Financial Market Structure and Risk Concentration*

Briana Chang† Shengxing Zhang‡

First Draft: January 2021, This Draft: Aug, 2022

Abstract

We propose a framework that jointly determines bilateral trading networks and risk allocation between banks. Banks use their bilateral connections to share and concentrate their exposures to idiosyncratic risks. Even when banks are ex-ante homogeneous and risk-averse, they may take risks collectively by concentrating risks on a small set of banks. A structural shift in the market structure in response to a small change in fundamentals and regulations is possible, causing discontinuous changes in aggregate risks and transaction prices. The framework is useful for deriving implications of financial market structure on asset price and bank size distribution and evaluating the responses of the market structure to regulations.

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.

‡London School of Economics, Peking University HSBC Business School; szhang@phbs.pku.edu.cn.

1 Introduction

In this paper, we develop a framework that jointly determines interbank trading links and risk allocations through them. Even when banks are ex-ante homogeneous and risk averse, we find that they could trade not only to share risks but also to concentrate them, creating an asymmetric interbank network consistent with empirical regularities: a few banks have large balance sheets, have large gross trading volume, and bear more risks in their asset positions. 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.

Banks are averse to uncertain asset positions and their initial asset positions are subject to idiosyncratic shocks. Interbank trade serves to diversify and reallocate their risky asset positions. In this environment, we connect trading frictions in the OTC market to the limited information that banks have about other banks' asset positions when they form a finite number of trading links.

Specifically, we assume banks are committed to with whom, when, and how they trade (dynamic matchings with their counterparties and terms of trade) before they observe their realized initial asset positions. We require that at any point in time bilateral matchings and terms of trade with their trading counterparties in current and future trading rounds be stable, thus allowing multiple pairwise deviations. The collection of banks' counterparties and trades over all trading rounds forms an endogenous trading network.

A bank's final payoff depends on the riskiness of its asset position after bilateral trades in the OTC market. We allow for a general payoff function and analyze how it affects the equilibrium network. When banks' private marginal cost of bearing risks is diminishing in risk level, the standard risk-sharing strategies can be suboptimal even though banks are risk averse and ex ante identical. Instead, banks may concentrate the risk exposure on one side of bilateral matches.

The diminishing marginal cost of bearing risks is relevant in many applications. For example, when bank have limited liability thus do not internalize the social cost of default, and when banks can invest in trading technologies to lower their risk-bearing cost, by accessing a faster trading platform or searching for trading counterparties more intensively. The marginal cost is diminishing in the case of limited liability because banks whose liability size is more uncertain are more likely to default. In the case of optional investment in trading technologies, banks facing more uncertainty on their asset positions are more likely to invest in faster trading technologies.

Having a few extremely risky banks may be optimal for the interbank network as a whole because they have lower marginal costs of bearing risks and take on risks from other banks. We can think of a bank's marginal cost of bearing risks as its risk bearing capacity. Risky banks have greater risk bearing capacity. But concentrating risks on a subset of banks generates greater overall risk exposure than sharing risks evenly across banks because less risks are diversified away through interbank trade. The equilibrium network is determined by the trade-off between risk concentration and risk sharing. An asymmetric network emerges when the benefit of risk-concentration outweighs the loss from less diversification.

As banks decide the sequence of bilateral trades under limited information, the belief about a bank's asset position distribution at the beginning of each trading round is the key dynamic element. It depends on with whom the bank has traded and how it has traded with the counterparties. And how it evolves over future trading rounds affects the expected payoff from matching and trade in the current trading round. We show that the variance of a bank's position, which can be interpreted as the risk that a bank bears, is a sufficient static for network formation. The main equilibrium objects are thus the evolution of banks' risk exposures (how they trade over times) and their counterparties (with whom they trade).

To understand how risks are concentrated through dynamic connections, we first show that, when banks have diminishing marginal cost of bearing risks, the interbank network concentrates risk via positive sorting. Riskier banks, who bear more risks through past transaction, are matched with riskier banks. We then show how banks allocate risks within a match endogenously depends on their future connections, because future connections and trade determine their marginal cost of bearing risks, or risk bearing capacity. A bank's risk-bearing capacity evolves over time: its current-period capacity equals the harmonic mean of the next-period capacities of the bank and its period-t counterparty. Lastly, because the benefit from concentrating risks materializes at the end of the dynamic trading game, we show that the globally optimal network delays risk concentration and thus, banks have more symmetric risk capacity in earlier trading rounds.

Having established the general properties of the equilibrium network, we apply the framework to a few application. The first application studies the market structure when banks have limited liability thus their payoffs are misaligned against their contribution to the society. There we show that a small increase in the balance sheet cost of holding the asset, either because of tightening regulation or because the asset becomes riskier, can result in a regime shift in the interbank network, whereupon banks switch from sharing risks with each other 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 the interbank trading networks.

Endogenous default risks in this application are different from those in standard theories of financial contagion. Standard theories take the interbank network as given and analyze how network amplifies and dampens the propagation of default risks ex-post. We study equilibrium network formation and highlight the ex-ante aggregate default risks can increase when banks systematically change their trading behaviors through the interbank network. Hence, even without contagion through interbank credit, we highlight another source of systematic risks through the network.

The second application studies how the option of investing in faster trading technologies affects the bilateral interbank market structure. Specifically, banks can choose whether or not to access a multilateral platform to increase their risk-bearing capacity at a fixed entry fee in the final period. There the property of delayed risk concentration is enough to pin down the unique optimal network given any size of core banks, those who invest in faster trading technologies. The optimal network is then reduced to choosing the optimal core size, trading off the cost of bearing risks versus the entry fee. Core size as a 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[1] contributes a tractable approach to studying the formation of trading network. It differs from existing approaches in the network formation literature[2] 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 her characteristic, which becomes the state variable governing how she trades in later periods. By imposing sequential rationality, we can solve the network formation problem through backward induction.

While we use pairwise stability to characterize the equilibrium matching in a subgame, a deviating agent in a subgame can change all her future links, not just one link as in the static setup that the literature usually adopts. 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.

Popular approaches to modeling OTC markets are based on random matching (e.g., Duffie, Gârleanu, and Pedersen 2005) or exogenous networks.[3] Relative to the litera-

[1]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., her 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.

[2]See the survey in Jackson 2005 for overview. Specifically, papers that have studied network formation in the financial market include Hojman and Szeidl (2008), 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.

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

ture 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.

One of our applications is on the joint determination of the bilateral trading network and platform access. So, our paper also sheds new lights on the literature on the costs and benefits of centralized vs. decentralized markets.[4] Instead of focusing on the trade-off between these two markets, we allow for nonexclusive participation in both markets and emphasize how the participation decision in the centralized markets interact with the structure of the bilateral OTC market. 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 environment where agents can access multiple types of platforms.

2 A Model of Trading Network Formation

2.1 Model Setup

The economy lasts $N+1$ periods, indexed by $t = 1, 2, \ldots, N+1$. It is populated by a continuum of banks of total measure 1, indexed by identity $i \in \mathbb{I} = [0, 1]$. Each bank is managed by a banker whose preference and choices governs the bank. There are two types of consumption goods, numeraire goods and dividend goods, and one type of asset. The asset is a claim to a unit of dividend good in each period.

A bank i receives a random initial asset position $a_{i,1}$, which is independently and identically distributed across banks, drawn from distribution $\pi_1(a)$. The randomness in the initial asset position represents a liquidity shock that shifts the bank's asset position away from its ideal position. It could be withdrawal of deposits from its customers. Or it could be a shock to a bank's valuation over an asset. We discuss the latter interpretation in more detail after specifying the bank's preference. Banks trade bilaterally from period 1 to period N. They have deep pockets in numeraire goods. In each period from $t = 1$ to $t = N$, there is market place where banks meet bilaterally and exchange the asset with

[4] 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.

numeraire goods.

The preference of a bank i is

$$\mathbb{E}_1 \left\{ \sum_{t=1}^{N} [u_t(\widetilde{a}_{i,t}) - x_{i,t}] + u_{N+1}(a_{i,N+1}) \right\}, \tag{1}$$

where $\mathbb{E}_1$ denotes the expectation in the beginning of period 1, $u_t : \mathbb{R} \to \mathbb{R}$ is a concave utility function, $\widetilde{a}_{i,t}$ denotes the amount of dividend goods bank i consumes in period $t = 1, 2, \ldots, N$, which equals the bank's posttrade asset position in that period, $x_{i,t}$ denotes the amount of numeraire goods the bank pays in exchange for assets.

The curvature in the utility function can be associated with the balance sheet costs of holding assets, which can be affected by regulations, the liquidity preference of a bank's depositors, for example. The heterogeneity in asset positions is the source of gains from trade because banks are risk averse. Transfers $x_{i,t}$ is a result of the transaction. $x_{i,t} = p_{i,t}(\widetilde{a}_{i,t} - a_{i,t})$, where $a_{i,t}$ denotes the pretrade asset position and $p_{i,t}$ denotes the price of the asset. The pretrade position in the following period, $a_{i,t+1}$, for $t = 1, 2, \ldots, N$, equals the posttrade asset position in period t, $\widetilde{a}_{i,t}$. So consumption of dividend goods in period $N+1$ equals the posttrade asset position in the penultimate period, $a_{i,N+1} = \widetilde{a}_{i,N}$.

General preferences allowing both preference shocks and asset endowment shocks, like in the frontier model in the literature on the OTC market, Üslü (2019), are allowed in our setup. Agents' flow utility from asset position $a_{i,t}$ in Üslü (2019) is $-\varepsilon_{i,t} a_{i,t} - a_{i,t}^2$, where $\varepsilon_{i,t}$ is an idiosyncratic preference shock. It is equivalent to $-(a_{i,t} - \bar{a}_{i,t})^2$ where $\bar{a}_{i,t} = -\varepsilon_{i,t}/2$. So the preference shock, $\varepsilon_{i,t}$, is equivalent to a shock to the ideal asset position, $\bar{a}_{i,t}$. Agents receive shocks to their ideal asset positions rather than their asset holding as in our setting. If we think of the asset position in a bank's preference (1) as the deviation of from the ideal position, analysis in the rest of the paper applies to the more general setting.

Formation of Ex Ante Trading Network At the beginning of period 1, banks choose and commit to bilateral trading counterparties for periods $t = 1, 2, \ldots, N$ and their trading strategies with their counterparties. Denote the trading counterparty of a bank i in period t $j_{i,t}$. The collection of a bank i's counterparties $j_{i,t}$ over N rounds of trade forms her trading links. We assume that banks form their trading links unconditional on their realized asset holdings and valuations. Therefore, our setup effectively has a network formation stage ex ante. We can interpret trading links as permanent trading relationships between banks when we repeat the trading game with a fresh draw of shocks.

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 because trading needs can be banks' private information at the trading stage. Without this assumption and if trading needs were banks' private information, 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.

Contacting Frictions in Bilateral Trades Because trading counterparties are determined before banks receive shocks on their asset positions, bilateral trading counterparties are chosen subject to limited information which prevents banks from locating ideal trading counterparties. Banks face uncertainty about their counterparty's asset position *before* contacting their counterparties in the corresponding.

But information is only limited at the matching stage in the beginning of period 1. Matched banks have complete information about each other's asset positions after they make contact, upon which they observe their counterparties' pretrade asset positions in the corresponding period.

If all banks could observe each other's realized positions before they choose their matches, the economy could achieve perfect risk sharing with one round of trade. For example, if banks' utility has a bliss point at 0, is symmetric around the bliss point, and the asset distribution is symmetric around 0, 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 that prevents banks from locating their ideal trading counterparties.

Terms of Trade: Contingent Asset Flows and Prices While the connections are determined ex ante, trades depend on the realized asset positions of a bank and her counterparties, because trading takes place after she and her counterparties 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, banks' realized asset positions change how they trade (i.e., the asset flows) within the network from day to day, even though the network remains the same.

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 $y(i,j) = \{\widetilde{a}_k(a_i,a_j), \widetilde{x}_k(a_i,a_j), k \in \{i,j\}\}$ be the terms of trade within the match (i,j), where $\widetilde{a}_k(a_i,a_j)$ denotes the posttrade asset holding of bank k, and $\widetilde{x}_k(a_i,a_j)$ denotes the transfer to bank k, $k \in \{i,j\}$. The within-match transfers sum up to zero,

$$\widetilde{x}_i(a_i,a_j) + \widetilde{x}_j(a_i,a_j) = 0. \tag{2}$$

The within-match asset allocation is feasible if

$$\widetilde{a}_i(a_i,a_j) + \widetilde{a}_j(a_i,a_j) = a_i + a_j. \tag{3}$$

The allocation of asset positions is associated with the allocation of risks from uncertain asset positions because given a distribution of banks i and j's pretrade asset positions, the posttrade positions also follow a distribution, which is the key characteristic that governs bilateral matching.

While the terms of trades are contingent on the realized positions within a pair, banks are committed to the contingent terms of trade ex ante. This is a strong assumption.

Sequential Choices of Trading Links and Terms of Trade Banks play a sequential game at the beginning of period one when they 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. They constitute the information set contingent on which banks' period t strategies are chosen.

Notice that the information set for period t strategy does not include the realized trading history contingent on that realized asset positions of a bank and her counterparties. This is consistent with our assumption that both the trading network and trading strategies are decided ex ante not contingent on the realized trading history.

A bank i's strategy in period t includes the choice of her counterparty, $j_{i,t}$, and the terms of trade with the counterparty, $y_t(i,j)$ for $j = j_{i,t}$ conditional on the information set for that period. Period-t strategies are sequentially optimal given the common information set.

The common information set for period t strategies can be summarized by the joint distribution of banks' asset positions. As we will see later, the gains from trade from period t onwards depend on the trading history only through the joint distribution. We thus study a dynamic matching model with the joint distribution of banks' asset holdings and the marginal asset distribution as the evolving characteristics.

Evolving Characteristics Now that banks' strategies are contingent on the public belief, characterizing its evolution over time is an essential part of our analysis. 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 in period 1. That is, her asset position in the next period equals the sum of her and her counterparty j's current asset positions, $a_{i,2} = a_{i,1} + a_{j,1}$. 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 $\pi_{i,t}(a) : \mathbb{R} \to [0,1]$. Her posttrade asset distribution $\pi_{i,2}(a)$ now has mean zero and variance $2v_1$ 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,

$$\pi_{i,t+1}(a) = \int \int \mathbb{I}(\widetilde{a}_{i,t}(a_i, a_j) \leq a) \boldsymbol{\pi}_{i,j,t}(da_i, da_j), \text{ for } a \in \mathbb{R}, \tag{4}$$

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, $\widetilde{a}_{i.t}(a_i, a_j)$.

2.2 Equilibrium Definition

Denote the joint payoff between two banks i and j conditional on equilibrium trading strategies before the trading round $\Omega_t(i,j)$. Given the aggregate distribution in period t, the joint payoff maximizes their joint flow utility and continuation value posttrade,

$$\Omega_t(i,j) \equiv \max_{\widetilde{a}_{i,t},\widetilde{a}_{j,t}} \mathbb{E}_1 \left[u_t(\widetilde{a}_{i,t}) + u_t(\widetilde{a}_{j,t}); \left(y^t, j^t\right)\right] + \widehat{W}_{t+1}(i) + \widehat{W}_{t+1}(j) \tag{5}$$

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 (5) because they sum up to zero. $\widehat{W}_{t+1}(i)$ denotes the bank's maximum payoff in the next period for marginal distribution $\pi_{i,t+1}(a)$ that resulted from asset allocations $\widetilde{a}_{i,t}$ and $\widetilde{a}_{j,t}$ and joint distribution with other banks' asset holding.

Taking the aggregate distribution $\boldsymbol{\pi}_{t+1}$ and other banks' equilibrium payoffs $W_t(j)$ as given,

$$\widehat{W}_t(i) \equiv \max_j \Omega_t(i,j) - W_t(j). \tag{6}$$

for $t \leq N$. On the equilibrium path, a bank's payoff is given by $W_t(i)$, which equals $\widehat{W}_t(i)$ for a bank i that adopts equilibrium strategies before period t.

The equilibrium in our model can be understood as competitive equilibrium in the literature of large games (McAfee 1993). Because there is a continuum of banks, a bank's decisions affect her own payoff taking as given the aggregate distribution of matching and trading decisions in the market and have negligible effect on the aggregate distribution. A bank's deviating decision thus does not affect her counterparties' outside option and their payoff from the deviation.

Definition 1. Given the initial distribution of asset positions $\boldsymbol{\pi_1}$, an equilibrium consists of strategies $\{s^*_{i,t}\}_{\forall i,t}$, market utilities $W_t(i)$, and a path of common beliefs $\boldsymbol{\pi}^*_t$ such that the following properties hold for all $t \in \{1, \ldots, N+1\}$:

1. Bilateral matches are stable. For any period $t \leq N$, if bank j is bank i's optimal counterparty, $j \in j_t(i)$, it solves (6), where the post-trade position $\{\tilde{a}_{i,t}, \tilde{a}_{j,t}\}$ maxi-

mizes Equation (5) and there is no profitable deviation when a measure ϵ of traders simultaneously deviate.

2. Feasibility of bilateral matching in any trading round $t \leq N$.

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

The equilibrium can be understood as multiple rounds of pairwise stable matching and trading. Our solution is stronger than the static pairwise stability in the following sense. First, our sequential setting allows for agents to deviate across periods. When an agent deviates in a period t, she can switch her own *subsequent* trading partners accordingly, provided that she promises her counterparties at least their equilibrium payoff $W_{t+1}(j)$. Second, we allow for a simultaneous deviation among a measure ϵ of agents in period t. Therefore, the deviation in a period t will not be constrained to the equilibrium distribution of $\boldsymbol{\pi_t}$, as such agents are free to create any $\tilde{\pi}_{t+1}$ among this group. The minimal number of agents in the deviation in period t so that the deviation is flexible is 2^{N-t+1}. This allows agents in a joint deviation to form matches flexibly not only in the current period but also in later periods. Since the measure of any finite number of traders is zero, the measure ϵ can be any positive value. We assume that the measure ϵ is sufficiently small so that the effect of a single pairwise deviation on the aggregate distributions is negligible in our environment with a continuum of agents, and thus agents can take equilibrium market utility as given.

2.3 Equivalence and Uniqueness

We first establish that the equilibrium outcome is unique and maximizes the aggregate payoff. Denote the aggregate payoff of the economy in period t to be Π_t, which depends on the joint asset distribution $\boldsymbol{\pi}_t$. Given a strategy s_t in period t, the aggregate payoff equals

$$\Pi_t(\boldsymbol{\pi}_t) = \mathbb{E}_1 \int_0^1 u_t(\widetilde{a}_{i,t}) di + \Pi_{t+1}(\boldsymbol{\pi}_{t+1}). \tag{7}$$

where $\mathbb{E}_1(u_t(\widetilde{a}_{i,t})) = \int\int u_t(\widetilde{a}_{i,t})\pi_{i,j_t(i),t}(da_i, da_{j_t(i)})$ and the terminal payoff is given by $\Pi_{N+1}(\boldsymbol{\pi}_{N+1}) = \mathbb{E}_1 \int_0^1 u_{N+1}\left(\widetilde{a}_{i,N+1}\right) di$.

The following proposition first shows that 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.[5] Second, when a deviation arises for various 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.

Although the equilibrium is constrained efficient taking banks' preferences as given, the equilibrium is socially optimal only if banks' private payoff is aligned with the social payoff. When there is a gap between the private payoff and the social payoff, we can use our framework to evaluate the divergence of the equilibrium market structure from the socially optimal structure.

3 Risk Distribution and Network Structure

We now focus on a more specific banks' risk preferences and analyze the resulting risk distribution and network structure.

Assumption 1. *For all trading round $t \leq N$, the flow utility is a quadratic function with bliss point at 0, $u_t(a_{i,t}) = -\kappa_t a_{i,t}^2$ where $\kappa_t \geq 0$ is a parameter for the flow cost in period t.*

Assumption 2. *In the final period $N+1$, the expected payoff of bank i, denoted by $W_{N+1}(v_{i,N+1})$, is a decreasing function of post-trade risk-exposure, where $v_{i,N+1}$ is the variance of $\pi_{i,N+1}(a)$.*

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

Assumption A1 can be understood as mean-variance utility from dividend goods, where the mean (i.e., ideal asset position of a bank) is normalized to zero.[6] Parameter κ_t then represents the balance sheet cost of holding nonzero asset positions during trading period t, which can be associated with the riskiness of the asset.

Assumption A2 allows the expected terminal payoff to be a general decreasing function of the variance of the post-trade asset position. We assume a decreasing function to avoid trivial risk-taking behaviors. A1 and A2 together imply that holding risks is fundamentally costly for all banks, which captures the standard risk-sharing incentives.

As we show below, Assumptions A1 and A2 allow us to reduce the relevant state variable to the *variance* of the distribution $\pi_{i,t}(a)$, which represents a bank's risk exposure over time.

Despite that all banks being risk-averse, we show below that whether banks would engage in risk-sharing crucially depends on the convexity of $W_{N+1}(v)$ – a convex $W_{N+1}(v)$ represents diminishing marginal costs of bearing risks. Note that the mean-variance utility for the terminal payoff can be nested as a special case in which the terminal payoff function $W_{N+1}(v)$ is linear in v thus represents a constant marginal cost of bearing risks. Below are two examples, which we discuss in more detail in Section 4 and 5, that naturally results in a convex terminal payoff function $W_{N+1}(v)$.

Application 1 (Limited Liability) When banks are protected by limited liability, the marginal cost of taking additional risks can be lower for riskier banks because they are more likely to default and thereby offload the cost of holding low asset positions to such external creditors as depositors.

For example, if the utility from final consumption of dividend goods is a CARA utility function, $u(c) = 1 - e^{-c}$, the expected value given the variance of asset position v is

$$W_{N+1}(v) = \int [1 - \exp(\max(a_{N+1}, -D))] \, d\pi(a_{N+1}) \tag{8}$$

where $D > 0$ denotes the face value of debt that the bank owes to depositors. We normalize the expected value of a_{N+1} to 0 and assume that the expected position is fixed due to regulation. One can then verify that $W_{N+1}(v)$ can be decreasing and convex in v

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

for $v \leq 2^N v_0$ for some value of debt face value D.

Application 2 (Optional Investment in Risk-Sharing Technologies) Banks in practice have options to improve their risk-bearing capacity by investing in superior but more costly trading technologies. These options lower their marginal cost of bearing risks and are more valuable for banks more exposed to uncertain asset positions.

For example, if banks can access a competitive trading platform with probability $1-\eta$ by paying fixed cost ϕ, their expected payoff from accessing the platform can be expressed as $-\phi-\eta\kappa_{N+1}a^2$ if we normalize the market price to 0 and assume that banks' final period payoff is quadratic in their asset holding, $u_{N+1}(a) = -\kappa_{N+1}a^2$.[7]

We can then show that the terminal payoff function

$$W_{N+1}(v) = \max\{-\kappa_{N+1}v, -\phi - \eta\kappa_{N+1}v\}.$$

If a bank does not pay the fixed cost to access the trading platform, the expected payoff is $Eu_{N+1}(a) = -\kappa_{N+1}v$. If she does, she reaches the ideal asset position 0 by trading the asset at market price 0 with probability $1-\eta$. The bank chooses to invest in the technology and reduce the marginal cost of bearing risks to $\eta\kappa_{N+1}$ if and only if the pretrade variance v is greater than $\frac{\phi}{(1-\eta)\kappa_{N+1}}$. So the value function is piece-wise linear and convex in v.

The convexity of the terminal payoff function $W_{N+1}(v)$ holds more generally when a bank has multiple options of investment in faster trading technologies and the decision can be contingent on realized asset positions. We provide such an example in Section A.1 in the appendix.

[7] Under a centralized platform, participating agents only care about the the market-clearing price, denoted by p. Specifically, if an agent chooses to enter CM with position $a_{i,N+1}$, he can buy and sell assets at the price p. Her expected payoff yields

$$-\phi + \int \max_{\tilde{a}} \left[p\left(a_{i,N+1} - \tilde{a}\right) - \kappa_{N+1}\tilde{a}^2\right] d\pi_{i,N+1}(a_{i,N+1}),$$

By the law of large numbers, the market clearing price must be such that all agents that participate CM can adjust their positions to the target level (i.e, $\tilde{a} = 0$), conditional on all participants' asset distribution is symmetric around zero. In this case, the final payoff of a participant is reduced to $-\phi$.

3.1 Risk Sharing vs. Risk Concentration in Bilateral Trade

Reformulation: Variance Representation Instead of working with asset allocations, we first simplify the analysis by reformulating the problem to work with the variance of asset position distribution (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, denote the variance of posttrade positions $\tilde{v}_k \equiv Var(\tilde{a}_k(a_i, a_j))$ and the variance of the sum of pretrade positions $V_{ij} \equiv Var(a_i + a_j)$. The feasibility constraint on bilateral trade, Equation (3), implies the following connection between pretrade and posttrade risk:

$$\tilde{v}_i + \tilde{v}_j + 2\tilde{\rho}_{ij}\sqrt{\tilde{v}_i\tilde{v}_j} = V_{ij}, \tag{9}$$

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

Lemma 1. *Under Assumptions 1 and 2, the optimal marginal distributions of the posttrade position for all banks have zero mean, and the posttrade positions for any two matched banks are perfectly positively correlated ($\tilde{\rho}_{ij} = 1$). The pretrade positions of any two matched banks in the efficient solution are uncorrelated.*

Since banks' payoff decreases with the variance and mean of their asset positions, it is optimal to keep the means of their posttrade positions at zero and change only the correlation and variances of their postrade positions.

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 (9). So, all else equal, it is optimal to match banks whose asset positions are not correlated (negative correlation is not available when banks trade optimally). This observation allows us to focus 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 her 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 her characteristic $v_{i,t}$.

Allocation of Risks Within Matches Given that an agent's continuation value can be summarized by her posttrade variance and posttrade correlation is one (Lemma 1), the optimal asset/risk allocation within any match (i, j) can thus be reformulated as choosing optimally the share of the risks within a pair (i, j), given their pre-trade variance V_{ij}, 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. By abusing the notation, we use $\Omega_t(V_{ij})$ to represent the joint payoff between any two agents with pre-trade variance V_{ij} in period t. The optimal share $\alpha \in [0, 1]$ maximizes thus solves the joint expected payoff

$$\Omega_t(V_{ij}) = \max_{\alpha\in[0,1]} \left\{-\kappa_t \left(\alpha^2 + (1-\alpha)^2\right) V_{ij} + W_{t+1}(\alpha^2 V_{ij}) + W_{t+1}((1-\alpha)^2 V_{ij})\right\} \tag{10}$$

Given any match, the optimal risk allocation within the pair must satisfy the FOC condition from Equation (10), so the share of total asset positions allocated to an agent

$$\alpha = \frac{\kappa_t + W'_{t+1}((1-\alpha)^2 V)}{\left(\kappa_t + W'_{t+1}((1-\alpha)^2 V\right) + \left(\kappa_t + W'_{t+1}(\alpha^2 V)\right)}. \tag{11}$$

Observe that share $\alpha = \frac{1}{2}$ represents the standard risk-sharing solution where agents share their exposure equally with any match, which also minimizes the sum of post-trade variance. Any concentration of risks ($\alpha \neq \frac{1}{2}$) is costly in the sense that it leads to higher post-trade pair-wise variance and such a cost increases with banks' balance cost κ_t.

3.2 Dynamic Risk Concentration via Sorting on Risk Exposures

In our framework, agents not only choose the risk allocation within a match but also whom they trade with. The joint determination of these two decisions pins down the underlying trading network.

The optimal solution of (10) crucially depends on the property of the value function $W_t(v)$, which endogenously depends on the optimal choice of counterparties and can be expressed as

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

where we use the fact that $V_{ij} = v_i + v_j$ as pretrade positions of any two matched banks are uncorrelated (Lemma 1). The proposition below first establishes how the property of the terminal payoff function $W_{N+1}(v)$ affects the dynamic matching outcomes.

Proposition 2. *(1) (Full Risk-sharing and Random Matching) When the terminal payoff function $W_{N+1}(v)$ is concave in variance v, the unique trading network is full risk-sharing and the matching outcome is equivalent to random matching. (2) (Positive Assortative Matching (PAM) on Risk Exposure) When the terminal payoff function $W_{N+1}(v)$ is convex in variance v, the optimal sorting outcome is PAM on variance v_t in any trading round t. Within any match in which agents have the same variance v_t, the optimal share maximizes joint payoff $\Omega_t(2v_t)$.*

With a concave terminal payoff function $W_{N+1}(v)$, the solution is well understood: risk-sharing ($\alpha = \frac{1}{2}$) is the unique global maximum as the objective function in Equation (10) is concave in α. Agents share their exposure equally with any match over time and thus $v_{i,t} = \frac{1}{2}v_{i,t-1} = \left(\frac{1}{2}\right)^t v_0 \forall i, t$. 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 $W_{N+1}(v) = -\kappa_{N+1}v$.[8] [9]

We focus on the case in which $W_{N+1}(v)$ is convex throughout the rest of the paper. 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 risk exposure, thus on the equilibrium path, $V_{ij} = 2v_i$. In other words, agents that take on risks from others (higher post-trade variance $v_{i,t+1}$) are matched with each other. Through this channel, these agents on average handle more risks compared with random matching, where the risk exposure of their counterparties next period is drawn randomly.

Figure 1 illustrates an example where risk concentration and PAM arise. Within a match, we use the arrow to point toward the agent with higher post-trade variance

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

[9]More generally, concavity in $W_{N+1}(v)$ predicts negative assortative matching (NAM). 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.

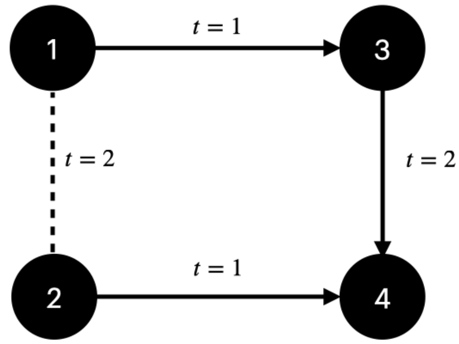

Figure 1: Risk-concentrating Network with PAM ($N = 2$)

if asymmetric allocation arises and the dashed line to represent equal risk-sharing. In this example, Agents 3 and 4 take on more risks from Agents 1 and 2 in period 1, thus have higher post-trade variance. In period 2, PAM implies that Agent 4 (Agent 1) is matched with Agent 3 (Agent 2). That is, an agent who gets a higher share of risks from her counterparty in period $t-1$ then matches with another agent who also holds more risks from past transactions. Through this dynamic matching process, risks can be concentrated on a smaller set of agents within the network.

3.3 General Properties

3.3.1 Risk-bearing Capacity

According to Equation (11), the allocation within the pair in period t crucially depends on the marginal cost of bearing risks $-W'_{t+1}(v_k)$. We refer the marginal cost of holding risk for an agent with v_t as her risk-bearing capacity in period t. The capacity depends on her connections and trades from period $t+1$ onwards. The lemma below establishes that the risk-bearing capacity can be characterized recursively.

Lemma 2. *The marginal cost of holding risks for agent with position v in period t is given by*

$$W'_t(v) = \frac{1}{2}H\left(\kappa_t + W'_{t+1}(\tilde{v}_h(v)), \kappa_t + W'_{t+1}(\tilde{v}_l(v))\right) \ \forall t \leq N. \tag{12}$$

Equation (12) has a simple interpretation: the risk-bearing cost of an agent i in period t depends on the harmonic mean[10] of the post-trade risk-bearing cost of bank i and her

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

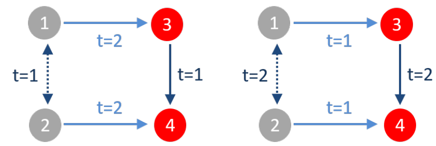

Figure 2: Late vs. early Concentration ($N = 2$)

counterparty $j_t(i)$. It also shows that, while currently matched agents can have different capacities next period, they currently must have the same capacity because matching is positive assortative and matched agents allocate the risks jointly taking into their future connections.

3.3.2 Delayed Risk Concentration

When $W_{N+1}(v)$ is convex in the variance of asset position distribution v, there in general exists multiple locally optimal solution that satisfy first order necessary conditions and PAM. They are not necessarily global optima. While finding the globally optimal solution – which is also the unique equilibrium outcome (Proposition 1) – is generally difficult analytically, we now further establish a general property for the globally optimal network. As shown in Section 5, this property is sufficient to pin down the unique solution in the case for binary choices at the end of the trading game.

Consider Figure 2 with the total number of trading rounds $N = 2$ and thus four banks could potentially be connected. Suppose that the bilateral trading outcome under both networks is such that Banks 3 and 4 (Banks 1 and 2) have higher (lower) post-trade variance, given by $v_{i,N+1} = v_H$ for $i = 3, 4$ ($v_{i,N+1} = v_l$ for $i = 1, 2$). They, however, differ in terms of the timing of the bilateral connections. In the left graph of Figure 2, Bank 1 is first connected to Bank 2 and then Bank 3, but this order is reversed in the right graph.

Observe that the ordering of matching outcomes must result in different dynamic paths of $v_{i,t}$ despite of having the same final outcome of $v_{i,N+1}$. Specifically, in order to concentrate risks on Agents 3 and 4, risk concentration takes place in period 2 for the left graph but in period 1 for the right graph. Since concentration necessarily results in higher total variance and is costly, it is optimal to delay the risk concentration whenever the flow marginal cost of bearing risks in period $t \leq N$ κ_t is strictly positive. Thus, any solution that violates back-loading property is dominated.

The lemma below establishes an additional necessary condition for the optimal path of $v_{i,t}$. Given that the solution could be asymmetric, we let $\tilde{v}_\theta(v)$ denote the post-trade variance within the pair v, where $\theta \in \{h, l\}$ and $\theta = h$ ($\theta = l$) represents the agent who takes on more (less) risks within the pair who both have pretrade variance v. For agents that begin with v_t in period t, there could be at most 2 post-trade variances, given by $\tilde{v}_{\theta_t}(v_t)$ for θ_t in $\{h, l\}$. Then, there could be four paths of asset position variances in period $t+2$, $v_{t+2} = \tilde{v}_{\theta_{t+1}}\left(\tilde{v}_{\theta_t}(v)\right)$ for θ_t and θ_{t+1} in $\{h, l\}$.

Lemma 3. *When the flow marginal cost of bearing risks $\kappa_t > 0$, the optimal solution must satisfy*

$$\tilde{v}_h(\tilde{v}_h(v_t)) \geq \tilde{v}_h(\tilde{v}_l(v_t)) \geq \tilde{v}_l(\tilde{v}_l(v_t)) \geq \tilde{v}_l(\tilde{v}_h(v_t)).$$

To understand the lemma, consider banks i and j who both take on more risks in period t (i.e., $v_{i,t+1} = v_{j,t+1} = \tilde{v}_h(v_t)$). Because of PAM, both of them are matched in period $t+1$. Lemma 3 implies that, if bank j unloads more risks to bank i , then her post-trade variance must be the lowest among all other values of $\tilde{v}_{\theta_{t+1}}\left(\tilde{v}_{\theta_t}(v)\right)$. In other words, while bank j takes more risks in period t, her post-trade variance must be the lowest after unloading more risks to bank i in period $t+1$. Bank j in fact becomes safer after trading with bank i. The right graph violates the condition above, as bank 3 takes on more risks in period 1 but still his final risk-position remain higher that others (as $v_{3,N+1} = v_H > v_{1,N+1}$).

Formally, we prove this by showing that, if this condition is violated, fixing $v_{k,t+2}$ but changing the ordering of the matches among these four banks lowers the total variance of $v_{k,t+1}$. Hence, whenever $\kappa_t > 0$, such a deviation is profitable.

3.3.3 Mapping to Trading Network

Let the solution onward for banks with variance v_t in period t be summarized by $g_t(v_t) = \left\{\tilde{v}_{\theta_N}\left(\tilde{v}_{\theta_{N-1}}\left(...\tilde{v}_{\theta_{t+1}}\left(\tilde{v}_{\theta_t}(v_t)\right)\right)\right)\right\}_{\forall\theta_\tau\in\{h,l\},\tau=t,t+1,...,N}$. It is the set of dynamic paths of variance that begin with variance v_t in period t. We now map the solution $g_t(v_t)$ and the bilateral links that implement the paths in the set to the underlying network.

Ex-ante Network While all agents are ex-ante homogeneous, the set of paths $g_1(v_1)$ summarizes at most 2^N different paths of variance from period 1 to period N. It can be interpreted as the network among 2^N types of agents, where each type of agents is a fraction $\frac{1}{2^N}$ of the group and is characterized a vector $i^k = \{\theta_\tau^k\}_{1 \le \tau \le N}$, where $\theta_\tau^k \in \{h, l\}$ indicates if agent k takes on higher or lower variance within her match in period τ. Note that variance $\tilde{v}_h(v_t)$ is only weakly greater than variance $\tilde{v}_l(v_t)$. While there are 2^N types of agents, this definition also allows for agents to have the same realization.

Corollary 1. *Given the set of dynamic paths of variance that begins with variance v_1 in period 1, $g_1(v_1)$, the bilateral links for each type $k \in \{1, 2, ..2^N\}$ can be constructed as follows: in a trading round t,if her counterparty is of type k', $j_t(i^k) = i^{k'}$ then her path overlaps with her counterparty's path before the period, $\theta_\tau^k = \theta_\tau^{k'}$ for $\forall \tau < t$ and their path diverges in period t $\theta_t^k \neq \theta_t^{k'}$.*

The condition $\theta_\tau^k = \theta_\tau^{k'}$ for $\forall \tau < t$ guarantees that any two matched agents in period t have adopted the same path in the past and thus have the same variance v_t. Thus matching them in period t satisfies PAM. Moreover, if they meet in period t, by definition, their posttrade risk exposure characterized by $\theta_t^k \in \{h, l\}$ must differ, which also implies that they will not be matched a second time.

Dynamic Evolution of Connected Banks Our sequential formulation implies that variance v_t summarizes the effect of earlier connections and trading outcomes. Conditional on variance v_t, the allocation in period t only depends on the future direct and indirect connections moving forward, summarized by the set of dynamic paths of variance that begin with variance v_t in period t, $g_t(v_t)$. [11] In other words, from the viewpoint of an agent with v_t, the allocation of any other agent with different value of v_t is no longer relevant. In this sense, while asset allocation among the 2^N types of agents are interdependent ex-ante, only 2^{N-t+1} type of agents are interdependent in trading round t. Equivalently, the network among 2^N agents is divided into different submarkets in period t, where each sub network has 2^{N-t+1} agents.

We now formally define the notion of connected agents in any period t. In period N, only two agents are connected, which is characterized by the bilateral link $j_N(i)$. In any

[11] 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 spanning from the current match. Nevertheless, the actual network does not need to be a tree. For example, according to Figure 1, the network graph contains loops. This is because the network itself can be static, in which case we are just solving a static network formation problem using a sequential method.

period t, an agent i is directly connected to counterparty $j_t(i)$,and thus is also indirectly to *future* counterparties of agent $j_t(i)$. Denote a set of agents I and their counterparties in period t $J_t(I)$, where $J_t(I) = \cup_{i\in I}\{i, j_t(i)\}$ is a list of period t link of agents in set I.

Definition 2. An agent i is connected to an agent j in period t or later iff

$$j \in \Psi_t(i) \equiv J_N(J_{N-1}(\ldots(J_{t+1}(J_t(i)))\ldots)).$$

$\Psi_t(i)$ denotes the set of agents who are connected to agent i in period t or later trading rounds. The set of agents that agent i is connected to from period t onwards can be understood as a tree with its root at the current match $J_t(i) = \{i, j_t(i)\}$.[12] By definition, an agent i is connected to at most 2^{N-t+1} agents from trading round t onwards.

3.3.4 Expected Transfers and Prices

Given our model predicts that risk can be concentrated on a few agents, we now turn to study the price of risks. Since holding risks is costly, an agent within the pair that holds more post-trade variance needs to be compensated by receiving transfers from her counterparty. The expected transfer makes any two banks matched under PAM receive the same expected payoff, once we know the asset allocation between the match. Hence, the expected transfer within a pair of banks with variance v matched in period t, denoted by $x_t(v)$, solves

$$-\kappa_t\tilde{v}_l(v) + W_{t+1}(\tilde{v}_l(v)) - x_t(v) = -\kappa_t\tilde{v}_h(v) + W_{t+1}(\tilde{v}_h(v)) + x_t(v), \tag{13}$$

where RHS (LHS) represents the payoff of the agent with lower (higher) post-trade variance $\tilde{v}_l(v)$ $(\tilde{v}_h(v))$.

Note that the equilibrium transfer $x_t(v)$ within the pair can be implemented as a constant bid-ask spread times the expected trading volume. The bank that holds higher post-trade variance within the pair commits to bid and ask prices regardless of their realized asset positions, denoted $P_t^A(v)$ and $P_t^B(v)$ respectively. The bid-ask spread

[12]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 spanned from the current match. Nevertheless, the actual network does not need to be a tree. For example, according to Figure 1, the network graph contains loops.

$S_t(v) \equiv P_t^A(v) - P_t^B(v)$ then solves

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

where $\vartheta_t(v) \equiv \mathbb{E}|\alpha^*(v)(a_{i,t} + a_{j,t}) - a_{i,t}|$ represents the expected volume between the pair of banks i and j.

Note that equation (13) immediately implies that the higher the degree of concentration, the higher the expected transfer. Hence, whenever the risk allocation changes discontinuously with the underlying parameters, the bid and ask prices and the bid-ask spread also change discontinuously.

3.3.5 Remarks on the Tractability of Our Framework

Our sequential approach admits a tractable numerical algorithm for any $W_{N+1}(v)$. This is primarily because the optimal solution is distribution-free under PAM. Since it is optimal for an agent to match with a counterparty of the same type, her matching and trading strategies holds for any distribution of risk exposures across agents.

The numerical solution can be found through the following procedure: First, solve backward value functions, $W_t(v)$, and policy functions for risk allocation, $\alpha_t(v)$, from the last period of bilateral trade, period N to the first period. $W_t(v)$ and $\alpha_t(v)$ for risk exposure v is the solution to a one-dimensional optimization problem specified in Eq (10) which uses the equilibrium property that matching is positive assortative. When the next-period continuation value $W_{t+1}(v)$ in Eq (10) is strictly convex, optimal posttrade risk exposures of matched agents may be different. We solve the expected transfers within a match so that matched agents are indifferent between posttrade risk exposures. Second, given the policy functions and positive assortative matching in equilibrium, solve for the evolution of risk exposure for any agent i, v_t^i from the first period to period N. Because the numerical algorithm involves only one dimensional optimization, it is easy to solve even if the objective function is convex in risk exposure.

Even when PAM does not hold in a more general environment, our framework gains tractability by solving the network formation game recursively. Our framework breaks a network formation game into a sequence of bilateral matchings as long as matching and trading in previous trading rounds affects the continuation value of an agent through her risk exposure. Bilateral matching (and trading) in each round depends on agents' current risk exposure and affects further network formation through the risk exposure distribution in the following period.

Although we solve a sequential network formation problem, our methodology can be applied to solve a static network formation problem by breaking it down into a sequence of bilateral linkage formation games.

An important simplifying assumption that makes our framework tractable is that we consider a large game with a continuum of agents. As we argued when defining the equilibrium,

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 2).[13] 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.

4.1 Structural Shift: from Full Risk-Sharing to Risk-Concentration to One Core

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 move the interbank network from the standard risk sharing to risk concentration, generating a discontinuously large increase in aggregate risks. In this sense, our model predicts that a small increase in risk-taking incentives can trigger a financial crisis through the network.

[13] 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.

Simple Case with Minimum Core Size We focus on the case when asymmetric allocation arises, it is optimal to concentrate on at most *one* of the 2^N connected agents in this application. We refer to such an agent as the core agent, since his final risk position and trading volume will be higher than others. In the next application, we further consider the possibility of multiple core agents.

Assumption 3. *(1) The terminal payoff function* $W_{N+1}(v)$ is $k-$times continuously differentiable with bounded derivatives, has a marginal cost of bearing risk converging to 0 at infinite variance, $\lim_{v\to\infty} W'_{N+1}(v) \to 0$, and is strictly convex, $W''_{N+1}(v) > 0$; (2) Function $\chi(v) \equiv \frac{1}{2}W'_{N+1}(v) + W''_{N+1}(v)v$ is concave for all variance $v \in [0, 2^N v_1]$.

The first part of A3 guarantees that the solution is interior and the convexity is bounded. The benefit of concentration is derived from the convexity $W''_{N+1}(v)$ and the cost of holding risks is captured by $-W'_{N+1}(v)$. $\chi(v) \equiv \frac{1}{2}W'_{N+1}(v) + W''_{N+1}(v)v$ can be interpreted as the net benefit of taking risks. Observe that $\chi(v) < 0$ when v is small so the benefit of concentration is relatively small.

We focus on the case when $\chi(v)$ is concave, which is a sufficient condition to guarantee that (1) the objective function of Equation (10) for period N is single-peaked, and (2) the corresponding solution $\alpha_N(v)$ is continuous. In other words, discontinuous change only arises when there is more than one round of bilateral trade. This further highlights the effect of interconnectedness.

Lemma 4. *(Minimum Core Size) Under Assumption A3 and the assumption that the flow utility in all trading rounds is zero,* $\kappa_t = 0$ *for all* $t \leq N$, *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* $v^c_{N+1} \geq v^0_{N+1}$. *The optimal risk-allocation solves*

$$\Pi(v_1) = \frac{1}{2^N} \max_{\alpha \geq \frac{1}{2^N}} \left\{ W_{N+1}\left(\alpha^2\left(2^N v_1\right)\right) + \left(2^N - 1\right) W_{N+1}\left(\left(\frac{1-\alpha}{2^N-1}\right)^2 \left(2^N v_1\right)\right)\right\}. \quad (14)$$

Thanks to Lemma 4, the allocation problem among 2^N agents can be greatly reduced to a one-dimensional problem. The aggregate payoff can be understood as 2^N agents sharing a total risk of $V = 2^N v_1$, where one "core" agent may hold more risks than the rest of $2^N - 1$ agents who bear even risks. Observe that $\alpha = \frac{1}{2^N}$ represents the case with full risk-sharing among all agents; and thus post-trade position $v^i_{N+1} = \left(\frac{1}{2^N}\right)^2 V = \frac{v_1}{2^N}$ $\forall i$. Under assumption A3 and the assumption that the cost of bearing risk before the terminal period is zero, $\kappa_t = 0$, there exists a cutoff v^* such that the equilibrium features full risk-sharing when banks' initial risk exposure is below the cutoff, $v_1 \leq v^*$, and features risk-concentration to one core bank among connected banks when their initial risk exposure is above the cutoff, $v_1 \geq v^*$. When there are multiple rounds of trade, $N > 1$, the aggregate posttrade risk exposure, $\int v_{i,N+1} di$, and thus risk premium in the asset price increases discontinuously at a threshold initial risk exposure v^*.

The proposition highlights that for small initial risk exposure $v_1 \leq v^*$, it is optimal for banks to use their network to share risks, consequently less aggregate risk exposure. For higher initial risk exposure $v_1 \geq v^*$, it is privately optimal for banks to shift to a concentrated structure, where $\frac{1}{2^N}$ fraction of banks (i.e., the only core agent among 2^N interconnected banks) bear disproportionately large risks, resulting in greater aggregate risks. Importantly, when $N > 1$, the solution to Equation (14) exhibits discontinuous jump.

Figure 3 illustrates the result of regime shift using the specification in Equation 8 with normal distribution. 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 on the core, which results in higher aggregate probability of default (which is proportional to the total variance).

4.2 Distribution and Flow of Risks

When risk concentration arises, the core agent collects risk from others. While the final allocation can be understood from a static model, our sequential setting further gives predictions regarding how the asset flows through the bilateral network. Intuitively, since the marginal cost of bearing risk is lower for the core agent, an agent i connected to the core in period t will then also have a lower marginal cost of bearing risk and thus can take on more risks from her counterparties. We now show that the risk-bearing capacity of an agent can be conveniently summarized by his core access.

Definition 3. (Core Access) Let the number of core access $c_{i,N+1} = 1$ iff agent i is the

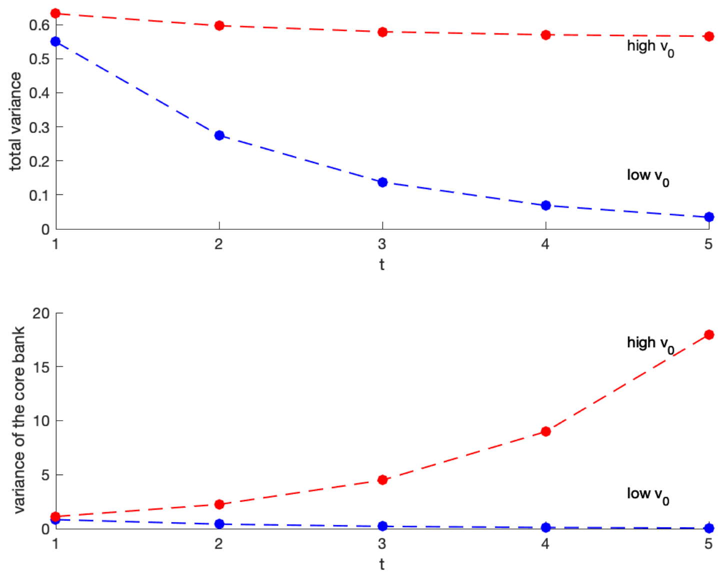

Figure 3: Regime Shift: $W_{N+1}(v) = -1 + e^{-cv}$, $c = 1.0196$, $v_{0L} = 1.02$ and $v_{0H} = 1.03$.

core agent and $c_{i,N+1} = 0$ otherwise. The core access of an agent from period t onwards is given by

$$c_{i,t} \equiv \Sigma_{k \in \Psi_t(i)} c_{k,N+1}. \tag{15}$$

In general, the core agents can be defined as the ones whose final risk position $v_{i,N+1}$ is above a certain percentile. In this specific example, since there is at most 1 core agent out of 2^N interconnected agents ex ante. The core is thus the one in the top $\frac{1}{2^N}$ percentile. Since there can be at most one connected core agent, the core access is always binary in any period t: $c_{i,t} \in \{0, 1\}$. At most one agent within the pair (i, j) can maintain the core access in period $t+1$. That is, $c_{i,t+1} = 1$ if the cost agent i_c is in the set of agents connect to agent i from period t onwards, $i_c \in \Psi_{t+1}(i)$ and $c_{i,t+1} = 0$ otherwise. Moreover, in general, $c_{i,t}$ must be (weakly) decreasing over time because the set of agents connected to agent i from period $t + 1$ onwards, $\Psi_{t+1}(i)$, is a subset of the set of agents connected to agent i from period t onwards, $\Psi_t(i)$.

The corollary below shows that the risk position of an agent $\{v_{i,t}\}$ can then be characterized as when she loses core access. Intuitively, an agent holds more risks from her counterparty until she loses her core access.

Corollary 2. *The dynamic path of core access for any agent i can be characterized by core access $\{c_{i,t}\}$ for all trading rounds $t = 1, \ldots, N$. For an agent whose $c_{i,t} = 1$ $\forall t \leq \tau$ and*

$c_{i,t} = 0\ \forall t \geq \tau+1$, *she collects risks from her counterparties for* $\tau-1$ *periods, unloads her risks to her counterparty in period* τ, *and shares risks with her counterparties afterwards. Given the final risk exposure of the core agents and other agents* $\left\{v^c_{N+1}, v^0_{N+1}\right\}$, *the risk-bearing capacity of an agent in period* t *can be expressed as a function of core access* $c_{i,t}$, *denoted by* $\gamma_t(c_{i,t})$,$\forall t \leq N$,

$$\begin{aligned}\gamma_t(1) &= \left(\left[W'_{N+1}(v^c_{N+1})\right]^{-1} + \left(2^{N-t+1}-1\right)\left[W'_{N+1}(v^0_{N+1})\right]^{-1}\right)^{-1} \\ &> \frac{1}{2^{N-t+1}} W'_{N+1}(v^0_{N+1}) = \gamma_t(0)\end{aligned}$$

This expression of the risk-capacity can be seen from Lemma 2. When the flow utility in all trading rounds is zero, the risk-capacity of an agent in period t can be further expressed as the harmonic mean of the marginal disutility of bearing risks in the terminal period, $W'_{N+1}(v_{k,N+1})$, of all her connected counterparties from period t onwards.

$$W'_t(v_{i,t}) = \left\{\Sigma_{k\in\Psi_t(i)}\left[W'_{N+1}(v_{k,N+1})\right]^{-1}\right\}^{-1}. \tag{16}$$

The core access is a key static in characterizing the risk-bearing capacity. For an agent not connected to the core in period t, the capacity is $W'_{N+1}(v^0_{N+1})$. For an agent connected to a core agent from period t onwards, one of terminal marginal cost, $W'_{N+1}(v_{k,N+1})$, is valued at the risk-capacity of the core agent, $W'_{N+1}(v^c_{N+1})$. Hence, an agent with a core access from period t onwards will have a lower marginal cost of bearing risks.

Agents that collect risks for more periods thus have greater expected trading volume. So agents with longer core access have greater expected volume. However, they are not riskier at the end, because they eventually unload their risks to the core agent. According to Lemma 4, they have the same final risk exposures and thus as “safe” as other banks.

When the flow cost of bearing risks is positive, $\kappa_t > 0$ for a period $t \leq N$, Lemma 3 shows that the noncore bank directly matched to the core in the final trading period in fact has the lowest posttrade risk exposure. In this sense, while these banks have been collecting risks over time and are “closest” to the core, they actually become least risky in the end.

4.3 Normative Implications

In this application, any risk taking is socially 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 We find that banks can take risks *collectively* by concentrating risks to a small set of banks through the interbank network. The collective moral hazard of risk taking is in the spirit of Farhi and Tirole (2012) but our theory points out a new channel of risk taking, through concentrating risk via interbank trade.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. In the literature on financial contagion, default propagates through banks' *gross* credit positions with each other. In our analysis, banks default over their *net* credit positions. Even so, default risk is interconnected through the endogenous interbank trading 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.[14] Consistent with our model, an asymmetric structure arises naturally when banks have options to invest in trading technologies to reduce their risk-bearing cost. We now apply our framework to study the positive and normative implications of reforms, taking into the equilibrium response of the market structure.

Assumption 4. *Piece-wise linear with binary action. The terminal payoff function solves the optimal decision of accessing a faster trading technology:*

$$W_{N+1}(v) = \max_{c_{N+1} \in \{0,1\}} \left\{ \gamma_{N+1}(c_{N+1})v - \varphi_{N+1}(c_{N+1}) \right\}, \tag{17}$$

where the marginal disutility of bearing risks when the bank does not access the faster technology is denoted $\gamma_{N+1}(0) = -\kappa_{N+1}$, the disutility when the bank accesses the faster technology is denoted $\gamma_{N+1}(1) = -\eta\kappa_{N+1}$, with $\eta \in [0, 1)$, and the fixed cost of accessing the technology is denoted $\varphi_{N+1}(1) = \phi > \varphi_{N+1}(0) = 0$.

Assumption A4 allows banks to choose a binary action, $c_{i,N+1} \in \{0, 1\}$ in period $N+1$, which naturally gives rise to a convex payoff function $W_{N+1}(v)$. That is, $c_{N+1} = 1$ represents that a bank invests a superior but more expensive trading technology. We refer agents that invest as core agents in this application. The cost of holding assets for non-core banks is given by $\gamma_{N+1}(0) = -\kappa_{N+1}$, and hence $\eta < 1$ captures the benefit of using the platform, which lowers the cost of holding risks. A fully competitive centralized market is a platform that allows fully risk-sharing with $\eta = 0$.

Remark 1. More generally, the usage cost can have variable components beyond the fixed cost. For example, consider the required collateral may be higher with larger positions conditional on entering the platform. This case could be nested by setting $\gamma_{N+1}(1) > 0$ and the same characterization can be applied.

[14] 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).

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 Structural Shift with Different Core Sizes

Our framework allows us to study how the market structure may respond to the underlying parameters of the economy which may be influenced by regulations. In this application, we are interested in the size of core and thus consider the general case with more than one core agent among those interconnected from bilateral trade. The market structure can thus be understood on two margins: First, the measure of banks that have access to the platform (i.e, the core size). Second, how other banks are connected and trade among each other.

Using Lemma 3, we first show that, given any core size, connections to the core is uniquely pinned down. Specifically, the optimal network must distribute the core access within the pair as even as possible next period, as it delays the needs for concentration. For example, if Agents i and j are matched in period t and are connected to two core agents from that period onwards, either of them maintains one core access from period $t+1$ onwards. That is, if $c_{i,t} = c_{j,t} = 2$, then $c_{i,t+1} = c_{j,t+1} = 1$.[15]

Lemma 5. *For any two agents i and j matched in period t, their posttrade core access is adjacent integers, $c_{i,t+1} = \lfloor \frac{c_{i,t}}{2} \rfloor$ and $c_{j,t+1} = \lceil \frac{c_{j,t}}{2} \rceil$. Under Assumption A4, the core access $c_{i,t}$ is the sufficient static for agent i's risk capacity in period t, where*

$$\gamma_t(c_{i,t}) = \frac{1}{2} H \left(-\kappa_t + \gamma_{t+1}(\lfloor \frac{c_{i,t}}{2} \rfloor), -\kappa_t + \gamma_{t+1}(\lceil \frac{c_{i,t}}{2} \rceil) \right) \; \forall t \leq N, \tag{18}$$

and $\gamma_t(c_{i,t})$ increase in $c_{i,t}$.

Assumption A4 implies that an agent's risk-capacity only depends on her access in period $N+1$ (piece-wise linear). Similar to Application 1, using Lemma 2, one can

[15] Recall that an agent's core access is defined in Equation 15, which must be (weakly) decreasing over time. The only difference is that $c_{i,t}$ can now be more than one, since an agent could be connected to more than one core agent.

show that the core access is again a sufficient static for agent $i's$ risk capacity. The only difference is that core access $c_{i,t}$ can be larger than one. As before, given $\gamma_t(c_{i,t})$, one can then pin down variance $\{v_{i,t}\}$ for all agents: since $\gamma_t(c_{i,t})$ increases in $c_{i,t}$, agents who have more core access posttrade bear more risks.

Optimal Core Size We have established that there is a unique optimal market structure given any core size c, which equals the initial core access for all agents connected from period 1 onwards, $c_{i,1}$. The optimal network can then be further reduced to choosing the number of core agents in the beginning of the trading game among 2^N connected agents from period 1 onwards. The expected ex ante payoff of an agent solves

$$\Pi(v_1) = \max_c \left\{ \gamma_1(c) v_1 - \left(\frac{c}{2^N} \right) \phi \right\}. \tag{19}$$

Given any core size c, $\gamma_1(c)$ represents the risk-bearing capacity for all agents, taking into account future connections according to Equation 18. If there are c core agents among 2^N agents, the total measure of core agents would be $\frac{c}{2^N}$; hence, the second term captures the total entry costs.[16]

The trade-off of core size can be seen from Equation 19: a larger core size results in higher total entry costs but lower risk-bearing costs. To explore how the core size depends on the underlying parameters, we further assume the parameter for the flow cost of bearing risks, $\kappa_t = \delta \kappa_{N+1}$ $\forall t \le N$, where the parameter δ represents the flow cost of bearing risks in a trading period relative to the terminal period.

Proposition 3. *Under Assumption A4 and positive flow marginal cost of bearing risks, the optimal measure of cores weakly decreases in the cost of accessing faster trading technology relative to the terminal marginal cost of bearing risks,* $\frac{\phi}{\kappa_{N+1} v_1}$. *When* $\delta = 0$ *and* $\eta = 0$, *the core size is* $\frac{1}{2^N}$ *if* $\frac{\phi}{\kappa_{N+1} v_1} \ge \bar{\phi}$ *and zero otherwise.*

We prove this result by showing that, given any (δ, η), the risk-capacity $\gamma_t(c)$ is a homogeneous function of degree 1 in κ_{N+1}. Recall that κ_{N+1} 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. Hence, the ratio, $\frac{\phi}{\kappa_{N+1} v_1}$, captures the entry cost *relative*

[16] 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 $1/2^N$. Then, there are $1/2^N$ identical replica of the finite network of size 2^N.

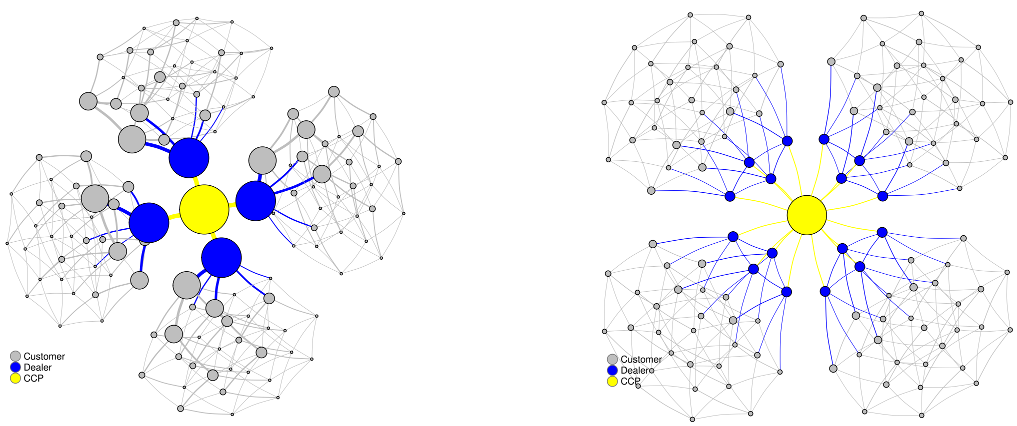

Figure 4: Pre vs. Post-regulation Market Structure.
Each panel shows the graph of the equilibrium trading network. In the network graph, each node represents a bank. The area of the node represents the gross trading volume involving the bank. The edges between nodes represent bilateral trading relationships.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

to level of risks. The higher the ratio means a relatively higher costs of using the platform, and thus the lower the optimal core size. The special case where $\delta = 0$ and $\eta = 0$ represents that both intraday holding costs and the core agent's marginal cost of holding risks are zero. Hence, similar to Application 1, there is no need to have more than one core agent.

The Effect of Reforms We think of 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). Hence, the equilibrium response is then characterized by Proposition 3 with a 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.[17] 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.

5.2 Normative Implications

Concentration Can be Efficient Our results highlight that the optimal intervention should not be targeting all-to-all trading or reducing risk concentration because the existence of exclusive core members and a high concentration of risks and volume *can* be efficient, 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 $\frac{c_0}{2^N}$ have built relationships among themselves and collectively operate the trading platform at cost ϕ. The incumbent agents jointly own the platform and decide whether to charge a new entrant to the platform an exogenous fee $\Delta > 0$.

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

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

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

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

A.1 Diminishing Marginal Cost of Bearing Risks and Endogenous Search Intensity: An Example

Suppose that banks can pay a quadratic cost $-\frac{c}{2}\gamma^2$ to have access to the competitive market with probability γ and that they choose the search intensity, γ, conditional their realized asset holding. Denote

$$W_{N+1}(v) = \int \widehat{W}(a) d\pi_{N+1}(a)$$

where $\widehat{W}(a)$ is a bank's expected payoff conditional on pretrade asset holding being a.

$$\widehat{W}(a) = \max_{\gamma} -\frac{c}{2}\gamma^2 - (1-\gamma)a^2$$

So the optimal search intensity conditional on pretrade asset holding a is

$$\gamma(a) = c^{-1}a^2$$

and

$$\widehat{W}(a) = -\frac{c}{2}\gamma(a)^2 - (1-\gamma(a))a^2 = -a^2 + c^{-1}a^4.$$

If the asset holding follows a normal distribution with mean 0 and variance v, the Kurtosis of the distribution is $Ea^4 = 3v^2$.

$$W_{N+1}(v) = E_v\widehat{W}(a) = -v + 3c^{-1}v^2$$

The marginal risk-bearing cost is decreasing in v, $W'(v) = -1 + 6c^{-1}v$.

A.2 Efficiency, Uniqueness, and Variance Representation

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 in 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 in period t. The planner's value function in period t has the joint asset distribution across agents as its state variable and can be characterized as

$$\Pi_t(\pi_t) = \int E_1 u_t(\tilde{a}_{i,t}(a_{i,t}, a_{j_{i,t},t}))di + \beta\Pi_{t+1}(\pi_{t+1}), \text{ for } t \leq N,$$

$$\Pi_{N+1}(\pi_{N+1}) = \int E_1 u_{N+1}(a_{i,N+1})di.$$

The constraints that the planner faces include:

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

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

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

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

Proposition 1 holds because the equilibrium value of a bank i in period t equals the shadow value of adding bank i to the planner's optimization problem in period t. For a more detailed proof, see for example Chang and Zhang (2022).

Under risk preferences specified in Section 3.1, 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.

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

Proof. Assumption (3) 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

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

where

$$\Pi_{N+1}(\boldsymbol{m}_{N+1}, \boldsymbol{v}_{N+1}, \boldsymbol{\rho}_{N+1}) = \int W_{N+1}(v_{i,N+1}) di. \tag{A.4}$$

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

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

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

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

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

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

for all $t \leq N$, where $\lambda^m_{i,j_{i,t},t}$ refers to the Lagrangian multiplier for constraint (A.5) for agent i and her counterparty $j_{i,t}$, $\lambda^v_{i,j_{i,t},t}$ refers to the Lagrangian multiplier for constraint (A.6).

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

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

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

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

Lemma 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.3 Network Properties

A.3.1 Proposition 2

For result (1): From Equation 10, let

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

We thus have

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

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

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

Hence, if $W_{N+1}(V)$ is concave in V, $\alpha = \frac{1}{2}$, which satisfies the FOC, is the global maximizer. Thus

$$\Omega_N(v_i + v_j) = -\kappa_N \left(\frac{v_i + v_j}{2}\right) + W_{N+1}(\frac{v_i}{2}) + W_{N+1}(\frac{v_j}{2}),$$

Given that $W_N(v_i) = \max_j \Omega_N(v_i + v_j) - W_N(v_j)$, we thus have

$$W_N'(v_i) = -\kappa_N + \frac{1}{2} W_{N+1}'(\frac{v_i}{2})$$

and hence, $W_N''(v_i) < 0$ if $W_{N+1}''(\frac{v_i}{2}) < 0$. By backward induction, we have $W_t''(v) < 0\ \forall t, v$ and thus risk-sharing is always the optimal solution. We thus have $v_{i,t} = \frac{1}{2} v_{i,t-1} = \left(\frac{1}{2}\right)^t v_0\ \forall i$. Since all agents are symmetric over time, it is WLOG to assume random matching.

For result (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

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

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.3.2 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

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

, 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.3.3 Proof For Lemma 3

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 in period $t+1$ only affects changes the flow payoff in 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.3.4 Proof for Lemma 4

Proof. Let $F(\alpha) \equiv \frac{1}{2^N}\left\{W_{N+1}\left(\alpha^2\left(V\right)\right) + \left(2^N - 1\right) W_{N+1}\left(\left(\frac{1-\alpha}{2^N-1}\right)^2 V\right)\right\}$,the FOC thus yields

$$\begin{aligned} F^{(1)}(\alpha) =& 2\sqrt{V}\left\{W'\left(\left(\alpha^2 V\right)\right)\sqrt{\alpha^2 V} - W'\left(\left(\frac{1-\alpha}{2^N-1}\right)^2 V\right)\sqrt{\left(\frac{1-\alpha}{2^N-1}\right)^2 V}\right\}. \\ =& 2V\left\{W'\left(v_h(\alpha)\right)\alpha - W'\left(v_l(\alpha)\right)\left(\frac{1-\alpha}{2^N-1}\right)\right\}. \end{aligned} \tag{A.8}$$

Note that, $F^{(1)}(1) = W'(V) < 0$ for any finite V, which means that the solution can't be at the boundary $\alpha = 1$. Moreover let $g(v) \equiv \left\{\left(\frac{dW'(v)\sqrt{v}}{dv}\right)\sqrt{v}\right\} = \frac{1}{2}W'(v) + W''(v)v$, SFOC can be rewritten as

$$F^{(2)}(\alpha) = 4V\left(g(\alpha^2 V) + g\left(\left(\frac{1-\alpha}{2^N-1}\right)^2 V\right)\frac{1}{2^N-1}\right). \tag{A.9}$$

Since $g(v)$ is concave, $F^{(2)}(\alpha)$ is concave. Let $g(\bar{v})$ such that $g'(\bar{v}) = 0$, we thus have $g'(v) > 0$ iff $v < \bar{v}$. Moreover, $g(0) = \frac{1}{2}W'(0) < 0$ and $\lim_{v\to\infty} g(v) > 0$, we have $g(\bar{v}) > 0$. There thus exists $\hat{v} < \bar{v}$ such that $g(\hat{v}) = 0$, and $g(v) < 0$ iff $v < \hat{v}$. Lastly, since $\frac{dW'(v)\sqrt{v}}{dv} > 0$ iff $v > \hat{v}$, $W'(v)\sqrt{v}$ is a unimodal function with the minimum at $\hat{v}$. Hence, for any asymmetric root that satisfies FOC , it must be the case that $v_l(\alpha) < \hat{v} < v_h(\alpha)$. We now use the next two lemma to establish that there can be at most one core.

□

Lemma 8. *For period N, $\alpha^*_N(v)$ continuously increases in v*

Proof. Observe that

$$F^{(3)}\left(\frac{1}{2^N}\right) = \left\{g'\left(\frac{1}{2^N}V\right)\sqrt{\frac{1}{2^N}V}\right\}\left(1 - \left(\frac{1}{2^N-1}\right)^2\right), \tag{A.10}$$

hence when $N=1$, $F^{(3)}(\frac{1}{2})=0$. As $F^{(2)}(\alpha)$ is concave, we thus have $F^{(2)}(\frac{1}{2})$ is the maximum of $F^{(2)}(\alpha)$, and $F^{(2)}(\alpha)<0$ and $F^{(3)}(\alpha)<0$, $\forall\alpha\geq\frac{1}{2}$. This means that (1) if $F^{(2)}(\frac{1}{2})=g(\frac{V}{4})<0$, then $\frac{1}{2}$ is the unique local maximum. To see this, suppose that there are two local maxima, then there must exist a local minimum α_{min} where $F^{(2)}(\alpha_{min})>0$, which contradicts that $F^{(2)}(\alpha)<0$ $\forall\alpha\geq\frac{1}{2}$. (2) if $F^{(2)}(\frac{1}{2})=g(\frac{V}{4})>0$, then $\frac{1}{2}$ is the local minimum. Moreover, then can be at most one local maximum. Suppose that there are two maximum (α_1,α_2), then there must exists $\alpha_{min}\in(\alpha_1,\alpha_2)$,where $F^{(2)}(\alpha_{min})>0$, which again contradicts that $F^{(2)}(\alpha)<0$ $\forall\alpha\geq\frac{1}{2}$.

By implicit theorem, for any $\alpha_N^*(V)\geq\frac{1}{2}$,since $g(v)=\frac{1}{2}W'(v)+W''(v)v$, and we have

$$\begin{aligned}
\frac{d\alpha_N^*(V)}{dV} &= -\frac{F_{\alpha V}(\alpha,V)}{F_{\alpha\alpha}(\alpha,V)}\,|_{\alpha=\alpha^*}\propto 2\left\{W''_{N+1}(\alpha^2V)\alpha^3V-W''_{N+1}((1-\alpha)^2V)(1-\alpha)^3V\right\}\\
&=2\left\{g(v_h(\alpha)-\frac{1}{2}W'(v_h(\alpha))\right\}\alpha-\left\{g(v_l(\alpha)-\frac{1}{2}W'(v_l(\alpha))\right\}(1-\alpha)\\
&=2g(v_h(\alpha))\alpha-g(v_l(\alpha))(1-\alpha)\geq 0.
\end{aligned}$$

The first equality uses the fact that $g(v)=\frac{1}{2}W'(v)+W''(v)v$, and the second uses $W'(v_h(\alpha))\sqrt{\alpha}=W'(v_h(\alpha))\sqrt{(1-\alpha)}$ at α^*. The last inequality uses the fact that, for any $\alpha_N^*(V)>\frac{1}{2}$, it must be the case that $v_l(\alpha)<\hat{v}<v_h(\alpha)$; hence, $g(v_l(\alpha))<0<g(v_h(\alpha))$.

□

Lemma 9. *Under A3 and* $\kappa_t=0$, *(1) there can be at most two different values of* v_{N+1};*and (2) if* $\max\{v_{N+1}^k\}>\min\{v_{N+1}^k\}$, *there can be at most one core agent when* $\kappa_t=0$.

Proof. First of all, from Equation A.8, $F^{(1)}(1)=W'(V)<0$, 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, where $\{v_{k,N+1}\}$ maximizes

$$\max\Sigma_{k=1}^{2^N}W_{N+1}(v_{k,N+1}) \tag{A.11}$$

$$\left[\Sigma_k\sqrt{v_{k,N+1}}\right]^2=2^Nv_1. \tag{A.12}$$

Hence,

$$\sqrt{v_{k,N+1}}\left(W'_{N+1}(v_{k,N+1})\right)=\lambda\sqrt{2^Nv_1}, \tag{A.13}$$

where λ is Lagrange multiplier of the constraint A.12. Since $W'(v)\sqrt{v}$ is a unimodal function with the minimum at $\hat{v}$, Hence, there can be at most two roots for Equation A.13.

For Result (2): let $v_{N+1}^{c} = \max\{v_{N+1}^{k}\}$ and $v_{N+1}^{0} = \min\{v_{N+1}^{k}\}$. This statement holds automatically when $N = 1$. We now show this holds when $N \geq 2$. Suppose that there are more than one agent with v_{N+1}^{c}. Given that the outcome can be achieved under any ordering of matching, first consider the case that the core is matched with a non-core agent in period N, which means that $v_N^1 = \frac{(\sqrt{v_{N+1}^{c}}+\sqrt{v_{N+1}^{0}})^2}{2}$ and it must be the case that $\alpha_N^*(v_N^1) > \frac{1}{2}$. The same outcome, however, can be achieved by have two core agents meet in period N,which implies their $v_N^2 = \frac{(\sqrt{v_{N+1}^{c}}+\sqrt{v_{N+1}^{c}})^2}{2}$ and they adopt risk-sharing, where $\alpha_N^*(v_N^2) = \frac{1}{2}$. Since $v_N^1 < v_N^2$, the fact that $\alpha_N^*(v_N^2) = \frac{1}{2}$ but $\alpha_N^*(v_N^1) > \frac{1}{2}$ violates the fact that that $\alpha_N^*(v_N)$ increases in v_N. Contradictions.

Lastly, given that there can be at most one core agent and the problem is identical to a static allocation, Equation 14 thus follows from Equation A.11, using the constraint that $\left[\sqrt{\alpha^2 V} + \left(2^N - 1\right)\sqrt{v_l(\alpha)}\right]^2 = 2^N v_1$, we thus have $v_l(\alpha) = \left(\frac{\sqrt{V}-\sqrt{\alpha^2 V}}{(2^N-1)}\right)^2$.

□

A.3.5 Proof for Proposition 4.1

Proof. Step 1: We first show that, for any $N > 1$, $F(\alpha)$ has at least two local maxima (α_e^*, α_c^*) from some mid-range of $V \in (V_\ell, V_h)$, where $\alpha_e^* = \frac{v_1}{2^N}$ and $\alpha_c^* \in (\frac{v_1}{2^N}, 1)$. To do so, we show that $F(\alpha)$ is convex in some region (α_1, α_2), where $\frac{1}{2^N} < \alpha_1 < \alpha_2 < 1$.

First of all, in order to guarantee that full risk-sharing is a local maximum, we need

$$F''(\frac{1}{2^N}) = 4V\left\{g(\frac{V}{(2^N)^2})\frac{2^N}{2^N-1}\right\} < 0,$$

hence, this condition holds whenever $\left(\frac{1}{2^N}\right)^2 V < \hat{v}$. Hence, we set $V_h = 2^N v_1 = \left(2^N\right)^2 \hat{v}$. Moreover, from Equation A.10,

$$F^{(3)}(\frac{1}{2^N}) = \left\{g'(\left(\frac{1}{2^N}\right)^2 V)\sqrt{\frac{1}{2^N}V}\right\}\left(1 - \left(\frac{1}{2^N-1}\right)^2\right) > 0,$$

as $g'(v) > 0$ for $v < \hat{v} < \bar{v}$. To show that $F^{(2)}(\alpha) > 0$ for some interior range of α, it is sufficient to show that the maximum value of $F''(\alpha)$ is large than zero. Let $\Gamma(V) = \max_\alpha F''(\alpha|V)$. One can show that $\Gamma(v)$ increases in v. To see this, let $\hat{\alpha}(V)$ be the solution above. For any $V' > V$, let $\tilde{\alpha} = \hat{\alpha}(V)\sqrt{\frac{V}{V'}}$, and thus

$$\begin{aligned}
\Gamma(V') &\geq g\left(\left(\frac{\hat{\alpha}^2(V)V}{V'}\right)V'\right) + g\left(\left(\frac{1-\left(\frac{\hat{\alpha}(V)V}{V'}\right)}{2^N-1}\right)^2 V'\right)\left(\frac{1}{2^N-1}\right) \\
&= g\left(\hat{\alpha}^2(V)V\right) + g\left(\left(\frac{\sqrt{V'}-\left(\hat{\alpha}(V)\sqrt{V}\right)}{2^N-1}\right)^2\right)\left(\frac{1}{2^N-1}\right) \\
&\geq g\left(\hat{\alpha}^2(V)V\right) + g\left(\left(\frac{\sqrt{V}-\left(\hat{\alpha}(V)\sqrt{V}\right)}{2^N-1}\right)^2\right)\left(\frac{1}{2^N-1}\right) = \Gamma(V),
\end{aligned}$$

where the last inequality uses the fact that $g'(v_l(\alpha)) \geq 0$. Let $\bar{V}$ such that $\Gamma(\bar{V}) = 0$, we thus have $\Gamma(V) > 0$, for $V > \bar{V}$, and by continuity, there exists a region where $F''(\alpha|V) > 0$. Moreover, since

$$\Gamma(V_h + \epsilon) > g(\hat{v} + \frac{\epsilon}{2^N})\frac{2^N}{2^N-1} > 0,$$

it must be the case that to $\bar{V} < 2^N\hat{v}$. Lastly, we need to have

$$F''(1|V) = g(V) + g\left(0\right)\frac{1}{2^N-1} < 0, \tag{A.14}$$

so that $F''(\alpha|V)$ is concave when α is large enough. Note that since $F'(1|V) < 0$, together with $F''(1|V) < 0$, it then guarantees the existence of another local maximum $\alpha_c^* < 1$. Condition A.14 is possible as $2^N v_1 > \bar{v}$, $g'(v) < 0$ for $v < \bar{v}$ and $g(0) < 0$, this condition is thus guarantees when v_1 is large enough. Let $\tilde{V} > \bar{v}$ such that $F''(1|\tilde{V}) = 0$, we thus have $F''(1|V) < 0$ for $V > \tilde{V}$. Hence, set $V_\ell = \max\{\bar{V}, \tilde{V}\}$, there exists a region where $F''(\alpha|V) > 0$ when $V \in (V_\ell, V_h)$.

Step 2: We now show that the exists $V^* \in (V_\ell, V_h)$ such that α_e^* is the global optimal iff $V < V^*$.

$$D(V,N) \equiv \max_{\alpha > \frac{1}{2^N}}\left\{W_{N+1}\left(\alpha^2 V\right) + \left(2^N-1\right)W_{N+1}\left(\left(\frac{1-\alpha}{2^N-1}\right)^2 V\right)\right\} - 2^N W_{N+1}\left(\left(\frac{1}{2^N}\right)^2 V\right)$$

$$
\begin{aligned}
\frac{\partial D(V,N)}{\partial V} &= \left\{ \left\{ W'_{N+1}\left(v_h(\alpha)\right) \frac{v_h(\alpha)}{V} + \left(2^N-1\right) W'_{N+1}\left(v_l(\alpha)\right) \frac{v_l(\alpha)}{V} \right\} - 2^N W'(\left(\frac{1}{2^N}\right)^2 V) \left(\frac{\left(\frac{1}{2^N}\right)^2 V}{V} \right) \right\} \\
&= W'_{N+1}\left(v_l(\alpha)\right) \sqrt{v_l(\alpha)} \frac{1}{V} \left\{ \sqrt{v_h(\alpha)} + \left(2^N-1\right) \sqrt{v_l(\alpha)} \right\} - W'(\left(\frac{1}{2^N}\right)^2 V) \frac{1}{2^N} \\
&= W'_{N+1}\left(v_l(\alpha)\right) \sqrt{v_l(\alpha)} - W'(\left(\frac{1}{2^N}\right)^2 V) \\
&= \frac{1}{\sqrt{V}} \left\{ W'_{N+1}\left(v_l(\alpha)\right) \sqrt{v_l(\alpha)} - W'(\left(\frac{1}{2^N}\right)^2 V) \sqrt{\frac{V}{\left(2^N\right)^2}} \right\} > 0
\end{aligned}
$$

where the first equality uses FOC and thus $W'\left(\left(\alpha^2 V\right)\right)\sqrt{\alpha^2 V} - W'\left(\left(\frac{1-\alpha}{2^N-1}\right)^2 V\right)\sqrt{\left(\frac{1-\alpha}{2^N-1}\right)^2 V} = 0$, the second equality uses the variance constraint $\sqrt{v_h(\alpha)} + \left(2^N-1\right)\sqrt{v_l(\alpha)} = \sqrt{V}$, and the last inequality uses the fact that $\frac{d\left(W'(v)\sqrt{v}\right)}{dv} < 0$ for $v < \hat{v}$.

□

A.3.6 Proof for Corollary 2

We first show that, according to Lemma 3, when $\kappa_{t=0}$, we have $W'_t(v_{i,t}) = \frac{1}{2^{N-t+1}} \left\{ \frac{2^{N-t+1}}{\Sigma_{k\in\Psi_t(i)}\left(\frac{1}{W'_{N+1}(v_{k,N+1})}\right)} \right\}$. This holds for period N. Assume that $W'_{t+1}(v_{i,t+1}) = \left\{ \frac{1}{\Sigma_{k\in\Psi_t(i)}\left(\frac{1}{W'_{N+1}(v_{k,N+1})}\right)} \right\}$,by backward induction, we thus have

$$
\begin{aligned}
W'_t(v_{i,t}) &= \frac{1}{2}\left\{ \frac{2}{\Sigma \frac{1}{W'_{t+1}(v_{i,t+1})}} \right\} = \frac{1}{\Sigma_{k\in\Psi_{t+1}(i)}\left(\frac{1}{W'_{N+1}(v_{k,N+1})}\right) + \Sigma_{k\in\Psi_{t+1}(j)}\left(\frac{1}{W'_{N+1}(v_{k,N+1})}\right)} \\
&= \frac{1}{\left\{\Sigma_{k\in\Psi_t(i)}\left(\frac{1}{W'_{N+1}(v_{k,N+1})}\right)\right\}^{-1}}
\end{aligned}
$$

Hence, the value above only depends on whether $i_c \in \Psi_t(i)$. If so, $c_{i,t} = 1$, and thus have

$$
\gamma_t(1) = \frac{1}{\frac{1}{W'_{N+1}(v^c_{N+1})} + \frac{(2^{N-t+1}-1)}{W'_{N+1}(v^0_{N+1})}} > \frac{1}{\frac{1}{W'_{N+1}(v^0_{N+1})} + \frac{(2^{N-t+1}-1)}{W'_{N+1}(v^0_{N+1})}} = \frac{W'_{N+1}(v^0_{N+1})}{2^{N-t+1}} = \gamma_t(0),
$$

where the inequality uses the fact that $\frac{1}{W'_{N+1}(v^c_{N+1})} < \frac{1}{W'_{N+1}(v^0_{N+1})}$.

A.3.7 Proof for Lemma 5

Proof. Given the payoff in the final period $N+1$, $W_{N+1}(v) = \max_{c_{N+1}} \gamma_{N+1}(c_{N+1})v - \phi(c_{N+1})$, where $\gamma_{N+1}(c_{N+1})$ increase 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 $v_{i,N+1} > v_{j,N+1}$. Since $c_{i,N} = c_{i,N+1} + c_{j_t(i),N+1} \in \{0,1,2\}$,the value of $\gamma_N(c)$ is given by Equation 12, which inreases in c. Thus, $c^*_{N+1}(v)$ must increase in v.

For any period $t = N-1$, suppose that $c_{j,N} - c_{i,N} \geq 2$. which is only possible when $c_{j,N} = 2$ and $c_{i,N} = 0$, as $c_{k,N} \in \{0,1,2\}$. Since $\gamma_N(c)$ increases in c, Agent j must hold strictly higher post-trade varaince (i.e., $v_{j,N} > v_{i,N}$). Moreover, as $c_{k,N} \in \{0,1,2\}$, $c_{j,N} - c_{i,N} \geq 2$ is only possible when $c_{j,N} = 2$ and $c_{i,N} = 0$. This thus means that $c_{j,N+1} = c_{j_N(j),N+1} = 1$ and $c_{i,N+1} = c_{j_N(i),N+1} = 0$. Since $c^*_{N+1}(v)$ must increase in v, it thus implies that

$$\min\{v_{j,N+1}, v_{j_N(j),N+1}\} > \min\{v_{i,N+1}, v_{j_N(i),N+1}\},$$

which contracts Lemma 3. Hence, for any $c_{N-1} \in \{0,1,2,3,4\}$,the connections are unique, where $c_{i,N} = \left\{\lfloor \frac{c_{i,N-1}}{2} \rfloor, \lceil \frac{c_{i,N-1}}{2} \rceil\right\}$ 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 $v_{i,t} > v_{j,t}$. Hence, by similar logics, if $c_{j,t+1} - c_{i,t+1} \geq 2$, then

$$\min\left\{c_{j,t+2}, c_{j^*_{t+1}(j),t+2}\right\} > \min\left\{c_{i,t+2}, c_{j^*_{t+1}(i),t+2}\right\}$$

and thus

$$\min\{v_{j,t+2}, v_{j^*_{t+1}(j),t+2}\} > \min\{v_{i,t+2}, v_{j^*_{t+1}(i),t+2}\},$$

which violates Lemma 3. 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 in period t. This thus establishes that Lemma 5 must hold for any t.

□

A.3.8 Proof of Proposition 3

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 18, 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 (19) 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 $\frac{\phi}{\kappa v_1}$.

References

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

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 (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.

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

Farhi, E. and J. Tirole (2012). "Collective moral hazard, maturity mismatch, and systemic bailouts". *American Economic Review* 102.1, pp. 60–93.

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.*

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.

Üslü, S. (2019). "Pricing and liquidity in decentralized asset markets". *Econometrica* 87.6, pp. 2079–2140.

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.