

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

# Delay in Trade Networks

Thành Nguyen

Krannert School of Management, Purdue University, West Lafayette, IN 60208, nguyet161@purdue.edu

Vijay Subramanian

EECS Department, University of Michigan, Ann Arbor, MI 48109, vgsubram@umich.edu

Randall Berry

EECS Department, Northwestern University, Evanston IL 60201, rberry@eecs.northwestern.edu

We study decentralized markets involving producers and consumers that are facilitated by middlemen. We do this by analyzing a non-cooperative networked bargaining game. We assume a complete information set-up wherein all the agents know the structure of the network, the values of the consumers and the transaction costs involved but allow for some search friction when either producers or consumers trade with middlemen. In such a setting, we show that sunk cost problems and a heterogeneous network can give rise to delay or failure in negotiation, and therefore, reduce the total trade capacity of the network. In the limiting regime of extremely patient agents, we provide a sharp characterization of the trade pattern and the segmentation of these markets.

*Key words:* Noncooperative Bargaining, Supply Chain Networks, Trade Volume

---

## 1. Introduction

Trade delay is a wide-spread phenomenon of markets. In several manufacturing industries, for example, retailers often postpone contracting with certain suppliers to search for cheaper and more reliable sources. The real estate market is another well-known example where sellers (vacant house owners) and buyers may coexist but may fail to negotiate a trade. Such delay, on the one hand, might allow agents to get more information and find better trading partners, but on the other hand, it slows down economic transactions and influences the overall trade volume and liquidity of

the market, typically leading to a reduction in both. Studying the causes of trade delay, therefore, has been an important question in economics and finance as well as in operations management.

A non-cooperative bargaining model between buyers and sellers is perhaps the most common framework to study trade delay in the literature. When sellers and buyers have complete information, the seminal work of Rubinstein (1982) shows that agents will reach an agreement immediately to avoid the surplus loss due to delay. Prior literature has extended Rubinstein (1982) in several directions to show that in some cases trading partners may choose not to reach an agreement and so strategically delay trade. (See the discussion on related literature in Section 1.1.) However, because of complexity, the literature has primarily focused on simple settings that often abstract away other aspects of markets. In this paper, we consider two important such aspects: middlemen and a heterogeneous interconnection network between producers and consumers involving these middlemen. Studying middlemen is important because, in most markets, trade does not involve just sellers and buyers but also includes one or more middlemen serving as intermediaries. For example, brokers and market makers fill this role in financial markets as do wholesalers and retailers in many manufacturing industries. Similarly, there is a wide variety of interconnection networks seen in real-world markets. The question we ask in this paper is the following: can middlemen and the complex network structure that connects them with sellers and buyers give rise to strategic delay? If yes, what are the properties of the network that lead to such delay?

Our paper provides answers to these questions. We consider a bargaining game with complete information but enrich the environment by adding middlemen, a heterogeneous network and some search friction when either producers or consumers trade with middlemen. Since we exclude information frictions, our results show that, in this setting, middlemen and heterogeneous networks are the sources of strategic delay. After arguing that delays in trade lead to inefficiency in terms of a reduction in the trade volume, we consider the limiting regime of extremely patient agents and provide a sharp characterization of networks where strategic delay emerges based on their *directed cuts*. Essentially, this result shows that a trading network, involving producers, middlemen and

consumers, gives rise to trade delay if there exists a partition of the network into two disjoint set of nodes,  $\mathcal{V}_1$  and  $\mathcal{V}_2$ , where goods can only be traded from  $\mathcal{V}_1$  to  $\mathcal{V}_2$ , i.e., a directed cut, such that

$$\frac{(\text{total trade surplus} - \text{total sunk costs}) \text{ of } \mathcal{V}_1}{\text{size of } \mathcal{V}_1} > \frac{(\text{total trade surplus} - \text{total sunk costs}) \text{ of } \mathcal{V}_2}{\text{size of } \mathcal{V}_2}.$$

We formally define these terms later in the paper, but intuitively, our results highlight two new features in networked bargaining, which are explained next.

First, because goods are traded from sellers (producers) to middlemen and then from middlemen to buyers (consumers), there is a sunk cost problem. Namely when middlemen bargain with consumers, previous transaction costs between producers and middlemen have been sunk, which makes middlemen more conservative in trade. Several papers, such as Wright and Wong (2014) and Nguyen (2014), have shown similar effects in bargaining. However, these papers show that in equilibrium trade either occurs at the maximum possible volume or there is no trade at all. Here we show that there is a *continuous* decrease in trade volume when the sunk cost increases. This is because we consider a large economy with search friction. In particular, if middlemen delay trade with producers, then the fraction of middlemen holding goods to trade with consumers becomes small. This enables middlemen to charge a higher price to consumers, and therefore they can partially overcome the sunk cost problem. This insight is illustrated in an example in Section 5.1.

The second insight in our paper is the fact that network structure can have an impact on strategic delays and therefore influences trade volume. To see this, we provide an example in Section 5.2 where delay occurs on a subset of the links of the network as the discount factor increases to 1; we isolate the impact of network structure by setting the transaction costs to zero in this example.

The intuition for both these insights can be seen from the cut condition above. In particular, the ratios in that inequality can be interpreted as the strength of each of the submarkets,  $\mathcal{V}_1$  and  $\mathcal{V}_2$ . The strength of a submarket measures the combined effect of sunk cost and network structure. Agents in a stronger submarket can trade faster and often get a higher surplus. Owing to this, when the submarket  $\mathcal{V}_1$  is stronger than  $\mathcal{V}_2$ , agents in  $\mathcal{V}_1$  will hesitate to *sell* to agents in  $\mathcal{V}_2$ . This leads to market segmentation because of delay and failure in negotiation across the two submarkets.

Section 1.1 below discusses the connection with the literature. Section 2 gives the formal bargaining model, and Section 3 lays out the solution concepts. Section 4 shows some basic properties of an equilibrium, which we use in later sections. Section 5 illustrates the main findings with two examples. Section 6 analyzes general networks and derives the main structural conditions for trade delay. Section 7 concludes. Technical proofs are given in the Appendix.

### 1.1. Related literature

Our paper is closely related to the recent literature on networked bargaining: Rubinstein and Wolinsky (1987b), Wright and Wong (2014), Gofman (2011), Manea (2011), Nguyen (2014, 2015). The question we ask in this paper is, however, different. Our goal is to understand what network structures lead to trade delay. In addition, the model we use in this paper has several important differences with these papers. Specifically, Rubinstein and Wolinsky (1987b) considers bargaining with middlemen in a 2-link network and assumes agents always trade. Unlike our paper, Wright and Wong (2014) and Gofman (2011) consider an environment with single item to trade. As a result, in these models, at equilibrium agents will either trade right away or they do not trade at all. Manea (2011) investigates a bargaining model without middlemen. Nguyen (2015) generalizes Manea (2011) to a coalitional bargaining game, that can model middlemen. However, unlike the present paper, in Nguyen (2015) all agents are short-lived, and therefore there is no sunk cost problem. Nguyen (2014) analyzes a local bargaining model in a *chain*, while our current paper consider a network with heterogeneous connections among producers, middlemen and consumers. A more fundamental difference between our model and Nguyen (2014) is that here in every period we choose a set of pairs of agents to trade with the effects of search friction modeled (so that if an agent is not matched with the right partner, trade is not possible), while in Nguyen (2014) a single agent is selected and that agent can always find a feasible trading partner. This causes a discontinuous change in the equilibrium of Nguyen (2014): limit stationary equilibrium might not always exist. Here, due to search friction, we show that a stationary equilibrium always exists, and

when a unique equilibrium exists, then it changes continuously as the parameters of the model change.

Our paper is also related to a different line of non-cooperative bargaining literature that investigates strategic delay. Since the seminal work of Rubinstein (1982), there have been multiple studies extending Rubinstein's bargaining games to capture delay in trade. This literature considers many modifications of Rubinstein (1982) in terms of information structure and strategy space. These modifications include information asymmetry (Gul and Sonnenschein (1988), Admati and Perry (1987) and Cho (1990)); the possibility of simultaneous multiple offers (Sakovics (1993)); constraints on reaction time (Perry and Reny (1993)); and reputation effects (Atakan and Ekmekci (2013)). Another way to extend Rubinstein (1982) to study strategic delay is to consider multilateral bargaining and the possibility of externalities. It is shown, for example, in Jehiel and Moldovanu (1995), that delay will occur if the bargaining outcome of a group of agents affects an outside member. In contrast with this literature, our paper adapts the simple bilateral bargaining protocol of Rubinstein and Wolinsky (1987a), assumes a complete information structure, allows at most one bargaining opportunity for an agent and does not include any reputation effects. Thus, it shuts down the channels discussed in the literature above to isolate the effects of middlemen and network structure.

There is also another line of literature concerning middlemen: Li (1998), Biglaiser (1993), Yavaş (1994), Johri and Leach (2002) and Masters (2007), just to name a few. The focus of this literature is on the following question: "why do middlemen exist?" This literature provides multiple answers ranging from economy of scale, economy of scope to advantages in information and inventory, or middlemen existing simply because of physical and institutional constraints. Our paper follows the recent literature on network games by studying middlemen from a different perspective. Namely, we assume that middlemen exist and that they form networks with buyers and sellers for a myriad of reasons, but knowing these networks, we seek to determine the trade patterns and the economy's over-all efficiency.

The literature on network games is too large for an extensive survey here. However, Blume et al. (2009) is the closest work from this broad literature to the current paper. In fact, the use of middlemen networks in our paper is borrowed from Blume et al. (2009). However, Blume et al. (2009) assumes that middlemen have full bargaining power on the prices charged to sellers and buyers (who then use prices to determine their trade partners resulting in Bertrand competition). The market, therefore, is frictionless, and all equilibria in Blume et al. (2009) are efficient. In our model no agent has full bargaining power, and our results highlight the market friction caused by network structures and bargaining incentives.

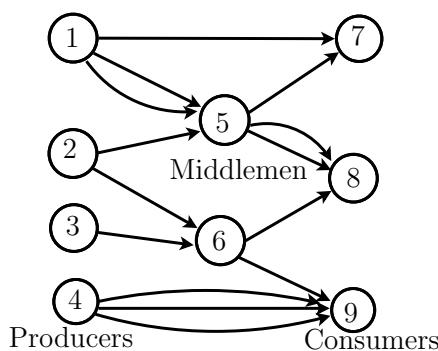
## 2. The Model

In this section we introduce the model that we will use. We start by defining the concept of a trading network.

### Trading Network

We consider a group of producers, consumers and middlemen interconnected by an underlying trading network, which is modeled as a directed (multi)graph,  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  (see Figure 1). Without loss of generality, we assume that the network is connected when the direction of the links are ignored, i.e., it cannot be partitioned into two set of nodes that are not connected by any link. Each node  $i \in \mathcal{V}$  represents a population of an  $N_i$  mass of agents, all of which are either consumers, producers or middlemen. Hence, we can partition the set of vertices into the following three disjoint sets: a set of producers denoted by  $\mathcal{P}$ , a set of middlemen denoted by  $\mathcal{M}$ , and a set of consumers denoted by  $\mathcal{C}$ . An agent from the population at a node  $i$  will sometimes be referred to as a type  $i$  agent. Trade occurs over directed edges, i.e., a directed edge  $(i, j) \in \mathcal{E}$  indicates that a type  $i$  agent can potentially directly trade with any type  $j$  agent with the good trading hands from  $i$  to  $j$  as a result of the trade. With a slight abuse of terminology, we often refer to two such agents as being connected by the edge  $(i, j)$ . We will allow for multiple links between  $i$  and  $j$ . The multiplicity of a link can be seen as a capacity between the two ends. This will be described precisely later when we specify the bargaining process.

For a consumer to acquire a good from a producer, there must be a (directed) path from the consumer to the producer. If this path has length 1, then the two can directly trade; otherwise, they must rely on middlemen to facilitate the trade. As mentioned earlier, for simplicity, we consider networks in which any path between a consumer and producer contains at most one middleman, i.e., all such paths are either length 1 or 2. An example of such a network is shown in Figure 1. With this assumption, the set of directed edges,  $\mathcal{E}$ , can also be partitioned into three disjoint sets: those that directly connect producers to consumers (denoted by  $\mathcal{E}_1$ ), those that connect producers to middlemen (denoted by  $\mathcal{E}_2$ ), and those that connect middlemen to consumers (denoted by  $\mathcal{E}_3$ ).



**Figure 1** A network among producers, consumers and middlemen.

We assume that there is one type of indivisible good in this economy. All producers produce identical goods (with the same production costs, although the model can be easily extended to a setting where production costs are based on the type of producer) and all consumers want to acquire these goods. The value that each consumer of type  $c \in \mathcal{C}$  gets from an item is  $V_c \geq 0$ , which can be different for different consumers. In every period, each agent can hold at most one unit of the good (an item). Thus, in every time period, a middleman either has an item or does not have one. Hence, if there is a directed edge from node  $i$  to node  $j$ , a specific agent of type  $i$  can only trade with an agent of type  $j$  if the type  $i$  agent has a copy of the good and the type  $j$  agent does not; we refer to such a pair of agents as *feasible trading partners*. Note that producers are assumed to always have a good available to trade and consumers are always willing to purchase a good.

So, for example, any two agents connected by an edge in the set  $\mathcal{E}_1$  are always feasible trading partners. For every edge  $(i, j) \in \mathcal{E}$ , we associate a non-negative transaction cost  $C_{ij} \geq 0$ ; this cost is incurred when trade occurs between an agent at node  $i$  and one at node  $j$ , which, in the case of multiple links, can be different on different links that connect a pair of nodes  $i$  and  $j$ .

Next we discuss the bargaining process that determines the trading patterns for how goods move through the network.

### Bargaining Process

We assume in every period there is a matching function that matches pairs of feasible trading partners that are connected in the network. In particular, we assume that in every period, on each link of the network, there is one unit mass of pairs of agents, selected at random, who meet one another. If there are  $k$  parallel links between  $i$  and  $j$ , then  $k$  unit mass of such pairs of agents will be selected. We assume that no agent is selected on more than one link. Here, in order to guarantee the existence of such a matching process, we assume that the population mass at each node  $i$ ,  $N_i$ , is at least its degree (sum of incoming and outgoing links) in the network.

For some pairs of agents, trade is not possible because the agents are either both empty handed or both hold a good. Some pairs of agents, on the other hand, are feasible trading partners. For such a pair, one of the agents is selected at random to be the proposer. The proposer makes a take-it-or-leave-it offer of a price at which he is willing to trade. The trading partner at the other end of the proposal can either refuse or accept the offer. If she accepts, the two agents will trade: the agent holding the item gives it to the agent desiring it and receives the payment, with the proposer paying the transaction cost  $C_{ij}$ .

If a consumer or producer participates in a trade, after trade they exit the game and are replaced by a clone. On the other hand, middlemen are long lived and do not produce nor consume; they earn money solely by flipping the good.

Let  $0 < \delta < 1$  be the discount rate used by the agents for valuing their returns in the game. The game is denoted by  $\Gamma(\mathcal{G}, C, V, N, \delta)$ , where  $C$  denotes the vector of links costs,  $V$  denotes the vector



of consumer valuations, and  $N$  denotes the vector of population sizes at each node. Sometimes, we will simply refer to this game as  $\Gamma$ .

### **Discussion of the model**

We assume middlemen are long lived, and on the other hand, producers and consumers exit the game after trading. This captures an extreme contrast between different types of agents: middlemen often stay in the market for a long period, while producers and consumers have limited supply and demand for a certain good. This is a reasonable scenario in many markets among small producers and consumers, who are faced with search problems and need to trade through middlemen (farmers, consumers and grocery stores; sellers, buyers and e-commerce companies like Ebay or Amazon; investors, borrowers and banks.)

We follow Rubinstein and Wolinsky (1987a), Gale (1987) and Manea (2011) in assuming the replacement of producers and consumers to capture a steady-state of the economy. As Manea (2011) explains: “The steady state assumption captures the idea that in many markets agents face stationary distributions of bargaining opportunities. In such cases, some agents take similar positions in relationships and transactions at different points in time.” Similar to these papers, our approach focuses on the influence of the heterogeneous network structure on the stationary outcome. Of course, it abstracts away the convergence of markets to such stationary equilibria, which raises more complex issues beyond the scope of our paper. On the other hand, as Gale (1987) points out this stationary assumption can be made endogenous in an extended economy, where there are incoming flows of producers and consumers at certain rates. To see this, assume each node of producers and consumers has an incoming flow of agents, who also has an outside option with certain payoff. In models like ours, if the population of a node increases, because of competition, the expected payoff decreases, eventually leading agents to instead take the outside option. Likewise, if the outside option is smaller than payoff received, then agents will move in to the node and the population will increase. It can be argued that this process will stabilize and will

lead to an inflow rate of goods at each producer that is equal to the rate of successful trade at that node, resulting in the replacement assumption of our model.

Besides the tractability of the model, another reason to use the steady-state assumption is because of its practicality. In real-world applications, most economic markets are in flow (agents come and leave the market), and we can only observe some snapshot of the market, maybe by taking some particular average. Therefore, the analysis of these markets assumes that the sampled snapshot is the steady-state of the economy, and then proceeds to make inferences using it.

### 3. Stationary Equilibrium

Next we turn to the solution concept considered in this paper. We will use the concept of stationary equilibrium. We start by making a few important definitions.

**DEFINITION 1.** The **state of the economy** is a vector  $\mu \in [0, 1]^{|M|}$ , where  $\mu_m$  denotes the fraction of middlemen at node  $m$  that hold an item.

**DEFINITION 2.** A strategy profile (possibly mixed strategy) is called a **stationary strategy** if it only depends on an agent's identity, his state (owning or not owning an item) and the play of the game (to whom the agent is matched with, who the proposer is and what is proposed). More precisely, suppose that agent  $i$  and agent  $j$  are selected to bargain, and assume  $i$  owns an item,  $j$  does not, and furthermore  $i$  is the proposer. In this case, a stationary strategy of agent  $i$  is a distribution of proposed prices to agent  $j$  and a stationary strategy of agent  $j$  is a probability of accepting the offer.

We are especially interested in the strategy in which all agents immediately trade whenever they meet feasible trading partners. We call such a strategy the “always-trade” strategy.

**DEFINITION 3.** A stationary strategy is called **always-trade** if whenever an agent  $i$  owning an item meets an agent  $j$  not owning an item, and the pair  $(i, j) \in \mathcal{E}$ , then  $i$  and  $j$  trade with probability 1, i.e, the proposer's offer is accepted with probability 1 by his trading partner.

In general, an always-trade strategy might not be the outcome of the game as some agents might strategically delay trade. To account for the possibility of strategic delays, in the rest of the paper, given a stationary strategy, we let  $\lambda_{ij}$  (for each link  $(i, j) \in \mathcal{E}$ ) denote the conditional probability that  $i$  and  $j$  trade when they are matched and trade is feasible (i.e.,  $i$  owns an item and  $j$  does not); note that an always-trade strategy corresponds to  $\lambda_{ij} \equiv 1$  for all  $(i, j) \in \mathcal{E}$ .

As is standard in the literature, to define a stationary equilibrium, we first derive necessary conditions on the two-tuple of a stationary strategy profile and state profile if they are to constitute an equilibrium; we divide these into two distinct sets: incentive constraints and stationary constraints. We will show that these conditions are also sufficient for an equilibrium. We start with the incentive constraints.

### Incentive Constraints

To define these constraints, we first introduce the expected pay-offs of agent  $i$  depending on whether he possesses or does not possess a good, which we denote by  $u_0(i)$  and  $u_1(i)$ , respectively. In a stationary equilibrium, agents believe that the state of the economy is captured by  $\mu$ . For this subsection, we will assume that  $\mu$  is given. After deriving the incentive conditions depending on  $\mu$ , we will discuss the second type of conditions in which  $\mu$  is determined by the stationary strategy profile.

The basic structure of the incentive constraints can be captured by the following argument. Assume two agents  $i$  and  $j$  meet on a given edge of  $\mathcal{E}$ , where  $i$  holds the good and  $j$  wants it. Also assume that  $i$  is the proposer. If the trade is successfully completed, then in the next period  $j$  possesses the item; if trade is not successful, then  $j$  is in the same state and his payoff is discounted. Thus, in equilibrium, agent  $i$  will demand from agent  $j$  the difference of the payoffs between these two scenarios. Say the price is  $\bar{p}_{ij}$ , then we have  $\delta u_0(j) = -\bar{p}_{ij} + \delta u_1(j)$  leading to  $\bar{p}_{ij} = \delta(u_1(j) - u_0(j))$ . Note that the state of  $i$  also changes upon completion of trade. Therefore, if trade is successfully completed, then  $i$ 's payoff is

$$\delta u_0(i) + \delta(u_1(j) - u_0(j)) - C_{ij}.$$

However, agent  $i$  has the option of not proposing a trade (or proposing something that will necessarily be rejected by the other party) and earning a payoff of  $\delta u_1(i)$ . Thus, in this situation, the continuation payoff of agent  $i$  is

$$\max\{\delta u_1(i), \delta u_0(i) + \delta(u_1(j) - u_0(j)) - C_{ij}\}.$$

It is important to note that, in an equilibrium, the continuation payoff of agent  $j$ , who is matched for trade and needs a good but is not a proposer, continues to be  $\delta u_0(j)$  irrespective of whether trade happens or not. Therefore, only the proposer of a trade stands to gain from the trade.

For ease of exposition we let  $z_{ij}$  be the **relative gain of trade** from agent  $i$  to agent  $j$ , which is the gain (or loss) seen by the proposer for trading compared to that for not trading. This is simply the difference between the two terms in the previous maximization, i.e.,

$$z_{ij} := \delta\left((u_1(j) - u_0(j)) - (u_1(i) - u_0(i))\right) - C_{ij}. \quad (1)$$

The continuation payoff of agent  $i$  when he is proposing to  $j$  is then  $\delta u_1(i) + \max\{z_{ij}, 0\}$ . From this we also obtain the following conditions on the dynamics of trade:

1. If  $z_{ij} < 0$ , then agent  $i$  will never sell an item to agent  $j$  and will wait for a future trade opportunity;
2. If  $z_{ij} > 0$ , then agent  $i$  will sell the item to agent  $j$  with probability one whenever they are matched; and finally
3. If  $z_{ij} = 0$ , then agent  $i$  is indifferent between selling and waiting, thus, the trade can occur with some probability  $\lambda_{ij} \in [0, 1]$ . Conversely, if trade between agents  $i$  and  $j$  occurs with probability  $0 < \lambda_{ij} < 1$ , then we must have  $z_{ij} = 0$ .

Note that in the third case ( $z_{ij} > 0$ ) though  $j$  is indifferent between trading or not when  $i$  proposes  $p_{ij}$ , in any equilibrium it must agree with probability one. This is because if  $j$  only agrees with a probability  $0 < p < 1$ ,  $i$  can improve his payoff by decreasing the proposing price by a small  $\epsilon > 0$ , in which case  $j$  would no longer be indifferent and accept with probability one. However,

for any such  $\epsilon > 0$ ,  $i$  again has a better deviation by decreasing the proposing price by a smaller amount, say  $\epsilon/2$ .

Similarly, assume now that instead of  $i$ , agent  $j$  is the proposer, then the continuation payoff of  $j$  in this case is  $\delta u_0(j) + \max\{z_{ij}, 0\}$ . Note that when  $j$  is the proposer, again the relative gain of trade is given by  $z_{ij}$ , i.e., this quantity depends on the direction of trade not on which agent is the proposer. Furthermore, the same conditions concerning the dynamic of trade between  $i$  and  $j$ , depending on  $z_{ij}$ , hold as above.

These conditions can be delineated for the general network model introduced in the previous section by considering each link, type of agent, the state of the agent in terms of holding a good or not, and the probability that she is selected as a proposer. In our model, we have three types of agents: producers, consumers and middlemen. Furthermore, middlemen can be separated into two types: those that have the item and those that do not. Thus, we will need four types of equations expressing the expected payoff of these agents given their states.

We consider these conditions in detail for the case of producers; the other types follow in a similar fashion. For each producer of type  $p \in \mathcal{P}$  who has an item to sell in each period, their continuation payoff depends on which type of link is selected, the pair of agents that are selected to trade, and whether  $p$  is selected as the proposer. Thus, agent  $p$ 's expected continuation payoff is

$$\begin{aligned} \sum_{c:(p,c) \in \mathcal{E}_1} \frac{1}{2N_p} (\delta u_1(p) + \max\{z_{pc}, 0\}) + \sum_{m:(p,m) \in \mathcal{E}_2} \frac{1}{2N_p} (1 - \mu_m) (\delta u_1(p) + \max\{z_{pm}, 0\}) + \quad (2) \\ + \left( 1 - \sum_{c:(p,c) \in \mathcal{E}_1} \frac{1}{2N_p} - \sum_{m:(p,m) \in \mathcal{E}_2} \frac{1}{2N_p} (1 - \mu_m) \right) \delta u_1(p). \end{aligned}$$

Furthermore,  $z_{pc}$  and  $z_{pm}$  are defined as in (1). The first term of (2) represents the case where  $p$  is the proposer to a consumer  $c$ . The second term represents  $p$  proposing to a middlemen  $m$ , who currently does not own a good. Finally, the last term describes the case where  $p$  is not a proposer. Note that in the summations in (2), in the case there are parallel links between a pair of nodes, each link will correspond to a separate term in the sum.<sup>1</sup>

Since in every period, one unit mass of pairs of agents is selected on each link, it follows that  $\frac{1}{N_p}$  is the probability that the specific agent of type  $p$  is chosen in the matching process. Once an agent

$p$  is selected, he will be the proposer with probability  $1/2$ . Hence,  $\frac{1}{2N_p}$  is the *ex-ante* probability that an agent  $p$  becomes the proposer for a consumer of type  $c$ . On the other hand, because only a fraction of middlemen are looking to buy, for every middlemen node  $m \in \mathcal{M}$ ,  $\frac{1}{2N_p}(1 - \mu_m)$  is the probability that  $p$  is matched with  $m$ ,  $m$  does not hold a good, and  $p$  is the proposer. One can interpret this as a form of *search friction*, that is the probability that a producer can find a trade-able middlemen depends on the state of the economy, which, in turn, impacts the transaction dynamics between the producer and the middleman.

Now, because  $u$  are assumed to be values of a stationary equilibrium,  $u_1(p)$  needs to equal the expression in (2). After some algebraic manipulation, this is equivalent to

$$u_1(p) = \sum_{c:(p,c) \in \mathcal{E}_1} \frac{\max\{z_{pc}, 0\}}{2N_p(1-\delta)} + \sum_{m:(p,m) \in \mathcal{E}_2} \frac{(1-\mu_m)\max\{z_{pm}, 0\}}{2N_p(1-\delta)}. \quad (3)$$

Similarly, for the two type of middlemen (either owning an item or not) and the consumers, we have the following set of equations:

$$\forall m \in \mathcal{M}, \quad u_0(m) = \sum_{p:(p,m) \in \mathcal{E}_2} \frac{\max\{z_{pm}, 0\}}{2N_m(1-\delta)}, \quad (4)$$

$$\forall m \in \mathcal{M}, \quad u_1(m) = \sum_{c:(m,c) \in \mathcal{E}_3} \frac{\max\{z_{mc}, 0\}}{2N_m(1-\delta)}, \quad (5)$$

$$\forall c \in \mathcal{C}, \quad u_0(c) = \sum_{p:(p,c) \in \mathcal{E}_1} \frac{\max\{z_{pc}, 0\}}{2N_c(1-\delta)} + \sum_{m:(m,c) \in \mathcal{E}_3} \frac{\mu_m \max\{z_{mc}, 0\}}{2N_c(1-\delta)}, \quad (6)$$

where  $z_{pm}, z_{mc}, z_{pc}$  are defined as

$$z_{ij} = \delta \left( u_1(j) - u_0(j) - (u_1(i) - u_0(i)) \right) - C_{ij}, \quad \forall (i, j) \in \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3. \quad (7)$$

Once again, because of search friction, the state  $\mu$  appears in the incentive equations above for the consumers, owing to the particular search model that we consider. However, for the middlemen, (4) and (5) do not involve  $\mu$ . This is because in our model, the producers always have an item to sell and the consumers can always consume; thus, middlemen are not faced with search friction.

Since producers and consumers exit the game after trading successfully, we have

$$\forall p \in \mathcal{P}, \quad u_0(p) = 0, \quad \text{and} \quad \forall c \in \mathcal{C}, \quad u_1(c) = V_c. \quad (8)$$

## Stationary Constraints

Recall that  $0 \leq \lambda_{ij} \leq 1$  is the conditional probability that a given feasible trading pair on the link  $(i, j)$  will trade if they are selected.

Given an intermediary node  $m$ , the mass of agents selling to  $m$  is  $\sum_{p:(p,m) \in \mathcal{E}_2} (1 - \mu_m) \lambda_{pm}$ . The mass of agents buying from  $m$  is  $\sum_{c:(m,c) \in \mathcal{E}_3} \mu_m \lambda_{mc}$ . The stationary constraints require that for every middlemen node  $m$ :

$$\sum_{p:(p,m) \in \mathcal{E}_2} (1 - \mu_m) \lambda_{pm} = \sum_{c:(m,c) \in \mathcal{E}_3} \mu_m \lambda_{mc}. \quad (9)$$

## Stationary equilibrium

We are now ready to define a stationary equilibrium and state our basic existence result.

DEFINITION 4. Given a finite game  $\Gamma(\mathcal{G}, C, V, N, \delta)$  and a state  $\mu$ , a stationary strategy profile is a **stationary equilibrium** if and only if there exists  $u, z, \lambda$  satisfying (3)-(9), and furthermore,

- If  $z_{ij} < 0$ , then irrespective of who the proposer is, agent  $i$  will never sell an item to agent  $j$ , so that he will wait for a future trade opportunity, that is  $\lambda_{ij} = 0$ .
- If  $z_{ij} > 0$ , then irrespective of who the proposer is, agent  $i$  will sell the item to agent  $j$  with probability one whenever they are matched, that is  $\lambda_{ij} = 1$ .
- If trade between agents  $i$  and  $j$  occurs with probability  $0 < \lambda_{ij} < 1$ , then agents are indifferent between trading and waiting, which implies  $z_{ij} = 0$ .

Next, we give an equilibrium existence result for our model.

THEOREM 1. *A stationary equilibrium always exists for the game  $\Gamma$ .*

*Proof of Theorem 1.* See Appendix EC.1.

## 4. Detecting Trade Delay in General Networks

In this section, we discuss the “always-trade” benchmark of our model. Notice in a flow economy model like ours, delay leads to smaller trade volume and thus, in certain scenarios causes welfare inefficiency. However, trade delay also allows agents to choose optimal trade opportunities. It is

not obvious *a priori* which of the two effects will dominate given a certain network. To answer this question, we first show that the capacity of trade is maximized when there is no delay. Then, we describe a method to detect if a given network can support an always-trade strategy as an equilibrium. Building on this result, in the next two sections we illustrate and analyze different types of trade delay and how they are caused by the underlying network structure.

We start by defining the trade volume given an equilibrium.

DEFINITION 5. Given a stationary equilibrium for the state  $\mu$  with corresponding trade probabilities  $\lambda$ , the **volume of trade**,  $\mathcal{V}$ , is given by the total volume of goods traded per unit time by the producers, or alternately received by the consumers, i.e.,

$$\mathcal{V} = \sum_{(p,c) \in \mathcal{E}_1} \lambda_{pc} + \sum_{m:(p,m) \in \mathcal{E}_2} (1 - \mu_m) \lambda_{pm} = \sum_{(p,c) \in \mathcal{E}_1} \lambda_{pc} + \sum_{m:(m,c) \in \mathcal{E}_3} \mu_m \lambda_{mc},$$

where the equality of the two expressions follows from the stationary constraints in (9).

Having defined the volume of trade for a stationary equilibrium, we present a general result.

LEMMA 1. *The volume of trade,  $\mathcal{V}$ , is maximized only when there is no delay in trade.*

The result sounds intuitive, however, it is not totally straightforward because delay might reduce the fraction of middlemen holding goods, and thus making it easier for some other producers to find trading partners. The proof of Lemma 1 is given in Appendix EC.2. The result in Lemma 1 is important because it says that a stationary equilibrium is efficient, in terms of maximizing the trade volume, if and only if it is an always-trade strategy, so that any equilibrium with endogenous delay is inefficient. Therefore, determining easily verifiable conditions for a given network to support an always-trade strategy in equilibrium is an important objective of the remainder of the paper.

Next we describe an algorithm to detect if a given network can support an always-trade strategy in equilibrium. To do that, we will first assume that the network can support the equilibrium in which trade always occur on every link. Under this assumption, due to the stationary condition, the fraction of middlemen  $m$  holding an item is



$$\mu_m^* = \frac{\text{indegree}(m)}{\text{indegree}(m) + \text{outdegree}(m)}. \quad (10)$$

Here,  $\text{indegree}(m)$  denotes the number of links from producers to  $m$ , and  $\text{outdegree}(m)$  denotes the number of links from  $m$  to consumers.

To simplify the notations let

$$d_i := u_1(i) - u_0(i).$$

Based on the equilibrium conditions derived in Section 3, we set up a system of linear equations with the variables  $(d, z)$ . Specifically from the incentive constraints (3)-(9) we have:

$$z_{ij} = \delta(d_j - d_i) - C_{ij} \quad (11)$$

$$d_p = \frac{1}{2N_p(1-\delta)} \left( \sum_{c:(p,c) \in \mathcal{E}_1} z_{pc} + \sum_{m:(p,m) \in \mathcal{E}_2} (1 - \mu_m^*) z_{pm} \right) \quad (12)$$

$$d_c = V_c - \frac{1}{2N_c(1-\delta)} \left( \sum_{p:(p,c) \in \mathcal{E}_1} z_{pc} + \sum_{m:(m,c) \in \mathcal{E}_3} \mu_m^* z_{mc} \right) \quad (13)$$

$$d_m = \frac{1}{2N_m(1-\delta)} \left( \sum_{c:(m,c) \in \mathcal{E}_3} z_{mc} - \sum_{c:(p,m) \in \mathcal{E}_2} z_{pm} \right). \quad (14)$$

We obtain the following result, which follows directly from the definition of a stationary equilibrium.

**THEOREM 2.** *The game has an always-trade stationary equilibrium if and only if the set of linear equations (11)-(14) above has a nonnegative solution.*

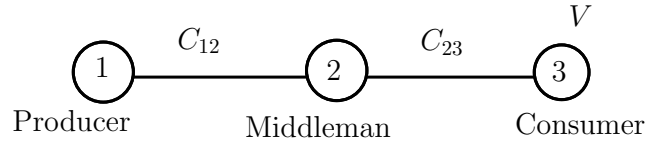
Note that Theorem 2 naturally provides an algorithm to detect if a trade network can support an always-trade equilibrium by solving a linear program corresponding to (11)-(14) with additional constraints  $d_i \geq 0$  and  $z_{ij} \geq 0$ .

## 5. Two Illustrative Examples

As discussed earlier, always-trade maximizes trade volume, which is one natural benchmark for a flow economy. However, in many networks, it might not be welfare efficient because different trade routes can have different costs. In this section, we provide two examples, where trade surplus is

the same on all trade routes. Thus, if the trade surplus is positive, then always-trade will also be the efficient market outcome. The main point of these examples is to illustrate how networks and bargaining incentives influence trade delay. Subsequently in Section 6 we give results for a general network including, in Section 6.2, the case where the costs of the trade routes are heterogeneous.

### 5.1. The Impact of Transaction Costs



**Figure 2** A simple two-link network to illustrate endogenous delay.

Consider a simple network that consists of two links as illustrated in Figure 2. Assume  $C_{12}$  and  $C_{23}$  are the transaction costs of the first and second link, respectively. Let  $V$  be the value of the consumption of the good. In Appendix EC.3, we show that this network has a unique equilibrium, which we can completely characterize in closed form.

This network represents the simplest example where producers and consumers cannot trade directly. Even in this simple network, we observe an interesting phenomenon of endogenous delay as part of the equilibrium. For example, in Appendix EC.3 we show the following:

**PROPOSITION 1.** *If  $0 < V - C_{12} - C_{23} < C_{12} \frac{N_m}{2N_c}$ , then there exists  $\delta^* < 1$  such that for all  $\delta > \delta^*$  no always-trade strategy is an equilibrium. If  $V - C_{12} - C_{23} \geq C_{12} \frac{N_m}{2N_c}$ , then there exists  $\delta^{**} < 1$  such that for all  $\delta > \delta^{**}$  the unique equilibrium is always-trade.*

The intuition for trade delay in this example is the following. In an efficient market if agents were always perfectly matched with a feasible trading partner and if  $\delta V > C_{12} + C_{23}$ , so that trade is beneficial, producers would always trade with middlemen and middlemen would trade with consumers whenever these agents meet one another. However, there are two key difference in our model: first after buying a good from a producer, a middleman has to wait to be matched with a

feasible trading partner, which due to discounting reduces the expected gain from trade and second once a middleman is matched it needs to bargain with a consumer to resell the good. At this point, the transaction cost  $C_{12}$  is sunk and is irrelevant in the negotiation. This type of sunk-cost problem is well known in the literature. For example, Wright and Wong (2014) considers a model for such a setting, where each node represents a single agent. In such a model, the expected payoff of the middlemen from the re-sale might not be enough to recover the sunk cost, leading to market failure.

Even without the strategic considerations due to bargaining, the additional delay due to the matching process can make trade infeasible even when  $\delta V > C_{12} + C_{23}$ . In Appendix EC.6 we analyze such a scenario for the example in Figure 2 and show that all trades generate non-negative returns only when

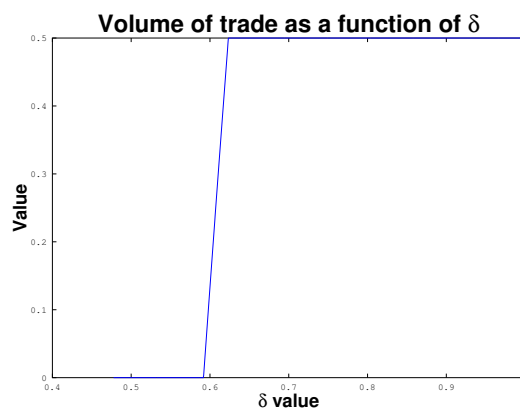
$$\delta V > C_{23} + (N_m/\delta - (N_m - 1))C_{12}. \quad (15)$$

For a given set of parameters, we will use this quantity as a benchmark for when always-trade is feasible. Note that the value of  $\delta V$  in (15) is strictly larger than when the matching is perfect (unless  $\delta = 1$ ). The analysis in Appendix EC.6 can easily be extended to show that the required value of  $\delta V$  for trade to be feasible is increasing in the amount of non-strategic delay incurred due to matching on the second link. However, the matching delay on the first link does not impact the feasibility of trade. This is due to the fact that the transaction cost for trades over this link are not incurred until a trade is made.

In our model each node consists of a large population of agents and the “bargaining power” of a middleman agent at location 2 compared with a consumer agent at location 3 depends on the competition with other middlemen that are also trying to sell. In particular, if the fraction of middlemen that are selling is small, then when negotiating with consumers, they obtain a higher payoff, which will overcome the sunk cost problem of trading with 1. However, to maintain the small fraction of middlemen looking to sell, the rate of trade between 1 and 2 needs to be smaller than the rate between 2 and 3. This then implies that when producers and middlemen meet, they

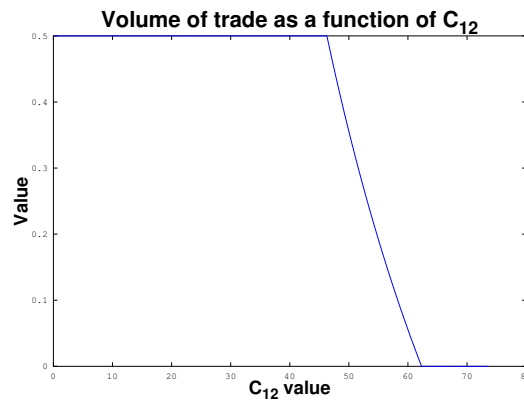
do not trade with probability 1. This can only be rationalized if the surplus of trade is the same as the producers' outside option, which we normalize to be 0. In other words, in this case the producers are indifferent between trading and not trading. When two agents meet, even though they can potentially trade, they may only enact a successful negotiation with a probability strictly in  $(0, 1)$ ; we interpret this as endogenous delay.

For concreteness, we discuss two specific numerical examples. For the first, we fix  $V = 50$ ,  $C_{12} = C_{23} = 1$ ,  $N_m = 20$  and  $N_p = N_c = 2$ , and then we vary  $\delta$  and investigate the change in trade volume, particularly as  $\delta$  increases towards 1. Our goal is to demonstrate that the discount factor can contribute to delay in trade. In Figure 3 we plot the volume of trade as a function of  $\delta \in [0.2, 1]$ . Note that from (15), the always-trade strategy is infeasible for  $\delta < 0.4776$ . However, in Figure 3, we see that there are a range of feasible  $\delta$  for which the always-trade strategy is not sustainable in equilibrium. Note also that in part of the range the trade on link  $(1, 2)$  occurs with probability strictly less than 1.



**Figure 3** Trade volume as a function of  $\delta$  for  $V = 50$ ,  $C_{12} = C_{23} = 1$ ,  $N_m = 20$  and  $N_p = N_c = 2$ .

For the second numerical example, we set  $V = 100$ ,  $C_{23} = 0$ ,  $\delta = 0.9$  and  $N_m = N_p = N_c = 2$ , and then vary  $C_{12}$  from 0 to 73.64. From (15), feasibility of always-trade requires  $C_{12} \leq 73.64$ . Again we present numerical calculations of the volume of trade as a function of  $C_{12}$ . Here by setting  $C_{23} = 0$ , we isolate the impact of transaction cost  $C_{12}$  of the first link. In Figure 4, we plot the volume of trade as a function of  $C_{12}$ . As  $C_{12}$  is increased, we find that there is a maximum value such that an always-trade strategy is an equilibrium. This value is considerably smaller than  $\delta V = 90$  or even the feasibility threshold of 73.64, and increasing the cost further results in a decrease in the volume of trade with the complete cut-off of the trade well before  $C_{12} = 73.64$ .

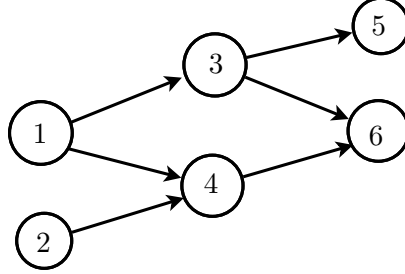


**Figure 4** Trade volume as function of  $C_{12}$  for  $V = 100$ ,  $C_{23} = 0$ ,  $\delta = 0.9$  and  $N_m = N_p = N_c = 2$ .

## 5.2. The Impact of Network Structure on Trade Capacity

As we have seen in the previous example, one important “friction” causing delay in trade is transaction cost. We now show another example, where the heterogeneity of the network structure is the main cause of delay. The network we consider is the one illustrated in Figure 5. In this network, 1

and 2 represent the sellers, 3 and 4 are middlemen, 5 and 6 are buyers. We assume the population at each node  $i$  is  $N_i \equiv 3$ . To isolate the effect of the network, we assume that all transaction costs are 0 and each consumer's valuation of the good is  $V = 1$ .



**Figure 5** Example to illustrate the effect of network structure. Here transaction costs at every links are 0 and the value for all buyers for obtaining a good is  $V = 1$ .

As all the transaction costs are 0, from Example 1 one might think that there should not be any delay in trade here, but it is not the case. In this example, middlemen at location 3 have a competitive advantage on the consumer's side: they have access to more consumers than middlemen at location 4. Thus, when holding a good, middlemen at 3 can find a consumer easier than those at 4. This knowledge, in return, influences trade in the previous round(s) between producers and middlemen. Specifically, consider the competition between middlemen at locations 3 and 4 when they are buying from 1 and 2. As middlemen at 3 can sell the goods faster than those at 4, the middlemen at 3 can offer a higher price to the producers at 1 than those at 4 can. As a result, when the discount rate is close to 1, agents at node 1 have incentive to hesitate trading with 4 in order to wait for trade opportunities with agents at 3. This causes delay in trade and influences the total trade volume.

We numerically compute the (unique) stationary equilibrium for this example. First, from Theorem 2, we can characterize when an always-trade equilibrium exists by computing the range of  $\delta$  in which (11)-(14) have a non-negative solution. Specifically, we have the following. *If  $\delta < 0.92473$ , then the unique pure equilibrium is an always-trade strategy, in which case the payoff of middlemen*

3 is greater than that of middlemen 4. This also results in the largest trade volume. Note all the  $z_{ij}$  variables are positive in this regime.

When  $\delta \geq 0.92473$ , the value  $\delta(d_4 - d_1)$  becomes negative. Hence we add another variable  $0 < \lambda_{14} < 1$  capturing the trade probability on the link (1, 4). As agents at location 1 are indifferent between trade and no trade with agents at node 4, the trade surplus of link (1, 4) is  $z_{14} = 0$ . With this new variable, we search for the range of  $\delta$  that can support this type of equilibrium. We have the following.

*If  $0.92473 \leq \delta \leq 0.94642$ , then in the unique equilibrium, trade on link (1, 4) gets delayed and occurs with probability  $\lambda_{14} \in [0, 1]$ ; the exact value depends in a complex manner on the discount factor with  $\lambda_{14} = 0$  when  $\delta = 0.94642$ . The volume of trade decreases as the probability of trade  $\lambda_{14}$  decreases to 0.*

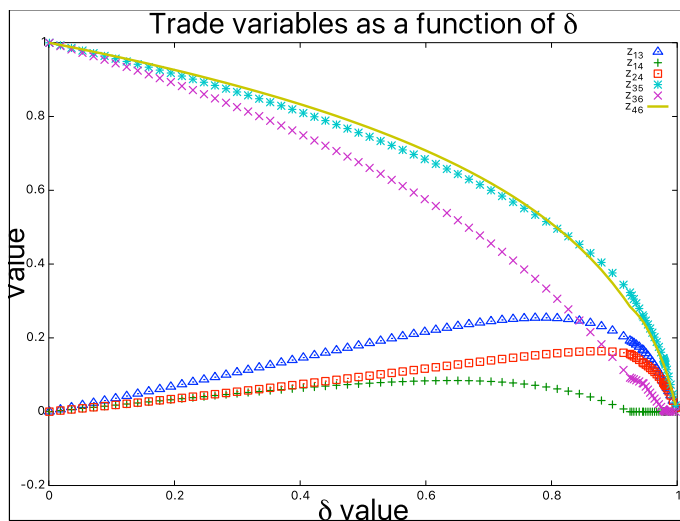
Similarly, when we increase  $\delta$  further we have the following: *If  $0.94642 \leq \delta \leq 0.98169$ , then in the unique equilibrium there is no trade whatsoever on link (1, 4) but trade occurs with probability 1 on all other links. The volume of trade remains a constant for this range of  $\delta$ . Note that  $z_{14}$  is negative in this regime.*

Finally, the last regime has the following behavior: *If  $\delta \geq 0.98169$ , then in the unique equilibrium there is no trade whatsoever on link (1, 4) and trade gets delayed on link (3, 6) with the probability of trade going to 0 in the limit of  $\delta$  increasing to 1. As before  $z_{14}$  is negative in this regime. Again the volume of trade decreases as  $\lambda_{36}$  goes to 0 with the lowest trade volume observed at  $\delta = 1$ . Also, since  $z_{36} = 0$  we also find  $d_3 = d_6$ .*

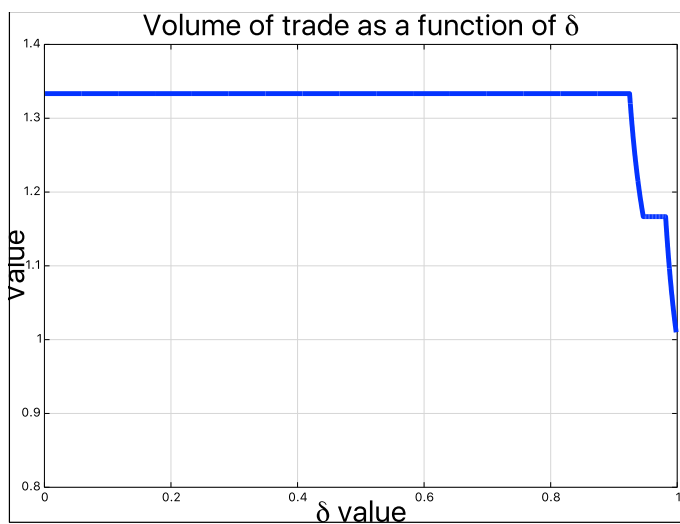
These results are illustrated in the following figures: Figure 6 displays the trade variable  $z_{ij}$  of every link  $(i, j)$  as a function of  $\delta$  from which one can discern the conclusions described above; and Figure 7 shows the volume of total trade as a function of  $\delta$ .

## 6. Network Cuts and Trade Delay

Even though Theorem 2 provides a general method to check if a network can support the maximum trading capacity (under an always-trade strategy), it lacks an intuitive explanation on the main factors that lead to trade delay. This is our goal in this section.



**Figure 6** The trade variable  $z_{ij}$  for each link  $(i, j)$ .



**Figure 7** Total volume of trade in the network.



We start our analysis in this section with the following observation, which captures another channel of delay in networked bargaining.

**THEOREM 3.** *Given a producer  $p$ , a consumer  $c$ , and any  $\epsilon > 0$ , there exists  $\delta^*$ , such that for all  $\delta > \delta^*$  and at any equilibrium the following is true. If  $\lambda_{pm} > 0$  and  $\lambda_{mc} > 0$  for a middlemen  $m$ , that is trade occurs along the route  $p \rightarrow m \rightarrow c$ , then the cost  $C_{pm} + C_{mc}$  is the smallest among all trading routes between  $p$  and  $c$ .*

*Proof of Theorem 3.* See Appendix EC.4.

An immediate consequence of Theorem 3 is that if there are multiple links in  $\mathcal{E}$  between nodes  $i$  and  $j$ , then in the limit of  $\delta$  approaching 1 only the lowest cost links will be used. This also leads to trade-delay. However, unlike the examples in Section 5, this is a “good” type of delay because it allows agents to find better trading routes and thus may improve welfare. In particular, in the limit as  $\delta$  approaches 1 this type of delay will improve the welfare gained by each good sold.

Given multiple sources of delay, it becomes challenging to analyze them separately. It is also unclear *a priori* if there are other factors that may cause trade delay. Our approach is to consider the limiting regime in which  $\delta$  approaches 1. Our main result in this section is a sharp characterization of trade delay in such a limit regime. The characterization consists of three conditions, which correspond to the three channels that we discussed so far.

To capture the equilibrium notion in the limit of  $\delta$  approaching 1 in a precise manner, we make the following definition.

**DEFINITION 6.** The game has an **always-trade equilibrium in the limit** if there exists  $\delta^* < 1$  such that for all  $\delta > \delta^*$  the game has an equilibrium with always-trade strategies.

We make two additional assumptions in this section to derive a complete characterization of networks that has an always-trade equilibrium in the limit. Henceforth, we assume  $\text{indegree}(m) = \text{outdegree}(m)$  for every middlemen  $m$ , and  $V_c \equiv V$  for all consumers  $c$ . Notice that the assumption  $V_c \equiv V$  is not crucial, and can be relaxed. The assumption  $\text{indegree}(m) = \text{outdegree}(m)$ , however,

is important because it allows us to derive the characterization from a network flow problem. With these additional assumptions, we can capture (11)-(14) as a network flow problem and thus derive intuitive explanation on trade delay based on the cuts of the underlying network.

Before we proceed, we derive the equations for stationary equilibria in the limit of  $\delta$  approaching 1. With the assumption that  $\text{indegree}(m) = \text{outdegree}(m)$  for every middlemen  $m$ , we have  $\mu_m^* \equiv 1/2$ . Using this consequence,  $V_c \equiv V$ , and setting

$$w_{ij} = \frac{z_{ij}}{1-\delta},$$

we can rewrite (11)-(14) as

$$d_j - d_i - \frac{(1-\delta)w_{ij}}{\delta} = \frac{C_{ij}}{\delta} \quad (16)$$

$$d_p = \frac{1}{2N_p} \left( \sum_{c:(p,c) \in \mathcal{E}_1} w_{pc} + \sum_{m:(p,m) \in \mathcal{E}_2} \frac{1}{2} w_{pm} \right) \quad (17)$$

$$d_c = V - \frac{1}{2N_c} \left( \sum_{p:(p,c) \in \mathcal{E}_1} w_{pc} + \sum_{m:(m,c) \in \mathcal{E}_3} \frac{1}{2} w_{mc} \right) \quad (18)$$

$$d_m = \frac{1}{2N_m} \left( \sum_{c:(m,c) \in \mathcal{E}_3} w_{mc} - \sum_{c:(p,m) \in \mathcal{E}_2} w_{pm} \right). \quad (19)$$

Since for each agent  $i$ ,  $d_i$  is bounded independent of  $\delta$ , then it can be seen from (11)-(14) that  $z_{ij}$  must approach zero as  $\delta$  approaches one and so  $\lim_{\delta \rightarrow 1} (1-\delta)w_{ij} = 0$ . Therefore, (16)-(19) transform to

$$d_j - d_i = C_{ij} \quad (20)$$

$$d_p = \frac{1}{2N_p} \left( \sum_{c:(p,c) \in \mathcal{E}_1} w_{pc} + \sum_{m:(p,m) \in \mathcal{E}_2} \frac{1}{2} w_{pm} \right) \quad (21)$$

$$d_c = V - \frac{1}{2N_c} \left( \sum_{p:(p,c) \in \mathcal{E}_1} w_{pc} + \sum_{m:(m,c) \in \mathcal{E}_3} \frac{1}{2} w_{mc} \right) \quad (22)$$

$$d_m = \frac{1}{2N_m} \left( \sum_{c:(m,c) \in \mathcal{E}_3} w_{mc} - \sum_{c:(p,m) \in \mathcal{E}_2} w_{pm} \right). \quad (23)$$

We have the following straightforward result that we present without proof.

PROPOSITION 2. *The following statements are true:*

1. *If for every  $\delta^* < 1$ , there exists  $\delta > \delta^*$  such that (16)-(19) have a non-negative solution, then (20)-(23) also have a non-negative solution. This means that if the game has an always-trade equilibrium in the limit, then (20)-(23) have a non-negative solution*

2. *If (20)-(23) have a positive solution, then there exists a  $\delta^{**}$  such that (16)-(19) also have a positive solution for all  $\delta > \delta^{**}$ . This implies that if (20)-(23) have a positive solution, then the game has an always-trade equilibrium in the limit.*

Using this result, we will now analyze conditions for the existence of non-negative solutions to (20)-(23).

### Network-flow reformulation

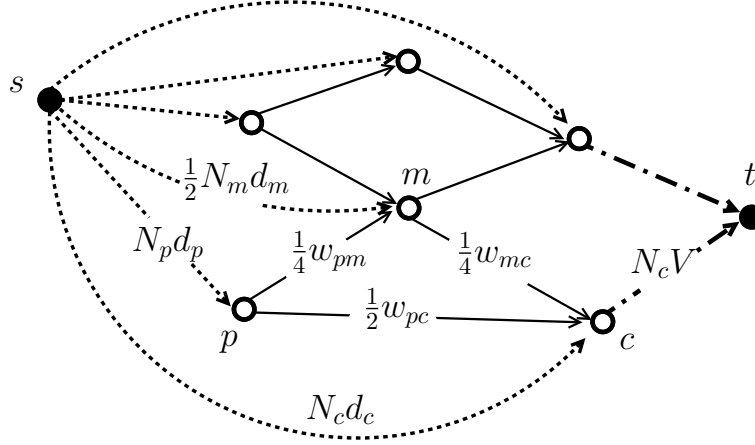
We will start by showing that the characterization of an always-trade equilibrium in the limit can be captured by a network flow formulation, i.e., the solution to the set of equations (21)-(23) can be derived from the solution of a related network flow problem.

To see this, we add a “source”  $s$  and a “sink”  $t$ . From  $s$  we make a directed link to every node of our network. Then from each consumer node, we make a directed link to the sink  $t$ . Let the flow value on the link  $(s, p)$  be  $N_p d_p$ , the flow value on the link  $(s, m)$  be  $\frac{1}{2} N_m d_m$ , the flow value on the link  $(s, c)$  is  $N_c d_c$ , and the flow on the link  $(c, t)$  is  $N_c V$ . The flow value on link  $(p, c)$  be  $\frac{1}{2} w_{pc}$  and finally, the flow value on link  $(p, m)$  and  $(m, c)$  will be  $\frac{1}{4} w_{pm}$  and  $\frac{1}{4} w_{mc}$ , respectively. See Figure 8 for an illustration. Notice that if there are  $k$  parallel links between two nodes, then the **total** flow value between them is  $k$  times the value above. It is clear that equations (21)-(23) correspond to a flow constraints in this network.

We use this network flow representation and its dual to obtain our main results in the following two subsections.

#### 6.1. Two transaction costs $C_{pm} = C_1$ ; $C_{mc} = C_2$

We start with the scenario in which transaction costs take two values. The cost between any producers and middlemen is assumed to be  $C_1$  and the cost between any middlemen and consumer



**Figure 8** Network flow representation of (21)-(23).

is assumed to be  $C_2$ . We will also assume that there are no direct producer and consumer links. Naturally, we assume trade is beneficial, that is,  $V > C_1 + C_2$ . By assuming this special cost structure, we eliminate the effect of trade routes with different costs. We will consider the general case in Section 6.2.

From (20),  $d_j - d_i = C_{ij}$ , and the assumption that the network is connected, we obtain  $d_p = d$ ,  $d_m = d + C_1$  and  $d_c = d + C_1 + C_2$  for some  $d$  to be determined as yet. (Recall that we consider the network to be connected when we ignore the direction of the links, i.e., it cannot be partitioned into two set of nodes that are not connected by any link.) Since the solution of (20)-(23) corresponds to the solution of the network flow problem, the total flow going out of  $s$  is equal to the flow going into  $t$ , hence

$$\sum_p N_p d + \sum_c N_c (d + C_1 + C_2) + \frac{1}{2} \sum_m N_m (d + C_1) = \sum_c N_c V.$$

This implies

$$d = \frac{(V - C_1 - C_2) \sum_c N_c - \frac{1}{2} \sum_m C_1 N_m}{\sum_p N_p + \sum_c N_c + \frac{1}{2} \sum_m N_m}. \quad (24)$$

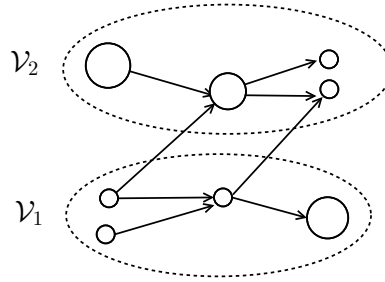
From this we obtain the first condition for the existence of a nonnegative solution,  $d \geq 0$ . Concretely,

$$(V - C_1 - C_2) \sum_c N_c \geq \frac{1}{2} \sum_m C_1 N_m. \quad (25)$$

As in Example 1, condition (25) captures the impact of sunk costs. (Note that (25) is exactly the condition of Proposition 1.) The left hand side of (25) is the total trade surplus. The right hand

side is half the total transaction costs between producers and middlemen. These costs is sunk when middlemen bargain with consumers. As illustrated in Example 1, when (25) is violated, middlemen will be more conservative and delay trade with producers.

However, (25) alone does not guarantee that (20)-(23) have a nonnegative solution. To derive another condition, consider a partition of network into two sets of vertices  $\mathcal{V}_1$  and  $\mathcal{V}_2$ , such that there are no directed links going from  $\mathcal{V}_2$  to  $\mathcal{V}_1$ . In other words, no producers in  $\mathcal{V}_2$  are linked with middlemen and consumers in  $\mathcal{V}_1$ , no middlemen in  $\mathcal{V}_2$  are linked with consumers in  $\mathcal{V}_1$ . We call such a partition a **directed cut** of the network. See Figure 9.



**Figure 9** A directed cut of the network

Assume there exists a nonnegative solution of (20)-(23) and consider the corresponding network flow as described in Figure 8. For a directed cut  $(\mathcal{V}_1, \mathcal{V}_2)$ , there can only be flow from  $\mathcal{V}_1$  to  $\mathcal{V}_2$ . Therefore, the total flow from  $s$  to  $\mathcal{V}_1$  has to be at least the total flow from  $\mathcal{V}_1$  to  $t$ . Hence,

$$\sum_{p \in \mathcal{V}_1} N_p d + \sum_{c \in \mathcal{V}_1} N_c (d + C_1 + C_2) + \frac{1}{2} \sum_{m \in \mathcal{V}_1} N_m (d + C_1) \geq \sum_{c \in \mathcal{V}_1} N_c V.$$

This implies

$$r(\mathcal{V}_1) := \frac{(V - C_1 - C_2) \sum_{c \in \mathcal{V}_1} N_c - \frac{1}{2} \sum_{m \in \mathcal{V}_1} C_1 N_m}{\sum_{p \in \mathcal{V}_1} N_p + \sum_{c \in \mathcal{V}_1} N_c + \frac{1}{2} \sum_{m \in \mathcal{V}_1} N_m} \leq d. \quad (26)$$

Furthermore, the total flow from  $s$  to  $\mathcal{V}_2$  has to be at most the total flow from  $\mathcal{V}_2$  to  $t$ . Thus,

$$\sum_{p \in \mathcal{V}_2} N_p d + \sum_{c \in \mathcal{V}_2} N_c (d + C_1 + C_2) + \frac{1}{2} \sum_{m \in \mathcal{V}_2} N_m (d + C_1) \leq \sum_{c \in \mathcal{V}_2} N_c V.$$

This implies

$$r(\mathcal{V}_2) := \frac{(V - C_1 - C_2) \sum_{c \in \mathcal{V}_2} N_c - \frac{1}{2} \sum_{m \in \mathcal{V}_2} C_1 N_m}{\sum_{p \in \mathcal{V}_2} N_p + \sum_{c \in \mathcal{V}_2} N_c + \frac{1}{2} \sum_{m \in \mathcal{V}_2} N_m} \geq d. \quad (27)$$

From (26) and (27), we obtain another condition for the existence of a nonnegative solution for (20)-(23), namely,

$$r(\mathcal{V}_2) \geq r(\mathcal{V}_1) \text{ for every directed cut } (\mathcal{V}_1, \mathcal{V}_2). \quad (28)$$

Condition (28) provides important economic intuition. Roughly speaking,  $r(\mathcal{V}_i)$  is the ratio between the total surplus minus the total sunk cost and the market size. (Here the contribution of a middleman towards the market size is half that of a producer or a consumer.) Thus, we can think of  $r(\mathcal{V}_i)$  as the “strength” of the submarket  $\mathcal{V}_i$ . For a directed cut  $(\mathcal{V}_1, \mathcal{V}_2)$ , if (25) does not hold, i.e, the strength of  $\mathcal{V}_1$  is larger than the strength of  $\mathcal{V}_2$ , then, when holding an item, agents in  $\mathcal{V}_1$  would not want to sell to agents in  $\mathcal{V}_2$ , at least when the discount rate is close to 1 (i.e., if they’re patient). The intuition that emerges is the following: agents in  $\mathcal{V}_1$  are better off waiting to trade with other agents in  $\mathcal{V}_1$ . Such a scenario can be interpreted as market segmentation: agents in a better submarket hesitate to trade with outsiders.

We have derived two necessary conditions (25) and (28) for the game to have an always-trade equilibrium in the limit (of  $\delta$  approaching 1). In fact, these condition are also “almost” sufficient. This is captured by the following result.

**THEOREM 4.** *If either (25) or (28) are violated, then the game does not have an always-trade equilibrium in the limit. Moreover, if both (25) and (28) are satisfied with strict inequalities, then the game has an always-trade equilibrium in the limit.*

Owing to the network flow reformulation, the proof of Theorem 4 can be derived from the duality theorem of the network flow problem. Instead of presenting the details, we will prove a more general result in the next section, Theorem 5.

## 6.2. Heterogeneous Transaction Costs

We next consider the general case of heterogeneous transaction costs. For ease of presentation, we illustrate the result in this section by the network in Figure 10. This is a slight modification of the network in Figure 5 so that the in-degree is equal to the out-degree for each middleman node.

As mentioned in the remarks after Theorem 3, the preference for cheaper trade routes in the limit of  $\delta$  approaching one yields a new type of delay in the trade. This new type of delay due to heterogeneous transaction costs can also be seen from the limit linear program (20)-(23). Namely, consider (20), which imposes

$$d_j - d_i = C_{ij}.$$

Note that if there are two paths of different costs between a producer  $p$  and a consumer  $c$ , then (20) cannot be satisfied, and thus, no always-trade strategy can be an equilibrium in the limit.

Constraint (20) is actually equivalent to the following condition on the transaction costs, which we call the 0-cycle condition.

DEFINITION 7. Given a directed network  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  and a cost function  $C_{ij}$  on the links, we say  $C$  satisfies the **0-cycle condition** if for any cycle  $(1, 2, \dots, k, k+1 \equiv 1)$  of the network

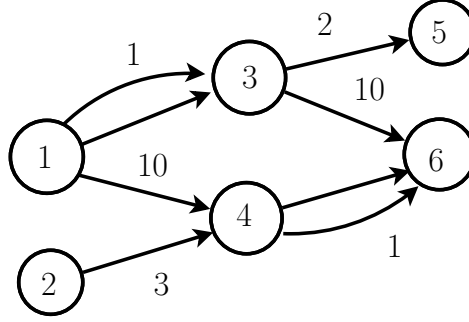
$$\sum_{i=1}^k \text{sign}_{i,i+1} \cdot C_{i,i+1} = 0,$$

where  $\text{sign}_{i,i+1} = 1$  if the direction of the link from  $i$  to  $i+1$  and  $\text{sign}_{i,i+1} = -1$ , otherwise.

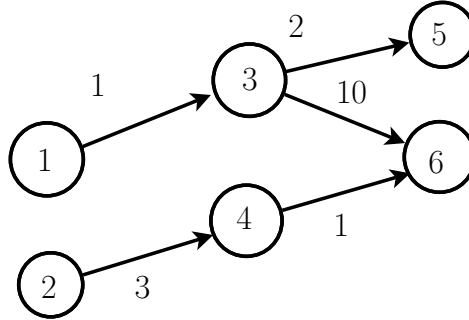
It is clear that if  $d_j - d_i = C_{ij}$  for all links  $(i, j)$ , then the 0-cycle condition must hold. In Section 6.1, there are only two types of transaction costs  $C_{pm} = C_1$  and  $C_{mc} = C_2$ , the 0-cycle condition is automatically satisfied.

In fact, the reverse is also true: if  $C$  satisfies the 0-cycle condition, then there exists a vector  $d$  such that  $d_j - d_i = C_{ij}$ . In particular, recall that in Section 6.1, from  $d_j - d_i = C_{ij}$  we derived a formulation for all the variables  $d$ . Specifically, in that example,  $d_p = d; d_m = d + C_1; d_c = d + C_1 + C_2$  for a specific choice of number  $d$ .

For the general case where the transaction costs satisfy the 0-cycle condition, we can obtain a similar calculation as in the previous section. To illustrate, consider the network in Figure 10, which satisfies the 0-cycle condition. Let  $T$  be a spanning tree of this network (See Figure 11). Note that here we ignore the direction of the links when constructing the spanning tree.



**Figure 10** An example of trading network



**Figure 11** An example of computing  $\Delta_i$ .

Assume we know the value of  $d_2$ . Because  $d_4 - d_2 = 3$ ,  $d_4 = d_2 + 3$ . Similarly,  $d_6 - d_4 = 1$ , thus,  $d_6 = d_2 + 4$ . Consider now the link (3,6), we have  $d_6 - d_3 = 10$ , therefore,  $d_3 = d_2 + 4 - 10 = d_2 - 6$ . Using this method, we have

$$d_1 = d_2 - 7; d_2 = d_2; d_3 = d_2 - 6; d_4 = d_2 + 3; d_5 = d_2 - 4; d_6 = d_2 + 4.$$

The lowest value among these values is  $d_1$ . We will determine these variable by letting  $d := d_1 = d_2 - 7$ . With this we have:

$$d_1 = d; d_2 = d + 7; d_3 = d + 1; d_4 = d + 10; d_5 = d + 3; d_6 = d + 11.$$

Moreover, it is not hard to see that because of the 0-cycle condition, all the solutions satisfying  $d_j - d_i = C_{ij}$  will be of this form, regardless of the choice of a spanning tree  $T$ . This is captured by the following result.



LEMMA 2. *If  $C$  satisfies the 0-cycle condition and the network is connected, then there exist unique  $\Delta_i$ -s such that  $\min_{i \in \mathcal{V}} \{\Delta_i\} = 0$  and all solutions satisfying  $d_j - d_i = C_{ij}$ ,  $d_i \geq 0$  are of the form  $d_i = d + \Delta_i$  for  $d \geq 0$ .*

*Proof of Lemma 2.* See Appendix EC.5.

Assume that the game has an always-trade equilibrium in the limit, then (20)-(23) have a non-negative solution. Using Lemma 2, we again have a network flow reformulation, by adding a source  $s$  and a sink  $t$  with the same connectivity as before. Thereafter we set the flow values from  $s$  to  $p$  and  $c$  to be  $N_p(d + \Delta_p)$  and  $N_c(d + \Delta_c)$ , respectively. The flow values from  $s$  to  $m$  is set to be  $\frac{1}{2}N_p(d + \Delta_p)$ . Finally, the flow from  $c$  to  $t$  is set to be  $N_cV$ .

There are two constraints for such a flow. First, we need to have  $d \geq 0$ . Second, the flow value on each link should be non-negative. The first condition gives us an inequality among  $V$  and all the  $\Delta_i$ , which we interpreted as the relationship between trade surplus and sunk costs in the previous section. The second implies the directed cut condition, which we interpreted as a condition on the strength of submarkets. Formally, consider the following definition.

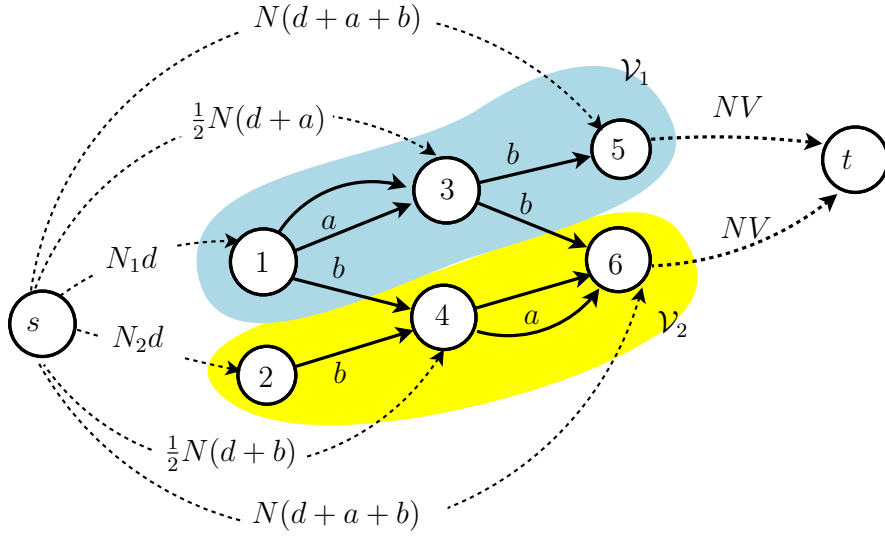
DEFINITION 8. Let  $\Delta_i$  be the value determined in Lemma 2. For every subset of vertices  $\mathcal{S} \subset \mathcal{V}$ , let the **strength** of the submarket  $\mathcal{S}$  be

$$r(\mathcal{S}) := \frac{\sum_{c \in \mathcal{S}} (V - \Delta_c) N_c - \sum_{p \in \mathcal{S}} \Delta_p N_p - \frac{1}{2} \sum_{m \in \mathcal{S}} \Delta_m N_m}{\sum_{p \in \mathcal{S}} N_p + \sum_{c \in \mathcal{S}} N_c + \frac{1}{2} \sum_{m \in \mathcal{S}} N_m}.$$

In the example shown in Figure 12, assume the population at every node is  $N$ . The strength of the submarket  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are

$$\frac{V - 3/2a - b}{5/2} \text{ and } \frac{V - a - 3/2b}{5/2}, \text{ respectively.}$$

From this, we can see that if  $b > a$ , then submarket  $\mathcal{V}_1$  is stronger than  $\mathcal{V}_2$ . The reason is that the cost of the first link from the producer to middlemen  $\mathcal{V}_2$  is  $b$ , which is higher than that in  $\mathcal{V}_1$ . Thus, the agents in  $\mathcal{V}_2$  suffer from higher sunk cost. Our next theorem will show that because of this, in the limit (of  $\delta$  approaching 1) trade from  $\mathcal{V}_1$  to  $\mathcal{V}_2$  is either delayed or unsuccessful.



**Figure 12** An example illustrating the strength of  $\mathcal{V}_1$  and  $\mathcal{V}_2$

On the other hand, consider  $a = b$ , but the populations at node 1 and 2 are different from  $N$ : they are  $N_1$ , and  $N_2$ , respectively. In this case the strength of the submarkets  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are

$$\left(V - \frac{5}{2}a\right) \frac{N}{\frac{3N}{2} + N_1} \text{ and } \left(V - \frac{5}{2}a\right) \frac{N}{\frac{3N}{2} + N_2}, \text{ respectively.}$$

Thus, if  $N_2 > N_1$ , then market  $\mathcal{V}_1$  is stronger than  $\mathcal{V}_2$ . This is also intuitively clear because now market  $\mathcal{V}_2$  has more producers while the total surplus and the transaction cost are the same as market  $\mathcal{V}_1$ . This causes increased competition in  $\mathcal{V}_2$ . Our next theorem will also show that when agents in  $\mathcal{V}_1$  are matched with agents in  $\mathcal{V}_2$ , they will not trade immediately.

**THEOREM 5.** *The game cannot have an always-trade equilibrium in the limit of  $\delta$  approaching 1 if one of the following is not satisfied:*

- (i.)  $C$  satisfies the 0-cycle condition
- (ii.)  $r(\mathcal{V}) \geq 0$
- (iii.) for every directed cut  $(\mathcal{V}_1, \mathcal{V}_2)$ :  $r(\mathcal{V}_1) \leq r(\mathcal{V}_2)$

*Conversely, if all three conditions hold, and the last two are satisfied with strict inequality, there exists an always-trade equilibrium in the limit.*

Notice that our main result, Theorem 5, implies Theorem 4 as a special case with  $\Delta_p = 0$ ,  $\Delta_m = C_1$  and  $\Delta_c = C_1 + C_2$ .

*Interpretation of Theorem 5.* Together with the examples in Section 5, Theorem 5 gives us a clear economic interpretation of the sources of delay and the behavior of our trading network. In particular, the three conditions in Theorem 5 essentially represents three channels leading to delay in our model. Condition (i.) requires the total transaction cost to be the same for two different trading routes, otherwise, agents will postpone trade to choose the cheapest trading route. Besides this “useful” type of type of delay, that helps to improve welfare, Theorem 5 identifies the remaining two sources of “harmful” delay. Specifically, when (i.) is satisfied, in an efficient economy, agents should trade immediately upon meeting with the right partner. However, if (ii.) is violated, then the total sunk cost is too high compared with the total trade surplus. In such a scenario, middlemen will trade conservatively with producers as we have seen in the example in Section 5.1. Finally, if (iii.) is violated, then in the limit, the market is segregated. Agents in the better submarket will hesitate to sell to outsiders.

### Proof of Theorem 5

In the remainder of this section, we give the proof of Theorem 5. The first part of the Theorem is analogous to the derivation of Condition (25) and (28) in Section 6.1. The reverse direction is the consequence of Hoffman’s Theorem, which is the duality result of a network circulation problem.

*Proof of Theorem 5.* To see the first part of the result, assume that the game has an always-trade equilibrium in the limit. Then the set of limit equations (20)-(23) has a nonnegative solution. Owing to the arguments made in detail earlier, the vector of transaction costs  $C$  needs to satisfy the 0-cycle condition.

To see (ii.) and (iii.) we will use the flow representation of a nonnegative solution of (20)-(23).

First because the outflow at  $s$  is equal to the inflow at  $t$ . This means that

$$\sum_{p \in \mathcal{V}} N_p(d + \Delta_p) + \sum_{c \in \mathcal{V}} N_c(d + \Delta_c) + \frac{1}{2} \sum_{m \in \mathcal{V}} N_m(d + \Delta_m) = \sum_{c \in \mathcal{V}} N_c V.$$

From this, we have  $d = r(\mathcal{V})$ . Thus, we need  $r(\mathcal{V}) \geq 0$ .

Second, to derive the condition for a directed cut  $(\mathcal{V}_1, \mathcal{V}_2)$ , observe that because all links are directed from  $\mathcal{V}_1$  to  $\mathcal{V}_2$ , the inflow to  $\mathcal{V}_1$  is at most the outflow of  $\mathcal{V}_1$ , while the inflow to  $\mathcal{V}_2$  is at least the outflow of  $\mathcal{V}_2$ . Using similar a similar derivation as in Section 6.1, we obtain that for every directed cut  $(\mathcal{V}_1, \mathcal{V}_2)$ :  $r(\mathcal{V}_1) \leq r(\mathcal{V}_2)$ .

To prove the reverse direction, observe that if the set of limit equations (20)-(23) has a positive solution, then there exists  $\delta^* < 1$  such that for all  $\delta > \delta^*$ , (16)-(19) also has a positive solution. This shows that always-trade is an equilibrium in the limit. Thus, it remains to prove that if (i.) holds and (ii.) and (iii.) are satisfied with strict inequality, then the set of limit equations (20)-(23) has a positive solution.

To prove this we will use Hoffman's circulation theorem (see, for example, Schrijver (2002), Theorem 11.2).

**LEMMA 3 (Hoffman).** *A weight vector  $\mathbf{f}$  on the links of a directed network is called a **circulation** if for every node of the network, the total weights on the links coming into any vertex is equal to the total weights on the links going out of the same vertex. Given a directed network, such that every link  $e$  has an lower and an upper capacity  $0 \leq \underline{k}_e \leq \bar{k}_e \leq +\infty$ . Then there exists a circulation  $\mathbf{f}$  satisfying  $\underline{k}_e \leq f_e \leq \bar{k}_e$  if and only if for every node partition  $(\mathcal{V}_1, \mathcal{V}_2)$*

$$\sum_{e \text{ from } \mathcal{V}_1 \text{ to } \mathcal{V}_2} \underline{k}_e \leq \sum_{e \text{ from } \mathcal{V}_2 \text{ to } \mathcal{V}_1} \bar{k}_e.$$

From the network flow representation in Figure 8, create the following instance of the circulation problem. For every link  $(c, t)$ , let both the lower and upper capacity be  $N_c V$ . For the links  $(s, p)$  and  $(s, c)$ , let both the lower and upper capacity be  $N_p \cdot (r(\mathcal{V}) + \Delta_p)$  and  $N_c \cdot (r(\mathcal{V}) + \Delta_c)$ , respectively. Similarly, let both the lower and upper capacity of  $(s, m)$  be  $\frac{1}{2} N_m \cdot (r(\mathcal{V}) + \Delta_m)$ . For the remaining links, let the upper capacity be  $+\infty$  and the lower capacity be  $\epsilon > 0$ . Finally, merge  $s$  and  $t$  into a single node  $st$ .

It remains to show that if (ii.) and (iii.) are satisfied with strict inequality, there exists an  $\epsilon$  small enough such that the condition in the Hoffman's circulation theorem (Lemma 3) is satisfied.

Thus, there exists a circulation such that the amount of flow on each link is at least  $\epsilon$ . This shows that (20)-(23) has a positive solution.

Consider a partition,  $(st + \mathcal{V}_1, \mathcal{V}_2)$ , of the network constructed above. (The proof is analogous for the partition  $(\mathcal{V}_1, st + \mathcal{V}_2)$ .) We need to show that the total lower capacity of links from  $st + \mathcal{V}_1$  to  $\mathcal{V}_2$  is at most the total upper capacity of the links in the opposite direction.

First notice that if there are links of the original trading network going from  $\mathcal{V}_2$  to  $\mathcal{V}_1$ , then the Hoffman's circulation condition holds because every link in the original trading network has an upper capacity of  $+\infty$ . Thus, we only need to consider a directed cut  $(\mathcal{V}_1, \mathcal{V}_2)$ .

In this case the total lower capacity of links from  $st + \mathcal{V}_1$  to  $\mathcal{V}_2$  is

$$\sum_{p \in \mathcal{V}_2} N_p(r(\mathcal{V}) + \Delta_p) + \sum_{c \in \mathcal{V}_2} N_c(r(\mathcal{V}) + \Delta_c) + \frac{1}{2} \sum_{m \in \mathcal{V}_2} N_m(r(\mathcal{V}) + \Delta_m) + \epsilon \cdot \# \text{ links from } \mathcal{V}_1 \text{ to } \mathcal{V}_2.$$

The total upper capacity of links from  $\mathcal{V}_2$  to  $st + \mathcal{V}_1$  is  $\sum_{c \in \mathcal{V}_2} N_c V$ .

To show that there exists  $\epsilon > 0$  such that the former is at most the latter, we need to show

$$\sum_{p \in \mathcal{V}_2} N_p(r(\mathcal{V}) + \Delta_p) + \sum_{c \in \mathcal{V}_2} N_c(r(\mathcal{V}) + \Delta_c) + \frac{1}{2} \sum_{m \in \mathcal{V}_2} N_m(r(\mathcal{V}) + \Delta_m) < \sum_{c \in \mathcal{V}_2} N_c V.$$

This is equivalent to  $r(\mathcal{V}) < r(\mathcal{V}_2)$ . Now,  $r(\mathcal{V}_2)$  is the average of trade minus sunk cost over  $\mathcal{V}_2$ ;  $r(\mathcal{V})$  is the average of trade minus sunk cost over  $\mathcal{V}_1 \cup \mathcal{V}_2$ . Furthermore, because of (iii.)  $r(\mathcal{V}_1) < r(\mathcal{V}_2)$ . Hence  $r(\mathcal{V}) < r(\mathcal{V}_2)$ , which is what we needed to prove. ■

## 7. Conclusions

Middlemen and networks are prevalent features of markets. Our paper shows that these are possible channels of strategic delay in trade that can reduce trade volume and market liquidity. We prove a sharp characterization on the kind of networks that lead to such problems. We believe that our results are relevant to a wide range of economic settings and provide a step toward understanding market dynamics and liquidity from a network perspective.

The financial over-the-counter market is a specific market that our model can apply to. Here financial intermediaries and the network on which they trade are important considerations for

analyzing how fast can agents borrow and invest. This fits well with questions addressed in our paper. In future work, we plan to study this market deeper by extending the model to include other trade frictions such as information and risks.

Another potential application of our model includes wireless spectrum sharing markets and other dynamic supply chain systems, where markets are highly decentralized and middlemen are common. In these settings, our work naturally raises the question of designing networks to optimize trade capacity and welfare, which potentially lead to policy implications for these markets.

## References

- Admati, Anat R, Motty Perry. 1987. Strategic delay in bargaining. *The Review of Economic Studies* **54**(3) 345–364.
- Atakan, Alp E, Mehmet Ekmekci. 2013. Bargaining and reputation in search markets. *The Review of Economic Studies* 21.
- Biglaiser, Gary. 1993. Middlemen as experts. *The RAND journal of Economics* 212–223.
- Blume, L.E., D. Easley, J. Kleinberg, E. Tardos. 2009. Trading networks with price-setting agents. *Games and Economic Behavior* **67**(1) 36–50.
- Cho, In-Koo. 1990. Uncertainty and delay in bargaining. *The Review of Economic Studies* **57**(4) 575–595.
- Gale, Douglas. 1987. Limit theorems for markets with sequential bargaining. *Journal of Economic Theory* **43**(1) 20–54.
- Gofman, Michael. 2011. A network-based analysis of over-the-counter markets. *Working paper, Wisconsin School of Business* .
- Gul, Faruk, Hugo Sonnenschein. 1988. On delay in bargaining with one-sided uncertainty. *Econometrica: Journal of the Econometric Society* 601–611.
- Jehiel, Philippe, Benny Moldovanu. 1995. Cyclical delay in bargaining with externalities. *The Review of Economic Studies* **62**(4) 619–637.
- Johri, Alok, John Leach. 2002. Middlemen and the allocation of heterogeneous goods. *International Economic Review* 347–361.

- Li, Yiting. 1998. Middlemen and private information. *Journal of Monetary Economics* **42**(1) 131–159.
- Manea, Mihai. 2011. Bargaining in stationary networks. *The American Economic Review* **101**(5) 2042–2080.
- Masters, Adrian. 2007. Middlemen in search equilibrium\*. *International Economic Review* **48**(1) 343–362.
- Nguyen, Thành. 2014. Local bargaining and supply chain instability. *Working paper, Krannert School of Management, Purdue University* .
- Nguyen, Thành. 2015. Coalitional bargaining in networks. *Operations Research* **63** 501–511.
- Perry, Motty, Philip J Reny. 1993. A non-cooperative bargaining model with strategically timed offers. *Journal of Economic Theory* **59**(1) 50–77.
- Rubinstein, A., A. Wolinsky. 1987a. Middlemen. *The Quarterly Journal of Economics* **102**(3) 581–593.
- Rubinstein, Ariel. 1982. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society* 97–109.
- Rubinstein, Ariel, Asher Wolinsky. 1987b. Middlemen. *The Quarterly Journal of Economics* **102**(3) 581–593.
- Sakovics, Jozsef. 1993. Delay in bargaining games with complete information. *Journal of Economic Theory* **59**(1) 78–95.
- Schrijver, Alexander. 2002. *Combinatorial Optimization (3 volume, A, B, & C)*. Springer.
- Wright, Randall, Yuet-Yee Wong. 2014. Buyers, sellers, and middlemen: Variations on search-theoretic themes. *International Economic Review* **55**(2) 375–397.
- Yavaş, Abdullah. 1994. Middlemen in bilateral search markets. *Journal of Labor Economics* 406–429.

## Proofs of Statements

### EC.1. Proof of Theorem 1

For every  $\delta \in (0, 1)$  we need to show that there exist vectors  $\lambda, \mu, u, z$  satisfying the following conditions:

1. Stationarity: given the trading dynamics defined by  $\lambda$ , the stationary distribution of goods for each agent  $\mu$  satisfies (9);
2. Payoff-state consistency:  $u$  and  $z$  need to satisfy equations (3)-(7); and
3. Payoff-dynamic consistency: if  $z_{ij} > 0$  then  $\lambda_{ij} = 1$ ; if  $z_{ij} < 0$  then  $\lambda_{ij} = 0$ ; and if  $z_{ij} = 0$ , then  $0 \leq \lambda_{ij} \leq 1$ .

Given  $(\lambda, \mu, u, z)$ , using equations (3)-(9), we can obtain the following correspondence  $(\Lambda, \mu', u', z')$ , which we can write as follows,

$$F(\lambda, \mu, u, z) = (\Lambda, \mu', u', z'),$$

where

$$\forall p \in \mathcal{P} \quad \mu'_p = 1, \quad \forall c \in \mathcal{C} \quad \mu'_c = 0,$$

and

$$\forall m \in \mathcal{M} \quad \mu'_m = \frac{\sum_{p \in \mathcal{P}: (p, m) \in \mathcal{E}_2} \lambda_{pm}}{\sum_{p \in \mathcal{P}: (p, m) \in \mathcal{E}_2} \lambda_{pm} + \sum_{c \in \mathcal{C}: (m, c) \in \mathcal{E}_3} \lambda_{mc}},$$

and for all  $(i, j) \in \mathcal{E}$

$$\Lambda_{ij} = \{1\} \text{ if } z_{ij} > 0, \Lambda_{ij} = \{0\} \text{ if } z_{ji} < 0, \Lambda_{ij} = [0, 1] \text{ if } z_{ij} = 0.$$

Furthermore, we have:

For all  $p \in \mathcal{P}$

$$u'_0(p) = 0 \text{ and } u'_1(p) = \sum_{c: (p, c) \in \mathcal{E}_1} \frac{\max\{z_{pc}, 0\}}{2N_p(1-\delta)} + \sum_{m: (p, m) \in \mathcal{E}_2} \frac{\max\{z_{pm}, 0\}}{2N_p(1-\delta)} (1 - \mu_m).$$



For all  $m \in \mathcal{M}$

$$u'_0(m) = \sum_{p:(p,m) \in \mathcal{E}_2} \frac{\max\{z_{pm}, 0\}}{2N_m(1-\delta)}, \text{ and } u'_1(m) = \sum_{c:(m,c) \in \mathcal{E}_3} \frac{\max\{z_{mc}, 0\}}{2N_m(1-\delta)}.$$

For all  $c \in \mathcal{C}$ :

$$u'_1(c) = V_c, \text{ and } u'_0(c) = \sum_{p:(p,c) \in \mathcal{E}_1} \frac{\max\{z_{pc}, 0\}}{2N_c(1-\delta)} + \sum_{m:(m,c) \in \mathcal{E}_3} \frac{\max\{z_{mc}, 0\}}{2N_c(1-\delta)} \mu_m.$$

Furthermore,

$$z'_{ij} = \delta \left( u_1(j) - u_0(j) - (u_1(i) - u_0(i)) \right) - C_{ij} \quad \forall (i, j) \in \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3.$$

It is straightforward to check that the function  $F(\cdot, \cdot, \cdot, \cdot)$  above satisfies all the requirements for Kakutani's fixed-point theorem. The domain is a non-empty, compact and convex subset of a finite-dimensional Euclidean space. The mapping/correspondence has a closed-graph: since the mappings from  $\lambda$  to  $\mu'$  and  $(\mu, z)$  to  $u'$  and  $u$  to  $z'$  are single-valued and continuous, we only need to satisfy this for the mapping from  $z$  to  $\Lambda$ . For any sequence  $(z_n, \Lambda_n)$  (in the domain) such  $\lim_{n \rightarrow \infty} (z_n, \Lambda_n) = (z, \Lambda)$ , it is easy to see that  $\Lambda$  must lie in the image of  $z$ . Finally, the image of any point in the domain is non-empty, closed and convex. Therefore, there must be a fixed-point, and furthermore, by definition, any fixed point of this mapping is a limit stationary equilibrium.

## EC.2. Proof of Lemma 1

We will prove this result by considering a given vector of trade probabilities  $\{\lambda_{ij}\}_{(i,j) \in \mathcal{E}}$  and the middlemen good holding probabilities  $\{\mu_m\}_{m \in \mathcal{M}}$  derived from  $\{\lambda_{ij}\}_{(i,j) \in \mathcal{E}}$  as a result of the stationary constraints in (9).

Consider a middlemen node  $m$  with in-degree of  $a$  and out-degree of  $b$ . Let

$$a' = \sum_{p:(p,m) \in \mathcal{E}_2} \lambda_{pm}, \text{ and } b' = \sum_{c:(m,c) \in \mathcal{E}_3} \lambda_{mc}.$$

Then  $a' \leq a$  and  $b' \leq b$  with equality if and only if trade occurs on all links associated with middlemen  $m$  with probability 1. The stationary constraints in (9) then result in

$$\mu_m = \frac{a'}{a' + b'},$$

and the volume of trade that middleman  $m$  participates in is given by

$$(1 - \mu_m)a' = \mu_m b' = \frac{a'b'}{a' + b'} = \frac{1}{\frac{1}{a'} + \frac{1}{b'}} \leq \frac{1}{\frac{1}{a} + \frac{1}{b}} = \frac{ab}{a + b},$$

with equality if and only if trade occurs on all links associated with middlemen  $m$  with probability 1. The above analysis can be repeated for every middleman. It is also obvious that the trade volume on the producer-consumer direct links is also maximized when there no delay in the trade. This completes our proof.

### EC.3. Calculations of Example 1

The equilibrium equations for this case are as follows:

$$\begin{aligned} u_0(p) &= 0, & u_1(c) &= V, \\ u_1(p) &= \frac{(1 - \mu_m^*)(z_{pm})^+}{2N_p(1 - \delta)}, & u_0(m) &= \frac{(z_{pm})^+}{2N_m(1 - \delta)}, \\ u_1(m) &= \frac{(z_{mc})^+}{2N_m(1 - \delta)}, & u_0(c) &= \frac{\mu_m^*(z_{mc})^+}{2N_c(1 - \delta)}, \\ z_{pm} &= \delta(u_1(m) - u_0(m) - u_1(p)) - C_{12}, \\ z_{mc} &= \delta(V - u_0(c) - (u_1(m) - u_0(m))) - C_{23}, \\ \lambda_{pm} &\in \begin{cases} \{1\} & z_{pm} > 0 \\ \{0\} & z_{pm} < 0 \\ [0, 1] & z_{pm} = 0 \end{cases}, & \lambda_{mc} &\in \begin{cases} \{1\} & z_{mc} > 0 \\ \{0\} & z_{mc} < 0 \\ [0, 1] & z_{mc} = 0 \end{cases}, & \mu_m^* &= \frac{\lambda_{pm}}{\lambda_{pm} + \lambda_{mc}} \end{aligned}$$

From the above it is clear that

$$u_1(p) = \frac{(1 - \mu_m^*)N_m}{N_p} u_0(m), \text{ and } u_0(c) = \frac{\mu_m^* N_m}{N_c} u_1(m).$$

Therefore, we effectively have four parameters  $u_1(m)$ ,  $u_0(m)$ ,  $\lambda_{pm}$  and  $\lambda_{mc}$  using which we can solve for all other relevant quantities.

First consider the assumption that trade occurs with probability 1 on both links, so that  $\lambda_{pm} = \lambda_{mc} = 1$  and  $\mu_m^* = 0.5$ . For this case define

$$K_p = 2N_m(1 - \delta) + \delta \left( 1 + \frac{N_m}{2N_p} \right), \text{ and } K_c = 2N_m(1 - \delta) + \delta \left( 1 + \frac{N_m}{2N_c} \right).$$

Then by solving the resulting linear equations we obtain

$$u_1(m) = \frac{K_p(\delta V - C_{23}) - \delta a}{K_p K_c - \delta^2}, \text{ and } u_0(m) = \frac{\delta^2 V - \delta C_{23} - K_c C_{12}}{K_p K_c - \delta^2}.$$

We can rewrite the above as

$$u_1(m) = \frac{\delta \left( \left(1 + \frac{N_m}{2N_p}\right) (\delta V - C_{23}) - C_{12} \right) + (1 - \delta) 2N_m (\delta V - C_{23})}{K_p K_c - \delta^2}$$

$$u_0(m) = \frac{\delta \left( \delta V - C_{23} - \left(1 + \frac{N_m}{2N_c}\right) C_{12} \right) - (1 - \delta) 2N_m C_{12}}{K_p K_c - \delta^2}$$

From this it follows that  $u_1(m)$  is always non-negative if  $\delta V \geq \frac{C_{12}}{1 + \frac{N_m}{2N_c}} + C_{23}$ , so that our assumption on the trade probabilities is true if and only if  $u_0(m) \geq 0$ , i.e., if and only if

$$\delta V \geq C_{23} + \left(1 + \frac{N_m}{2N_c} + 2N_m \frac{(1 - \delta)}{\delta}\right) C_{12}. \quad (\text{EC.1})$$

If the inequality in (EC.1) does not hold, then  $\lambda_{pm}$  cannot be 1 leading to a delay in the trade.

Thus, next we analyze the case when  $\lambda_{pm} \in (0, 1)$  and  $\lambda_{mc} = 1$ . Here we obtain

$$u_0(p) = u_1(p) = u_0(m) = 0, \quad u_1(c) = V$$

$$u_1(m) = \frac{(z_{mc})^+}{2N_m(1 - \delta)}, \quad u_0(c) = \frac{\mu_m^*(z_{mc})^+}{2N_c(1 - \delta)},$$

$$z_{pm} = \delta u_1(m) - C_{12} = 0, \quad z_{mc} = \delta(V - u_0(c) - u_1(m)) - C_{23} \geq 0$$

$$\mu_m^* = \frac{\lambda_{pm}}{\lambda_{pm} + 1}$$

Solving we get

$$u_1(m) = \frac{C_{12}}{\delta}, \quad \lambda_{pm} = \frac{\mu_m^*}{1 - \mu_m^*},$$

$$\mu_m^* = \frac{\delta V - C_{23} - \left(1 + 2N_m \frac{1 - \delta}{\delta}\right) C_{12}}{\frac{N_m}{N_c} C_{12}}$$

It is easily observed that  $\mu_m^* \leq 1/2$ . Note that such an equilibrium is sustainable if and only if we have

$$\delta V \in \left( C_{23} + \left(1 + 2N_m \frac{1 - \delta}{\delta}\right) C_{12}, C_{23} + \left(1 + \frac{N_m}{2N_c} + 2N_m \frac{(1 - \delta)}{\delta}\right) C_{12} \right).$$

If, instead,  $\delta V$  is at most

$$C_{23} + \left(1 + 2N_m \frac{1-\delta}{\delta}\right) C_{12},$$

then no trade happens on the first link, so that trade stops completely. Note that the bargaining procedure adds a penalty term of  $2N_m \frac{1-\delta}{\delta} C_{12}$  to the value in order for trade to even be feasible.

We conclude by considering the limit of  $\delta \uparrow 1$ . Here, owing to the increased patience of the agents, trade is feasible whenever  $V > C_{12} + C_{23}$ . However, if

$$V \in \left(C_{12} + C_{23}, \left(1 + \frac{N_m}{2N_c}\right) C_{12} + C_{23}\right),$$

then there is a delay in the trade, and we have

$$\begin{aligned} u_0(p) = u_1(p) = u_0(m) = 0, & & u_1(c) = V, \\ u_1(m) = C_{12}, & & u_0(c) = V - C_{23} - C_{12}, \\ \lambda_{pm} = \frac{V - C_{12} - C_{23}}{\left(1 + \frac{N_m}{N_c}\right) C_{12} - C_{23} - V}, \lambda_{mc} = 1, & & \mu_m^* = \frac{V - C_{23} - C_{12}}{\frac{N_m}{N_c} C_{12}}, \\ d_1 = 0, d_2 = C_{12}, & & d_c = C_{12} + C_{23}. \end{aligned}$$

If, instead,  $V \geq \left(1 + \frac{N_m}{2N_c}\right) C_{12} + C_{23}$ , then always-trade strategy is the unique equilibrium and

$$\begin{aligned} u_0(p) = 0, & & u_1(p) = \frac{N_m}{2N_p} \frac{V - C_{23} - C_{12} \left(1 + \frac{N_m}{2N_c}\right)}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}, \\ u_0(m) = \frac{V - C_{23} - C_{12} \left(1 + \frac{N_m}{2N_c}\right)}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}, & & u_1(m) = \frac{\left(1 + \frac{N_m}{2N_p}\right) (V - C_{23}) - C_{12}}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}, \\ u_0(c) = \frac{N_m}{2N_c} \frac{\left(1 + \frac{N_m}{2N_p}\right) (V - C_{23}) - C_{12}}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}, & & u_1(c) = V, \\ \lambda_{pm} = \lambda_{mc} = 1, & & \mu_m^* = 0.5, \\ d_1 = \frac{N_m}{2N_p} \frac{V - C_{23} - C_{12} \left(1 + \frac{N_m}{2N_c}\right)}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}, & & d_2 = \frac{\frac{N_m}{2N_p} (V - C_{23}) - \frac{N_m}{2N_c} C_{12}}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}, \\ d_3 = \frac{\frac{N_m}{2N_p} V + \frac{N_m}{2N_c} \left(1 + \frac{N_m}{2N_p}\right) C_{23} + \frac{N_m}{2N_c} C_{12}}{\left(1 + \frac{N_m}{2N_p}\right) \left(1 + \frac{N_m}{2N_c}\right) - 1}. & & \end{aligned}$$

### EC.4. Proof of Theorem 3

Consider equations (4)–(6). As  $\delta$  approaches 1, the  $1 - \delta$  term approaches 0. Since  $u_1(p) \in [0, \max_{c \in \mathcal{C}} V_c]$  for all  $i \in \mathcal{P}$ , and  $u_0(c) \in [0, \max_{c \in \mathcal{C}} V_c]$  for all  $c \in \mathcal{C}$ , it has to be that given any  $\epsilon > 0$ , there exists  $\delta^*$  such that for all  $\delta > \delta^*$ , we have

$$z_{pc} \leq \epsilon \quad \forall (p, c) \in \mathcal{E}_1,$$

$$z_{pm} \leq \epsilon \quad \forall (p, m) \in \mathcal{E}_2,$$

$$z_{mc} \leq \epsilon \quad \forall (m, c) \in \mathcal{E}_3.$$

Now consider a pair of agents, one producer  $p$  and consumer  $c$ . We have three cases then:

1. All trade routes from  $p$  to  $c$  have to visit some middleman. Let  $m \in \mathcal{M}$  be one such middleman so that  $(p, m) \in \mathcal{E}_2$  and  $(m, c) \in \mathcal{E}_3$ . The inequalities above then imply the following:

$$\begin{aligned} u_1(m) - u_0(m) &\geq u_1(c) - u_0(c) - \frac{C_{mc} + \epsilon}{\delta}, \\ u_1(m) - u_0(m) &\leq u_1(p) - u_0(p) + \frac{C_{pm} + \epsilon}{\delta}. \end{aligned}$$

These with  $u_0(p) = 0$  and  $u_1(c) = V_c$  imply

$$u_1(p) + u_0(c) \geq V_c - \frac{C_{pm} + C_{mc} + 2\epsilon}{\delta}.$$

Note that this inequality holds for every  $m \in \mathcal{M}$  that lies along a trade route from  $p$  to  $c$ . Therefore,

$$u_1(p) + u_0(c) \geq V_c - \frac{\min_{\{m:(p,m) \in \mathcal{E}_2 \text{ and } (m,c) \in \mathcal{E}_3\}} (C_{pm} + C_{mc}) + 2\epsilon}{\delta}.$$

Since  $\epsilon$  can be chosen arbitrarily small, thus for any middleman  $m$  who is not on a smallest transaction cost path from  $p$  to  $c$ , we can choose  $\delta$  close enough to 1 such that either  $z_{pm}$  or  $z_{mc}$  is strictly negative and so no trade can occur on the corresponding edge;

2. Notice that the same argument also works for the case, where if in addition to the middlemen, there also exists a direct link between  $p$  and  $c$ . Then

$$u_1(p) + u_0(c) \geq V_c - \min \left( C_{pc}, \min_{\{m:(p,m) \in \mathcal{E}_2 \text{ and } (m,c) \in \mathcal{E}_3\}} (C_{pm} + C_{mc}) \right).$$

Again it is clear that no trade occurs over links that are not part of a smallest transaction cost path from  $p$  to  $c$ ;

3. If  $p$  and  $c$  only have a direct route between them, then that is the only route via which trade can occur between this producer and consumer pair. Also, if no routes exist between  $p$  and  $c$ , then obviously no trade occurs between these two agents.

### EC.5. Proof of Lemma 2

Since the network is connected, if we fix a value of  $d_1$ , then because of the constraints  $d_b - d_a = C_{ab}$ , all other  $d_i$ -s are determined uniquely. Furthermore, if we change  $d_1$  by  $x$ , then all other  $d_i$ -s are changed by the same amount. Therefore, all solutions  $d_i$  will be of the form  $d_i = d + \Delta_i$  for  $d \geq 0$ . To find  $\Delta_i$  we need to find  $d_1$  such that the minimum value  $\min_i \Delta_i = 0$ .

The following is an algorithm to determine  $\Delta_i$ . Let  $T$  be a spanning tree. Fix a node 1 in  $T$ . For a node  $i \in T$ , assume the unique path connecting 1 and  $i$  is  $1 = i_1, i_2, \dots, i_k, i_{k+1} = i$ . Let  $X_i = \sum_{j=1}^k \text{sign}_{i_j, i_{j+1}} \cdot C_{i_j, i_{j+1}}$ . Clearly,  $d_i = d_1 + X_i$ . Let  $X = \min_i X_i$ . Thus,  $d_i = (d_1 + X) + (X_i - X)$ . Let  $\Delta_i := X_i - X$ . These are the values of  $\Delta_i$ -s that we needed to determine. ■

### EC.6. Feasibility of the always-trade strategy for the example in Section 5.1

Here we present analysis of the always-trade strategy for the example in Section 5.1 for a scenario where feasible trading partners always trade if they are matched (i.e. there is no strategic bargaining). For such a setting, we seek to understand when the always-trade strategy generates non-negative returns for each trade. In the stationary regime, under the always-trade strategy, with probability 0.5 a middleman will hold the good. Hence, it takes  $N_1 + N_2$  units of time for a tagged good to make it from the producer to the consumer where  $N_1$  and  $N_2$  are independent geometric random variables with

$$\mathbb{P}(N_1 = k) = \frac{1}{2N_p} \left(1 - \frac{1}{2N_p}\right)^{k-1} \quad \text{and} \quad \mathbb{P}(N_2 = k) = \frac{1}{N_m} \left(1 - \frac{1}{N_m}\right)^{k-1}.$$

Note that  $N_1$  depends on both identifying a middleman without any good and picking the producer with the tagged good, and  $N_2$  depends on picking the specific middleman with the tagged good;

the given expressions then follow from the definition of the matching process. At the end of  $N_1$ , the good reaches a middleman and a transaction cost of  $C_{12}$  is paid. While at the end of  $N_1 + N_2$ , the good reaches a consumer and a transaction cost of  $C_{23}$  needs to be paid. The value of  $V$  is obtained at the end of the transaction.

Define by  $u_p$  the return from a producer-middleman trade and  $u_m$  the return from a middleman-consumer trade. These can be obtained using the following equations

$$\begin{aligned} u_p &= \left(1 - \frac{1}{2N_p}\right) \delta u_p + \frac{1}{2N_p} (\delta u_m - C_{12}), \\ u_m &= \left(1 - \frac{1}{N_m}\right) \delta u_m + \frac{1}{N_m} (\delta V - C_{23}), \end{aligned}$$

where the Bellman equations follow in a simple manner by considering the probability of getting a feasible match. Solving we get

$$u_m = \frac{\delta V - C_{23}}{N_m - (N_m - 1)\delta}, \text{ and } u_p = \frac{\delta u_m - C_{12}}{2N_p - (2N_p - 1)\delta} = \frac{\delta^2 V - \delta C_{23} - (N_m - (N_m - 1)\delta)C_{12}}{(2N_p - (2N_p - 1)\delta)(N_m - (N_m - 1)\delta)}.$$

We define always-trade to be feasible if both  $u_m$  and  $u_p$  are non-negative. It follows that this is the case if and only if

$$\delta V \geq C_{23} + \left(\frac{N_m}{\delta} - (N_m - 1)\right) C_{12}.$$

If there are any delays in trade on the second link (owing to non-strategic reasons), then the parameters of the  $N_2$  random variable change leading to a reduction in the expected reward from the trade. This would then require even greater value for trade to be feasible. Note that delay on the first link (due to the matching process) does not impact the feasibility of trade; . this is because the transaction cost  $C_{12}$  is not paid until the trade is completed.

Given  $V$ ,  $C_{12}$  and  $C_{23}$  this means that any  $\delta < \delta^*$  yields negative return (and non-negative return otherwise) where

$$\delta^* = \frac{C_{23} - (N_m - 1)C_{12} + \sqrt{(C_{23} - (N_m - 1)C_{12})^2 + 4N_m V C_{12}}}{2V}.$$

Similarly if  $V$ ,  $\delta$  and  $C_{23}$  are given, then we need  $C_{12} \leq C_{12}^*$  to obtain non-negative return where

$$C_{12}^* = \frac{\delta V - C_{23}}{\frac{N_m}{\delta} - (N_m - 1)}.$$

## Endnotes

1. Notice that it is straightforward to generalize the model such that the transaction costs vary among the multiple links, the  $z_{ij}$  and the resulting  $\lambda_{ij}$  variables could be different for each of these links. For ease of presentation, we consider these values to be the same for parallel links.

## Acknowledgments

The authors thank the associate editors and two anonymous referees for their insightful comments and suggestions. Thanh Nguyen's work was partly support by NSF grants AST-1343381 and AST-1443965. Vijay Subramanian's work was partly support by NSF via grants IIS-1538827, AST-1343381 and AST-1516075. Randall Berry's work was partly support by NSF grants IIS-1538827 and AST-1343381.