ for a given settlement time $\tau$ is

$$
\begin{aligned}
U_i(\tau, D) =& e_i - s_i - L(\tau) + (1 - F(\tau)) \left[ \sum_{j \in N} q_{ij} \hat{d}_j - \hat{d}_i \right] \\
&+ F(\tau)\, E \left[ \sum_{j \in N} q_{ij} x_j - x_i - \epsilon_i - \gamma \min \left\{ \beta \xi_i(\tau, D, \epsilon), A_i(\tau, D, \epsilon) \right\} \right],
\end{aligned} \tag{5}
$$

where the payments and values relevant to deadweight losses, $(x_i, \xi_i, A_i)_{i \in N}$, are determined by the payment equilibrium when liquidity shocks arrive, and $E[\cdot]$ is the expectation over different realizations of these liquidity shocks with probability $G(\epsilon)$.

Define the ex ante (expected) social welfare losses for a given $\tau$ and network $D$ as

$$
W(\tau, D) \equiv \sum_{i \in N} \left( L(\tau) + F(\tau) E \left[ \min \left\{ \beta \xi_j(\tau, D, \epsilon), A_j(\tau, D, \epsilon) \right\} \right] \right), \tag{6}
$$

where the first-term is the liquidity cost for settling trades at $\tau$, and the second term is the sum of the total deadweight losses for internal network agents ($\gamma$ proportion) and for external agents ($(1-\gamma)$ proportion) multiplied by the probability of shock arrival $F(\tau)$. Note that the payment shortfall to the liquidity shock $\epsilon$ is not included in the welfare loss, as the payments to $\epsilon$ are simple transfers of wealth from the shocked agents to some unmodeled external agents (senior creditors), i.e., it is zero sum.[12] See Appendix C for a detailed discussion.

Equation (6) implies that the ex ante social welfare loss is determined by the liquidity costs, and the expected deadweight losses from payment shortfalls, multiplied by the probability of liquidity shock arrival. As $\tau$ increases, the likelihood of liquidity shock state increases, the inter-agent liabilities decrease, and the liquidity cost decreases. Hence, the resulting ex ante social welfare loss depends on the tradeoff and functional forms of $\alpha(\tau)$ and

[12] This definition of social welfare loss generalizes the typical definition of ex post social surplus in the literature. The ex post social surplus in our model is consistent with that of Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015) and Chang and Chuan (2024).

$L(\tau)$ as well as the structure of the network $D$ and resulting spillovers. Note that increase in $\tau$ also decreases the liquidity cost when there is no liquidity shock.

# 3. Settlement Speed and Contagion

## 3.1. Discontinuous Contagion Patterns

We establish our first main result, which shows the discontinuous contagion patterns in our model. When the settlement time $\tau$ changes, the social welfare may drastically increase or decrease depending on the network structure. To illustrate the main mechanism clearly, we impose a few simplifying assumptions in this subsection.

Following the literature, we focus on *regular networks*, which have homogeneous agents, i.e., $e_j = \bar{e}$ and $s_j = \bar{s}$ for any $j \in N$, with the same in-degree and out-degree of payments, i.e., $\sum_{i \in N} d_{ij} = \sum_{i \in N} d_{ji} = \bar{d}$ (Acemoglu, Ozdaglar, and Tahbaz-Salehi, 2015; Donaldson, Piacentino, and Yu, 2022).[13] Further, we assume the full netting matrix would eliminate all cycles in the network.[14] Specifically, for any connected regular network, the full netting matrix will be $\underline{D} = \mathbf{0}$, where $\mathbf{0}$ is the zero matrix. To simplify the feedback from deadweight losses, we focus on the indirect deadweight loss case in which agents in the network do not suffer the deadweight loss from payment shortfalls directly, i.e. $\gamma = 0$. Finally, we assume that only one of the $n$ agents is hit by the liquidity shock once it arrives, and the size of the liquidity shock is fixed as a single value $\bar{\epsilon}$. Hence, the liquidity shock vector will be $\epsilon_i = \bar{\epsilon}$ and $\epsilon_j = 0$ for any $j \neq i$. Also, we focus on the case with $\bar{\epsilon} > \bar{e} - \bar{s} - L(\bar{T})$ to exclude trivial cases of no default even after receiving a liquidity shock. This simplified setup enables us to compare

[13]In addition to the benefits of bringing simplicity and tractability due to symmetry, this setup allows us to focus on the network connection itself rather than the size or asymmetry of payment flows. Given that payment networks tend to have net zero flows in a steady state (because otherwise an agent would be depleted of cash over time), this assumption is relatively in line with real-world payment networks.

[14]The notion of netting or portfolio compression here is "conservative compression" defined by D'Errico and Roukny (2021). A "nonconservative compression" is possible if new edges of liabilities can be formed. For a practical example, if transactions are centrally cleared, they can be netted even without cycles. See Section B for more discussion.

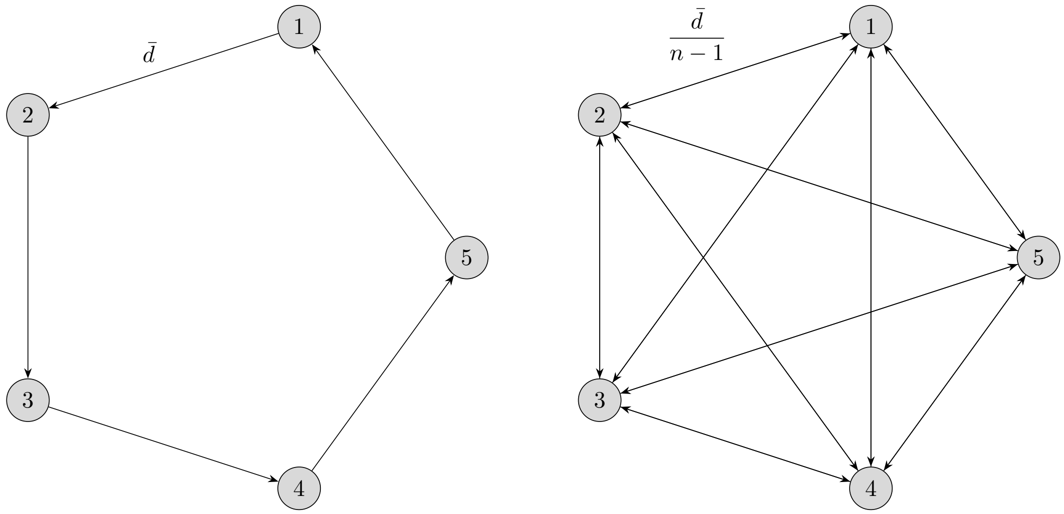



Figure 2: The ring network and the complete network

Note: The left graph depicts the ring network, in which each agent owes the full liability amount of $\bar{d}$ to the next agent. The right graph depicts the complete network, in which each agent owes $\dfrac{\bar{d}}{n-1}$ to every other agent.

our model directly with other key models in the literature, most notably that in Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015).

First, we focus on two special cases of regular networks: ring and complete networks, depicted in Figure 2. The ring network is a structure in which agent $i$ is the only agent to be paid by agent $i-1$ for each $i > 1$ (for the entirety of $\bar{d}$), and agent 1 is the only agent to be paid by agent $n$. The complete network is a structure in which every agent owes the same amount to each other agent, $\bar{d}/(n-1)$. Denote the ring network and complete network as $D_r$ and $D_c$, respectively.

The first proposition establishes the discontinuous patterns of contagion in payment equilibrium for the complete network, which contrasts with the pattern of contagion in the ring network. In particular, we consider two moving parts, $L(\tau)$ and $\alpha(\tau)$, to prepare for comparative statics with the settlement time $\tau$. As $\tau$ increases, $L(\tau)$ decreases, and agents have more liquidity. Therefore, agents can absorb larger shocks with their liquidity buffer. In addition, as $\tau$ increases, more netting can be done (lower $\alpha(\tau)$), reducing the total inter-agent liability amount $\alpha(\tau)\bar{d}$. This further reduces the inter-agent exposures, decreasing the contagion throughout the network. Therefore, understanding the roles of the liquidity cost, $L(\tau)$, and the netting function, $\alpha(\tau)$, is crucial for the ultimate comparative statics with the settlement time, $\tau$.

The proposition shows the following: For the complete network, if either the size of the liquidity shock is relatively small (compared to the liquidity buffer) or the size of the total liabilities is relatively small (compared to the liquidity buffer), then only one agent defaults. In comparison, the number of defaulting agents is greater than one in the ring network under the same conditions. This result aligns with the results by Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015), but extends their results by considering different levels of the total liabilities.

In contrast, when both the relative size of the liquidity shock and the relative size of the total liabilities are large, all agents in the complete network default, and the same happens in the ring network. The complete network has the highest number of defaults in a certain parameter region, while it has the lowest number of defaults in the rest of the parameter space. This abrupt change in contagion pattern, or phase transition, is in line with existing literature (Acemoglu, Ozdaglar, and Tahbaz-Salehi, 2015; Chang and Chuan, 2024).

However, there is a subtle difference between the results in our model and those in the literature. In typical models, such as those in Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015) and Chang and Chuan (2024), the complete network and the ring network have the same level of social welfare when full contagion (i.e., all agents defaulting) occurs, because both networks have the same number of defaults. In contrast, our model shows that the complete network has a smaller total welfare loss than the ring network when the deadweight losses don't hit their upper bound, i.e., not all of the assets disappear due to deadweight losses. This novel result further identifies a subtle difference in contagion pattern even when the number of defaults is the same.

The main factor is our definition of deadweight loss, which is different from simply counting the number or the amount of defaults in Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015) and Chang and Chuan (2024). Even when all agents default, the total deadweight loss depends on the distribution of payment shortfalls in our model. In the ring network each defaulting agent is simply a conduit of the payments from previous agents to the next agent, until the payment flow reaches the agent under shock. However, in the complete network, each agent is using its payment to pay everyone else, who then reuses the payment to pay everyone else, and so on. The excess propagation of payments reduces payment shortfalls in the complete network, reducing the deadweight loss. This novel mechanism is relevant to deadweight loss in the real world, as deadweight losses would be smaller in more complexly interconnected payment systems.

**Proposition 1.** *For a fixed settlement time $\tau$, suppose that a liquidity shock $\bar{\epsilon}$ hits one agent, and denote $\epsilon^*(\tau) \equiv n(\bar{e} - \bar{s} - L(\tau))$ and $d^*(\tau) \equiv (n-1)(\bar{e} - \bar{s} - L(\tau))$.*

1. *Suppose that $D = D_c$, the complete network.*

   (a) *If either the relative size of the liquidity shock is small or the relative size of the total liabilities is small, i.e. $\bar{\epsilon} \leq \epsilon^*(\tau)$ or $\alpha(\tau)\bar{d} \leq d^*(\tau)$, then only one agent defaults and the total deadweight loss from default is*

$$\min\left\{\beta\left(\bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)\right), \bar{e} + \alpha(\tau)\bar{d}\right\}. \tag{7}$$

   (b) *If both the relative size of the liquidity shock is large and the relative size of the total liabilities is large, i.e. $\bar{\epsilon} > \epsilon^*(\tau)$ and $\alpha(\tau)\bar{d} > d^*(\tau)$, then all agents default and the total deadweight loss from defaults is*

$$\begin{aligned} &\min\left\{\beta\left[\alpha(\tau)\bar{d} + \bar{\epsilon} - n(\bar{e} - \bar{s} - L(\tau))\right], \bar{e} + (n-1)(\bar{e} - \bar{s} - L(\tau))\right\} \\ &+ (n-1)\min\left\{\beta\left[\alpha(\tau)\bar{d} - (n-1)(\bar{e} - \bar{s} - L(\tau))\right], \bar{e} + (n-2)(\bar{e} - \bar{s} - L(\tau))\right\}. \end{aligned} \tag{8}$$

2. *Suppose that $D = D_r$, the ring network.*

(a) *If either the relative size of the liquidity shock or the relative size of the total liabilities is small, i.e.,* $\bar{\epsilon} \leq \epsilon^*(\tau)$ *or* $\alpha(\tau)\bar{d} \leq d^*(\tau)$*, then* $\psi$ *other agents default in addition to the shocked agent, and the total deadweight loss from defaults is*

$$\begin{aligned} &\min\left\{\beta(\bar{s}+\bar{\epsilon}-\bar{e}+L(\tau)), \bar{e}+\alpha(\tau)\bar{d}\right\} \\ &+\sum_{m=1}^{\psi} \min\left\{\beta\left(\alpha(\tau)\bar{d}-m(\bar{e}-\bar{s}-L(\tau))-\phi\right), \bar{e}+(m-1)(\bar{e}-\bar{s}-L(\tau))+\phi\right\}, \end{aligned} \tag{9}$$

*where* $\phi = \left[\bar{e}-\bar{s}-L(\tau)+\alpha(\tau)\bar{d}-\bar{\epsilon}\right]^+$*, and* $\psi = \min\left\{n-1, \left\lfloor \frac{\alpha(\tau)\bar{d}-\phi}{\bar{e}-\bar{s}-L(\tau)} \right\rfloor\right\}$.[15]

(b) *If both the relative size of the liquidity shock is large and the relative size of the total liabilities is large, i.e.* $\bar{\epsilon} > \epsilon^*(\tau)$ *and* $\alpha(\tau)\bar{d} > d^*(\tau)$*, then all agents default, and the total deadweight loss from default is*

$$\begin{aligned} &\min\left\{\beta\left[\alpha(\tau)\bar{d}+\bar{\epsilon}-n(\bar{e}-\bar{s}-L(\tau))\right], \bar{e}+(n-1)(\bar{e}-\bar{s}-L(\tau))\right\} \\ &+\sum_{m=1}^{n-1} \min\left\{\beta\left(\alpha(\tau)\bar{d}-m(\bar{e}-\bar{s}-L(\tau))\right), \bar{e}+(m-1)(\bar{e}-\bar{s}-L(\tau))\right\}, \end{aligned} \tag{10}$$

*which is greater than* (8) *for a small enough* $\beta$.

All proofs are relegated to the appendix.

First, consider the case in which either the shock size relative to the amount of liquidity is small or the relative size of the total liabilities is small. For the complete network, the deadweight loss amount (7) comes from only the shocked agent, as all other agents can pay each other in full by dispersing the liquidity shock among themselves equally. In contrast, there can be other defaulting agents in the ring network, resulting in additional deadweight losses as in the second term of (9). This is because the shock size can be still greater than what a single agent can bear with its own cash buffer.

Suppose, without loss of generality, that agent 1 is the agent under liquidity shock. Agent 1 receives the payment in full as $\alpha(\tau)\bar{d}$ but still defaults and only pays whatever is left over

[15] $\lfloor\cdot\rfloor$ denotes the floor function, i.e. the greatest integer that is less than or equal to $\cdot$.

after paying out the liquidity shock, i.e., the amount of $\phi$. Agent 2 has its cash buffer net of liquidity cost $\bar{e} - \bar{s} - L(\tau)$ plus any payment from agent 1, $\phi$; however that might not be enough for its payment, so agent 2 pays agent 3 the amount of $\bar{e} - \bar{s} - L(\tau) + \phi$. Similarly, agent 3's cash buffer net of liquidity cost plus the payment from agent 2, $2(\bar{e} - \bar{s} - L(\tau)) + \phi$, still may not be enough to pay in full. Therefore, defaults will continue until agent $\psi + 1$, who would have enough cash buffer $\psi(\bar{e} - \bar{s} - L(\tau)) + \phi$ to pay its obligation $\alpha(\tau)\bar{d}$ in full.

Now suppose that both the shock size relative to the amount of liquidity and the relative size of the total liabilities are large. All agents default for both the ring and complete networks. However, the two networks experience different levels of changes in deadweight losses. The difference in deadweight loss for the complete network when all agents default, the difference between (8) and (7), primarily stems from the second term in (8), which is

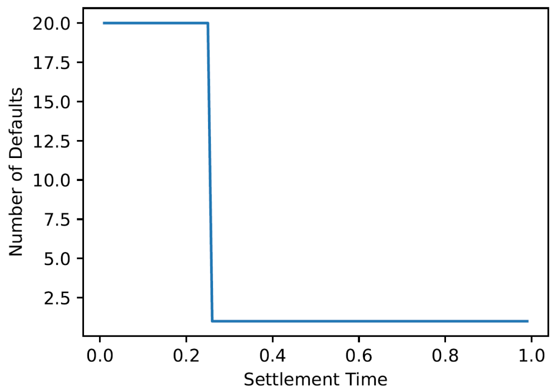
$$(n-1)\min\left\{\beta\left[\alpha(\tau)\bar{d} - (n-1)(\bar{e} - L(\tau) - \bar{s})\right], \bar{e} + (n-2)(\bar{e} - \bar{s} - L(\tau))\right\}.$$

For the ring network, the change in deadweight loss between the two cases, represented by the difference between (10) and (9), is relatively smaller as

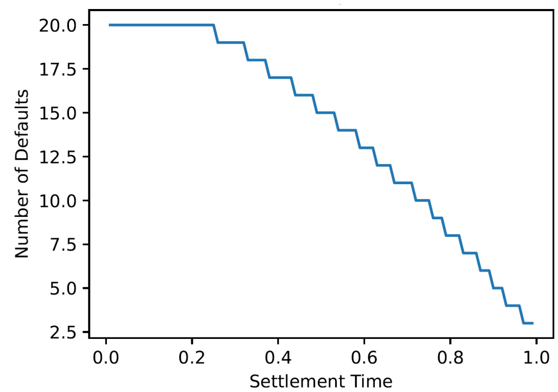
$$\sum_{m=\psi+1}^{n-1}\min\left\{\beta\left(\alpha(\tau)\bar{d} - m(\bar{e} - \bar{s} - L(\tau))\right), \bar{e} + (m-1)(\bar{e} - \bar{s} - L(\tau))\right\},$$

where $\phi = 0$ in the case of 2.(b) in Proposition 1. This is mainly because there are already additional $\psi$ number of agents (other than the shocked agent) who are defaulting in the ring network even in case 2.(a), so only the additional $n - \psi - 1$ agents are defaulting, whereas the additional number of defaulting agents is $n - 1$ in the complete network. Figure 3 illustrates the results in Proposition 1.

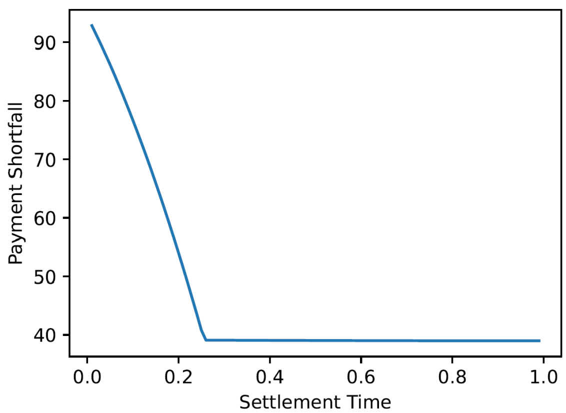
Note that full contagion happens only if both the shock size relative to the amount of liquidity *and* the relative size of the total liabilities are large. This property of full contagion requiring both conditions to be simultaneously satisfied is very similar to that of Chang and Chuan (2024), in which full contagion occurs only if the liquidity shock size is large *and* the



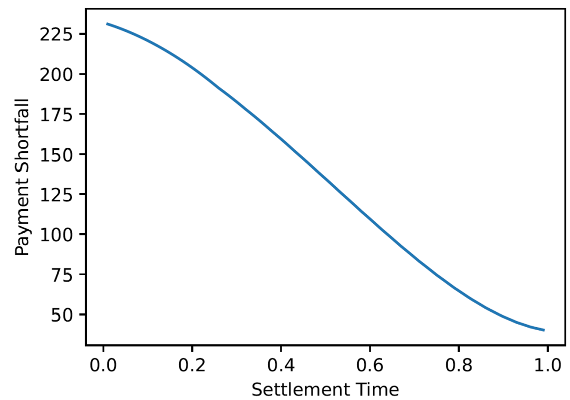

(a) Number of defaults: Complete network




(b) Number of defaults: Ring network




(c) Aggregate payment shortfall: Complete network




(d) Aggregate payment shortfall: Ring network

Figure 3: Number of defaults and aggregate payment shortfall: Complete vs Ring

Note: This figure illustrates the results in Proposition 1. The top panels have the number of defaults on the y-axis and the settlement time $\tau$ on the x-axis. The top-left panel illustrates the phase transition property of contagion in the complete network, as the number of defaults jumps at the threshold $\tau$ value. In contrast, the top-right panel illustrates the gradual increase in the number of defaults for the ring network as $\tau$ decreases. The bottom panels have the aggregate payment shortfall on the y-axis and the settlement time $\tau$ on the x-axis. The bottom-left panel illustrates the phase transition property in the complete network. The bottom-right panel illustrates smooth changes in the payment shortfall in the ring network as $\tau$ decreases.

collateral amount is small. Hence, our results align with the general findings in the financial network literature—contagion is rare, as multiple conditions have to hold, and monitoring systemic risk due to contagion requires *collecting multi-dimensional information.*

Finally, we discuss the two key functions, the netting function $\alpha(\tau)$ and the liquidity cost function $L(\tau)$, and how they interact with each other.

First, the liquidity cost function $L(\tau)$ plays a role in determining the threshold liquidity shock size $\epsilon^*(\tau) = n(\bar{e} - \bar{s} - L(\tau))$. Hence, as $\tau$ increases, $\epsilon^*(\tau)$ increases; i.e., the given liquidity shock $\bar{\epsilon}$ is less likely to exceed the threshold, due to the decrease in $L(\tau)$. In other words, agents will have a relatively larger cash buffer if they are paying less liquidity cost when the settlement is slower. Similarly, $L(\tau)$ also plays a role in increasing the threshold total liability amount $d^*(\tau) = (n-1)(\bar{e} - \bar{s} - L(\tau))$ in a similar fashion. As the liquidity cost decreases with slower settlement, the given total liability amount is less likely to exceed the threshold.

The netting function $\alpha(\tau)$ plays a role in determining only the total liability threshold. As $\tau$ increases, the effective netting increases, i.e., $\alpha(\tau)$ decreases, effectively increasing the required total liability amount. This is because the inequality

$$\bar{d} \leq \frac{d^*(\tau)}{\alpha(\tau)}$$

becomes more likely to hold as $\alpha(\tau)$ decreases for a given $\bar{d}$.

Both $L(\tau)$ and $\alpha(\tau)$ move in the same direction, making cases 1.(a) and 2.(a) more likely as $\tau$ increases. We can rearrange the threshold conditions and combine the two as presented in the following corollary.

**Corollary 1.** *All agents default in the complete network if and only if $\tau > \tau^*$, where $\tau^*$ is such that*

$$L(\tau^*) \equiv \bar{e} - \bar{s} - \min\left\{\frac{\bar{\epsilon}}{n}, \alpha(\tau^*)\frac{\bar{d}}{n-1}\right\} \tag{11}$$

*holds.*

We refer to $\tau^*$ as the threshold point (for the complete network) from now on. Corollary 1 implies that the threshold point is decreasing in the cash buffer net of senior debt $\bar{e} - \bar{s}$, as the required liquidity cost amount for the threshold condition increases with it. For example, if there is no net cash buffer at all ($\bar{e} - \bar{s} = 0$), then full contagion occurs in the complete network even if there is no liquidity cost ($L(\tau) = 0$), regardless of other values. Moreover, the threshold point, $\tau^*$, decreases, as the shock size divided by the number of agents decreases. In other words, if the shock can be absorbed by the aggregate net cash buffer of all agents, $n(\bar{e} - \bar{s}) > \bar{\epsilon}$, then some degree of liquidity cost $L(\tau)$ can still be absorbed by the system, thus there is no full contagion with a lower settlement time.

Finally, Corollary 1 implies that there is an interaction between $L(\tau)$ and $\alpha(\tau)$ when determining the threshold condition. As the degree of netting increases (i.e., $\alpha(\tau)$ decreases), the network can withstand a larger value of liquidity cost $L(\tau)$, allowing for a faster settlement speed (lower $\tau^*$) without full contagion. However, if the degree of netting is small ($\alpha(\tau)$ is relatively large) and the total liabilities for each agent are large ($\bar{d}$ is large), then the interaction between $\alpha(\tau)$ and $L(\tau)$ becomes irrelevant. This is because the size of the liquidity shock becomes the minimum of the last term in the right-hand side of (11). In other words, in this case, the relative size of the total liabilities is already large, so the size of the liquidity shock is the marginal factor that determines the value of the liquidity cost that the network can bear while absorbing the liquidity shocks and preventing full contagion.

## 3.2. Heterogeneous Ex Ante Social Welfare Implications

Proposition 1 shows the phase transition in ex post contagion patterns. Now we utilize the phase transition property to evaluate ex ante social welfare implications of different networks and settlement times. The next proposition builds on Proposition 1 and shows that decreasing the settlement time $\tau$ has different effects across different network structures. In particular, we show that faster settlement time can improve ex ante social welfare of the ring network but worsen ex ante social welfare of the complete network.

For expositional simplicity, we focus on the case in which the upper bound on each agent's deadweight loss is not binding, i.e., $\beta$ and $\xi_i$ are small enough so that the defaulting agent's assets $A_i$ are not completely wiped out due to bankruptcy losses. While this assumption is mainly for tractability, it is also consistent with reality. Legal costs of bankruptcy or costs of delays in allocating remaining assets of a defaulting agent are unlikely to be as large as the total assets themselves. Nevertheless, our model can be easily extended to incorporate cases with $\beta\xi_i > A_i$ in a numerical model.

A marginal change in the settlement time $\tau$ affects the ex ante welfare loss in three different ways. It affects:

(a) The aggregate liquidity cost $nL(\tau)$;

(b) The likelihood of the liquidity shock and resulting deadweight losses due to counterparty defaults by changing the shock arrival probability $F(\tau)$; and

(c) The total deadweight losses $\sum_{j \in N} \beta\xi_j(\tau, D, \epsilon)$.

Therefore, for a given $\tau$ and network $D$, the marginal change in the ex ante welfare loss is:

$$\frac{\partial W(\tau, D)}{\partial \tau} = nL'(\tau) + F(\tau) \sum_{j \in N} \beta \frac{\partial \xi_j(\tau, D, \epsilon)}{\partial \tau} + F'(\tau) \sum_{j \in N} \beta\xi_j(\tau, D, \epsilon). \tag{12}$$

The first term of (12) represents the decrease in liquidity cost for all $n$ agents. The second term represents the marginal change of the deadweight losses multiplied by the default probability. The third term represents an increase in the default probability multiplied by the deadweight losses due to default contagion. We later show that the second term is always negative in Lemma 1.

The following proposition demonstrates the heterogeneous changes in ex ante welfare loss when $\tau$ changes.

**Proposition 2.** *For a fixed $\tau \in (0, \bar{T})$, and the complete network $D_c$ and the ring network $D_r$, the following statements hold:*

1. *Suppose that $L(\tau) < L^*(\tau)$. The comparison between the marginal change in the ex ante welfare loss for the complete network and that for the ring network is ambiguous, i.e., both $\frac{\partial W(\tau, D_c)}{\partial \tau} > \frac{\partial W(\tau, D_r)}{\partial \tau}$ and $\frac{\partial W(\tau, D_c)}{\partial \tau} \leq \frac{\partial W(\tau, D_r)}{\partial \tau}$ are possible depending on the parameter values and functional forms.*

2. *Suppose that $L(\tau) \geq L^*(\tau)$. Then, the marginal change in the ex ante welfare loss for the complete network is less than that for the ring network, i.e., $\frac{\partial W(\tau, D_c)}{\partial \tau} < \frac{\partial W(\tau, D_r)}{\partial \tau}$. Moreover, an increase in $\tau$ can make the complete network better off, while making the ring network worse off, i.e.,*

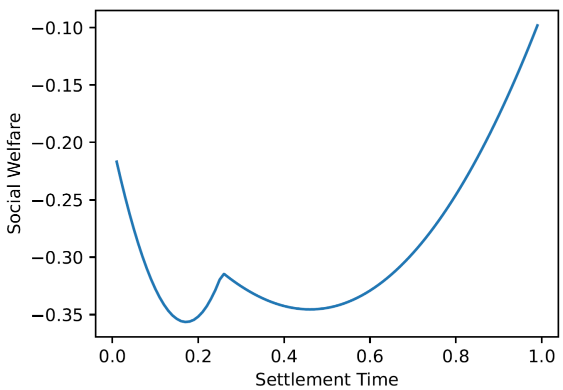
$$\frac{\partial W(\tau, D_c)}{\partial \tau} < 0 < \frac{\partial W(\tau, D_r)}{\partial \tau}. \tag{13}$$

Proposition 2 implies that there can be cases, where faster settlement speed can make the ring network better off while making the complete network worse off. In particular, such concrete ex ante social welfare comparison is possible once the liquidity cost exceeds the phase transition threshold $L^*(\tau)$. If the liquidity cost is below the threshold, or equivalently $\tau$ is large enough, then the comparison depends on all parameter values and functional forms, as we have many moving parts. However, this intricacy surprisingly disappears once the liquidity cost reaches its threshold, or equivalently $\tau$ is small enough. Thus, our result highlights the critical role of sudden jumps in the number of defaulting agents when considering the ex ante social welfare of a given network and settlement speed $\tau$.

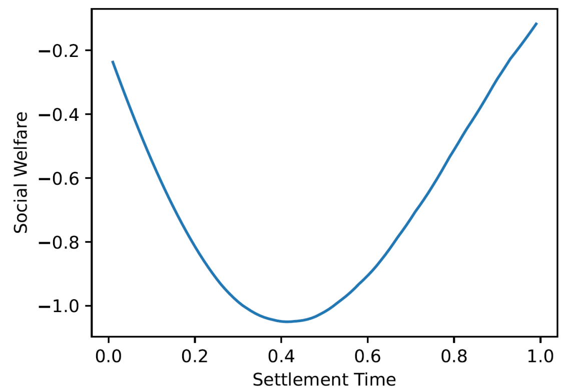
Figure 4 shows that a decrease in settlement time can either increase or decrease the ex ante social welfare depending on the current settlement time $\tau$. Although the global maximum of the ex ante social welfare is attained at the maximum $\tau$ value in both cases under this specific parameter space of simulations, it's important to note that the global maximum depends on the functional form and parameter values of $\alpha(\tau)$, $L(\tau)$, $F(\tau)$, and others, as stated in the first part of Proposition 2.

### 3.3. General Networks and Ex Ante Social Welfare

We continue to explore the effect of jumps in the number of defaulting agents on marginal ex ante social welfare in a general setup. In this subsection, we allow for the case with $\gamma > 0$, i.e., deadweight losses from defaults further reduce the payments from defaulting agents. For




(a) Ex ante Social Welfare: Complete network




(b) Ex ante Social Welfare: Ring network

Figure 4: Ex ante Social Welfare: Complete vs Ring

Note: This figure illustrates the results in Proposition 2. Both panels have the ex ante social welfare on the y-axis and the settlement time $\tau$ on the x-axis. The left panel shows the changes in ex ante social welfare for the complete network. The ex ante social welfare shows a kink, a discrete jump in the slope, at the threshold value. The right panel shows the changes in the ex ante social welfare for the ring network that changes smoothly. Thus, if $\tau$ is just below the threshold value, the marginal change in the ex ante social welfare is positive for the complete network, while it is negative for the ring network.

simplicity, we focus on the case of $\gamma = 1$; however, our results can be easily extended to any $0 < \gamma < 1$. Similarly, we focus on cases in which $\epsilon_j < e_j$ for any $j \in N$ and any realization of $\epsilon$, for expositional simplicity.[16]

### 3.3.1. Aggregate Deadweight Loss and Welfare Decomposition

Consider any general regular network $D$. Denote the set of defaulting agents and solvent agents as $\mathcal{D}(\tau, \epsilon)$ and $\mathcal{S}(\tau, \epsilon)$, respectively, for a given settlement time $\tau$ and liquidity shock vector $\epsilon$. First, we show that the deadweight loss from payment shortfall of any $j \in N$,

$$\beta\xi_j(\tau, D, \epsilon) = \beta \left[ \alpha(\tau)\bar{d} + \bar{s} + \bar{\epsilon}_j - \bar{e} + L(\tau) - \sum_{k \in N} q_{jk} x_k(\tau, D, \epsilon) \right]^+,$$

is decreasing in $\tau$. In other words, for a fixed probability of shock arrival, $F(\tau)$, the ex ante welfare loss always decreases as $\tau$ increases.

[16] This assumption is similar to the assumption on liquidity shock in Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015).

**Lemma 1.** *The aggregate deadweight loss conditional on the arrival of a liquidity shock decreases as $\tau$ increases. Furthermore, the number of defaulting agents weakly decreases as $\tau$ increases, i.e., $\mathcal{D}(\tau', \epsilon) \subseteq \mathcal{D}(\tau, \epsilon)$ for any $\tau < \tau'$, $D$, and $\epsilon$.*

The intuition is that the payment shortfall for each agent decreases as its liability and liquidity cost decrease through decreases in $\alpha(\tau)$ and $L(\tau)$. Even though the payments from others also decrease in $\sum_{k \in N} q_{jk} x_k$ if $x_k = \alpha(\tau)\bar{d}$, the sum of $q_{jk}$ over $j \in N$ cannot exceed 1, which is the coefficient for the liability of agent $j$ itself. In other words, the decrease in liabilities to be paid by $j$ exceeds the decrease in payments from solvent counterparties.

Lemma 1 clarifies the tradeoff associated with changes in $\tau$. As $\tau$ increases, the aggregate deadweight loss always decreases, either through a smaller amount of payment shortfalls or a reduced number of defaulting agents. However, an increase in the likelihood of shock arrival counteracts that effect. The following result summarizes the decomposition of net welfare effect of an increase in $\tau$:

**Proposition 3.** *For a given settlement time $\tau$, network structure $D$, and the liquidity shock state $\epsilon$, suppose that $\mathcal{D}(\tau, \epsilon) = \mathcal{D}(\tau + \delta, \epsilon)$ for a small enough $\delta > 0$. Then, the effect of an increase in the settlement lag on the ex ante welfare loss depends on netting benefits, liquidity costs, and counterparty risks, as shown in the below decomposition:*

$$\underbrace{\underbrace{nL'(\tau)}_{\substack{\text{decrease in}\\ \text{liquidity cost}}} + \underbrace{F(\tau)}_{\substack{\text{probability of}\\ \text{shock arrival}}} \underbrace{\sum_{j \in \mathcal{D}(\tau,\epsilon)} \beta \frac{\partial \xi_j(\tau, D, \epsilon)}{\partial \tau}}_{\substack{\text{decrease in aggregate}\\ \text{deadweight loss}}}}_{\text{Benefit}} + \underbrace{\underbrace{F'(\tau)}_{\substack{\text{increased likelihood}\\ \text{of shock arrival}}} \underbrace{\sum_{j \in \mathcal{D}(\tau,\epsilon)} \beta \xi_j(\tau, D, \epsilon)}_{\substack{\text{aggregate}\\ \text{deadweight loss}}}}_{\text{Cost}}. \tag{14}$$

The proposition provides us several insights. For example, if a default does not incur deadweight losses, i.e., $\beta = 0$, the aggregate welfare loss is decreasing in $\tau$ due to decreases in liquidity costs $nL'(\tau)$. Similarly, if the payment shortfall, $\xi_j(\tau, D, \epsilon)$, is very small, the ex ante welfare loss given $\epsilon$ decreases as $\tau$ increases. The welfare implication becomes ambiguous only if the conditional payment shortfall is large, and the increase in the likelihood of shock arrival $F'(\tau)$ is significant.

### 3.3.2. Default Threshold Points, Centrality, and Ex Ante Welfare Loss

Now, we can finally analyze the changes in the ex ante welfare loss by utilizing the monotone effect of the settlement time $\tau$ from Lemma 1 and the decomposition of net welfare effect from Proposition 3. In particular, we show that default threshold points, the settlement times at which one or more agents are on the verge of default, play an important role in comparing ex ante welfare losses across two different settlement times. Moreover, we show that the node depth centrality, which represents the systemic importance of an agent, is key for measuring the effect of payment shortfalls. The formal definitions and statements are below.

First, we compute the total amount of deadweight loss for a given shock vector $\epsilon$.

**Node depth centrality.** Recall that the difference between the total liability and the actual payment of a defaulting agent $j$ is

$$\alpha(\tau)\bar{d} - x_j(\tau, D, \epsilon) = (1+\beta)\xi_j(\tau, D, \epsilon), \tag{15}$$

i.e., the fundamental payment shortfall

$$\xi_j(\tau, D, \epsilon) \equiv \left[\alpha(\tau)\bar{d} + \bar{s} + \epsilon_j + L(\tau) - \bar{e} - \sum_{k\in N} q_{jk} x_k(\tau, D, \epsilon)\right]^+,$$

multiplied by $(1+\beta)$ due to the deadweight loss. We can use (15) to represent the aggregate deadweight losses, similar to Capponi, Corell, and Stiglitz (2022), as:

$$\beta \sum_{j\in\mathcal{D}(\tau,\epsilon)} \xi_j(\tau, D, \epsilon) \left[1 + (1+\beta) \sum_{k\in\mathcal{D}(\tau,\epsilon)} q_{kj} + (1+\beta)^2 \sum_{k\in\mathcal{D}(\tau,\epsilon)} \sum_{l\in\mathcal{D}(\tau,\epsilon)} q_{lk} q_{kj} \right.$$
$$\left. + (1+\beta)^3 \sum_{k\in\mathcal{D}(\tau,\epsilon)} \sum_{l\in\mathcal{D}(\tau,\epsilon)} \sum_{m\in\mathcal{D}(\tau,\epsilon)} q_{ml} q_{lk} q_{kj} + \cdots \right]. \tag{16}$$

The expression in (16) captures the following:

1. The initial deadweight loss for a defaulting agent $j$: $\beta\xi_j(\tau, D, \epsilon)$.

2. The direct first-order effect of payment shortfalls and the deadweight loss on agent $j$'s counterparties: $\beta\xi_j(\tau, D, \epsilon)(1+\beta)\sum_{k\in\mathcal{D}(\tau,\epsilon)} q_{kj}$.

3. The second-order effect of these counterparties affecting their own counterparties: $\beta\xi_j(\tau, D, \epsilon)(1+\beta)^2\sum_{k\in\mathcal{D}(\tau,\epsilon)}\sum_{l\in\mathcal{D}(\tau,\epsilon)} q_{lk}q_{kj}$.

4. All the higher-order effects from iterating this process infinitely.

Denote $Q_{\mathcal{D}(\tau,\epsilon)}$ as a submatrix of $Q$ for the subset of $\mathcal{D}(\tau, \epsilon)$. If the spectral radius of $(1 + \beta)Q_{\mathcal{D}(\tau,\epsilon)}$ is less than one, then we can represent the sum as a Neumann series

$$\beta \sum_{j\in\mathcal{D}(\tau,\epsilon)} \xi_j(\tau, D, \epsilon) \underbrace{\left(1 + (1+\beta)\sum_{k\in\mathcal{D}(\tau,\epsilon)} q_{kj} + (1+\beta)^2 \sum_{k\in\mathcal{D}(\tau,\epsilon)}\sum_{l\in\mathcal{D}(\tau,\epsilon)} q_{lk}q_{kj} + \cdots\right)}_{C_j(\mathcal{D}(\tau,\epsilon))}, \quad (17)$$

where $C_j(\mathcal{D}(\tau, \epsilon))$ is the *node depth centrality* among $\mathcal{D}(\tau, \epsilon)$ as defined by Glasserman and Young (2015). Note that $C_j(\mathcal{D}(\tau, \epsilon))$ only depends on the set of defaulting agents.[17]

We assume that the spectral radius assumption, the spectral radius of $(1+\beta)Q_{\mathcal{D}(\tau,\epsilon)}$ is less than one, holds for the following analysis, as it allows us to derive crisp analytical results, which bring useful insights. In particular, the node depth centrality is useful to derive a closed-form solution for the defaulting amount in matrix notation. However, the insights can be generalized to cases in which the spectral radius assumption does not hold.

As it can be seen from Lemma A.2 in the Appendix, the aggregate deadweight loss can be represented in terms of the node depth centrality. The (node depth) centrality $C_j(\mathcal{D}(\tau, \epsilon))$

[17] The technical condition for the spectral radius being less than one also has a clear economic interpretation. For the node depth centrality to be well defined, if there are cycles in the subnetwork $\mathcal{D}(\tau, \epsilon)$, the entries of the relative liability matrix $Q_{\mathcal{D}(\tau,\epsilon)}$ have to be sufficiently small. In other words, if a payment within the subnetwork can be amplified through cycles, then it must be sufficiently small, because otherwise there would be too much amplification of welfare losses within the set of defaulting agents and the infinite sum (16) does not converge, resulting in a complete depletion of one or more agents' assets. For example, the ring network has a spectral radius greater than one, and its node depth centrality could not be computed using matrix inversion.

can be interpreted as a "default contagion multiplier." This multiplier captures the amplification effect of defaults on the aggregate deadweight losses throughout the network, as $C_j(\mathcal{D}(\tau, \epsilon))$ shows how the initial impact of a default by agent $j$ is magnified across the network of defaulting agents. The magnitude of $C_j(\mathcal{D}(\tau, \epsilon))$ represents the importance of $j$, in terms of amplification of deadweight losses, in the given network structure $D$ for a given settlement time $\tau$. Therefore, the multiplier $C_j(\mathcal{D}(\tau, \epsilon))$ quantifies the systemic importance of agent $j$ in terms of its payment shortfall-induced deadweight losses.

Finally, we consider comparative statics of decreasing $\tau$ to $\tau'$ on the ex ante welfare loss. There can be two different cases following this effect: with or without the changes in the set of defaulting agents.

**Case with no additional defaults.** First, suppose that $\mathcal{D}(\tau', \epsilon) = \mathcal{D}(\tau, \epsilon)$, i.e., the set of defaulting agents under the same liquidity shock state $\epsilon$ would be the same as that under $\tau$. The centrality of an agent $C_j(\mathcal{D}(\tau, \epsilon))$ would remain the same, but the payment shortfall amount will change due to the following factors:

(a) The initial decline in $x_j(\tau, D, \epsilon)$ due to $L(\tau)$ increasing to $L(\tau')$, resulting from higher liquidity cost to facilitate quicker trades;

(b) The decrease in inter-agent payments $x_j(\tau, D, \epsilon)$ to $x_j(\tau', D, \epsilon)$ that amplifies the initial increase in cost of $L(\tau') - L(\tau)$; and

(c) The increase from $\alpha(\tau)\bar{d}$ to $\alpha(\tau')\bar{d}$ due to a lesser degree of netting of liabilities.

**Case with additional defaults.** Second, suppose that $\mathcal{D}(\tau, \epsilon) \subset \mathcal{D}(\tau', \epsilon)$, i.e., there are additional agents who default under this new regime under the same $\epsilon$. Then, there will be an additional jump in deadweight losses due to reaching this default threshold for previously solvent agents. This impacts the deadweight losses in two ways:

1. There will be more agents included in the summation of $\sum_{j \in \mathcal{D}(\tau', \epsilon)}$ compared to $\sum_{j \in \mathcal{D}(\tau, \epsilon)}$.

2. The multiplier for the payment shortfalls, $C_j(\mathcal{D}(\tau,\epsilon))$, will jump up to $C_j(\mathcal{D}(\tau',\epsilon))$ as the summation of the weighting matrix $\sum_{k\in\mathcal{D}(\tau',\epsilon)} Q_{kj}$ increases for each $j \in \mathcal{D}(\tau',\epsilon)$.

Consequently, even if the payment shortfall for each agent $\alpha(\tau)\bar{d} - x_j(\tau, D, \epsilon)$ is continuously decreasing in $\tau$, there will be discontinuous effects from the change in the set of defaulting agents $\mathcal{D}(\tau',\epsilon)$. This effect increases as the number of additional defaults increases.

After considering both changes, we also discount the deadweight losses by the lower likelihood of liquidity shocks, as $F(\tau') < F(\tau)$. Hence, the tension between the two factors—an increase in deadweight losses versus a decrease in the likelihood of liquidity shocks—will determine whether faster payments (lower $\tau'$) increase or decrease social welfare.

**Default threshold points.** Combining the three factors, we argue that focusing on the *default threshold points*—the values of $\tau$ where a marginal decrease in $\tau$ would increase the number of defaulting agents—is critical in evaluating comparative statics with ex ante welfare losses. Formally, we define the set of default threshold points for a given network $D$ and liquidity shock $\epsilon$ as $\mathcal{T}(D,\epsilon)$, where $\forall\tilde{\tau} \in \mathcal{T}(D,\epsilon)$, $\mathcal{D}(\tilde{\tau},\epsilon) \subset \mathcal{D}(\tilde{\tau}-\delta,\epsilon)$ and $\mathcal{D}(\tilde{\tau}-\delta,\epsilon) \nsubseteq \mathcal{D}(\tilde{\tau},\epsilon)$, while $\mathcal{D}(\tilde{\tau}+\delta,\epsilon) = \mathcal{D}(\tilde{\tau},\epsilon)$, for any small $\delta > 0$. Define the union of the set of default threshold points for every possible realization of $\epsilon$ (i.e., $\forall\epsilon$ with positive probability) as the default threshold points $\mathcal{T}(D)$. We can define the set of settlement times such that the number of defaulting agents remains constant over the set up to the default threshold points. In other words, for any $\tilde{\tau} \in \mathcal{T}(D)$, we define the set of constant defaults as $\underline{\mathcal{T}}(\tilde{\tau}, D)$ such that $\mathcal{D}(\tau,\epsilon) = \mathcal{D}(\tilde{\tau},\epsilon)$ if and only if $\tau \in \underline{\mathcal{T}}(\tilde{\tau}, D)$ for every possible $\epsilon$.

**Theorem 1.** *If the settlement time $\tau$ is a default threshold point, i.e., $\tau \in \mathcal{T}(D)$ for a given network $D$, the left-hand derivative of the ex ante welfare loss is strictly less than the right-hand derivative of the ex ante welfare loss at $\tau$, i.e., $\frac{\partial W(\tau,D)}{\partial\tau^-} < \frac{\partial W(\tau,D)}{\partial\tau^+}$, due to additional defaults. Furthermore, there can be cases in which an increase in $\tau$ improves welfare exactly at a default threshold point.*

Theorem 1 implies that evaluating the changes in ex ante social welfare should account for both the continuous changes in the likelihood of shocks and the resulting aggregate

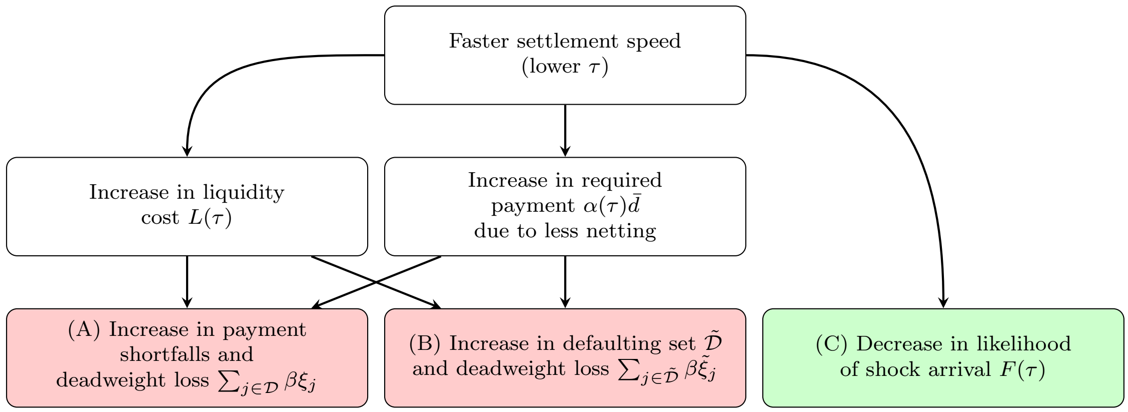



Figure 5: Impact of a faster settlement speed on financial stability

Note: This figure visualizes the effect of a faster settlement speed (lower $\tau$) on financial stability through three different channels—the cost channels, (A) and (B), and the benefit channel (C). Faster settlement increases liquidity cost and required payment due to less netting, while decreasing the likelihood of shock arrival, i.e., the benefit channel (C). However, the first two effects lead to an increase in payment shortfalls and deadweight loss, i.e., the cost channel (A). Moreover, if $\tau$ is a default threshold point, the first two effects can lead to an increase in the number of defaults, leading to a further increase in deadweight loss, i.e., the cost channel (B). Therefore, whether $\tau$ is a default threshold point or not is important, as it determines the importance of the cost channel (B), highlighted by Theorem 1.

deadweight loss and the *discrete jump* in the aggregate deadweight loss due to changes in the number of defaulting agents at the default threshold points in $\mathcal{T}(D)$. An extreme case of the effect of a default threshold point is the phase transition of the complete network shown in Proposition 2. There is only one threshold point for the complete network due to its perfect symmetry. Figure 5 visualizes the tradeoff and the role of default threshold points highlighted by Theorem 1.

Note that if $\tau$ is not a default threshold point, the effect of faster settlement speed is ambiguous. A marginal change in ex ante welfare loss $W(\tau, D)$ with respect to $\tau$ can be positive or negative, depending on the functional forms ($\alpha(\tau)$, $L(\tau)$, and $F(\tau)$), the parameters $(\beta, \bar{e}, \bar{s}, G(\epsilon))$, and the centrality of each agent under each realization of $\epsilon$ ($C_j(\mathcal{D}(\tau, \epsilon))$). In other words, the tradeoff depends on many moving parts—netting benefits, liquidity costs, counterparty risks, and network structure—as characterized by Proposition 3.

Figure 4 illustrates a great example of Theorem 1 and the tradeoff of faster settlement speed. First, it shows that the settlement time does not have a monotone effect on the ex ante social welfare. When settlement time $\tau$ is close to zero, an increase in $\tau$ leads to a decrease in welfare. The ex ante social welfare starts to increase around $\tau \approx 0.18$. This is because of the interaction between the cost and benefit of a marginal change in $\tau$, summarized in Proposition 3. The cost of increased likelihood of shock arrival dominates the benefits of decreases in liquidity cost and aggregate deadweight loss, resulting in a U-shaped graph. However, the ex ante social welfare forms another larger U-shaped graph once $\tau$ reaches the default threshold point at which the set of defaulting agents becomes just one shocked agent. Therefore, the ex ante social welfare has a stark kink at the threshold point, where the left-hand derivative and the right-hand derivative have different signs. This is due to the sudden shift in the set of defaulting agents from all agents to only one agent.

### 3.3.3. Identification of Default Threshold Points

Given the importance of default threshold points on the evaluation of ex ante social welfare with respect to settlement speed, finding the default threshold points of a given network is crucial. However, if identification of default threshold points can be computationally complex, our results, especially Theorem 1, would be difficult to utilize in practice. Hence, we characterize the tractability of default threshold point identification. In particular, we show that identification of default threshold points is solvable in polynomial time (i.e., a problem in complexity class P).

First, we introduce an finite discretization of a continuous distribution $G(\epsilon)$. This is, in practice, the only feasible way to simulate and compute continuous distributions. Formally, a finite discrete approximation of $G(\epsilon)$ is $\tilde{G}(\epsilon)$, which is supported on a finite set $\left\{\epsilon^{(1)}, \ldots, \epsilon^{(M)}\right\}$. This approximation can be done by several methods including quantization (divide the support into bins and take representative values) and Monte Carlo simulation. For any desired level of precision, we can approximate the shock distribution $G(\epsilon)$ arbitrarily close to that level of precision by adjusting the finite number $M$.

**Proposition 4.** *Identification of the set of default threshold points is solvable in polynomial* *time. Moreover, each node's default indicator is monotone in $\tau$, so the number of threshold points is at most $n$ per shock realization $\epsilon$.*

Proposition 4 implies that the seemingly complex problem of identifying default threshold points of a network is computationally tractable in polynomial time. In addition to simply checking which points are the default threshold points in $[0, \bar{T}]$, we can also see how many additional defaults are occurring at each default threshold point. Thus, we can identify whether a network would have a sudden jump in deadweight loss, as in the case of the complete network, or not. Therefore, the key insight of considering default threshold points for the ex ante welfare calculations, highlighted by Theorem 1, can be utilized in practical application to solving for the optimal settlement speed for financial stability, or evaluating the effect of faster settlement speed on financial stability.

# 4. Empirical Implications

The model suggests that faster settlement can reduce the likelihood of stress events but can increase the severity of stress events when they happen. This could be tested by comparing the likelihood and severity of systemic events across markets or time periods with different settlement speeds, controlling for other factors.

The model's insights on how network structure affects contagion patterns could be empirically tested. For instance, researchers could analyze payment system data to identify whether more interconnected networks show a greater degree of contagion during stress times.

Although we do not model government intervention explicitly, the role of the government or central bank is crucial in maintaining stable payment systems, and our results provide insights on this as well. An implicit benefit of a slower settlement speed (larger $\tau$) is that the government has less incentive to intervene in the ex-post event of a shock arrival. Conversely, if settlement speed is fast and a shock arrives, despite its lower likelihood, the government has a much stronger incentive to intervene, as the ex-post cost of the systemic event is much larger. This implies that faster settlement systems might paradoxically lead to more frequent government interventions. Researchers could investigate this by comparing the frequency and scale of government interventions in markets with different settlement speeds.

Lemma A.2 provides a justification for ex-post liquidity injection to central agents when a shock arrives and causes a systemic event. This implies that policymakers and central banks should focus on identifying and supporting the most systemically important institutions during stress times. Historical examples of various emergency facilities for central entities in payment systems, such as the Primary Dealer Credit Facility, are well justified by our results and by various previous papers in the literature on bailouts.

Our model's focus on ex ante welfare, incorporating both the likelihood and severity of shocks, suggests a new framework for empirical assessments of payment system stability. Further studies could compare traditional ex post measures with ex ante measures that account for the probability of shock arrivals.

On a related note, our results suggest that optimal settlement speed may vary depending on network structure and other parameters. This implies that a one-size-fits-all approach to settlement speed regulation may be suboptimal. Empirical studies could identify the parameter values and analyze whether a faster settlement speed can improve or worsen the financial stability of a payment system.

These empirical implications suggest a set of empirically testable hypotheses and new approaches to analyzing payment systems and financial networks. They also highlight the importance of considering network structure, settlement speed, and systemic risk in an integrated framework when designing and evaluating financial market infrastructures.

# 5. Conclusion

This paper provides a comprehensive analysis of the implications of settlement speed in payment systems, taking into account the complex interactions between network structure, netting efficiency, liquidity costs, and counterparty risks. We highlight the potential trade-offs and identify key factors that influence the effect of a change in settlement speed. Our results contribute to the ongoing debate about optimal settlement speeds in financial markets and offer several key insights in the design and regulation of payment systems.

First, we show that faster settlements do not always lead to improved financial stability. While quicker settlement can reduce the likelihood of counterparty defaults, it can also increase liquidity costs and reduce netting opportunities. Our model shows that the optimal settlement speed depends critically on the structure of the payment network, the distribution of liquidity shocks, and the rates of changes in liquidity costs and netting efficiency.

Second, our results highlight the importance of network structure in determining the systemic risk implications of settlement speed. We show that different network topologies, such as the ring and complete networks, can respond very differently to changes in settlement speed. Our results highlight the need for policymakers to consider the specific structure of payment networks when designing settlement systems or implementing changes to settlement speeds.

Third, we identify the existence of default threshold points—settlement times at which the number of defaulting agents in the network changes discontinuously. These threshold points play a crucial role in determining the ex-ante welfare implications of changes in settlement speed. Our results show that the welfare effects of changing settlement speed can differ dramatically depending on whether the system is operating at or above these threshold points. Furthermore, we show that identification of the default threshold points of a given network can be solved in polynomial time, thus, our results can be utilized in practical applications.

Fourth, we show that the node depth centrality is a key measure of an agent's systemic importance in the context of payment systems, consistent with the literature on general financial networks. This measure captures how the initial impact of a default by one agent is amplified through the network of defaulting agents, providing a useful tool for identifying systemically important institutions in payment networks.

Our findings have several important policy implications. First of all, a one-size-fits-all approach to settlement speed may be suboptimal, given the heterogeneity in network structures and other relevant parameters across different financial systems. Instead, policymakers should carefully consider the underlying characteristics of each payment system when determining optimal settlement speeds.

Moreover, our results provide a theoretical underpinning for targeted liquidity support to central entities of payment systems during times of stress. This aligns with and provides a formal basis for various emergency facilities implemented by central banks during financial crises.

Relatedly, our results imply that faster settlements can decrease the likelihood of crisis events but increase the severity of the crises. Thus, faster settlement systems might paradoxically lead to more frequent government interventions due to stronger incentives to intervene when a crisis is more severe. This insight should be further accounted for when deciding settlement speeds in payment systems.

While our model provides significant insights, it also opens up several avenues for future research. Extensions could include incorporating strategic behavior by agents, analyzing more general network structures, such as core-periphery networks, or considering multi-period dynamics. In addition, empirical work testing the predictions of our model across different payment systems and market structures would be valuable to expand our understanding and provide concrete policy implications.

## References


Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015): "Systemic risk and stability in financial networks," *American Economic Review*, 105(2), 564–608.

Acharya, V. V., and O. Merrouche (2013): "Precautionary Hoarding of Liquidity and Inter-Bank Markets: Evidence from the Sub-prime Crisis," *Review of Finance*, 17(1), 107–160.

Afonso, G., and H. S. Shin (2011): "Precautionary demand and liquidity in payment systems," *Journal of Money, Credit and Banking*, 43, 589–619.

Allen, F., and D. Gale (2000): "Financial contagion," *Journal of Political Economy*, 108(1), 1–33.

Andolfatto, D. (2020): *Hot Money for a Cold Economy*. Federal Reserve Bank of St. Louis, Research Division.

Berman, A., and R. J. Plemmons (1979): *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, Cambridge, MA.

Bernard, B., A. Capponi, and J. E. Stiglitz (2022): "Bail-ins and bailouts: Incentives, connectivity, and systemic stability," *Journal of Political Economy*, 130(7), 1805–1859.

Brunnermeier, M. K., and L. H. Pedersen (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22(6), 2201–2238.

Capponi, A., F. Corell, and J. E. Stiglitz (2022): "Optimal bailouts and the doom loop with a financial network," *Journal of Monetary Economics*, 128, 35–50.

Capponi, A., J. J. Wang, and H. Zhang (2022): "A Theory of Collateral Requirements for Central Counterparties," *Management Science*, 68(9), 6993–7017.

Carapella, F., and D. Mills (2011): "Information insensitive securities: the true benefits of central counterparties," Discussion paper, mimeo, Board of Governors.

Chang, J.-W. (2021): "Collateralized debt networks with lender default," *Working paper*, p. SSRN. https://ssrn.com/abstract=3468267.

Chang, J.-W., and G. Chuan (2024): "Contagion in debt and collateral markets," *Journal of Monetary Economics*, 148, 103600.

——— (2025): "Collateral Reuse and Financial Stability," *FEDS Working Paper*, 2025(35).

Chang, J.-W., E. Klee, and V. Yankov (2025): "Rewiring repo," *FEDS Working Paper*, 2025(013).

Copeland, A., and R. Garratt (2019): "Nonlinear pricing and the market for settling payments," *Journal of Money, Credit and Banking*, 51(1), 195–226.

Csóka, P., and P. J.-J. Herings (2021): "An axiomatization of the proportional rule in financial networks," *Management Science*, 67(5), 2799–2812.

Ding, D., R. Gonzalez, Y. Ma, and Y. Zeng (2025): "The Effect of Instant Payments on the Banking System: Liquidity Transformation and Risk-Taking," *Working paper*.

Donaldson, J. R., G. Piacentino, and X. Yu (2022): "Systemic Risk in Financial Networks Revisited: The Role of Maturity," *Working paper*.

Duffie, D., M. Scheicher, and G. Vuillemey (2015): "Central Clearing and Collateral Demand," *Journal of Financial Economics*, 116(2), 237–256.

Duffie, D., and H. Zhu (2011): "Does a Central Clearing Counterparty Reduce Counterparty Risk?," *Review of Asset Pricing Studies*, 1(1), 74–95.

D'Errico, M., and T. Roukny (2021): "Compressing over-the-counter markets," *Operations Research*, 69(6), 1660–1679.

Eisenberg, L., and T. H. Noe (2001): "Systemic risk in financial systems," *Management Science*, 47(2), 236–249.

Elliott, M., B. Golub, and M. O. Jackson (2014): "Financial networks and contagion," *American Economic Review*, 104(10), 3115–3153.

Glasserman, P., and H. P. Young (2015): "How likely is contagion in financial networks?," *Journal of Banking & Finance*, 50, 383–399.

Hancock, D., and J. A. Wilcox (1996): "Intraday management of bank reserves: the effects of caps and fees on daylight overdrafts," *Journal of Money, Credit and Banking*, 28(4), 870–908.

Jackson, M. O., and A. Pernoud (2024): "Credit freezes, equilibrium multiplicity, and optimal bailouts in financial networks," *The Review of Financial Studies*, p. hhad096.

Kahn, C. M., J. McAndrews, and W. Roberds (2003): "Settlement risk under gross and net settlement," *Journal of Money, Credit and Banking*, pp. 591–608.

Kahn, C. M., and W. Roberds (1998): "Payment system settlement and bank incentives," *The review of financial studies*, 11(4), 845–870.

Khapko, M., and M. Zoican (2020): "How fast should trades settle?," *Management Science*, 66(10), 4573–4593.

Kotidis, A., and S. Schreft (forthcoming): "The propagation of cyberattacks through the financial system: Evidence from an actual event," *Journal of Finance*.

McAndrews, J., and S. Rajan (2000): "The timing and funding of Fedwire funds transfers," *Economic Policy Review*, 6(2).

Mills Jr, D. C., and T. D. Nesmith (2008): "Risk and concentration in payment and securities settlement systems," *Journal of Monetary Economics*, 55(3), 542–553.

Rochet, J.-C., and J. Tirole (1996): "Controlling risk in payment systems," *Journal of Money, Credit, and Banking*, 28(4), 832.

Rogers, L. C., and L. A. Veraart (2013): "Failure and rescue in an interbank network," *Management Science*, 59(4), 882–898.

VanHoose, D., and G. H. Sellon Jr (1989): "Daylight overdrafts, payments system risk, and public policy," *Economic Review*, 74(5), 9–29.

Veraart, L. A. M. (2022): "When does portfolio compression reduce systemic risk?," *Mathematical Finance*, 32(3), 727–778.

# Appendix (for online publication only)

## A. Omitted Proofs

### A.1. Proof of Proposition 1

**Proof.** Without loss of generality, assume that agent 1 receives the liquidity shock.

1. First, consider the complete network.

   (a) **Case 1.1.** Suppose that the relative size of the liquidity shock is small as $\bar{\epsilon} \leq n(\bar{e} - \bar{s} - L(\tau))$. First consider the case in which the relative size of the total liabilities is small as $\alpha(\tau)\bar{d} \leq \bar{\epsilon} - (\bar{e} - \bar{s} - L(\tau))$. Hence, even with all the payments received, agent 1 does not have any excess cash to cover the liquidity shock and the senior debt, i.e. agent 1 fully defaults on all the liabilities towards other agents in the network.

   Then, for any agent $j \geq 2$ to survive, the following should hold:

$$\bar{e} - \bar{s} - L(\tau) + \frac{n-2}{n-1}\alpha(\tau)\bar{d} \geq \alpha(\tau)\bar{d}. \tag{18}$$

   Rearranging (18) yields

$$(n-1)(\bar{e} - \bar{s} - L(\tau)) \geq \alpha(\tau)\bar{d}.$$

   This holds because

$$\alpha(\tau)\bar{d} \leq \bar{\epsilon} - (\bar{e} - \bar{s} - L(\tau)) \leq (n-1)(\bar{e} - \bar{s} - L(\tau)),$$

   where the first inequality comes from the relative size of the total liabilities, and the second inequality comes from the relatively small size of the liquidity shock.

Therefore, (18) holds, so only agent 1 defaults.

Since all other agents are paying in full, $A_1 = \bar{e} + \alpha(\tau)\bar{d}$. Then,

$$\xi_1 = \alpha(\tau)\bar{d} + \bar{s} + \epsilon - \bar{e} + L(\tau) - \alpha(\tau)\bar{d}$$
$$= \bar{s} + \epsilon - \bar{e} + L(\tau),$$

and the total deadweight loss from default in this case is

$$\min\left\{\beta\left(\bar{s} + \epsilon - \bar{e} + L(\tau)\right), \bar{e} + \alpha(\tau)\bar{d}\right\}.$$

**Case 1.2.** Now consider the case in which the relative size of the total liabilities is large as $\alpha(\tau)\bar{d} > \bar{\epsilon} - (\bar{e} - \bar{s} - L(\tau))$. Then, all the payments received from agents 2 through $n$ will exceed the cash amount required to pay for the liquidity shock for agent 1. Then, all agents survive if

$$n(\bar{e} - \bar{s} - L(\tau)) \geq \bar{\epsilon},$$

which is immediately satisfied by the small shock assumption. Therefore, only agent 1 defaults in this case as well.

Since agent 1 is receiving all the liabilities from other agents in full, $A_1 = \bar{e} + \alpha(\tau)\bar{d}$. Therefore,

$$\xi_1 = \alpha(\tau)\bar{d} + \bar{s} + \bar{\epsilon} - \bar{e} + L(\tau) - \alpha(\tau)\bar{d}$$
$$= \bar{s} + \bar{\epsilon} - \bar{e} + L(\tau).$$

Hence, the total deadweight loss from default is

$$\min\left\{\beta\left(\bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)\right), \bar{e} + \alpha(\tau)\bar{d}\right\}.$$

**Case 1.3.** Next, suppose that the relative size of the liquidity shock is large as $\bar{\epsilon} > n(\bar{e} - \bar{s} - L(\tau))$ but the relative size of the total liabilities is small as $\alpha(\tau)\bar{d} \leq (n-1)(\bar{e} - \bar{s} - L(\tau))$. Because the liquidity shock exceeds the aggregate cash in the network, agent 1 will be fully defaulting on its payments to other agents. Therefore, (18) should hold for the remaining agents to survive. Rearranging (18) implies

$$(n-1)(\bar{e} - \bar{s} - L(\tau)) \geq \alpha(\tau)\bar{d},$$

which is immediately satisfied by the assumption on the relative size of the total liabilities. Therefore, no other agents default in this case.

Since all other agents are paying in full, $A_1 = \bar{e} + \alpha(\tau)\bar{d}$. Then,

$$\begin{aligned}\xi_1 &= \alpha(\tau)\bar{d} + \bar{s} + \epsilon - \bar{e} + L(\tau) - \alpha(\tau)\bar{d} \\ &= \bar{s} + \epsilon - \bar{e} + L(\tau),\end{aligned}$$

and the total deadweight loss from default in this case is

$$\min\left\{\beta\left(\bar{s} + \epsilon - \bar{e} + L(\tau)\right), \bar{e} + \alpha(\tau)\bar{d}\right\}.$$

(b) Suppose that the relative size of the liquidity shock is large as $\bar{\epsilon} > n(\bar{e} - \bar{s} - L(\tau))$, and the relative size of the total liabilities is large as $\alpha(\tau)\bar{d} > (n-1)(\bar{e} - \bar{s} - L(\tau))$. Again, due to the size of the liquidity shock, agent 1 fully defaults on its payments to other agents, and the no default condition for other agents is (18). Again, (18) is equivalent to

$$(n-1)(\bar{e} - \bar{s} - L(\tau)) \geq \alpha(\tau)\bar{d},$$

which contradicts the condition for the relative size of the total liabilities. Therefore, all agents default in this case.

Finally, we compute the total deadweight loss amount from defaults in this case. First, note that $x_1 = 0$, as the liquidity shock is larger than the total amount of cash in the network, i.e. there will be no remaining asset value even if all other agents pays their full cash amount to agent 1. Second, due to symmetry, all other agents will pay the same amount, i.e. $x_2 = x_3 = \cdots = x_n = \bar{x}$. Plugging this into (4) implies

$$\bar{x} = \bar{e} - \bar{s} - L(\tau) + \frac{(n-2)\bar{x}}{n-1}.$$

Thus, we derive

$$\bar{x} = (n-1)(\bar{e} - \bar{s} - L(\tau)).$$

Then, plugging this into the definition $A_j$ yields

$$\begin{aligned} A_j &= \bar{e} + \frac{n-2}{n-1}(n-1)(\bar{e} - \bar{s} - L(\tau)) \\ \Rightarrow A_j &= \bar{e} + (n-2)(\bar{e} - \bar{s} - L(\tau)) \end{aligned} \tag{19}$$

for any $j > 1$. Thus, the total payment shortfall is

$$\begin{aligned} \xi_j &= \alpha(\tau)\bar{d} + \bar{s} + L(\tau) - A_j \\ \Rightarrow \xi_j &= \alpha(\tau)\bar{d} + \bar{s} - (n-1)(\bar{e} - L(\tau)) + (n-2)\bar{s} \\ \Rightarrow \xi_j &= \alpha(\tau)\bar{d} - (n-1)(\bar{e} - L(\tau) - \bar{s}) \end{aligned} \tag{20}$$

for each $j > 2$ after plugging in (19). By construction, $A_1 = \bar{e} + (n-1)(\bar{e} - \bar{s} - L(\tau))$, and

$$\xi_1 = \alpha(\tau)\bar{d} + \bar{s} + L(\tau) + \bar{\epsilon} - A_1$$
$$= \alpha(\tau)\bar{d} + \bar{\epsilon} - n(\bar{e} - \bar{s} - L(\tau)).$$

Thus, the total deadweight loss from defaults is

$$\min\left\{\beta\left[\alpha(\tau)\bar{d} + \bar{\epsilon} - n(\bar{e} - \bar{s} - L(\tau))\right], \bar{e} + (n-1)(\bar{e} - \bar{s} - L(\tau))\right\}$$
$$+(n-1)\min\left\{\beta\left[\alpha(\tau)\bar{d} - (n-1)(\bar{e} - L(\tau) - \bar{s})\right], \bar{e} + (n-2)(\bar{e} - \bar{s} - L(\tau))\right\}.$$

2. Now consider the ring network.

   (a) **Case 2.1.** Suppose that the relative size of the liquidity shock is small as $\bar{\epsilon} \leq n(\bar{e} - \bar{s} - L(\tau))$. Hence, if all agents combine their available liquidity buffers, they can either pay the promised payment in full even without any payment from agent 1, when the total liabilities for each agent is small, or can reuse the payments received by agent 1 and exceed the liquidity shock amount by combined payments, when the total liabilities for each agent is large. In either case, agent 1 will receive the full payment $\alpha(\tau)\bar{d}$ from agent $n$. Therefore, agent 1 pays

$$x_1 = \phi \equiv \left[\bar{e} - \bar{s} - L(\tau) + \alpha(\tau)\bar{d} - \bar{\epsilon}\right]^+$$

   to agent 2. Then, agent 2 can reuse this payment amount on top of its own cash buffer $\bar{e} - \bar{s} - L(\tau)$ to pay agent 3, and agent 3 can reuse the whole payment on top of its own cash buffer $\bar{e} - \bar{s} - L(\tau)$, and so on. In other words, by the same iterative steps,

$$x_2 = \bar{e} - \bar{s} - L(\tau) + \phi$$
$$x_3 = 2(\bar{e} - \bar{s} - L(\tau)) + \phi$$
$$x_4 = 3(\bar{e} - \bar{s} - L(\tau)) + \phi$$
$$\vdots$$
$$x_k = (k-1)(\bar{e} - \bar{s} - L(\tau)) + \phi.$$

By the initial assumption on the shock size, there exists an integer $k$ such that

$$k(\bar{e} - \bar{s} - L(\tau)) + \phi \geq \alpha(\tau)\bar{d},$$

and rearranging the condition implies the smallest $k$ is equivalent to

$$\psi = \min\left\{n-1, \left\lfloor \frac{\alpha(\tau)\bar{d} - \phi}{\bar{e} - \bar{s} - L(\tau)} \right\rfloor\right\},$$

where $\lfloor \cdot \rfloor$ denotes the floor function, i.e., the greatest integer that is less than or equal to the argument $\cdot$. Now we calculate the total deadweight loss from the defaults of $\psi$ agents and agent 1. First, note that agent 1 receives the full payment amount, so $A_1 = \bar{e} + \alpha(\tau)\bar{d}$, and $\xi_1 = \bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)$. Thus, the deadweight loss from agent 1's default is

$$\min\left\{\beta(\bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)), \bar{e} + \alpha(\tau)\bar{d}\right\}.$$

Second, note that agent 2 receives $\phi$ from agent 1, so $A_2 = \bar{e} + \phi$. Thus, $\xi_2 = \alpha(\tau)\bar{d} + \bar{s} - \bar{e} + L(\tau) - \phi$. From this, we can also compute $A_3 = \bar{e} + x_2 =$

$2\bar{e}-\bar{s}-L(\tau)+\phi$, and $\xi_3=\alpha(\tau)\bar{d}-2(\bar{e}-\bar{s}-L(\tau))-\phi$. Similarly, for any $m\leq\psi$,

$$A_{m+1}=\bar{e}+(m-1)(\bar{e}-\bar{s}-L(\tau))+\phi$$
$$\xi_{m+1}=\alpha(\tau)\bar{d}-m(\bar{e}-\bar{s}-L(\tau))-\phi.$$

Therefore, the total deadweight loss from defaults is

$$\min\left\{\beta(\bar{s}+\bar{\epsilon}-\bar{e}+L(\tau)),\bar{e}+\alpha(\tau)\bar{d}\right\}$$
$$+\sum_{m=1}^{\psi}\min\left\{\beta\left(\alpha(\tau)\bar{d}-m(\bar{e}-\bar{s}-L(\tau))-\phi\right),\bar{e}+(m-1)(\bar{e}-\bar{s}-L(\tau))+\phi\right\},$$

**Case 2.2.** Finally, suppose that the relative size of the liquidity shock is large as $\bar{\epsilon}>n(\bar{e}-\bar{s}-L(\tau))$, but the relative size of the total liabilities is small as $\alpha(\tau)\bar{d}\leq(n-1)(\bar{e}-\bar{s}-L(\tau))$. Then, agent $n$ can still pay its full liabilities to agent 1, and the rest of the payment structure would be the same. Therefore, all the previous arguments go through and the total deadweight loss from defaults is the same.

(b) Suppose that the relative size of the liquidity shock is large as $\bar{\epsilon}>n(\bar{e}-\bar{s}-L(\tau))$, and the relative size of the total liabilities is also large as $\alpha(\tau)\bar{d}>(n-1)(\bar{e}-\bar{s}-L(\tau))$.

Then, all agents default, and agent 1 cannot even pay the liquidity shock and senior debt, as shown in the complete network case (Statement 1). Hence, $\phi=0$, and even agent $n$ cannot pay its payment obligation in full as

$$(n-1)(\bar{e}-\bar{s}-L(\tau))<\alpha(\tau)\bar{d}.$$

Thus, all agents default in this case.

Finally, we compute the total deadweight loss in this case. As agents 2 through

$n$ will accumulate all their available liquidity and pay agent 1, $x_n = (n-1)(\bar{e} - \bar{s} - L(\tau))$, implying

$$\begin{aligned} A_1 &= \bar{e} + (n-1)(\bar{e} - \bar{s} - L(\tau)) \\ \xi_1 &= \alpha(\tau)\bar{d} + \bar{\epsilon} - n(\bar{e} - \bar{s} - L(\tau)). \end{aligned}$$

Thus, the deadweight loss from agent 1's default is

$$\min\left\{\beta\left[\alpha(\tau)\bar{d} + \bar{\epsilon} - n(\bar{e} - \bar{s} - L(\tau))\right], \bar{e} + (n-1)(\bar{e} - \bar{s} - L(\tau))\right\}.$$

Since agent 2 receives nothing from agent 1,

$$\begin{aligned} A_2 &= \bar{e} \\ \xi_2 &= \alpha(\tau)\bar{d} - (\bar{e} - \bar{s} - L(\tau)). \end{aligned}$$

Similarly, agent 3 receives $x_2 = \bar{e} - \bar{s} - L(\tau)$, so

$$\begin{aligned} A_3 &= \bar{e} + (\bar{e} - \bar{s} - L(\tau)) \\ \xi_2 &= \alpha(\tau)\bar{d} - 2(\bar{e} - \bar{s} - L(\tau)). \end{aligned}$$

Iteratively, we obtain

$$\begin{aligned} A_{m+1} &= \bar{e} + (m-1)(\bar{e} - \bar{s} - L(\tau)) \\ \xi_{m+1} &= \alpha(\tau)\bar{d} - m(\bar{e} - \bar{s} - L(\tau)). \end{aligned}$$

Therefore, the total deadweight loss from default is

$$\min\left\{\beta\left[\alpha(\tau)\bar{d}+\bar{\epsilon}-n(\bar{e}-\bar{s}-L(\tau))\right],\bar{e}+(n-1)(\bar{e}-\bar{s}-L(\tau))\right\}$$
$$+\sum_{m=1}^{n-1}\min\left\{\beta\left(\alpha(\tau)\bar{d}-m(\bar{e}-\bar{s}-L(\tau))\right),\bar{e}+(m-1)(\bar{e}-\bar{s}-L(\tau))\right\},$$

which is greater than (8) for a small enough $\beta$, because $m \leq n-1$.

■

### A.2. Proof of Corollary 1

**Proof.** Recall that $\epsilon^*(\tau) = n(\bar{e} - \bar{s} - L(\tau))$ and $d^*(\tau) = (n-1)(\bar{e} - \bar{s} - L(\tau))$ and the threshold conditions are

$$\bar{\epsilon} > n(\bar{e} - \bar{s} - L(\tau))$$
$$\alpha(\tau)\bar{d} > (n-1)(\bar{e} - \bar{s} - L(\tau)).$$

Thus, by rearranging both threshold conditions we obtain

$$L(\tau) > \bar{e} - \bar{s} - \frac{\bar{\epsilon}}{n}$$
$$L(\tau) > \bar{e} - \bar{s} - \alpha(\tau)\frac{\bar{d}}{n-1},$$

where the first inequality comes from rearranging the threshold condition for $\bar{\epsilon}$ and the second inequality comes from the rearranged threshold condition for $\bar{d}$. Since both conditions should be simultaneously satisfied, we can combine the two conditions as the following

$$L(\tau) > \max\left\{\bar{e} - \bar{s} - \frac{\bar{\epsilon}}{n}, \bar{e} - \bar{s} - \alpha(\tau)\frac{\bar{d}}{n-1}\right\}$$
$$\Rightarrow L(\tau) > \bar{e} - \bar{s} - \min\left\{\frac{\bar{\epsilon}}{n}, \alpha(\tau)\frac{\bar{d}}{n-1}\right\}. \tag{21}$$

Note that $\tau^*$ such that $L(\tau^*) \equiv \bar{e} - \bar{s} - \min\left\{\frac{\bar{\epsilon}}{n}, \alpha(\tau^*)\frac{\bar{d}}{n-1}\right\}$ holds is unique, because $L(\tau)$ is strictly decreasing in $\tau$, while $\alpha(\tau)$ is decreasing in $\tau$. Therefore, for any $\tau > \tau^*$, the combined threshold condition, (21), is satisfied, and vice versa. ■

### A.3. Proof of Proposition 2

**Proof.** First, note that the identity of the agent hit by a shock is irrelevant due to the symmetry of the network structure for both the complete and the ring networks. Suppose that $\tau$ is large enough so that $L(\tau) < L^*(\tau)$. If $D$ is the complete network, then the marginal change in the ex ante welfare loss is

$$\frac{\partial W(\tau, D_c)}{\partial \tau} = nL'(\tau) + F'(\tau)\beta\left(\bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)\right) + F(\tau)\beta L'(\tau). \tag{22}$$

If $D$ is the ring network, then the marginal change in the ex ante welfare loss is

$$\begin{aligned}\frac{\partial W(\tau, D_r)}{\partial \tau} =& nL'(\tau) + F'(\tau)\Bigg[\beta\left(\bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)\right) \\ &+ \sum_{m=1}^{\psi}\beta\left(\alpha(\tau)\bar{d} - m(\bar{e} - \bar{s} - L(\tau))\right)\Bigg] \\ &+ F(\tau)\left[\beta L'(\tau) + \sum_{m=1}^{\psi}\beta\left(\alpha'(\tau)\bar{d} + mL'(\tau)\right)\right],\end{aligned} \tag{23}$$

if $\phi \equiv \left[\bar{e} - \bar{s} - L(\tau) + \alpha(\tau)\bar{d} - \bar{\epsilon}\right]^+ = 0$, and

$$\begin{aligned}\frac{\partial W(\tau, D_r)}{\partial \tau} =& nL'(\tau) + F'(\tau)\Bigg[\beta\left(\bar{s} + \bar{\epsilon} - \bar{e} + L(\tau)\right) \\ &+ \sum_{m=1}^{\psi}\beta\left(\bar{\epsilon} - (m+1)(\bar{e} - \bar{s} - L(\tau)) - \phi\right)\Bigg] \\ &+ F(\tau)\left[\beta L'(\tau) + \sum_{m=1}^{\psi}\beta\left((m+1)L'(\tau) - \alpha'(\tau)\bar{d}\right)\right],\end{aligned} \tag{24}$$

if $\phi > 0$, where $\psi \equiv \min\left\{n-1, \left\lfloor\frac{\alpha(\tau)\bar{d} - \phi}{\bar{e} - \bar{s} - L(\tau)}\right\rfloor\right\}$.

The comparison between the marginal change in the ex ante welfare loss for the complete network and that for the ring network is ambiguous. For example, (23) subtracted by (22) implies

$$F'(\tau)\sum_{m=1}^{\psi}\beta\left[\alpha(\tau)\bar{d}-m\left(\bar{e}-\bar{s}-L(\tau)\right)\right]+F(\tau)\sum_{m=1}^{\psi}\beta\left[\alpha'(\tau)\bar{d}+mL'(\tau)\right],$$

which can be positive or negative depending on the parameter values $\beta, \bar{e}, \bar{s}, \bar{d}$ as well as the functional forms of $\alpha(\tau)$, $F(\tau)$, and $L(\tau)$, because $F'(\tau) < 0$ and $F(\tau) > 0$.

Now suppose that $\tau$ is small enough so that $L(\tau) \geq L^*(\tau)$. We first show that every agent in the ring network defaults even if $L(\tau) = L^*(\tau)$.

**Case 1.** Suppose that $\frac{\bar{\epsilon}}{n} \leq \alpha(\tau)\frac{\bar{d}}{n-1}$. Then, $n(\bar{e}-\bar{s}-L(\tau) = \bar{\epsilon}$ and $(n-1)(\bar{e}-\bar{s}-L(\tau)) \leq \alpha(\tau)\bar{d}$. Therefore, $\phi = 0$ and $\psi = n-1$.

**Case 2.** Suppose $\frac{\bar{\epsilon}}{n} > \alpha(\tau)\frac{\bar{d}}{n-1}$. Then, $(n-1)(\bar{e}-\bar{s}-L(\tau)) = \alpha(\tau)\bar{d}$ and $n(\bar{e}-\bar{s}-L(\tau)) < \epsilon$. Therefore, $\phi = 0$ and $\psi = n-1$.

Therefore, even at the verge of the threshold condition, the ring network already has full contagion.

Recall that the aggregate deadweight loss in the complete network is

$$\begin{aligned}&\beta\left[\alpha(\tau)\bar{d}+\bar{\epsilon}-n(\bar{e}-\bar{s}-L(\tau))\right]+(n-1)\beta\left[\alpha(\tau)\bar{d}-(n-1)(\bar{e}-L(\tau)-\bar{s})\right]\\ =&\beta\left[n\alpha(\tau)\bar{d}+\bar{\epsilon}-\left(n+(n-1)^2\right)(\bar{e}-\bar{s}-L(\tau))\right].\end{aligned}$$

The marginal change in ex ante welfare loss with respect to $\tau$ is

$$\begin{aligned}\frac{\partial W(\tau, D_c)}{\partial \tau} =&nL'(\tau)+F'(\tau)\beta\left[n\alpha(\tau)\bar{d}+\bar{\epsilon}-\left(n+(n-1)^2\right)(\bar{e}-\bar{s}-L(\tau))\right]\\ &+F(\tau)\beta\left[n\bar{d}\alpha'(\tau)+\left(n+(n-1)^2\right)L'(\tau)\right].\end{aligned} \tag{25}$$

Now recall that the aggregate deadweight loss in the ring network is

$$\begin{aligned}
&\beta\left[\alpha(\tau)\bar{d}+\bar{\epsilon}-n(\bar{e}-\bar{s}-L(\tau))+\sum_{m=1}^{n-1}\left(\alpha(\tau)\bar{d}-m(\bar{e}-\bar{s}-L(\tau))\right)\right]\\
=&\beta\left[n\alpha(\tau)\bar{d}+\bar{\epsilon}-n(\bar{e}-\bar{s}-L(\tau))-\sum_{m=1}^{n-1}m(\bar{e}-\bar{s}-L(\tau))\right]\\
=&\beta\left[n\alpha(\tau)\bar{d}+\bar{\epsilon}-n(\bar{e}-\bar{s}-L(\tau))-\frac{n(n-1)}{2}(\bar{e}-\bar{s}-L(\tau))\right]\\
=&\beta\left[n\alpha(\tau)\bar{d}+\bar{\epsilon}-\frac{n(n+1)}{2}(\bar{e}-\bar{s}-L(\tau))\right],
\end{aligned}$$

Therefore, the marginal change in the ex ante welfare loss with respect to $\tau$ is

$$\begin{aligned}
\frac{\partial W(\tau,D_r)}{\partial\tau}=&nL'(\tau)+F'(\tau)\beta\left[n\alpha(\tau)\bar{d}+\bar{\epsilon}-\frac{n(n+1)}{2}(\bar{e}-\bar{s}-L(\tau))\right]\\
&+F(\tau)\beta\left[n\bar{d}\alpha'(\tau)+\frac{n(n+1)}{2}L'(\tau)\right],
\end{aligned}\tag{26}$$

because $\phi\equiv\left[\bar{e}-\bar{s}-L(\tau)+\alpha(\tau)\bar{d}-\bar{\epsilon}\right]^{+}=0$ by the assumption.

Subtracting (25) from (26) implies

$$\begin{aligned}
\frac{\partial W(\tau,D_r)}{\partial\tau}-\frac{\partial W(\tau,D_c)}{\partial\tau}=&F'(\tau)\beta\left[\left(n+(n-1)^2\right)-\frac{n(n+1)}{2}\right](\bar{e}-\bar{s}-L(\tau))\\
&-F(\tau)\beta\left[\left(n+(n-1)^2\right)-\frac{n(n+1)}{2}\right]L'(\tau).
\end{aligned}\tag{27}$$

Since $(\bar{e}-\bar{s}-L(\tau))$ is positive and $-L'(\tau)$ is positive, (27) is positive, because

$$\left[\left(n+(n-1)^2\right)-\frac{n(n+1)}{2}\right]>0.\tag{28}$$

The previous inequality (28) holds because

$$n + (n-1)^2 > \frac{n(n+1)}{2}$$
$$2n^2 - 2n + 2 > n^2 + n$$
$$n^2 + 2 > 3n,$$

which holds for any $n > 2$. Therefore, $\frac{\partial W(\tau, D_c)}{\partial \tau} < \frac{\partial W(\tau, D_r)}{\partial \tau}$ holds.

Moreover, since there is a gap between $\frac{\partial SW(\tau, D_c)}{\partial \tau}$ and $\frac{\partial SW(\tau, D_r)}{\partial \tau}$, the following inequalities

$$\begin{aligned}
&nL'(\tau) + F'(\tau)\beta\left[n\alpha(\tau)\bar{d} + \bar{\epsilon} - \frac{n(n+1)}{2}(\bar{e} - \bar{s} - L(\tau))\right] \\
&+ F(\tau)\beta\left[n\bar{d}\alpha'(\tau) + \frac{n(n+1)}{2}L'(\tau)\right] > 0 \\
>&nL'(\tau) + F'(\tau)\beta\left[n\alpha(\tau)\bar{d} + \bar{\epsilon} - \left(n + (n-1)^2\right)(\bar{e} - \bar{s} - L(\tau))\right] \\
&+ F(\tau)\beta\left[n\bar{d}\alpha'(\tau) + \left(n + (n-1)^2\right)L'(\tau)\right]
\end{aligned} \tag{29}$$

hold, depending on the parameters and functional forms. Thus, for the same $\tau$ increase, the complete network can be better off while the ring network can be worse off. ■

### A.4. Proof of Lemma 1

**Proof.** Recall that the payment vector is

$$x = \left[\min\left\{\alpha(\tau)\bar{d}\mathbf{1}, Qx + \bar{e}\mathbf{1} - \bar{s}\mathbf{1} - L(\tau)\mathbf{1} - \epsilon - \beta\xi\right\}\right]^+,$$

where the payment shortfall is

$$\xi(\tau) = \left[\alpha(\tau)\bar{d}\mathbf{1} + \bar{s}\mathbf{1} + \epsilon - \bar{e}\mathbf{1} + L(\tau)\mathbf{1} - Qx\right]^+.$$

Suppose that $\tau < \tau'$. We compare the payment shortfalls and the sets of defaulting agents for these two different settlement times.

**Case 1.** If $x_j = \alpha(\tau)\bar{d}$, then $\xi_j = 0$, and $x_j$ remains $\alpha(\tau)\bar{d}$ as $\tau$ increases, and thus, $\xi_j$ remains to be zero.

**Case 2.** Suppose that the agent under liquidity shock defaults on its senior liabilities with $\tau$ but not with $\tau'$. If $x_j < \alpha(\tau)\bar{d}$ and $x_j = \alpha(\tau')\bar{d}$ due to the decrease in liabilities through more netting by $\alpha(\tau')$, then $\xi_j$ decreases to zero under $\tau'$.

**Case 3.** Now consider the last case such that $x_j < \alpha(\tau')\bar{d}$. As noted previously, $\mathcal{S}(\tau) \subseteq \mathcal{S}(\tau')$ and $\mathcal{D}(\tau') \subseteq \mathcal{D}(\tau)$, as agents who are solvent in $\tau$ remains to be solvent in $\tau'$.

**Case 3.1.** Suppose that the agent hit by liquidity shock $j$ does not default on its senior liabilities with $\tau'$. We can permutate the network into a matrix such that the payment equilibrium vector is $x = \begin{pmatrix} x_{\mathcal{D}} \\ \alpha(\tau')\bar{d}\mathbf{1} \end{pmatrix}$, where $x_{\mathcal{D}}$ is payments by agents in the defaulting set $\mathcal{D}(\tau')$. The payments by agents in the defaulting set is

$$x_d = \alpha(\tau')\bar{d}Q_{ds}\mathbf{1} + Q_{dd}x_d\mathbf{1} + \left[\left(\bar{e} - \bar{s} - L(\tau')\right)\mathbf{1} - \epsilon_d - \beta\xi_d\right],$$

where $Q_{ds}$ is the liability weight matrix from $\mathcal{S}(\tau')$ to $\mathcal{D}(\tau')$, $Q_{dd}$ is the liability weight matrix within $\mathcal{D}(\tau')$, and $\epsilon_d$ and $\xi_d$ are subvectors of $\epsilon$ for $\mathcal{D}(\tau)$ and $\xi$ for $\mathcal{D}(\tau')$, respectively. Because

$Q$ is a stochastic matrix, $Q_{ds} + Q_{dd} = 1$, so the payment vector becomes

$$\begin{aligned}
x_d &= (I - Q_{dd})^{-1} \left[\alpha(\tau')\bar{d}Q_{ds}\mathbf{1} + (\bar{e} - \bar{s} - L(\tau'))\,\mathbf{1} - \epsilon_d - \beta\xi_d\right] \\
\Rightarrow x_d &= (I - Q_{dd})^{-1} \left[\alpha(\tau')\bar{d}\,(I - Q_{dd})\,\mathbf{1} + (\bar{e} - \bar{s} - L(\tau'))\,\mathbf{1} - \epsilon_d - \beta\xi_d\right] \\
\Rightarrow x_d &= \alpha(\tau')\bar{d}\mathbf{1} + (I - Q_{dd})^{-1} \left[(\bar{e} - \bar{s} - L(\tau'))\,\mathbf{1} - \epsilon_d - \beta\xi_d\right] < \alpha(\tau')\bar{d}\mathbf{1}, \qquad (30)
\end{aligned}$$

where $I$ denotes the identity matrix of the appropriate dimension, and the last inequality of (30) holds, because agents in $\mathcal{D}(\tau')$ default.

Now consider the case with $\tau$. We show that even when there are no additional defaults, i.e., $\mathcal{D}(\tau') = \mathcal{D}(\tau)$, the aggregate deadweight loss is greater than that under $\tau'$, resulting in even greater payment shortfall due to an increase in $\beta\xi_d$.

If any agent in $\mathcal{D}(\tau)$ can no longer meet their senior liabilities, then we have an increase in deadweight losses. Suppose the contrary and agents can still meet their senior liabilities. By $L(\tau) > L(\tau')$, we have

$$\begin{aligned}
&(I - Q_{dd})^{-1} \left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \epsilon_d - \beta\xi_d\right] \\
&< (I - Q_{dd})^{-1} \left[(\bar{e} - \bar{s} - L(\tau'))\,\mathbf{1} - \epsilon_d - \beta\xi_d\right] < 0,
\end{aligned}$$

where the last inequality holds by (30). Then, it is immediate that $x_d(\tau) < \alpha(\tau)\bar{d}\mathbf{1}$. In addition, the payment shortfall increases as

$$\begin{aligned}
\alpha(\tau)\bar{d}\mathbf{1} - x_d &= (I - Q_{dd})^{-1} \left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \epsilon_d - \beta\xi_d(\tau)\right] \\
&\geq (I - Q_{dd})^{-1} \left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \epsilon_d - \beta\xi_d(\tau')\right] \\
&> (I - Q_{dd})^{-1} \left[(\bar{e} - \bar{s} - L(\tau'))\,\mathbf{1} - \epsilon_d - \beta\xi_d(\tau')\right],
\end{aligned}$$

where $\xi_d(\tau)$ and $\xi_d(\tau')$ denote the payment shortfall amounts with $\tau$ and $\tau'$, respectively. Therefore, the aggregate deadweight loss is greater with $\tau$.

**Case 3.2.** Suppose that the agent under liquidity shock defaults on its senior liabilities for both $\tau$ and $\tau'$. First, we show the following lemma.

**Lemma A.1.** *Suppose that the agent under liquidity shock $f$ defaults on its senior liabilities. If $\mathcal{D}(\tau)$ denotes the set of all other defaulting agents, then*

$$(I - Q_{dd})^{-1}\left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau) - \alpha(\tau)\bar{d}Q_{df}\right] < 0. \tag{31}$$

*Further, if (31) is satisfied for a subset of agents $\mathcal{D}(\tau)$, then all agents in $\mathcal{D}(\tau)$ default.*

**Proof of Lemma A.1.** For the first statement, suppose that the set of defaulting agents except agent $f$ is $\mathcal{D}(\tau)$ and the complement set, i.e., the set of solvent agents is $\mathcal{S}(\tau)$. Then, by the assumption,

$$\begin{aligned} x_d &= Q_{dd}x_d + \alpha(\tau)\bar{d}Q_{ds}\mathbf{1} + (\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau) \\ \Rightarrow x_d &= (I - Q_{dd})^{-1}\left[\alpha(\tau)\bar{d}Q_{ds}\mathbf{1} + (\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau)\right]. \end{aligned}$$

Since $Q$ is a stochastic matrix, $Q_{ds}\mathbf{1} + Q_{dd}\mathbf{1} + Q_{df} = \mathbf{1}$, implying that

$$\begin{aligned} Q_{ds}\mathbf{1} + Q_{df} &= (I - Q_{dd})\,\mathbf{1} \\ \Rightarrow Q_{ds}\mathbf{1} - (I - Q_{dd})\,\mathbf{1} &= -Q_{df}. \end{aligned}$$

Since agents in $\mathcal{D}(\tau)$ default, the payment vector $x_d$ is less than the promised payment $\alpha(\tau)\bar{d}$. Therefore,

$$\begin{aligned} (I - Q_{dd})^{-1}\left[\alpha(\tau)\bar{d}Q_{ds}\mathbf{1} + (\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau)\right] &< \alpha(\tau)\bar{d}\mathbf{1} \\ \Rightarrow (I - Q_{dd})^{-1}\left[\alpha(\tau)\bar{d}\left(Q_{ds}\mathbf{1} - (I - Q_{dd})\mathbf{1}\right) + (\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau)\right] &< 0 \\ \Rightarrow (I - Q_{dd})^{-1}\left[-\alpha(\tau)\bar{d}Q_{df} + (\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau)\right] &< 0, \end{aligned}$$

which shows the first statement.

For the second statement, suppose that (A.1) is satisfied for the agents in $\mathcal{D}(\tau)$. Again using the fact that $Q$ is a stochastic matrix, we obtain

$$\begin{aligned}&(I - Q_{dd})^{-1}\left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau) - \alpha(\tau)\bar{d}Q_{df}\right] < 0\\ \Rightarrow\ &(I - Q_{dd})^{-1}\left[\alpha(\tau)\bar{d}Q_{ds}\mathbf{1} + (\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau)\right] < \alpha(\tau)\bar{d}\mathbf{1},\end{aligned}$$

which implies that even if no other agents in $N \backslash \mathcal{D}(\tau)$ default, all agents in $\mathcal{D}(\tau)$ cannot pay their liabilities. Therefore, all agents in $\mathcal{D}(\tau)$ default. ■

As in the assumption of Lemma A.1, denote the agent under liquidity shock, the set of defaulting agents other than the agent under liquidity shock, and the set of solvent agents as $f$, $\mathcal{D}(\tau')$, and $\mathcal{S}(\tau')$, respectively. By Lemma A.1,

$$(I - Q_{dd})^{-1}\left[(\bar{e} - \bar{s} - L(\tau'))\,\mathbf{1} - \beta\xi_d(\tau') - \alpha(\tau')\bar{d}Q_{df}\right] < 0.$$

Because entries in $Q_{dd}$ is a sub-stochastic matrix with non-negative real numbers, $I - Q_{dd}$ is an M-matrix, which is a matrix whose off-diagonal entries are less than or equal to zero and whose eigenvalues have nonnegative real parts.[18] Therefore, $(I - Q_{dd})^{-1}$ is an inverse M-matrix and has nonnegative elements, implying

$$(I - Q_{dd})^{-1}\left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau') - \alpha(\tau)\bar{d}Q_{df}\right] < 0,$$

as $L(\tau) > L(\tau')$ and $\alpha(\tau) > \alpha(\tau')$. Then, the payment shortfalls are greater under $\tau$, hence,

$$\begin{aligned}&(I - Q_{dd})^{-1}\left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau) - \alpha(\tau)\bar{d}Q_{df}\right]\\ <\ &(I - Q_{dd})^{-1}\left[(\bar{e} - \bar{s} - L(\tau))\,\mathbf{1} - \beta\xi_d(\tau') - \alpha(\tau)\bar{d}Q_{df}\right] < 0,\end{aligned}$$

and the set of defaulting agents under $\tau$ is a superset of the set of defaulting agents with $\tau'$,

[18] If matrix $A$ can be expressed as $A = sI - B$, $s \geq \rho(B)$, $B \geq 0$, where $\rho(B)$ is the spectral radius of $B$, then $A$ is an M-matrix (Berman and Plemmons, 1979, p. 133). An M-matrix is non-singular if $s > \rho(B)$. See the appendix of Chang and Chuan (2024) for more details of how the concept of M-matrix can be applied to financial contagion.

as in the second statement of Lemma A.1. Hence, the aggregate deadweight loss is greater with $\tau$.

**Case 3.3.** Finally, suppose that the agent under liquidity shock defaults on its senior liabilities with $\tau'$ but not with $\tau$. In this case, the agent under liquidity shock, agent $f$, receives more payments from others, as $\alpha(\tau)\bar{d} > \alpha(\tau')\bar{d}$ for the payments from solvent agents, and able to pay its senior liabilities under $\tau$. From Cases 1 and 2, $\mathcal{D}(\tau') \subseteq \mathcal{D}(\tau)$ as noted in the beginning of Case 3. Then, by following the same steps in Case 3.1,

$$\alpha(\tau)\bar{d}\mathbf{1} - x_d(\tau) > \alpha(\tau')\bar{d}\mathbf{1} - x_d(\tau') \geq \sum_{j \in N} \xi_j,$$

where $x_d(\tau')$ includes the negative payment from agent $f$ with $x_f(\tau') = 0$ by the initial assumption of Case 3.3. Therefore, the aggregate deadweight loss is greater with $\tau$. ■

### A.5. Statement and Proof of Lemma A.2

**Lemma A.2.** *The vector of defaulting amounts* $\alpha(\tau)\bar{d}\mathbf{1} - x_d$*, where each entry is* $\alpha(\tau)\bar{d} - x_j(\tau, D, \epsilon)$ *for each* $j \in \mathcal{D}(\tau, \epsilon)$ *can be represented as*

$$\alpha(\tau)\bar{d}\left(\mathbf{1} + \beta C(\mathcal{D}(\tau,\epsilon))\mathbf{1}\right) + (1+\beta)\left(C(\mathcal{D}(\tau,\epsilon))L(\tau)\mathbf{1} - C(\mathcal{D}(\tau,\epsilon))(e-s-\epsilon)\right), \tag{32}$$

*where* $C(\mathcal{D}(\tau,\epsilon)) \equiv \left(I - (1+\beta)Q_{\mathcal{D}(\tau,\epsilon)}\right)^{-1}$*, and* $C(\mathcal{D}(\tau,\epsilon))\mathbf{1}$ *is the* $|\mathcal{D}(\tau,\epsilon)| \times 1$ *vector of node depth centrality.*

**Proof.** We can express (15) in a matrix notation as

$$\alpha(\tau)\bar{d}\mathbf{1} - x_d = (1+\beta)\left(\alpha(\tau)\bar{d}\mathbf{1} + s + \epsilon + L(\tau)\mathbf{1} - e - Q_{\mathcal{D}(\tau,\epsilon)}x_d\right).$$

Rearranging the expression to collect $x_d$ term yields

$$\begin{aligned}
\left(I - (1+\beta)Q_{\mathcal{D}(\tau,\epsilon)}\right)x_d &= -\beta\alpha(\tau)\bar{d}\mathbf{1} + (1+\beta)\left(e - s - \epsilon - L(\tau)\mathbf{1}\right)\\
\Rightarrow x_d &= -\beta\alpha(\tau)\bar{d}\left(I - (1+\beta)Q_{\mathcal{D}(\tau,\epsilon)}\right)^{-1}\mathbf{1} + (1+\beta)\left(I - (1+\beta)Q_{\mathcal{D}(\tau,\epsilon)}\right)^{-1}\left(e - s - \epsilon - L(\tau)\mathbf{1}\right)\\
\Rightarrow x_d &= -\beta\alpha(\tau)\bar{d}C(\mathcal{D}(\tau,\epsilon))\mathbf{1} + (1+\beta)C(\mathcal{D}(\tau,\epsilon))\left(e - s - \epsilon - L(\tau)\mathbf{1}\right).
\end{aligned} \tag{33}$$

Plugging (33) into $\alpha(\tau)\bar{d}\mathbf{1} - x_d$ implies

$$\begin{aligned}
&\alpha(\tau)\bar{d}\mathbf{1} - x_d\\
&= \alpha(\tau)\bar{d}\left[\mathbf{1} + \beta C(\mathcal{D}(\tau,\epsilon))\mathbf{1}\right] + (1+\beta)L(\tau)C(\mathcal{D}(\tau,\epsilon))\mathbf{1} - (1+\beta)C(\mathcal{D}(\tau,\epsilon))(e-s-\epsilon),
\end{aligned}$$

i.e., the expression in (32). ■

### A.6. Proof of Theorem 1

**Proof.** For the first part, by Proposition 3, we know that the cost and benefits of an increase in $\tau$ depends on the functional forms of the liquidity cost, the probability of shock arrival, and the netting efficiency. From (32), we know that the conditional aggregate deadweight loss depends on the centrality of agents in the defaulting set.

Furthermore, from (32), the decrease in the aggregate deadweight loss $\frac{\partial \xi_j(\tau, D, \epsilon)}{\partial \tau}$ equals

$$\frac{1}{1+\beta} \sum_{j \in \mathcal{D}(\tau,\epsilon)} \left[ (1 + \beta C_j(\mathcal{D}(\tau,\epsilon)))\, \alpha'(\tau)\bar{d} + (1+\beta) C_j(\mathcal{D}(\tau,\epsilon)) L'(\tau) \right].$$

Then, the expression (14) has to be multiplied by the probability of the realization of the liquidity shock $\epsilon$, $G(\epsilon)$.

For the second part, we first show that an increase in the set of defaulting agents results in a discontinuous jump in the aggregate deadweight loss. The definition of node depth centrality in (17) implies that an increase in the set of defaulting agents leads to a discrete increase in the centrality, because

$$\begin{aligned} & 1 + (1+\beta) \sum_{k \in \mathcal{D}(\tau,\epsilon)} q_{kj} + \cdots \\ < & 1 + (1+\beta) \sum_{k \in \mathcal{D}(\tau',\epsilon)} q_{kj} + \cdots, \end{aligned}$$

if $\mathcal{D}(\tau, \epsilon) \subset \mathcal{D}(\tau', \epsilon)$ and $\mathcal{D}(\tau', \epsilon) \not\subset \mathcal{D}(\tau, \epsilon)$. Furthermore, the initial summation of (17) includes more agents in $\mathcal{D}(\tau', \epsilon)$. Therefore, the aggregate deadweight loss shows a discrete jump when there is an increase in the set of defaulting agents.

Because $\tau$ is a default threshold point, by definition, $\mathcal{D}(\tau, \epsilon) \subset \mathcal{D}(\tau - \delta, \epsilon)$ and $\mathcal{D}(\tau - \delta, \epsilon) \not\subset \mathcal{D}(\tau, \epsilon)$ for any small $\delta > 0$. Therefore, a change from $\tau - \delta$ to $\tau$ will result in a change in the defaulting set and discrete decrease in the aggregate deadweight loss. Such a change is multiplied by the shock realization probability $G(\epsilon)$ and the probability of shock arrival

$F(\tau)$.

Because all functions are continuously changing in $\tau$ and are linearly combined for the ex ante welfare loss and payment shortfall function, they are continuous in $\tau$. Thus, a marginal increase in $\tau$ from a threshold point does not result in an additional solvency, as there is already a discrete jump in payments due to the reduction of defaulting set of agents. Therefore, the right-hand derivative of the ex ante welfare loss does not involve a discrete change in welfare loss. Hence, $\frac{\partial W(\tau, D)}{\partial \tau^-} < \frac{\partial W(\tau, D)}{\partial \tau^+}$. ■

### A.7. Proof of Proposition 4

**Proof. Step 1. Solving payment equilibrium is polytime.** For any given $\tau$, computing the payment equilibrium clearing vector $x(\tau, D, \epsilon)$ reduces to solving a fixed-point iteration in the framework of Eisenberg and Noe (2001) with a default cost extension. Rogers and Veraart (2013) show that this algorithm finds a fixed point at most in $n$ steps of iteration. Therefore, $\mathcal{D}(\tau, \epsilon)$ can be computed in $n$ steps or in polynomial time for any $\tau$.

**Step 2. Monotonicity.** By Lemma 1, $\mathcal{D}(\tau, \epsilon)$ is monotonically (weakly) decreasing in $\tau$. Thus, each agent's status of being default or solvent switches only once as $\tau$ increases.

**Step 3. Checking existence of a default threshold point.** For the given network $D$ and parameters and functional forms $(\beta, \bar{e}, \bar{s}, \alpha(\cdot), L(\cdot))$, and shock realization $\epsilon$, fix an interval $[a, b] \subset \mathbb{R}^+$. We first ask whether there exists $\tilde{\tau} \in [a, b]$ such that $\tilde{\tau} \in \mathcal{T}(D, \epsilon)$. This is equivalent to computing $\mathcal{D}(a, \epsilon)$ and $\mathcal{D}(b, \epsilon)$, which is polytime as shown in Step 1. If $\mathcal{D}(a, \epsilon) = \mathcal{D}(b, \epsilon)$, then by Step 2, no threshold exists in $[a, b]$. If $\mathcal{D}(a, \epsilon) \neq \mathcal{D}(b, \epsilon)$, then at least one default threshold point exists. This entire procedure is again solvable in polynomial time.

**Step 4. Binary search.** If an agent $j$'s status switches within the interval $[a, b]$, i.e., $j \in \mathcal{D}(a, \epsilon)$ and $j \notin \mathcal{D}(b, \epsilon)$, then there exists a default threshold point. By Step 2, there are at most $n$ thresholds. Each can be located to precision $\delta$ by binary search: iteratively evaluating $\mathcal{D}(\tau, \epsilon)$ at midpoints until the switch point is bracketed within a subinterval. This requires $O\left(\log\left(\frac{b-a}{\delta}\right)\right)$ of payment equilibrium solves per agent. Thus, for any arbitrary precision $\delta > 0$, the overall procedure is solvable in polynomial time.

**Step 5. Approximation by discretization.** By step 4, for any arbitrary precision $\delta > 0$, we can compute the default threshold points for a given $\epsilon$, $\mathcal{T}(D, \epsilon)$. If the distribution $G(\epsilon)$ is discrete, then we are done. Now suppose that $G(\epsilon)$ is a continuous distribution. Then, an approximation of $G(\epsilon)$ is $\tilde{G}(\epsilon)$, which is supported on a finite set $\{\epsilon^{(1)}, \ldots, \epsilon^{(M)}\}$. Then, the union of default threshold points for the approximation, i.e.,

$$\tilde{\mathcal{T}}(D) \equiv \bigcup_{m=1}^{M} \mathcal{T}(D, \epsilon^{(m)}) \tag{34}$$

is computable in polynomial time. ■

## B. Details of Netting

In this section, we discuss the details of various netting procedures and the resulting full netting matrices.

The simplest netting procedure is bilateral netting—netting the payments between two agents in all existing pairs. The full netting matrix under bilateral netting of the payment network $D$ is

$$\underline{D} = \left[D - D^T\right]^+ . \tag{35}$$

There are multiple ways to perform multilateral netting or portfolio compression. One of the more relevant and realistic netting procedure is simply eliminating cycles that D'Errico and Roukny (2021) defined as conservative portfolio compression. For example, Veraart (2022) builds upon D'Errico and Roukny (2021) and considers the effect of eliminating all cycles by a degree of $\mu$. Therefore, our model is generalizing the netting (or compression) in her model by allowing time-varying degree of netting, $\alpha(\tau)$. The exact procedure of obtaining the full netting matrix under this cycle elimination is the following:[19]

1. Fix the index $\kappa = 1$ and define $D^0 \equiv D$.

2. From the matrix $D^{\kappa-1}$, find a cycle, indexed by $\kappa$, which is a sequence of agents $i_0, \ldots, i_K$, for some $K \geq 2$ such that $i_0 = i_K$, $i_\ell \neq i_0$, and $d^{\kappa-1}_{i_{\ell+1} i_\ell} \geq \underline{d}^\kappa > 0$ for each $\ell < K$, where $\underline{d}^\kappa \equiv \max_\delta \left\{ \delta \leq d^{\kappa-1}_{i_{\ell+1} i_\ell} : \forall 0 \leq \ell < K \right\}$.

3. Subtract $\underline{d}^\kappa$ from each $d^{\kappa-1}_{i_{\ell+1} i_\ell}$ along the cycle $\kappa$. Denote the new matrix as $D^\kappa$.

4. If there is a cycle in $D^\kappa$, return to step 2. If there is no cycle, define the resulting matrix as the full netting matrix $\underline{D} \equiv D^\kappa$.

[19] This procedure is equivalent to obtaining the $\mu$-compressed liabilities matrix with $\mu = \mu^{\max}$ in Veraart (2022). Note that instead of changing the parameter $\mu$, we set the time varying value of $1-\alpha(\tau)$ for the degree of netting. Our procedure is closer to how to design numerical procedures, which might be computationally intensive as shown by Jackson and Pernoud (2024).

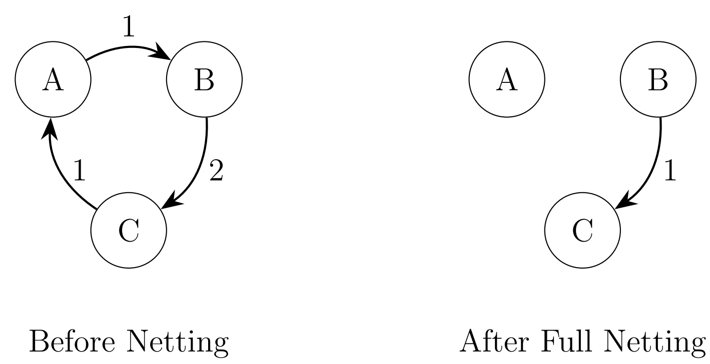



Figure 6: An example of non-zero full netting matrix

Note: The left graph depicts the network, in which agent $A$ owes 1 to agent $B$, who owes 1 to agent $C$, before netting. The right graph depicts the same network after full netting, in which agent $B$ still owes 1 to agent $C$.

If a network is regular, then the full netting matrix under this procedure will be $\underline{D} = \mathbf{0}$.

However, the full netting matrix can be a non-zero matrix for non-regular networks. For example, consider a simple ring network as depicted in the left panel of Figure 6. Agent $A$ owes 1 to agent $B$, agent $B$ owes 2 to agent $C$, and agent $C$ owes 1 to agent $A$. The netting procedure can eliminate the cycle up to the maximum of 1. Therefore, as in the right panel of Figure 6, agent $B$ still owes 1 to agent $C$ after the full netting procedure. The full netting matrix is not a zero matrix in this example.

## C. Discussion of Social Welfare

In this section, we define the ex ante (expected) social welfare and discuss how it relates to the ex ante welfare loss defined in (6).

Define the expected transfer from agent $i$ to the sources of liquidity shock as $T_i \equiv F(\tau)1/n\epsilon$. As previously mentioned, we consider transfers to liquidity shock as simple transfers of wealth from agents to senior creditors and vice versa. The sum of agents' expected utilities plus the expected transfers to liquidity shock net of the deadweight loss incurred to external agents is

$$\begin{aligned}
&\sum_{i\in N}\left(U_i(\tau,D)+T_i-F(\tau)(1-\gamma)E\left[\min\left\{\beta\xi_i(\tau,D,\epsilon),A_i(\tau,D,\epsilon)\right\}\right]\right)\\
&=\sum_{i\in N}(e_i-s_i)-nL(\tau)+(1-F(\tau))\left[\sum_{i\in N}\sum_{j\in N}q_{ij}\hat{d}_j-\sum_{i\in N}\hat{d}_i\right]\\
&\quad+F\left(\tau\right)E\left[\sum_{i\in N}\sum_{j\in N}q_{ij}x_j-\sum_{i\in N}x_i-\sum_{i\in N}(\epsilon_i-\epsilon_i)-\sum_{i\in N}\min\left\{\beta\xi_i,A_i\right\}\right]\\
&=\sum_{i\in N}(e_i-s_i)-nL(\tau)-F\left(\tau\right)E\left[\sum_{j\in N}\min\left\{\beta\xi_j,A_j\right\}\right],
\end{aligned}\tag{36}$$

because $\sum_{j\in N}q_{ij}=1$.

Since the first term of (36) does not depend on $\tau$, $D$, or $\epsilon$, we can focus on the second and third terms only. Therefore, the ex ante (expected) social welfare losses, defined by (6), represent the relevant changes in social welfare due to changes in the settlement time $\tau$ and the network structure $D$.