Risk Concentration and Interconnectedness in OTC Markets*

Briana Chang† Shengxing Zhang‡

First Draft: January 2021, This Draft: November 2021

Abstract

We develop a tractable framework that jointly determines trading connections and risk allocations banks in over-the-counter (OTC) markets. In an environment with ex-ante homogenous and risk-averse banks, a concentrated structure arises when banks are endowed with limited liability or options of investing technologies. The equilibrium network is payoff unique, trading off between the benefit of risk-sharing vs. risk-concentration. A continuous change in asset riskiness or policy parameters can trigger a structural shift, resulting in discontinuous changes in aggregate risks and transaction prices. We use this framework to evaluate the effect of different interventions on the market structure.

Keywords: Network, Over-the-Counter Market, Regulations

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 seminar participants at Wharton, 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

How does banks' risk valuations affect the interbank market structure? To answer this question, this paper develops a novel framework that jointly determines interbank links and risk allocations through them. Even when banks are ex-ante homogeneous, we find that they trade not only to share risks but also to concentrate them, creating an asymmetric interbank network consistent with empirical regularities: a few banks grow large and bear more risks. Our approach also provides a tractable framework for policy analysis and comparing market structures for assets with varied riskiness, providing new insights on how and when the the interbank network and the risk exposure of large banks may change.

We connect trading frictions in the OTC market to the limited information that banks have about other banks' asset positions when they form finite trading links and the uncertainty or risk it entails. Formally, they choose their trading partners sequentially for multiple round of bilateral trades, based on potential counterparties' identity and, more importantly, beliefs about their asset positions. At each trading round, we require that the bilateral matchings at the round and future rounds be stable, allowing multiple deviations. The collection of banks' counterparties and trades over all trading rounds represents the underlying trading network.

The final payoff of an individual bank depends on its risk position after OTC trades. We allow for a general payoff function at the individual level and analyze how it affects the aggregate network. At the individual level, we assume that banks' payoffs decrease in the risks (i.e., they all dislike holding risks). We show that the standard risk-sharing strategies may be overturned when banks face diminishing marginal cost of bearing risks. The diminishing cost is relevant in many applications, for example, when bank have limited liability and/or options of investing technologies to lower their risk-bearing cost(e.g., entering a trading platform or different risk-management technologies).

Having a few extremely risky banks could be optimal because they take on risks from other banks and have lower marginal risk-bearing costs (as they are in turn likely to default or access other trading platforms). But concentrating risks to a subset of banks generates higher overall risk exposure than sharing risks evenly across banks. The equilibrium network is determined by this trade-off. The asymmetric network emerges when the benefit of risk-concentration dominates.

The key dynamic element in our model is the public belief about a bank's asset positions at each trading round. It endogenously depends on with whom the bank has traded with and how it has traded with the counterparties. We show that the variance of a bank's position, which can be interpreted as the risk that a bank bears, is the sufficient static for network formation. The main equilibrium objects are thus the evolution of banks' risk exposures (i.e., how each bank trades over times) and their counterparties (i.e., whom they trade with).

To understand how risks are concentrated through dynamic connections, we first establish that, with diminishing marginal cost of bearing risks, risks are concentrated through positive sorting over times. That is, riskier banks, who hold more risks in the past, are matched with each other.

How banks allocate risks within a match depends on their marginal cost of bearing additional risk, which we refer to as their risk-bearing capacity. A bank's risk-bearing capacity evolves endogenously based on the trading network and trading dynamics. Given any network that is locally optimal, a bank's current-period capacity equals the harmonic mean of the next-period capacities of the bank and its period-t counterparty.

In the first application with limited liability, we show that a small increase in the riskiness of the asset, thus the balance sheet cost of holding it, can result in a regime 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 "systematic risks" through trading networks.

Our notion of default risks in this application is different from standard theories of financial contagion, which analyze how bank defaults propagate through given network connections.[1] In our framework, the aggregate default risks increase because banks systematically change their trading behaviors through the interbank network. Hence, even without contagion, we highlight another source of systematic risks through the network.

We show that the globally optimal network features more symmetric risk capacity in

[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.

earlier matches. Because risk concentration increases the overall risk exposure, it is better to have it later. In the simple case in which banks can take binary actions in the final period, we show that, given any number of core banks who choose to take costly actions to increase their risk bearing capacities in the final period, the optimal network must distribute the core access evenly within a match. The optimal network is then reduced to choosing the optimal core size.

We use this analytical characterization to study our second application, where banks can choose whether to access a multilateral platform to increase their risk-bearing capacity for a fixed entry fee. The aggregate market structure in this case depends on the cost of bearing risks relative to the entry fee. This summary statistic allows us to derive 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.

Consistent with empirical evidence, our model predicts that policies that increase balance sheet costs relative to the entry fee could result in a more symmetric market structure. Nevertheless, it can have ambiguous effects on transaction costs measured by volume-weighted average bid-ask spreads.

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

[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.

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 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 determination of the bilateral trading 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.

[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.

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 his payoff at the terminal period $N+1$, depends on his positions after bilateral trades.

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,

[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.

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.

Post-Trade Risk Exposures 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 $W_{N+1}(v_{i,N+1}) = Eu_{N+1}(a_{i,N+1}) = -\kappa_{N+1}v_{i,N+1}$. More generally, we allow for any final payoff $W_{N+1}(v)$

Assumption 1. *(Risk-Averse)* $W_{N+1}(v)$ *is a decreasing function of post-trade risk-exposure* $v_{i,N+1}$.

Assumption 1 means that, at the individual level, all agents would like to reduce their risk-holdings. We impose this assumption to avoid trivial risk-taking behaviors.

As discussed later in details, while it's costly to hold risks for all banks ($W'_{N+1}(v) \leq 0$), what matters for the aggregate network is the convexity of $W_{N+1}(v)$, where a convex $W_{N+1}(v)$ represents diminishing marginal costs of holding risks.

This can happens when, for example, when banks are protected by limited liability; and thus, the marginal cost of taking additional risks is lower for riskier banks that are likely to default. Or, one can also imagine that banks have options to invest with superior but more expensive trading or risk-management technologies. Since the value of doing so is higher for riskier banks, whose marginal cost of holding risks thus can be lower. To proceed, we first establish the general results for any $W_{N+1}(v)$ and then apply it to two specified applications: platform access (Section 4) and limited liability (Section 5).

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. Note that, while the terms of trades are only contingent on the realized positions within a pair, agents choose it optimally according to the network in equilibrium, which we will explain in detail later in the equilibrium definition.

We make the following assumption on posttrade asset allocations.

Assumption 2. *The posttrade asset allocations between matched agents i and j are quasilinear in their pretrade total asset holdings, $a_i + a_j$.*

$$\tilde{a}_k(a_i, a_j) = \alpha_k(a_i + a_j) + \beta_k, \ \textit{for } k = i, j.$$

Assumption 2 simplifies our analysis but is not restrictive on the allocation of risks. Under this assumption, the feasibility constraint on posttrade asset holdings, (2) for all a_i and a_j, is reduced to two constraints, $\alpha_i + \alpha_j = 1$ and $\mu_i + \mu_j = 0$, where α_i controls the riskiness of agent i's asset holding, μ_i controls the expected amount of her holding. For example, when $\alpha_i = 1$, $\alpha_j = 0$, the variance of agent i's holding equals to the variance of total asset holdings, $a_i + a_j$. When $\alpha_i = \alpha_j = 0.5$, the allocation diversifies risks, $\text{Var}(\tilde{a}_i) = \text{Var}(\tilde{a}_j) = \frac{1}{4}\text{Var}(a_i + a_j)$. Assumption 2 does rule out nonlinear allocations, for example, $\tilde{a}_i(a_i, a_j) = \max(a_i + a_j, 0)$, $\tilde{a}_j(a_i, a_j) = \min(a_i + a_j, 0)$. Such nonlinear rules may allow agents to coordinate their matching in a different manner, which may be relevant for some applications.

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 the Bayes' rule,

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

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.

[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.

2.2 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.

Let $\hat{W}_{t+1}(i)$ 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.

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

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 $t+1$.

In other words, for any round t, an agent i takes the future aggregate evolution of π^*_τ and hence the individual payoff $\hat{W}_\tau(i)$ $\forall\tau \geq t$ as given, and choose his trading partner and the term of trades at round t optimally. Our formulation thus means that each agent is negligible relative to the aggregate distribution; however, each agent is non-negligible in pursuit of his own interest through matching and contracting within a pair.[9]

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 $t \in \{1, \dots, N+1\}$:

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

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

[9] In this sense, it can be understood as a competitive equilibrium as in the literature on large games (McAfee 1993), where the payoff of each individual is determined only by his own decision and by the aggregate distribution of trading decisions in the market.

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.

3 General Properties

3.1 Equivalence and Uniqueness

We first establish that the equilibrium outcome is unique and maximizes the aggregate payoff. 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 $E_t(\tilde{a}_{i,t}{}^2) = \int\int \tilde{a}_{i,t}(a_i, a_{j_t(i)})^2 \pi_{i,j_t(i),t}(da_i, da_{j_t(i)})$ and the terminal payoff depends on the post-trade variance at period N, which is given by $\Pi_{N+1}(\boldsymbol{\pi}_{N+1}) = \int_0^1 W_{N+1}(v_{i,N+1})\, di$,where $v_{i,N+1} = \int a_{i,N+1}^2 \pi_{N+1}(da_{i,N+1})$.

Proposition below first shows that, given any κ_t and $W_{N+1}(v)$, the equilibrium strategies - including agents' bilateral connections and the terms of trade thin each match - maximizes the aggregate payoffs.

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.[10] 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.

3.2 Risk Allocation

Reformulation: Variance Representation Instead of working with asset allocations, we first reformulate the problem in the space of variance (i.e., risks). 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 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.

[10] Because agents have quasilinear preferences, this is equivalent to solving for Pareto optimal allocations.

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 now 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, given any pre-trade variance V_{ij} for the agent i and j, the optimal share $\alpha \in [0,1]$ solves*

$$\Omega_t(V_{ij}) = \max_{\alpha\in[0,1]} \left\{-\kappa_t \left(\alpha^2 + (1-\alpha)^2\right) V_{ij} + W_{t+1}(\alpha^2 V_{ij}) + W_{t+1}((1-\alpha)^2 V_{ij})\right\} \tag{8}$$

In other words, the optimal asset/risk allocation with any pair with post-trade variance V_{ij} can thus be reformulated as choosing the share of the risks within a pair (i, j), where bank i holds a share $\alpha_i \in [0, 1]$ of total position, so that $\tilde{a}_i(a_i, a_j) = \alpha_i(a_i + a_j)$ and bank j holds $\alpha_j = 1 - \alpha_i$ share. 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, as $\tilde{v}_k = \alpha_k^2 V_{ij}$.

Allocation of Risks Within the matches Given any match, the risk allocation within the pair is thus pins down by the FOC condition from Equation (8). Assuming $W_t(v)$ is differentiable $\forall t, v$, we thus have

$$\alpha_t(V) = \frac{\kappa_t + W'_{t+1}((1-\alpha)^2 V)}{\left(\kappa_t + W'_{t+1}((1-\alpha)^2 V\right) + \left(\kappa_t + W'_{t+1}(\alpha^2 V)\right)}. \tag{9}$$

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 pins down the underlying trading network. We thus now analyze the matching outcome.

3.3 Sorting Dynamics

First of all, observe that when $W_{t+1}(v)$ is concave, one can show that the objective function in Equation (8) is concave; and since $\alpha = \frac{1}{2}$ satisfies the FOC condition, it is also the unique global maximum. In other words, the standard predictions on risk-sharing are obtained in this case: agents share their exposure equally with any match and thus $v_{i,t+1} = \frac{v_{i,t}+v_{j,t}}{4}$. 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 $W_{N+1}(v) = -\kappa_{N+1}v$.[11,12]

Lemma 2. *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.*

Given that a concave $W_{N+1}(v)$ is well-understood, we focus on the case when $W_{t+1}(v)$ is convex throughout the rest of the paper. Observe that the FOC conditions in this case are generally not sufficient, and asymmetric risk-allocation *could* be optimal. That is, it is possible to exist $\alpha \in (\frac{1}{2}, 1]$ that also satisfies the FOC where $W'_{t+1}(\alpha^2 V) > W'_{t+1}((1-\alpha)^2 V)$. That is, Agent i unloads more risk to her counterparty j when Agent j has a lower marginal cost of risk-bearing the next period.

Whether the asymmetric solution is indeed optimal, it will then generally depend on the convexity of $W_t(v)$. In the static model ($N = 1$), the solution depends on the specified property of $W_{N+1}(v)$ and the given initial condition $v_{i,N} = v_1 \forall i$.

In our dynamic environment, the post-trade variance $v_{i,N}$ depends on how an agent traded through bilateral network over times; moreover, at any period t, the value func-

[11] 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 ($v^i_t = \left(\frac{1}{2}\right)^t v_0 \ \forall i$).

[12] 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.

tion $W_t(v)$ endogenously depends on the optimal choice of counterparties, which can be expressed as

$$W_t(v_i) = \max_j \left\{ \Omega_t(V_{ij}) - W_t(v_j) \right\}.$$

In other words, through the dynamic trades, our framework captures how the risk can accumulated through the trading network. Proposition below first establishes risks are in fact concentrated over time through positive associative matching.

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

This is because that, with convex $W_{N+1}(v)$, one can show that $\Omega_t(V_{ij})$ is also convex in $V_{ij} = v_i + v_j$ $\forall t$. Hence, given any distribution of $v_{i,t}$, agents are matched with counterparties that hold the same level of risk, thus on the equilibrium path, $V_{ij} = 2v_i$. In other words, agents that accumulate risks from others (higher post-trade variance $v_{i,t+1}$) are matched among with each other. Through this channel, 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.

Our approach admits a tractable characterization with a convex payoff function in the final period, $W_{N+1}(v)$, as shown later in our detailed analytical characterization. Our setup also admits a tractable numerical algorithm for any $W_{N+1}(v)$. A feature of PAM is that the optimal solution is distribution-free.: the matching and trading strategies of an agent holds for any distribution of risk exposures across agents. Under PAM, the network formation problem can be greatly simplified. First, solve agents' value functions, $W_t(v)$, and policy functions, $\alpha_t(v)$, backward from period N to period 1. Given equilibrium matching is positive assortative, $W_t(v)$ and $\alpha_t(v)$ for a value v is the solution to a one-dimensional optimization problem in Eq (8). As a byproduct of the optimization problem, we also solve the transfers within a match. Second, given the policy functions and positive assortative matching in equilibrium, solve for the evolution of variances v_t^i over time. Because the numerical algorithm involves only one dimensional optimization, it is easy to solve even if the objective function is convex. As a result, unlike the typical network formation problem that often suffers from the curse of dimensionality and may be hard to track, our recursive approach admits a rather tractable solution. Note that this is true even when the model is static – which can be obtained by setting $\kappa_t = 0$ for $t \leq N$ so

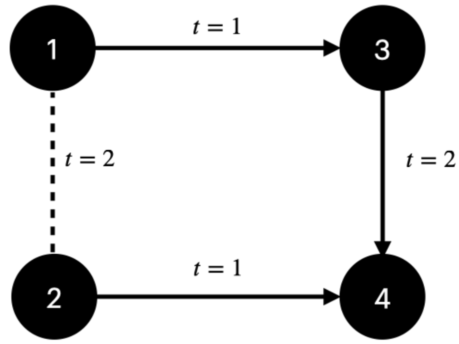

Figure 1: Risk-concentrating Network ($N = 2$)

that only the final allocation matters.

3.4 Trading Network

Figure 1 illustrates an example of trading network that involves risk-concentration. Within a match, the allocation must solves Equation 8, which means that two matching agents can potentially have different post-trade variance. We use the arrow points toward the agent with higher post-trade variance if asymmetric allocation arises. In this example, all agents start with the same initial risk position v_1. At period 1, Agent 3 and 4 take on more risks from Agent 1 and 2. At period 2, PAM implies that Agent 3 and 4 are matched while Agent 1 and 2 are matched.

Our sequential formulation of networks implies that the effect of earlier connections and trading outcomes is summarized by the state variable v_t. Conditional on v_t, the allocation at period t only depends on the future connections moving forward. That is, in this example, when Agent 1 trades with Agent 3 at period 1, he takes into account that she is connected to Agent 4 at period 2. In this sense, Agent 1 is indirectly connected to Agent 4 through Agent 3 at period 1.

To define the indirect connections, denote a set of agents I and their counterparties at period t $J_t(I)$. $J_t(I) = \cup_{i \in I}\{i, j_t(i)\}$. We say that an agent i is connected to an agent j from period t onwards if $j \in J_N(J_{N-1}(\ldots(J_{t+1}(J_t(i)))\ldots))$. Under this definition, the set of agents that Agent i is connected to from period t onwards is a tree with its root at the current match $J_t(i) = \{i, j_t(i)\}$.[13] At period t, there are $N-t+1$ rounds of bilateral matching that remain, and an agent can be connected to at most 2^{N-t+1} agents from period t onwards. Let $g_t(v_{i,t})$ summarizes the period-t network for an agent with $v_{i,t}$, which includes risk positions of all his connected counterparites

$$g_t(v_{i,t}) = \{v_{k,\tau}, j_{k,\tau}\}_{\forall k \in J_t(i), \forall t \leq \tau \leq N}$$

Two notes: First, since Agent i can have higher or lower post-trade variance within the pair v_t at each round, there can be at most 2^N types of dynamic path of risk-positions over N trading rounds. It is thus convenient to interpret that $g_1(v)$ represents the *ex-ante* trading networks among 2^N types of agents. Mapping to a continuum of traders, each type then has a measure of $\frac{1}{2^N}$.

Second, the network at period $t+1$ can be understood as deleting bilateral links $\{j_t(i)\}$ at period t from network $g_t(v_{i,t})$. Let $\tilde{v}_\theta(v)$ denote the post-trade variance within the pair v, where $\theta \in \{h, l\}$ and $\tilde{v}_h(v) \geq \tilde{v}_l(v)$. In this example, after Agent 3 holds takes on more risks $v_{3,t+1} = \tilde{v}_h(v) > v_{1,t+1} = \tilde{v}_l(v)$, at period $t = 1$, these two agents are no longer connected and now in two different sub-networks, given by $g_{t+1}(\tilde{v}_h(v))$ and $g_{t+1}(\tilde{v}_l(v))$, respectively.

Risk-bearing Capacity and Dynamic Connections The dynamics of connections thus affect agents' risk-bearing capacity over times, which in turns determines the risk-allocation within a pair. Let $\hat{W}_t(v|g_t(v))$ represent the joint payoff and payoff of an agent with pre-trade position v under network $g_t(v)$.

Given that two matching agents are identical at period t (because of PAM), the equilibrium payoff for each agent must be identical, which yields

$$\hat{W}_t(v|g_t(v)) = \frac{1}{2}\left(\Sigma_\theta\left\{-\kappa_t \tilde{v}_\theta(v) + \hat{W}_{t+1}(\tilde{v}_\theta(v)|g_{t+1}(\tilde{v}_\theta(v))\right\}\right).$$

We refer the marginal cost of holding risk for an with v_t as his risk-bearing capacity at period t, which is denoted by $\hat{W}_t'(v_t|g_t(v))$. Recall that any solution $\tilde{v}_\theta(v)$ must satisfy

[13]Due to the dynamic nature of our framework, the future links are the specific factor that matters for current trading decisions. Thus, the relevant connections for an agent can be understood as a tree. Nevertheless, the actual network $g_1 = \{j_\tau(i), A_{i,N+1}\}_{\forall i, 1 \leq \tau \leq N}$ does not need to be a tree. For example, according to Figure **??**, the network graph contains loops.

FOCs in Equation 9. Lemma below establishes that the risk-bearing capacity can be characterized recursively for any network that satisfies FOCs and PAM.

Lemma 3. *For any network $g_t(v)$ that satisfies the FOC conditions and PAM, the marginal cost of holding risks for agent with position v at period t is given by*

$$\hat{W}_t'(v|g_t(v)) = \frac{1}{2}H\left(\kappa_t + \hat{W}_{t+1}'(\tilde{v}_h(v)|g_{t+1}(\tilde{v}_h(v))), \kappa_t + \hat{W}_{t+1}'(\tilde{v}_l(v)|g_{t+1}(\tilde{v}_l(v)))\right) \; \forall t \leq N. \tag{10}$$

Equation 10 has a simple interpretation: the risk-bearing cost of Agent i at period t is the harmonic mean[14] 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 period, they must have the same capacity at period t, given they allocate the risks jointly, taking into their future connections.

In general, observe that when $W_t(v)$ is convex, there could exist multiple solutions that satisfy FOCs and PAM. In other words, these two are only necessary conditions for the optimal network. To proceed, we consider two applications and establish properties of the global optimal network. First, we provide conditions and full characterization when there is at most one "core" agent and $\kappa_t = 0$. We apply this result to highlight the excess risk-taking through network as a result of limited liability. Second, we allow for $\kappa_t > 0$ and analyze the optimal core size in an environment where banks have options to invest better technologies (such as, entering a centralized platform).

4 Application 1: Limited Liability

A prevalent concern in financial intermediation is the risk-taking incentive that results from limited liability. We now show that banks might collectively use their network to concentrate risks instead of sharing risks. This result holds despite that, banks are risk-averse (i.e., under Assumption 1).[15] Since default effectively offloads downside risks to outside creditors, any risk-taking is inefficient from viewpoint of planner. We then consider how interventions can correct such incentives.

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

[15]Note that, the standard risk-taking behavior arises where banks' payoffs are convex in their asset positions and thus banks might prefer higher variance, which gives higher upsides. Our result here goes beyond this channel as we assume that $W_{N+1}(v)$ decreases in v.

4.1 Regime Shift: Full Risk–sharing vs. Maximum Concentration

In this application, we are interested in the interaction between any given banks' individual payoff $W_{N+1}(v)$ and the outcome of bilateral networks. In particular, we show that 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.

To illustrate this point, we set $\kappa_t = 0$ for simplicity; hence, this environment is equivalent to a static model in the sense that only the final risk positions matter. Observe that, a static problem, which involves allocating risks $v_{i,N+1}$ among 2^N agents, is a highly multidimensional optimization problem. Below, we first identify the sufficient conditions that guarantees the if asymmetric allocation arises, risks are concentrated only at *one* of the 2^N agents. We thus refer this agent as the "core".

Condition 1. $W_{N+1}(v)$ is twice differentiable, and for $\forall v$ $W'_{N+1}(v) < 0$, $W''_{N+1}(v) > 0$, $W'_{N+1}(v)$ is bounded

Condition 2. (Monotonicity) $c(v) \equiv -\frac{W''_{N+1}(v)}{W'_{N+1}(v)} > 0$ weakly increases in $v \in [0, 2^N v_0]$.

Condition 1 guarantees that the solution is interior. Since the benefit of concentration is driven by the convexity $W''_{N+1}(v)$ and the cost of holding risks is captured by $W'_{N+1}(v)$, one can be interpret that Condition 2 means the relative benefit of concentration increases in the risk positions. It thus implies that concentration are more likely to happen for large risk positions.

Lemma 4. *(Concentration w/ One Core) Under Condition 1, 2, and* $\kappa_t = 0$, *risk allocations among* 2^N *agents involve at most one core agents with post-trade variance* v^c_{N+1}, *and the rest of non-core agents have the same terminal variance* v^0_{N+1}, *where* $v^c_{N+1} \geq v^0_{N+1}$.

Lemma thus reduces the multidimensional problem to one-dimensional. The aggregate payoff can be understood as 2^N agents share a total risk of $V \equiv 2^N v_0$,where one "core" agent may hold more risks than the rest of $2^N - 1$ agents, which yields

$$\Pi(v_0, N) = \frac{1}{2^N} \max_{\alpha \geq \frac{1}{2^N}} \left\{ W_{N+1}\left(\alpha^2 V\right) + \left(2^N - 1\right) W_{N+1}\left(\left(\frac{1-\alpha}{2^N - 1}\right)^2 V\right) \right\}. \tag{11}$$

Observe that $\alpha = \frac{1}{2^N}$ represent the case with fully risk-sharing among all agents; and thus $v^i_{N+1} = \left(\frac{1}{2^N}\right)^2 V = \frac{v_0}{2^N}$ $\forall i$.

4.1.1 Aggregate Implications: Regime Shifts

Let $\alpha^*(V)$ denote the optimal risk-allocation that solves Equation 11. The solution can be understood as the following: Agents engage fully risk-sharing ($\alpha^*(V) = \frac{1}{2^N}$) for lower initial risk-position v_0. However, there exists a marginal value v^* such that it is optimal for agents to shift to maximum concentration to only core agent, i.e., $\alpha^*(V) > \frac{1}{2^N}$.

Moreover, when this happens, $\alpha^*(V)$ involves discontinuous jump with $N \geq 2$. We thus refer this discontinuous changes as regime shifts. In fact, such a discontinuous change will not exist if $N = 1$ (i.e., with minimum interconnectedness). In other words, $\alpha^*(V)$ is a continuous and increasing function in V. Moreover, a higher N implies that the core agent can collect more risks from more counterparties. Hence, higher N leads to higher degree of concentration. Our result thus highlights the risk-concentration incentives are relevant when banks are highly interconnected, or have higher risk positions.

Expected Transfers (Prices of Risks) The regime shift does not imply discontinuous change in the aggregate risks, but also affects the transaction price. To see this, we now look at the price/transfer that implements the optimal network in equilibrium. Note that, because of positive sorting, any two matching agents have are homogeneous when they meet (which is characterized by $v_{i,t}$) at period t on the equilibrium path. Hence, while the optimal allocation of risk is generally asymmetric, the transfers must be such that they are indifferent.

Given that holding risk is costly, agent j that holds more risk needs to be compensated so that the agent will be indifferent. Specifically, the expression of $W_t(v)$ shows that an agent's maximum payoff is decreasing in v. Thus, the expected transfer from agent i to j solves

$$-\kappa_t \tilde{v}_i + W_{t+1}(\tilde{v}_i) - x_t = -\kappa_t \tilde{v}_j + W_{t+1}(\tilde{v}_j) + x_t, \tag{12}$$

where $\tilde{v}_k$ is given by the optimal allocation and $\tilde{v}_i = \tilde{v}_l(v) \leq \tilde{v}_j = \tilde{v}_h(v)$ and $x_t \geq 0$.

The equilibrium transfer $x_t(i, j)$ within the pair can be implementedxx don't talk about implmentation of BA.

as linear bid-ask spread times the expected volume. Specifically, the agent who holds higher post-trade variance within the pair charges linear bid and ask prices, $P_t^A(i,j)$ and $P_t^B(i,j)$, respectively. The spread $S_t(i,j) \equiv P_t^A(i,j) - P_t^B(i,j)$ then solves

$$\left(\frac{S_t(i,j)}{2}\right)\vartheta_t(i,j) = x_t(i,j),$$

where $\vartheta_t(i,j) \equiv \mathbb{E}|\alpha_{i,t}(a_{i,t-1}+a_{j,t-1})-a_{i,t-1}|$ represents the expected volume between the pair (i,j).

This immediately implies that when banks engage in risk-sharing, the bid-ask spread is zero within the pair. Hence, when the network shifts to risk-concentration, bid-ask spread must become positive. Moreover, such increase is discontinuous with discontinuous increase in the risk concentration.

Proposition 3. *Under Condition 1 and 2, there exists a cutoff v^* such that the equilibrium is fully risk-sharing for $v \leq v^*$ and features concentration with one core agent with $v \geq v^*$. For any $N > 1$, the aggregate post-trade exposures, $\int v_{i,N+1}di$, and bid-ask spread discontinuously increase at v^*.*

Figure 2 illustrates this result using $W_{N+1}(v) = -1 + e^{-cv}$, which implies constant $\frac{W''_{N+1}(v)}{W'_{N+1}(v)} = c$ and thus Condition 1 and 2 are satisfied. The red line represents the outcome where banks choose to share risks. Hence, each of them has low final risk exposure and default probability. The blue line, on the other hand, represents the case when it becomes optimal for banks to concentrate risks to the core, 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.

4.1.2 Distribtuion of Risks in Banking Network

builing up risks and then get rid of risks "better"

When risk concentration arises, the core agent collects risk from others. While the final allocation can be understood from a static model, our sequence setting further gives predictions regarding how the asset flows through the bilateral network. Intuitively, agents who are directly or indirectly connected to the core agent will have lower cost of

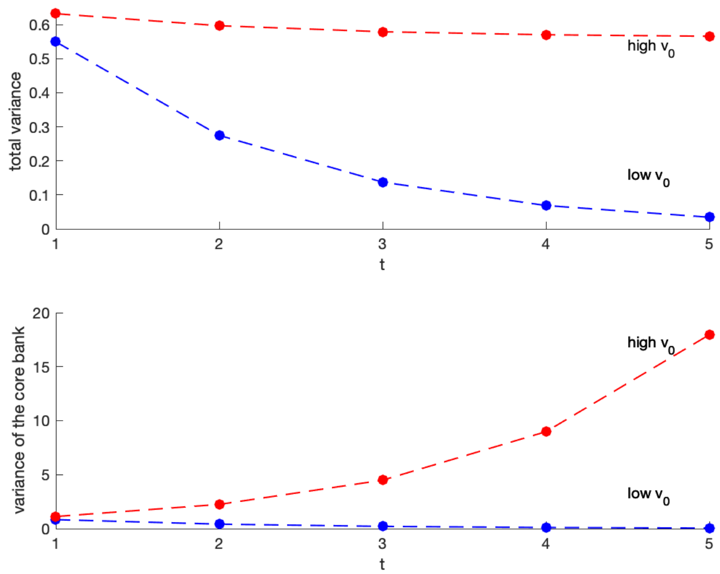

Figure 2: Regime Shift: $W_{N+1}(v) = -1 + e^{-cv}$, $c = 1.0196$, $v_{0L} = 1.02$ and $v_{0H} = 1.03$.

holding risks and thus can take on more risks from his counterparties. Importantly, as shown in Lemma 3, the model dynamics imply that the risk-bearing cost of Agent i is time-varying.

Dynamic Cores Access (to understand how risks flow) Let $c_{i,N+1} = 1$ iff $v_{i,N+1} = v^c_{N+1}$ denote the agent is the core and zero otherwise. We define the core access of an agent at time t as the number of core agents that Agent i is directly and indirectly connected at period t :

$$c_{i,t} \equiv \Sigma_{k \in J_t(i)} c_{k,N+1}.$$

Recall that the network of an agent at time t can be understood as adding the bilateral $j_t(i)$ to their $t+1$ networks. Hence, the core access can be defined recursively: $c_{i,t} = c_{i,t+1} + c_{j_t(i),t+1}$. That is, by connecting his counterparty $j_t(i)$, Agent i obtains the future core access of agent $j_t(i)$ at period t.

When there is only one core agent among 2^N agents, then the core access is always binary at any t: $c_{i,t} = 1$ iff agent i is directly or indirectly connected to the core agent at period t and $c_{i,t} = 0$ otherwise. When $\kappa_t = 0$,Lemma 3 further implies that the risk-capacity of an agent at period t only depends on the harmonic mean of $W'_{N+1}(v_{k,N+1})$ of all his connected counterparties at period t, which yields

$$W_t'(v_{i,t}) = \frac{1}{2^{N-t}} \left\{ \Sigma_{k\in J_t(i)}^{2^{N-t}} \left(\frac{1}{W_{N+1}'(v_{k,N+1})} \right) \right\}^{-1}.$$

In other words, the risk-capacity for an agent at period t only depends on whether he has core access at that period. If the agent is connected to the core at period t, then one of $W_{N+1}'(v_{k,N+1})$ is valued at the core agent $W_{N+1}'(v_{N+1}^c)$, while the rests are valued at $W_{N+1}'(v_{N+1}^0)$. Since $W_{N+1}'(v_{N+1}^c) \geq W_{N+1}'(v_{N+1}^0)$, it means that the agent with core access will have lower cost of holding risks.

Recall that there can be at most one agent within the pair can maintain the core access at period $t+1$,or equivalently, one of them must loss core access. Hence, according to the FOCs, the share to agent i within the pair is thus strictly higher if and only if agent i has core access. Agents share risk if and only if both of them do not have core access at period $t+1$.

To summarize, when there is one core agent (i.e., $c_{i,1} = 1\ \forall i$), the dynamic path for any agent i can be understood as when he loss his core access. For an agent where $c_{i,t} = 1\ \forall t \leq \tau$ and $c_{i,t} = 0\ \forall t \geq \tau + 1$, he will collect risks from his counterparties for $\tau - 1$ periods, unload his the risks to his counterparties at period τ, and then engage fully risk-sharing afterward.

Expected Volume Agents that collect risks for more periods will thus then have higher expected volume. Formally, the expression of $\vartheta_t(i,j)$ highlights that the expected volume depends on their pre-trade exposure, where PAM implies that $v_{i,t} = v_{j,t}$, and the optimal allocation of risks, captured by $\alpha_{i,t}$. Suppose that, for example, the pretrade position $a_{i,t}$ follows a normal distribution, given any $(v_{i,t}, \alpha)$,the volume is then given by $\vartheta_t(i,j) = (2/\pi)^{1/4}\sqrt{((1-\alpha_{i,t})^2 + \alpha_{i,t}^2)v_{i,t}}$. This thus shows that agents that hold more risks over time are likely to higher expected volume, as $\vartheta_t(i,j)$ increases with $v_{i,t}$.

Note that while agent with longer core access will have higher expected volume; they are not riskier at the end, as they will ultimately unload their risks to the core. According to Lemma 4, they are in fact as "safe" as other banks, measured by their final risk exposures. This result holds more generally even when $\kappa_t > 0$ as discussed in Section 5.1.

4.2 Policy Implications

Normative Implications In this application, any risk-taking is social inefficient because default effectively offloads downside risks to outside creditors. Since the social planner prefers risk-sharing, 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 $\kappa_t(1+\tau^\kappa)$.

Formally, the objective of the social planner is

$$-\int_0^1 \left[\sum_{t=1}^{N+1} \kappa_t(1+\tau^\kappa)v_{i,t}\right] di + T$$

where T is a lumpsum transfer from the planner. The planner maximizes the objective subject to the government budget constraint,

$$\int_0^1 \left[\sum_{t=1}^{N+1} \kappa_t \tau^k v_{i,t}\right] di - T \geq 0.$$

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.

5 Application 2: 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.[16] Motivated by this, we now consider the environment where the convexity arises when banks have options to invest better technologies in order to reduce their risk-bearing

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

cost. We model this option as banks can choose a binary action, $c_{i,N+1} \in \{0,1\}$ at period $N+1$.

Assumption 3. *Piece-wise linear with Binary Action*

$$W_{N+1}(v) = \max_{c_{N+1}\in\{0,1\}} \{-\gamma_{N+1}(c_{N+1})v - \phi(c_{N+1})\}, \tag{13}$$

where $\gamma_{N+1}(0) > \gamma_{N+1}(1)$ and $\phi(0) < \phi(1)$.

One can interpret $c_{N+1} = 1$ represents that banks have access to a centralized platform, where $\gamma_{N+1}(1)$ means agents can have lower post-trade exposure in that platform and $\phi_{N+1}(1)$ represents its corresponding entry cost. A fully competitive centralized market can be understood as a platform that allows fully risk-sharing, where $\gamma_{N+1}(1) = 0$. In case when banks do not have access, $\gamma_{N+1}(0)$ is then the cost of holding assets and we normalize $\phi(0) = 0$. We now apply our framework to study the positive and normative implications of reforms, taking into the equilibrium response of the market structure.

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.

5.1 Full Characterization: Delayed Concentration and Core Size (shorten)

In this application, we take into account that holding risks on the balance sheets can be costly within the trading window, not just at the terminal period. We thus consider the general case where $\kappa_t > 0$. Theoretically, this implies that dynamic connections now matter.

To see this, consider $N = 2$ and thus there were four banks that could potentially be connected. Suppose that, the bilateral trading outcome is such that, at period $N+1$,

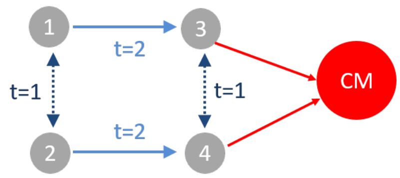

Figure 3: Late vs. early Concentration ($N = 2$)

both Agents 3 and 4 have higher post-trade variance and chooses the action to reduce their risk-bearing costs (such as, paying a fee to access to a centralized market). Agent 1 and 2, on the other hand, have lower post-trade variance and do not have access.

Figure 3 shows two possible network graph which differ in terms of their dynamic bilateral connections $\{j_\tau(i)\}_{\tau=1,2}$ and thus risk-allocations over times. In the left graph of Figure 3, an agent is first connected with another with the same platform access and then connected with another agent with with different platform access at period 2. Intuitively, this connection implies that all four agents have direct or indirect access to the platform at period 2. Indeed, according to Lemma 3, their effective risk-bearing capacity are in fact symmetric under the left network is given by $\frac{1}{2}H\left(\kappa_N + \gamma_{N+1}(0), \kappa_N + \gamma_{N+1}(1)\right)\right) \forall i$.

Hence, the optimal allocation at period $N-1$ must also be symmetric. In other words, all agents first adopt risk-sharing at period 1 for the left network. This order of matching however is reversed under the right network. According to 3, the marginal cost of holding risk for Agent 3 and 4 is now lower than the one of Agent 1 and 2. Hence, under the right network, the risk-allocation between $\{1, 3\}$ and $\{2, 4\}$ at period 1 must be asymmetric, given that their risk-capacity at period 2 differs.

We now show that, for any $\kappa_t > 0$, the right network, which results higher asymmetric risk-capacity at period $t + 1$ and thus earlier concentration, is dominated by the left network, which results in more symmetric risk-capacity at period $t + 1$ and thus late concentration. Intuitively, this is because that concentration necessarily results in higher total variance; and thus, any network that violates back-loading property is dominated.

Lemma 5. *When $\kappa_t > 0$, for any optimal network, if $v_{4,t+2} \geq v_{3,t+2} > v_{2,t+2} \geq v_{1,t+2}$, it must be the case that agents with $v_{1,t+2}$ ($v_{2,t+2}$) are matched with agents with $v_{4,t+2}$ ($v_{3,t+2}$) at period $t + 1$.*

We prove this by showing that, if this condition is violated, fixing $v_{k,t+2}$ and all the networks onward, re-matching type 1 (2) agent with type 4 (3) agent at period $t+1$ lowers the total variance of $v_{k,t+1}$. Hence, whenever $\kappa_t > 0$, such a deviation is profitable.

We now show that Lemma 3 and 5 together pins down the unique market structure. We refer agents that have platform access as "core agents". The definition of core access $c_{i,t}$ is the same as before, with the difference that $c_{i,t}$ can be more than one. Lemma 5 implies that, given any core access within a pair, the optimal network must distribute the core access as even as possible next period. Hence, the core access $c_{i,t}$ becomes the sufficient statistic for the network. We thus use $\gamma_t(c_{i,t})$ to denote the risk capacity for agent with core access $c_{i,t}$ at time t under the optimal connections.

Corollary 2. *Under A3 (binary actions), core access $c_{i,t}$ is the sufficient statics for agent i's risk capacity at period t. The risk-capacity, $\gamma_t(c_{i,t})$, decreases in $c_{i,t}$ and*

$$\gamma_t(c_{i,t}) = \frac{1}{2} H\left(\kappa_t + \gamma_{t+1}(\lfloor \frac{c_{i,t}}{2} \rfloor), \kappa_t + \gamma_{t+1}(\lceil \frac{c_{i,t}}{2} \rceil)\right) \; \forall t \leq N. \tag{14}$$

Optimal Core Size Since we have established that there is a unique optimal market structure given any core size c.[17] 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

$$\Pi = \max_c \left\{ -\gamma_1(c) v_1 - \frac{c}{2^N} \phi \right\}. \tag{15}$$

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 4. *(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* (15). *For any two matching matching i and j at period t, 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 $v_{i,t+1} = \left(\frac{\kappa_t + \gamma_{t+1}(c_{j,t+1})}{\Sigma_k \left(\kappa_t + \gamma_{t+1}(c_{k,t+1}) \right)} \right)^2 (2 v_{i,t})$.*

[17] Recall that, an agent i can connect, directly or indirectly, to at most 2^N agents in N rounds of trade, where each type has 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.

5.2 Intensive Margin

5.3 Extensive Margin (Core Size)

5.3.1 Intensive Margin (fixing core size)

- crossional implications across banks
- volume change respect to δ

5.3.2 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.[18] Importantly, agents in our framework can respond in two margins. First, for fixed agents' connections, 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 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 impose following parameterizations:

P1: $\gamma_{N+1}(1) = \kappa$, $\kappa_t = \delta\kappa \; \forall t \leq N$, and $\eta = \frac{\gamma_{N+1}(0)}{\gamma_{N+1}(1)} \in [0, 1)$.

The parameter δ captures the cost of holding risk in an earlier period relative to the terminal period, and η represents the benefit of using the platform, where $\eta = 0$ can be understood as a fully competitive market. Given any (δ, η), one can show that the risk-capacity $\gamma_t^*(c)$ is a homogeneous function of degree 1 in κ. Hence, the optimal core

[18] 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.

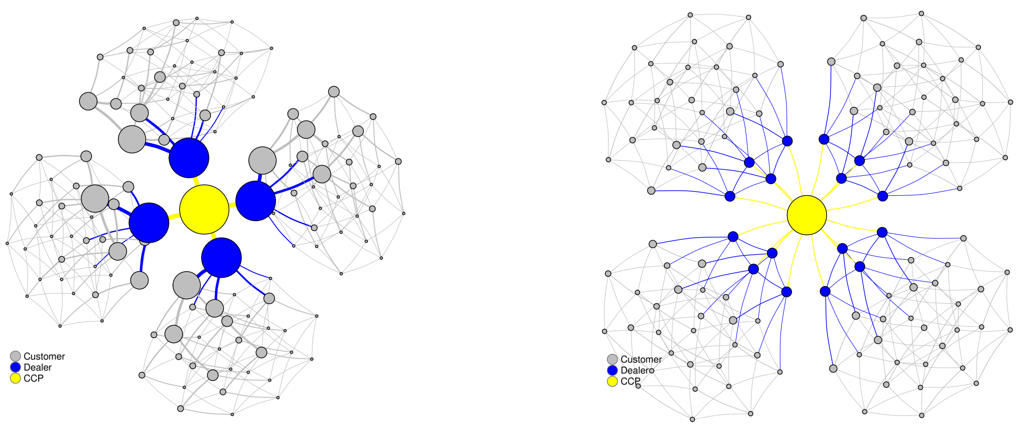

Figure 4: Pre vs. Post-regulation Market Structure.
Each panel shows the graph of the equilibrium trading network. In the network graph, each node represents a bank. The area of the node represents the gross trading volume involving the bank. The edges between nodes represent bilateral trading relationships. 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.

size depends on the entry cost *relative* to level of risks $\frac{\phi}{\kappa v_1}$, where κ represents the balance cost of holding the assets and can be mapped to riskiness of the underlying assets and v_1 represents the ex-ante exposure. 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 when entry costs are lower or when the cost of holding risks are higher.

Proposition 5. *Under P1, given any* (δ, η), *the optimal measure of cores (weakly) decreases with* $\frac{\phi}{\kappa v_1}$.

The effect of reforms that subsides entry or increasing balance sheet costs can thus be understood as resulting in lower $\frac{\phi}{\kappa v_1}$. Figure 4 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).

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.[19] The standard 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.

5.4 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, if there is no gap between private incentives of risk taking and entry-cost.

Welfare-maximizing Policy On the other hand, whenever there are frictions that lead to a deviation between private incentives of risk taking and entry-cost, the equilibrium can be inefficient. According to Proposition 1, such an inefficiency (if it exists) can be corrected by aligning private and social value of risk-taking and/or entry.

Entrenchment by Incumbent Cores One common concern, for example, 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 $\frac{c_0}{2N}$ 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

[19] 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.

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.

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

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

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

The constraints that the planner faces include:

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

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

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

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

Proof. Assumption (2) can be translated into controlled changes in the mean and variance of an agent's asset holding. 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. 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. 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

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

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

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 $(m^2_{i,N+1} + v_{i,N+1})\kappa_{i,N+1} > \phi_i$.

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 $t \le N$, 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,

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

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 $\lambda_{i,j_{i,t},t}^m < 0$ 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 $\alpha_{i,t}^m$, such that $m_{i,t+1} = \alpha_{i,t}^m(m_{i,t} + m_{j,t})$, $m_{j,t+1} = (1 - \alpha_{i,t}^m)(m_{i,t} + m_{j,t})$. 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 $\alpha_{i,t}^m$ should be between 0 and 1. If $\alpha_{i,t}^m$ 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 $\alpha_{i,t}^m > 1$, by setting $\alpha_{i,t}^m$ to 1 reduces $m_{i,t+1}^2$ to $(m_{i,t} + m_{j,t})^2$ and $m_{j_{i,t},t+1}^2$ 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 $0 \leq \alpha_{i,t}^m \leq 1$. 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 $0 \leq \alpha_{i,t}^m \leq 1$.

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

Lemma 6 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 7. *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 6, 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 7 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 Network Properties

A.2.1 Proof for Lemma 2

From Equation 8, let

$$F(\alpha) \equiv -\kappa_t \left\{\alpha^2 + (1-\alpha)^2\right\} V + W_{t+1}(\alpha^2 V) + W_{t+1}((1-\alpha)^2 V)$$

We thus have

$$F'(\alpha) = \left(-\kappa_t + W'_{t+1}(\alpha^2 V)\right) 2\alpha V - \left(-\kappa_t + W'_{t+1}((1-\alpha)^2 V)\right) 2(1-\alpha)V.$$

If $W''_{t+1} < 0$,$F(\alpha)$ is a concave function in α, as

$$\begin{aligned} F''(\alpha) = &\left(-\kappa_t + W'_{t+1}(\alpha^2 V)\right) 2V + \left(-\kappa_t + W'_{t+1}((1-\alpha)^2 V)\right) 2V \\ &+ W''_{t+1}(\alpha^2 V)(2\alpha V)^2 + W''_{t+1}((1-\alpha)^2 V)\left(2(1-\alpha)V\right)^2 < 0. \end{aligned}$$

Hence, $\alpha = \frac{1}{2}$, which satisfies the FOC, is the global maximizer; and thus $v_{i,t+1} = \frac{1}{2}v_{i,t}$ and since all agents are symmetric over time, it is WLOG to assume random matching. Note that, more generally, if all agents start with different $v_{i,1}$, one can show that the sorting is generally NAM as as $\Omega_t(V)$ is concave when $W_{N+1}(v)$ is concave.

A.2.2 Proposition 2

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 $\alpha = \alpha^*(V)$ denote the optimal allocation under V.

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

where the first inequality follows that the surplus under optimal allocation $\alpha^*(\lambda V)$ and $\alpha^*((1-\lambda)V)$ is higher than using the allocation rule $\alpha^*(V)$. 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 $\forall t$.

A.2.3 Proof for Lemma 3

Proof. For any $\alpha(V)$ that satisfies the FOC condition and PAM,we thus have

$$\Omega_t(V|g_t) = \Sigma_k \left\{-\kappa_t\alpha_k^2 V + W_{t+1}(\alpha_k^2 V|g_{t+1}(\alpha_k^2 V))\right\},$$

where $\alpha_i = \alpha(V) = 1 - \alpha_j$.

By Envelop, and $v = 2V$,$W_t(v|g_t) = \frac{1}{2}\Omega_t(2v|g_t)$, we have

$$\begin{aligned} W_t'(v|g_t) = \Omega_t'(2v|g_t) &= \left\{-\kappa_t + W_{t+1}'(\alpha^2 V|g_{t+1}(\alpha^2 V))\right\}\alpha^2 + \left\{-\kappa_t + W_{t+1}'((1-\alpha)^2 V|g_{t+1}((1-\alpha)^2 V))\right\}(1 \\ &= \frac{\prod_{k\in\{i,j\}}\left(-\kappa_t + W_{t+1}'(\alpha_k^2 V|g_{t+1}(\alpha_k^2 V))\right)}{\Sigma_{k\in\{i,j\}}\left(-\kappa_t + W_{t+1}'(\alpha_k^2 V|g_{t+1}(\alpha_k^2 V))\right)} = \frac{1}{2}H(-\kappa_t + W_{t+1}'(\alpha^2 V|g_{t+1}(\alpha^2 V)), -\kappa_t + W_{t+1}'((1-\alpha \end{aligned}$$

, where using the fact that from FOC $\alpha_k = \frac{-\kappa_t + W_{t+1}'(\alpha_{-k}^2 V|g_t(\alpha_{-k}^2 V))}{\Sigma_k\left(-\kappa_t + W_{t+1}'(\alpha_k^2 V|g_t(\alpha_k^2 V))\right)}$. □

A.2.4 Proof for Lemma 4

We prove this result by the following two lemmas.

Lemma 8. *When $\kappa_t = 0$, the sequential setting is equivalent to the static optimization problem, where $\{v_{k,N+1}\}$ maximizes*

$$\begin{aligned} &\max \Sigma_{k=1}^{2^N} W_{N+1}(v_{k,N+1}) \\ &\left[\Sigma_k \sqrt{v_{k,N+1}}\right]^2 = 2^N v_1 \end{aligned} \tag{A.8}$$

Proof. Observe that any $\tilde{v}_k$ that satisfy the pair-wise constraints and PAM must satisfy

$$(\sqrt{v_{i,t+1}} + \sqrt{v_{j,t+1}})^2 = 2v_t$$

Hence,

$$\sqrt{v_t} = \frac{\sqrt{v_{i,t+1}} + \sqrt{v_{j,t+1}}}{\sqrt{2}} = \left(\frac{1}{\sqrt{2}}\right)^{N-t+1} \left\{\Sigma_k^{2^{N-t+1}} \sqrt{v_{k,N+1}}\right\}$$
$$\left(\sqrt{2}\right)^N \sqrt{v_1} = \Sigma_k^{2^N} \sqrt{v_{k,N+1}}$$
$$2^N v_1 = \left[\Sigma_k^{2^N} \sqrt{v_{k,N+1}}\right]^2$$

We now shows that any solution that satisfies Equation A.8, there exists $j^*_\tau(k)$ and $\{v_{k,\tau}\}_{\tau \leq N}$ such that pair-wise constraints and PAM are satisfied. That is, if (i, j) are matched at period t, $j^*_t(i) = j$, then

$$v_{k,t} = \frac{(\sqrt{v_{i,t+1}} + \sqrt{v_{j,N+1}})^2}{2}, \forall k \in \{i, j\}$$

Moreover, we choose $j^*_\tau(k)$ so that if (i, j) are matched on period τ, they cannot be connected at period $\tau + 1$. This thus guarantees that at each period τ, each agent is connected to $2^{N-\tau}$ agents. Since such construction guarantees each agent k is connected to 2^N agents at period 1, $v_{k,1}$ must be the same for all agents, and

$$v_{k,1} = \frac{(\sqrt{v_{k,t+1}} + \sqrt{v_{j^*_t(k),t+1}})^2}{2} = \frac{1}{2}^N \left\{\Sigma_k^{2^N} \sqrt{v_{k,N+1}}\right\}^2 = v_1.$$

□

Lemma 9. *Under Condition 1 and 2, there can be at most two different values of* v_{N+1}*;and (2) if* $\max\{v^k_{N+1}\} > \min\{v^k_{N+1}\}$*, there can be at most one core agent when* $\kappa_t = 0$.

Proof. First of all, at period N, the FOC of 8 yields

$$F_\alpha(\alpha, V) = 2V\{W'(\alpha^2 V)\alpha - W'((1-\alpha)^2 V)(1-\alpha)\}$$

Hence, under Condition 1, at $\alpha = 1$,$F_\alpha(1, V) < 0$, which means that the solution must be interior.

For Result (1), observe that any v_{N+1} must satisfy the FOC from the static problem

$$\sqrt{v_{k,N+1}} \left(W'_{N+1}(v_{k,N+1})\right) = \lambda\sqrt{2^N v_1}, \tag{A.9}$$

where λ is Lagrange multiplier of the constraint A.8. Let $z(v) = \sqrt{v}W'_{N+1}(v)$. Since $z'(v) = \frac{1}{2}\frac{W'_{N+1}(v)}{\sqrt{v}} + \sqrt{v}W''_{N+1}(v) = \frac{W'_{N+1}(v)}{\sqrt{v}}(\frac{1}{2} + v\frac{W''_{N+1}(v)}{W'_{N+1}(v)}) = \frac{W'_{N+1}(v)}{\sqrt{v}}(\frac{1}{2} - vc(v))$. Under the assumption that $c(v)$ is monotonic in v, if there exists $\hat{v} \in [0, 2^N v_1]$ such that $\hat{v}c(\hat{v}) = \frac{1}{2}$, then $z'(v) < 0$ for $v < \hat{v}$, and $z'(v) > 0$ for any $v > \hat{v}$. Hence, there can be at most two roots. If such $\hat{v}$ doesn't exist, then $z(v)$ is monotonic and can only be one root.

For Result (2), let $v^c_{N+1} = \max\{v^k_{N+1}\}$ and $v^0_{N+1} = \min\{v^k_{N+1}\}$. This statement holds automatically when $N = 1$. We now show this holds when $N \geq 2$. Let $v^1_N = \frac{(\sqrt{v^c_{N+1}}+\sqrt{v^0_{N+1}})^2}{2}$, we have $x_N(v^1_N) > \frac{1}{2}$. Suppose that there are $k > 1$, the same outcome is achieved setting $v^2_N = \frac{(\sqrt{v^c_{N+1}}+\sqrt{v^c_{N+1}})^2}{2} > v^1_N$, and let $x_N(v^2_N) = \frac{1}{2}$, which violates that $x_N(v_N)$ increases in v_N. □

A.2.5 Proof For Proposition 3

Proof. We now show that □

A.2.6 Proof For Lemma 5

Let $g_t(v)$ be the set of solutions that satisfies FOC. We now show that if $g_t(v)$ violates the condition, there exists a network $\hat{g}_t$ such that $\Omega_t(v|g_t) < \Omega_t(v|\hat{g}_t)$ for any $\kappa_t > 0$.

Given that the constraint yields $(\sqrt{v_{i,t+2}} + \sqrt{v_{j,t+2}})^2 = 2v_{i,t+1}$, we thus have,

$$\begin{aligned}
\Omega_t(v|g_t) &= -\kappa_t \frac{1}{2}\left\{\left[\sqrt{v_{i,t+2}} + \sqrt{v_{j,t+2}}\right]^2 + \left[\sqrt{v_{i,t+2}} + \sqrt{v_{j,t+2}}\right]^2\right\} + \Sigma_k\left(-\kappa_{t+1}v_{k,t+2} + W_{t+2}(v_{k,t+2})\right) \\
&\leq -\kappa_t \frac{1}{2}\left\{\underbrace{\left[\sqrt{v_{1,t+2}} + \sqrt{v_{4,t+2}}\right]^2}_{v_{1,t+1}} + \underbrace{\left[\sqrt{v_{2,t+2}} + \sqrt{v_{3,t+2}}\right]^2}_{v_{2,t+1}}\right\} + \Sigma_k\left(-\kappa_{t+1}v_{k,t+2} + W_{t+2}(v_{k,t+2})\right) \\
&= \Omega_t(v|\hat{g}_t)
\end{aligned}$$

The first inequality uses the fact that $f(v_i, v_j) \equiv \left[\sqrt{v_i} + \sqrt{v_j}\right]^2$ and $f_{12}(v_i, v_j) > 0$; hence NAM sorting minimizes the flow payoff. The last equality uses the fact that

$$v_t = \left(\sqrt{v_{1,t+1}} + \sqrt{v_{2,t+1}}\right)^2 = \frac{1}{4}\left[\sqrt{v_{1,t+2}} + \sqrt{v_{4,t+2}} + \sqrt{v_{2,t+2}} + \sqrt{v_{3,t+2}}\right]^2.$$

In other words, different matching plan at period $t+1$ only affects changes the flow payoff at period t. Hence, if the condition is violated, then there exists $\hat{g}_{t+2}$ that are identical with g_{t+2} from period $t+2$ onward but its matching plan satisfies negative sorting.

A.2.7 Proof for Lemma 2

Proof. For period $N+1$, the payoff can be rewritten as $W_{N+1}(v) = \max_{c_{N+1}} -\gamma_{N+1}(c_{N+1})v - \phi(c_{N+1})$, where $\gamma_{N+1}(c_{N+1})$ decrease in $c_{N+1} \in \{0,1\}$, and thus if $c_{i,N+1} > c_{j,N+1}$, then it must be the case that $v_{i,N+1} > v_{j,N+1}$. Define $c_{i,N} = c_{i,N+1} + c_{j_t(i),N+1} \in \{0,1,2\}$,the value of $\gamma_N(c)$ is given by Equation 10, which decreases in c.

For $t = N-1$, let

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

We now show that $\{(1,1),(0,0)\}$ is dominated by $\{(1,0),(1,0)\}$. This is because that if $\gamma_{N+1}(c_{i,N+1}) = 1$ and $\gamma_{N+1}(c_{j,N+1}) = 0$, then it must be the case that $v_{i,N+1} > v_{j,N+1}$. Hence, $\{(1,1),(0,0)\}$ implies that $\{(v_{4,N+1}, v_{3,N+1}), (v_{2,N+1}, v_{1,N+1})\}$ and $v_{4,N+1} \geq v_{3,N+1} > v_{2,N+1} \geq v_{1,N+1}$, which thus violated Lemma 5. 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. Lastly, since $\gamma_N(c)$ decrease in c, $\gamma_{N-1}(c)$ thus also increases in c.

By backward induction, assume that $c_{i,t} = \left\{\lfloor\frac{c_{i,t}}{2}\rfloor, \lceil\frac{c_{i,t}}{2}\rceil\right\}$ and let $\gamma_{t+1}(c)$ denote its corresponding risk-capacity, which decrease in c and the value function yields

$$W_t(v) = \max_c \gamma_t(c)v + \phi(c),$$

and hence if $c_{i,t} > c_{j,t}$, then it must be the case that $v_{i,t} > v_{j,t}$.

Suppose that at period t,$c_{i,t} = (m,n)$ where $m - n \geq 2$,then at period $t+2$, we thus have $c_{k,t+2}$ such that

$$c_{4,t+2} = \lceil\frac{m}{2}\rceil \geq c_{3,t+2} = \lfloor\frac{m}{2}\rfloor > c_{2,t+2} = \lceil\frac{n}{2}\rceil \geq c_{1,t+2} = \lfloor\frac{n}{2}\rfloor.$$

Given that $\gamma_{t+2}(c)$ increases in c, this implies that $v_{4,t+2} \geq v_{3,t+2} > v_{2,t+2} \geq v_{1,t+2}$ is satisfied and thus Lemma 5 is violated. This network is thus dominated by rematching agent 1 and 4 at period $t+1$. Hence, 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 2 must hold for any t. $\square$

A.2.8 Proof for Proposition 4

Proof. Since Lemma 2 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

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

$\square$

A.2.9 Proof of Proposition 5

We first show that $\gamma_t^*(c|\delta,\eta,\kappa) = \kappa\gamma_t^*(c|\delta,\eta,1)$ is a homogeneous function of κ. This holds for $N+1$, as $\gamma_{N+1}(1) = \eta\kappa$ and $\gamma_{N+1}(0) = \kappa$. Given the expression of $\gamma_t^*(c|\delta,\eta,\kappa)$ from equation 14, we thus have

$$\begin{aligned}
\gamma_t^*(c|\delta,\eta,\kappa) &= \frac{1}{2}H\left\{\kappa(\delta + \gamma_{t+1}^*(\lfloor\frac{c}{2}\rfloor|\delta,\eta,1)), \kappa(\delta + \gamma_{t+1}^*(\lceil\frac{c}{2}\rceil|\delta,\eta,1))\right\} \\
&= \kappa\frac{1}{2}\left\{H\left(\delta + \gamma_{t+1}^*(\lfloor\frac{c}{2}\rfloor|\delta,\eta,1)\right), \left(\delta + \gamma_{t+1}^*(\lceil\frac{c}{2}\rceil|\delta,\eta,1)\right)\right\}.
\end{aligned}$$

Hence, Equation (15) can be rewritten as $\Pi = \kappa v_1 \max_c \left\{-\hat{\gamma}_1(c) - \frac{c}{2^N}\left(\frac{\phi}{\kappa v_1}\right)\right\}$, where $\hat{\gamma}_1(c) = \gamma_t^*(c|\delta,\eta,1)$. 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 $\gamma_t(1) = \frac{1}{2}H(\delta + \gamma_{t+1}^0, \delta + \gamma_{t+1}^1) \to 0$ $\forall t$ 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.

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".

McAfee, R. P. (1993). "Mechanism design by competing sellers". *Econometrica: Journal of the econometric society*, pp. 1281–1312.

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.