Risk Concentration and Interconnectedness in OTC Markets*

Briana Chang† Shengxing Zhang‡

This Draft: September 2021

Abstract

We develop a tractable framework that jointly determines trading connections and risk allocations among banks. It links the risk-taking incentives at the individual level to the aggregate trading network. The network is payoff unique and generally features risk concentration and heterogeneous risk exposures across banks, despite they're ex-ante homogeneous. A change in the risk-taking incentives can result in regime shifts and thus discontinuous change in aggregate risks. It further delivers new insights on regulatory reforms, taking into account the equilibrium response in networks.

Keywords: Trading Network, Over-the-Counter Market, Reforms

JEL classification: C70, G1, G20

---

*We would like to thank Manuel Amador, Scott Bauguess, Piero Gottardi, Peter Kondor, Cecilia Parlatore, Guillaume Rocheteau, Marzena Rostek, Gideon Saar, Neil Wallace, and Pierre-Olivier Weill for their useful discussions and comments. We also thank participants at Penn State, LSE, UT Austin, Houston, the Wisconsin School of Business, the Carlson School of Management, the Federal Reserve Bank of Minneapolis, Philadelphia Fed Conference – the Impact of Post-Crisis Regulation on Financial Markets, the 2019 Society for Economic Dynamics Annual Meeting, and the Virtual Finance Theory Seminar.

†University of Wisconsin–Madison, Wisconsin School of Business, Finance Department, 975 University Avenue, Madison, WI 53706, USA; briana.chang@wisc.edu.

‡Department of Economics, London School of Economics; s.zhang31@lse.ac.uk. Zhang thanks the Centre for Macroeconomics at the London School of Economics for its financial support.

1 Introduction

This paper develops a tractable framework that determines jointly trading connections and risk allocation among banks. We analyze how risk-taking incentives at the individual level affect the aggregate trading network. In an environment where all banks are ex-ante homogeneous and risk-averse, we show that the network is unique and generally asymmetric, where risks are concentrated among a subset of banks, consistent with observed empirical regularities in OTC markets.

The degree of network asymmetry is determined by the tension between the benefit of risk-sharing vs. risk-concentration. Our tractable characterization allows us to analyze how the underlying network responds to varied policies and/or the riskiness of underlying assets. We are thus able to provide new insights on how and when the interconnections and risk exposure of large financial institutions may change.

The framework contributes to the literature on OTC markets by explicitly modeling the formation of bilateral trading relationships under limited information. The information friction arises from the uncertainty on other banks' asset positions when bankers form trading relationships and their limited capacity in creating trading relationships.

We further allow banks to choose an individual action to minimize their private costs of bearing risks. These actions include, for example, whether to access a multilateral trading platform that allows participants to better share risks or whether to default to offload downside risks to outside creditors.

Banks choose their trading partners sequentially for multiple trading rounds as well as their final actions, matching with one counterparty in each round. The bilateral matching in each round is based on potential counterparties' identity and, more importantly, beliefs about their asset positions. We apply pairwise stability to determine equilibrium matching for each trading rounds. The collection of a bank's counterparties over multiple trading rounds forms its trading links.

Our framework highlights the interdependence between bilateral connections, the risk taking actions that each bank takes, and the risk allocation among them. Despite all banks being ex ante homogeneous and risk averse, they may use their bilateral networks to concentrate risks on certain banks which in turn have higher incentives to take actions to reduce their final private risk exposures. As a result, the resulting network features risk concentration, and banks end up with heterogeneous risk exposures.

Formally, the key dynamic element in our model is the public belief about a bank's asset positions at a trading round, which endogenously depends on with whom it has traded with and how it has traded with its counterparties. We establish that the sufficient statics is the variance of a bank's position, which can be interpreted as the risk that a bank bears. When banks' payoffs are convex in their risk positions (which arises naturally due to their optimal decision at the terminal period), those with higher risk positions are matched with each other and the risk-allocation within any pair is generally asymmetric. That is, in each trading round, there is positive sorting matching on the risk banks bear; moreover, the distribution of risk across banks becomes more dispersed after trades.

Defining the risk-bearing capacity of a bank as its marginal value over the variance of its asset holding, we establish that, fixing any connections, the effective risk-bearing capacity of a bank at period t can be characterized recursively as the harmonic mean of its and its period-t counterparty's next period risk-capacity. Moreover, the optimal network must result in more symmetric risk capacity in earlier matches so that risk-concentration arise later rather than earlier.

Thanks to this property, in the simple case with binary actions for risk taking (where we refer banks that take the action to reduce their costs of holding risks as core banks), the number of core banks an agent is directly or indirectly connected with at any period becomes the sufficient static of an agent's effective risk-capacity at that period. The optimal network is then reduced to choosing the optimal core size.

We then use this analytical characterization to study the positive and normative implications of reforms that promote central clearing and/or discourage risk taking, taking into account the equilibrium response of the underlying market structure. When banks can choose whether to have platform access, the aggregate market structure can be summarized by the ratio between the cost of bearing risks and the entry cost to access multilateral clearing platform to reduce risk exposure. Consistent with empirical evidence, our model predicts that policies that increase balance sheet costs relative to the entry cost could result in a more symmetric market structure. Nevertheless, it can have ambiguous effects on transaction costs measured by bid-ask spreads.

In the second application, we consider banks' default behaviors and the interbank lending network. We assume that the default probability increases in their final risk-positions. We show that a small increase banks' risk-taking incentives can result in a dramatic shift in the interbank network, whereupon banks switch from sharing risks with each other through the network to concentrating risks to a small set of banks. The switch to risk concentration results in a discontinuously large increase in aggregate default probability. In this sense, a small shock can trigger "crises" through financial networks.

In this application, our notion of default risks is distinct from standard theories of financial contagion, which highlight how bank defaults propagate through given network connections.[1] In our framework, the aggregate default risks increase because banks systematically change their lending behaviors through the interbank network.

Related Literature Methodologically, our dynamic framework with repeated bilateral matching[2] contributes a tractable approach to studying the formation of trading network. Our method differs from the existing network formation literature[3] as it breaks down a complex network formation game into a sequence of subgames, each of which involves one round of bilateral matching together with asset trading, and a subsequent sub-game. How an agent traded in the past is summarized by his characteristic, which becomes the state variable governing how he trades in later periods. By imposing sequential rationality, we can solve the network formation problem through backward induction.

While we use pairwise stability to characterize the equilibrium matching in a subgame, a deviating agent in a subgame can change all his future links, not just one link as in the static setup that the literature adopt. This method derives a unique solution. It is thus in sharp contrast to the standard network formation problem where agents form multiple

[1]A growing literature focuses on the role of the architecture of financial systems as an amplification mechanism. For example, Allen et al. (2000)(Allen and D. Gale 2000), Acemoglu et al. (2014)(Acemoglu, Ozdaglar, and Tahbaz-Salehi 2013), Elliott et al. (2014)(Elliott, Golub, and Jackson 2014), Cabrales et al. (2014)(Cabrales, Gottardi, and Vega-Redondo 2014), and Gofman (2014) (Gofman 2014) study the financial contagion in given networks.

[2]Most works in the matching literature involve a static environment, with only a few exceptions. Corbae, Temzelides, and Wright (2003) introduced directed matching into the money literature, where the key state variable is the traders' money holding. Because there are no information frictions in Corbae, Temzelides, and Wright (2003), belief updating is not essential for their analysis, whereas it is a key component of our theory. With regard to the labor market, Anderson and Smith (2010) analyzed the dynamic matching pattern for which the public belief about a trader's skill (i.e., his reputation) evolves according to matching decisions. In our trading environment, the updating process depends endogenously on both the traders' matching decisions and the terms of trade within a match.

[3]See the survey in Jackson 2005 for overview. Specifically, papers that have studied network formation in the financial market include Hojman and Szeidl (2008), D. M. Gale and Kariv (2007), Babus and Hu (2017), and Cabrales, Gottardi, and Vega-Redondo (2017), Farboodi (2014), Wang (2016)), where the last two papers in particular focuses on the core-periphery structure.

links simultaneously, which is often subject to the curse of dimensionality and prone to multiple equilibria, because pairwise stability allows for the deviation of only one pair of traders even though traders form multiple links.

A similar approach has been used in our previous work, Chang and Zhang (2018), where we consider a pure bilateral OTC market with risk-neutral agents and an indivisible asset. This paper allows for risk-averse agents and unrestricted asset holdings, which allows us to analyze risk concentration within the network.

The common approaches to modeling OTC markets are based on random matching (e.g., Duffie, Gârleanu, and Pedersen (2005)) or exogenous networks.[4] Relative to the literature that takes the network as given, our model provides a formal analysis of how the underlying structure of the OTC market might respond to policies.

Because one of our applications is on the joint determinations of bilateral network and platform access, our paper also sheds new lights on the literature on the costs and benefits of centralized vs. decentralized markets.[5] Instead of focusing on the trade-off between these two markets, we allow for nonexclusive participation and emphasize the interdependence between these two choices. The paper is related to recent works that studies the co-existence of these two venues and market fragmentation, including Dugast, Üslü, and Weill (2019) and Babus and Parlatore (2017). Our framework is designed to analyze the network response and the results can be generalized to multiple types of platforms.

2 Model

We consider a trading game in an economy that lasts $N+1$ periods and is populated by a set of banks, each with a fixed identity $i \in \mathbb{I} = [0, 1]$. Banks trade among each other for N periods and allow to take an optimal action to minimize their risk exposure at the terminal period $N+1$. We allow for a general payoff function at the terminal period and characterize how the payoff function affects the trading network through N rounds of bilateral connections.

[4]For example, see Gofman (2011), Babus and Kondor (2018), and Malamud and Rostek (2014).

[5]Specifically, existing studies (e.g., Malamud and Rostek (2014), Glode and Opp (2019), and Yoon (2017)) consider other dimensions such as price impact and asymmetric information. They show that OTC markets can be beneficial for certain types of traders. In our model, a centralized platform is assumed to be a superior trading technology but requires a higher participation cost.

There are two types of consumption goods, numeraire goods and dividend goods, and one type of asset. The asset generates a unit stream of dividend goods in each period. All banks are endowed with an initial asset position which is an i.i.d. draw from a symmetric distribution with mean zero, variance v_1, and distribution function $\pi_1(a)$.

The heterogeneity in asset positions is the source of gains from trade in the economy. Banks can trade their asset positions with numeraire goods, of which they have deep pockets. The flow utility at period t of a bank i that has asset position $a_{i,t} \in \mathbb{R}$ and receives transfer $x_{i,t} \in \mathbb{R}$ is $u_t(a_{i,t}) + x_{i,t}$. We assume that a bank derives mean-variance utility from dividend goods and normalize the mean to zero; above, $u_t(a_{i,t}) = -\kappa_t a_{i,t}^2$ with $\kappa_t \geq 0$ for all $t \leq N$.[6] In other words, the ideal asset position of a bank is normalized zero. Parameter κ_t represents the balance sheet cost of holding nonzero asset positions at period t, which can be associated with the riskiness of the asset.

Contacting Frictions in Bilateral Trades From period 1 to period N, banks can connect sequentially to N counterparties with no extra cost to engage in N rounds of bilateral trades. Bilateral trades are subject to limited information that prevents banks from locating ideal trading counterparties.

We explicitly model this friction by assuming that an bank can only observe another bank's asset position after the two have contacted one another. In other words, each bank faces uncertainty about the counterparty's asset position *before* making the contact. Thus, there is limited information at the matching stage but complete information between matched banks after they make contact.

Observe that, given the assumed payoff structure, if all banks could observe each other's realized positions before they choose their matches, it is straightforward to show that the economy achieves perfect risk sharing with one round of trade. In this case, banks with position a are matched with banks with the opposite position $-a$, and their posttrade positions would net out to zero (i.e., there would be perfect negative sorting on asset positions.) Hence, the assumed contacting frictions aim to capture the spirit of conventional search frictions.

[6] More generally, $u_t(a_{i,t}) = \kappa_{0,t} a_{i,t} + \kappa_{1,t} a_{i,t}^2$. Because $\kappa_{0,t}$ does not contribute to the heterogeneity in marginal utility, it is without loss of generality to set it to zero.

Post-Trade Risk Exposures To study the interaction of banks' risk-taking behavior and the OTC network, we assume that the expected payoff of an bank depends on the variance of his asset position after OTC trades, denoted by $W_{N+1}(v_{i,N+1})$, where $v_{i,N+1}$ is the variance of $\pi_{i,N+1}(a)$. Observe that, if there were no additional actions at $N+1$, then the final payoff for a bank is simply $W_{N+1}(v_{i,N+1}) = Eu_{N+1}(a_{i,N+1}) = -\kappa_{N+1}v_{i,N+1}$. In general, we assume that $W_{N+1}(\cdot)$ is a decreasing function of post-trade risk-exposure $v_{i,N+1}$.

As will be explained later, the interesting case is when $W_{N+1}(\cdot)$ is convex. The convexity naturally arises when banks can potentially choose different actions given their risk positions $v_{i,N+1}$. Assumption 1 below allows finite possible actions, denoted by z, and assumes that, give any action z, the payoff is linear in their risk exposure v.

Assumption 1: Given any action z, the payoff is linear in their risk exposure v. Banks' final payoff $W_{N+1}(v)$ is given by

$$W_{N+1}(v) = \max_{z} \left\{ -\gamma_{N+1}(z)v - \phi_{N+1}(z) \right\}.$$

For example, one can interpret z as different trading platforms, where lower $\gamma_{N+1}(z)$ means agents can have lower post-trade exposure in that platform and $\phi_{N+1}(z)$ represents its corresponding entry cost. For example, a fully competitive centralized market can be understood as a platform that allows fully risk-sharing; hence, $\gamma_{N+1}(z) = 0$.

We will study two applications where banks' risk-taking behavior interacts with the bilateral OTC market structure. First, allowing banks to access multilateral clearing platforms by paying a fixed costs. Second, the moral hazard of taking risks when the asset positions represent a bank's net debt positions and banks have limited commitment to debt repayment, where we assume that banks' default probability increases in banks' final exposures. In both cases, $W_{N+1}(v)$ is convex in v.

2.1 Matching and Trading Decisions

Given the uncertainty, the matching decisions are thus based on the identities of their counterparties. Formally, the choice of counterparties is modeled as choosing N counterparties sequentially at $t = 0$; that is, banks decide ex ante bilateral matches for each trading round.

Ex Ante Network Denote the trading counterparty of a bank i at period t $j_{i,t}$. The collection of a bank i's counterparties $j_{i,t}$ over N rounds of trade forms his trading links. We assume that banks form their trading links before their asset holdings and valuations are realized. Therefore, our setup effectively has a network formation stage ex ante, and we can interpret trading links as permanent trading relationships between banks. Since trading needs are banks' private information at the trading stage, the assumption that banks form trading links ex ante and cannot be contingent on realized trading needs also avoids some technical complications in matching models under asymmetric information.[7]

Terms of Trade: Contingent Asset Flows and Prices While the connections are determined ex ante, trades are contingent on the realized asset positions of a bank and her counterparty in a match, because trading takes place after she and her counterparty make their contact and observe each other's realized asset positions. Thus, if we think of the economy as a trading game within a trading day and repeat it over time, even though the network remains the same, banks' realized asset positions change how they trade (i.e., the asset flows) within the network from day to day.

Formally, the terms of trade within a match, including both asset allocations and transfers of numeraire goods, are contingent on the realized positions of a bank i and her counterparty j, denoted by a_i and a_j respectively. Let $y(i,j) = \{\tilde{a}_k(a_i, a_j), \tilde{x}_k(a_i, a_j), k \in \{i,j\}\}$ be the terms of trade within the match (i,j), where $\tilde{a}_k(a_i, a_j)$ denotes the posttrade asset holding of bank k, and $\tilde{x}_k(a_i, a_j)$ denotes the transfer to bank k, $k \in \{i,j\}$. The within-match transfers sum up to zero,

$$\Sigma_{k=i,j}\tilde{x}_k(a_i, a_j) = 0. \tag{1}$$

The within-match asset allocation is feasible if

$$\Sigma_{k=i,j}\tilde{a}_k(a_i, a_j) = a_i + a_j. \tag{2}$$

The allocation of asset positions is associated with the allocation of risks from uncertain

[7] Without this assumption, banks can in theory signal their types through different matching decisions and the equilibrium would depend on how we specify off-equilibrium beliefs and require heavier notations. One can in theory impose off-equilibrium beliefs that support a pooling equilibrium and obtain the same outcome.

asset positions because given a distribution of banks i and j's pretrade asset positions, the posttrade positions also follow a distribution. This is the key characteristic that governs bilateral matching.

Sequential Choices of Trading Links and Terms of Trade When banks decide trading links and terms of trade ex ante, they make decisions for earlier trading rounds first. All trading links and terms of trade before a period t are public information when banks decide matching and within-match terms of trade for the period. Thus, links and terms of trade are sequentially optimal in the sense that when a bank chooses his counterparty and terms of trade for a period t, he takes into account all banks' matches and terms of trade before the period.

A bank i's strategy at period t conditional on the public information at that period includes the choice of his counterparty, $j_{i,t}$, and the terms of trade with the counterparty, $y_t(i,j)$ for $j = j_{i,t}$.

We can summarize the public information for period t strategies by the public belief of joint distribution of banks' asset positions.[8] Now that banks' strategies are contingent on the public belief of banks' trading needs, characterizing its evolution over time is an essential part of our analysis. Denote the joint distribution of banks' asset holdings at the beginning of period t $\boldsymbol{\pi}_t : \mathbb{R}^{[0,1]} \to [0,1]$ and the marginal distribution of bank i's asset position at the beginning of period $\pi_{i,t}(a) : \mathbb{R} \to [0,1]$.

Evolving Characteristics To understand how a bank's asset holding distribution evolves over time, consider the following example: suppose a bank i bears all position exposures within her match at period 1. That is, her asset position in the next period equals the sum of her and her counterparty j's current asset positions, $a_{i,2} = a_{i,1} + a_{j,1}$. Her posttrade asset distribution $\pi_{i,2}(a)$ now has mean zero and a variance of $2v_1$ when her pretrade position is uncorrelated with her counterparty's. On the other hand, under this first-period strategy, her counterparty's posttrade asset position is always zero, $a_{j,2} = 0$ (i.e., $\pi_{j,2}(a)$ is degenerate with both its mean and variance being zero).

In general, the law of motion of the asset distribution of a bank i, $\pi_{i,t}(a)$, is given by

[8] As we will show later, the gains from trade from period t onwards depend on the trading history only through the public belief of banks' asset positions.

the Bayes' rule,

$$\pi_{i,t+1}(a) = \int\int \mathbb{I}(\tilde{a}_{i,t}(a_i, a_j) \leq a)\boldsymbol{\pi}_{i,j,t}(da_i, da_j), \text{ for } a \in \mathbb{R}, \tag{3}$$

where $\boldsymbol{\pi}_{i,j,t}(a_i, a_{-i})$ denotes the joint distribution of bank i and her counterparty j's period-t pretrade asset positions. This again highlights the fact that bank i's posttrade asset distribution, $\pi_{i,t+1}(a)$, depends on the the joint distribution of the pretrade asset positions of bank i and her optimally chosen counterparty, and on how she trades with her counterparty, $\tilde{a}_{i.t}(a_i, a_j)$.

To sum up, we study a dynamic matching model with evolving characteristics; the marginal asset distribution $\pi_{i,t}(a)$ and the correlation pattern between the marginal distributions depend on past matching and trading decisions. We can think of the joint distribution $\boldsymbol{\pi}_t$ of all banks' asset positions as the aggregate state variable.

2.2 Equilibrium

2.2.1 Equilibrium Definition

Denote the joint payoff between two banks i and j, $\Omega_t(i, j)$. Given the aggregate distribution at period t as

$$\Omega_t(i, j) \equiv \max_{\tilde{a}_{i,t}, \tilde{a}_{j,t}} -\kappa_t \int\int [(\tilde{a}_{i,t})^2 + (\tilde{a}_{j,t})^2]\boldsymbol{\pi}_{i,j,t}(da_i, da_j) + \hat{W}_{t+1}(i) + \hat{W}_{t+1}(j) \tag{4}$$

subject to feasibility constraints, which depends on the pretrade joint asset distribution of banks i and j, $\pi_{i,j,t}(a_i, a_j)$. The within-match transfers do not show up in (4) because they sum up to zero.

Here, we use $\hat{W}_{t+1}(i)$ to denote the bank's maximum payoff in the next period with any marginal distribution $\pi_{i,t}(a)$ and joint distribution with other banks' asset holding, taking the aggregate distribution $\boldsymbol{\pi}_{t+1}$ and other banks' equilibrium payoffs $W_{j+1}(j)$ as given, which yields

$$\hat{W}_{t+1}(i) \equiv \max_j \Omega_{t+1}(i, j) - W_{j+1}(j). \tag{5}$$

On the equilibrium path, a bank's payoff is given by $W_{t+1}(i)$, which equals $\hat{W}_{t+1}(i)$ for a bank i that adopts equilibrium strategies before period $t+1$.

Definition 1. Given $\boldsymbol{\pi_0}$, an equilibrium consists of strategies $\{s^*_{i,t}\}_{\forall i,t}$, market utilities $W_t(i)$, and a path of common beliefs $\boldsymbol{\pi}^*_t$ such that the following properties hold for all $t \in \{1, \ldots, N+1\}$:

1. Pairwise stability at $t \leq N$: if $j \in j_t(i)$,

$$W_t(i) = \max_j \Omega_t(i,j) - W_t(j),$$

   where the post-trade position $\{\tilde{a}_{i,t}, \tilde{a}_{j,t}\}$ maximizes Equation (4).

2. Feasibility of bilateral matching at $t \leq N$.

3. Dynamic Bayesian consistency: The joint asset distributions evolves following the Bayes rule given banks' strategies.

Our equilibrium notion can be understood as multiple rounds of pairwise stabile matching. Bilateral matches across all banks at period t are stable if no individuals in a match can be better off by forming new matches, conditional on providing the counterparty at least the latter's equilibrium market utility, denoted by $W_t(j)$.

Our notion, however, does allow for joint deviations with multiple banks that occur sequentially, which is thus different from the standard pairwise stability in simultaneous-move network formation games. Specifically, when a bank deviates at period t, the bank is also allowed to switch own *future* trading partners accordingly, conditional on providing own counterparties with equilibrium payoff $W_{t+1}(j)$. The deviation payoff is described by Equation (5), which allows banks to re-optimize their future counterparties.

2.2.2 Equivalence and Uniqueness

Denote the aggregate payoff of the economy at period t to be Π_t, which depends on the joint asset distribution $\boldsymbol{\pi}_t$. Given a strategy s_t at period t, the aggregate payoff equals

$$\Pi_t(\boldsymbol{\pi}_t) = -\kappa_t \int_0^1 E_t(\tilde{a}_{i,t}{}^2)di + \Pi_{t+1}(\boldsymbol{\pi}_{t+1}). \tag{6}$$

where $E_t(\tilde{a}_{i,t}{}^2) = \int\int \tilde{a}_{i,t}(a_i, a_{j_t(i)})^2 \pi_{i,j_t(i),t}(da_i, da_{j_t(i)})$. Efficient strategies maximize $\Pi_1(\boldsymbol{\pi}_1)$ subject to the contacting frictions.

Proposition 1. *Strategies $\{s_{i,t}\}_{\forall i,t}$ are equilibrium strategies if and only if they maximize $\Pi_1(\boldsymbol{\pi}_1)$.*

Proposition 1 has three implications. First, without any deviation between private and social values, the equilibrium is efficient. Second, when a deviation arises for varied reasons, one can implement the social planner's solution through taxes by simply aligning costs. Third, it implies that the equilibrium market structure and asset allocations through the market structure are payoff unique. The multiplicity that often makes it hard to characterize financial networks does not show up in our framework. This gives the theoretical foundation to solve the trading network numerically.

3 Characterization

3.1 Variance Representation

Within a match (i, j), the posttrade positions $\tilde{a}_k(a_i, a_j)$ depend on the realized positions of the two banks (a_i, a_j). Given any allocation rule, let $\tilde{v}_k \equiv Var(\tilde{a}_k(a_i, a_j))$ denote the variance of posttrade positions and $V_{ij} \equiv Var(a_i + a_j)$ denote the variance of the sum of pretrade positions. The feasibility constraint on bilateral trade, Equation (2), implies the following connection between pretrade and posttrade risk:

$$\tilde{v}_i + \tilde{v}_j + 2\tilde{\rho}_{ij}\sqrt{\tilde{v}_i\tilde{v}_j} = V_{ij}, \tag{7}$$

where $\tilde{\rho}$ denotes the correlation of posttrade positions of two banks, which depends endogenously on the allocation rule.

Lemma 1. *The socially optimal posttrade positions must have zero mean for all banks, and the posttrade positions for any two matched banks are perfectly positively correlated. Moreover, the pretrade positions of any two matched banks in the efficient solution are uncorrelated.*

Under the quadratic utility, the aggregate payoff decreases with the variance and mean, which explains why it is optimal to maintain the mean of posttrade positions at zero and change only their correlation and variances.

Moreover, positive correlation between pretrade positions of two matched banks necessarily increases the variance of their total pretrade positions, which is the right-hand-side of the feasibility constraint for variance allocation, Equation (7). This implies that, all else equal, it is optimal to match banks with zero correlations. This observation allows us to solve the model by focusing on the variance of individual banks' positions. It also implies that it is not optimal to match two banks twice because asset positions of any two previously matched banks are positively correlated. The pretrade variance on any path of optimal matches can thus be simplified to $V_{ij} = v_i + v_j$.

Given that the asset positions for all agents are uncorrelated on the path, the sufficient statics of an agent's characteristic is his pre-trade variance $v_{i,t}$. In other words, $v_{i,t}$ is the state variable and thus, we use $W_t(v_{i,t})$ to denote the bank's maximum payoff given his characteristic $v_{i,t}$.

Corollary 1. *At each period t, within any pair (i, j), the post-trade variance $(\tilde{v}_i, \tilde{v}_j)$ maximizes*

$$\Omega_t(i,j) = \max_{\tilde{v}_i,\tilde{v}_j} \Sigma_{k=i,j} \left\{-\kappa_t \tilde{v}_k + W_{t+1}(\tilde{v}_k)\right\} \tag{8}$$
$$s.t.\ \tilde{v}_i + \tilde{v}_j + 2\sqrt{\tilde{v}_i \tilde{v}_j} = v_{i,t} + v_{j,t}.$$

In other word, we can reformulate the asset allocation problem as choosing the posttrade variances of the marginal asset distributions for banks in a match, denoted by $(\tilde{v}_i, \tilde{v}_j)$. This is because any variance allocation $(\tilde{v}_i, \tilde{v}_j)$ can be mapped to an asset allocation rule, where bank i holds a share $\alpha_i \in [0, 1]$ of total position, so that $\tilde{a}_i(a_i, a_j) = \alpha_i(a_i + a_j)$. A bank who holds a larger share of the total position will then have a higher variance on her posttrade asset position than her counterparty. In the special case where two matched banks have the same cost of absorbing risk, one would expect that they will share risks equally, i.e., $\alpha_i = \alpha_j = \frac{1}{2}$, as this minimizes the total variance within the match, $\tilde{v}_i + \tilde{v}_j$.

3.2 Risk Concentration and Interconnectedness

In our framework, not only agents change the risk allocation within the match but also whom they trade with. As shown in Corollary (1), choosing different agents results in different per-trade variance V_{ij}, which in turns affect the post-trade variance of an agent. The joint determination of these two decisions thus allows us to characterize the endogenous network effect on risk-allocation.

Allocation of Risks We first look at the risk allocation given any match. Let $v^*_\theta(V_{ij})$ denote the post-trade variance within the pair with $V_{ij} = v_{i,t} + v_{j,t}$, where $\theta \in \{h, l\}$ and $v^*_h(V_{ij}) \geq v^*_l(V_{ij})$. From Equation (8), the FOC condition yields

$$v^*_\theta(V_{ij}) = \left(\frac{\kappa_t + W'_{t+1}(v^*_{-\theta}(V_{ij}))}{\Sigma_\theta \kappa_t + W'_{t+1}(v^*_\theta(V_{ij}))} \right)^2 V_{ij}. \tag{9}$$

First of all, observe that, when $W_{t+1}(v)$ is concave, the standard predictions on risk-sharing are obtained: agents share their exposure equally with any match and thus $v^*_h(V_{ij}) = v^*_l(V_{ij}) = \frac{V_{ij}}{4}$. Moreover, since all agents share the risk equally, there is no cross-sectional dispersion of $v_{i,t}$, the matching outcome is equivalent to random matching. In this sense, the trading outcome is the same as in Afonso and Lagos (2015), which can be nested in our framework as $W_{N+1}(v) = -\kappa_{N+1} v$. [9] [10]

We thus focus on the case when $W_{t+1}(v)$ is convex throughout the rest of the paper. In this case, asymmetric risk-allocation can be optimal: Agent i unloads more risk to her counterparty j if Agent j has a lower marginal cost of risk-bearing the next period, captured by $W'_{t+1}(v^*_h(V_{ij})) \geq W'_{t+1}(v^*_l(V_{ij}))$.

In the static case with $N = 1$, the amount of concentration thus only depends on the exogenous convexity of $W_{N+1}(v)$. In our dynamic environment, the value function $W_t(v)$ further depends on the optimal choice of counterparties, which is given by

$$W_t(v_i) = \max_j \left\{ \Omega_t(i,j) - W_t(v_j) \right\}.$$

Proposition 2. *(Sorting) When $W_{N+1}(v)$ is convex in v, the optimal sorting outcome is*

[9] Afonso and Lagos (2015) predicts that post-trade exposure is given by $a^k_{t+1} = \frac{a^i_t + a^j_t}{2}$, which implies that the post-trade variance is reduced to half, $v^i_{t+1} = \frac{v^i_t + v^j_t}{4}$. Since all agents share the risk equally, their characteristics remains the same ($v^i_t = \left(\frac{1}{2}\right)^t v_0 \ \forall i$).

[10] More generally, concavity in $W_{N+1}(v)$ predicts negative sorting. Even if the economy starts with two different initial values (say half of agents start with low (high) exposure v^L_0 (v^H_0)), all agents again become homogeneous next periods under NAM.

PAM on v_t $\forall t$. When $W_{N+1}(v)$ is concave in v, the unique trading network is full risk-sharing, where $v_{i,t} = \frac{1}{2}v_{i,t-1} = \left(\frac{1}{2}\right)^t v_0$ $\forall i, t$, and the matching outcome is equivalent to random matching.

The positive sorting establishes how risks are concentrated over the networks: agents that accumulate risks from others (higher post-trade variance $v_{i,t+1}$) are matched among with each other. Hence, on the equilibrium path, $V_{ij} = 2v_i$ for any match. In particular, compare to the random matching, where the risk exposure of his counterparty next period is drawn randomly, these agents thus handle more risks on average.

Effective Risk-Bearing Costs through Connections The key object in our framework is the risk-bearing costs of an agent through his connections. While an agent might have higher cost γ_{N+1} at the final period, he can connect counterparties with lower cost γ_{N+1} and unload more risks to them. In other words, by using his network, he can effectively enjoys lower risk-bearing costs.

The risk-bearing cost of Agent i at period t should depend on his own action as well as the actions of all agents that he directly or indirectly connected. To highlight the relationship between risk-bearing costs and network, we define the network at time t as $g_t \equiv \{j_\tau(i), z_i\}_{\forall i, t \leq \tau \leq N}$. That is, a network is determined by agents' connections and their final action z_i. For any given g_t, let $\hat{W}_t(v_{i,t}|g_t)$ represent the payoff of an agent with $v_{i,t}$. Lemma 3 shows that the network effect on risk-bearing cost can be characterized recursively.

Lemma 2. *Under A1, given any g_t, $\hat{W}'_t(v_{i,t}|g_t) = \gamma_{i,t}$ $\forall t$, where the risk-capacity of agent i at period t is given by*

$$\gamma_{i,t} = \frac{1}{2} H\left(\kappa_t + \gamma_{i,t+1}, \kappa_t + \gamma_{j_t(i),t+1}\right) \ \forall t \leq N, \tag{10}$$

where $\gamma_{i,N+1} = \gamma_{N+1}(z_i)$.

Equation 10 has a simple interpretation: the risk-bearing cost of Agent i at period t is the harmonic mean[11] of the post-trade risk-bearing cost of Agent i and her counterparty $j_t(i)$. It also shows that, while two matching agents can have different capacity next

[11] The harmonic mean of any two variables γ_j and γ_j is $\frac{2}{\gamma_i^{-1}+\gamma_j^{-1}}$.

Figure 1: Late vs. early Concentration ($N = 2$)

period, they must have the same capacity at period t, given they allocate the risks jointly, taking into their future connections.

For any given network g_t, we have now characterized agents' risk-capacity and hence the corresponding risk allocations among this network. This is because the optimal risk-allocation within any pair must satisfy the FOC condition (Equation (9)), where $\hat{W}'_{t+1}(v_{i,t+1}|g_{t+1})$ is valued at $\gamma_{i,t+1}(g_{t+1})$ (given by Lemma 3) and $V_{ij} = 2v_{i,t}$ (given PAM). Next, we analyze which network is optimal.

3.3 Properties for Optimal Network

We now turn to analyze the equilibrium network, which, according to Prop 1, is unique. Specifically, given that $W_t(v)$ is convex, the set of networks characterized in Section 3.2 corresponds to the set of multiple local optimal solutions.[12]

To given a concrete example, 1 considers binary actions with $N = 2$ with four banks. Agents 3 and 4 (Agent 1 and 2) chooses an action with lower (higher) risk-bearing costs, denoted by γ^l_{N+1} (γ^h_{N+1}) at period $N + 1$. For example, only Agent 3 and 4 have access to a centralized market at period $N + 1$.

These two graphs differ in terms of their dynamic bilateral connections $\{j_t(i)\}$. In the left graph of Figure 1, an agent is first connected with another with the same platform access (i.e., $\{(1, 2), (3, 4)\}$ at $t = 1$) and then connected with another agent with with different platform access i.e., $\{(1, 3), (2, 4)\}$at $t = 2$). This order is reversed in the right graph.

For both graphs, the effective risk-bearing costs for all agents at period $t = 1$ is given by Lemma 2, which crucially depends on the timing of the matching plan as it results in different $\gamma_{k,t+1}$. To see this, in the illustrative example, since agents with different access

[12]Under A1, one can see that second order conditions are also satisfied, which yields $\left\{-\kappa_t + W'_{t+1}(\tilde{v}_j)\right\}(\tilde{v}_i)^{-\frac{3}{2}}\sqrt{V_{ij}} < 0$.

are matched at period 2 in the left graph, their effective risk-bearing capacity at period 2 is thus symmetric, with $\bar{\gamma}_{t+1} = \frac{1}{2}H\left(\kappa_{t+1} + \gamma^h_{N+1}, \kappa_{t+1} + \gamma^l_{N+1})\right)$.

In the right graph, on the other hand, since agents with the same access are matched at period 2, their effective risk-bearing capacity at period $t = 2$ are thus heterogeneous, with $\gamma^\theta_{t+1} = \frac{1}{2}\left(\kappa_{t+1} + \frac{1}{2}H(\gamma^\theta_{N+1}, \gamma^\theta_{N+1})\right)$ and $\theta \in \{h, l\}$. One can show that the the effective risk-bearing cost at period $t = 1$ under the right graph is higher than the left graph, as for any $\kappa_t > 0$,

$$\frac{1}{2}H(\kappa_t + \gamma^l_{t+1}, \kappa_t + \gamma^h_{t+1}) > \frac{1}{2}\left(\kappa_t + \bar{\gamma}_{t+1}\right).$$

Intuitively, the right graph results in more asymmetric γ_{t+1} and thus early risk concentration. The left graph, on the other hands, implies more symmetric γ_{t+1}. For any $\kappa_t > 0$, early risk-concentration is dominated by late concentration. Hence, it's optimal to have a more symmetric risk capacity γ_{t+1}, which in turn lowers the needs for risk concentration at period t and thus results in lower effective risk-bearing cost γ_t.

In the special case where $\kappa_t = 0\ \forall t$, one can show that ordering no long matters. The effective risk-bearing cost in this case is reduced to the harmonic mean of the risk-capacity at $N + 1$ of all connected agents (i.e., $\gamma_{i,t} = \left(\Sigma_k \frac{1}{\gamma_{k,t+2}}\right)^{-1}$) and thus different timing does not matter. $g_t \equiv \{j_\tau(i), z_i\}_{\forall i, t \le \tau \le N}$.

Lemma 3 generalizes the result for any period t. Fixing any network g_{t+2} and thus $\gamma_{k,t+2}$, the optimal matching at period $t + 1$ must be such that agents have similar $\gamma_{k,t+1}$ at period $t + 1$, which implies negative sorting on $\gamma_{k,t+2}$.

Lemma 3. *Under A1, for any $v_{k,t+2}(v_t)$ and the corresponding $\gamma_{k,t+2}$, any matching plan that violates negative sorting on $\gamma_{k,t+2}$ is dominated.*

3.4 Full Characterization for Binary Choices

We now show that, with binary choices $z_i \in \{0, 1\}$, Lemma 3 pins down the unique connections, given the set of agents that choose $z_i = 1$. We refer these agents are core agents $i \in C$ iff $z_i = 1$, as these agents have lower costs of holding risks and thus would absorb more risks and thus have higher trading volume from his counterparties.

Optimal Connections to Cores Give g_t, let $c_{i,t}$ denote the number of core agents that Agent i is directly and indirectly connected, which is defined recursively as $c_{i,t} =$

$c_{i,t+1} + c_{j_t(i),t+1}$,with $c_{i,N+1} = z_i$. We thus refer $c_{i,t}$ as the core access, in the sense that agents have higher $c_{i,t}$ at time t are connected to more core agents. By definition, only the core agent has access at period $N+1$. However, for any $t \leq N$, any non-core agent can obtain core access through bilateral connections.

In general, the risk capacity of Agent i at time t generally depends on the underlying network g_t according to Equation 10. We now shows that, given any $c_{i,t}$, there is an unique network that satisfies Lemma 3. Hence, the core access c_{it} becomes the sufficient statistic. We thus use $\gamma_t^*(c_{i,t})$ denote the risk capacity for agent with core access $c_{i,t}$ at time t under the optimal connections.

Lemma 4. *With binary actions, core access $c_{i,t}$ is the sufficient statics for agent i's risk capacity at period t, where $\gamma_t^*(c_{i,t})$ decreases in $c_{i,t}$ and*

$$\gamma_t^*(c_{i,t}) = \frac{1}{2} H\left(\kappa_t + \gamma_{t+1}^*(\lfloor \frac{c_{i,t}}{2} \rfloor), \kappa_t + \gamma_{t+1}^*(\lceil \frac{c_{i,t}}{2} \rceil)\right) \; \forall t \leq N. \tag{11}$$

Note that, by definition, if two agents (i,j) are matched at period t, they have the same pre-trade core-access, $c_{i,t} = c_{j,t}$, as it represents the sum of their core access next period. Equation 11 thus implies that the optimal core access must be distributed evenly within the pair. That is, $c_{i,t+1} = \lfloor \frac{c_{i,t}}{2} \rfloor$ and $c_{j,t+1} = \lceil \frac{c_{j,t}}{2} \rceil$. The intuition is the same as shown in the illustrative example in Figure 1. Negative sorting on $\gamma_{k,t+2}$ minimizes the difference in $\gamma_{k,t+1}$ and thus reduces the needs for earlier concentration.

Optimal Core Size Since we have established that there is a unique optimal market structure given any core size c. Recall that, an agent i can connect, directly or indirectly, to at most 2^N agents in N rounds of trade. The problem can thus be solved as if solving the optimal network among 2^N agents. To map the finite network among 2^N agents to the market structure for a continuum of agents of total measure 1, we interpret our results here as if there were 2^N "types" of agents and each had a measure of $\frac{1}{2^N}$. Thus, if there are c cores among 2^N agents, the total measure of core agents would be $\frac{c}{2^N}$. Then, there are $1/2^N$ identical replica of the finite network of size 2^N.

The optimal network can be further reduced to choosing the number of core agents in the beginning of the trading game among 2^N agents, which can be expressed as

$$\Pi = \max_c \left\{ -\gamma_1^*(c) v_1 - \frac{c}{2^N}\phi \right\}. \tag{12}$$

Given any c, $\gamma_1^*(c)$ represents the risk capacity for all agents, taking into account the future connections. It again highlights that while each agent might have asymmetric access over time, their effective risk exposures are the same ex ante, as they optimize jointly the allocation over their core access.

Proposition 3. *(Optimal Network and Risk-Allocation with Binary Actions) All agents start with core access* $c_{i,1} = c^* \; \forall i$*, where the optimal core size* c^* *solves Equation* (12)*. For any two matching matching* i *and* j *at period* t*, they must have the same core access* $c_{i,t} = c_{j,t}$ *and their post-trade core access are adjacent integers,* $c_{i,t+1} = \lfloor \frac{c_{i,t}}{2} \rfloor$ *and* $c_{j,t+1} = \lceil \frac{c_{j,t}}{2} \rceil$*. Their posttrade variance is given by* $v_{i,t+1} = \left(\frac{\kappa_t + \gamma_{t+1}^*(c_{j,t+1})}{\Sigma_k \left(\kappa_t + \gamma_{t+1}^*(c_{k,t+1}) \right)} \right)^2 (2v_{i,t})$.

4 Application 1: Platform Access

Many financial over-the-counter (OTC) markets operate as classical two-tiered markets where a few core banks have exclusive access to an exchange-like interdealer market. Such a structure have been the focus of regulation and policy debates after the 2007-08 financial crisis. In particular, post crisis reforms have increased dealer banks' balance sheet costs through tightened capital requirements and additional liquidity requirements and have promoted all-to-all exchanges.[13]

We now apply our framework to study the positive and normative implications of reforms, taking into the equilibrium response of the market structure. Specifically, the underlying environment is an example of having a binary action, where $z_i = 1$ if banks enter the platform and zero otherwise.

We think of the trading platform as a superior technology but more expensive trading technology. That is, it gives better risk-sharing and thus reduce the cost of holding risks $\gamma_{N+1}(1) < \gamma_{N+1}(0)$. We normalize the cost of a bilateral relationship to be zero, and assume a fixed cost of using platform, which can be interpreted as the additional cost of

[13]See detailed discussions in Yellen (2013) and Duffie (2018).

trading multilaterally with a larger set of market participants.[14] Thus, $\phi_{N+1}(1) = \phi > \phi_{N+1}(0) = 0$.

Remark 1. More generally, the usage cost can have variable components beyond the fixed cost. For example, consider the required collateral may be higher with larger positions (given by $\phi_c v$). This would effectively lead to higher $\gamma_{N+1}(1)$.

Remark 2. While the timing of our framework implies that the platform entry is at the end, this assumption can be relaxed as long as there is a fixed cost associated with each entry. If there is no delay cost, it is indeed optimal to postpone the access until the end, as agents would prefer to accumulate as much risk as possible from bilateral trades first before joining the platform. [15]

4.1 Equilibrium Response of Market Structure

We model the polices that promote central clearing and/or discourage risk taking as providing subsidy of platform participation and/or taxing banks' net exposure. In other words, the policy can be understood as increasing κ_t (i.e., making it more costly for banks to hold risks) and/or decreasing the entry cost of the platform (ϕ).

Since these policies change agents' incentives to hold risks and/or the entry cost, the equilibrium response can thus be understood through comparative statics on κ_t and ϕ.[16] Importantly, agents in our framework can respond in two margins. First, for fixed agents' connections g_t, the asset and risk allocation can differ. Such a change is hence similar to the existing literature with exogenous networks.

The key advantage of our framework is that agents can change their connections and

[14]This cost can be interpreted as a fixed cost of setting up the platform or more stringent regulatory requirements, membership, or a collateral requirement associated with a more centralized market. Indeed, as summarized in the report "Incentives to centrally clear over-the-counter (OTC) derivatives"by Board 2018, "almost all respondents agreed that high fixed costs make clearing services expensive, and that these costs are substantially higher than the benefits that a smaller participant might accrue from central clearing". The report further mentions that "survey responses state that providing connectivity to CCPs requires incurring high fixed costs, which are likely passed on to clients through minimum fees and other charges, increasing clients' costs of central clearing. For smaller, lower activity clients in particular, this can raise their cost of cleared trades, and thus can have a material impact on their incentives to centrally clear."

[15]If delay costs are sufficiently high, agents might choose to obtain access earlier at period $n^* < N$. In this case, the result can be understood as applying our characterization for n^* rounds of bilateral trades together with $N - n^*$ rounds of random matching.

[16]Recall that $\kappa \equiv \kappa_{N+1}$ and $\kappa_t = \delta\kappa \; \forall t \leq N$. Since we assume that the tax τ^κ applies to all periods, for private agents it is equivalent to a higher κ.

access optimally. Moreover, since our predicts that the market structure is unique and the core size is the sufficient statics. The change in the market structure, which includes the set of agents who choose to have platform access and peripheral connections, can be summarized by the core size at the aggregate level.

To explore how the core size depends on the underlying parameters, we now assume that $\kappa_t = \delta\kappa \ \forall t \leq N$ and $\kappa_{N+1} = \kappa$. In other words, parameter δ captures the cost of holding risk in an earlier period relative to the terminal period. Given any δ, one can show that the risk-capacity $\gamma_t^*(c)$ is a homogeneous function of degree 1 in κ. Hence, the optimal core size depends on the entry cost *relative* to balance sheet costs, $\frac{\phi}{\kappa v_1}$. Since agents face the trade-off between the cost of risk concentration and that of entry, the model thus predicts a (weakly) larger core size with a lower relative entry costs.

Proposition 4. *Given any δ, the optimal measure of cores (weakly) decreases with $\frac{\phi}{\kappa v_1}$. If $\delta = 0$, the optimal core size is $\frac{1}{2^N}$ if and only if $\phi < \kappa v_1$, and is zero otherwise. If $\delta \to \infty$, the optimal core size is 1 if and only if $\phi < \left(\frac{1}{2}\right)^N \kappa v_1$, and is zero otherwise.*

Figure 2 illustrates the change in the market structure before and after such a policy, which induces an increase in participation in the central platform (i.e., a larger core size), as tax and subsidy effectively increase (decrease) the private cost of holding risk (respectively, entry).

Our model predicts that the structure becomes more symmetric; nevertheless, the two-tier market structure persists. This explains why, as discussed in Collin-Dufresne, Junge, and Trolle (2018) and Duffie (2018), all-to-all trading has not materialized and the provision of clearing services remains concentrated.

Moreover, as the size of cores increases, banks transit from risk-concentrating, market-making trades towards risk-sharing trades. Since trades among customers share risks on asset positions symmetrically and have zero spread, such a structural change could result in lower average transaction costs despite the increase in the spread that market-makers charge.

Our prediction is consistent with the empirical findings in Choi and Huh (2018) and rationalizes the seemingly contradicting evidence in the post-Volcker rule era.[17] The stan-

[17]Bao, O'Hara, and Zhou (2016) and Bessembinder et al. (2018) show that the Volcker rule leads to lower inventories and capital commitment for bank-affiliated dealers. Such a decline, however, does not worsen the overall market liquidity measured by the bid-ask spread.

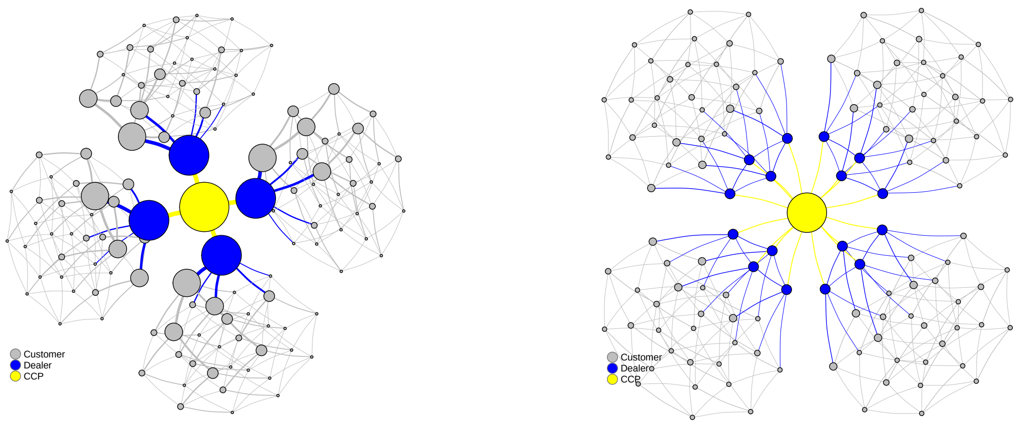

Figure 2: Pre vs. Post-regulation Market Structure.
Each panel shows the graph of the equilibrium trading network. In the network graph, each node represents a bank. The area of the node represents the gross trading volume involving the bank. The edges between nodes represent bilateral trading relationships. The width of an edge represents the bilateral trading volume. The left panel illustrates the pre-regulation market structure. The right panel illustrates the post-regulation market structure with increased balance sheet costs and lowered cost of accessing the centralized trading platform.

dard results that banks' balance sheet cost increases the bid-ask spreads and transaction costs may not hold when the market structure changes in response. Our result further suggests that under an endogenous market structure, transaction costs are generally no longer a sufficient measure of welfare.

4.2 Normative Implications

Concentration can be Efficient Our results highlight that the optimal intervention should not be targeting all-to-all trading or reducing risk concentration because the existence of exclusive core members and a high concentration of risks and volume *can* be efficient.

Proposition 4 presents two extreme cases that further highlight the benefit of having an asymmetric market structure. Such a benefit is highest if there is no cost of risk concentration before the terminal period ($\delta = 0$). In this case, it is optimal to concentrate all variance in the core and thus predicts highest risk-concentration.

Welfare-maximizing Policy More generally, the optimal intervention in our model is very simple: whenever there are frictions that lead to a deviation between private incentives of risk taking and entry-cost, the equilibrium can be inefficient.

Our model also provides a simple guideline for correcting such an inefficiency (if it exists), taking into account the equilibrium response of the trading network. According to Proposition 1, the optimal policy policy (such as tax and/or subsidy) can restore the efficient market structure by aligning private and social value of risk-taking and/or entry.

For example, one common concern is that the platform might be controlled or entrenched in by the incumbent dealers. One can capture this in our environment by assuming that a set I_0 of agents with exogenous measure $\frac{c_0}{2^N}$ have built relationships among themselves and collectively operate the trading platform at cost ϕ. The incumbent agents jointly own the platform and decide whether to charge a new entrant to the platform an exogenous fee $\Delta > 0$.

Given any fee, this setup can thus be understood as our trading game with heterogeneous costs ϕ_i where $\phi_i \equiv \phi + \Delta$ for potential entrants $i \notin I_0$ and $\phi_i = \phi$ for incumbent banks $i \in I_0$. That is, the incumbent cores have a lower entry cost than the rest of the market. The existence of the fee thus generate the wedge between private and social value of platform.

Our model thus predicts that by setting the subsidy for entry so that $c^*(\phi + \Delta - s^c) = c^*(\phi)$, or introducing a new platform with entry cost ϕ will restore the efficient market structure.

5 Application 2: Limited liability

A prevalent concern in financial intermediation is the risk-taking incentive that results from limited liability. We use our framework to show that such incentives can be amplified through the trading network.

We assume that the loss to a bank is capped by ϕ when it defaults because of limited liability. Banks' final payoffs are thus specified as

$$W_{N+1}(v) = -[(1 - p(v))\kappa_{N+1}v + p(v)\phi],$$

where $p(v)$ denotes the probability of default when the variance of the bank's asset holding is v.

One can see that when the probability of default is increasing in the uncertainty of the asset holding $p'(v) > 0$, $W_{N+1}(v)$ can become convex.[18] This formulation thus captures the well known channel that limited liability encourages risk-taking.

Regime Shifts in Trading Networks Our framework demonstrates how the risk-taking incentives affect the trading network. Since the network can be asymmetric, the effect on each bank thus depends on their network positions.

Formally, this is captured by banks' risk positions $v_{i,N}$, which depends on the outcomes of bilateral trades. Without any interconnectedness or concentration, all banks will have relative low and symmetric risk positions and thus such incentives might not be relevant.

When banks can use their networks to shift risks, a small change in such incentives at the individual level can shift the aggregate network from sharing risks to concentrating risks. It thus can generate a discontinuously large increase in aggregate default probability.

We think of lower convexity as the normal regime (represented by the red line): banks choose to share risks and thus each of them has low final risk exposure and default probability.

Facing a higher convexity, it becomes optimal for banks to concentrate risks to a few banks (denoted by the blue line), which results in higher aggregate probability of default (which is proportional to the total variance). In this sense, our model predicts that a small increase in risk-taking incentives can trigger a financial crisis through the network connections.

Normative Implications In this application, any risk-taking is inefficient from viewpoint of planner, as default effectively offloads downside risks to outside creditors. Since the planner prefers risk-sharing, for the similar logic as before, the efficient network can be restored by increasing the cost of holding risks – such as setting a tax to increase banks' flow costs of holding risks $\kappa_t(1 + \tau^\kappa)$.

[18] This is true as long as $W''_{N+1}(v) \propto 2p'(v)\kappa_{N+1} - p''(v)\phi > 0$.

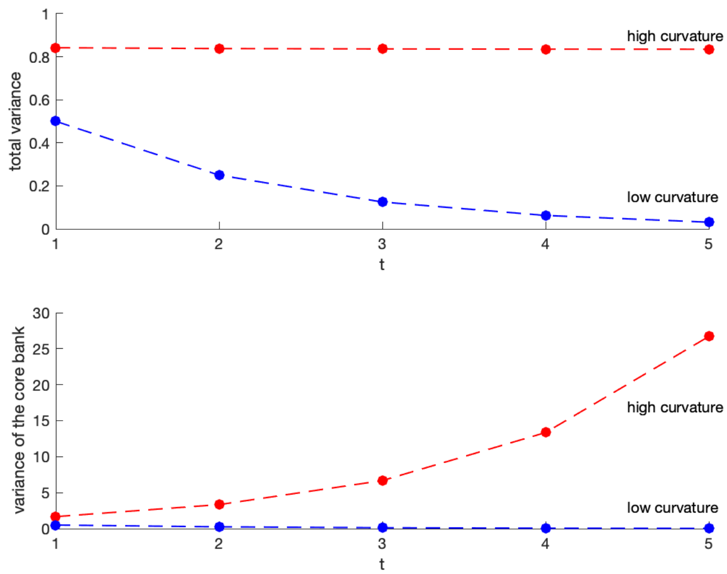

Figure 3: Regime Shift: $W_{N+1}(v) = -1 + e^{-cv}$ and $\frac{c_H}{c_L} = \frac{1.1097}{1.1096}$

Relation to Systematic Risk in Networks In the existing literature on financial networks, banks use their links to diversify the risks, while the systemic risk could arise from cascading failures among banks interconnected through a predetermined financial network. We point out that, apart from the ex post contagion, the aggregate default risk can increase as banks can change their risk-taking behaviors by changing how banks are connected and concentrate risks ex ante.

6 Conclusions

In this paper, we develop a tractable framework of endogenous trading networks and use it to analyze how the market structure may respond to underlying parameters and/or regulatory changes. Exactly because banks can accumulate risks from others, any policy must take into account the network effect of risk-taking behaviors among banks. Although the network structure seems complex, our framework provides a tractable and unique characterization as well as a simple guideline for possible interventions when private incentives are distorted relative to the social cost.

A Appendix: Omitted Proofs

A.1 Efficiency and Uniqueness

Because agents' utility is quadratic in their asset holding, only the mean and variance of a distribution are relevant to their payoff. In general, we can represent the joint distribution by the means and variances of agents' asset holdings and covariances between their asset holdings. To do this, we first show that it is optimal to keep the means of individual asset holding at zero. We then show that it is optimal to match agents whose asset holdings are not correlated.

Because agents have quasilinear utility, Pareto optimal allocations are the solution to a simple social planner's optimization problem where the planer maximizes the present value of total utility of the economy. The planner's choices at period t include any agent i's counterparty $j_{i,t}$, asset allocation within a match, $\tilde{a}_{i,t+1}(a_{i,t}, a_{j_{i,t},t})$ and $\tilde{a}_{j_{i,t},t+1}(a_{i,t}, a_{j_{i,t},t})$. The planner chooses period-t counterparties given period-0 information and asset distribution at period t. The planner's value function at period t has the joint asset distribution across agents as its state variable and can be characterized as

$$\Pi_t(\pi_t) = -\int \kappa_{i,t} E_t(\tilde{a}_{i,t}^2(a_{i,t}, a_{j_{i,t},t}))di + \beta \Pi_{t+1}(\pi_{t+1}), \text{ for } t \leq N,$$

$$\Pi_{N+1}(\pi_{N+1}) = \int \max\{-\phi_{i,t}, -E_{N+1}(a_{i,N+1}^2)\kappa_{i,N+1}\}di.$$

The constraints that the planner faces include:

(1) Given π_t, the planner's period-t is feasible if and only if

$$\int_0^i \Pr(j_{\iota,t} \leq \iota)d\iota \leq i, \tag{A.1}$$

$$\tilde{a}_{i,t}(a_i, a_{j_{i,t}}) + \tilde{a}_j(a_i, a_{j_{i,t}}) = a_i + a_{j_{i,t}}, \tag{A.2}$$

where (A.1) is the feasibility constraint of the matching allocation of the planner, $\Delta(\pi_{i,t})$ refers to the support of the marginal distribution $\pi_{i,t}$; (2) The joint distribution evolves consistently with the counterparty assignment and within match asset allocations.

Lemma 5. *It is optimal to keep the means of individual asset holding at zero.*

Proof. Because the utility function of the agent is quadratic, the marginal asset distribution for Agent i enter the social planner's objective through its expected value and variance. Denote $E_t a_{i,t} = m_{i,t}$, $E_t(a_{i,t} - m_{i,t})^2 = v_{i,t}$ and $\rho_{i,j,t} = \frac{Cov(a_{i,t+1}, a_{j,t+1})}{\sqrt{v_{i,t+1}v_{j,t+1}}}$ for all i, j, and t. Let $\boldsymbol{m}_t =$

$\{m_{i,t}\}_{\forall i}$, $\boldsymbol{v}_t = \{v_{i,t}\}_{\forall i}$, $\boldsymbol{\rho}_t = \{\rho_{i,j,t}\}_{\forall i,j}$. Then the period-$t$ state variable of the social planner can be summarized by $(\boldsymbol{m}_t, \boldsymbol{v}_t, \boldsymbol{\rho}_t)$.

The planner's objective function is then

$$\Pi_t(\boldsymbol{m}_t, \boldsymbol{v}_t, \boldsymbol{\rho}_t) = -\int \kappa_{i,t}\left(m_{i,t+1}^2 + v_{i,t+1}\right) di + \beta \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1}), \text{ for } t \leq N, \tag{A.3}$$

$$\Pi_{N+1}(\boldsymbol{m}_{N+1}, \boldsymbol{v}_{N+1}, \boldsymbol{\rho}_{N+1}) = \int \max\{-\phi_i, -(m_{i,N+1}^2 + v_{i,N+1})\kappa_{i,N+1}\} di, \tag{A.4}$$

given optimal choices for $(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})$. The choices at period $N+1$ are obvious: the planner chooses to access multilateral clearing for Agent i if and only if $(m_{i,N+1}^2 + v_{i,N+1})\kappa_{i,N+1} > \phi_i$.

The feasibility of within-match asset allocation between agent i and her counterparty j implies that $a_{i,t+1} + a_{j,t+1} = a_{i,t} + a_{j,t}$ for all $t \leq N$, which is translated into two separate constraints for the mean and the variance of asset allocation to Agents i and j

$$m_{i,t+1} + m_{j,t+1} = m_{i,t} + m_{j,t}, \tag{A.5}$$

$$v_{i,t+1} + v_{j,t+1} + 2\sqrt{v_{i,t+1}v_{j,t+1}}\rho_{i,j,t+1} = v_{i,t} + v_{j,t} + 2\sqrt{v_{i,t}v_{j,t}}\rho_{i,j,t}. \tag{A.6}$$

Notice that the choice over the expected asset holding is subject to a separate constraint, (A.5), from the choice over its variance, (A.6). And the law of motion of asset holding variance and correlation does not depend on the expected asset holding.

The planner's optimization problem at period t can be summarized by the following Lagrangian,

$$\begin{aligned}\mathcal{L}_t(\boldsymbol{m}_t, \boldsymbol{v}_t, \boldsymbol{\rho}_t) = & -\int \kappa_{i,t}\left(m_{i,t+1}^2 + v_{i,t+1}\right) di + \beta \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1}) \\ & + \int \lambda_{i,j_{i,t},t}^m \left(m_{i,t} - m_{i,t+1}\right) di \\ & + \int \lambda_{i,j_{i,t},t}^v (v_{i,t} + \sqrt{v_{i,t}v_{j_{i,t}t}}\rho_{i,j_{i,t},t} - v_{i,t+1} - \sqrt{v_{i,t+1}v_{j_{i,t+1},t+1}}\rho_{i,j_{i,t},t+1}) di\end{aligned} \tag{A.7}$$

for all $t \leq N$, where $\lambda_{i,j_{i,t},t}^m$ refers to the Lagrangian multiplier for constraint (A.5) for agent i and his counterparty $j_{i,t}$, $\lambda_{i,j_{i,t},t}^v$ refers to the Lagrangian multiplier for constraint (A.6).

For period $N+1$, $\frac{\partial \Pi_{N+1}(\boldsymbol{m}_{N+1}, \boldsymbol{v}_{N+1}, \boldsymbol{\rho}_{N+1})}{\partial m_{i,N+1}}$, $\frac{\partial \Pi_{N+1}(\boldsymbol{m}_{N+1}, \boldsymbol{v}_{N+1}, \boldsymbol{\rho}_{N+1})}{\partial v_{i,N+1}} \leq 0$ and $\frac{\partial \Pi_{N+1}(\boldsymbol{m}_{N+1}, \boldsymbol{v}_{N+1}, \boldsymbol{\rho}_{N+1})}{\partial \rho_{i,j,N+1}} = 0$ for all i, j.

Using mathematical deduction, we can then show that $\frac{\partial \Pi_t(\boldsymbol{m}_t, \boldsymbol{v}_t, \boldsymbol{\rho}_t)}{\partial m_{i,t}} \leq 0$ for all i and all $t \leq N$, where the inequality is strict if and only if there exits $t \leq t' \leq N$ such that $\kappa_{t'} > 0$. This is because given the counterparty choices, $j_{i,t}$, the first order condition with respect to $m_{i,t+1}$

implies that $\lambda^m_{i,j_{i,t},t} < 0$ when $\kappa_t > 0$ or $\frac{\partial \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial m_{i,t+1}} < 0$.

The effect of within-match asset allocation on Agent i's expected asset holding can be summarized by $\alpha^m_{i,t}$, such that $m_{i,t+1} = \alpha^m_{i,t}(m_{i,t} + m_{j,t})$, $m_{j,t+1} = (1 - \alpha^m_{i,t})(m_{i,t} + m_{j,t})$. If $\frac{\partial \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial m_{i,t+1}} < 0$, it is clear that $\alpha^m_{i,t}$ should be between 0 and 1. If $\alpha^m_{i,t}$ were greater than 1 or less than 0, the planner can strictly increase either agent i or her counterparty $j_{i,t}$'s marginal contribution to the planner's period t objective function without reducing other agents' contribution. For example, if $\alpha^m_{i,t} > 1$, by setting $\alpha^m_{i,t}$ to 1 reduces $m^2_{i,t+1}$ to $(m_{i,t} + m_{j,t})^2$ and $m^2_{j_{i,t},t+1}$ to 0. If $\frac{\partial \Pi_t(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial m_{i,t+1}} = 0$, but $\kappa_{i,t} > 0$, the same argument applies so that $0 \leq \alpha^m_{i,t} \leq 1$. If $\frac{\partial \Pi_t(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial m_{i,t+1}} = 0$, and $\kappa_{i,t} = 0$, it is without loss to the social planner to impose $0 \leq \alpha^m_{i,t} \leq 1$.

Because the expected value of agents' initial marginal asset distribution is zero, the fact that $0 \leq \alpha^m_{i,t} \leq 1$ implies that $m_{i,t} = 0$ for all i and all period. □

Lemma 5 is the first step in characterizing the efficient asset allocation. It implies that the socially optimal asset distribution in any period can be represented by the variance of individual agents' asset holdings and the correlation of their asset holdings.

Lemma 6. *In the socially optimal matching assignments and asset allocations, the post trade asset holdings of two matched Agents i and j are perfectly correlated, and the planner always match agents with uncorrelated asset holding. That is, $\rho_{i,j_{i,t},t} = 0$, and $\rho_{i,j_{i,t},t+1} = 1$, for any agent i and their optimal counterparty $j_{i,t}$.*

Proof. The proof takes two steps. First, we show that if $\rho_{i,j_{i,t},t} = 0$ for for any agent i and their optimal counterparty $j_{i,t}$, it is optimal to have within match asset allocation perfectly correlated.

If $\rho_{i,j_{i,t+1},t+1} = 0$, then for all i, j such that $\rho_{i,j,t+1} > 0$, we can show by differentiating the planner's Lagrangian, (A.7), that $\frac{\partial \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial \rho_{i,j,t+1}} = 0$. Following similar argument to that in the proof for Lemma 5, we can see that the marginal value of increasing an agent's variance is negative $\frac{\partial \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial v_{i,t+1}} \leq 0$.

The feasibility of within-match asset allocation implies that variances of asset allocations satisfy (A.6). According to (A.6), increasing the correlation between the asset allocations to matched agents reduces the total variance of asset allocation to them, $v_{i,t+1} + v_{j_{i,t},t+1}$. Because $\frac{\partial \Pi_{t+1}(\boldsymbol{m}_{t+1}, \boldsymbol{v}_{t+1}, \boldsymbol{\rho}_{t+1})}{\partial \rho_{i,j,t+1}} = 0$, it is then optimal to set $\rho_{i,j_{i,t},t+1} = 1$.

The second step is to show $\rho_{i,j_{i,t},t} = 0$. Because the initial asset holdings are not correlated, if $\rho_{i,j_{i,t},t+1} = 1$, then the asset allocations are either uncorrelated or perfectly positively correlated. Because there is a continuum of agents in the economy, for any agent i, if the planner is to match him with an agent with variance v', there always exists such an agent whose asset holdings are uncorrelated with agent i. According to (A.7), this shadow value of $\rho_{i,j_{i,t},t}$ equals $\lambda^v_{i,j_{i,t},t}$, which is weakly negative. It is then optimal to match two agents whose asset holdings are not correlated.

□

Lemma 6 implies that even though agents have the option to trade repeatedly with a counterparty, repeated trade without receiving new asset holding shocks is suboptimal. Trading once, the asset holdings of Agent i and the counterparty become positively correlated. Then, trading twice is dominated by trading with a new counterparty with the same asset holding variance but whose asset holding is not correlated with Agent i's. Thus, we can characterize the equilibrium using a representation of the aggregate asset holding distribution by the variances of individual agents' asset holding distribution.

A.2 General Properties

A.2.1 Proof for Proposition 2

Proof. Given any variance constraint V_{ij}, the optimization problem can be expressed as

$$\Omega_t(V_{ij}) = \max_{\tilde{v}_i} \underbrace{-\kappa_t \left\{V_{ij} - 2\sqrt{\tilde{v}_i}\left(\sqrt{V_{ij}} - \sqrt{\tilde{v}_i}\right)\right\} + W_{t+1}(\tilde{v}_i) + W_{t+1}\left(\left(\sqrt{V_{ij}} - \sqrt{\tilde{v}_i}\right)^2\right)}_{\equiv F(\tilde{v}_i)}, \tag{A.8}$$

where FOC yields

$$\frac{dF(\tilde{v}_i)}{d\tilde{v}_i} = -\kappa_t + W'_{t+1}(\tilde{v}_i) + \left\{-\kappa_t + W'_{t+1}(\tilde{v}_j)\right\}\left(1 - \sqrt{\frac{V_{ij}}{\tilde{v}_i}}\right) = 0.$$

Since the allocation within a pair can be understood as choose the ratio of total variance, Equation A.8 can be further rewritten as

$$\Omega_t(V) = \max_{\omega} \left\{\kappa_t \left(\omega^2 V\right) + W_{t+1}(\omega^2 V) + \kappa_t \left(((1-\omega)^2 V\right) + W_{t+1}((1-\omega)^2 V)\right\}.$$

Let $\omega^*(V) = \frac{\left(-\kappa_t + W'_{t+1}(v^*_j(v,v'))\right)}{\Sigma_k\left\{-\kappa_t + W'_k(v^*_k(v,v'))\right\}}$ denote the solution that satisfies the FOC condition.

Convex $W_{N+1}(v)$ Given that $V_{ij} = v_i + v_j$, to establish PAM, it is sufficient to show that $\Omega_t(V)$ is convex in V $\forall t$. Let $\omega = \omega^*(V)$ denote the optimal allocation under V.

$$\begin{aligned}
&\Omega_t(\lambda V) + \Omega_t((1-\lambda)V) \\
\geq&\kappa_t \left\{(\omega^2 + (1-\omega)^2)V\right\} + W_{t+1}(\omega^2 \lambda V) + W_{t+1}((1-\omega)^2 \lambda V) \\
&+ W_{t+1}(\omega^2(1-\lambda)V) + W_{t+1}((1-\omega)^2(1-\lambda)V) \\
\geq& \left\{\kappa_t(\omega^2 + (1-\omega)^2)V + W_{t+1}\left(\omega^2\left(\lambda V + (1-\lambda)V\right)\right) + W_{t+1}\left((1-\omega)^2\left(\lambda V + (1-\lambda)V\right)\right)\right\} = \Omega_t(V).
\end{aligned}$$

where the first inequality follows that the surplus under optimal allocation $\omega^*(\lambda V)$ and $\omega^*((1-\lambda)V)$ is higher than using the allocation rule $\omega^*(V)$. The second follows that $W_{t+1}(v)$ is convex in v, which is true for $W_{N+1}(v)$. Assume that $W_{t+1}(v)$ is convex, it thus implies that $\Omega_t(V_{ij})$ is convex in $V_{ij} = v_i + v_j$. Moreover, since

$$W_t(v_i) = \max_j \left\{\Omega_t(v_i + v_j) - W_t(v_j)\right\},$$

it thus shows that $W_t(v)$ is convex in v $\forall t$. Hence, by backward induction, $\Omega_t(v_i + v_j)$ is convex in $v_i + v_j$ and hence PAM $\forall t$.

Concave $W_{N+1}(v)$ Observe that if $W_{t+1}(v)$ is concave, we thus have $F''(v) = W''(\tilde{v}_i) + W''_{t+1}(\tilde{v}_j)\left(1 - \sqrt{\frac{V_{ij}}{\tilde{v}_i}}\right)^2 + \left\{-\kappa_t + W'_{t+1}(\tilde{v}_j)\right\}\frac{\sqrt{V_{ij}}}{(\tilde{v}_i)^{\frac{3}{2}}} < 0$, and thus the unique global optimal is given by $\tilde{v}_i = \tilde{v}_j = \frac{v_i + v_j}{4}$.

For a similar logic, we now show that the sorting is NAM, as $\Omega_t(V)$ is concave when $W_{N+1}(v)$ is concave.

$$\begin{aligned}
\Omega_t(\lambda V) + \Omega_t((1-\lambda)V) &= \kappa_t \left\{\frac{V}{4}\right\} + 2W_{t+1}(\frac{1}{4}\lambda V) + 2W_{t+1}(\frac{1}{4}(1-\lambda)V) \\
&\leq \kappa_t \left\{\frac{V}{4}\right\} + 2W_{t+1}(\frac{1}{4}V) = \Omega_t(V),
\end{aligned}$$

where the inequality uses the fact that $W_{t+1}(V)$ is concave in V. While the sorting is generally NAM in this case, given that all agents start with the same variance v_0, we thus have $v_{i,t+1} = \frac{v_{i,t} + v_{j,t}}{4} = \frac{v_{i,t}}{2}$.

□

A.2.2 Proof for Lemma 2

Proof. At period N, given $\gamma_{i,N+1}$, the optimal allocations within any matched pair (i,j) the optimal allocations within any matched pair (i,j) thus solve, for $t=N$,

$$\Omega_t(i,j) = \max_{\tilde{v}_k} -\Sigma_{k=i,j}\left\{\kappa_t \tilde{v}_k + \gamma_{k,t+1}\tilde{v}_k\right\}$$

subject to the variance constraint (7). Hence, the optimal posttrade variance is described by the FOC in Equation (9), for $t=N$.

$$\hat{W}_t'(v_{i,t}|g_t) = \frac{d\Omega_t(i,j)}{dv_{i,t}} = \Sigma_k\left\{(\kappa_t+\gamma_{k,t+1})\left(\frac{\partial\tilde{v}_k}{\partial v_i}\right)\right\} \tag{A.9}$$

Given that γ_{t+1} is linear, let $\hat{\gamma}_i \equiv \kappa_t + \gamma_{i,t+1}$; according to Equation (9), we thus have $\omega_{ij} = \frac{\hat{\gamma}_j^2}{(\hat{\gamma}_i+\hat{\gamma}_j)^2}$ and

$$\begin{aligned}\gamma_{i,t} &= \hat{\gamma}_i\omega_{ij}^2 + \hat{\gamma}_j(1-\omega_{ij})^2 = \frac{\hat{\gamma}_i\hat{\gamma}_j^2+\hat{\gamma}_j\hat{\gamma}_i^2}{(\hat{\gamma}_i+\hat{\gamma}_j)^2}\\ &= \frac{\hat{\gamma}_i\hat{\gamma}_j}{(\hat{\gamma}_i+\hat{\gamma}_j)} = \frac{1}{2}H(\kappa_t+\gamma_{i,t+1},\kappa_t+\gamma_{j,t+1})\end{aligned}$$

which shows that this Lemma holds for period N. By backward induction, given $\gamma_{i,t+1}$ holds, Equations (9) and (A.9) can be applied to any t.

□

A.2.3 Proof For Lemma 3

Proof. Given PAM on v_t,for any solution $v_\theta^*(v_t)$ and the corresponding $v_{\theta'}^*(v_\theta^*(v_t))$, the total surplus among these four banks who all start with v_t can be expressed as

$$2\Omega_t(v|g_t) = 2\left(\Sigma_\theta\kappa_t v_\theta^*(v_t)\right) + \Sigma_{\theta,\theta'}\left\{\kappa_{t+1}v_{\theta'}^*(v_\theta^*(v_t)) + W_{t+1}\left(v_{\theta'}^*(v_\theta^*(v_t)|g_t)\right)\right\}.$$

Given that $W_t(v)$ is piece wise linear $\forall t$, let $\gamma_{k,t+2}(v_t)$ be the corresponding risk-capacity given $v_{\theta'}^*(v_\theta^*(v_t)$. Given $\gamma_{k,t+2}$, any optimal g_t must result in lower $\gamma_{i,t}(g_t)$,which can be expressed as

$$\gamma_{i,t} = \frac{1}{2}H\left(\kappa_t+\gamma_{i,t+1},\kappa_t+\gamma_{j_t(i),t+1}\right) = \Gamma\left\{\left(\gamma_{i,t+2},\gamma_{j_{t+1}(i),t+2}\right),\left(\gamma_{j_t(i),t+2},\gamma_{j_{t+1}(j_t(i)),t+2}\right)\right\},$$

which depends on $\gamma_{k,t+2}$ for four agents and the matching at period $t+1$.

We now show that any g_t that violates negative sorting on $\gamma_{k,t+2}$ is dominated. Rank agents by their cost γ_{t+2}^k, where $\gamma_{t+2}^1 \leq \gamma_{t+2}^2 < \gamma_{t+2}^3 \leq \gamma_{t+2}^4$. Let

$$f(\gamma_{t+2}, \gamma'_{t+2}) \equiv \frac{1}{\kappa_t + \left\{ \frac{1}{\kappa_{t+1}+\gamma_{t+2}} + \frac{1}{\kappa_{t+1}+\gamma'_{t+2}} \right\}^{-1}},$$

we thus have

$$\gamma_t^i = \left\{ \frac{1}{\kappa_t + \gamma_{t+1}^i} + \frac{1}{\kappa_t + \gamma_{t+1}^j} \right\}^{-1} = \underbrace{\left\{ f\left(\gamma_{t+2}^i, \gamma_{t+2}^{j_{t+1}^*(i)}\right) + f\left(\gamma_{t+2}^j, \gamma_{t+2}^{j_{t+1}^*(j)}\right) \right\}^{-1}}_{=\Gamma\left\{\left(\gamma_{t+2}^i, \gamma_{t+2}^{j_{t+1}^*(i)}\right), \left(\gamma_{t+2}^j, \gamma_{t+2}^{j_{t+1}^*(j)}\right)\right\}}$$

Since

$$\frac{\partial^2 f}{\partial\gamma\partial\gamma'} = \frac{-2\kappa_t\gamma\gamma'}{(\gamma\gamma' + \kappa_t(\gamma' + \gamma'))^3} \leq 0,$$

Hence, given γ_{t+2}^k, negative sorting at $t+1$ minimizes γ_t. □

A.3 Binary Actions

A.3.1 Proof for Lemma 4

Proof. Given that $\gamma_{N+1}(z)$ decrease in $z \in \{0,1\}$, Equation 10 gives the value of $\gamma_N(c)$, where $c_{i,N} = z_i + z_{j_t(i)} \in \{0,1,2\}$, and $\gamma_N(c)$ decrease in c. For $t = N-1$,

$$c_{i,N-1} = \left\{c_{i,N}, c_{j_N(i),N}\right\} = \left\{\left\{c_{i,N+1}, c_{j_N(i),N+1}\right\}, \left\{c_{j_{N-1}(i),N+1}, c_{j_N(j_{N-1}(i))N+1}\right\}\right\}.$$

By Lemma 3, $\{(1,1),(0,0)\}$ is dominated by $\{(1,0),(1,0)\}$ and hence, for any $c_{N-1} \in \{0,1,2,3,4\}$,the connections are unique, where $c_{i,N-1} = \left\{\lfloor\frac{c_{i,N-1}}{2}\rfloor, \lceil\frac{c_{i,N-1}}{2}\rceil\right\}$ and thus c_{N-1} is sufficient statics. Given that $\gamma_N(c)$ decrease in c, $\gamma_{N-1}(c)$ thus also increases in c.

By backward induction, assume that $c_{i,t} = (\lfloor\frac{c_{i,t}}{2}\rfloor, \lceil\frac{c_{i,t}}{2}\rceil)$ and let $\gamma_{t+1}(c)$ denote its corresponding risk-capacity, which decrease in c. Suppose that at period t,$c_{i,t} = (m,n)$ where $m-n \geq 2$,then $\lceil\frac{m}{2}\rceil \geq \lfloor\frac{m}{2}\rfloor > \lceil\frac{n}{2}\rceil \geq \lfloor\frac{n}{2}\rfloor$. Hence,

$$\gamma_t(m,n) \geq \gamma_t\left(\lfloor\frac{m}{2}\rfloor, \lceil\frac{m}{2}\rceil\right), \left(\lfloor\frac{n}{2}\rfloor, \lceil\frac{n}{2}\rceil\right) > \gamma_t\left(\lfloor\frac{m}{2}\rfloor, \lceil\frac{n}{2}\rceil\right), \left(\lfloor\frac{n}{2}\rfloor, \lceil\frac{m}{2}\rceil\right),$$

which is dominated. Thus thus shows that the optimal core access within any pair must be evenly distributed. Lastly, since $\gamma_{t+1}(c)$ is decreasing in c and, under the optimal access, $\gamma_t(c) = \frac{1}{2}H(\kappa_t+\gamma_{t+1}(\lfloor\frac{c}{2}\rfloor), \kappa_t+\gamma_{t+1}(\lceil\frac{c}{2}\rceil))$ is thus increasing in c at period t. This thus establishes that Lemma 4 must hold for any t.

□

A.3.2 Proof for Proposition 3

Proof. Since Lemma 4 has shown that, given any $c_{i,t}$, the optimal connections must distributed core access evenly within any pair, we thus have $c_{i,t+1} = \lfloor\frac{c_{i,t}}{2}\rfloor$ and $c_{j,t+1} = \lceil\frac{c_{j,t}}{2}\rceil$ with each pair, where by definition $c_{i,t} = c_{j,t}$. Given Proposition 2, we thus have $v_{i,t} = v_{j,t}$ and, within the pair, Equation 9 is thus reduced to

$$v_{i,t+1} = \left(\frac{\kappa_t + \gamma_{t+1}(c_{j,t+1})}{(\kappa_t + \gamma_{t+1}(c_{i,t+1})) + (\kappa_t + \gamma_{t+1}(c_{j,t+1}))}\right)^2 (2v_{i,t})\,.$$

□

A.3.3 Proof of Proposition 4

Given the expression of γ_t, we have $\gamma_t(\theta\kappa) = \theta\gamma_t(\kappa)$. Hence, Equation (12) can be rewritten as $\Pi = \kappa v_1 \max_c \left\{-\gamma_1^*(c) - \frac{c}{2^N}\left(\frac{\phi}{\kappa v_1}\right)\right\}$, where by comparative statics, $c^*\left(\frac{\phi}{\kappa v_1}\right)$ increases in $\frac{\phi}{\kappa v_1}$.

We now show that if $\delta = 0$, $\gamma_t(c) = 0\ \forall t, c \geq 1$. As $\gamma_{N+1}(1) = 0$ and $\gamma_{N+1}(0) = \kappa$, we thus have $\gamma_N(1) = \frac{1}{2}H(\delta + \gamma_{N+1}(1), \delta + \gamma_{N+1}(1)) = 0$ if $\delta = 0$. Assume that $\gamma_{t+1}(1) \to 0$, then $\gamma_t(1) = \frac{1}{2}H(\delta + \gamma^0_{t+1}, \delta + \gamma^1_{t+1}) \to 0\ \forall t$ by backward induction. Now we show that this property also holds for any $c > 1$. Assume that $\gamma_{t+1}(c) = 0$ holds for any (t, c); we thus have $\gamma_t(c) = \frac{1}{2}H(\delta + \gamma_{t+1}(\lfloor\frac{c}{2}\rfloor), \delta + \gamma_{t+1}(\lceil\frac{c}{2}\rceil)) = 0, \forall(t, c)$.

As $\delta \to \infty$, $\alpha_t(c) \to \frac{1}{2}$. Hence, regardless of the core access, the allocation is always symmetric, and thus $v_{i,t+1} = \frac{1}{2}v_{i,t}$ for all $i, t \leq N$. Hence, $\sum_{t\leq N}\int v_{i,t}di = \delta\left(\frac{1}{2} + \left(\frac{1}{2}\right)^2 + ..\left(\frac{1}{2}\right)^N\right) = \delta\left(1 - \left(\frac{1}{2}\right)^N\right)$, and thus

$$\Pi_0 = \max_c -\left\{\kappa\left[\left(1 - 2^{-N}\right)\delta v_1 + (1 - 2^{-N}c)2^{-N}v_1\right] + 2^{-N}c\phi\right\}.$$

Therefore, $c = 2^N$ iff $\left(\frac{1}{2}\right)^N v_0 > \phi$.

References

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2013). *Systemic risk and stability in financial networks*. Tech. rep. National Bureau of Economic Research.

Afonso, G. and R. Lagos (2015). "Trade dynamics in the market for federal funds". *Econometrica* 83.1, pp. 263–313.

Allen, F. and D. Gale (2000). "Financial contagion". *Journal of Political Economy* 108.1, pp. 1–33.

Anderson, A. and L. Smith (2010). "Dynamic matching and evolving reputations". *The Review of Economic Studies* 77.1, pp. 3–29.

Babus, A. and T.-W. Hu (2017). "Endogenous intermediation in over-the-counter markets". *Journal of Financial Economics* 125.1, pp. 200–215.

Babus, A. and P. Kondor (2018). "Trading and Information Diffusion in Over-the-Counter Markets".

Babus, A. and C. Parlatore (2017). "Strategic fragmented markets". *Available at SSRN 2856629*.

Bao, J., M. O'Hara, and X. A. Zhou (2016). "The Volcker rule and market-making in times of stress". *Journal of Financial Economics, Forthcoming*.

Bessembinder, H. et al. (2018). "Capital commitment and illiquidity in corporate bonds". *The Journal of Finance* 73.4, pp. 1615–1661.

Board, F. S. (2018). "Evaluation of incentives to centrally clear OTC derivatives".

Cabrales, A., P. Gottardi, and F. Vega-Redondo (2014). "Risk-sharing and contagion in networks".

— (2017). "Risk sharing and contagion in networks". *The Review of Financial Studies* 30.9, pp. 3086–3127.

Chang, B. and S. Zhang (2018). "Endogenous market making and network formation". *Available at SSRN 2600242*.

Choi, J. and Y. Huh (2018). "Customer liquidity provision: Implications for corporate bond transaction costs". *Available at SSRN 2848344*.

Collin-Dufresne, P., B. Junge, and A. B. Trolle (2018). "Market structure and transaction costs of index CDSs". *Swiss Finance Institute Research Paper* 18-40.

Corbae, D., T. Temzelides, and R. Wright (2003). "Directed matching and monetary exchange". *Econometrica* 71.3, pp. 731–756.

Duffie, D. (2018). "Post-crisis bank regulations and financial market liquidity". *Lecture, Baffi.*

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005). "Over-the-Counter Markets". *Econometrica* 73.6, pp. 1815–1847.

Dugast, J., S. Üslü, and P.-O. Weill (2019). *A Theory of Participation in OTC and Centralized Markets.* Tech. rep. National Bureau of Economic Research.

Elliott, M., B. Golub, and M. O. Jackson (2014). "Financial networks and contagion". *American Economic Review* 104.10, pp. 3115–53.

Farboodi, M. (2014). "Intermediation and voluntary exposure to counterparty risk". *Available at SSRN 2535900.*

Gale, D. M. and S. Kariv (2007). "Financial networks". *American Economic Review*, pp. 99–103.

Glode, V. and C. C. Opp (2019). "Over-the-Counter versus Limit-Order Markets: The Role of Traders' Expertise". *The Review of Financial Studies.*

Gofman, M. (2011). "A network-based analysis of over-the-counter markets". In: *AFA 2012 Chicago Meetings Paper.*

— (2014). "Efficiency and Stability of a Financial Architecture with Too-Interconnected-to-Fail Institutions". *Available at SSRN 2194357.*

Hojman, D. A. and A. Szeidl (2008). "Core and periphery in networks". *Journal of Economic Theory* 139.1, pp. 295–309.

Jackson, M. O. (2005). "A survey of network formation models: stability and efficiency". *Group Formation in Economics: Networks, Clubs, and Coalitions*, pp. 11–49.

Malamud, S. and M. Rostek (2014). "Decentralized Exchange".

Wang, C. (2016). "Core-periphery trading networks".

Yellen, J. (2013). "Interconnectedness and systemic risk: Lessons from the financial crisis and policy implications". *Board of Governors of the Federal Reserve System, Washington, DC.*

Yoon, J. H. (2017). *Endogenous market structure: Over-the-counter versus exchange trading.* Tech. rep. Working Paper, University of Wisconsin-Madison.