Contents lists available at ScienceDirect

# ELSEVIER



journal homepage: www.elsevier.com/locate/jedc

# The formation of a core-periphery structure in heterogeneous financial networks<sup> $\star$ </sup>



Daan in 't Veld<sup>a,b</sup>, Marco van der Leij<sup>a,c</sup>, Cars Hommes<sup>a,c,d,\*</sup>

<sup>a</sup> CeNDEF, University of Amsterdam, Netherlands

<sup>b</sup> PBL Netherlands Environmental Assessment Agency, Netherlands

<sup>c</sup> Tinbergen Institute Amsterdam, Netherlands

<sup>d</sup> Bank of Canada, Ottawa, Canada

# ARTICLE INFO

Article history: Received 8 November 2019 Revised 29 June 2020 Accepted 6 August 2020 Available online 20 August 2020

JEL classification: D85 G21 L14

Keywords: Financial networks Core-periphery structure Network formation models Over-the-counter markets Interbank market

# 1. Introduction

ABSTRACT

Recent empirical evidence suggests that financial networks exhibit a core-periphery network structure. This paper aims at giving an explanation for the emergence of such a structure using network formation theory. We propose a general, stylized model of the interbank trading market, in which banks compete for intermediation benefits. Focusing on the role of bank heterogeneity, we find that a core-periphery network cannot be unilaterally stable when banks are homogeneous. A core-periphery network structure can form endogenously, however, if we allow for heterogeneity among banks in size. Moreover, size heterogeneity may arise endogenously if payoffs feed back into bank size.

> © 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

The extraordinary events of 2007 and 2008 in which the financial system almost experienced a global meltdown, has led to increased interest in the role of the financial network – the network of trading relationships and exposures between financial institutions<sup>1</sup> – in systemic risk: the risk that liquidity or solvency problems in one institution spread to the entire sector. Building on pre-crisis work by Allen and Gale (2000) and Eisenberg and Noe (2001), an extensive body of theoretical, simulation and empirical research has shown that the structure of the network of interbank liabilities matters for the likelihood and the extent of financial contagion.<sup>2</sup>

\* Corresponding author.

E-mail addresses: daan.intveld@pbl.nl (D. in 't Veld), m.j.vanderleij@uva.nl (M. van der Leij), c.h.hommes@uva.nl (C. Hommes).

<sup>1</sup> In this paper we use the words 'financial institutions' and 'banks' interchangeably.

https://doi.org/10.1016/j.jedc.2020.103972

<sup>\*</sup> The authors would like to thank Darrell Duffie, Ester Faia, Filomena Garcia, Michael Gofman, Matthew Jackson, Christian Julliard and Mariëlle Non for comments. We also thank participants at several workshops and conferences for their attendance and comments. The research has been supported by the Netherlands Organisation for Scientific Research (NWO) under the Complexity program Understanding financial instability through complex systems (grant number: 645.000.015).

<sup>&</sup>lt;sup>2</sup> Concerning the literature of financial contagion, we refer to Glasserman and Young (2016) for a comprehensive review, and Gai et al. (2011), Elliott et al. (2014) and Acemoglu et al. (2015) for seminal papers.

<sup>0165-1889/© 2020</sup> The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Importantly, however, until recently almost all of this work assumed that the network of financial interconnections is *exogenously fixed*. This assumption ignores the fact that financial networks do not come out of the blue. Financial relations are formed consciously by financial institutions who borrow, lend and trade financial assets with each other in order to maximize profits. This is important, as a change in the risk or regulatory environment may incentivize financial institutions to rearrange their financial links. This change may in itself constitute a financial crisis, for example in the case of an interbank market freeze, which may be interpreted as a sudden shift from a connected to an empty financial interbank network.

It is therefore important to better understand the formation process of financial networks. A natural starting point is to try to explain stylized facts about financial network structure. Regarding financial networks, one consistent empirical finding is that financial networks of interbank markets have a structure close to a *core-periphery structure*, which is defined as a connected network that has two tiers, a core and a periphery, the core forming a fully connected clique, whereas peripheral banks are only connected to the core (Borgatti and Everett, 1999). For example, Craig and Von Peter (2014) and In 't Veld and Van Lelyveld (2014) find such a structure in interbank markets in respectively Germany and the Netherlands. Moreover, In 't Veld and Van Lelyveld (2014) show that a core-periphery network structure fits the actual Dutch interbank network better than alternative network structures, such as a scale-free network or a nested split graph.<sup>3</sup>

This paper aims to contribute to our understanding of why such core-periphery networks are formed. This aim is similar to recent work by Castiglionesi and Navarro (2019), Chang and Zhang (2019), Craig and Ma (2020), Farboodi et al. (2020), and Wang (2018). We consider our contribution relative to this previous literature in more detail in Section 1.1.

We present a network formation model of a financial market with an explicit role for *intermediation*. The model builds on work of Goyal and Vega-Redondo (2007). They show that, starting from ex ante identical agents, the star network with a single intermediating counterparty, quickly arises in an environment in which relations are costly.<sup>4</sup> Intuitively, this result arises from network effects for intermediation; it becomes more attractive to link to an intermediator if the intermediator has already many links. However, in practice, rather than simple stars, we observe core-periphery networks that have *multiple banks* in the core. Moreover, core banks tend to form a fully connected *clique*. Such core-periphery networks are not stable in the framework of Goyal and Vega-Redondo (2007). An important reason for that is their assumption of perfect competition for intermediation benefits. This assumption drives payoffs to zero for any network in which there are two or more members in the core. On the other hand, *imperfect* competition for intermediation benefits allows for positive payoffs for competing core players, which may allow for the possibility of non-trivial stable core-periphery networks.

In order to see if core-periphery networks are stable in an environment of imperfect competition for intermediation benefits, we adapt the framework of Goyal and Vega-Redondo (2007). We propose a general yet stylized trading model, in which pairs of banks divide a trading surplus. We assume that trade can only take place between costly long-term trading relationships, allowing for the possibility of intermediation benefits. Unlike (Goyal and Vega-Redondo, 2007) competition between intermediators is imperfect, opening up the possibility that multiple core players benefit from intermediation. With the benefits from this trading network as a second stage, we then consider first-stage network formation of long-term trading relationships. Rather than the equilibrium concept of bilateral equilibrium (Goyal and Vega-Redondo, 2007), we focus in our framework on the concept of *unilateral stability* (Buskens and Van de Rijt, 2008). This concept allows for entry and exit of intermediators in the core.

We ask ourselves if the core-periphery network is stable in this model. To our surprise, in a homogeneous situation, the answer is generally no. We provide two results on that. First, Proposition 1 shows that a core-periphery network, in which the set of connections of one core player contains the set of another core player, is not pairwise (let alone unilaterally) stable. The intuition behind this result is as follows: a stable core-periphery network implies that periphery banks prefer to trade indirectly via intermediating core banks, rather than trade directly. However, given that periphery and core banks have identical technologies, the core bank with the larger set of connections should have an incentive to trade indirectly via peripheral banks as intermediators as well. Hence, the core player does not have an incentive to maintain all direct trading relationships in the core, in contradiction to the definition of a core-periphery network.

Second, Proposition 2 states that, when the periphery becomes very large compared to the core, a core-periphery network cannot be unilaterally stable. For large enough networks, the payoff inequality between core and periphery banks becomes unsustainably large, as intermediation benefits for core banks grow quadratically with the number of periphery banks. Periphery banks therefore have an incentive to enter the core, even if competition between intermediators reduces their benefits.

Key to these findings is that banks are ex ante identical; periphery banks have the same trading surplus and the same intermediation technology as core banks. This puts a limit on inequality, and therefore excludes the stability of a coreperiphery network. We therefore investigate the role of heterogeneity in our model. We analyze a version with two types of banks, big banks and small banks, and allow big banks to have higher trading surpluses. Proposition 3 shows that for sufficiently large differences between big and small banks, it becomes beneficial for large banks to have direct lending relationships with all other large banks in the core, such that the core-periphery network becomes a stable structure.

<sup>&</sup>lt;sup>3</sup> For theoretical network models generating nested split graphs, see Cohen-Cole et al. (2015) and König et al. (2014).

<sup>&</sup>lt;sup>4</sup> A star network is a network in which one and only one node, the center, is connected to all periphery nodes, and no other links exist. Formally, the star may be considered a trivial case of a core-periphery network. Our interest is in core-periphery networks that are not star networks.

Which network structures do arise, if they are not core-periphery networks? Theorem 1 is concerned with a bestresponse dynamic framework a la Kleinberg et al. (2008) to show which type of networks are likely to arise. This analysis confirms that homogeneous network dynamics never converge to a core-periphery structure. Interestingly, instead of core-periphery networks, we find that *multipartite* networks may be a stable outcome. These type of networks are two- or multi-tiered as well; however, unlike core-periphery networks, they do not have links within a (core) tier.

Finally, we show that a stable core-periphery network may arise *endogenously* with ex ante identical banks, if one allows for a feedback loop from inequality in payoffs to inequality in size. This process works as follows. Starting from identical banks, best-response dynamics may converge to an unequal multipartite network, such that one side earns more than the other side of the network. Due to the feedback from payoff to size, the banks on the side that earn more increase their scale, until it finally becomes attractive for the largest banks to trade directly, forming a core-periphery network structure.

Overall, the main message of our paper is that bank heterogeneity matters crucially for the formation of core-periphery networks. This suggests that, in order to understand the financial system and its (systemic) risks, policy makers would first need to understand the sources of this heterogeneity and its relation with the financial network structure.

This paper is organized as follows. Section 1.1 places our contribution in the literature. In Section 2 we introduce our model with the basic network structures, the pay-off function and the stability concepts. Our main results are presented separately in terms of stability (Section 3) and dynamics (Section 4). In Section 5 we provide an application of our model to the interbank market of the Netherlands. Section 6 concludes.

# 1.1. Relation to earlier literature

We now place our contribution in the literature on financial network formation.<sup>5</sup> We distinguish between two sets of results in the literature. A first set of results starts off from ex ante heterogeneity to generate core-periphery networks (e.g. Craig and Ma, 2020; Farboodi, 2017; Farboodi et al., 2019). A second set of results assumes ex ante homogeneity and reproduces certain stylized facts of the network structure (e.g. Babus and Hu, 2017; Castiglionesi and Navarro, 2019; Chang and Zhang, 2019; Farboodi et al., 2020). However, the proposed models in this set are unable to explain a core-periphery network as defined in the empirical literature on financial networks (Craig and Von Peter, 2014). Our paper is one of the first to generate a 'pure' core-periphery structure from ex ante homogeneity (cf. Wang, 2018).

In the first set of papers, various sources of heterogeneity have been proposed as a driving mechanism to create a core of intermediators in a trading network. In Farboodi (2017) banks are heterogeneous in their investment opportunities, and they compete for intermediation benefits. A core-periphery network is formed with investment banks forming the core, as they are able to offer better intermediation rates. In Farboodi et al. (2019) agents with superior bargaining skills form the core of the trading network. In Bedayo1 et al. (2016) intermediaries brokering between two traders bargain bilaterally for the intermediation benefits. Here agents are heterogeneous in their time discounting. They find that a core-periphery network is formed with impatient agents in the core. Craig and Ma (2020) consider Nash bargaining between lending, borrowing and intermediary banks. They estimate German banks' preferences revealed by the observed interbank market, verifying the pairwise stability conditions for possible new links. Our paper (see Proposition 3) provides a different source of heterogeneity, namely trade surplus through bank size ( $\alpha$ ), as a possible driving force to separate between core and periphery.

Models in the second set of papers start off from ex ante homogeneity and reproduce some stylized facts of the network structure, but without leading to a core-periphery network. In Castiglionesi and Navarro (2019) heterogeneity in investments arises endogenously. Some banks invest in safe projects, and others in risky projects. Links are created as a coinsurance to liquidity shocks. Safe banks link freely with each other, but the incentives to link to risky banks is limited, leading to a structure close to, but different from a core-periphery structure. In particular, 'periphery' banks in the model of Castiglionesi and Navarro (2019) may form links with each other. This contradicts the definition of a core-periphery network of Borgatti and Everett (1999), Craig and Von Peter (2014) and In 't Veld and Van Lelyveld (2014), which we follow. In Chang and Zhang (2019) a financial system is formed in which the most stable banks have many trading relations and intermediate for other banks. The outcome is a multi-layered hierarchy network in which agents are ordered by trading activity. Again, this structure violates the definition of a two-layer core-periphery network that we apply. Babus and Hu (2017) consider a financial network formation model that builds on Goyal and Vega-Redondo (2007). In their model an interlinked star network with 2 members in the core may be stable. Similarly as in Goyal and Vega-Redondo (2007), core-periphery networks with 3 or more nodes in the core are not stable in the model of Babus and Hu (2017). In Farboodi et al. (2020) traders can choose their rate of contacting counterparties. They find that the emerging network is scale-free.

The paper of Wang (2018) is closest to ours in generating a core-periphery network from ex ante homogeneity. Wang (2018) looks for Bayesian equilibria under ex ante identical traders in which some of them act as dealers who manage an asset inventory and provide price quotes. He shows that core-periphery networks form one possible outcome of multiple equilibrium outcomes. Wang (2018) provides two (static) selection criteria that select a unique core-periphery structure.

<sup>&</sup>lt;sup>5</sup> Some papers in the social science literature also explain core-periphery networks. These network formation models are typically concerned with optimal effort levels to account for peer effects. Galeotti and Goyal (2010) and Hiller (2017) provide conditions under which core-periphery networks are the only stable network structure. See also Persitz (2016) who introduces heterogeneity in the connections model of Jackson and Wolinsky (1996). This literature cannot easily be translated to financial networks, because of the different interpretation of links as (channels for) financial transactions.



**Fig. 1.** A core-periphery network with n = 8 players, of which k = 3 are in the core.

Compared to this paper, our methodology is quite different and, arguably, more intuitive. We use best-response dynamics as an equilibrium selection tool. Also, we model linking costs (c) and competition between intermediators ( $\delta$ ) as the main parameters. Using only these ingredients, our model (see Theorem 1) places the core-periphery network within a general perspective between other (hypothetical) outcomes of the complete, star and empty networks.

# 2. Model

Our goal is to model the formation of a network of long-term trading relationships between banks. We denote the set of banks by N, and the number of banks by n. There are two stages. In the first stage, t = 0, banks form an undirected network, g, of these trading relationships. Denote by  $g_{ij} = g_{ji} = 1$  the existence of a trading relationship, and by  $g_{ij} = g_{ji} = 0$ the absence of a link. After forming their long-term trading relationships, trade takes place through these relationships at stage t = 1. Payoffs from forming trading relationships at stage t = 0 depend on the benefits from interbank trades along the network g and the costs of maintaining relationships.

# 2.1. Basic structures

Before discussing the payoff structure of the model, we first define the relevant network structures around which our analysis revolves. Denote the *empty network*,  $g^e$ , as the network without any links, i.e.  $\forall i, j \in N$ :  $g_{ii} = 0$ , and the *complete network*,  $g^c$ , as the network with all possible links, i.e.  $\forall i, j \in N$ :  $g_{ij} = 1$ . A star network,  $g^s$ , has a single player, the center of the star, that is connected to all other nodes, while no other links exist, i.e.  $\exists i$  such that  $\forall j \neq i : g_{ij} = 1$  and  $\forall j, k \neq i : g_{ik} = 0$ .

A core-periphery network is a network, in which the set of agents can be partitioned in a core and a periphery, such that all agents in the core are completely connected within and are linked to some periphery agents, and all agents in the periphery have at least one link to the core, but no links to other periphery agents.

**Definition 1.** A network g is a core-periphery network, if there exists a set of core agents  $K \subset N$  and periphery agents  $P = N \setminus K$ , such that:

- (a)  $\forall i, j \in K : g_{ij} = 1$ , and  $\forall i, j \in P : g_{ij} = 0$ ; (b)  $\forall i \in K : \exists j \in P$  with  $g_{ij} = 1$ , and  $\forall j \in P : \exists i \in K$  with  $g_{ij} = 1$ .

This definition follows Borgatti and Everett (1999), Craig and Von Peter (2014) and In 't Veld and Van Lelyveld (2014). See Fig. 1 for an example of a core-periphery network.

A special case of a core-periphery network is the *complete core-periphery network*, where each agent in the core K is linked to all agents in the periphery  $P: \forall i \in K$  and  $\forall j \in P$  it holds that  $g_{ij} = 1$ . See Fig. 2 for an example. We denote a complete core-periphery network with k = |K| agents in the core as  $g_{com}^{CP(k)}$ .<sup>6</sup> In Section 3.2, we use this special case to show the stability of core-periphery networks under heterogeneous traders.

Finally, a (complete) multipartite network is a network in which the agents can be partitioned into q groups, i.e.  $N = \{K_1, K_2, \dots, K_q\}$ , such that nodes do not have links within their group, but are connected to (all) nodes outside their own group. Formally, in a complete multipartite network it holds that  $\forall m \in \{1, 2, ..., q\}$ :  $\forall i \in K_m$  we have  $\forall j \in K_m$ :  $g_{ij} = 0$ and  $\forall j \notin K_m$ :  $g_{ij} = 1$ . All relevant multipartite networks in this paper are *complete* multipartite networks and therefore, for brevity, we will drop the word 'complete'. Multipartite networks will be denoted as  $g_{k_1,k_2,...,k_q}^{mp(q)}$ , where  $k_m \equiv |K_m|$  is the size of the *m*-th group. Multipartite networks are called *balanced* if the group sizes are as close as possible to each other, i.e.  $|k_m - k_{m'}| \le 1$  for all m, m'. Fig. 7 in Section 4.1 presents examples of multipartite networks that arise in our model.

<sup>&</sup>lt;sup>6</sup> Borgatti and Everett (1999) call this architecture a perfect core-periphery network. By Definition 1, empty, star and complete networks are special cases of core-periphery networks with cores of size k = 0, k = 1 and k = n respectively. A complete core-periphery networks with k = n - 1 is also identical to a complete network. In discussing our results we will make clear when we are speaking of non-trivial core-periphery networks with  $k \in \{2, 3, ..., n-2\}$ .



**Fig. 2.** A complete core-periphery network with n = 8 players, of which k = 3 are in the core.

#### 2.2. Trading benefits

We now turn to the payoffs from trading at stage t = 1. A network g results in a payoff  $\pi_i(g)$  for every bank *i*. We assume the following payoff function, with agent *i* in graph g benefitting from trades it is involved in, minus the cost of linking:<sup>7</sup>

$$\pi_i(g,\delta,c) = \sum_{j\in N_i^1(g)} \left(\frac{1}{2}\alpha_i\alpha_j - c\right) + \sum_{j\in N_i^2(g)} \alpha_i\alpha_j f_e(m_{ij}(g),\delta) + \sum_{k,l\in N_i^1(g)|g_{kl}=0} \alpha_k\alpha_l f_m(m_{kl}(g),\delta).$$
(1)

direct trade indirect trade intermediation benefits

In this equation,  $N_i^d(g)$  denotes the set of nodes at distance *d* from *i* in the network *g*; *c* denotes the costs of linking;  $\alpha_i \alpha_j$  the size of the surplus from trade between *i* and *j*;  $f_e$  and  $f_m$  the shares of trade surplus distributed to end nodes (i.e. borrowers and lenders) and to middlemen;  $m_{ij}$  the number of middlemen connecting *i* and *j*; and  $\delta$  the level of competition between middlemen. The payoff function can be partitioned in benefits from trading directly minus the cost of maintaining a link, benefits from trading indirectly, and benefits from intermediating between two banks, and will be discussed in detail below.

We start with two general comments on how to interpret our payoff function.

Concerning benefits, the interbank links in this model are undirected (although loans do have a direction) and are to be interpreted as established preferential trading relationships. We assume that trade can only be realized if the agents are linked either directly or indirectly through intermediation by mutual trading relationships. This is a strong, but not implausible assumption for interbank trade. The existence of preferential trading relationships has been shown by Cocco et al. (2009) in the Portuguese interbank market and by Bräuning and Fecht (2017) in the German interbank market. Afonso and Lagos (2015) document how some commercial banks act as intermediaries in the U.S. federal funds market. A bank that attempts to borrow outside its established trading relationships may signal that it is having difficulties to obtain funding and, as a consequence, may face higher borrowing costs. Hence, banks have incentives to use their established trading relationships.

Concerning costs, it is assumed that the preferential trading relationship comes at a fixed cost of c. This cost follows from maintaining mutual trust and from monitoring, i.e. assessing the other bank's risks. In principle, it is possible that these costs are not constant over banks, e.g. economies of scale may decrease linking costs in the number of relationships that are already present. The possible heterogeneity in linking costs is most likely smaller than the heterogeneity in trading surpluses, and we therefore assume that all banks pay an equal cost for a trading relationship. We assume that each link between i and j involves a cost c to both i and j.

We now discuss the interbank payoffs in more detail. Each pair of banks *ij* creates a potential trade surplus  $\alpha_i \alpha_j$ . So every trade surplus depends only on one parameter for *i* ( $\alpha_i$ ) and one for *j* ( $\alpha_j$ ). The network determines for which pairs *ij* trade can be realized, but does not influence the size of the trade surplus. Instead, the network *g* determines the distribution of the trade surplus.

To simplify the exposition, we assume that trade between *i* and *j* can only be realized if *i* and *j* have a direct (long-term) trading relationship,  $g_{ij} = 1$ , or an indirect trading relationship through one or more middlemen, who are directly connected to both *i* and *j*. Denote the set of these middlemen in *g* as  $M_{ij}(g) = \{k : g_{ik} = g_{jk} = 1\}$ , and the number of middlemen as  $m_{ij} \equiv m_{ij}(g) = |M_{ij}(g)|$ . For trade to be realized, we require that these middlemen are directly connected to *i* and *j*. In network parlor, it implies that trade between *i* and *j* can only be realized if the network distance, that is the shortest path length, between *i* and *j* in *g* is at most 2. However, we emphasize that our main result – core-periphery networks are generally unstable under homogeneity, but can be stable if agents are heterogeneous – does *not* depend on this assumption;

<sup>&</sup>lt;sup>7</sup> In 't Veld et al. (2019) derive this payoff function from a liquidity trade model with an infinite number of periods (Proposition 1 in the working paper), as a possible microfoundation.

Table 1

Assumptions about the payoff shares  $f_l(m_{ii}, \delta)$ ,  $f_b(m_{ii}, \delta)$  and  $f_m(m_{ii}, \delta)$ .

Share of payoffs for:	lender	borrower	middlemen
Number of middlemen <i>m</i>	$\frac{\partial f_l}{\partial m} > 0$	$\frac{\partial f_b}{\partial m} > 0$	$\frac{\partial f_m}{\partial m} < 0$
Level of competition $\boldsymbol{\delta}$	$\frac{\partial f_l}{\partial \delta} > 0$	$\frac{\partial f_b}{\partial \delta} > 0$	$\frac{\partial f_m}{\partial \delta} < 0$ $f_m(., 1) = 0$

this result also holds if i and j are connected at distance more than 2. This is shown in the proofs of Propositions 1 and 2, in Appendix A.

If the network distance between *i* and *j* is at most 2, trade is realized, and agents divide the surplus of  $\alpha_i \alpha_j$  in the following way. Consider a trade between a lender *i* and a borrower *j*. We let the share that each agent obtains depend, first, on the network structure *g*, and second, on a parameter  $\delta \in [0, 1]$ . This parameter captures the level of *competition* between the number of middlemen  $m_{ij}$  of a certain trade between *i* and *j*. If  $\delta = 0$  ('no competition'), then middlemen collude and act as if there were only a single middleman between *i* and *j*. If  $\delta = 1$  ('perfect competition'), then middlemen engage in Bertrand competition whenever  $m_{ii} > 1$ ; then banks *i* and *j* share the surplus.

More concretely, if *i* and *j* are directly connected, then the share for bank *i* is assumed to be  $f_l(0, \delta) > 0$ , and the share for *j*:  $f_b(0, \delta) > 0$ , such that  $f_l(0, \delta) + f_b(0, \delta) = 1$ . Note that the level of competition between middlemen is irrelevant for this trade (as the trade is direct), so  $f_l(0, \delta)$  and  $f_l(0, \delta)$  are independent of  $\delta$ .

If *i* and *j* are indirectly connected in *g* by  $m_{ij}$  middlemen, then bank *i* (the lender) receives a share of  $f_l(m_{ij}, \delta)$  and bank *j* (the borrower) receives a share of  $f_b(m_{ij}, \delta)$ . Each of the  $m_{ij}$  middlemen receives a share of  $f_m(m_{ij}, \delta)$  if  $k \in M_{ij}(g)$  and 0 otherwise. Note that by definition:

$$f_l(m_{ij}, \delta) + m_{ij}f_m(m_{ij}, \delta) + f_b(m_{ij}, \delta) = 1.$$

If there is one middleman,  $m_{ij} = 1$ , then we have  $f_l(1, \delta) < f_l(0, \delta)$ ,  $f_b(1, \delta) < f_b(0, \delta)$ , and  $f_m(1, \delta) = 1 - f_l(1, \delta) - f_b(1, \delta) > 0$ , and then the shares are independent of  $\delta$ . If there is more than one middleman, then the distribution over agents depends on the level of competition. If  $\delta = 0$ , the middlemen collude, and *i* and *j* obtain the same share as if there were one intermediator,  $f_l(m_{ij}, 0) = f_l(1, \delta)$  and  $f_b(m_{ij}, 0) = f_b(1, \delta)$ . The middlemen share the intermediation benefits evenly, i.e.  $\forall m_{ij} \in \{2, ..., n-2\}$ :  $f_m(m_{ij}, 0) = f_m(1, \delta)/m_{ij}$ .

If  $\delta = 1$  and there are  $m_{ij} > 1$  intermediaries, perfect competition drives the intermediary shares to 0, and  $f_l(m_{ij}, 1) = f_l(0, \delta)$ ,  $f_b(m_{ij}, 1) = f_b(0, \delta)$  and  $f_m(m_{ij}, 1) = 0$ . A further straightforward assumption is monotonicity with respect to the competitiveness and number of middlemen. Table 1 summarizes the assumed dependencies of the surplus distributions  $f_l(\cdot), f_b(\cdot)$  and  $f_m(\cdot)$  to the parameters.

The effect of competition parameter  $\delta$  on the division of the trade surplus and intermediation benefits can be thought of as arising from a bargaining process, in which  $\delta$  is the discount factor of players, such that bargaining power of the banks *i* and *j* increases with  $\delta$  and intermediation benefits decrease. We do not model this bargaining process explicitly. However, our assumptions on  $f_l(\cdot)$ ,  $f_b(\cdot)$  and  $f_m(\cdot)$  generalize an explicit Rubinstein-type of bargaining process developed by Siedlarek (2015).<sup>8</sup> From his bargaining protocol the distribution of the surplus is given by

$$f_l(m,\delta) = f_b(m,\delta) = \frac{m-\delta}{m(3-\delta)-2\delta} \text{ and } f_m(m,\delta) = \frac{1-\delta}{m(3-\delta)-2\delta}.$$
(2)

It is easily checked that (2) satisfies the assumptions we made on  $f_l(\cdot)$ ,  $f_b(\cdot)$  and  $f_m(\cdot)$ . Below we will use this explicit function to illustrate our (more general) results in Figs. 6, 9 and 10.

Fig. 3 gives examples of payoff shares received by different agents involved in a trade between *i* and *j* for different levels of competition  $\delta$  and different numbers of middlemen  $m_{ij}$ , given equations (2). In Appendix A of the working paper version of this paper – In 't Veld et al. (2019) – we give a detailed explanation of the specification of payoffs by Siedlarek (2015).

In the rest of this paper, we do not distinguish between borrowers and lenders, but focus on the difference between middlemen and *end nodes*. As we denote the share of end nodes as  $f_e(\cdot) \equiv (f_l(\cdot) + f_b(\cdot))/2$ , this assumption does not impose restrictions on bargaining power between borrowers and lenders.

More specifically, we also fix  $f_e(\delta, 1) = f_m(\delta, 1) = \frac{1}{3}$ . In other words, we assume that middlemen with a monopoly position between pairs of lenders and borrowers get a share of one third of their trades. This equality assumption simplifies the exposition without qualitatively changing the results.<sup>9</sup>

<sup>&</sup>lt;sup>8</sup> In this bargaining process,  $\delta$  has the usual interpretation of a discount factor, such that, if  $\delta$  is higher, intermediary agents are forced to make more competitive offers to end players.

<sup>&</sup>lt;sup>9</sup> In the formulas in Theorem 1, the assumption is visible in the recurring factor  $\frac{1}{3}$ .



Fig. 3. Examples of payoff shares received by endnodes i and j and intermediaries k, depending on the parameters  $\delta$  and m under the specification of  $f_{b}$ ,  $f_l$  and  $f_m$  in equation (2), as in Siedlarek (2015).

# 2.3. Network stability concepts

Given the setup of the payoffs discussed above, we analyze which networks arise if agents form links strategically in stage t = 0.<sup>10</sup> Here we assume that, in order to establish a link between two agents, both agents have to agree and both agents face the cost of a link, a version of network formation that is called *two-sided network formation*.<sup>11</sup> Network formation theory has developed stability or equilibrium concepts to analyze the stability of a network. Here, stability does not refer to systemic risk, but to the question whether an agent or a pair of agents has an incentive and the possibility to modify the network in order to receive a higher payoff.

There are many stability concepts, which differ in the network modifications allowed. For an overview of these stability concepts, we refer to Jackson (2005) or Goyal (2009). For our purposes we consider two stability concepts.

The first concept is *pairwise stability*, a standard concept in the literature (Jackson and Wolinsky, 1996). A network is pairwise stable if, for all the links present, no player benefits from deleting the link, and for all the links absent, one of the two players does not want to create a link. Denote the network  $g + g_{ij}$  as the network identical to g except that a link between i and j is added. Similarly, denote  $g - g_{ii}$  as the network identical to g except that the link between i and j is removed. While focusing on comparing networks, we drop the arguments of  $\delta$  and c in the function  $\pi(\cdot)$ . Then the definition of pairwise stability is as follows:

**Definition 2.** A network g is pairwise stable if for all  $i, j \in N$ ,  $i \neq j$ :

- (a) if  $g_{ij} = 1$ , then  $\pi_i(g) \ge \pi_i(g g_{ij}) \land \pi_j(g) \ge \pi_j(g g_{ij})$ ; (b) if  $g_{ij} = 0$ , then  $\pi_i(g + g_{ij}) > \pi_i(g) \Rightarrow \pi_j(g + g_{ij}) < \pi_j(g)$ .

The concept of pairwise stable networks only allows for deviations of one link at a time. This concept is often too weak to draw distinguishable conclusions, i.e. in many applications including ours, there are many networks that are pairwise stable.

In our application, we consider it relevant that agents may consider to propose many links simultaneously in order to become an intermediary and establish a client base. The benefits from such a decision may only become worthwhile if the agent is able to create or remove more than one link. This leads us naturally to the concept of *unilateral stability*, originally proposed by Buskens and Van de Rijt (2008).<sup>12</sup>

A network is unilaterally stable if no agent *i* in the network has a profitable unilateral deviation: a change in its links by either deleting existing links such that *i* benefits, or proposing new links such that *i* and all the agents to which it proposes a new link benefit. Denote  $g^{iS}$  as the network identical to g except that all the links between i and every  $j \in S$  are altered by  $g_{ii}^{iS} = 1 - g_{ii}$ , i.e. are added if absent in g or are deleted if present in g.

**Definition 3.** A network g is *unilaterally stable* if for all i and for all subsets of players  $S \subset N\{i\}$ :

- (a) if  $\forall j \in S : g_{ij} = 1$ , then  $\pi_i(g) \ge \pi_i(g^{iS})$ ; (b) if  $\forall j \in S : g_{ij} = 0$ , then  $\pi_i(g^{iS}) > \pi_i(g) \Rightarrow \exists j : \pi_j(g^{iS}) < \pi_j(g)$ .

<sup>&</sup>lt;sup>10</sup> In this paper we do not discuss efficiency. See our working paper version (In 't Veld et al., 2019) for a characterization of the efficient network under homogeneity (Theorem 2 in the working paper).

<sup>&</sup>lt;sup>11</sup> See Goyal (2009) for a text book discussion.

<sup>&</sup>lt;sup>12</sup> Bilateral equilibrium, as in Goyal and Vega-Redondo (2007), does not allow for the addition of multiple links. Therefore, bilateral equilibrium excludes the possibility of intermediators entering into the core, and is less appropriate in our setting than unilateral stability. Under bilateral equilibrium, a pair of agents that add a link, may simultaneously delete some of their other links. Note that bilateral equilibrium implies pairwise stability: a network that is a bilateral equilibrium is also pairwise stable, but not vice versa.



Fig. 4. A core-periphery network with n = 8 and k = 3, and two potential deviations to the structure, one adding a link (+) and one deleting a link (-).

Note that, being a stronger concept, unilateral stability implies pairwise stability. That is, a network that is unilaterally stable is also pairwise stable, but not vice versa. This can be easily verified by considering subsets *S* that consist of only one node  $j \neq i$ .<sup>13</sup>

# 3. Stability and heterogeneity

Under which conditions do core-periphery networks constitute a stable architecture? We consider the two stability concepts described in Section 2.3: pairwise and unilateral stability. Section 3.1 contains our main result that core-periphery networks are not unilaterally stable under the assumption of homogeneity.<sup>14</sup> In Section 3.2 we introduce heterogeneity into the model and discuss the implications for stability.

# 3.1. Homogeneous banks

We first analyze the model for the baseline homogeneous case where all pairs generate the same trade surplus,  $\alpha_i = 1$  for all  $i \in N$ .<sup>15</sup> Proposition 1 provides conditions under which, in the homogeneous case, core-periphery networks are not pairwise stable.

**Proposition 1.** Let  $\alpha_i = 1$  for all  $i \in N$ , and let c > 0 and  $0 < \delta < 1$  be given. Suppose that g is a core-periphery network with  $K \subset N$  the set of core agents, and that for the number of core agents k = |K| it holds that  $2 \le k \le n - 3$ . Denote the set of agents connected to some player i as  $N_i$  with size  $n_i = |N_i|$ . If there are two core agents  $i, j \in K$  with  $N_i \supseteq N_j$  and  $n_j \ge k + 1$ , then g is not pairwise stable.

# **Proof.** See Appendix A.

By definition, Proposition 1 also holds for any stability concept that is stronger than pairwise stability. Hence, under the conditions mentioned in the proposition, the network is also not unilaterally stable (nor a bilateral equilibrium, etc.).

The intuition behind Proposition 1 is illustrated by the following example. Suppose that the core-periphery network in Fig. 1 (see Section 2.1) is pairwise stable. Such a stable network would imply that each pair of periphery banks, for example 4 and 8, does *not* have an incentive to trade directly with each other, and at the same time, every pair of core banks *does* have an incentive to trade directly, such as 1 and 3. Fig. 4 shows the two associated deviations from the core-periphery structure. The reason that this structure is not pairwise stable, is because periphery agent 4 has only one core agent as intermediator (namely 1), while agent 1 has three potential intermediators to connect to core bank 3 (namely 2, 7 and 8). After all, periphery banks may act as intermediators between core banks. Given that two periphery banks trade indirectly, then two core banks have an incentive to do the same. Hence, core bank 1 should delete its link with 3, contradicting the pairwise stability of the network.

The condition in Proposition 1 that there exist two core agents *i* and *j* with  $N_i \supseteq N_j$  and  $n_j \ge k + 1$ , guarantees that core bank *i*, after deleting a link with core bank *j*, retains the same access to the periphery. In the example,  $N_1 \supseteq N_3$ , so that player 1 can delete its link with 3, as all periphery players connected to 3 are also connected to 1. Player 1 may not wish to delete its link to 2, however, as 2 provides access to periphery player 6.

<sup>&</sup>lt;sup>13</sup> Our definition of unilateral stability is slightly less restrictive than the definition in Buskens and Van de Rijt (2008). While we are following Buskens and Van de Rijt (2008) in considering deviations of simultaneously deleting *or* proposing multiple links, we do not allow simultaneously deleting *and* proposing of multiple links. This adaptation eases the exposition, but does not affect our results qualitatively.

<sup>&</sup>lt;sup>14</sup> We emphasize that the results in this section do not depend on the assumption that trade surpluses between *i* and *j* are only realized if the path length between *i* and *j* is less than 3, as shown in Appendix A.

<sup>&</sup>lt;sup>15</sup> The normalization  $\alpha = 1$  in the homogeneous case goes without loss of generality. If  $\alpha_i = \alpha$  for all *i* with  $\alpha > 0$ , then the payoffs in equation (1) are proportional to  $\alpha$  when costs *c* are considered as a fraction of  $\alpha$ .



Fig. 5. A core-periphery network with n = 8 and k = 3, and a potential deviations to the structure with multiple links added (+).

The condition in Proposition 1 is fairly mild. For a core-periphery networks to be pairwise stable, *all* core agents are necessarily connected to a periphery agent with a single link. In this situation, every core player has local monopoly power. Empirically, relevant interbank structures typically show levels of connectivity between core and periphery that exclude this local monopoly power.<sup>16</sup>

Even core-periphery networks not excluded by Proposition 1 are often unstable. So the condition is sufficient but not necessary. We now consider unilateral deviations, in which agents are allowed to add or delete multiple links. We show that all core-periphery networks are unstable, as long as there are enough periphery banks.<sup>17</sup>

**Proposition 2.** Let the payoff function be homogeneous ( $\alpha_i = 1$  for all  $i \in N$ ) and let c > 0 and  $0 < \delta < 1$  be given. Then, there is a function  $F(c, \delta)$ , such that, if  $n - k > F(c, \delta)$ , a core-periphery network with k core and n - k periphery players is not unilaterally stable.<sup>18</sup>

# **Proof.** See Appendix A. $\Box$

The intuition for Proposition 2 is that, for n - k large enough, periphery banks have an incentive to enter the core; e.g. player 4 in Fig. 5. Because we allow for multiple links to be added simultaneously, peripheral players can take a share of the intermediation benefits by replicating the position of core players. Unequal payoffs between core and periphery players makes the core-periphery networks unstable.

# 3.2. Heterogeneous banks

We found that core-periphery networks were generally not stable, when agents are ex ante identical. In real interbank markets we do observe links within the core. We now try to explain this discrepancy.

Key in Propositions 1 and 2 is the assumed homogeneity: periphery banks have the same capabilities as core banks in terms of profit generation, intermediation or linking, such that they can easily replace or imitate a core bank. In practice, we see large differences between banks, in particular banks in the core are much bigger than banks in the periphery (Craig and Von Peter, 2014). It is natural to think that these big banks have a strong incentive to have tight connections within the core as well as to the periphery for intermediation reasons.

Now we analyze the consequences of heterogeneity within our model. We start by considering exogenous heterogeneity in the trade surplus of pairs of banks within our model.<sup>19</sup> We interpret this heterogeneity in trade surplus as arising from differences in bank size. It seems likely there is a relation between bank size and trade surplus, as larger banks are able to use a larger leverage. Moreover, a larger staff may enable a bank to detect more trading opportunities and act on it. Also, a larger bank has more clients that generate more trading requests than a smaller bank.

We introduce two types of banks, k big banks and n - k small banks, and model the trade surplus between i and j as proportional to their size,  $\alpha_i$  and  $\alpha_j$ . If a bank i is big we assume  $\alpha_i = \alpha \ge 1$ , and if i is small we assume  $\alpha_i = 1.2^{20}$  So the difference in size is captured by a parameter  $\alpha \ge 1$  quantifying the relative size of a big bank. In this subsection we take their size as exogenously given. Given heterogeneity in the size of banks, we will show that a stable core-periphery network can form.

<sup>&</sup>lt;sup>16</sup> Proposition 1 considers core-periphery networks with  $n - k \ge 3$ , i.e. the number of periphery agents is at least three, which is generally satisfied for actual financial networks. If n - k = 2, and the set of connections of core player *j* is a subset of the connections of core player *i* are both connected to two distinct core agents, then given  $\delta$  there is a unique value of c > 0 for which stability cannot be excluded.

<sup>&</sup>lt;sup>17</sup> Propositions 1 and 2 still leave open the possibility for unilaterally stable core-periphery networks with local monopoly power and small network size. Indeed, in Appendix E of In 't Veld et al. (2019) we do give examples of such stable core-periphery networks. Below we show that these pairwise stable networks cannot be the outcome of best-response dynamics (Theorem 1 in Section 4.1).

<sup>&</sup>lt;sup>18</sup> The proposition does not rule out that the considered networks are pairwise stable, or a bilateral equilibrium.

<sup>&</sup>lt;sup>19</sup> In Section 4.2, heterogeneity is endogenized by extending a dynamic process to include feedback of profits on the surplus of trading.

<sup>&</sup>lt;sup>20</sup> The normalization  $\alpha_i = 1$  for small banks goes without loss of generality. Compare footnote 15.

The number of big banks, k, has become a new exogenous parameter, and we consider a complete core-periphery network where the core consists of the k big banks. Proposition 3 states that the complete core-periphery network can be unilaterally stable for a sufficient level of heterogeneity.

**Proposition 3.** Consider the model with the following form of heterogeneity:  $\alpha_i = \alpha$  for  $i \in \{1, 2, ..., k\}$  ('big banks'), and  $\alpha_i = 1$  for  $i \in \{k + 1, k + 2, ..., n\}$  ('small banks'). Then, if

$$c \ge \frac{1}{2} - f_e(k,\delta) + (n-k-2)\min\left\{\frac{1}{2}f_m(k+1,\delta), f_e(k+1,\delta) - f_e(k,\delta)\right\},\$$

there exists an  $\overline{\alpha} > 1$ , such that for all  $\alpha > \overline{\alpha}$  the complete core-periphery network with k big banks is unilaterally stable.

**Proof.** See Appendix A.

In the limits  $\delta \downarrow 0$ ,  $\delta \uparrow 1$  and  $k \to \infty$ , the required level of heterogeneity  $\overline{\alpha}$  is arbitrarily close to 1, as shown in Appendix A. Proposition 3 combined with Propositions 1 and 2 implies that heterogeneity is crucial in understanding core-periphery networks. Complete core-periphery networks are never stable under homogeneous players, but can be stable for arbitrary small levels of heterogeneity.

# 4. Best-response dynamics

So far, core-periphery networks were shown to be generally unstable under homogeneous agents, and possibly stable under heterogeneous agents. The question remains what happens in homogeneous networks, and, if not core-periphery networks, what other kind of network architectures arise. This motivates us to consider a simple dynamic process of network formation and analyze its stable states.

Section 4.1 introduces a best-response dynamic process to find what structures arise in the homogeneous case. In Section 4.2, heterogeneity is endogenized by extending the dynamic process to include feedback of profits on trade surplus.

# 4.1. Dynamics under homogeneous banks

We consider a round-robin best-response-like dynamic process as in Kleinberg et al. (2008). We order nodes 1, 2, ..., n. In the first round, we start from the empty graph and let nodes consecutively try to improve their position by taking a *best feasible action*. An action of player *i* is defined as feasible if *i* either proposes links to a subset of players *S* such that every  $j \in S$  accepts a link with *i*, or if *i* deletes a subset of its links. The best feasible action of player *i* is then the feasible action that delivers *i* the highest payoff under the current graph (so in a myopic way). The formal definition of a best feasible action, based on the concept of unilateral stability, is given in Appendix A. Note that a unilaterally stable network is a network in which all players choose a best feasible action.<sup>21</sup>

After node n has chosen its best feasible action, the second round restarts with node 1, and again each player consecutively chooses its best feasible action. Next, the third round starts, and so on, until convergence. The process converges if n - 1 consecutive players cannot improve their position.

The assumptions about the dynamic process allow for a sharp characterisation of stable network structures. The advantage of starting in an empty network is that, initially, nodes can only add links to the network. A fixed round-robin order limits the number of possibilities that has to be considered for every step in the process. Simulations for our model indicate that the results below also hold for a random order of agents.<sup>22</sup> Excluding border line cases, we find that the dynamics always converge to a unilaterally stable network whose architecture is well described for every choice of the model parameters.<sup>23</sup>

Theorem 1 below shows which network structures result from the dynamics. The empty, star and complete networks are prominent for large parameter regions. In the remaining parameter regions the attracting steady state cannot be a coreperiphery network (in line with Propositions 1 and 2) and turns out to be a multipartite network. The theorem singles out one special type of multipartite networks, called maximally unbalanced bipartite networks, for parameters satisfying condition IV, while parameters under condition V can lead to various types of multipartite networks.

**Theorem 1.** Consider the homogeneous baseline model with  $\alpha_i = 1$  for all  $i \in N$ . From an empty graph, the round-robin best-feasible-action dynamics converge to a unilaterally stable network with the following network architecture for the following parameter regions:

<sup>&</sup>lt;sup>21</sup> If a player's best feasible action is not unique, then we assume that a player chooses randomly from the set of best feasible actions. In Theorem 1 below we focus on parameter regions in which the best feasible action is always unique.

<sup>&</sup>lt;sup>22</sup> Houy (2009) considers a dynamic framework for the model of Goyal and Vega-Redondo (2007), in which randomly drawn agents make best responses to the network.

<sup>&</sup>lt;sup>23</sup> In principle, a dynamic process may lead to cycles of improving networks (cf. Jackson and Watts, 2002; Kleinberg et al., 2008), but Theorem 1 shows that this is not the case in our model.

- I: for  $c > \frac{1}{2} + \frac{1}{6}(n-2)$ the empty network,
- II: for  $c \in \left(\frac{1}{6} + (n-3)\min\{\frac{1}{2}f_m(2,\delta), f_e(2,\delta) \frac{1}{3}\}, \frac{1}{2} + \frac{1}{6}(n-2)\right)$ the star natural
- the star network, III: for  $c < \frac{1}{2} f_e(n-2, \delta)$ the complete network.
- IV: for  $c \in \left(\frac{1}{2} f_e(2,\delta) + (n-4)\min\{\frac{1}{2}f_m(3,\delta), f_e(3,\delta) f_e(2,\delta)\}, \frac{1}{6} + (n-3)\min\{\frac{1}{2}f_m(2,\delta), f_e(2,\delta) \frac{1}{3}\}\right)$  the maximum in mn(2)

mally unbalanced bipartite network 
$$g_{2,n-2}^{mp(2)}$$
,

V: for  $c \in \left(\frac{1}{2} - f_e(n-2,\delta), \frac{1}{2} - f_e(2,\delta) + (n-4)\min\{\frac{1}{2}f_m(3,\delta), f_e(3,\delta) - f_e(2,\delta)\}\right)$ 

a multipartite network  $g_{k_1,k_2,...,k_q}^{mp(q)}$  with  $q \ge 2$  and  $|k_m - k_{m'}| < n-4$  for all  $m, m' \in \{1, 2, .., q\}$ .

# **Proof.** See Appendix B.

Fig. 6 illustrates the parameter regions specified by Theorem 1 for n = 4 and n = 8 under the specification of  $f_e$  and  $f_m$ in equation (2), as in Siedlarek (2015). The possible network outcomes range intuitively from empty to complete networks as the cost of linking decreases.

The star is an important outcome in between empty and complete networks. Here, the outcome of Goyal and Vega-Redondo (2007), although under a different concept of network stability, arises as a special case in our model. Goyal and Vega-Redondo (2007) show that for  $c > \frac{1}{6}$ ,  $\delta = 1$  and *n* sufficiently large, the star network is the unique non-empty *bilateral equilibrium* network. Theorem 1 implies that for  $\delta = 1$ , the star network is the outcome of the *unilateral* best-response dynamics in the range  $c \in \left(\frac{1}{6}, \frac{1}{2} + \frac{1}{6}(n-2)\right)$ .

The star, however, cannot be a unilaterally stable outcome for intermediate competition  $\delta$  and relatively low *c*. For intermediate  $\delta$  the incentives to enter the core and the incentives for periphery players to accept the proposal of the new core player are both high. The parameter area between complete networks and stars gives multipartite networks as the stable outcomes, and this area increases with *n*; see Fig. 10a for the results for n = 100.

It is noteworthy that for large *n* the set of attained multipartite networks is quite diverse. Possible outcomes for n = 8 are the maximally unbalanced bipartite network  $g_{2,6}^{mp(2)}$ , but also, for example, an unbalanced bipartite network  $g_{3,5}^{mp(2)}$  or a balanced multipartite network  $g_{2,3,3}^{mp(3)}$  consisting of 3 groups. Fig. 7 presents these three examples of multipartite networks and the values of  $\delta$  and c for which they are formed.

Interestingly, multipartite networks are also the equilibrium outcome in the models of two closely related papers: Buskens and Van de Rijt (2008) and Kleinberg et al. (2008). Although the payoff functions differ, as we do in our paper, they analyze models in which intermediators in a social network, or network brokers, obtain a higher payoff. Hence, the appearance of multipartite networks seems to be a natural outcome in environments in which network brokerage or intermediation takes place.

Appearance of multipartite networks may have implications for network processes taking place in the system. Golub and Jackson (2012) show that the speed of learning and best-response dynamics depend crucially on the amount of homophily in the network, that is, to what degree similar people link to each other. Learning will be slower in networks that are more homophilous. However, multipartite networks are perfectly heterophilous; that is, the nodes can be partitioned into groups, and within a group no links are present. This suggests that the speed of learning in multipartite networks is likely to be rather high. For example, traders in financial over-the-counter markets may learn equilibrium prices and quantities relatively guickly in a multipartite network.<sup>24</sup>

Overall, Theorem 1 implies that empty, star, complete and multipartite networks can all arise within a network formation model with intermediation and imperfect competition. Depending on the network size, linking costs and level of competition, the attained equilibrium in the dynamic process is unique. Our model thus places earlier results of homogeneous network formation models into a more general perspective.

Fig. 8 shows the possible routes of the dynamics for n = 4, depending on the remaining parameters c and  $\delta$ . For n = 4, the only possible multipartite network is one that consists of q = 2 groups of 2 nodes, which coincides with a ring of all 4 players. The steps towards this multipartite network are as follows. Starting from an empty network, in the first round a star is formed. One of the periphery players then has an incentive to join the core, such that in the second round a complete core-periphery network is formed. By Proposition 1, this core-periphery network is not stable, as core players have an incentive to trade indirectly with each other. Hence, the link within the core is dropped and the multipartite network is formed.

<sup>&</sup>lt;sup>24</sup> See Babus and Kondor (2018) for a model of trade in over-the-counter markets. In Section 6 of their paper, they present a dynamic price discovery model that converges to the equilibrium outcomes of their one shot over-the-counter model. The structure of the dynamics of their price discovery model is similar to that of the model of Golub and Jackson (2012).



**Fig. 6.** Attained equilibria after best-feasible-action dynamics from an empty network in  $(\delta, c)$ -space for  $\alpha_i = 1$  for all  $i \in N$  and  $n \in \{4, 8\}$  under the specification of  $f_e$  and  $f_m$  in equation (2). The roman numbers correspond with those in Theorem 1. The symbols ( $\blacksquare$ ,  $\blacktriangle$  and  $\star$ ) correspond to examples of multipartite networks in Fig. 7.

# 4.2. Endogenous heterogeneity

From the analysis in Section 3.2 it follows that heterogeneity is a necessary condition for a stable core-periphery network. We have shown that under the assumption of ex ante heterogeneity in trade surpluses, stable core-periphery networks arise for large regions in the parameter space. The assumption of heterogeneity is quite realistic, given the amount of heterogeneity between banks in practice.

Nevertheless, it is of interest whether the formation process in itself may generate sufficient payoff differences between banks as to form an endogenous core-periphery network structure. To this end we extend the dynamic process to allow for



(a) Black square  $\blacksquare$ :  $(\delta, c) = (0.8, 0.3)$ . The maximally unbalanced bipartite network:  $g_{26}^{mp(2)}$ .



(b) Black triangle  $\blacktriangle$ : ( $\delta$ , c) = (0.4, 0.15). An unbalanced bipartite network:  $g_{35}^{mp(2)}$ .



(c) Black star  $\bigstar$ : ( $\delta$ , c) = (0.8, 0.06). A balanced multipartite network:  $g_{2,3,3}^{mp(3)}$ 

**Fig. 7.** Examples of attained multipartite networks after best-feasible-action dynamics from an empty graph for n = 8 players. The symbols ( $\blacksquare$ ,  $\blacktriangle$  and  $\star$ ) correspond to locations in the ( $\delta$ , c)-space in Fig. 6.

feedback of profits on bank size. Using simulations, we will show that core-periphery networks can be the outcome of a dynamic process, when bank size is updated according to profits.

The extended dynamic process starts off at time  $\tau = 1$  as the homogeneous baseline model with  $\alpha_i^{(1)} = 1$  for all  $i \in N$ . From an empty graph, round-robin best-feasible-action dynamics converge to the empty network, the star network, the complete network or multipartite networks, as in Theorem 1. In this attained network the profits are  $\pi_i^{(1)}$ . From then on, at the beginning of time  $\tau = 2, 3, 4, \ldots$  banks sizes are updated as

$$\alpha_i^{(\tau)} = \frac{\pi_i^{(\tau-1)}}{\min_k [\pi_k^{(\tau-1)}]}$$
(3)

As discussed in Section 3.2, trade surplus is assumed to be proportional to size. In this updating, we assume that profits are reinvested in the bank to scale up the size of the bank. We normalize bank sizes with  $\min_k[\pi_k]$  to assure that trade surpluses remain finite in the long run; without a normalization, the complete network would be the unique outcome in the long run. In general, finite market demand and banking competition limits the scale of the banks.

Having updated the bank sizes  $\{\alpha_i\}_{i \in N}$  at the beginning of time  $\tau$ , we re-initiate the process of the round-robin best-feasible-action dynamics, starting from the attained network  $g^{(\tau)}$ . This leads potentially to a different network structure and other payoffs at time  $\tau + 1$ . At this point, the bank sizes are again updated using (3), and so on.

We simulate this dynamic process. As a stopping rule for the simulations, we impose that the process stops at time  $\mathcal{T} = 25$ , or before at time  $\tau$  if the bank sizes do not alter any more given a tolerance level  $\Delta \alpha$ , i.e. if  $|\alpha_i^{(\tau)} - \alpha_i^{(\tau-1)}| < \Delta \alpha$  for all  $i, j \in N$ . We choose  $\Delta \alpha = 0.1$ . Theoretically, it is possible that the system exhibits recurring cycles. In our simulations presented below, we have checked that the dynamics converge within the tolerance level for all  $\delta \leq 0.8$ . For  $\delta$  close to 1, dynamics do not converge before time  $\mathcal{T}$  and we cannot exclude cyclic behavior in that case.

Fig. 9 plots the results for n = 8 and  $\Delta \alpha = 0.1$  with  $\mathcal{T} = 1$  (left panel) and after 25 runnings  $\mathcal{T} = 25$  (right panel), using a 25x25 grid in the  $(\delta, c)$ -space ( $\delta \in [0, 1]$  and  $c \in [0, 0.4]$ ). The black regions correspond to complete networks, green regions to star networks, blue and purple to multipartite networks and different shades of red to core-periphery networks. In the case of  $\mathcal{T} = 1$ , banks are still homogeneous, that is,  $\alpha_i^{(1)} = 1$  for all  $i \in N$ . The left panel of Fig. 9 therefore repeats Fig. 6b. As we have seen in Subsection 4.1, in that case, core-periphery networks do not occur.



**Fig. 8.** Map of the dynamics from an empty graph for n = 4 leading to one out of 4 stable possible structures. The roman numbers correspond with the conditions as in Theorem 1 under which the best-feasible-action route follows the direction of the arrows.

The right panel of Fig. 9 shows the results after a maximum of T = 25 updates in the parameters  $\alpha_i$ . The black regions remain unchanged, as the banks are in symmetric positions. All banks obtain the same payoffs, and bank sizes do not change. Also the star network cannot change: profit feedback increases the size  $\alpha_i$  of the center of the star i = 1. Links with the center thus generate higher payoffs, and will not be severed. Reversely, links between periphery players do not generate higher payoffs, so no links are added. Hence, if a star network is formed under homogeneity at time  $\tau = 1$ , then the network architecture remains a star network after any future rounds.

This is not the case, however, if at time  $\tau = 1$  a multipartite network is formed. In that case, we observe significant changes over time. In fact, in many simulations, we observe a transition from multipartite networks to core-periphery networks. This happens intuitively. For example, if at time  $\tau = 1$  a bipartite network with two banks in tier 1 and six banks in tier 2 are formed  $(g_{2,6}^{mp(2)})$ , lighter blue region), the payoffs of the tier-1 banks is much higher than the payoffs of the banks in tier 2. This is, because the two tier-1 banks intermediate between the six tier-2 banks, receiving much more intermediation benefits than the tier-2 banks. With the feedback mechanism specified in the dynamics, these higher payoffs for the tier-1 banks result in larger bank size  $\alpha_i$ , such that in the end the two banks in tier-1 have an incentive to form a direct link. The network architecture then converts into a core-periphery network. Similarly, many multipartite networks in the darker blue regions evolve into red regions with a core-periphery networks architecture.<sup>25</sup>

We conclude that for many parameter values, core-periphery networks can arise endogenously in the extended model with ex ante homogeneous banks and feedback of the payoffs on trade surpluses.

# 5. Applying the model to the Dutch interbank market

To gain some insight into the general applicability of the model, we calibrate the model to the Dutch interbank market. The network structure of bilateral exposures between banks in the Netherlands, a relatively small market with approximately

 $<sup>^{25}</sup>$  For  $\delta$  close to 1, the results have to be interpreted with care. The reason is that given the high level of competition, central players may earn less than peripheral players. After the updating of bank size, central players may be induced to remove several links, and cycles with different players taking central positions may occur.



**Fig. 9.** Attained equilibria after best-feasible action dynamics from an empty graph in ( $\delta$ , c)-space for n = 8, extended with profit feedback using  $\Delta \alpha = 0.1$  and  $\mathcal{T} = 1$  (left panel) or  $\mathcal{T} = 25$  (right panel). The simulations are run under the specification of  $f_e$  and  $f_m$  in equation (2). The black regions correspond to complete networks, green regions to star networks, blue and purple to multipartite networks and different shades of red to core-periphery networks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

100 financial institutions, has been investigated before by In 't Veld and Van Lelyveld (2014). The attracting networks in the dynamic homogeneous model with n = 100 are indicated in Fig. 10a. Under homogeneous banks our model predicts multipartite networks for most parameter values. In contrast, In 't Veld and Van Lelyveld (2014) found that the observed network contains a very densely connected core with around k = 15 core banks.

We choose a level of heterogeneity  $\alpha = 10$  to capture in a stylized way the heterogeneity of banks in the Netherlands. In reality, banks in the core as well as in the periphery of the Dutch banking system are quite diverse. A few very large banks reach a total asset size of up to  $\in 1$  *trillion*, while the asset value of some investment firms active in the interbank market may not be more than a few million euro. The median size of a core bank lies around  $\in 8$  billion; for periphery banks the median size is approximately  $\in 300$  million (In 't Veld and Van Lelyveld, 2014).<sup>26</sup> Ignoring the exceptionally large size of some core banks, a relative difference of  $\alpha = 10$  is a reasonable order of magnitude. The larger  $\alpha$ , the wider the area in which the core-periphery network is stable.

Fig. 10 shows the parameter region given by Proposition 3 for which a complete core-periphery network of the fifteen big banks is a unilaterally stable network. The stability of the complete network does not depend on  $\alpha$  as also indicated in Fig. 10b. These complete networks and complete core-periphery networks would also be the outcome of best-feasible-action dynamics starting with the big banks.<sup>27</sup> The observed core-periphery structure in the Netherlands can be reproduced for many reasonable choices of linking costs *c* and competitiveness  $\delta$ .

This application suggests that our model is suitable to explain stylized facts of national or perhaps even international interbank networks. It should be noted that observed core-periphery networks don't necessarily satisfy the (theoretical) concept of complete core-periphery networks. Empirical studies of interbank markets often rely on either balance sheet data measuring the total exposure of one bank on another (such as In 't Veld and Van Lelyveld, 2014), or overnight loan data specifying the actual trades. We interpret the undirected links in our model as established preferential lending relationships, which are typically unobserved in practice. Trades and exposures are executed on this structure of long-term relationships, approximated by a core-periphery network. The empirical (observed) core-periphery structure from these trade networks is a subset of the (unobserved) relationship network, and hence, not necessarily contains all connections between core and periphery. In any case, the densely connected core of a subset of the banks is a well-documented empirical fact that is reproduced by our model.

<sup>&</sup>lt;sup>26</sup> See Fig. 9 of In 't Veld and Van Lelyveld (2014) for the plotted distribution of total asset size over core and periphery banks.

<sup>&</sup>lt;sup>27</sup> We do not present a full description of the outcomes of these dynamics, specifying all other areas in Fig. 10b. Simulations show that for linking costs so high that CP networks are not stable, many different structures arise. As none of them are core-periphery networks, we restrict ourselves to Proposition 3. Proposition 5 in In 't Veld et al. (2019) provides an example of the possible outcomes of the dynamics for n = 4.



(a) Homogeneous banks  $(\alpha_{ij} = 1 \text{ for all } i, j \in N)$ 



(b) Heterogeneous banks: k = 15 big banks (with relative size  $\alpha = 10$ ) and n - k = 85 small banks

**Fig. 10.** Application of the model to the Dutch interbank market. Attained equilibria after best-response dynamics from an empty network in ( $\delta$ , *c*)-space for n = 100, k = 15 and  $\alpha \in 1$ , 10 under the specification of  $f_e$  and  $f_m$  in equation (2). The roman numbers correspond with those in Theorem 1 and Proposition 3. In the shaded area, the complete core-periphery network with k = 15 big banks is the unique unilaterally stable outcome of the dynamics.

# 6. Conclusion

In this paper we propose a way to explain the formation of financial networks by intermediation. We focus on the coreperiphery network because it is found to give a fair representation of the complex empirical structures (In 't Veld and Van Lelyveld, 2014), while at the same time being relatively simple and intuitively appealing. In our model, brokers strive to intermediate between their counterparties and compete with each other. Our results suggest that heterogeneity is crucial, that is, the core-periphery structure of the interbank market cannot be understood separately from the heterogeneity and inequality in the intrinsic characteristics of banks. We explore these results further in a dynamic extension of the model. We endogenize heterogeneity by updating the size of each bank with the payoffs received from trades in the network. Better connected banks receive higher payoffs by intermediation, which could feed back on the balance sheet and thus on potential surplus from trading. We find that core-periphery networks arise endogenously in the extended model with ex ante homogeneous agents and feedback of the network structure on trade surpluses.

We would like to make three suggestions for future research. First, introducing balance sheets would be one potential direction to extend our model. Research that analyzes financial contagion typically take the balance sheet and network structure as exogenously fixed. However, this forces them to make arbitrary assumptions on the size of the balance sheets, that is, total (interbank) assets and liabilities, when analyzing the effect of heterogeneity or core-periphery structure on financial contagion. For example, Nier et al. (2007) keep the size of banks' balance sheet constant, when analyzing the effect of a two-tier system on financial contagion. Our research suggests that one cannot impose such arbitrary regularities, and instead one has to think carefully on how heterogeneity in balance sheets and heterogeneity in network structure co-evolve. This point was made by Glasserman and Young (2016) as well.

Second, our model does not involve any risk of banks defaulting. Hence, our model is a model of a riskless interbank market, such as the interbank trading market approximately before the financial crisis of 2007. However, it is possible to analyze the effects of an increase of default risk on interbank lending as well by introducing a probability that a bank and its links default. Doing so allows us to understand network formation in stress situations. This seems relevant in light of findings that the fit of the core-periphery network in the interbank market deteriorated during the recent financial crisis (Fricke and Lux, 2015; In 't Veld and Van Lelyveld, 2014; Martínez-Jaramillo et al., 2014).

Finally, while we focus on financial networks, we believe that our model could be applied more broadly. Basically, the model is applicable to any setting in which trade is constrained by long term trading connections and in which traders imperfectly compete for intermediation. One interesting application would be markets in the developing world, as trade there is often constrained by social or cultural ties (Fafchamps, 2003; Fafchamps and Minten, 1999).

# Appendix A. Proofs of Section 3

Definition of best feasible action. We first introduce the concept of best feasible (unilateral) action of a certain player *i*. An action of player *i* is feasible if its proposed links to every  $j \in S$  are accepted by every  $j \in S$ , or if *i* deletes all links with every  $j \in S$ . This action changes the network *g* into  $g^{iS}$ . The best feasible action is formally defined as follows.

**Definition 4.** A feasible action for player *i* in network *g* is represented by a subset  $S \subseteq N \setminus \{i\}$  with:

(a) 
$$\forall j \in S : g_{ij} = 1$$
, or:  
(b)  $\forall j \in S : g_{ij} = 0$  and  $\pi_i(g^{iS}) \ge \pi_i(g)$ .

A best feasible action  $S^*$  for player *i* in network *g* is a feasible action in *g* that gives *i* the highest payoffs:

$$\forall S \subseteq N \setminus \{i\}: \pi_i(g^{iS^*}) \ge \pi_i(g^{iS}).$$

A network g is unilaterally stable if and only if, for all players,  $S^* = \emptyset$  is a best feasible action in g. In other words, no player has an incentive to take a feasible action that changes the network.

We continue with the proofs of Section 3.

**Proof of Proposition 1..** Suppose *g* is a core-periphery network, such that there are two core agents *i*,  $j \in K$  with  $N_i \supseteq N_j$  and  $n_j = |N_j| \ge k + 1$ . Note that by definition of a core-periphery network:  $g_{ij} = 1$ , as both *i* and *j* are in the core.

For the marginal benefit of any two peripheral players  $i_1$  and  $i_2$  to connect directly, it holds that:

$$\pi_{i_l}(g + g_{i_1 i_2}) - \pi_{i_l}(g) \ge \left(\frac{1}{2} - f_e(k, \delta)\right) - c \tag{A.1}$$

for l = 1, 2, as two peripheral players have by definition at most k intermediaries.

For the marginal benefit of core player *i* to delete its link to *j*, it holds that:

$$\pi_i(g - g_{ij}) - \pi_i(g) \ge c - \left(\frac{1}{2} - f_e(k, \delta)\right).$$
(A.2)

The value  $(\frac{1}{2} - f_e(k, \delta))$  equals the difference in payoff *i* receives from its link to *j*. After the removal of their direct link, player *i* is indirectly linked to *j* via at least *k* intermediaries, that is, k - 2 core and at least 2 periphery agents to which both *i* and *j* are connected. Note that *i* and *j* must have at least 2 periphery agents in common as  $n_i \ge n_i = |N_i| \ge k + 1$ .

Suppose now, contrary to the proposition, that g is pairwise stable. In order for the network to be pairwise stable, it is required that  $\pi_{i_l}(g + g_{i_li_2}) - \pi_{i_l}(g) \le 0$  and  $\pi_i(g - g_{i_j}) - \pi_i(g) \le 0$  hold simultaneously, which implies that

$$\pi_{i_i}(g + g_{i_1 i_2}) - \pi_{i_i}(g) = \pi_i(g - g_{ij}) - \pi_i(g) = \left(\frac{1}{2} - f_e(k, \delta)\right) - c = 0.$$
(A.3)

Hence,  $c = \left(\frac{1}{2} - f_e(k, \delta)\right)$ .

Consider now a third periphery node  $i_3 \in P$ . This node exists, since  $k \leq n - 3$ . Node  $i_3$  cannot be linked to both i or j, as in that case, i and j have at least 3 periphery neighbors in common, such that the marginal benefit for i to delete its link with j would be

$$\pi_i(g-g_{ij})-\pi_i(g) \geq c - \left(\frac{1}{2}-f_e(k+1,\delta)\right) > c - \left(\frac{1}{2}-f_e(k,\delta)\right),$$

which would violate (A.3). Consider now the marginal benefit for  $i_1$  and  $i_3$  of creating a direct link. As  $i_3$  is linked to at most k - 1 core nodes, the marginal benefit is

$$\pi_{i_l}(g+g_{i_1i_3})-\pi_{i_l}(g) \ge \left(\frac{1}{2}-f_e(k-1,\delta)\right)-c > \left(\frac{1}{2}-f_e(k,\delta)\right)-c,$$

for l = 1, 3. This violates (A.3). Hence, g is not pairwise stable if  $0 < \delta < 1$ .  $\Box$ 

**Proof of Proposition 2..** Consider a periphery player *i* proposing  $l_i = n - k - 1$  links in order to reach all other periphery players. The marginal benefits<sup>28</sup> for *i* of this action are bounded by:

$$M_{i}(g^{CP(k)}, +l_{i}) \geq -(n-k-1)\left(c-\frac{1}{6}\right) + \binom{n-k-1}{2}f_{m}(2,\delta).$$
(A.5)

Then positive marginal benefits for *i* are implied by:

$$M_{i}(g^{CP(k)}, +l_{i}) > 0 \ll n-k > F_{1}(c, \delta) \equiv 2 + \frac{c-\frac{1}{6}}{\frac{1}{2}f_{m}(2,\delta)}$$
(A.6)

For the deviation of  $l_i$  links to be executed, the peripheral players  $j \in (P \setminus i)$  also have to agree with the addition, that is, they should not receive a lower payoff. The marginal benefits for j are bounded by:

$$M_{j}(g^{CP(k)}, +l_{i}) \geq -c + (n-k-2)\left(f_{e}(2,\delta) - \frac{1}{3}\right)$$
(A.7)

Positive marginal benefits for *j* are implied by:

$$\begin{array}{ll} M_j(g^{CP(k)},+l_i) &\geq & 0\\ & \Leftarrow n-k &> & F_2(c,\delta) \equiv 2 + \frac{c}{f_e(2,\delta)-1/3} \end{array}$$

$$(A.8)$$

Combining the conditions for *i* (A.6) and *j* (A.8), a sufficient lower bound for the number of periphery nodes n - k for any core-periphery network to be unilaterally unstable is:

$$n - k > F(c, \delta) \equiv \max \{F_1(c, \delta), F_2(c, \delta)\}$$
(A.9)

*Remarks on Proposition 2.* In the derivation of this sufficient lower bound on *n*, we have not made any assumptions about the path lengths on which trade is allowed. It is sufficient to consider the parts of the marginal benefits for *i* and *j* that depend on the constant costs *c* and the new intermediation route is formed between *j*,  $l \in (P \setminus i)$ .

A core-periphery network is generally unstable because the inequality between core and periphery becomes large for increasing *n*, and a periphery player can always benefit by adding links to all other players. An important assumption for this result is therefore that multiple links can be added at the same time. It is crucial that  $\delta > 0$  because for  $\delta = 0$  intermediated trade always generates  $f_e(m, 0) = f_e^0 \forall m$ , i.e. additional intermediation paths do not increase profits for endnodes. Moreover it is crucial to impose  $\delta < 1$  because for  $\delta = 1$  intermediation benefits disappear, i.e.  $f_m(m, 1) = 0 \forall m > 1$ .

**Proof of Proposition 3..** We start with possible deviations of a peripheral player by adding links to other periphery players. Note that the change in payoffs by these deviations do not depend on  $\alpha$ , because they only concern trade surpluses between small banks. Lemma 2 (see below) holds in a heterogeneous setting as well: if adding one link improves the payoff of a small periphery player, the best feasible action is to connect to all other small banks.

Starting from a complete core-periphery network with the k big banks in the core, if one peripheral player adds links to all other periphery players, the core is extended to k + 1 players, k big banks and 1 small bank. For the new core member i to have positive marginal benefits of supporting all new links, it is required that

$$M_{i}(g_{com}^{CP(k)}, +(n-k-1)) > 0$$
  
$$\Rightarrow c < \frac{1}{2} - f_{e}(k,\delta) + \frac{1}{2}(n-k-2)f_{m}(k+1,\delta)$$
(A.10)

<sup>28</sup> Throughout the appendix, marginal benefits of an action S by player *i* for player *j* are defined as:

 $M_i(g, S) \equiv \pi_i(g^{iS}) - \pi_i(g).$ 

(A.4)

Suppose g is a core-periphery network, such that there exist two distinct periphery agents  $i_1$ ,  $i_2 \in P$  that are connected by two distinct core agents  $j_1$ ,  $j_2 \in K$ ;  $g_{i_1j_1} = g_{i_1j_2} = g_{i_2j_1} = g_{i_2j_2} = 1$ . Let k = |K| be the number of core nodes. Note that by definition of a core-periphery network.

and every remaining peripheral player  $j \neq i$  requires

$$M_{j}(g_{com}^{CP(k)}, +(n-k-1) \ge 0$$
  
$$\Rightarrow c \le \frac{1}{2} - f_{e}(k,\delta) + \frac{1}{2}(n-k-2) + (n-k-2)(f_{e}(k+1,\delta) - f_{e}(k,\delta)).$$
(A.11)

Conversely, the deviation of player i is not a best feasible action if c exceeds the minimum of the two values in equations (A.10) and (A.11):

$$c \ge \frac{1}{2} - f_e(k,\delta) + (n-k-2)\min\{\frac{1}{2}f_m(k+1,\delta), f_e(k+1,\delta) - f_e(k,\delta)\}.$$
(A.12)

Given a complete core-periphery network and a sufficiently high c satisfying (A.12), it is not beneficial for periphery players to add links.

We will now derive an  $\overline{\alpha}$ , such that for all  $\alpha > \overline{\alpha}$  the complete core-periphery network with *k* big banks is unilaterally stable. In order to derive  $\overline{\alpha}$  we consider the possible deletion of one or multiple links by either core or periphery players. After deriving  $\overline{\alpha}$ , we will rewrite the condition of stable complete core-periphery networks in terms of *c*.

First, consider a core player *i*. Player *i* can delete links with other core players and/or links with periphery players. The marginal benefit of deleting  $l^c$  core links and  $l^p$  periphery links is:

$$M_{i}(g_{com}^{CP(k)}, -(l^{c}+l^{p})) = l^{c} \left(c - \alpha^{2} \left(\frac{1}{2} - f_{e}(n-l^{c}-l^{p}-1,\delta)\right)\right) + l^{p} \left(c - \alpha \left(\frac{1}{2} - f_{e}(k-l^{c}-1,\delta)\right) - (2n-2k-l^{p}-1)f_{m}(k,\delta)\right)$$
  
$$\equiv M_{i}^{c} + M_{i}^{p}$$
(A.13)

The marginal benefit of deleting  $l^c + l^p$  links can be separated in benefits from deleting links with the core  $M_i^c$  and benefits from deleting links with the periphery  $M_i^p$ . The cross-over effects of deleting links with both groups of banks are negative:  $M_i^c$  is decreasing in  $l^p$  and  $M_i^p$  is decreasing in  $l^c$ . To find the conditions under which the core player does not want to delete any link, it is therefore sufficient to consider deletion of links in each group separately.

The marginal benefit of deleting  $l^c$  core links is:

-

$$M_{i}(g_{com}^{CP(k)}, -l^{c}) = l^{c} \left( c - \alpha^{2} \left( \frac{1}{2} - f_{e}(n - l^{c} - 1, \delta) \right) \right)$$
(A.14)

If  $M_i(g_{com}^{CP(k)}, -l^c) > 0$ , it must hold that  $M_i(g_{com}^{CP(k)}, -1) > 0$ , because  $f_e(n - l^c - 1, \delta)$  decreases in  $l^c$ . Player *i* thus has a beneficial unilateral deviation if

$$M_{i}(g_{com}^{CP(k)}, -l^{c}) > 0$$

$$\Rightarrow M_{i}(g_{com}^{CP(k)}, -1) > 0$$

$$\Leftrightarrow \alpha < \sqrt{c/(\frac{1}{2} - f_{e}(n-2, \delta))}$$
(A.15)

The marginal benefit for a core player i of deleting  $l^p$  links with the periphery is :

$$M_{i}(g_{com}^{CP(k)}, -l^{p}) = l^{p} \left( c - \alpha \left( \frac{1}{2} - f_{e}(k-1, \delta) \right) - (2n - 2k - l^{p} - 1) f_{m}(k, \delta) \right)$$
(A.16)

If  $M_i(g_{com}^{CP(k)}, -l^p) > 0$ , it must be a best feasible action to choose  $l^p = n - k$  and delete all links with the periphery, as the function is convex in  $l^p$ . Player *i* thus has a beneficial unilateral deviation if

$$M_{i}(g_{com}^{CP(k)}, -l^{p}) > 0$$
  

$$\Rightarrow M_{i}(g_{com}^{CP(k)}, -(n-k)) > 0$$
  

$$\Rightarrow \alpha < \frac{c - \frac{1}{2}(n-k)(n-k-1)f_{m}(k,\delta)}{\frac{1}{2} - f_{e}(k-1,\delta)}$$
(A.17)

Second, consider a player  $i \in P$  in the periphery. This marginal benefit for a periphery bank  $i \in P$  to delete l links is positive if:

$$M_{i}(g_{com}^{CP(k)}, -l) = l\left(c - \alpha\left(\frac{1}{2} - f_{e}(k - l, \delta)\right)\right) - (n - k - 1)(f_{e}(k, \delta) - f_{e}(k - l, \delta)) > 0$$
  
$$\Leftrightarrow \alpha < \frac{c - \frac{n - k - 1}{l}(f_{e}(k, \delta) - f_{e}(k - l, \delta))}{\frac{1}{2} - f_{e}(k - l, \delta)}$$
(A.18)

The network is unilaterally stable if neither of these three deviations is beneficial, i.e. if  $\alpha$  exceeds all values given in equations (A.15), (A.17) and (A.18):

$$\overline{\alpha} = \max\left\{ \begin{array}{l} \sqrt{c/(\frac{1}{2} - f_e(n-2,\delta))}, \frac{c - \frac{1}{2}(n-k)(n-k-1)f_m(k,\delta)}{\frac{1}{2} - f_e(k-1,\delta)}, \\ \max_{l \le k} \left( \frac{c - \frac{n-k-1}{l}(f_e(k,\delta) - f_e(k-l,\delta))}{\frac{1}{2} - f_e(k-l,\delta)} \right) \right\}$$

$$(A.19)$$

*Remarks on Proposition 3.* By rewriting equation (A.19) we get, given a sufficiently large level of heterogeneity  $\alpha > 1$ , the following condition for unilaterally stable core-periphery networks in terms of linking costs:

$$c \in \left(\frac{1}{2} - f_e(k,\delta) + (n-k-2)\min\{\frac{1}{2}f_m(k+1,\delta), f_e(k+1,\delta) - f_e(k,\delta)\}, \\ \min\left\{\alpha^2(\frac{1}{2} - f_e(n-2,\delta)), \alpha(\frac{1}{2} - f_e(k-1,\delta)) + \frac{1}{2}(n-k)(n-k-1)f_m(k,\delta), \\ \min_{l \le k}\left\{\alpha(\frac{1}{2} - f_e(k-l,\delta)) + \frac{n-k-1}{l}(f_e(k,\delta) - f_e(k-l,\delta))\right\}\right\}.$$

Notice that, for given *c* and  $\delta$  and for *n* sufficiently large, the level of heterogeneity  $\overline{\alpha}$  as given in (A.19) equals

$$\sqrt{c/\frac{1}{2} - f_e(k-1,\delta)}.$$
 (A.20)

For such a value of  $\alpha$ , the complete core-periphery network with *k* big banks in the core is unilaterally stable if condition (A.12) is satisfied. To minimize  $\overline{\alpha}$  we take the smallest vale of *c* satisfying (A.12). A lower bound for  $\overline{\alpha}$ , given values of  $\delta$ , *k* and sufficiently large *n*, is thus given by:

$$\overline{\alpha} \ge \overline{\alpha}_{min} \equiv \sqrt{\frac{\frac{1}{2} - f_e(k,\delta) + (n-k-2)\min\{\frac{1}{2}f_m(k+1,\delta), f_e(k+1,\delta) - f_e(k,\delta)\}}{\frac{1}{2} - f_e(k-1,\delta)}}.$$
(A.21)

Using the assumptions we made about the distribution of intermediated trades in Section 2.2, one can verify that

$$\lim_{k \to \infty} \overline{\alpha}_{min} = \lim_{\delta \downarrow 0} \overline{\alpha}_{min} = \lim_{\delta \uparrow 1} \overline{\alpha}_{min} = 1, \tag{A.22}$$

showing that an arbitrary small level of heterogeneity can be sufficient to have a unilaterally stable core-periphery network.

# Appendix B. Proofs of Section 4

**Proof of Theorem 1..** To proof this Theorem we start with two Lemma's, followed by the main part of the proof that makes use of these Lemma's. First, in an empty network, the best feasible action for a player is to connect either to all other players or to none, as shown in the following lemma.

**Lemma 1.** Consider the homogeneous baseline model with  $\alpha_i = 1$  for all  $i \in N$ . If the empty network  $g^e$  is not unilaterally stable, then the best feasible action for a player in  $g^e$  is to add links to all other nodes.

**Proof of Lemma 1..** The marginal benefits of adding *l* links to an empty network are:

$$M_i(g^e, +l) = l(\frac{1}{2} - c) + \binom{l}{2}\frac{1}{3}$$
(B.1)

As this function is convex in *l*, maximizing the marginal benefits over *l* results in either  $l^* = 0$  or  $l^* = n - 1$ .

Second, in a complete core-periphery network, if a peripheral agent has an incentive to add a link it is optimal to connect to all other players, thereby entering in the core himself.

**Lemma 2.** Consider the homogeneous baseline model with  $\alpha_i = 1$  for all  $i \in N$ , and suppose that the complete core-periphery network  $g_{com}^{CP(k)}$  (including the star network) with  $k \in \{1, 2, ..., n-2\}$  is not unilaterally stable because a peripheral player i can deviate by adding one or more links. Then the best feasible action for i in  $g_{com}^{CP(k)}$  is to add links to all l = n - k - 1 other periphery players.

**Proof of Lemma 2..** The marginal benefits of a periphery member for replicating the position of a core member of a complete core-periphery networks with *k* core members are:

$$M_{i}(g_{com}^{CP(k)}, +l) = l(\frac{1}{2} - f_{e}(k, \delta) - c) + \binom{l}{2} f_{m}(k+1, \delta)$$
(B.2)

As this function is convex in *l*, the maximum of this function is reached at  $l^* = 0$  or  $l^* = n - k - 1$ .

Now we have proven Lemma's 1 and 2, we continue the proof of Theorem 1. Part

I: By Lemma 1, after the move of player i = 1 the network is either empty or a star. In case it is empty the process converges to an empty network, as all other nodes face the same decision as i = 1. Player 1 has no incentive or possibility to form a star  $g^s$  with 1 the center if  $\pi_1(g^s) < 0$ , that is, if

$$c > \frac{1}{2} + \frac{1}{6}(n-2),$$

and part I directly follows. Part

II: If  

$$c < \frac{1}{2} + \frac{1}{6}(n-2),$$
(B.3)

then player 1 forms a star. Note that this is a feasible action, because  $\pi_1(g^s) > 0$  implies  $\pi_j(g^s) > 0$  for  $j \neq 1$ .

If player 2 does not want to add a subset of links, then the other players also do not want to, in which case the star is the outcome of the dynamic process (by part I, deleting her link is not profitable for player 2). If player 2 has an incentive to add links, then by Lemma 2 she will add all n - 2 links in order to form a complete core-periphery network  $g_{com}^{CP(2)}$  with 1 and 2 in the core. This happens if  $\pi_2(g_{com}^{CP(2)}) > \pi_2(g^s)$  and  $\pi_j(g_{com}^{CP(2)}) > \pi_j(g^s)$  for all j = 3, ..., n, that is,

$$c < \frac{1}{6} + \frac{1}{2}(n-3)f_m(2,\delta) \text{ and } c < \frac{1}{6} + (n-3)(f_e(2,\delta) - \frac{1}{3}).$$
 (B.4)

On the other hand, if either  $c > \frac{1}{6} + \frac{1}{2}(n-3)f_m(2,\delta)$  or  $c > \frac{1}{6} + (n-3)(f_e(2,\delta) - \frac{1}{3})$ , then neither player 2 nor any of the subsequent players change the network, and part II directly follows. Part

III: Under equation (B.4), after player 1 and 2 have chosen their links, a core-periphery network with 2 core players is formed. Player 3 does not delete any of her links under the second part of equations (B.3) and (B.4). By the same argument using Lemma 2, each next player  $i \in \{3, 4, ..., k\}$  adds links to all nodes not yet connected to i, as long as costs c are low enough. This might lead to a stable complete network, if the last two nodes have an incentive to form a link (which will be proposed by player i = n - 1), i.e.

$$c < \frac{1}{2} - f_e(n-2,\delta), \tag{B.5}$$

and part III directly follows.

For parameters not satisfying I, II or III, the first round of best feasible actions results in a complete core-periphery network with 1 < k < n - 1 core members. This complete core-periphery network is not stable by Lemma 1. Because the k + 1-th node did not connect to other periphery nodes, adding links in the periphery cannot be beneficial. Therefore the first core bank i = 1 must have an incentive to delete at least one within-core link. The marginal benefit of deleting *l* core links in the network is:

$$M_{i}(g_{com}^{CP(k)}, -l) = l(c - \frac{1}{2} + f_{m}(n - l - 1, \delta)).$$
(B.6)

As  $M_i(g_{com}^{CP(k)}, -0) = 0$  and  $M_i(g_{com}^{CP(k)}, -1) > 0$ , there is a unique choice  $l_1^* > 0$  of the optimal number of core links to delete. It will become clear that, for parameters outside  $I \cup II \cup III$ , attracting networks are multipartite networks of various sorts, depending on the choice  $l_1^*$ .

An important case is that a complete core-periphery network with k = 2 has arisen after the first round. For k = 2 the only possible solution is  $l_1^* = 1$ . A (complete) bipartite network arises with a small group of 2 players and a large group of n - 2 players, the maximal difference in group size possible for bipartite networks. We denote such a network as  $g_{2n-2}^{mp(2)}$ .

The case of k = 2 happens if the third player i = 3 does not enter the core, either because entering is not beneficial for himself or because some other periphery players j does not accept the offer of i. For k = 2 a positive marginal benefit of entering the core implies:

$$M_{i}(g_{com}^{CP(2)}, +(n-2)) > 0;$$
  
$$\Rightarrow c < \frac{1}{2} - f_{e}(2,\delta) + \frac{1}{2}(n-4)f_{m}(3,\delta),$$
(B.7)

and the periphery player *j* has an incentive to accept the offer if

$$M_{j}(g_{com}^{CP(2)}, +(n-2)) \ge 0;$$
  
$$\Rightarrow c \le \frac{1}{2} - f_{e}(2,\delta) + (n-4)(f_{e}(3,\delta) - f_{e}(2,\delta));$$
(B.8)

So if after the first round the result is k = 2 and the third player has not entered the core, it must be the case that:

$$c \ge \frac{1}{2} - f_e(2,\delta) + (n-4)\min\{\frac{1}{2}f_m(3,\delta), f_e(3,\delta) - f_e(2,\delta)\}.$$
(B.9)

The parameter values for which the maximally unbalanced network  $g_{2,n-2}^{mp(2)}$  is the attracting steady state are given by condition IV.

Alternatively, under the remaining condition V, the first round has resulted in a complete core-periphery network with k > 2. This network cannot be stable, and the first core bank i = 1 has an optimal choice of  $0 < l_1^* \le k - 1$  links to delete depending on the parameters. First consider that the core bank deletes all its within-core links, i.e.  $l_1^* = k - 1$ . This action necessarily implies that the linking costs exceed the loss in surplus associated with having an indirect connection to other core banks via the periphery rather than a direct connection:

$$c > \frac{1}{2} - f_e(n-k,\delta).$$
 (B.10)

Given this high level of linking costs, the next core banks  $i = \{2, ..., k\}$  have the same incentive to delete all their within-Given this high level of mixing costs, the next core banks - (z, ..., u) with two groups of size  $k_1 = k$  and  $k_2 = n - k$ .

Finally, consider the case  $0 < l_1^* < k - 1$ .<sup>29</sup> Denote  $K_1 \subset K$  as the set of core banks with at least one missing within-core link after the best feasible action of i = 1, including players i = 1 itself. The size of this set of banks is  $k_1 = l_1^* + 1$ . The best feasible action of i = 1 necessarily implies:

$$c > \frac{1}{2} - f_e(n - k_1, \delta).$$
 (B.11)

Given this high level of linking costs, all next banks  $i \in K_1$  with connections to some other  $j \in K_1$  with j > i have the same incentive to delete every link g<sub>ii</sub>. These banks do not have an incentive to remove any further link, because the number of intermediators for indirect connections to banks  $j \in K_1$  would become less than  $n - k_1$ . If the latter were worthwhile, i = 1would have chosen a higher number  $l_1^*$ . Therefore  $K_1$  becomes a group of players not connected within their group, but completely connected to all players outside their group.

Players  $i \in (K \setminus K_1)$  are connected to all other players  $j \in N$ . Because of the lower bound on c in equation (B.11), these players have an incentive to remove links to some j as long as the number of intermediators for the indirect connections to j stays above  $n - k_1$ . This will lead to other groups  $K_2, \ldots, K_{q-1}$  of players not connected within their group, but completely connected to all players outside their group. The remaining group  $K_q$  was the original periphery at the end of the first round. The result of the best-feasible-action dynamics if  $0 < l_1^* < k - 1$  is a multipartite network  $g_{k_1,k_2,...,k_q}^{mp(q)}$  with  $q \ge 3$ .

For all parameters under condition V, the resulting network is multipartite with  $q = \lfloor \frac{k}{k_1} \rfloor + 1$  groups. These multipartite

networks are more balanced than  $g_{2,n-2}^{mp(2)}$ , i.e. have group sizes  $|k_m - k_{m'}| < n-4$  for all  $m, m' \in \{1, 2, .., q\}$ .

#### References

Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A., 2015. Systemic risk and stability in financial networks. Am. Econ. Rev. 105 (2), 564-608. doi:10.1257/aer. 20130456

Afonso, G., Lagos, R., 2015. Trade dynamics in the market for federal funds. Econometrica 83 (1), 263-313. doi:10.3982/ECTA10586.

Allen, F., Gale, D., 2000. Financial contagion. J. Polit. Econ 108 (1), 1-33. doi:10.1086/262109.

Babus, A., Hu, T.-W., 2017. Endogenous intermediation in over-the-counter markets. J. Financ. Econ. 125 (1), 200-215. doi:10.1016/j.jfineco.2017.04.009.

Babus, A., Kondor, P., 2018. Trading and information diffusion in over-the-counter markets. Econometrica 86 (5), 1727–1769. doi:10.3982/ECTA12043.

Bedayo1, M., Mauleon, A., Vannetelbosch, V., 2016. Bargaining in endogenous trading networks. Math. Soc. Sci. 80, 70–82. doi:10.1016/j.mathsocsci.2016.02. 007.

Borgatti, S.P., Everett, M.G., 1999. Models of core/periphery structures. Soc. Networks 21 (4), 375-395. doi:10.1016/S0378-8733(99)00019-2.

Bräuning, F., Fecht, F., 2017. Relationship lending in the interbank market and the price of liquidity. Rev. Financ. 21 (1), 33-75. doi:10.1093/rof/rfw042. Buskens, V., Van de Rijt, A., 2008. Dynamics of networks if everyone strives for structural holes. Am. J. Sociol. 114 (2), 371-407. doi:10.1086/590674.

Castiglionesi, F., Navarro, N., 2019. (In)efficient interbank networks. J. Money Credit Bank. 52 (2-3), 365-407. doi:10.1111/jmcb.12664.

Chang, B., Zhang, S., 2019. Endogenous market making and network formation. Mimeo. University of Wisconsin-Madison and London School of Economics. Cocco, J.F., Gomes, F.J., Martins, N.C., 2009. Lending relationships in the interbank market. J. Finan. Intermediation 18 (1), 24-48. doi:10.1016/j.jfi.2008.06.003. Cohen-Cole, E., Patacchini, E., Zenou, Y., 2015. Static and dynamic networks in interbank markets. Network Science 3 (01), 98-123. doi:10.1017/nws.2015.1. Craig, B.R., Ma, Y., 2020. Intermediation in the interbank lending market. Working Paper 20-09. Federal Reserve Bank of Cleveland.

- Craig, B., Von Peter, G., 2014. Interbank tiering and money center banks. J. Finan. Intermediation 23 (3), 322-347. doi:10.1016/j.jfi.2014.02.003.
- Eisenberg, L., Noe, T.H., 2001. Systemic risk in financial systems. Manage. Sci. 47 (2), 236-249. doi:10.1287/mnsc.47.2.236.9835.

Elliott, M., Golub, B., Jackson, M.O., 2014. Financial networks and contagion. Am. Econ. Rev. 104 (10), 3115–3153. doi:10.1257/aer.104.10.3115.

Fafchamps, M., 2003. Market institutions in sub-Saharan Africa: Theory and evidence. MIT Press.

Fafchamps, M., Minten, B., 1999. Relationships and traders in Madagascar. J. Dev. Stud. 35 (6), 1–35. doi:10.1080/00220389908422600.

Farboodi, M., 2017. Intermediation and voluntary exposure to counterparty risk. Mimeo. Princeton University.

Gai, P., Haldane, A., Kapadia, S., 2011. Complexity, concentration and contagion. J. Monet. Econ. 58 (5), 453–470. doi:10.1016/j.jmoneco.2011.05.005.

Galeotti, A., Goyal, S., 2010. The law of the few. Am. Econ. Rev. 100 (4), 1468-1492. doi:10.1257/aer.100.4.1468.

Glasserman, P., Young, H.P., 2016. Contagion in financial networks. J. Econ. Lit. 54 (3), 779-831. doi:10.1257/jel.20151228.

Golub, B., Jackson, M.O., 2012. How homophily affects the speed of learning and best-response dynamics. Q. J. Econ. 1287-1338. doi:10.1093/qje/qjs021. Goyal, S., 2009. Connections: An introduction to the economics of networks. Princeton University Press.

Goyal, S., Vega-Redondo, F., 2007. Structural holes in social networks. J. Econ. Theory 137 (1), 460-492. doi:10.1016/j.jet.2007.01.006.

Hiller, T., 2017. Peer effects in endogenous networks. Games Econ. Behav. 105, 349-367. doi:10.1016/j.geb.2017.07.010.

Farboodi, M., Jarosch, G., Menzio, G., Wiriadinata, U., 2019. Intermediation as rent extraction. Mimeo. National Bureau of Economic Research. Farboodi, M., Jarosch, G., Shimer, R., 2020. The emergence of market structure. Mimeo. Massachusetts Institute of Technology.

Fricke, D., Lux, T., 2015. Core-periphery structure in the overnight money market: Evidence from the e-MID trading platform. Comput. Econ. 45 (3), 359-395. doi:10.1007/s10614-014-9427-x.

<sup>&</sup>lt;sup>29</sup> This is the only case in which the best feasible action is not unique in a parameter region with nonzero measure: the first core banks i = 1 has  $\binom{k}{l_{*}}$ best feasible actions deleting  $l_1^*$  links from the core. The resulting multipartite networks can consist of different sets  $K_1, K_2, \ldots, K_{q-1}$ , but all are isomorphic.

Houy, N., 2009. Structural holes in social networks: A remark. J. Econ. Theory 144 (1), 422-431. doi:10.1016/j.jet.2008.05.007.

- In 't Veld, D., Van der Leij, M., Hommes, C., 2019. The formation of a core periphery structure in heterogeneous financial networks. TI Discussion Paper 14–098/II, d.d. 2 May 2019. Tinbergen Institute.
- In 't Veld, D., Van Lelyveld, I., 2014. Finding the core: Network structure in interbank markets. J. Bank. Financ. 49, 27–40. doi:10.1016/j.jbankfin.2014.08.006. Jackson, M.O., 2005. A Survey of Network Formation Models: Stability and Efficiency. In: Demange, G., Wooders, M. (Eds.), Group Formation in Economics: Networks, Clubs and Coalitions. Cambridge University Press, Cambridge, UK, pp. 11–57.
- Jackson, M.O., Watts, A., 2002. The evolution of social and economic networks. J. Econ. Theory 106 (2), 265-295. doi:10.1006/jeth.2001.2903.
- Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. J. Econ. Theory 71 (1), 44-74. doi:10.1006/jeth.1996.0108.
- Kleinberg, J., Suri, S., Tardos, E., Wexler, T., 2008. Strategic network formation with structural holes. EC '08: P. 9th ACM Conference Electron. Commer. 480–489. doi:10.1145/1386790.1386835.
- König, M.D., Tessone, C.J., Zenou, Y., 2014. Nestedness in networks: A theoretical model and some applications. Theor. Econ. 9 (3), 695-752. doi:10.3982/ TE1348.
- Martínez-Jaramillo, S., Alexandrova-Kabadjova, B., Bravo-Benítez, B., Solórzano-Margain, J.P., 2014. An empirical study of the Mexican banking system's network and its implications for systemic risk. J. Econ. Dyn. Control 40, 242–265. doi:10.1016/j.jedc.2014.01.009.
- Nier, E., Yang, J., Yorulmazer, T., Alentorn, A., 2007. Network models and financial stability. J. Econ. Dyn. Control 31 (6), 2033-2060. doi:10.1016/j.jedc.2007. 01.014.
- Persitz, D., 2016. Core-periphery R&D collaboration networks. Technical Report. Tel Aviv University.
- Siedlarek, J.-P., 2015. Intermediation in networks. Working Paper 15-18. Federal Reserve Bank of Cleveland.
- Wang, C., 2018. Core-periphery trading networks. Mimeo. Wharton School, University of Pennsylvania.