

# The formation of a core-periphery structure in heterogeneous financial networks

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

<sup>a</sup>*CeNDEF, University of Amsterdam, Valckeniersstraat 65-67, 1018 XE Amsterdam, The Netherlands*

<sup>b</sup>*Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands*

<sup>c</sup>*Research Department, De Nederlandsche Bank, Westeinde 1, 1017 ZN Amsterdam, The Netherlands*

<sup>d</sup>*SEO Amsterdam Economics, Roetersstraat 29, 1018 WB Amsterdam, The Netherlands*

11 August 2016

---

## 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 simple model of the overnight interbank lending 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.

*Keywords:* financial networks, core-periphery structure, network formation models, over-the-counter markets, interbank market

*JEL classifications:* D85, G21, L14

---

## 1. Introduction

The extraordinary events of 2007 and 2008 in which the financial system almost experienced a global meltdown, led to an increased interest in the role of financial networks, the network

---

<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 workshops and conferences at the Banque de France (Paris), Bundesbank (Frankfurt), Castellón, CEMLA (Mexico City), EEA 2015 (Mannheim), Lorentz Center (Leiden), Norwich, NWO Complexity meeting, SAFE/Bundesbank (Frankfurt), Utrecht, and Tinbergen Institute (Amsterdam) and seminars at the CPB, De Nederlandsche Bank, GREQAM, Isaac Newton Institute (program Systemic Risk: Mathematical Modelling and Interdisciplinary Approaches), ITAM, London School of Economics, Stanford University, Tinbergen Institute, University of the Basque Country, and University of Siena. 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). Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

\*Corresponding author: Marco van der Leij, e-mail: M.J.vanderLeij@uva.nl

of trading relationships and exposures between financial institutions,<sup>1</sup> on systemic risk, the risk that liquidity or solvency problems in one financial institution spread to the whole financial 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>

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 maximise profits. This is important, as a change in the risk or regulatory environment may incentivise financial institutions to rearrange their financial linkages. 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 stylised 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, Peltonen et al. (2014) find such a structure for financial networks in derivative markets, Di Maggio et al. (2016) in the corporate bond markets, and Craig and Von Peter (2014) and in 't Veld and van Lelyveld (2014) 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 data better than 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 (2016), Farboodi (2015), Chang and Zhang (2016) and Bedayo et al. (2016). Differently from those papers, we aim to understand if such a core-periphery structure may arise *endogenously from ex-ante*

---

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

<sup>2</sup>See 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>3</sup>See Cohen-Cole et al. (2015) and König et al. (2014) for financial network models that have nested split graphs as outcome.

*identical banks.*

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 intermediary if the intermediary 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 networks are not stable in the framework of Goyal and Vega-Redondo (2007), as competition for intermediation benefits would drive core payoffs in such networks to zero.<sup>5</sup>

In order to allow for the possibility of non-trivial core-periphery networks, we extend the framework of Goyal and Vega-Redondo (2007). First, we propose a simple model of interbank overnight lending, in which banks receive positive or negative liquidity shocks on a daily basis, creating trading opportunities for banks on the interbank overnight lending market. We assume that trade can only take place between costly long-term trading relationships, allowing for the possibility of intermediation benefits. Intermediators compete for these benefits, but unlike Goyal and Vega-Redondo (2007) competition is *imperfect*, opening up the possibility that multiple core members 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. Apart from the usual equilibrium concept of pairwise stability (Jackson and Wolinsky, 1996), we also consider the stronger concept of *unilateral stability*. This concept allows for deviations of deleting or adding multiple links.

We ask ourselves if the core-periphery network is stable in this model. To our surprise, in general, the answer is no. We provide three results on that. First, we show that a complete core-periphery network, in which all possible links between a core and periphery are present, is not pairwise (let alone unilaterally) stable. The intuition behind this result is as follows; a stable complete core-periphery network implies that periphery banks prefer to trade indirectly via intermediating core banks, rather than trade directly. However,

---

<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.

<sup>5</sup>Babus and Hu (2016) consider a financial network formation model that also builds on Goyal and Vega-Redondo (2007). In their model an interlinked star network with 2 members in the core may be stable. However, 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 (2016).

given that periphery and core banks have identical technologies, core banks should have an incentive to trade indirectly with peripheral banks as intermediators as well. Hence, core members do not have an incentive to maintain direct trading relationships with each other, in contradiction to the definition of a core-periphery network.

Our second result is 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 unsustainable 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.

The first two results, still leaves open the possibility for unilaterally stable (incomplete) core-periphery networks, and indeed, in Appendix F we do give examples of such stable core-periphery networks. The third result, however, shows that, in a dynamic setting, such stable core-periphery networks are unlikely to arise. In particular, we show that, in a simple dynamic model a la Kleinberg et al. (2008), best-response dynamics never converge to a core-periphery network. 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, and may exhibit substantial inequality. However, unlike core-periphery networks, they do not have links within a (core) tier.<sup>6</sup>

Key to our finding is that banks are ex ante identical; periphery banks obtain the same trading opportunities as core banks and have the same technology to intermediate between other banks. This excludes the stability of a complete core-periphery network, and puts a limit on inequality. 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 more frequent trading opportunities. We find 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.

Finally, we show that this heterogeneity between banks, and in fact 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

---

<sup>6</sup>We also characterise efficient networks. These are the empty network or the star network. The core-periphery network and (stable) multipartite network are never efficient networks.

bipartite 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, become relatively bigger and bigger, until it finally becomes attractive for the big banks to trade directly, forming a core-periphery network structure.

We now review the literature on financial network formation. One may make a distinction between papers that are more concerned with the trade offs between contagion, risk sharing, efficiency and stability (Cabrales et al., 2013; Acemoglu et al., 2014; Babus, 2016), and papers that (among other things) rationalise the formation of a core-periphery structure in financial networks (Farboodi, 2015; Bedayo et al., 2016; Castiglionesi and Navarro, 2016; Chang and Zhang, 2016; Wang, 2016). Our paper belongs to the second category. It is interesting to note the role of heterogeneity in these papers. In papers of the first category (Cabrales et al., 2013; Acemoglu et al., 2014) ex-ante homogeneity is assumed, and indeed, the resulting efficient or stable networks do not correspond to core-periphery networks at all.<sup>7</sup> On the other hand, in papers of the second category heterogeneity plays a key role. In Farboodi (2015) 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 Bedayo et al. (2016) intermediation also plays a key role, with intermediaries bargaining sequentially and bilaterally on a path between two traders. Here agents are heterogeneous in their time discounting. They find that a core-periphery network is formed with impatient agents in the core. In Castiglionesi and Navarro (2016) 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 core-periphery like structure.<sup>8,9</sup>

---

<sup>7</sup>In Babus (2016) it is assumed that there are 2 regions with negatively correlated shocks, such that a bipartite network arises.

<sup>8</sup>In the model of Castiglionesi and Navarro (2016) periphery bank 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.

<sup>9</sup>In Chang and Zhang (2016) banks are heterogeneous in the volatility of their liquidity needs. More volatile banks trade with more stable banks, creating a multi-tier financial system with the most stable banks in the core. However, banks do not have incentives to link with other banks in the same tier, and hence, their network structure is more like a multipartite network than a core-periphery network. Wang (2016) has a model with 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 as well as bipartite networks can be equilibrium outcomes. There are other papers in the social science literature that 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 (2015) provide conditions under which core-periphery networks are the only stable network structure. See also Persitz (2012) who adopts 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.

Overall, the main message of our paper is that bank heterogeneity matters for the formation of financial core-periphery networks, and that, in order to understand the financial system and its (systemic) risks, it is crucial to understand which types of heterogeneity and which mechanisms are driving the core-periphery network structure. This is particularly relevant, because inefficiency results tend to depend on the particular type of heterogeneity.

This paper is organised as follows. 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 for homogeneous traders (Section 3) and heterogeneous traders (Section 4). In Section 5 we provide an application of our model to the interbank market of the Netherlands. Section 6 concludes.

## 2. Model

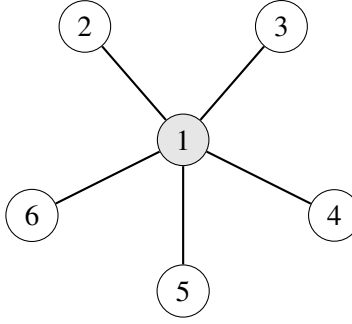
Our goal is to model the formation of a network of *long-term* trading relationships between a set of banks, denoted by  $N$ , having cardinality  $n = |N|$ . There are two stages. In the first stage, at time  $t = 0$ , banks form an undirected network,  $g$ , of these trading relationships. Denote by  $g_{ij} = g_{ji} = 1$  the existence of a trading relationship, and by  $g_{ij} = g_{ji} = 0$  the absence of it. After forming their long-term trading relationships, liquidity trade takes place through these relationships in an infinite number of periods,  $t = 1, 2, \dots$ . Payoffs from forming trading relationships at time 0 depend on expected present value trade benefits from these liquidity trades 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.  $\forall i, j \in N : g_{ij} = 0$ , 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*,  $g^s$  (see Figure 1) 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  $\forall j, k \neq i : g_{jk} = 0$ .

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. This definition is taken from 't Veld and van Lelyveld (2014), adapted to undirected networks, and follows the definitions of Borgatti and Everett (1999) and Craig and Von Peter (2014).

**Definition 1.** *A network  $g$  is a core-periphery network, if there exists a set of core agents*



**Figure 1:** A star network with  $n = 6$  players

$K \subset N$  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  $g_{ij} = 1$ .

See Figure 2 for an example. 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  $g_{ij} = 1$ .<sup>10</sup> See Figure 3 for an example. We denote a complete core-periphery network with  $k = |K|$  agents in the core as  $g_{com}^{CP(k)}$ .<sup>11</sup>

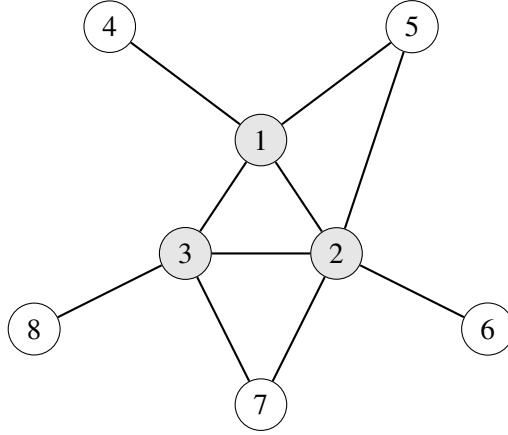
Finally, a *complete multipartite network* is a network, in which the agents can be partitioned into  $q$  groups,  $N = \{K_1, K_2, \dots, K_q\}$ , such that nodes do not have links within their group, but are completely connected to all nodes outside their own group. Formally, in a complete multipartite network it holds that  $\forall m \in \{1, 2, \dots, q\} : \forall i \in K_m$  we have  $\forall j \in K_m : g_{ij} = 0$  and  $\forall j \notin K_m : g_{ij} = 1$ . Complete multipartite networks will be denoted as  $g_{k_1, k_2, \dots, k_q}^{mp(q)}$ , 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 equal as possible, i.e.  $|k_m - k_{m'}| \leq 1$  for all  $m, m'$ . Figure 7 in Section 3.2 presents examples of complete multipartite networks that arise in our model.

## 2.2. Trading Benefits

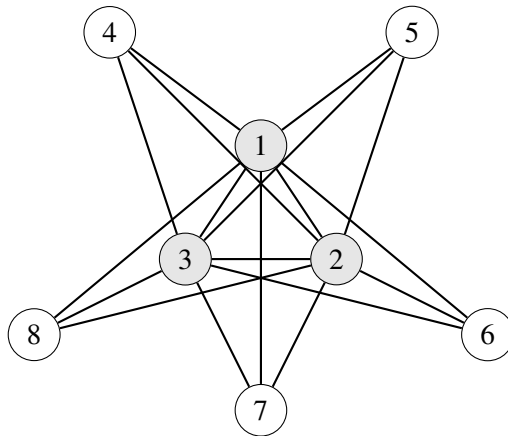
We now turn to the actual model. After forming a network of long-term trading relationships, liquidity trade takes place through these relationships in an infinite number of

<sup>10</sup>Borgatti and Everett (1999) call this architecture a perfect core-periphery network.

<sup>11</sup>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  $k \in \{2, 3, \dots, n - 2\}$ .



**Figure 2:** A core-periphery network with  $n = 8$  players, of which  $k = 3$  are in the core



**Figure 3:** A complete core-periphery network with  $n = 8$  players, of which  $k = 3$  are in the core



periods,  $t = 1, 2, \dots$ . We first discuss the payoffs in a single trading period. We assume that trades cannot take place outside the long-term trading relationships. Hence, in each period a (directed) trading network arises which is a subnetwork of the (undirected) relationship network  $g$ . Trade in each period arises as follows. At the beginning of period  $t \in \{1, 2, \dots\}$  each bank  $i$  receives a random unexpected liquidity shock  $S_{it} \in \{-s, 0, s\}$ , independently distributed with probabilities  $P[S_{it} = -s] = P[S_{it} = s] = \gamma_i$  and  $P[S_{it} = 0] = 1 - 2\gamma_i$ . Without loss of generality, we set  $s = 1/2$ . This liquidity shock has a temporary effect on the liquidity position of the bank, that is, in that period  $t$  it may have a liquidity surplus or deficit relative to a liquidity target, set by the Central Bank or by internal risk management procedures, but at the beginning of the next period, each bank is again in a neutral position. The effect of liquidity shocks are thus temporary, as in the intermediate run banks are able to manage their liquidity position by increasing or decreasing their asset size, for example by extending more or less loans. Hence, liquidity positions do not carry over to the next period, and this excludes the possibility of speculation or storage of liquidity.<sup>12</sup>

If bank  $i$  has a positive liquidity shock, it may set aside the resulting liquidity surplus at the Central Bank, receiving interest rate  $\underline{r}$ . On the other hand, if, say, bank  $j$  receives a negative shock, it may borrow liquidity from the Central Bank at cost  $\bar{r}$ . Assuming that there is an interest rate wedge  $\bar{r} - \underline{r} > 0$ , trading opportunities arise between banks that receive a positive liquidity shock and banks that receive a negative liquidity shock. For example, if  $S_{it} = 1/2$  and  $S_{jt} = -1/2$ , then bank  $j$  could borrow from bank  $i$  at interest rate  $r_{ij} \in [\underline{r}, \bar{r}]$ . Relative to the outside option of borrowing or depositing at the Central Bank, bank  $i$  would gain  $r_{ij} - \underline{r}$  from the trade, and bank  $j$  would gain  $\bar{r} - r_{ij}$ . Without loss of generality, we set  $\bar{r} - \underline{r} = 1$ .

In general, in each period there may be multiple banks with a positive and multiple banks with a negative liquidity shock, leading to a myriad of potential trades. In order to keep the analysis tractable, we analyze the model in case the only relevant trading opportunities arise from one (and only one) bank, say  $i$ , receiving a positive shock and one (and only one) other bank, say  $j$ , receiving a negative shock. This situation occurs when the probability of a liquidity shock becomes small, and is formerly worked out in Proposition 1 of Section 2.3.

Now, suppose that bank  $i$  receives a positive shock,  $S_{it} = 1/2$ , bank  $j$  receives a negative shock,  $S_{jt} = -1/2$ , and all other banks receive no shock,  $S_{kt} = 0$  for all  $k \neq i, j$ . We assume that a liquidity trade between  $i$  and  $j$  can only be realised if  $i$  and  $j$  have a direct (long-term) trading relationship,  $g_{ij} = 1$ , or an indirect trading relationship through one or more

---

<sup>12</sup>Our model differs from the search and matching models of over-the-counter markets, as introduced by Duffie et al. (2005), in which trading positions remain until a suitable trading partner is found.

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 realised, we require that these middlemen are directly connected to  $i$  and  $j$ . In network parlance, it implies that trade between  $i$  and  $j$  can only be realised if the network distance, that is the shortest path length, between  $i$  and  $j$  in  $g$  is at most 2. We make this assumption to simplify notation. However, we emphasise that our main results – core-periphery networks are generally unstable under homogeneity, but can be stable if agents are heterogeneous – do *not* depend on this assumption; the proposition also holds if  $i$  and  $j$  are connected at distance more than 2. This is explained further when discussing the particular propositions and theorems.<sup>13</sup>

If the network distance between  $i$  and  $j$  is at most 2, trade is realised, and agents divide the surplus of 1 in the following way. Let  $v_k^{ij}(g, \delta)$  be the share that  $k$  receives if  $i$  receives a positive shock and  $j$  a negative shock. Apart from the network structure  $g$ , we let the share that each agent obtains depend 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  $\delta = 0$ , then middlemen collude and act as if there were only a single middleman between  $i$  and  $j$ . If  $\delta = 1$ , then middlemen engage in Bertrand competition whenever  $m_{ij} > 1$ , such that  $i$  and  $j$  are able to retain all the surplus.

More concretely, if  $i$  and  $j$  are directly connected, then  $v_i^{ij}(g, \delta) = v_j^{ij}(g, \delta) = 1/2$ , that is, each receive half of the surplus. If  $i$  and  $j$  are indirectly connected in  $g$  by  $m_{ij}$  middlemen, then the endnodes  $i$  and  $j$  each receive a share of  $v_i^{ij}(g, \delta) = v_j^{ij}(g, \delta) = f_e(m_{ij}, \delta)$ . Each of the  $m_{ij}$  middlemen receives a share of  $v_k^{ij}(g, \delta) = f_m(m_{ij}, \delta)$  if  $k \in M_{ij}(g)$  and 0 otherwise. Note that by definition:

$$2f_e(m_{ij}, \delta) + m_{ij}f_m(m_{ij}, \delta) = 1.$$

If there is one middleman,  $m_{ij} = 1$ , then  $i$ ,  $j$ , and the middleman split the surplus evenly, hence,

$$f_e(1, \delta) = f_m(1, \delta) = \frac{1}{3}.$$

If there is more than only one middleman, then the share that the agents get depends on the amount of competition. If  $\delta = 0$ , the middlemen collude, and  $i$  and  $j$  obtain the same share as if there were one intermediary. The middlemen share the intermediation benefits

---

<sup>13</sup>The same assumption was made by Kleinberg et al. (2008) and Siedlarek (2013).

evenly, that is, for all  $m_{ij} \in \{2, \dots, n-2\}$ :

$$f_e(m_{ij}, 0) = \frac{1}{3} \text{ and } f_m(m_{ij}, 0) = \frac{1}{3m_{ij}}.$$

If  $\delta = 1$ , between  $m_{ij} > 1$  intermediaries, their payoffs disappear, hence

$$f_e(m_{ij}, 1) = \frac{1}{2} \text{ and } f_m(m_{ij}, 1) = 0.$$

The same happens in the limit of the number of middlemen to infinity. A further straightforward assumption is monotonicity with respect to the competitiveness and number of middlemen.

Share of payoffs for:	endnodes	middlemen
Level of competition $\delta$	$f_e(\cdot, 0) = \frac{1}{3}$ $\frac{\partial f_e}{\partial \delta} > 0$ $f_e(\cdot, 1) = \frac{1}{2}$	$f_m(\cdot, 0) = \frac{1}{3m}$ $\frac{\partial f_m}{\partial \delta} < 0$ $f_m(\cdot, 1) = 0$
Number of middlemen $m$	$f_e(1, \cdot) = \frac{1}{3}$ $\frac{\partial f_e}{\partial m} > 0$ $\lim_{m \rightarrow \infty} f_e(m, \cdot) = \frac{1}{2}$	$f_m(1, \cdot) = \frac{1}{3}$ $\frac{\partial f_m}{\partial m} < 0$ $\lim_{m \rightarrow \infty} f_m(m, \cdot) = 0$

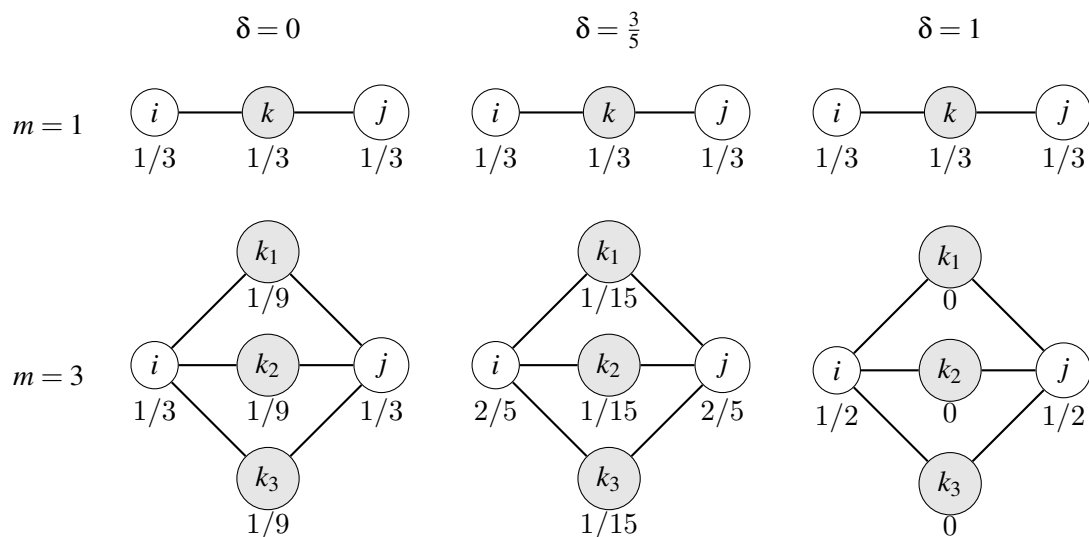
**Table 1:** Assumptions about the payoff shares  $f_e(m_{ij}, \delta)$  and  $f_m(m_{ij}, \delta)$ .

Table 1 summarises the assumed dependencies of the surplus distributions  $f_e(\cdot)$  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 end nodes (that is, the banks with the positive and negative liquidity shock) increases with  $\delta$  and intermediation benefits decrease. We do not model this bargaining process explicitly. However, our assumptions on  $f_e(\cdot)$  and  $f_m(\cdot)$  generalise an explicit Rubinstein-type of bargaining process developed by Siedlarek (2015). From his bargaining protocol, in which  $\delta$  has the interpretation of a discount factor, the distribution of the surplus is given by

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

It is easily checked that (1) satisfies the assumptions we made on  $f_e(\cdot)$  and  $f_m(\cdot)$ . Below we will use this explicit function to illustrate our (more general) results in Example 1 and Figures 6, 9 and 11.

Figure 4 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  $m_{i,j}$ . In Appendix A, we give a detailed explanation of the specification of payoffs by Siedlarek (2015).



**Figure 4:** Examples of payoff shares received by endnodes  $i$  and  $j$  and intermediaries  $k$ , depending on the parameters  $\delta$  and  $m$  under the specification of  $f_e$  and  $f_m$  in equation (1), as in Siedlarek (2015).

### 2.3. Network payoffs

We now turn to the payoff from network formation in period  $t = 0$ . In period zero, payoffs from network formation are determined by the expected net present value of payoffs from trades in subsequent periods minus the net present value of the costs of maintaining trading relationships.

Let  $S_t = \{S_{it}\}_{i \in N}$  and shock  $s_t \in \{-1/2, 0, 1/2\}^n$  a realization of  $S_t$ . Let  $\beta \in [0, 1)$  be the discount factor. Let  $b(s_t, g, \delta) \in [0, 1]^n$  the vector of trade benefits of the  $n$  agents given network  $g \in G$ , and shock realization  $s_t$ . Let  $\tilde{c}$  be the cost of maintaining a link for one period. Then the payoff function of bank  $i$  is given by

$$\tilde{\pi}_i(g, \delta, \tilde{c}) = \sum_{t=1}^{\infty} \beta^t (E[b_i(S_t, g, \delta)] - \eta_i(g)\tilde{c}),$$

where  $\eta_i(g) = |N_i^1(g)|$  the number of links of  $i$  in  $g$ .

We write the probability of a liquidity shock as  $\gamma_i = \rho\alpha_i$ .  $\rho$  is a scaling factor, and  $\alpha_i$  captures potential heterogeneity among banks. We assume that bigger banks (in terms of asset size) have higher  $\alpha_i$ , that is, they receive more trade opportunities. In the homogeneous case,  $\alpha_i = 1$  for all  $i$ . If we let  $\rho$  become small, then the payoffs can be approximated as follows:

**Proposition 1.** *Let  $N_i^d(g)$  denote the set of nodes at distance  $d$  from  $i$  in network  $g$ . The quadratic approximation of firm  $i$ 's payoff function around  $\rho = 0$  is given by*

$$\tilde{\pi}_i(g, \delta, c) = \frac{\beta\rho^2}{1-\beta}\pi_i\left(g, \delta, \frac{\tilde{c}}{\rho^2}\right) + O(\rho^3),$$

where the function  $\pi_i(g, \delta, c)$  is given by

$$\pi_i(g, \delta, c) = \underbrace{\sum_{j \in N_i^1(g)} \left(\frac{1}{2}\alpha_i\alpha_j - c\right)}_{\text{direct trade}} + \underbrace{\sum_{j \in N_i^2(g)} \alpha_i\alpha_j f_e(m_{ij}(g), \delta)}_{\text{indirect trade}} + \underbrace{\sum_{k, l \in N_i^1(g) | g_{kl}=0} \alpha_k\alpha_l f_m(m_{kl}(g), \delta)}_{\text{intermediation benefits}} \quad (2)$$

*Proof.* See Appendix B. □

Proposition 1 says that, for  $\rho$  small, the payoff function can be approximated by

$$\tilde{\pi}_i(g, \delta, \tilde{c}) \approx \frac{\beta\rho^2}{1-\beta}\pi_i\left(g, \delta, \frac{\tilde{c}}{\rho^2}\right),$$

in which the only trade opportunities that matter are those in which one and only one bank has a positive liquidity shock and one and only one bank a negative liquidity shock. The benefits from these trade opportunities 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. Trade opportunities arising from multiple banks receiving a shock, occur with infinitesimal probability. In the rest of the paper we therefore consider the function  $\pi_i(g, \delta, c)$  as in (2) with  $c \equiv \frac{\tilde{c}}{\rho^2}$ , as the payoff function from network formation. The payoff function of Goyal and Vega-Redondo (2007) is a special case with homogeneity ( $\alpha_i = 1 \forall i$ ) and perfect competition ( $\delta = 1$ ).

We now comment on the interpretation of these payoffs. First, the undirected links in this model should be interpreted as established preferential trading relationships. We assume that trade opportunities can only be realised if the agents are linked directly or indirectly through intermediation by mutual trading relationships. This is a strong, but not implausible assumption. The existence of preferential trading relationships has been shown by Cocco et al. (2009) in the Portuguese interbank market and by Bräuning and Fecht (2012) 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 liquidity funding and, as a consequence, may face higher borrowing costs. Hence, banks have incentives to use their established trading relationships.

Second, it is assumed that the preferential trading relationship comes at a fixed cost of  $\tilde{c}$ . 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.

#### **2.4. Network stability concepts**

Given the setup of the payoffs discussed above, we analyze which networks arise if agents form links strategically at time  $t = 0$ . 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>14</sup> Network for-

---

<sup>14</sup>See Goyal (2009) for a text book discussion.

mation 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* (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  $g - g_{ij}$  as the network identical to  $g$  except that the link between  $i$  and  $j$  is removed. Then the definition of pairwise stability is as follows (dropping the argument of  $\delta$  and  $c$  in the function  $\pi(\cdot)$ ):

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

- (a) if  $g_{ij} = 1$ , then  $\pi_i(g) \geq \pi_i(g - g_{ij}) \wedge \pi_j(g) \geq \pi_j(g - g_{ij})$ ;
- (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).

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  $g_{ij}^{iS} = 1 - g_{ij}$ , 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  $S \subseteq N \setminus \{i\}$ :

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

Note that 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>15</sup>

We illustrate the difference between pairwise stability and unilateral stability with the following example.

**Example 1.** Consider a star network  $g^s$  with  $n = 6$  banks with bank 1 in the center. Suppose  $\delta = 1/2$  and  $c = 1/5$ , and suppose surplus division functions  $f_e(\cdot)$  and  $f_m(\cdot)$  are given by (1). Then in a star network, the center earns  $\pi_1(g^s, 1/2, 1/5) = 29/6 \approx 4.833$ , and the periphery earns  $\pi_i(g^s, 1/2, 1/5) = 49/30 \approx 1.633$ . See Figure 5, left panel.

Removing one or multiple links would obviously hurt the players involved.<sup>16</sup> Moreover, adding a link between two peripheral banks, say between 2 and 3, does not benefit them, as in that case  $\pi_2(g^s + g_{23}, 1/2, 1/5) = \pi_3(g^s + g_{23}, 1/2, 1/5) = 6/5 = 1.6 < \pi_2(g^s, 1/2, 1/5)$ . See Figure 5, middle panel. Hence, the star network is pairwise stable in this case.

However, the star network is not unilaterally stable. Each peripheral bank, say bank 2, would benefit from creating relationships with all other banks at once, such that a complete core-periphery network with 2 banks in the core is formed. See Figure 5, right panel. In that case, player 2 receives additional benefits from intermediation, and obtains a payoff of, in total,  $\pi_2(g_{com}^{CP(2)}, 1/2, 1/5) = 9/4 = 2.25 > \pi_2(g^s, 1/2, 1/5)$ . Moreover, also the peripheral players benefit from this new core member, as there are now *two* members that compete for intermediation. Hence intermediation fees go down, and periphery payoffs go up, becoming  $\pi_i(g_{com}^{CP(2)}, 1/2, 1/5) = 69/40 = 1.725 > \pi_i(g^s, 1/2, 1/5)$ . Therefore, becoming a second core member is a feasible deviation for bank 2.

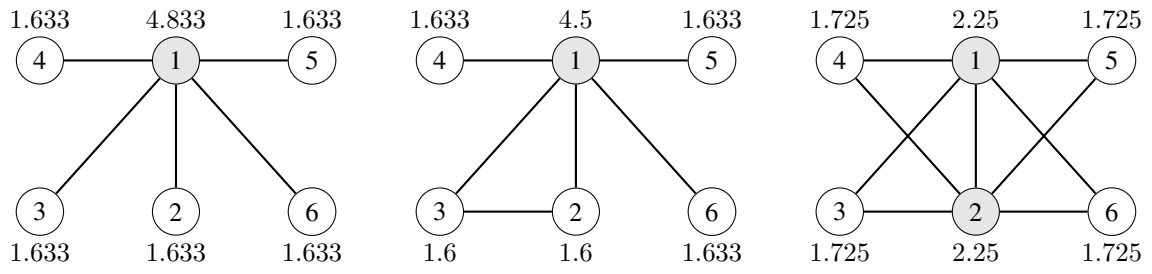
Note that, in order to make it feasible to deviate and become a second core member, it is crucial that  $\delta$  has an intermediate value of around  $\delta = 1/2$ . If  $\delta$  is too low, then becoming a second core member is not attractive, as intermediation benefits are too low. If  $\delta$  is too high, then peripheral players do not benefit from increased competition among middlemen.

<sup>15</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>16</sup>In a star network, the payoffs of the center are given by  $\pi_c(g^s, \delta, c) = (n-1) \left( \frac{1}{2} - c \right) + \binom{n-1}{2} \frac{1}{3}$ , such that the center always prefers a bigger star network, as long as  $c \leq \frac{1}{2}$ .



Hence,  $\delta$  must be in-between.



**Figure 5:** Examples of networks with  $n = 6$  players, with the network payoffs for  $\delta = 1/2$  and  $c = 1/5$ , under the specification of  $f_e$  and  $f_m$  in equation (1).

Apart from analyzing the stability of networks, for policy considerations it is also relevant to consider efficient networks. As usual in the literature, we define a network efficient, if it maximises the total sum of payoffs of the agents.

**Definition 4.** *A network  $g$  is efficient if there is no other network  $g'$ , such that*

$$\sum_{i \in N} \pi_i(g', \delta, c) > \sum_{i \in N} \pi_i(g, \delta, c).$$

### 3. Results for 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>17</sup> We consider the two stability concepts described above, pairwise and unilateral stability. Section 3.1 contains our main result that core-periphery networks are not unilaterally stable under the assumption of homogeneity. To find what structures arise in this homogeneous case, if not core-periphery networks, we investigate in Section 3.2 a best-response dynamic process. In Section 3.3 we relate these outcomes to the efficient networks.

#### 3.1. Stability of basic structures

To gain understanding of our trading model, we start by analyzing the stability of the empty, star and complete networks. Proposition 2 summarises the stability conditions for empty, star and complete networks under homogeneous agents; see Appendix C for the theoretical results underlying this proposition.

**Proposition 2.** *In the homogeneous baseline model with  $\alpha_i = 1$  for all  $i \in N$  the following holds:*

I: *The empty network is unilaterally stable if and only if  $c \geq \frac{1}{2} + \frac{1}{6}(n-2)$ .*

II: *The star network is unilaterally stable if and only if*

$$c \in \left[ \frac{1}{6} + (n-3) \min\left\{ \frac{1}{2} f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\}, \frac{1}{2} + \frac{1}{6}(n-2) \right]$$

III: *The complete network is unilaterally stable if and only if  $c \leq \frac{1}{2} - f_e(n-2, \delta)$ .*

---

<sup>17</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 Proposition 1 are proportional to  $\alpha$  when costs  $c$  are considered as a fraction of  $\alpha$ .

*Proof.* See Appendix D. □

Figure 6 in Section 3.2 depicts the stability regions I, II and III of the empty, star and complete networks in the  $(\delta, c)$ -space for  $n = 4$  and  $n = 8$ . The areas I and III are economically intuitive. When linking costs are (very) high, it is not beneficial to form links with other players, and the empty network is stable. When linking are (very) close to 0, linking to all other players is beneficial, and the complete network is stable.

We now consider the stability of core-periphery networks, starting with complete core-periphery networks. In a complete core-periphery network, in which  $k$  core members are connected to all other nodes, the profit function in Proposition 1 simplifies to:

$$\pi_i(g_{com}^{CP(k)}) = \begin{cases} (n-1)(\frac{1}{2}-c) + \binom{n-k}{2}f_m(k, \delta) & \text{if } i \in C \\ k(\frac{1}{2}-c) + (n-k-1)f_e(k, \delta) & \text{if } i \in P \end{cases} \quad (3)$$

The reason we focus on this special type of complete core-periphery networks is that they are formed naturally as peripheral agents try to improve their position by avoiding payments to core players and linking to other peripheral players.

We first show that, in the homogenous case, complete core-periphery networks can never be stable.

**Proposition 3.** *Let  $\alpha_i = 1$  for all  $i \in N$ , and let  $c > 0$  and  $0 < \delta < 1$  be given. Then complete core-periphery networks with  $k : 2 \leq k < n - 2$  core members are not pairwise stable, and thus also not unilaterally stable.*

*Proof.* See Appendix D. □

This result does not depend on the assumption that trade surpluses between  $i$  and  $j$  are only realised if the path length between  $i$  and  $j$  is less than 3, as shown in Appendix D.

The intuition behind this result is as follows. Suppose that a complete core-periphery network is pairwise stable. This implies that two periphery banks do *not* have an incentive to trade directly with each other, and at the same time, two core banks *do* have an incentive to trade directly. This constitutes a contradiction, because, in a core-periphery network, there are  $n - 2$  potential intermediators between two core banks versus  $k < n - 2$  potential intermediators between two periphery banks. After all, all periphery banks may act as an intermediary between core banks. Given that intermediation costs decrease with the number of intermediators, if two periphery banks trade indirectly, then two core banks have

an incentive to do the same. Hence, core banks should delete their core links, contradicting the pairwise stability of core-periphery networks.

Crucial in the proof is that the core-periphery network is (near) complete, and that, *ceteris paribus*, periphery banks have the same intermediation capabilities as core banks.<sup>18</sup> We next show that incomplete core-periphery networks too are unstable (in the sense of unilateral stability), as long as the network is large enough.

**Proposition 4.** *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 incomplete core-periphery networks with  $k \geq 1$  core member(s) are not unilaterally stable if  $n$  is sufficiently large. More precisely, there is a function  $F(c, \delta, k)$  such that if  $n > F(c, \delta, k)$ , the result holds.*

*Proof.* See Appendix D. □

We again emphasise that this result does not depend on the assumption that trade surpluses between  $i$  and  $j$  are only realised if the path length between  $i$  and  $j$  is less than 3, as shown in the proof.

The intuition for Proposition 4 is that, for  $n$  large enough, inequality in payoffs between core and periphery banks becomes so large that periphery banks have an incentive to replicate the linking strategies of core banks and enter the core. Consider for example the payoffs for complete core-periphery networks in equation (3). As the network size  $n$  increases for a fixed core size  $k$ , core members get more and more intermediation benefits (at a quadratic rate), while periphery payoffs fall behind (as it grows at a linear rate).

We now compare this result with that of Goyal and Vega-Redondo (2007). They show that for  $c > \frac{1}{6}$ ,  $\delta = 1$  and  $n$  sufficiently large, the star network (i.e. a core-periphery network with  $k = 1$ ) is the unique non-empty *bilaterally* stable network, stressing the importance of the star structure. Proposition 4 refines this result and indicates that for  $\delta < 1$  core-periphery networks with  $k$  core players *including* star networks with  $k = 1$  are *not unilaterally* stable, even for values of  $\delta$  arbitrary close but below 1.

We argue that the crucial assumption for our result is imperfect competition, that is,  $\delta < 1$ . Imperfect competition creates profitable deviations for peripheral players to circumvent (large) intermediation fees to core players and to start receiving intermediation benefits. A further important difference in this paper compared to Goyal and Vega-Redondo (2007)

---

<sup>18</sup>It is sufficient for the proof that each pair of core banks is connected by more intermediators than each pair of periphery banks.

is the stability concept: in the proof of Proposition 4, we make explicit use of the fact that periphery players may add multiple links at the same time, which is not allowed in Goyal and Vega-Redondo (2007). Because we allow for unilateral deviations, peripheral players can replicate the position of core players in order to benefit from intermediation. Inequality between core and periphery players makes the core-periphery networks unstable for sufficiently large  $n$ .

### 3.2. A dynamic process

So far, we analyzed the unilateral stability of empty, star, complete and core-periphery architectures. Core periphery networks were shown to be generally unstable under homogeneous agents, as long as the network is sufficiently large. The question remains what happens in small-sized networks, and, if not core-periphery networks, what other kind of network architectures do arise? This motivates us to consider a simple dynamic process of network formation and analyze its stable states.

In particular, we consider a round-robin best-response-like dynamic process as in Kleinberg et al. (2008). We order nodes  $1, 2, \dots, n$  and starting from the empty graph, in this order 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 leads to the highest payoff for  $i$ . The formal definition of a (best) feasible action is given in Appendix C. Note that a unilaterally stable network is a network in which all players chooses a best feasible action.

After node  $n$  has chosen its best feasible action the second round starts again with node 1, again each player consecutively choosing its best feasible action. Next, the third round start, 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 characterization 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>19</sup> The outcome of the dynamic process must be a unilaterally stable network. We find that the dynamics

---

<sup>19</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.

converge to a unique unilaterally stable network for every choice of the model parameters (except for borderline cases).<sup>20</sup>

Theorem 1 below shows which network structures result from the dynamics. The parameter regions for which empty, star and complete networks are the attracting steady states coincide with conditions I, II and III in Proposition 2 for which these networks are stable. In the remaining parameter regions the attracting steady state cannot be a core-periphery network (in line with Propositions 3 and 4) 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 the following unilaterally stable equilibria:*

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

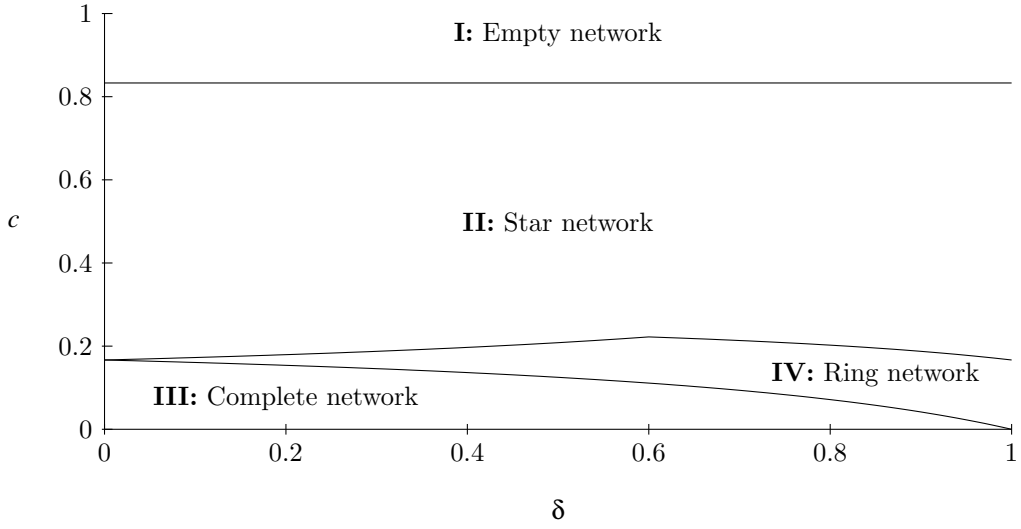
V: for  $c \in \left( \frac{1}{2} - f_e(n-2, \delta), \frac{1}{2} - f_e(2, \delta) + (n-4) \min\{\frac{1}{2}f_m(3, \delta), f_e(3, \delta) - f_e(2, \delta)\} \right)$   
a multipartite network  $g_{k_1, k_2, \dots, k_q}^{mp(q)}$  with  $q \geq 2$  and  $|k_m - k_{m'}| < n-4$  for all  $m, m' \in \{1, 2, \dots, q\}$ .

*Proof.* See Appendix D. □

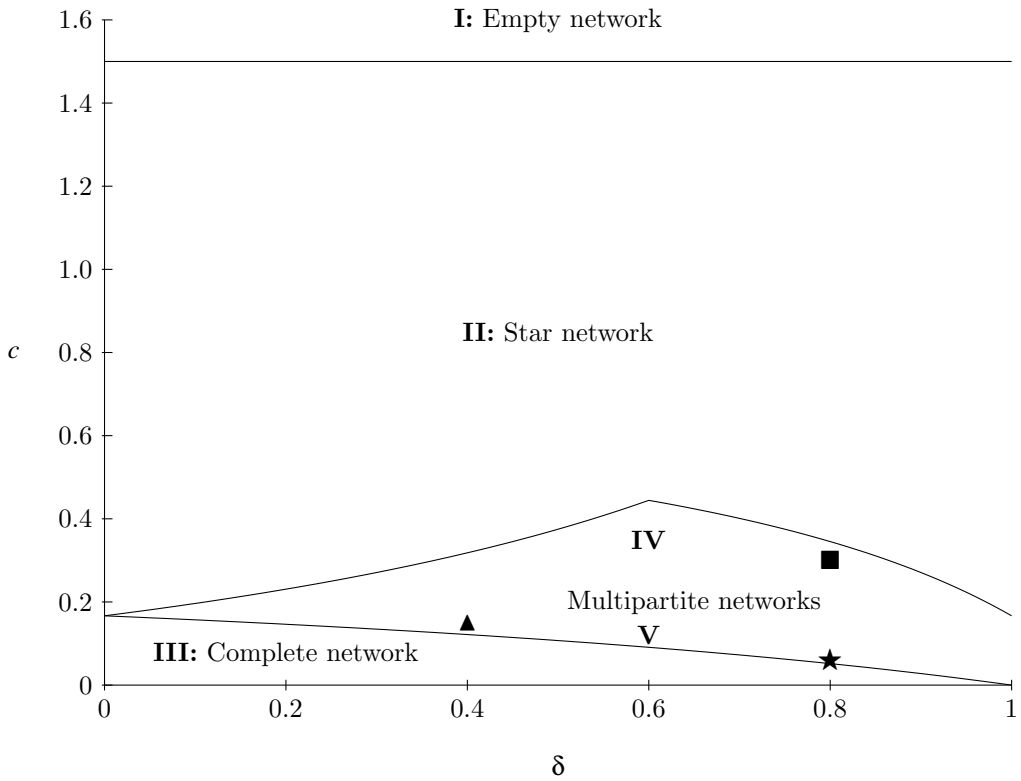
Figure 6 illustrates the parameter regions specified by Theorem 1 for  $n = 4$  and  $n = 8$  under the specification of  $f_e$  and  $f_m$  in equation (1), as in Siedlarek (2015). The possible network outcomes range intuitively from empty to complete networks as the cost of linking

---

<sup>20</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.

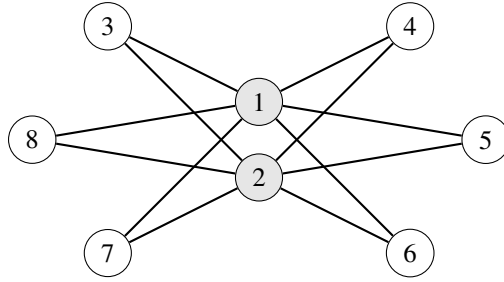


(a)  $n = 4$

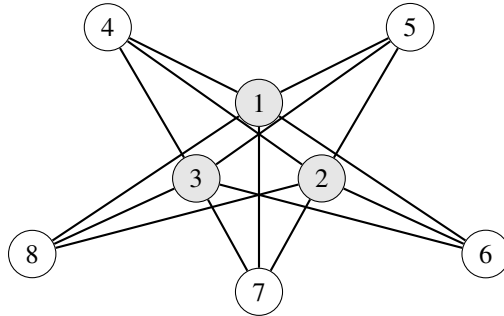


(b)  $n = 8$

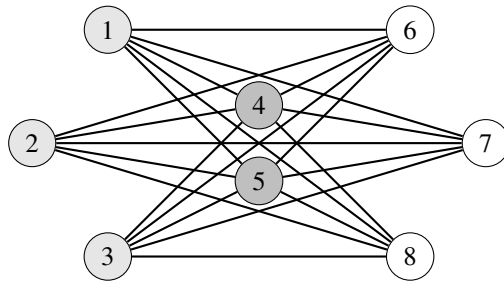
**Figure 6:** Attained equilibria after best-feasible-action dynamics from an empty network in  $(\delta, c)$ -space for  $\alpha_i = 1$  for all  $i \in N$  and  $n \in \{4, 8\}$  under the specification of  $f_e$  and  $f_m$  in equation (1). The roman numbers correspond with those in Theorem 1. The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



(a) Black square  $\blacksquare$ :  $(\delta, c) = (0.8, 0.3)$ . The maximally unbalanced bipartite network:  $g_{2,6}^{mp(2)}$ .



(b) Black triangle  $\blacktriangle$ :  $(\delta, c) = (0.4, 0.15)$ . A less than maximally unbalanced bipartite:  $g_{3,5}^{mp(2)}$ .



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

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



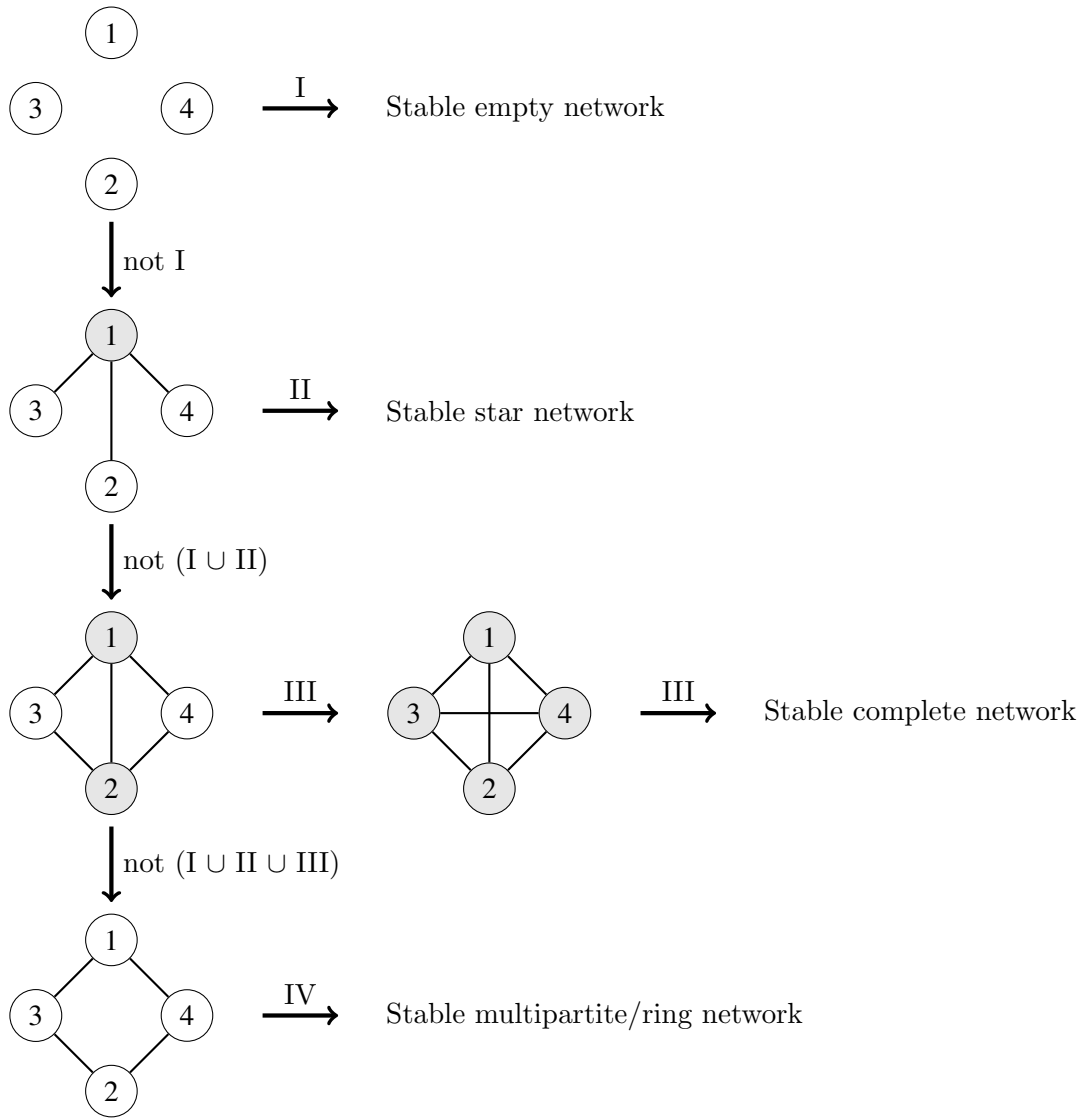
decreases. The star is an important outcome in between empty and complete networks, but cannot be a unilaterally stable outcome for intermediate competition  $\delta$  and relatively low  $c$ . As explained in Section 3.1, for intermediate  $\delta$  the incentives to enter the core and the incentives for periphery members to accept the proposal of the new core member are both high. The parameter area between complete networks and stars gives multipartite networks as the stable outcomes, and this area increases with  $n$ .

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, a less than maximally balanced bipartite network  $g_{3,5}^{mp(2)}$  or a balanced multipartite network  $g_{2,3,3}^{mp(3)}$  consisting of 3 groups. Figure 7 presents these three examples of multipartite networks and the values of  $\delta$  and  $c$  for which they are formed.

A comparison with the results of Goyal and Vega-Redondo (2007) can easily be made by setting  $\delta = 1$ . We observe that their assumption of perfect competition  $\delta = 1$  is a special case for which a complete network is not stable. In their model agents in a complete network always have incentives to remove links even for arbitrary low linking costs, because intermediated trades along multiple middlemen gives them exactly the same share of the surplus.

More importantly, we show that the star is less stable when competition is imperfect compared to the case of Goyal and Vega-Redondo (2007). For  $\delta < 1$  and a relatively low  $c$  multipartite networks arise instead of stars. Multipartite networks arise for larger regions of parameter choices if  $n$  increases; see Figure 11a for the results for  $n = 100$ . Interestingly, multipartite networks were found by Buskens and van de Rijt (2008) and Kleinberg et al. (2008) as the main equilibrium architecture. The current analysis shows that empty, star, complete and multipartite networks can all arise within a network formation model with intermediation and imperfect competition. Our model thus reproduces earlier results of homogeneous network formation models and places them in a more general perspective.

Figure 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. As in Example 1, 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 3, this core-periphery network is not stable, as core members have an incentive to trade indirectly with each other. Hence, the link within the core is dropped and the multipartite network is formed.



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

### 3.3. Efficient networks

In this section we compare the stable networks with efficient networks. Following the results of Goyal and Vega-Redondo (2007), minimally connected networks, i.e. with  $n - 1$  links, are efficient for  $c < \frac{n}{4}$ . Networks are efficient if all trade surpluses are realised irrespective of the distribution of these trading surpluses. For higher  $c$  it is efficient to have no network at all, i.e. an empty network. Because of the assumption that two agents only trade if they are at distance 1 or 2, the network should not have a maximal distance higher than 2, leaving the star as the unique efficient minimally connected network. This is summarised in the following theorem.

**Theorem 2.** *If the payoff function  $\pi(g, \delta, c)$  is given as in Proposition 1 with  $\alpha_i = 1$  for all  $i \in N$ , then:*

- (a) *If  $c \geq n/4$ , then the empty network is efficient*
- (b) *If  $c \leq n/4$ , then the star network is efficient*
- (c) *No other network structure than the empty or star network is efficient.*

Theorem 2 implies that core-periphery networks with  $k \geq 2$  are not efficient.

Comparing the results in Theorem 2 to Theorem 1, we observe that the star network is always efficient, whenever the star network is attained in the dynamic round-robin best-feasible-action process. However, for both low costs and an area of high costs  $c$ , the dynamic process leads to an inefficient outcome. When  $c$  is reasonably high ( $\frac{1}{2} + \frac{1}{6}(n-2) < c < \frac{n}{4}$ ), the dynamic process converges to the empty network, even though the star network is efficient. This is because of the marginal benefits from maintaining a link in the star network are unevenly distributed; the center benefits more than the periphery, such that the center may have an incentive to create link with the periphery, but not vice versa. Here, the possibility of transfers from center to periphery would resolve the inefficiency.

For low costs, under conditions III, IV or V in Theorem 1, the dynamic process converges to either multipartite or to complete networks, whereas the star network is the efficient network. In this case, networks are overconnected. The intuition is given in Example 1 in Subsection 2.4. The star network, although efficient, is not stable, as the center extracts high intermediation rents, which the other players try to circumvent.

The upperbound on  $c$  of area IV, below which stable networks are not efficient, is increasing in  $n$  as shown in Figures 6 and 11a. So for relatively large network sizes, stable networks can be expected to be overconnected.

## 4. Results for heterogeneous banks

We found that complete core-periphery networks were not stable, when agents are ex-ante identical. Instead, multipartite networks are formed, in which banks are connected to members of other groups, but not within their own group. In real interbank markets we do observe links within the core. We now try to explain this discrepancy.

Key in this result in Section 3 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.

For this reason we analyze the consequences of heterogeneity within our model. In Section 4.1 heterogeneity is introduced in terms of exogenous differences in bank size. In Section 4.2, heterogeneity is endogenised by extending the dynamic process to include feedback of profits on trading opportunities.

### 4.1. Exogenous heterogeneity

We start by considering exogenous heterogeneity in the trading opportunities of banks within our model. We interpret this heterogeneity in trading opportunities as arising from differences in bank size. We introduce two types of banks,  $k$  big banks and  $n - k$  small banks, and we assume that the probability of receiving a random unexpected liquidity shock is proportional to bank size. If a bank  $i$  is big we assume  $\alpha_i = \alpha \geq 1$ , and if  $i$  is small we assume  $\alpha_i = 1$ .<sup>21</sup> So the difference in size is captured by a parameter  $\alpha \geq 1$  quantifying the relative size of a big bank. In this subsection we do not explain why some banks are big or small, but rather takes their size as exogenously given. Given heterogeneity in the size of banks, we will show that a stable core-periphery network can form.

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 5 states that the complete core-periphery network can be unilaterally stable under heterogeneity.

**Proposition 5.** *Consider the model with the following form of heterogeneity:  $\alpha_i = \alpha$  for*

---

<sup>21</sup>The normalization  $\alpha_i = 1$  for small banks goes without loss of generality. Compare footnote 17.

$i \in \{1, 2, \dots, k\}$  ('big banks'), and  $\alpha_i = 1$  for  $i \in \{k+1, k+2, \dots, n\}$  ('small banks'). Then given a level of heterogeneity  $\alpha > 1$ , the complete core-periphery network with  $k$  big banks is unilaterally stable if and only if:

$$\text{VII: } c \in \left( \begin{aligned} & \frac{1}{2} - f_e(k, \delta) + (n-k-2) \min\{\frac{1}{2}f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta)\}, \\ & \min\left\{ \alpha^2(\frac{1}{2} - f_e(n-2, \delta)), \alpha(\frac{1}{2} - f_e(k-1, \delta)) + \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta), \right. \\ & \left. \min_{l \leq k} \left\{ \alpha(\frac{1}{2} - f_e(k-l, \delta)) + \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta)) \right\} \right\} \end{aligned} \right).$$

*Proof.* See Appendix E. □

Region VII in Proposition 5 gives the parameter combinations for which complete core-periphery networks are unilaterally stable. The lower bound for  $c$  in region VII does not depend on  $\alpha$ . The upperbound for  $c$  does depend on  $\alpha$ . If the condition for the lower bound is satisfied, the upperbound implicitly defines a minimum level of heterogeneity required for a stable complete core-periphery network. This is made explicit in the following corollary.

**Corollary 1.** *Let the payoff function be heterogeneous as in Proposition 5 and let  $c > 0$  and  $0 < \delta < 1$  be given. Then, if*

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

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

*Proof.* The upperbound for  $c$  in region VII in Proposition 5 is at least linearly increasing in  $\alpha$ . Given  $c > 0$  and  $0 < \delta < 1$ , region VII is nonempty if  $\alpha$  is large enough. □

For  $\delta \downarrow 0$ ,  $\delta \uparrow 1$  and  $k \rightarrow \infty$ , the required level of heterogeneity  $\bar{\alpha}$  is arbitrarily close to 1, as shown in Appendix E.

Proposition 5 combined with Proposition 3 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.

Heterogeneity also affects the results of the dynamics. Starting from an empty graph, complete core-periphery networks can arise because players replicate the position of the central players and become intermediators for periphery players. If the relative size  $\alpha$  is

sufficiently large and condition VII of Proposition 5 is fulfilled, the result will be unilaterally stable. Proposition 6 below specifies the seven possible attracting stable networks for the special case of  $n = 4$  and  $k = 2$  big banks. In this dynamic process, it is assumed that the  $k$  large banks are the first in the round-robin order. Given  $n = 4$ ,  $k = 2$  and any choice of the other parameters  $c$ ,  $\delta$  and  $\alpha$  (except for borderline cases), the proposition shows that the dynamics converge to a unique unilaterally stable network.<sup>22</sup>

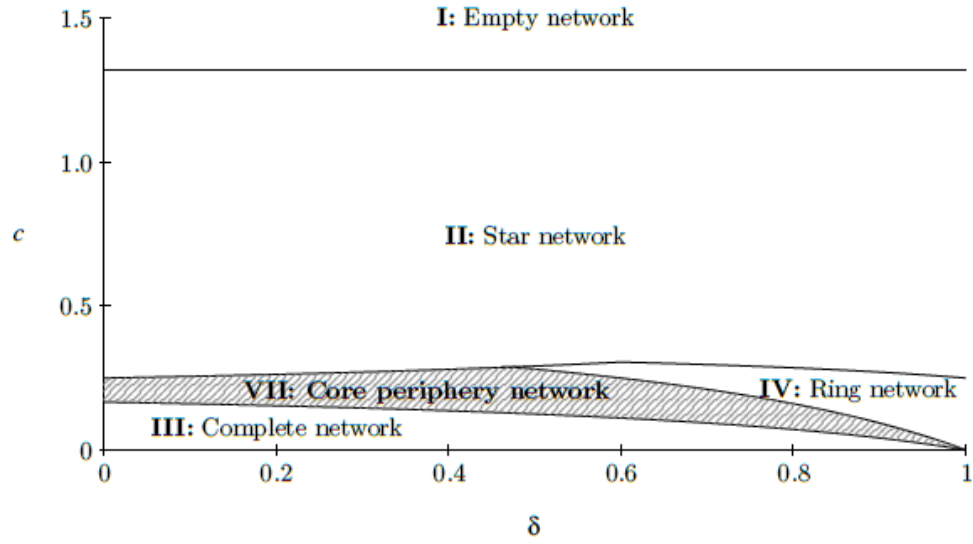
**Proposition 6.** *Consider the model with  $n = 4$ , of which  $k = 2$  big banks having size  $\alpha > 1$  and  $n - k = 2$  small banks having size 1. From an empty graph, the round-robin best-feasible-action dynamics starting with the two big banks converge to the following unilaterally stable equilibria:*

- I: for  $c > \max \left\{ \frac{1}{2}\alpha^2, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\}$   
the empty network,
- II: for  $c \in \left( \frac{1}{6}\alpha + \min \left\{ \frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\}, \min \left\{ \frac{5}{6}\alpha + \frac{1}{6}, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\} \right)$   
the star network,
- III: for  $c < \frac{1}{2} - f_e(2, \delta)$   
the complete network,
- IV: for  $c \in \left( \min \left\{ \alpha^2 \left( \frac{1}{2} - f_e(2, \delta) \right), \frac{1}{6}\alpha + f_m(2, \delta), \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3} \right\}, \frac{1}{6}\alpha + \min \left\{ \frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3} \right\} \right)$   
the multipartite (ring) network  $g_{2,2}^{mp(2)}$
- V: (other multipartite networks do not exist for  $n = 4$ ),
- VI: for  $c \in \left( \min \left\{ \frac{5}{6}\alpha + \frac{1}{6}, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\}, \max \left\{ \frac{1}{2}\alpha^2, \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9} \right\} \right)$   
the ‘single pair’ network  $g$  with  $g_{12} = 1$  and  $g_{ij} = 0$  for all  $(i, j) \neq (1, 2)$ ,
- VII: for  $c \in \left( \frac{1}{2} - f_e(2, \delta), \min \left\{ \alpha^2 \left( \frac{1}{2} - f_e(2, \delta) \right), \frac{1}{6}\alpha + f_m(2, \delta), \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3} \right\} \right)$   
the complete core-periphery network  $g_{com}^{CP(2)}$   
(other core-periphery networks do not exist for  $n = 4$ ).

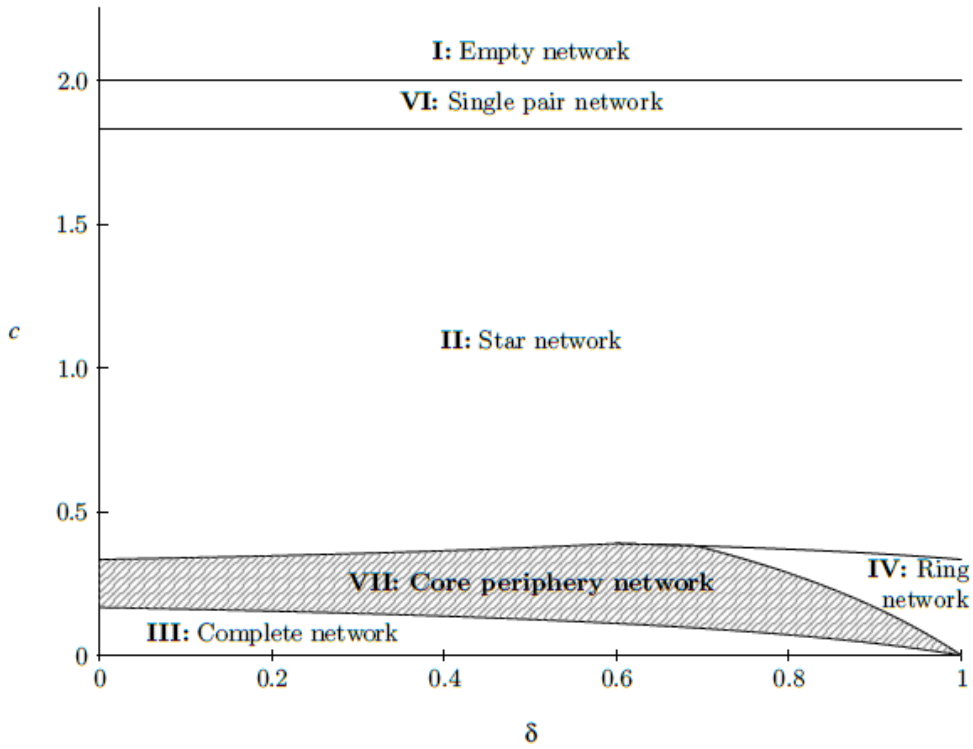
*Proof.* See Appendix E. □

---

<sup>22</sup>Under heterogeneous players the results of the dynamic process depend on the order in which agents choose their best feasible actions. If agents are randomly drawn to make best feasible actions, multiple attracting steady states may exist. Using simulations, we found that for certain parameter values in VII a multipartite ring network can arise. However, for a large subset of the parameter region VII, the complete core-periphery remains the unique attracting steady state even under a random order of agents.



(a)  $\alpha = 1.5$



(b)  $\alpha = 2$

**Figure 9:** Attained equilibria after best-response dynamics from an empty graph in  $(\delta, c)$ -space for  $n = 4$ ,  $k = 2$  and  $\alpha \in \{1.5, 2\}$  under the specification of  $f_e$  and  $f_m$  in equation (2). The roman numbers correspond with those in Theorem 3. In the shaded area, the complete core periphery network with  $k = 2$  big banks is the unique unilaterally stable outcome of the dynamics.

Figure 9 illustrates the different network outcomes for  $n = 4$ ,  $k = 2$  and two levels of heterogeneity,  $\alpha = 1.5$  and  $\alpha = 2$ . Proposition 6 introduces a new (simple) type of network, called a ‘single pair’ network, in which only the two big banks are linked and the small banks have no connections. The parameter region VI is nonempty if the level of heterogeneity is sufficiently high:  $\alpha > \frac{1}{6}(5 + \sqrt{37}) \approx 1.85$ . As observed in Figure 9a, for  $\alpha = 1.5$  the regions of empty and star networks share a border in the  $(\delta, c)$ -diagram:  $c = \frac{1}{6}\alpha^2 + \frac{5}{9}\alpha + \frac{1}{9}$ . For a larger value like  $\alpha = 2$  in Figure 9b, the single pair network arises under condition VI. This type of network structure in which only part of the nodes is connected intuitively arises because some connections are more worthwhile.

Let us now discuss the main network outcome of interest, namely core-periphery networks. The condition under which the complete core-periphery network is the attracting steady state is in Figure 9 indicated by the shaded regions. As expected, this region increases with the level of heterogeneity  $\alpha$ . Also observe that for complete core-periphery networks to arise it is necessary that competition is less than fully perfect, i.e.  $\delta < 1$ . For the special case of  $\delta = 1$  (as considered by Goyal and Vega-Redondo, 2007) core-periphery networks are never unilaterally stable, not even under large heterogeneity. Finally, complete core-periphery networks arise for a larger set of parameters if the number of players is larger than the minimal network size of  $n = 4$ . See Figure 11b for the shaded region VII given  $n = 100$ ,  $k = 15$  and  $\alpha = 10$ .



## 4.2. Endogenous heterogeneity

From the analysis in the above subsection 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 trading opportunities, 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 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.

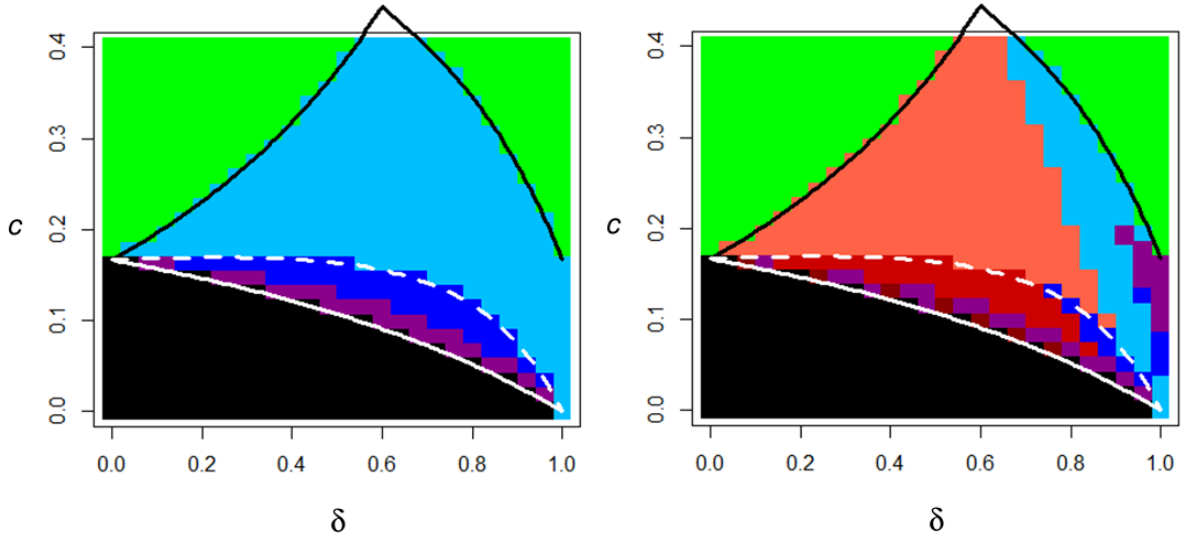
The extended dynamic process starts off at round  $\tau = 1$  as the homogeneous baseline model with  $\alpha_i^{(1)} = 1$  for all  $i \in N$ . From an empty graph, round-robin best-feasible-action dynamics converge to the empty network, the star network, the complete network or multipartite networks,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

$$\alpha_i^{(\tau)} = \frac{\pi_i^{(\tau-1)}}{\min_k [\pi_k^{(\tau-1)}]} \quad (4)$$

Trade is assumed to be proportional to size. In this updating, the additional assumption is made that size is proportional to interbank profits. The trading opportunities are rescaled with  $\min_k [\pi_k]$  to assure that trade surpluses of the smallest banks remain normalised to  $\alpha_i = 1$ .

Having updated the trading opportunities  $\{\alpha_i\}_{i \in N}$  at the beginning of round  $\tau$ , we perform a new round of the round-robin best-feasible-action dynamics, leading to potentially a new network structure and new payoffs. In the next round, the trading opportunities are again updated using (4), and so on.

We simulate this dynamic process. As a stopping rule for the simulations, we impose that the process stops after  $\mathcal{T} = 25$  rounds, or before at round  $\tau$  if the trading opportunities 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 within 25 rounds and we cannot exclude cyclic behaviour in that case.



**Figure 10:** 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 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.

Figure 10 plots the results for  $n = 8$  and  $\Delta\alpha = 0.1$  after 1 round (left panel) and after 25 rounds  $\mathcal{T} = 25$  (right panel), using a  $25 \times 25$  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  $\tau = 1$ , banks are still homogeneous, that is,  $\alpha_i^{(1)} = 1$  for all  $i \in N$ . The left panel of Figure 10 therefore repeats Figure 6b. As we have seen in Subsection 3.2, in that case, core-periphery networks do not occur.

The right panel of Figure 10 shows the results after a maximum of  $\mathcal{T} = 25$  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 trading opportunities remain equal to 1. Also the star network cannot change: profit feedback increases the trading opportunities  $\alpha_i$  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 members do not generate more, so no links are added. Hence, if a star network is formed under homogeneity after the first round, then the network architecture remains a star network after any future rounds.

This is not the case, however, if in the first round a multipartite network is formed. In that case, we observe significant changes after 25 rounds. In fact, in many simulations, we

observe a transition from multipartite networks to core-periphery networks. This happens intuitively. For example, if after round 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 higher trading opportunities  $\alpha_i$ , 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>23</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 the interbank market in the Netherlands, a relatively small market with approximately a hundred 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 Figure 11a. 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 stylised 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 €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 of in 't Veld and van Lelyveld (2014) lies around €8 billion. For periphery banks the median size is approximately €300 million (see Fig. 9, in 't Veld and van Lelyveld, 2014, for the plotted distribution of total asset size over core and periphery banks). Ignoring the exceptionally

---

<sup>23</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 trading opportunities central players may be induced to remove several links, and cycles with different players taking central positions may occur.

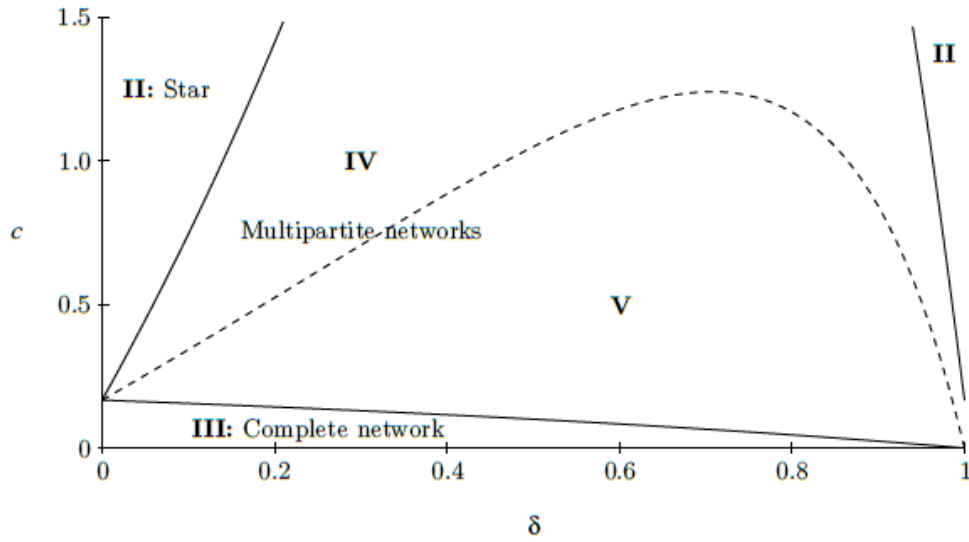
large size of some core banks, a relative difference of  $\alpha = 10$  seems a reasonable order of magnitude. The larger  $\alpha$ , the wider the area in which the core-periphery network is stable.

Figure 11b shows the parameter region given by Proposition 5 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 Figure 11b. These complete networks and complete core-periphery networks would also be the outcome of best-feasible-action dynamics starting with the big banks.<sup>24</sup> The observed core-periphery structure in the Netherlands can be reproduced for many reasonable choices of linking costs  $c$  and competitiveness  $\delta$ .

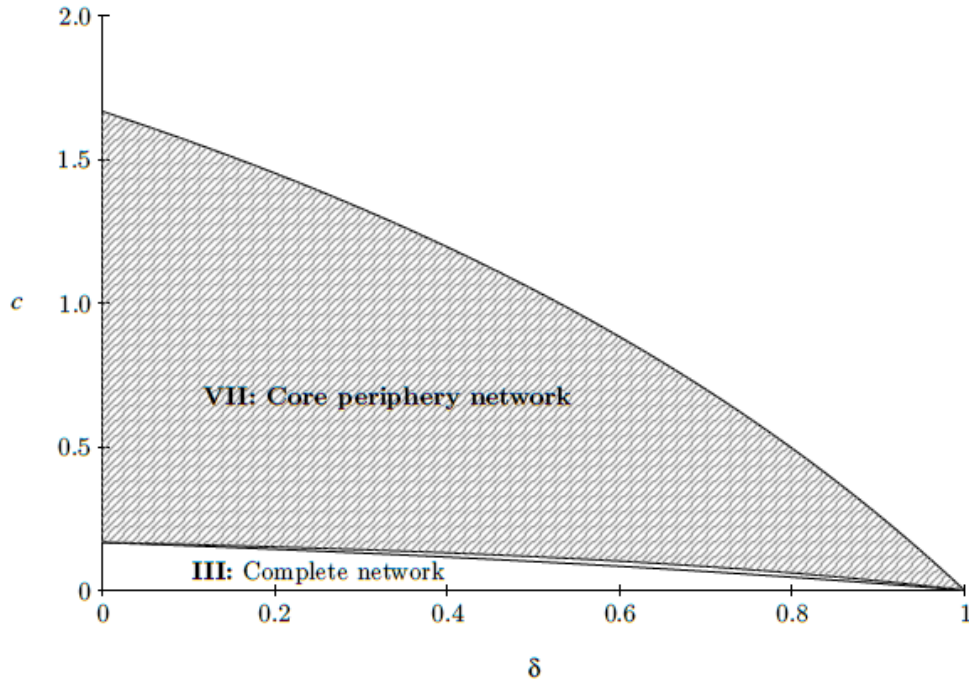
This application suggests that our model is suitable to explain stylised facts of national or perhaps even international interbank networks. It should be noted that observed core-periphery networks are not necessarily complete core-periphery networks, which can be explained as follows. Empirical studies of interbank markets often rely on either balance sheet data measuring the total exposure of one bank on another, 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. Given a theoretically complete core-periphery network of lending relationships, trades and exposures are executed on the same structure of connections; see Section 2.2 on how we model trade benefits. The empirically observed core-periphery structure from these trade networks is a subset of the unobserved relationship network, and hence, it can have less than complete connections between core and periphery depending on the realizations of trade opportunities. 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>24</sup>A full description of the outcomes of these dynamics, specifying all other areas in Figure 11b, is missing. This would require a generalization of Proposition 6 for all  $n \geq 4$ . 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 5 and Proposition 6.



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



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

**Figure 11:** Application of the model to the Dutch interbank market. Attained equilibria after best-response dynamics from an empty graph 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 2 and Proposition 3. In the shaded area, the complete core periphery network with  $k = 15$  big banks is the unique unilaterally stable outcome of best response dynamics starting with the large banks.

## 6. Conclusion

In this paper we propose a way to explain the formation of financial networks by intermediation. We focus on the core-periphery network because it is found to give a fair representation of the complex empirical structures, 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 endogenise 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 future trade opportunities of these banks. This is consistent with a recent empirical finding by Akram and Christophersen (2010) that banks with many financial linkages in the Norwegian interbank market face lower interest rates. 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 two suggestions for future research. First, 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. In this paper, we do not model the balance sheets. Introducing balance sheets in our model, would be one potential direction for future research.

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 overnight loan market approximately was before the financial crisis of 2007. However, we could analyze the effects of an increase of default risk on overnight 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 (in 't Veld and van Lelyveld, 2014; Martínez-Jaramillo et al., 2014; Fricke and Lux, 2015).

## References

- Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A., 2014. Systemic risk in endogenous financial networks. mimeo, MIT and Columbia University.
- Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A., 2015. Systemic risk and stability in financial networks. *American Economic Review* 105 (2), 564–608.
- Afonso, G., Lagos, R., 2015. Trade dynamics in the market for federal funds. *Econometrica* 83, 263–313.
- Akram, Q. F., Christophersen, C., 2010. Interbank overnight interest rates: Gains from systemic importance. Working Paper 2010-11, Norges Bank.
- Allen, F., Gale, D., 2000. Financial contagion. *Journal of Political Economy* 108, 1–33.
- Babus, A., 2016. The Formation of Financial Networks. *RAND Journal of Economics* 47 (2), 239–272.
- Babus, A., Hu, T.-W., 2016. Endogenous Intermediation in Over-the-Counter Markets. *Journal of Financial Economics*, to appear.
- Bedayo, M., Mauleon, A., Vannetelbosch, V., 2016. Bargaining in endogenous trading networks. *Mathematical Social Sciences* 80, 70–82.
- Borgatti, S. P., Everett, M. G., 1999. Models of core/periphery structures. *Social Networks* 21, 375–395.
- Bräuning, F., Fecht, F., 2012. Relationship Lending in the Interbank Market and the Price of Liquidity. Deutsche Bundesbank Discussion Paper 2012 (22).
- Buskens, V., van de Rijt, A., 2008. Strategic network formation with structural holes. *American Journal of Sociology* 114, 37–407.
- Cabrales, A., Gottardi, P., Vega-Redondo, F., 2013. Risk-sharing and contagion in networks. mimeo, Universidad Carlos III and European University Institute.
- Castiglionesi, F., Navarro, N., 2016. (in)efficient interbank networks. *Cahiers du GRETha* 2016-13.
- Chang, B., Zhang, S., 2016. 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. *Journal of Financial Intermediation* 18, 24–48.
- Cohen-Cole, E., Patacchini, E., Zenou, Y., 2015. Static and dynamic networks in interbank markets. *Network Science* 3 (01), 98–123.
- Craig, B., Von Peter, G., 2014. Interbank tiering and money center banks. *Journal of Financial Intermediation* 23 (3), 322–347.
- Di Maggio, M., Kermani, A., Song, Z., 2016. The value of trading relationships in turbulent times. *Journal of Financial Economics*, to appear.
- Duffie, D., Gârleanu, N., Pedersen, L. H., 2005. Over-the-Counter Markets. *Econometrica* 73, 1815–1847.
- Eisenberg, L., Noe, T. H., 2001. Systemic Risk in Financial Systems. *Management Science* 47 (2), 236–249.
- Elliott, M., Golub, B., Jackson, M. O., 2014. Financial networks and contagion. *American Economic Review* 104 (10), 3115–3153.
- Farboodi, M., 2015. Intermediation and voluntary exposure to counterparty risk. Mimeo, princeton university.
- Fricke, D., Lux, T., 2015. Core–periphery structure in the overnight money market: evidence from the e-mid trading platform. *Computational Economics* 45 (3), 359–395.
- Gai, P., Haldane, A., Kapadia, S., 2011. Complexity, Concentration and Contagion. *Journal of Monetary Economics* 58 (5).
- Galeotti, A., Goyal, S., 2010. The law of the few. *American Economic Review* 100 (4), 1468–92.
- Glasserman, P., Young, H. P., 2016. Contagion in financial networks. *Journal of Economic Literature*, to appear.
- 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. *Journal of Economic Theory* 137 (1), 460–492.
- Hiller, T., 2015. Peer effects in endogenous networks. mimeo, University of Bristol.



- Houy, N., 2009. Structural holes in social networks: A remark. *Journal of Economic Theory* 144 (1), 422–431.
- in 't Veld, D., van Lelyveld, I., 2014. Finding the core: Network structure in interbank markets. *Journal of Banking & Finance* 49, 27–40.
- Jackson, M., Wolinsky, A., 1996. A strategic model of social and economic networks. *Journal of Economic Theory* 71 (1), 44–74.
- 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. *Journal of Economic Theory* 106 (2), 265–295.
- Kleinberg, J., Suri, S., Tardos, E., Wexler, T., 2008. Strategic network formation with structural holes. *Proceedings 9th ACM Conference on Electronic Commerce*, 480–489.
- König, M. D., Tessone, C. J., Zenou, Y., 2014. Nestedness in networks: A theoretical model and some applications. *Theoretical Economics* 9 (3), 695–752.
- 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. *Journal of Economic Dynamics and Control* 40, 242–265.
- Merlo, A., Wilson, C. A., 1995. A stochastic model of sequential bargaining with complete information. *Econometrica* 63 (2), 371–99.
- Nier, E., Yang, J., Yorulmazer, T., Alentorn, A., 2007. Network models and financial stability. *Journal of Economic Dynamics and Control* 31, 2033–2060.
- Peltonen, T., Scheicher, M., Vuillemeij, G., 2014. The network structure of the cds market and its determinants. *Journal of Financial Stability* 13, 118–133.
- Persitz, D., 2012. Power and core-periphery networks. Tech. rep., Tel Aviv University.
- Siedlarek, J.-P., 2013. Making friends meet: Network formation with introductions. mimeo, University of Mannheim.
- Siedlarek, J.-P., October 2015. Intermediation in networks. Working Paper 15-18, Federal Reserve Bank of Cleveland.
- Wang, C., 2016. Core-periphery trading networks. mimeo, Stanford University.

## Appendix A. Siedlarek’s payoff function

Siedlarek (2015) applies the bargaining protocol introduced by Merlo and Wilson (1995) to derive the distribution of a surplus for a trade that is intermediated by competing middlemen. The underlying idea is that one involved agent is selected to propose a distribution of the surplus. If the proposer succeeds in convincing the other agents to accept his offer, the trade will be executed. Otherwise, the trade will be delayed and another (randomly selected) agent may try to make a better offer. A common parameter  $0 \leq \delta \leq 1$  is introduced with which agents discount future periods, that results in the level of competition as used in our paper. The relation between the discount factor  $\delta$  and the level of competition, is that if agents are more patient, then intermediaries are forced to offer more competitive intermediation rates when having the chance to propose a distribution, as trading partners are more willing to wait for the opportunity to trade with alternative intermediaries. As a special case, for  $\delta = 1$  the surplus is distributed equally among the essential players as in Goyal and Vega-Redondo (2007).

More formally, assume that the set of possible trading routes is known to all agents (i.e. complete information). Each period a route is selected on which the trade can be intermediated, and additionally one player (the ‘proposer’) on this route that proposes an allocation along the entire trading route.<sup>25</sup> Any state, which is a selection of a possible path and a proposer along that path, is selected with equal probability and history independent. The question now is: what is the equilibrium outcome, i.e. the expected distribution of the surplus, taking into account that every agent proposes optimally under common knowledge of rationality of other possible future proposers? Siedlarek (2015) shows that the unique Markov perfect equilibrium is characterised by the following payoff function for any player  $i$  in a certain state:

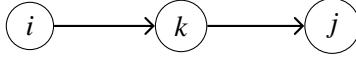
$$f_i = \begin{cases} 1 - \sum_{j \neq i} \delta E_j[f_j] & \text{if } i \text{ is the proposer in this state} \\ \delta E_i[f_i] & \text{else if } i \text{ is involved in this state} \end{cases} \quad (\text{A.1})$$

This equation shows that the proposer can extract all surplus over and above the outside option value given by the sum of  $E_j[f_j]$  over all other players  $j \neq i$ . All (and only) the players along the same route have to be convinced by offering exactly their outside option.

In the simple example of intermediary  $k$  connecting  $i$  and  $j$  (see Figure A.12), each of the

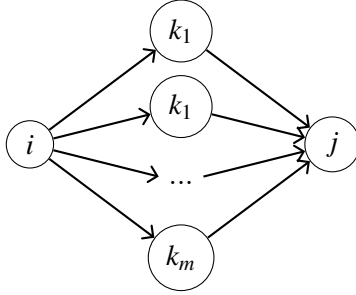
---

<sup>25</sup>Siedlarek (2015) assumes that only *shortest* paths are considered, an assumption not explicitly made by Goyal and Vega-Redondo (2007). However, in the model Goyal and Vega-Redondo (2007) only essential players, who by definition are part of shortest paths, receive a nonzero share of the surplus.



**Figure A.12:** A trading route with one intermediary

three players has equal probability of becoming the proposer and proposes an equal share to the other two, so the equilibrium distribution of the surplus will simply be the equal split  $(f_i, f_k, f_j) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  for all  $\delta$ .



**Figure A.13:** A trading network with  $m$  competing intermediaries

In case  $m$  intermediaries compete for the trade between  $i$  and  $j$  (as in Figure A.13), the extracted intermediation rents are reduced. Siedlarek (2015) shows that:

$$f_i = f_j = f_e(m, \delta) \equiv \frac{m - \delta}{m(3 - \delta) - 2\delta} \quad (\text{A.2})$$

$$f_k = f_m(m, \delta) \equiv \frac{1 - \delta}{m(3 - \delta) - 2\delta} \quad (\text{A.3})$$

As mentioned before, this distribution of payoffs satisfies our assumptions on  $f_e(m, \delta)$  and  $f_m(m, \delta)$ . The distribution of Siedlarek (2015) is used as the leading example in the results. All results hold for general  $f_e(m, \delta)$  and  $f_m(m, \delta)$  satisfying the assumptions of Table 1.

In a general network with longer intermediation chains  $d > 2$ , the payoffs cannot be calculated directly, but depend on specifics of (part of) the network. The vector of payoffs  $\vec{F}$  can be calculated indirectly from the network  $g$  using the following notation:

- $s$ : The number of shortest paths that support the trade,
- $d$ : The length of the shortest paths (i.e. the number of required intermediaries plus one),
- $\vec{P}$ :  $(n \times 1)$ -vector of shortest paths that run through any player,

- $K$ :  $(n \times n)$ -diagonal matrix of times that any player receives an offer,
- $S$ :  $(n \times n)$ -off-diagonal matrix of times that two players share a path.

As in any given state the distribution of profits is as given in equation (A.1), the vector of payoffs that averages over all possible states of the world is:

$$\vec{F} = \frac{1}{s \cdot d} (\vec{P} - \delta(S - K)\vec{F}) \quad (\text{A.4})$$

If  $\delta < 1$  this can be solved to:

$$\vec{F} = (s \cdot d \cdot I_n + \delta(S - K))^{-1} \vec{P} \quad (\text{A.5})$$

In Appendix F we use this general formula to investigate whether core-periphery networks can be stable when intermediation paths of lengths  $d = 3$  are allowed.

## Appendix B. Proofs of Section 2

*Proof of Proposition 1.* Writing out the payoff function, we obtain

$$\begin{aligned}
\tilde{\pi}_i(g, \delta, \tilde{c}) &= \sum_{t=1}^{\infty} \beta^t \left( \left( \sum_{s_t \in \{-1, 0, 1\}^n} P[S_t = s_t] b_i(s_t, g, \delta) \right) - \eta_i(g) \tilde{c} \right) \\
&= \frac{\beta}{1-\beta} \left( \left( \sum_{s_t \in \{-1, 0, 1\}^n} P[S_t = s_t] b_i(s_t, g, \delta) \right) - \eta_i(g) \tilde{c} \right) \\
&= \frac{\beta}{1-\beta} \left( \left( \sum_{s_t: n_t^+ = n_t^- = 1} P[S_t = s_t] b_i(s_t, g, \delta) + \sum_{s_t: n_t^+ + n_t^- \geq 3} P[S_t = s_t] b_i(s_t, g, \delta) \right) - \eta_i(g) \tilde{c} \right),
\end{aligned}$$

where  $n_t^- = |s_{it} < 0|$ ,  $n_t^+ = |s_{it} > 0|$ , and  $n_t^0 = n - n_t^+ - n_t^-$ . Note that  $b_i(s_t, g, \delta) = 0$  if  $n_t^+ = 0$  or  $n_t^- = 0$ , as in that case, there is no trade. Let  $A = \max_i \alpha_i$ .  $S_t$  has a multinomial distribution, and hence  $P[S_t = s_t]$  is bounded by  $P[S_t = s_t] < A^n \rho^{n_t^+ + n_t^-}$ . As  $b_i(s_t, g, \delta) < 1$ , and  $|s_t : n_t^+ + n_t^- \geq 3| < 2^n$  we have

$$\lim_{\rho \downarrow 0} \frac{1}{\rho^3} \sum_{s_t: n_t^+ + n_t^- \geq 3} P[S_t = s_t] b_i(s_t, g, \delta) < (2A)^n,$$

that is, benefits from periods in which three banks or more receive a liquidity shock are of order  $O(\rho^3)$ . Hence,

$$\tilde{\pi}_i(g, \delta, \tilde{c}) = \frac{\beta}{1-\beta} \left( \left( \sum_{s_t: n_t^+ = n_t^- = 1} P[S_t = s_t] b_i(s_t, g, \delta) \right) - \eta_i(g) \tilde{c} \right) + O(\rho^3).$$

Suppose that  $j$  receives a positive shock and  $k$  a negative shock. Then the benefit of  $i$  is  $b_i(s_t, g, \delta) = sv_i^{jk}(g, \delta) = v_i^{jk}(g, \delta)/2$ , as explained in the main text. Hence the payoff function

becomes

$$\begin{aligned}
\tilde{\pi}_i(g, \delta, \tilde{c}) &= \frac{\beta}{1-\beta} \left[ \left( \sum_j \sum_{k \neq j} P[S_{jt} = 1, S_{kt} = -1, \forall l \neq j, k : S_{lt} = 0] v_i^{jk}(g, \delta) / 2 \right) - \eta_i(g) \tilde{c} \right] + O(\rho^3) \\
&= \frac{\beta}{1-\beta} \left[ \left( \sum_j \sum_{k \neq j} \rho^2 \alpha_j \alpha_k \left( \prod_{l \neq j, k} (1 - 2\rho \alpha_l) \right) v_i^{jk}(g, \delta) / 2 \right) - \eta_i(g) \tilde{c} \right] + O(\rho^3) \\
&= \frac{\beta}{1-\beta} \left[ \left( \sum_j \sum_{k \neq j} \rho^2 \alpha_j \alpha_k v_i^{jk}(g, \delta) / 2 \right) - \eta_i(g) \tilde{c} \right] \\
&\quad - \frac{\beta}{1-\beta} \left[ \sum_j \sum_{k \neq j} \rho^2 \alpha_j \alpha_k \left( 1 - \prod_{l \neq j, k} (1 - 2\rho \alpha_l) \right) v_i^{jk}(g, \delta) / 2 \right] + O(\rho^3) \\
&= \frac{\beta \rho^2}{1-\beta} \left[ \left( \sum_j \sum_{k \neq j} \alpha_j \alpha_k v_i^{jk}(g, \delta) / 2 \right) - \eta_i(g) \frac{\tilde{c}}{\rho^2} \right] + O(\rho^3),
\end{aligned}$$

as

$$\rho^2 \left( 1 - \prod_{l \neq j, k} (1 - 2\rho \alpha_l) \right) = O(\rho^3).$$

The benefits for bank  $i$  from direct trade are:

$$\sum_{j \in N_i^1(g)} \alpha_i \alpha_j \left( v_i^{ij}(g, \delta) / 2 + v_i^{ji}(g, \delta) / 2 \right) = \sum_{j \in N_i^1(g)} \frac{1}{2} \alpha_i \alpha_j.$$

Note that we count two trades for  $i$  and  $j$ , one time  $i$  having a liquidity surplus, and one time  $i$  having a liquidity deficit. Similarly, the benefits from indirect trade are given by

$$\sum_{j \in N_i^2(g)} \alpha_i \alpha_j \left( v_i^{ij}(g, \delta) / 2 + v_i^{ji}(g, \delta) / 2 \right) = \sum_{j \in N_i^2(g)} \alpha_i \alpha_j f_e(m_{ij}(g), \delta),$$

and for  $d > 2$ :

$$\sum_{j \in N_i^d(g)} \alpha_i \alpha_j \left( v_i^{ij}(g, \delta) / 2 + v_i^{ji}(g, \delta) / 2 \right) = 0.$$

Finally, intermediation benefits are given by

$$\sum_{k, l \in N_i^1(g) | g_{kl} = 0} \alpha_k \alpha_l \left( v_i^{kl}(g, \delta) / 2 + v_i^{lk}(g, \delta) / 2 \right) = \sum_{k, l \in N_i^1(g) | g_{kl} = 0} \alpha_k \alpha_l f_m(m_{kl}(g), \delta),$$

for all pairs  $k, l$  at distance 2 and  $i$  in between. For all other pairs, there are no benefits to  $i$ . The payoff function then follows.  $\square$

## Appendix C. Theoretical results

To find under which conditions the basic network structures are stable, we introduce the 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 5.** *A feasible action for player  $i$  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  $\pi_j(g^{iS}) \geq \pi_j(g)$ .

The best feasible action  $S^*$  is the feasible action that gives  $i$  the highest payoffs:

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

A network  $g$  is unilaterally stable if every player chooses its best feasible action.

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.** *In an empty network that is not unilaterally stable, the best feasible action for a player is to add links to all other nodes.*

*Proof.* The marginal benefits<sup>26</sup> of adding  $l$  links to an empty network are:

$$M_i(g^e, +l) = l\left(\frac{1}{2} - c\right) + \binom{l}{2} \frac{1}{3}. \quad (\text{C.2})$$

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

---

<sup>26</sup>Throughout the paper, marginal benefits of an action  $S$  by player  $i$  for player  $j$  are defined as:

$$M_j(g, S) \equiv \pi_j(g^{iS}) - \pi_j(g). \quad (\text{C.1})$$

With some abuse of notation, we will informally denote the action  $S$  as  $+l$  or  $-l$ , where  $l = |S|$  is the number of links that is either added or deleted in the network. From the text it will be clear which player  $i$  and which set  $S$  are considered.

The star is unstable for low costs  $c$ , because periphery members gain incentives to add links. The gain of having a direct (rather than an intermediated) trade are in this case  $\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ . However, it is even better for one periphery member to propose multiple links, thereby creating more chains of neighbours through which it intermediates.

**Lemma 2.** *Consider a star network that is not unilaterally stable because a peripheral player  $i$  can deviate by adding one or more links. Then the best feasible action is to add links to all  $l = n - 2$  other periphery players.*

*Proof.* Consider the deviation of a peripheral player  $i$  to add  $l$  links. The marginal benefits consist of making trades direct instead of intermediated by the center of the star, and of creating middlemen benefits:

$$M_i(g^s, +l) = l\left(\frac{1}{6} - c\right) + \binom{l}{2} f_m(2, \delta) \quad (\text{C.3})$$

As in Lemma 1 this function is convex in  $l$ , so the maximum is reached at  $l^* = 0$  or  $l^* = n - 2$ .  $\square$

Lemma 2 implies that it is never a best feasible action to add  $0 < l < n - 2$  links to some other peripheral nodes. Together Lemmas 1 and 2 imply that to find the stability conditions of empty and star networks, it is sufficient to check whether adding or deleting all possible links is beneficial in these networks.

As it was optimal in an empty graph and in a star network to link immediately to all players (Lemmas 1 and 2), if a peripheral agent will add a link it is optimal to connect to all other players, thereby entering in the core himself.

**Lemma 3.** *Consider a complete core-periphery network with  $k \in \{2, 3, \dots, n - 2\}$  that is not unilaterally stable because a peripheral player  $i$  can deviate by adding one or more links. Then the best feasible action is to add links to all  $l = n - k - 1$  other periphery players.*

*Proof.* The marginal benefits for replicating the position of the center of the star was given in Lemma 2. It can easily be generalised to complete core-periphery networks with  $k$  core members:

$$M_i(g_{com}^{CP(k)}, +l) = l\left(\frac{1}{2} - f_e(k, \delta) - c\right) + \binom{l}{2} f_m(k + 1, \delta) \quad (\text{C.4})$$

The maximum of this function is reached at  $l^* = 0$  or  $l^* = n - k - 1$ .  $\square$



## Appendix D. Proofs of Section 3

*Proof of Proposition 2.* As in Goyal and Vega-Redondo (2007), the center of the star earns more than the periphery members, but receives less profit per link. The star therefore becomes unilaterally unstable if the center of the star would benefit from deleting all its links, i.e.:

$$\begin{aligned} M_i(g^s, -(n-1)) = -\pi_i(g^s) &< 0 \\ \Leftrightarrow c &> \frac{1}{2} + \frac{1}{6}(n-2). \end{aligned} \quad (\text{D.1})$$

If and only if this condition holds, then the empty network is unilaterally stable.

Lemma 2 states that if a peripheral player  $i$  in a star network wants to deviate by adding links, this player will become a member of a newly formed core. In other words, a complete core-periphery network arises with  $k = 2$ . Player  $i$  becomes directly connected to all other players, and also receives intermediation gains which depend on  $\delta$ . For the new core member  $i$  to have positive marginal benefits of supporting all new links, it is required that

$$\begin{aligned} M_i(g^s, +(n-2)) &> 0 \\ \Rightarrow c &< \frac{1}{6} + \frac{1}{2}(n-3)f_m(2, \delta), \end{aligned} \quad (\text{D.2})$$

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

$$\begin{aligned} M_j(g^s, +(n-2)) = \frac{1}{6} - c + (n-3)(f_e(2, \delta) - \frac{1}{3}) &\geq 0 \\ \Rightarrow c &\leq \frac{1}{6} + (n-3)(f_e(2, \delta) - \frac{1}{3}). \end{aligned} \quad (\text{D.3})$$

Conversely, the star network is stable if the deviation of player  $i$  is *not* beneficial to  $i$  and/or a peripheral player  $j \neq i$ , i.e. if  $c$  exceeds the minimum of the two values in equations (D.2) and (D.3):

$$c \geq \frac{1}{6} + (n-3) \min\left\{\frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3}\right\}. \quad (\text{D.4})$$

A complete network can become unstable when any two nodes decide to remove a mutual link. Note that the trade between those two links is sustained by intermediation through all other nodes. Formally we have therefore a core-periphery network with  $k = n - 2$  core

players. The complete network is therefore stable if and only if:

$$c \leq \frac{1}{2} - f_e(n-2, \delta) \quad (\text{D.5})$$

□

*Proof of Proposition 3.* Consider the deviation of a core player  $i$  to delete one link to another core member. The marginal benefit of doing so is:

$$M_i(g_{com}^{CP(k)}, -1) = c - \left(\frac{1}{2} - f_e(n-2, \delta)\right) \quad (\text{D.6})$$

The marginal benefit for a peripheral player  $j$  to add a link in the periphery is:

$$M_j(g_{com}^{CP(k)}, +1) = \left(\frac{1}{2} - f_e(k, \delta)\right) - c \quad (\text{D.7})$$

In order for the network to be pairwise stable, it is required that  $M_i(g_{com}^{CP(k)}, -1) \leq 0$  and  $M_j(g_{com}^{CP(k)}, +1) \leq 0$ , so:

$$f_e(n-2, \delta) \leq \frac{1}{2} - c \leq f_e(k, \delta)$$

However, as  $\frac{\partial f_e}{\partial k} > 0$  for  $0 < \delta < 1$  and  $k < n-2$ , this interval of  $\frac{1}{2} - c$  is empty. Complete core-periphery networks are never pairwise stable, and therefore also not unilaterally stable. □

*Proof of Proposition 4.* We will determine the benefits of a periphery player to propose links to all other periphery players. In Lemma 3 we showed that this is a best feasible action in *complete* core-periphery networks, if there exists a profitable deviation by adding links at all. In a general core-periphery network, adding to all other periphery players is not necessarily a best feasible action, but it will help to find the sufficient condition for instability of the networks, as claimed in the Proposition.

Consider a periphery player  $i$  proposing  $l_i = n - k - 1$  links in order to reach all other periphery players. The minimal marginal benefits for  $i$  of this action are bounded by:

$$M_i(g^{CP(k)}, +l_i) > -(n-k-1)c + \binom{n-k-1}{2} f_m(k+1, \delta). \quad (\text{D.8})$$

This lower bound for the marginal benefits consists of two parts. The first part,  $-(n-k-1)c$ , denotes the direct costs for adding the  $l_i$  links and is linear in  $n$ . The second part,  $\binom{n-k-1}{2} f_m(k+1, \delta)$ , denotes the minimal intermediation benefits from becoming a new intermediator between the  $n-k-1$  remaining periphery members and is quadratic in  $n$ . In

general these intermediation benefits for  $i$  can be higher if there are fewer  $m_{jl} < k$  intermediators in the original network  $g^{CP(k)}$  between  $j, l \in P$ . Benefits from trade (direct or indirect) do not have to be considered because, obviously, adding links to the network can only increase the access from  $i$  to other players, and the benefits from trade therefore weakly increase.

Let the parameters  $c > 0$ ,  $0 < \delta < 1$  and  $k \geq 1$  be given. Then positive marginal benefits for  $i$  imply:

$$\begin{aligned} M_i(g^{CP(k)}, +l_i) &> 0 \\ \Rightarrow c &< \frac{1}{2}(n-k-2)f_m(k+1, \delta) \\ \Leftrightarrow n &> F_i(c, \delta, k) \equiv k+2 + \frac{c}{\frac{1}{2}f_m(k+1, \delta)} \end{aligned} \quad (\text{D.9})$$

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  $j$  depend on its number of connections  $n_j$ , but are bounded by:

$$M_j(g^{CP(k)}, +l_i) > -c + (n-k-2) \min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\}. \quad (\text{D.10})$$

The second part of these marginal benefits indicates that the number of intermediation paths from  $j$  to any other periphery member  $l \in (P \setminus i, j)$  strictly increases, and therefore the benefits from trade strictly increase.<sup>27</sup> This function is similarly convex, and positive marginal benefits for  $j$  imply:

$$\begin{aligned} M_j(g^{CP(k)}, +l_i) &\geq 0 \\ \Rightarrow c &< (n-k-2) \min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\} \\ \Leftrightarrow n &> F_j(c, \delta, k) \equiv k+2 + \frac{c}{\min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\}} \end{aligned} \quad (\text{D.11})$$

Combining the conditions for  $i$  (D.9) and  $j$  (D.10), a sufficient lower bound for the network size  $n$  for any core-periphery network to be unstable is:

$$n > F(c, \delta, k) \equiv k+2 + \frac{c}{\min\left\{\frac{1}{2}f_m(k+1, \delta), \min_{n_j \leq k} \{f_e(m+1, \delta) - f_e(m, \delta)\}\right\}} \quad (\text{D.12})$$

---

<sup>27</sup>It is possible that player  $i$  pays intermediation benefits to  $j \in P$  if it can access some  $l \in C | g_{il} = 0$  better. The lower bound on  $M_j$  given in (D.10) is sufficient for the proof.

□

*Remarks on Proposition 4.* 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  $f_e(m, 0) = \frac{1}{3} \forall m$ , 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 Theorem 1.* By Lemma 1, after the move of player  $i = 1$  the network is either empty or a star. Obviously, in case it is empty it is stable, as all other nodes face the same decision as  $i = 1$ . If it is a star network and none of the other nodes wants to add a subset of links, this is the final stable outcome. Otherwise, by Lemmas 2 and 3, each next node  $i = 2, \dots, k$  adds links to all nodes not yet connected to  $i$ . There is a third and final possibility that the dynamic process is ended before the second round, namely if node  $i = n - 1$  fulfills the complete network by linking to node  $n$ . The conditions I, II and III for convergence to the empty, star or complete network coincide with the stability conditions of these three networks as in Proposition 2.

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 Proposition 3. 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:

$$M_i(g_{com}^{CP(k)}, -l) = l(c - \frac{1}{2} + f_m(n - l - 1, \delta)). \quad (\text{D.13})$$

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  $\text{I} \cup \text{II} \cup \text{III}$ , 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  $g_{2,n-2}^{mp(2)}$ .

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:

$$\begin{aligned} M_i(g_{com}^{CP(2)}, +(n-2)) &> 0; \\ \Rightarrow c &< \frac{1}{2} - f_e(2, \delta) + \frac{1}{2}(n-4)f_m(3, \delta), \end{aligned} \quad (D.14)$$

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

$$\begin{aligned} M_j(g_{com}^{CP(2)}, +(n-2)) &\geq 0; \\ \Rightarrow c &\leq \frac{1}{2} - f_e(2, \delta) + (n-4)(f_e(3, \delta) - f_e(2, \delta)); \end{aligned} \quad (D.15)$$

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 (compare equation (D.4)):

$$c \geq \frac{1}{2} - f_e(2, \delta) + (n-4) \min\left\{\frac{1}{2}f_m(3, \delta), f_e(3, \delta) - f_e(2, \delta)\right\}. \quad (D.16)$$

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  $0 < l_1^* \leq k - 1$  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:

$$c > \frac{1}{2} - f_e(n-k, \delta). \quad (D.17)$$

Given this high level of linking costs, the next core banks  $i = \{2, \dots, k\}$  have the same incentive to delete all their within-core links. The attracting network is therefore a complete multipartite network  $g_{k_1, k_2}^{mp(2)}$  with two groups of size  $k_1 = k$  and  $k_2 = n - k$ .

Finally, consider the case  $0 < l_1^* < k - 1$ .<sup>28</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:

$$c > \frac{1}{2} - f_e(n - k_1, \delta). \quad (\text{D.18})$$

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  $g_{ij}$ . 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,  $i = 1$  would 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 (D.18), 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  $K_2, \dots, K_{q-1}$  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  $g_{k_1, k_2, \dots, k_q}^{mp(q)}$  with  $q \geq 3$ .

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  $m, m' \in \{1, 2, \dots, q\}$ .  $\square$

## Appendix E. Proofs of Section 4

*Proof of Proposition 5.* 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. Hence Lemma 3 holds also in this heterogeneous setting: 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

---

<sup>28</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  $\binom{k}{l_1^*}$  best feasible actions deleting  $l_1^*$  links from the core. The resulting multipartite networks can consist of different sets  $K_1, K_2, \dots, K_{q-1}$ , but all are isomorphic.

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

$$\begin{aligned} M_i(g_{com}^{CP(k)}, +(n-k-1)) &> 0 \\ \Rightarrow c &< \frac{1}{2} - f_e(k, \delta) + \frac{1}{2}(n-k-2)f_m(k+1, \delta) \end{aligned} \quad (\text{E.1})$$

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

$$\begin{aligned} M_j(g_{com}^{CP(k)}, +(n-k-1)) &\geq 0 \\ \Rightarrow c &\leq \frac{1}{2} - f_e(k, \delta) + \frac{1}{2}(n-k-2) + (n-k-2)(f_e(k+1, \delta) - f_e(k, \delta)). \end{aligned} \quad (\text{E.2})$$

Conversely, the deviation of player  $i$  is *not* a best feasible action if  $c$  exceeds the minimum of the two values in equations (E.1) and (E.2):

$$c \geq \frac{1}{2} - f_e(k, \delta) + (n-k-2) \min\left\{\frac{1}{2}f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta)\right\}. \quad (\text{E.3})$$

Given a complete core-periphery networks and a sufficiently high  $c$  satisfying (E.3), it is not beneficial for periphery players to add links.

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

$$\begin{aligned} M_i(g_{com}^{CP(k)}, -(l^c + l^p)) &= l^c \left( c - \alpha^2 \left( \frac{1}{2} - f_e(n - l^c - l^p - 1, \delta) \right) \right) \\ &\quad + l^p \left( c - \alpha \left( \frac{1}{2} - f_e(k - l^c - 1, \delta) \right) - (2n - 2k - l^p - 1) f_m(k, \delta) \right) \\ &\equiv M_i^c + M_i^p \end{aligned} \quad (\text{E.4})$$

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:

$$M_i(g_{com}^{CP(k)}, -l^c) = l^c \left( c - \alpha^2 \left( \frac{1}{2} - f_e(n - l^c - 1, \delta) \right) \right) \quad (\text{E.5})$$

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

$$\begin{aligned} & M_i(g_{com}^{CP(k)}, -l^c) > 0 \\ \Rightarrow & M_i(g_{com}^{CP(k)}, -1) > 0 \\ \Leftrightarrow & \alpha < \sqrt{c / \left( \frac{1}{2} - f_e(n - 2, \delta) \right)} \end{aligned} \quad (\text{E.6})$$

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

$$M_i(g_{com}^{CP(k)}, -l^p) = l^p \left( c - \alpha \left( \frac{1}{2} - f_e(k - 1, \delta) \right) - (2n - 2k - l^p - 1) f_m(k, \delta) \right) \quad (\text{E.7})$$

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

$$\begin{aligned} & M_i(g_{com}^{CP(k)}, -l^p) > 0 \\ \Rightarrow & M_i(g_{com}^{CP(k)}, -(n - k)) > 0 \\ \Leftrightarrow & \alpha < \frac{c - \frac{1}{2}(n - k)(n - k - 1) f_m(k, \delta)}{\frac{1}{2} - f_e(k - 1, \delta)} \end{aligned} \quad (\text{E.8})$$

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:

$$\begin{aligned} & M_i(g_{com}^{CP(k)}, -l) = l \left( c - \alpha \left( \frac{1}{2} - f_e(k - l, \delta) \right) \right) - (n - k - 1) (f_e(k, \delta) - f_e(k - l, \delta)) > 0 \\ \Leftrightarrow & \alpha < \frac{c - \frac{n - k - 1}{l} (f_e(k, \delta) - f_e(k - l, \delta))}{\frac{1}{2} - f_e(k - l, \delta)} \end{aligned} \quad (\text{E.9})$$

The network is unilaterally stable if neither of these three deviations is beneficial, i.e. if



$\alpha$  exceeds all values given in equations (E.6), (E.8) and (E.9):

$$\bar{\alpha} = \max \left\{ \sqrt{c / \left( \frac{1}{2} - f_e(n-2, \delta) \right)}, \frac{c - \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta)}{\frac{1}{2} - f_e(k-1, \delta)}, \max_{l \leq k} \left( \frac{c - \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta))}{\frac{1}{2} - f_e(k-l, \delta)} \right) \right\} \quad (\text{E.10})$$

By rewriting 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:

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

□

*Remarks on Proposition 5.* Notice that, for given  $c$  and  $\delta$  and for  $n$  sufficiently large, the level of heterogeneity  $\bar{\alpha}$  as given in (E.10) equals

$$\sqrt{c / \left( \frac{1}{2} - f_e(k-1, \delta) \right)}. \quad (\text{E.11})$$

For such a value of  $\alpha$ , the complete core-periphery network with  $k$  big banks in the core is unilaterally stable if condition (E.3) is satisfied. To minimise  $\bar{\alpha}$  we take the smallest value of  $c$  satisfying (E.3). A lower bound for  $\bar{\alpha}$ , given values of  $\delta$ ,  $k$  and sufficiently large  $n$ , is thus given by:

$$\bar{\alpha} \geq \bar{\alpha}_{min} \equiv \sqrt{\frac{\frac{1}{2} - f_e(k, \delta) + (n-k-2) \min \left\{ \frac{1}{2} f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta) \right\}}{\frac{1}{2} - f_e(k-1, \delta)}}. \quad (\text{E.12})$$

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

$$\lim_{k \rightarrow \infty} \bar{\alpha}_{min} = \lim_{\delta \downarrow 0} \bar{\alpha}_{min} = \lim_{\delta \uparrow 1} \bar{\alpha}_{min} = 1, \quad (\text{E.13})$$

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

*Proof of Proposition 6.* Starting in an empty network, the first big bank  $i = 1$  has two relevant options: either connect to all players or connect only to big bank 2. Linking to only part of the small banks cannot be a best feasible action, because linking to an additional small bank pays off positive additional intermediation benefits (similar to the proof of Lemma 1). Player 1's payoffs depending on its action  $S$  are:

$$\pi_1(S) = \begin{cases} 0 & \text{if } S = \emptyset \\ \frac{1}{2}\alpha^2 - c & \text{if } S = \{2\} \\ \frac{1}{2}\alpha^2 - c + 2(\frac{1}{2}\alpha - c) + 2\frac{1}{3}\alpha + \frac{1}{3} \\ = \frac{1}{2}\alpha^2 + \frac{5}{3}\alpha + \frac{1}{3} - 3c & \text{if } S = \{2, 3, 4\} \end{cases} \quad (\text{E.14})$$

Under condition I, player 1's best feasible action is not to add any links. The resulting network is empty. In this case the network must be unilaterally stable. The reason is that second big bank faces the same decision as  $i = 1$ , and small banks have strictly lower payoffs from adding links, so no player will decide to change the network structure.

Under condition VI, the best feasible action is to add only one link to big bank 2. The second big can choose from the same resulting networks as  $i = 1$  could, so does not add or remove links. Small banks have strictly lower payoffs from adding links and also do not change the structure. So under condition VI the resulting network with only one link, namely  $g_{12} = 1$ , is stable.

If the first player adds links to all three other players, a star network is formed. Then the best feasible action for the second big bank  $i = 2$  is either to do nothing or to add links to both periphery players. If 2 does nothing, the periphery nodes 3 and 4 will likewise decide not to add links, and the star network is the final, stable outcome. For 2 to have positive marginal benefits of supporting two links, it is required that

$$\begin{aligned} M_2(g^s, +2) &> 0 \\ \Rightarrow c &< \frac{1}{6}\alpha + \frac{1}{2}f_m(2, \delta), \end{aligned} \quad (\text{E.15})$$

and each peripheral player  $j \in \{3, 4\}$  requires

$$\begin{aligned} M_j(g^s, +2) &\geq 0 \\ \Rightarrow c &\leq \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3}. \end{aligned} \quad (\text{E.16})$$

Conversely, the star network is stable if the deviation of player 2 is *not* beneficial to 2 and/or a peripheral player  $j$ , i.e. if  $c$  exceeds the minimum of the two values in equations

(E.15) and (E.16):

$$c \geq \frac{1}{6}\alpha + \min\left\{\frac{1}{2}f_m(2, \delta), f_e(2, \delta) - \frac{1}{3}\right\}, \quad (\text{E.17})$$

leading to condition II.

Consider the case that both big banks have added all possible links. Then there is a possibility that the dynamic process leads to a complete network if node 3 links to node 4. This happens if  $c$  is so small that the share  $f_e(2, \delta)$  from intermediated trade between 3 and 4 can be raised to  $\frac{1}{2}$  by creating a direct link. The complete network is therefore stable under condition III.

For parameters not satisfying I, II, III or VI, the first round of best feasible actions results in a complete core-periphery network with 2 core members. This complete core-periphery network can be stable because of the heterogeneity between core and periphery banks with  $\alpha > 1$ . As stated in proposition 5, this occurs under condition VII. For  $n = 4$  and  $k = 2$  condition VII reduces to:

$$\begin{aligned} c \in & \left( \frac{1}{2} - f_e(k, \delta) + (n-k-2) \min\left\{\frac{1}{2}f_m(k+1, \delta), f_e(k+1, \delta) - f_e(k, \delta)\right\}, \right. \\ & \min\left\{ \alpha^2\left(\frac{1}{2} - f_e(n-2, \delta)\right), \alpha\left(\frac{1}{2} - f_e(k-1, \delta)\right) + \frac{1}{2}(n-k)(n-k-1)f_m(k, \delta), \right. \\ & \left. \left. \min_{l \leq k} \left\{ \alpha\left(\frac{1}{2} - f_e(k-l, \delta)\right) + \frac{n-k-1}{l}(f_e(k, \delta) - f_e(k-l, \delta)) \right\} \right\} \right) \\ \in & \left( \frac{1}{2} - f_e(2, \delta), \min\left\{\alpha^2\left(\frac{1}{2} - f_e(2, \delta)\right), \frac{1}{6}\alpha + f_m(2, \delta), \frac{1}{6}\alpha + f_e(2, \delta) - \frac{1}{3}\right\} \right). \end{aligned}$$

Notice that when this condition is fulfilled, the second player always adds the two links to the periphery (cf. condition (E.17)).

Finally, in the remaining region IV, the complete core-periphery network cannot be stable. Because the node  $i = 3$  node did not connect to the last periphery node, adding links cannot be a best feasible action. Therefore the first core bank  $i = 1$  must have an incentive to delete the link with 2. The attracting steady state is the multipartite (ring) network consisting of the groups  $\{1, 2\}$  and  $\{3, 4\}$ .  $\square$

## Appendix F. Longer intermediation chains

In this appendix we generalise the model to allow for intermediation paths of lengths longer than two. We will analyze the general model for lengths up to distance three and under homogeneity (i.e.  $\alpha_i = 1$  for all  $i, j$ ), and check whether a (incomplete) core-periphery network may be stable.

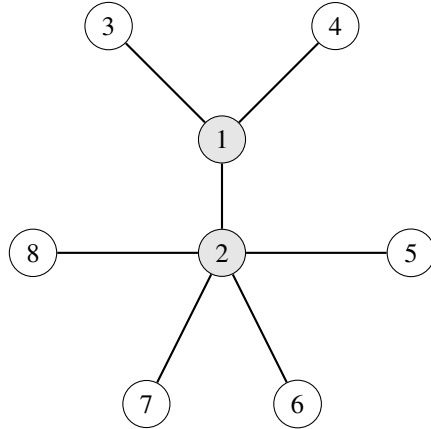
Before rewriting a more general form than the payoff function in Proposition 1 formally, we repeat that the proofs of Proposition 3 and 4 do not require any assumptions on the path lengths on which trade is allowed; see the remarks in Appendix D on these propositions. Proposition 3 states that complete core-periphery networks are not pairwise (or unilaterally) stable. Proposition 4 states that incomplete core-periphery networks are not unilaterally stable if  $n$  is sufficiently large. Also Proposition 5, specifying a level of heterogeneity sufficient for a complete core-periphery network to be unilaterally stable, holds for longer intermediation paths.

We introduce new, more general notation for intermediation over longer path lengths:  $F_e(g, \{i, j\}, \delta)$  denote the shares for the endnodes in the pair  $i$  and  $j$ ; and  $F_m(g, \{i, j, k\}, \delta)$  denotes how much middleman  $k$  receives. In Appendix A it is explained how such a distribution can be derived for long intermediation chains in the example of Siedlarek (2015). Note that if  $i$  and  $j$  at length three are assumed to generate a surplus, middlemen involved in a trade between  $i$  and  $j$  do not necessarily earn the same: if  $k$  lies on more of the shortest paths than  $k'$ ,  $k$  will earn more than  $k'$ . For this reason (part of) the graph  $g$  must be given as an argument in the function  $F_e$  and  $F_m$ . The payoff function becomes:

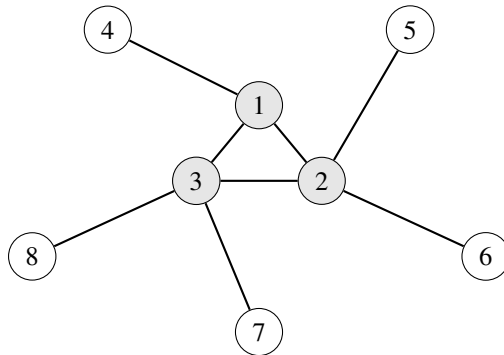
$$\pi_i(g) = \eta_i(g) \left( \frac{1}{2} - c \right) + \sum_{j \in N_i^3(g)} F_e(g, \{i, j\}, \delta) + \sum_{\substack{k, l \in N_i^2(g) \\ g_{kl} = 0, d_{kl} \leq 3}} F_m(g, \{k, l, i\}, \delta), \quad (\text{F.1})$$

where  $N_i^r(g)$  denotes the set of nodes at distance  $r$  from  $i$  in network  $g$ ,  $\eta_i(g) = |N_i^1(g)|$  the number of direct connections of  $i$ , and  $d_{kl}$  the distance between nodes  $k$  and  $l$ .

In incomplete core-periphery networks, shortest paths of three may exist between some periphery players, which were previously assumed not to generate any trading surplus. By allowing intermediation chains of three we found that some core-periphery networks can become unilaterally stable. For  $n = 8$ , we found that  $k = 2$  and  $k = 3$  are the only possibly stable core sizes. See Figure F.14 for two examples of networks that are stable for the given parameter values when paths of three are allowed. These two network structures are not stable in the baseline model.



(a)  $(\delta, c) = (0.8, 1.1)$ . A minimally connected core-periphery network with  $k = 2$ .



(b)  $(\delta, c) = (0.5, 0.9)$ . A minimally connected core-periphery network with  $k = 3$ .

**Figure F.14:** Examples of unilaterally stable core-periphery networks after allowing for intermediation chains of length 3, for  $n = 8$  players and given  $(\delta, c)$ .

We can safely interpret these examples as low-dimensional exceptions to the rule that the core-periphery structure in homogeneous networks is generally unstable. The examples in Figure F.14 show that for  $n = 8$  incomplete core-periphery networks with core sizes of  $k = 2$  and  $k = 3$  can be stable. For  $n$  sufficiently large, however, core-periphery networks are always unstable as stated by Proposition 4.

Moreover, even though exceptionally for small  $n$  core-periphery networks can be unilaterally stable, they are never the outcome of a dynamic process as described in Section 3.2. For  $n = 8$ , the core-periphery networks with  $k = 2$  and  $k = 3$  were found to be stable in parameter regions where the star networks is stable as well, cf. region II in Figure 6b. Exploring the parameter space by simulations, we found that this was always the case for such stable core-periphery networks. By Lemma 1, the star is created as a first step in the dynamic process whenever the initial empty network is not stable. Therefore the star network is the outcome of a dynamic process even when exceptional (low-dimensional) networks are unilaterally stable as well given the parameter values. This implies that the dynamic results of Theorem 1 do not depend on the assumption of maximal intermediation paths of length two, as was already shown for the static results of Propositions 3, 4 and 5.