Strategic Formation and Reliability of Supply Chain Networks

(Working Paper, January 10, 2020)

Victor Amelkin
Warren Center for Network and Data Sciences, Department of Electrical and Systems Engineering
University of Pennsylvania, Philadelphia, PA 19104, vctr@seas.upenn.edu, https://victoramelkin.com

Rakesh Vohra
Warren Center for Network and Data Sciences, Department of Electrical and Systems Engineering, Department of Economics
University of Pennsylvania, Philadelphia, PA 19104, rvohra@seas.upenn.edu, https://sites.google.com/site/quaerereverum9/

Supply chains are the backbone of the global economy. Disruptions to them can be costly. Centrally managed supply chains invest in ensuring their resilience. Decentralized supply chains, however, must rely upon the self-interest of their individual components to maintain the resilience of the entire chain.

We examine the incentives that independent self-interested agents have in forming a resilient supply chain network in the face of production disruptions and competition. In our model, competing suppliers are subject to yield uncertainty (they deliver less than ordered) and congestion (lead time uncertainty or, "soft" supply caps). Competing retailers must decide which suppliers to link to based on both price and reliability. In the presence of yield uncertainty only, the resulting supply chain networks are sparse. Retailers concentrate their links on a single supplier, counter to the idea that they should mitigate yield uncertainty by diversifying their supply base. This happens because retailers benefit from supply variance. It suggests that competition will amplify output uncertainty. When congestion is included as well, the resulting networks are denser and resemble the bipartite expander graphs that have been proposed in the supply chain literature, thereby, providing the first example of endogenous formation of resilient supply chain networks, without resilience being explicitly encoded in payoffs. Finally, we show that a supplier's investments in improved yield can make it worse off. This happens because high production output saturates the market, which, in turn lowers prices and profits for participants.

Key words: supply chain, supply chain network, strategic network formation, market clearing, disruptions, reliability, yield uncertainty, lead time uncertainty, congestion, pure strategy Nash equilibrium

---

(The most recent version of this paper is available at `https://victoramelkin.com/pub/supply-chains/`.)

Contents

1. Introduction

Supply chains are the backbone of the global economy. Disruptions to them can be costly. They happen because the individual components of the chain are subject to yield uncertainty (a supplier comes up short on the ordered product quantity) as well as lead time uncertainty (clients of overly congested suppliers experience delivery delays) (Snyder et al. 2016). The degree of uncertainty can be large. Bohn and Terwiesch (1999), for example, suggest that disk drive manufacturer Seagate experiences production yields as low as 50%. Centrally managed and controlled supply chains invest a great deal in mitigating these disruptions. Decentralized supply chains, however, must rely upon the self-interest of their individual components to maintain the resilience of the entire chain.

Cursory reflection suggests that the incentives of an individual component in the chain should align with the chain as a whole. A supplier, for example, will be rewarded with greater business if it invests in reducing the possibility of it being disrupted relative to its competitors. However, a supplier's customers can also hedge against disruption by multi-sourcing (Cachon and Terwiesch 2008, Chopra and Meindl 2016, Tomlin 2006). Thus, a potential customer may prefer to source from many low-cost unreliable suppliers rather than a few highly reliable but costlier suppliers. Cursory reflection also ignores the impact on output prices that would result from reducing the frequency of disruptions. If prices adjust to clear markets, increased throughput may result in lower prices. Thus, one must compare the profits earned from high volumes with low margins with those generated from lower volumes but with higher margins. It is not obvious which will dominate.

In this paper we examine the strategic formation of a two-tier[1] supply chain network by *independent* self-interested agents. Retailers occupy the first tier and suppliers the second tier. The price at which trade takes place between tiers is set to clear the market. Retailers decide which suppliers to source from. There is also a cost for linking to a supplier.[2]

Every agent present in the supply chain is subject to yield uncertainty which affects their capacity.[3] It is modeled as a Bernoulli random variable. Yield uncertainty of this kind can arise from the nature of the production process (e.g., farming); it can also arise from disruptions like a natural disaster or a union strike. The resulting random output of each tier is distributed among agents in the downstream tier in proportion to their demands, following the proportional rationing rule (Rong et al. 2017, Cachon and Lariviere 1999b). Every supplier in the chain is also subject to congestion, and the resulting congestion costs are borne by the retailers. Congestion in our model has at least two interpretations. One is a delay cost associated with lead time uncertainty. The second is a "soft" supply constraint.

We are interested in whether a *decentralized* supply chain network resilient to disruptions will form *endogenously* in the presence of competition and different types of uncertainty.

The *three major findings* of our analyses are as follows:

▷ With only yield uncertainty and no congestion, *retailers create a sparse network, with a single link per retailer, and concentrate links on a single supplier.* This generalizes to the case of more than two tiers where the corresponding supply chain network is almost a chain. Link concentration runs counter to the common wisdom about the benefits of multi-sourcing. *Link concentration helps retailers secure low upstream prices* in the presence of

[1] Our model straightforwardly generalizes to the case of an arbitrary number of tiers.

[2] According to Cormican and Cunningham (2007), it takes, on average, six months to a year to qualify a new supplier.

[3] To quote Yossi Sheffi: "The essence of most disruptions is a reduction in capacity and, therefore, inability to meet demand."

high upstream yield, and low expenditures in case of the target upstream supplier's failure. Therefore, retailers benefit from supply variance. It suggests that competition can amplify output uncertainty. The network formed in our model is dramatically different from the ones that are *assumed* in the existing literature. Bimpikis et al. (2019), for example, assume that, in the presence of yield uncertainty only, a k-tier supply chain network will take the form of a complete k-partite graph.

⊳ In the presence of yield uncertainty *and* congestion, the network formed is *sparse, yet well-connected* resembling an expander graph. Similar objects have been shown to have good resilience properties in the context of centrally organized supply chains, see for example, Chou et al. (2011). Congestion, unlike yield uncertainty, encourages retailers to split their demand across several suppliers to lower congestion costs. Thus, our work provides the first example of the endogenous formation of resilient supply chain networks, without an explicit concern for resilience being encoded in the payoffs.

⊳ *Yield uncertainty and congestion have fundamentally different implications for supply chains.* In the presence of yield uncertainty only, each supplier has a unilateral incentive to increase its average yield. With both yield uncertainty *and* congestion, a unilateral reduction in congestion costs unconditionally benefits that supplier, but increasing mean yield could make a supplier worse off! This is because high yield results in market saturation, which leads to low prices and profits for market members.

The rest of the paper is organized as follows. The next section discusses prior work. The subsequent section introduces notation. Sec. 4 describes the model of strategic formation of supply chains with costly links, competition, and yield uncertainty only. Sec. 5 augments the previous model with congestion, and provides its comprehensive characterization. Sec. 5.2 focuses on the case of a small two-tier supply chain. Sec. 5.3 provides a limited set of results for the general two-tier case. Finally, in Sec. 5.4, we describe the quality-investment behavior of competing heterogeneous suppliers and the qualitative differences between yield uncertainty and congestion / lead time uncertainty.

2. Prior Work

Our model has three features:

1. strategic network formation,
2. disruptions, and
3. competition.

In the extensive literature on supply chain networks, one will find models that possess some, but not all three features, with the exception of one recent model of Amelkin and Vohra (2019). Table 1 categorizes a sample of recent related works. A detailed comparison with prior work on network formation in supply chains follows.

There are many papers that study supply chains in the presence of competition. Examples are Carr and Karmarkar (2005) and Fang and Shou (2015) which use Cournot competition, while Chod et al. (2019) uses Bertrand competition. Some also incorporate disruptions, such as Deo and Corbett (2009), and Babich et al. (2007). However, none of considers endogenous network formation.

A significantly smaller set of papers compare supply chain performance across different network structures, but do not consider endogenous formation. Within this stream two papers are closely related to ours. The first is Bimpikis et al. (2019). We share the same price formation process in every tier of the supply chain and the same production model with Bernoulli yield. Our supply chain network, however, is endogenously formed while

Paper \ Feature	Network Formation	Disruptions	Competition
Present work	+	+	+
Amelkin and Vohra (2019)	+	+	+
Bimpikis et al. (2018)	+	+	−
Ang et al. (2016)	+	±	−
Tang and Kouvelis (2011)	±	+	±
Chod et al. (2019)	−	±	+
Bimpikis et al. (2019)	−	+	+
Fang and Shou (2015)	−	+	+
Adida and DeMiguel (2011)	−	+	+
Babich et al. (2007)	−	+	+
Carr and Karmarkar (2005)	−	−	+
Bernstein and Federgruen (2005)	−	+	+
Anupindi and Akella (1993)	−	+	−
Kotowski and Leister (2018)*	+	±	+
Ambrus et al. (2014)*	−	+	−
Deo and Corbett (2009)*	−	+	±
Kranton and Minehart (2001)*	+	±	+

Table 1 Prior work summary. Partial presence (±) of disruption means that they are present only in a part of the system, e.g., in the upstream tier of suppliers (Ang et al. 2016) or the tier of buyers / retailers (Chod et al. 2019), while in our model every agent can be disrupted. Also, none of the mentioned works considers anything analogous to our congestion or lead time uncertainty. Absence (−) of competition implies fixed prices. Works marked with an asterisk are not about supply chains.

theirs is exogenously fixed to be a complete k-partite graph. Our results cannot be deduced from the model of Bimpikis et al. (2019). In particular, our analysis in Sec. 4 demonstrates that the complete k-partite networks assumed in Bimpikis et al. (2019) need not arise endogenously in their setting.

The second work is Tang and Kouvelis (2011), with two competing reliable retailers sourcing from two non-competing suppliers subject to yield uncertainty that is correlated. Because supplier prices are fixed exogenously, the focus is on how order quantities change with the sourcing decisions of retailers. This paper compares outcomes across possible networks, but does not consider all possible networks. In particular, the case when retailers single-source from the *same* supplier is excluded. In one of our models, this configuration arises in equilibrium.

Tang and Kouvelis (2011) also study the effect of an exogenously given correlation between supplier yields. In our model, supplier yields are independent of each other. However, when different retailers in our model have overlap in their supplier bases, it results in an implicit correlation of their production outputs. This implicit correlation influences network formation in our model.

Papers that consider endogenous supply chain network formation can be numbered on a single hand. Amelkin and Vohra (2019) consider an endogenous supply chain network formation model, with yield uncertainty and competition. In their model, suppliers, having uncertain i.i.d. supplies, strategically announce wholesale prices to retailers, and the retailers, then, compete to sell the product to consumers. There are two major differences between their model and ours. In Amelkin and Vohra (2019), wholesale prices are set in advance while in the present paper they are set to clear the market as in Bimpikis et al. (2019). This seems particularly relevant to supply chains where prices are determined by a spot market, for example, the memory chip industry[4] (Bimpikis et al. 2019) and the

[4] `https://www.dramexchange.com`

US corn, soybeans, wheat, and tobacco industries (Mendelson and Tunca 2007). Even in markets where prices are set by contract, they may be pegged to an index, such as the average spot price on an agreed upon date. The equilibrium outcomes under a spot price are radically different. The second difference is that, in Amelkin and Vohra (2019), retailers are quantity-takers, that is, they have unconstrained demands, while in our model, there is a specific demand originating at the consumer tier, which, then, "propagates" through the supply chain network based on how the latter is structured. These two differences produce different network outcomes in equilibrium.

Another work capturing endogenous network formation is Ang et al. (2016), who consider a supply chain comprised of a manufacturer issuing orders to two suppliers, who, in turn, are linked to two higher-level suppliers. The manufacturer issues contracts incorporating quantities and prices, thereby, affecting sourcing decisions of intermediary suppliers. In this model only top-tier suppliers fail, while every agent can fail in our model. Further, unlike our paper, there is no competition: top-tier sourcing costs—different for reliable (higher) and unreliable (lower) suppliers—are exogenously fixed; so is the price at which the manufacturer sells a unit of product to consumers.

Ang et al. (2016) are interested in how the intermediate suppliers decide between single- vs. multi-sourcing decisions. In their paper, the manufacturer, primed by a deterministic exogenous consumer demand issues price-quantity contracts to the intermediary suppliers, and each of the intermediaries decides upon how much to order from each of the top-tier suppliers at the fixed prices. Thus, sourcing decisions of intermediate suppliers are the result of strategic price-setting by the manufacturer. In our model, prices are formed via competition in every tier of the supply chain. Thus, the formed networks we observe are a result of competition in all tiers.

Ang et al. (2016) characterize optimal sourcing strategies (optimal order quantities) in a network where intermediaries source from different top-tier suppliers (V-shaped network), and in the network where they source from the same top-tier supplier (diamond-shaped network). They also propose and analyze a game where intermediate suppliers strategically decide upon sourcing from the top tier, and characterize its pure equilibria, which happen to be the diamond- and V-shaped networks. V-shaped networks are equilibria in our network formation game as well, but we find other equilibria as well (see Sec. 5.2). It is important to emphasize that V-shaped equilibria are absent in our model when yield uncertainty is the *only* disruption present (see Sec. 4). Hence, if competition was incorporated into Ang et al. (2016), it would eliminate the V-shaped equilibria.

The presence of competition in our work (and its absence in Ang et al. (2016)) is crucial, as competition—additionally augmented with congestion in our model—largely drives strategic link formation in the supply chain network.

Finally, this paper is related to the literature on strategic network formation, including the work on buyer-seller networks of Kranton and Minehart (2001), trading networks with intermediaries of Kotowski and Leister (2018), and risk-sharing networks of Ambrus et al. (2014). However, the multipartite structure of our networks, the mechanics of our models, and the conclusions we arrive at are both different and distant from the ones in these works.

3. Preliminaries and Notation

In this section, we introduce notation—summarized in Table 2—and several useful definitions.

Supply Chain: A supply chain is a multi-tier network comprised of T tiers of agents—also known as firms or suppliers—where tier t is denoted with $\mathbb{T}_t = \{1, 2, \dots\}$. Most of our modeling efforts will target 2-tier supply chains, in which $T = 2$, $|\mathbb{T}_1| = n > 1$, $|\mathbb{T}_2| = m > 1$. The agents in tier $\mathbb{T}_1$ are referred to as *retailers*, who sell product to consumers; higher-tier agents are *suppliers*. When using notation independent of the total number of tiers, we may also refer to any agent in any tier of the supply chain as a supplier. Implicitly present is tier $\mathbb{T}_0$ of consumers, and another tier $\mathbb{T}_{T+1}$ corresponding to the raw material market[5]. The product will flow from higher numbered tiers (*upstream*) to lower numbered tiers (*downstream*). Throughout this paper, we assume that downstream agents strategically link to upstream agents in the adjacent tier. All retailers are linked to consumers, and all top-level suppliers are linked to raw materials producers.

Demands: Each retailer will experience a fixed consumer demand of $D > 0$. By considering equal consumer demand distribution over the retailers, we make sure that the retailers differ only in what suppliers they link to. $\Delta = nD$ is the total consumer demand. More generally, we denote by $D_{t,i} \in \mathbb{R}_+$ the demand experienced by agent $i \in \mathbb{T}_t$, indicating the amount of product collectively requested from supplier i by downstream agents from tier $\mathbb{T}_{t-1}$.

Prices: Each tier $\mathbb{T}_1, \dots, \mathbb{T}_{T+1}$ of the supply chain consists of agents competing to supply agents in the adjacent downstream tier. Within a tier, "supply" is the *realized* total quantity present in the tier. The realized quantity may be lower than the demanded quantity due to upstream production failures. Denote the supply of $i \in \mathbb{T}_t$ by $S_{t,i}$, and the total supply of tier $\mathbb{T}_t$ by

$$S_t = \sum_{i \in \mathbb{T}_t} S_{t,i}. \tag{1}$$

The market price p_t per unit of output of tier t is set so as to "clear" the market, i.e.,

$$p_t = \Delta - S_t. \tag{2}$$

Production: Each supplier, having received some product quantity, supplies the same amount of product downstream—provided there are downstream agents linked to this supplier—with probability $\lambda \in (0, 1)$, and fails to produce any output with the complementary probability $(1 - \lambda)$. We exclude $\lambda \in \{0, 1\}$ to avoid trivialities. If $\lambda = 0$, the agents clearly cannot make a profit. Setting $\lambda = 1$ entails the same degenerate outcome, the reasons for which are given in Theorem 2. Raw material producers never fail.

We use $R_{t,i} \in \mathbb{R}_+$ to denote the *realized demand* of supplier $i \in \mathbb{T}_t$, that is, how much supplier i receives from upstream suppliers in $\mathbb{T}_{t+1}$ in response to i's demand of $D_{t,i}$. The *production success indicator* $\omega_{t,i} \sim \text{Bernoulli}(\lambda)$ is a random variable that indicates whether supplier $i \in \mathbb{T}_t$ has succeeded in producing output.

Network: All suppliers together with their links comprise the *network* underlying the supply chain. Let $\mathcal{N}^-_{t,i} \subseteq \mathbb{T}_{t-1}$ and $\mathcal{N}^+_{t,i} \subseteq \mathbb{T}_{t+1}$ denote in- and out-neighborhoods of supplier $i \in \mathbb{T}_t$, that is, the sets of suppliers that source product from i or that i sources product from, respectively. Thus, the network formed by suppliers of tier $\mathbb{T}_t$ is $g_t = (\mathcal{N}^+_{t,1}, \dots, \mathcal{N}^+_{t,n_t})$, with $g = g_1$ being used in the analysis of two-tier chains. We also define

[5] The consumer and the raw material producer tiers will actually be represented by a single meta-consumer and meta-raw material producer, respectively.

$g_t^{-i} = (\mathcal{N}_{t,1}^+, \ldots, \mathcal{N}_{t,i-1}^+, \mathcal{N}_{t,i+1}^+, \ldots \mathcal{N}_{t,n_t}^+)$, and $g^{-i} = g_1^{-i}$ for two-tier chains. In- and out-degrees of i are $d_{t,i}^- = |\mathcal{N}_{t,i}^-|$ and $d_{t,i}^+ = |\mathcal{N}_{t,i}^+|$, respectively. $d_{i \cap i'}^+ = |\mathcal{N}_{t,i}^+ \cap \mathcal{N}_{t,i'}^+|$ stands for the size of out-neighborhood overlap of suppliers $i, i' \in \mathbb{T}_t$. Additionally, we introduce the effective out-degree $e_{t,i}^+ = \sum_{j \in \mathcal{N}_{t,i}^+} \omega_{t+1,j}$ of supplier $i \in \mathbb{T}_t$, that measures the number of its out-neighbors who successfully produced output. We will say that a supplier is *active* if its in- and out-degrees are both positive; other suppliers are inactive (and they cannot possibly earn profits due to their inability to either buy or sell). By $\mathbb{T}_t^a \subseteq \mathbb{T}_t$ we will denote the *subset of active suppliers* of tier $\mathbb{T}_t$, and the *number of active suppliers* is $n_t^a = |\mathbb{T}_t^a| \le n_t$.

Finally, we introduce the following expressions useful in the analysis of our models:

$$\rho_{t,i}^+ = \sum_{i' \in \mathbb{T}_t^a \setminus \{i\}} \frac{d_{i \cap i'}^+}{d_{t,i'}^+} = \sum_{i' \in \mathbb{T}_t^a \setminus \{i\}} \frac{|\mathcal{N}_{t,i'}^+ \cap \mathcal{N}_{t,i}^+|}{d_{t,i'}^+}, \tag{3}$$

$$F_{t,j} = \sum_{i \in \mathcal{N}_{t,j}^- \setminus \{i\}} \frac{1}{d_{t-1,i}^+}, \tag{4}$$

$$F_{t,j}^{-i} = \sum_{i' \in \mathcal{N}_{t,j}^- \setminus \{i\}} \frac{1}{d_{t-1,i'}^+}. \tag{5}$$

Here, $\rho_{t,i}^+$ measures the aggregate relative extent to which out-neighborhoods of active suppliers in tier $\mathbb{T}_t$ overlap with the out-neighborhood of supplier i, or, less formally, how well supplier i is "embedded" in its tier. Thus, we will refer to ρ^+ as the *degree of overlap of* i *with its peer suppliers.* As suppliers distribute their demand uniformly over out-neighborhoods, $F_{t,j}$ quantifies (scaled) *congestion at supplier* $j \in \mathbb{T}_t$ (where "congestion" is understood with respect to the demand coming from downstream agents), and $F_{t,j}^{-i}$ measures the same quantity excluding the impact of supplier i. Seemingly different, $\rho_{t,i}^+$ and $F_{t+1,j}^{-i}$ are actually closely related, as shown in the next Lemma.

LEMMA 1 **(About $\rho_{t,i}^+$ and $F_{t+1,j}^{-i}$)**.

$$\rho_{t,i}^+ = \sum_{j \in \mathcal{N}_{t,i}^+} F_{t+1,j}^{-i}.$$

Proof of Lemma 1:

$$\rho_{t,i}^+ = \sum_{i' \in \mathbb{T}_t^a \setminus \{i\}} \frac{d_{i \cap i'}^+}{d_{t,i'}^+} = \sum_{i' \in \mathbb{T}_t^a \setminus \{i\}} \frac{|\mathcal{N}_{t,i'}^+ \cap \mathcal{N}_{t,i}^+|}{d_{t,i'}^+} = \sum_{i' \in \mathbb{T}_t^a \setminus \{i\}} \sum_{j \in \mathcal{N}_{t,i'}^+ \cap \mathcal{N}_{t,i}^+} \frac{1}{d_{t,i'}^+} = \sum_{i \to j \leftarrow i' \neq i} \frac{1}{d_{t,i'}^+}$$
$$= \sum_{j \in \mathcal{N}_{t,i}^+} \sum_{i' \in \mathcal{N}_{t+1,j}^- \setminus \{i\}} \frac{1}{d_{t,i'}^+} = \sum_{j \in \mathcal{N}_{t,i}^+} F_{t+1,j}^{-i}.$$

■

When dealing with two-tier supply chains, we call the underlying network *left-regular* if all retailers have identical out-degrees, and *right-regular* if all suppliers have identical in-degrees.

We restrict attention to pure strategy Nash equilibrium.

$\mathbb{T}_t$	$\{1,\ldots,n_t\}$ – tier $t\in\{1,\ldots,T\}$ of the supply chain; $\lvert\mathbb{T}_1\rvert=n$, $\lvert\mathbb{T}_T\rvert=m$, $\lvert\mathbb{T}_t=n_t\rvert$
$D_{t,i}$	demand exerted upon supplier $i\in\mathbb{T}_t$ by downstream suppliers from $\mathbb{T}_{t-1}$
D	$D_{1,i}=\text{const}$ – consumer demand per retailer
$R_{t,i}$	realized demand of supplier $i\in\mathbb{T}_t$; $R_{t,i}\le D_{t,i}$
Δ	nD – total consumer demand
$S_{t,i}$	supply delivered by supplier $i\in\mathbb{T}_t$ to downstream suppliers sourcing from i
S_t	$\sum_{i\in\mathbb{T}_t} S_{t,i}$ – total supply of tier $\mathbb{T}_t$
$\omega_{t,i}$	$\omega_{t,i}\sim\text{Bernoulli}(\lambda)$ – production success indicator of supplier $i\in\mathbb{T}_t$
λ	$\mathbb{P}\{\omega_{t,i}=1\}\in(0,1)$ – production success likelihood
$\mathcal{N}^+_{t,i}$	$\mathcal{N}^+_{t,i}\subseteq\mathbb{T}_{t+1}$ – out-neighborhood of supplier $i\in\mathbb{T}_t$; ($\mathcal{N}^-_{t,i}\subseteq\mathbb{T}_{t-1}$ – in-neighborhood)
g_t	$(\mathcal{N}^+_{t,1},\ldots,\mathcal{N}^+_{t,n_t})$ – network between tiers $\mathbb{T}_t$ and $\mathbb{T}_{t+1}$; when $T=2$, $g=g_1$
$d^+_{t,i}$	$\lvert\mathcal{N}^+_{t,i}\rvert$ – out-degree of supplier $i\in\mathbb{T}_t$; ($d^-_{t,i}$ – in-degree)
$d^+_{t,i\cap i'}$	$\lvert\mathcal{N}^+_{t,i}\cap\mathcal{N}^+_{t,i'}\rvert$ – number of out-neighbors that $i,i'\in\mathbb{T}_t$ share
$e^+_{t,i}$	$\sum_{j\in\mathcal{N}^+_{t,i}}\omega_{t+1,j}$ – effective out-degree of supplier $i\in\mathbb{T}_t$
$\mathbb{T}^a_t$	$\{i\in\mathbb{T}_t \mid d^-_{t,i}\cdot d^+_{t,i}>0\}$ – subset of active suppliers in tier $\mathbb{T}_t$
n^a_t	$\lvert\mathbb{T}^a_t\rvert$ – number of active suppliers in tier $\mathbb{T}_t$; $0\le n^a_t\le n_t$
p_t	market price of a unit of product at which tier $\mathbb{T}_t$ sells downstream
$\rho^+_{t,i}$	aggregate relative extent of overlap of out-neighborhoods in $\mathbb{T}_t$ with $\mathcal{N}^+_{t,i}$
$F_{t,j}$	congestion at supplier $j\in\mathbb{T}_t$
$F^{-i}_{t,j}$	$F_{t,j}$ excluding the contribution of $i\in\mathbb{T}_{t-1}$
c	constant cost of linking to an upstream supplier
γ	constant congestion cost
$\pi_{t,i}$	payoff of supplier $i\in\mathbb{T}_t$

Table 2 Notation summary.

4. Strategic Formation of Supply Chain Networks Without Congestion

In this section, we consider a supply chain model, similar to Bimpikis et al. (2019), augmented with strategic link formation. Our equilibrium networks will differ from the networks exogenously imposed in Bimpikis et al. (2019).

At a high-level, the two-tier version of our model is as follows. In a supply chain, consumers and raw material producers are connected via two tiers—retailers (linked to consumers) and suppliers (linked to the raw material producers). Only retailers are strategic in that they decide which suppliers to source product from. In each tier of the supply chain, prices are determined via market clearing. When planning, retailers take into account production failures that may occur at any agent present in the system. Retailers pay a constant cost for each link they create. Each supplier is capable of delivering any amount of product—conditional upon production success in the chain—regardless of the collective demand retailers exert upon it.

4.1. Model Without Congestion

Let us consider a two-tier ($T=2$) supply chain model—illustrated in Fig. 1—in which tier $\mathbb{T}_1$ consists of n retailers, all linked to consumers, tier $\mathbb{T}_2$ consists of m suppliers, all linked to raw material producers. It is up to the retailers in $\mathbb{T}_1$ to decide which suppliers in $\mathbb{T}_2$ to link to.

4.1.1. Demands Demand in the supply chain originates in the consumer tier, and, propagates up the chain. Consumers exert a fixed demand of D units of product per retailer.[6] Consumer demand across retailers is equal consistent with the absence of differentiation.

[6] This is similar to the model of Ang et al. (2016), where the manufacturer is primed with a fixed consumer demand.

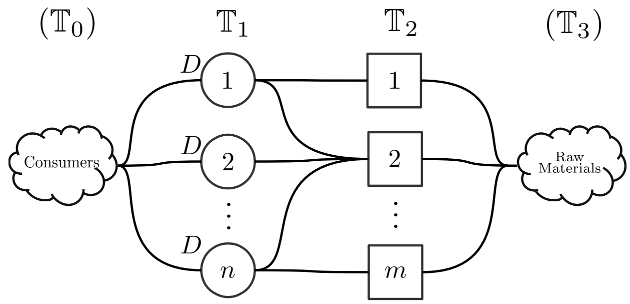


Figure 1 Two-tier supply chain with n retailers, m suppliers, and two implicitly present tiers of consumers and raw material producers. Links are directed from downstream to upstream agents, and appear only between adjacent tiers.

Each agent $i \in \mathbb{T}_t$ allocates its demand $D_{t,i}$ equally among its out-neighborhood. Thus, each upstream agent in i's out-neighborhood receives an order for $D_{t,i}/d^+_{t,i}$ units (which is true even for consumers, as long as tier $\mathbb{T}_0$ is represented by a single meta-consumer, with demand $\Delta = nD$), so the demand of a supplier j in tier $\mathbb{T}_2$ is

$$D_{2,j} = \sum_{i \in \mathcal{N}^-_{2,j}} \frac{D_{1,i}}{d^+_{1,i}} = D \sum_{i \in \mathcal{N}^-_{2,j}} \frac{1}{d^+_{1,i}}, \tag{6}$$

where $\mathcal{N}^-_{2,j}$ is the in-neighborhood of supplier $j \in \mathbb{T}_2$, and $d^+_{1,i}$ is the out-degree of retailer $i \in \mathbb{T}_1$; or, more generally,

$$D_{t,j} = \sum_{i \in \mathcal{N}^-_{t,j}} \frac{D_{t-1,i}}{d^+_{t-1,i}}. \tag{7}$$

Agent $i \in \mathbb{T}_t$, experiencing demand $D_{t,i}$, orders that exact amount of product upstream. An agent could order more than this to mitigate the uncertainty with upstream supply. However, as discussed at the end of Sec. 4.2.3, allowing agents to strategically set order quantities will not affect the equilibria networks arising in the model without congestion.

The demand formation process just described is equivalent to Bimpikis et al. (2019) who “prime” the supply chain with a fixed price for raw materials. Under the market clearing assumption this translated into demand for raw materials, and the latter propagates through the supply chain, eventually, determining consumer demand. In our model, demand propagates in the opposite direction—from the consumers towards the raw material producers.

4.1.2. Production, Failures, and Supplies Having received up to $D_{t,i}$ units of product from upstream, agent i passes all it receives from the upstream tier to the downstream tier with *production success probability* $\lambda \in (0,1)$, or fails to do so with the complementary probability $(1-\lambda)$.

To analyze production failures, we introduce random variables $\omega_{t,i} \sim \text{Bernoulli}(\lambda)$—*production success indicators*—indicating whether production at agent $i \in \mathbb{T}_t$ succeeds. Production failures at different suppliers are independent, so $\omega_{t,i}$ are i.i.d. Using these random variables, we define realized demand $R_{t,i}$ of agent $i \in \mathbb{T}_t$—the amount of product delivered to this agent by upstream suppliers in response to its demand $D_{t,i}$ and along its out-links $\mathcal{N}^+_{t,i}$—where each supplier allocates its available product (whose quantity may be lower than the one requested from the supplier due to disruptions) over its downstream agents proportionally[7] to the latters' demands:

$$R_{t,i} = \sum_{j \in \mathcal{N}^+_{t,i}} \underbrace{\left(\omega_{t+1,j} \frac{R_{t+1,j}}{D_{t+1,j}} \right)}_{\substack{\text{produced share of} \\ \text{supply of } j \in \mathbb{T}_{t+1}}} \underbrace{\frac{D_{t,i}}{d^+_{t,i}}}_{\substack{\text{amount } i \\ \text{requested} \\ \text{from } j}}, \tag{8}$$

$$R_{2,j} = D_{2,j}, \tag{9}$$

where the second expression implies never-failing raw material producers in the two-tier model. Substituting (9) into (8), we get the realized amount that a retailer receives is

$$R_{1,i} = D \frac{\sum_{j \in \mathcal{N}^+_{1,i}} \omega_{2,j}}{d^+_{1,i}} = D \frac{e^+_{1,i}}{d^+_{1,i}}, \tag{10}$$

where

$$e^+_{1,i} = \sum_{j \in \mathcal{N}^+_{1,i}} \omega_{2,j} \tag{11}$$

is the *effective out-degree* of i, that is, the number of i's out-neighbors whose production succeeded. We define supply $S_{t,i}$ of agent $i \in \mathbb{T}_t$ to its downstream customers as

$$S_{t,i} = \omega_{t,i} R_{t,i}. \tag{12}$$

Market clearing at every pair of adjacent tiers of the supply chain translates into the following equality

$$S_{t+1} = \Big(\sum_{j \in \mathbb{T}_{t+1}} S_{t+1,j} \Big) = \sum_{i \in \mathbb{T}_t} R_{t,i}, \tag{13}$$

that is, we assume that the entire amount of product S_{t+1} supplied by the upstream suppliers $j \in \mathbb{T}_{t+1}$ is consumed by downstream suppliers $i \in \mathbb{T}_t$, as expressed via $\sum_i R_{t,i}$.

4.1.3. Payoffs and Prices The payoff $\pi_{t,i}$ of agent $i \in \mathbb{T}_t$ is as follows:

$$\pi_{t,i} = \underbrace{S_{t,i} \cdot p_t}_{\substack{\text{selling} \\ \text{downstream}}} \underbrace{- R_{t,i} \cdot p_{t+1}}_{\substack{\text{buying} \\ \text{upstream}}} \underbrace{- c \cdot d^+_{t,i}}_{\substack{\text{linking} \\ \text{cost}}}, \tag{14}$$

$$\pi_{T,i} = S_{T,i} \cdot p_T - R_{T,i} \cdot p_{T+1}, \tag{15}$$

where $c \geq 0$ is a fixed linking cost, and

$$p_t = \Delta - S_t \tag{2}$$

is the market price in tier $\mathbb{T}_t$ being a function of the total output S_t of that tier (p_{T+1} is the market price of raw materials). This cost can be interpreted as the expense a retailer

[7] Proportional allocation is a widely used scheme; for justification, see, for example Rong et al. (2017)

incurs to establish a relationship with a new supplier. In a two-tier model, suppliers are not strategic[8] and, hence, do not pay for their links to the raw material producers. From (14), it is clear that an agent's payoff depends upon how agents are interlinked, how much product agent i requests from upstream suppliers, as well as the random production failures.

The price formation mechanism in (2) implies that every active supplier in tier $\mathbb{T}_t$ contributes to the tier's output S_t and, hence, to the market price p_t. This can be justified by assuming that negotiations happen *ex ante*, without any pre-existing relationships between buyers and sellers. A link corresponds to a contract, according to which a buyer promises to buy up to a given quantity of product from the supplier at the market price which is determined after production failures are realized and the total product quantity in the upstream market is established. This can be implemented, for example, through a price-matching clause in the contract, based on which suppliers would be discouraged to deviate from the market price, thereby—and due to the complete information assumption—establishing a single market price for the upstream market.

Each retailer in tier $\mathbb{T}_1$ experiences a fixed consumer demand of D units. However, the actual amount supplied by each retailer will, because of production failures, be less than D. If S_1 is the total realized supply at the retailer level, the price paid per unit by consumers will be given by $p_1 = \Delta - S_1$. When no production failures have occurred and, consequently, $S_1 = \Delta$, then, the retailers sell at the (zero) marginal cost.

This price formation mechanism is identical to Bimpikis et al. (2019), with one cosmetic difference. In Bimpikis et al. (2019), the supply chain is "bootstrapped" with a fixed price for raw materials, which, then, translates into demand for raw materials and propagates through the supply chain under market clearance, until it reaches consumers; while in our model, the supply chain is bootstrapped with consumer demand, which propagates up the chain, eventually, determining the market price for raw materials. Mathematically, both pricing mechanism are identical. Bimpikis et al. (2019) interpret this pricing model—where prices are not fixed in advance in a contract—as a spot market, which is a common mechanism occurring "in a number of real-world supply chains, such as those for semiconductors and microelectronics" (see also Mendelson and Tunca (2007)).

Finally, we note that the payoff (14) does not include an explicit penalty for product under-delivery (in addition to the lost opportunity itself). While such penalties can be included in a contract, they are complex to implement and are rarely used (Ang et al. 2016, Hwang et al. 2015).

4.1.4. Network Formation Game Without Congestion The major qualitative difference between our model without congestion from the model considered in Bimpikis et al. (2019) is that the agents in our model are allowed to choose their links. We model the agents' link formation behavior as a one-shot network formation game. We describe the game for the two-tier model but it generalizes to the multi-tier case.

Definition 1 (Strategic Network Formation Game Without Congestion). In a two-tier supply chain, every retailer is considered a player, with payoff (14), and whose pure strategy is its out-neighborhood $\mathcal{N}^+_{1,i}$, that is, which upstream suppliers in $\mathbb{T}_2$ to link to. The retailers (or, all the strategic agents in tiers $1, \ldots, T-1$ in the multi-tier case) simultaneously decide upon their pure strategies, rationally maximizing their expected payoffs.

[8] In Sec. 5.4 we allow suppliers to invest in their reliability.

We will be interested in pure strategy Nash equilibria of this game, defined with respect to arbitrary unilateral deviations in a standard fashion as follows.

DEFINITION 2 (PURE STRATEGY NASH EQUILIBRIUM). $g^* = (\mathcal{N}_{1,i}^{+\,*})_{i\in\mathbb{T}_1}$ is a *pure strategy Nash equilibrium* of the network formation game without congestion if for any retailer $i\in\mathbb{T}_1$ and for any $\mathcal{N}_{1,i}^+$, it holds that

$$\mathbb{E}[\pi_{1,i}(\mathcal{N}_{1,i}^+, (\mathcal{N}_{1,i'}^{+\,*})_{i'\in\mathbb{T}_1\setminus\{i\}})] \le \mathbb{E}[\pi_{1,i}(g^*)].$$

Throughout this work, whenever we refer to an equilibrium, we mean pure strategy Nash equilibrium.

4.2. Analysis

In this section we characterize the pure strategy Nash equilibria of the network formation game without congestion.

4.2.1. Expected Payoffs The first step in the analysis is to obtain the expected payoff of a retailer, based on equation (15).

PROPOSITION 1 **(Retailer's Expected Payoff)**. *For an active retailer $i\in\mathbb{T}_1$ (with $d_{1,i}^+ > 0$),*

$$\mathbb{E}[\pi_{1,i}] = \lambda(1-\lambda)D(\lambda D((1+\lambda)n_1^a - \lambda) - \Delta) + \lambda(1-\lambda)^2 D^2 \frac{1+(1+\lambda)\rho_{1,i}^+}{d_{1,i}^+} - cd_{1,i}^+, \qquad (16)$$

where λ is the production success likelihood, D is the consumer demand per retailer, $n_1^a \le n$ is the number of active retailers in tier $\mathbb{T}_1$, $\Delta = nD$ is the total consumer demand, and

$$\rho_{1,i}^+ = \sum_{i'\in\mathbb{T}_1\setminus\{i\}} d_{1,i'\cap i}^+ \;/\; d_{1,i'}^+ = \sum_{i'\in\mathbb{T}_1\setminus\{i\}} |\mathcal{N}_{1,i}^+ \cap \mathcal{N}_{1,i'}^+| \;/\; d_{1,i'}^+$$

measures the extent of overlap between the out-neighborhood of retailer i and those of i's peers $i'\in\mathbb{T}_1$, $i'\neq i$. For a retailer i having no out-links, $\mathbb{E}[\pi_{1,i}] = 0$.

Proof of Proposition 1: From equation (14), we have

$$\pi_{1,i} = S_{1,i}p_1 - R_{1,i}p_2 - cd_{1,i}^+ = (\text{from (12) and (2)}) = \omega_{1,i}R_{1,i}(\Delta - S_1) - R_{1,i}(\Delta - S_2) - cd_{1,i}^+$$

$$= (\text{from (1) and (13)}) = \omega_{1,i}R_{1,i}(\Delta - \sum_{i'\in\mathbb{T}_1^a}\omega_{1,i'}R_{1,i'}) - R_{1,i}(\Delta - \sum_{i'\in\mathbb{T}_1^a} R_{1,i'}) - cd_{1,i}^+$$

$$= R_{1,i}\Big(\sum_{i'\in\mathbb{T}_1^a}(1-\omega_{1,i}\omega_{1,i'})R_{1,i'} - (1-\omega_{1,i})\Delta\Big) - cd_{1,i}^+,$$

where $\mathbb{T}_1^a \subseteq \mathbb{T}_1$ is a subset of active retailers that have at least one out-link each. To compute expectation of the obtained expression for $\pi_{1,i}$, let us first compute expectations of its components.

$$\mathbb{E}[R_{1,i}] = (\text{from }\ (10)) = \mathbb{E}\Big[D\frac{e_{1,i}^+}{d_{1,i}^+}\Big] = \frac{D}{d_{1,i}^+}\mathbb{E}[e_{1,i}^+] = (\text{from }\ (11)) = \frac{D}{d_{1,i}^+}\mathbb{E}\Big[\sum_{j\in\mathcal{N}_{1,i}^+}\omega_{2,j}\Big]$$

$$= \frac{D}{d_{1,i}^+}\sum_{j\in\mathcal{N}_{1,i}^+}\mathbb{E}[\omega_{2,j}] = \frac{D}{d_{1,i}^+}\sum_{j\in\mathcal{N}_{1,i}^+}\lambda = \frac{D}{d_{1,i}^+}\lambda d_{1,i}^+ = \lambda D.$$

$$\mathbb{E}[R_{1,i}R_{1,i'}] = (\text{from } (10)) = \mathbb{E}\left[\left(D\frac{e^+_{1,i}}{d^+_{1,i}}\right)\left(D\frac{e^+_{1,i'}}{d^+_{1,i'}}\right)\right] = \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\,\mathbb{E}[e^+_{1,i}e^+_{1,i'}]$$

$$= \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\,\mathbb{E}\left[\Big(\sum_{j\in\mathcal{N}^+_{1,i}}\omega_{2,j}\Big)\Big(\sum_{j'\in\mathcal{N}^+_{1,i'}}\omega_{2,j'}\Big)\right]$$

$$= (\text{as } \omega_{t,i} \text{ are i.i.d., and } \mathbb{E}[X^2] = \mathbb{E}^2[X] + \mathrm{Var}[X])$$

$$= \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\Big(\sum_{j\in\mathcal{N}^+_{1,i}}\sum_{j'\in\mathcal{N}^+_{1,i'}}\mathbb{E}[\omega_{2,j}]\,\mathbb{E}[\omega_{2,j'}] + \sum_{j\in\mathcal{N}^+_{1,i}\cap\mathcal{N}^+_{1,i'}}\mathrm{Var}[\omega_{2,j}]\Big)$$

$$= \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\Big(\sum_{j\in\mathcal{N}^+_{1,i}}\sum_{j'\in\mathcal{N}^+_{1,i'}}\lambda^2 + \sum_{j\in\mathcal{N}^+_{1,i}\cap\mathcal{N}^+_{1,i'}}\lambda(1-\lambda)\Big) = \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\Big(\lambda^2 d^+_{1,i}d^+_{1,i'} + \lambda(1-\lambda)d^+_{1,i\cap i'}\Big)$$

$$= \lambda D^2\Big(\lambda + \frac{1-\lambda}{d^+_{1,i}}\cdot\frac{d^+_{1,i\cap i'}}{d^+_{1,i'}}\Big),$$

where $d^+_{1,i\cap i'} = |\mathcal{N}^+_{1,i}\cap\mathcal{N}^+_{1,i'}|$. In particular, when $i' = i$,

$$\mathbb{E}[R^2_{1,i}] = \lambda D^2\Big(\lambda + \frac{1-\lambda}{d^+_{1,i}}\cdot\frac{d^+_{1,i\cap i}}{d^+_{1,i}}\Big) = \lambda D^2\Big(\lambda + \frac{1-\lambda}{d^+_{1,i}}\cdot\frac{d^+_{1,i}}{d^+_{1,i}}\Big) = \lambda D^2\Big(\lambda + \frac{1-\lambda}{d^+_{1,i}}\Big).$$

Having computed expectations of expressions involving realized demands, we can now return to the computation of expectation of retailer payoff.

$$\mathbb{E}[\pi_{1,i}] = \mathbb{E}\left[R_{1,i}\Big(\sum_{i'\in\mathbb{T}^a_1}(1-\omega_{1,i}\omega_{1,i'})R_{1,i'} - (1-\omega_{1,i})\Delta\Big) - cd^+_{1,i}\right] = (\text{as } \omega_{t,i} \text{ and } R_{t+1,j} \text{ are indep.})$$

$$= (1-\mathbb{E}[\omega^2_{1,i}])\,\mathbb{E}[R^2_{1,i}] + \sum_{i'\in\mathbb{T}^a_1, i'\neq i}(1-\mathbb{E}[\omega_{1,i}\omega_{1,i'}])\,\mathbb{E}[R_{1,i}R_{1,i'}] - \Delta(1-\mathbb{E}[\omega_{1,i}])\,\mathbb{E}[R_{1,i}] - cd^+_{1,i}$$

$$= (1-\lambda)\lambda D^2\Big(\lambda + \frac{1-\lambda}{d^+_{1,i}}\Big) + (1-\lambda^2)\sum_{i'\in\mathbb{T}^a_1, i'\neq i}\lambda D^2\Big(\lambda + \frac{1-\lambda}{d^+_{1,i}}\cdot\frac{d^+_{1,i\cap i'}}{d^+_{1,i'}}\Big) - \Delta(1-\lambda)\lambda D - cd^+_{1,i}$$

$$= \lambda^2 D^2(1-\lambda) + \frac{\lambda(1-\lambda)^2 D^2}{d^+_{1,i}} + \lambda^2(1-\lambda^2)D^2(|\mathbb{T}^a_1| - 1)$$

$$+ \frac{\lambda(1-\lambda)(1-\lambda^2)D^2}{d^+_{1,i}}\sum_{i'\in\mathbb{T}^a_1\setminus\{i\}}\frac{d^+_{1,i'\cap i}}{d^+_{1,i'}} - \lambda(1-\lambda)\Delta D - cd^+_{1,i}$$

$$= \lambda^2 D^2(1-\lambda) + \frac{\lambda(1-\lambda)^2 D^2}{d^+_{1,i}} + \lambda^2(1-\lambda^2)D^2(n^a_1 - 1) + \frac{\lambda(1-\lambda)(1-\lambda^2)D^2}{d^+_{1,i}}\rho^+_{1,i}$$

$$- \lambda(1-\lambda)\Delta D - cd^+_{1,i}$$

$$= \lambda(1-\lambda)D(\lambda D((1+\lambda)n^a_1 - \lambda) - \Delta) + \lambda(1-\lambda^2)D^2\frac{1+(1+\lambda)\rho^+_{1,i}}{d^+_{1,i}} - cd^+_{1,i}.$$

∎

4.2.2. Bounding Costs To prevent trivial equilibrium outcomes such as an empty network, we need to ensure that costs are not excessive.

ASSUMPTION 1 **(Bounding Costs for Network Formation Without Congestion)**. *If the number of suppliers is at least as large as the number of retailers, that is, $m \geq n$, then, the network with parallel links—in which every retailer maintains a single link, pointing to an exclusive supplier yields each retailer positive expected payoff.*

Assumption 1 states that the model's parameters are such that the network with parallel links—illustrated in Fig. 2—is at least as good as the empty network. This network is the simplest and least cost—from the point of view of link maintenance cost—network in which retailers can turn a profit.

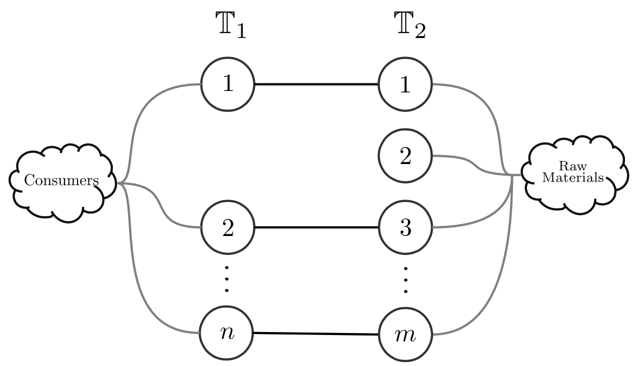


Figure 2 **A network with parallel links. Model's parameters are assumed to be bounded, so that in such a network, if $m \geq n$, every retailer would earn positive expected payoff.**

PROPOSITION 2 **(Bounding Costs in the Model Without Congestion)**. *For the model without congestion, Assumption 1 holds if and only if*

$$c < \lambda(1-\lambda)(n-1)(\lambda^2+\lambda-1)D^2, \tag{17}$$

$$\lambda > \frac{\sqrt{5}-1}{2} \approx 0.618. \tag{18}$$

Proof of Proposition 2: From Proposition 1, we know that the expected retailer payoff in a network with parallel links and enough suppliers ($m \geq n$)—shown in Fig. 2—is as follows.

$$\begin{aligned}
\mathbb{E}[\pi_{1,i}] &= \lambda(1-\lambda)D(\lambda D((1+\lambda)n_1^a - \lambda) - \Delta) + \lambda(1-\lambda)^2 D^2 \frac{1+(1+\lambda)\rho_{1,i}^+}{d_{1,i}^+} - c d_{1,i}^+ \\
&= \lambda(1-\lambda)D(\lambda D((1+\lambda)n - \lambda) - \Delta) + \frac{\lambda(1-\lambda)^2 D^2}{1} - c \cdot 1 \\
&= D^2 \cdot \lambda(1-\lambda)((n-1)\lambda^2 + (n-1)\lambda + 1) - D \cdot \lambda(1-\lambda)\Delta - c \\
&= (\Delta = nD) = \lambda(1-\lambda)(n-1)(\lambda^2+\lambda-1)D^2 - c.
\end{aligned}$$

The proposition statement's requirement $\mathbb{E}[\pi_{1,i}] \geq 0$ immediately translates into the upper bound for the linking cost

$$c < \lambda(1-\lambda)(n-1)(\lambda^2+\lambda-1)D^2.$$

For this upper bound to be well-defined, however, it must be non-negative, since $c \geq 0$. It is non-negative as long as $\lambda^2+\lambda-1 \geq 0$, which holds iff $\lambda \in (\frac{\sqrt{5}-1}{2}, 1)$.

■

4.2.3. Nash Equilibria Characterization In our analysis of equilibria, we will first deal with an empty network equilibrium. While such a network itself is not of particular interest to us, its being an equilibrium provides useful insights into the model without congestion and the impact of the presence of multiple active retailers upon the latters' profit-making.

Theorem 1 **(Empty Equilibrium Existence)**. *If the linking cost $c > 0$, then an empty network is a pure strategy Nash equilibrium of the network formation game without congestion.*

Proof of Theorem 1: From (14), we know

$$\pi_{t,i} = \underbrace{S_{t,i} \cdot p_t}_{\substack{\text{selling}\\ \text{downstream}}} \underbrace{- R_{t,i} \cdot p_{t+1}}_{\substack{\text{buying}\\ \text{upstream}}} \underbrace{- c \cdot d^+_{t,i}}_{\substack{\text{linking}\\ \text{cost}}}.$$

When only one retailer $i \in \mathbb{T}_1$ is active, $d^+_{1,i} > 0$, while all its peers have no links,

$$\begin{aligned} p_1 &= (\text{from (2)}) = \Delta - S_1 = (\text{from (1)}) = \Delta - S_{1,i} = (\text{from (12)}) = \Delta - \omega_{1,i} R_{1,i}, \\ p_2 &= (\text{from (2)}) = \Delta - S_2 = (\text{from (13)}) = \Delta - R_{1,i}, \end{aligned}$$

so,

$$\begin{aligned} \pi_{1,i} &= S_{1,i} \cdot p_1 - R_{1,i} \cdot p_2 - c \cdot d^+_{1,i} = \omega_{1,i} R_{1,i}(\Delta - \omega_{1,i} R_{1,i}) - R_{1,i}(\Delta - R_{1,i}) - c d^+_{1,i} \\ &\leq R_{1,i}(\Delta - R_{1,i}) - R_{1,i}(\Delta - R_{1,i}) - c d^+_{1,i} = -c d^+_{1,i} < 0 \end{aligned}$$

as long as $d^+_{1,i} > 0$. Hence, no retailer would prefer to unilaterally deviate from an empty network, making it an equilibrium. ■

In words, Theorem 1 states that a single active retailer cannot create and exploit a gap between upstream and downstream prices if no other active retailers are in the market. It is useful to note that a similar effect is present when there are multiple active retailers and no production failures, as Theorem 2 states.

Theorem 2 **(Empty Equilibrium Uniqueness When Nobody Fails)**. *If production never fails, that is, $\lambda = 1$, and $c > 0$, then, the empty network is the unique equilibrium of the network formation game without congestion.*

Proof of Theorem 2: If $\lambda = 1$, the expected payoff of a retailer having positive out-degree $d^+_{1,i} > 0$ is as follows:

$$\begin{aligned} \pi_{1,i} &= S_{1,i} \cdot p_1 - R_{1,i} \cdot p_2 - c \cdot d^+_{1,i} = (\text{from (2)}) = S_{1,i} \cdot (\Delta - S_1) - R_{1,i} \cdot (\Delta - S_2) - c \cdot d^+_{1,i} \\ &= (\text{from (12)}) = \omega_{1,i} R_{1,i} \cdot (\Delta - S_1) - R_{1,i} \cdot (\Delta - S_2) - c \cdot d^+_{1,i} = (\text{since no failures}) \\ &= R_{1,i} \cdot (\Delta - n^a_1 D) - R_{1,i} \cdot (\Delta - n^a_1 D) - c \cdot d^+_{1,i} = -c d^+_{1,i} < 0. \end{aligned}$$

Thus, a retailer cannot have a positive out-degree at an equilibrium, making the empty network—which is an equilibrium as per Theorem 1—a unique equilibrium. ■

Theorem 2 easily generalizes to a supply chain with an arbitrary number $T \geq 2$ of tiers. Theorem 2 states that production failures are essential for the agents' ability to make positive profit in the model. The latter is the result of price formation through competition under market clearance, as well as due to our stipulation that the agents function as "repeaters", at best reproducing the input quantity, and not actually transforming the product and/or adding any value to it.

From now on we will be interested in non-trivial equilibria. In the following Theorem 3, we characterize non-trivial equilibria of the supply chain network formation game without congestion. The networks from these equilibria are illustrated in Fig. 3.

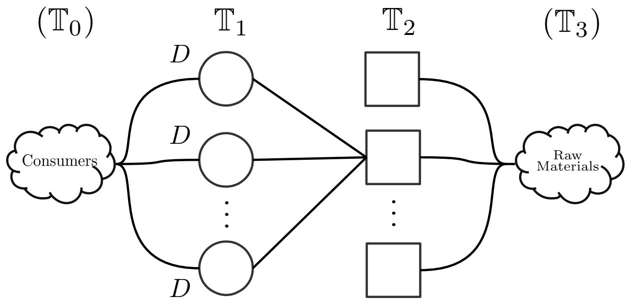


Figure 3 **Pure strategy Nash equilibrium in supply chain network formation game without congestion.**

Theorem 3 **(Non-empty Equilibria Characterization)**. *In the supply chain network formation game without congestion (Definition 1), under Assumption 1, if $c > 0$, then a* cone network*—in which every retailer $i \in \mathbb{T}_1$ maintains a single link, and all the retailers link to the same upstream supplier—is a pure strategy Nash equilibrium. This is a unique non-empty equilibrium of the game, up to supplier labeling.*

Proof of Theorem 3 We first show that under the assumptions about the linking cost, cone networks are Nash equilibria, and, then, show their uniqueness.

1) Cone networks are equilibria: Consider one such network—as shown in Fig. 3—where all the retailers maintain a single link each, linking to the same supplier. From (16), the expected payoff of retailer i is

$$\mathbb{E}[\pi_{1,i}] = \lambda(1-\lambda)D(\lambda D((1+\lambda)n_1^a - \lambda) - \Delta) + \lambda(1-\lambda)^2 D^2 \frac{1+(1+\lambda)\rho_{1,i}^+}{d_{1,i}^+} - cd_{1,i}^+.$$

In a cone network, in the expression above,

$$n_1^a = n, \quad d_{1,i}^+ = 1, \quad \rho_{1,i}^+ = \sum_{i' \in \mathbb{T}_1^a \setminus \{i\}} \frac{d_{1,i' \cap i}^+}{d_{1,i'}^+} = \sum_{i' \in \mathbb{T}_1^a \setminus \{i\}} 1 = n_1^a - 1 = n - 1,$$

so, in such a network,

$$\begin{aligned}
\mathbb{E}[\pi_{1,i}] &= \lambda(1-\lambda)D(\lambda D((1+\lambda)n - \lambda) - nD) + \lambda(1-\lambda)^2 D^2 (1 + (1+\lambda)(n-1)) - c \\
&= \lambda(1-\lambda)((1+\lambda)n - \lambda)D^2 - \lambda(1-\lambda)D^2 n - c \\
&= \lambda^2(1-\lambda)(n-1)D^2 - c.
\end{aligned}$$

This is non-negative as long as

$$c \le \lambda^2(1-\lambda)(n-1)D^2.$$

Simultaneously, we have an upper bound (16) on c from Proposition 2, coming from the assumption about the feasibility of a network with parallel links

$$c < \lambda(1-\lambda)(n-1)(\lambda^2+\lambda-1)D^2$$

that holds in the considered region $\lambda \in (\frac{\sqrt{5}-1}{2}, 1)$. The latter bound on c is tighter for such λ than the former one, so, under Assumption 1, for the considered network we have

$$\mathbb{E}[\pi_{1,i}] > 0.$$

It is also clear that no agent strictly prefers to unilaterally deviate from that network: (i) by dropping a link, a retailer would change its positive expected payoff to zero expected payoff; (ii.a) $\rho^+_{1,i}$ is already at its maximum (i's out-neighborhood completely overlaps with that of each of its peers, of all whom are active) and cannot be improved; and (ii.b) $\mathbb{E}[\pi_{1,i}]$ is strictly decreasing in $d^+_{1,i}$. Consequently, the considered network is a pure strategy Nash equilibrium.

2) Cone networks are the only equilibria: Now, we show that the cone networks—each corresponding to a different supplier $j \in \mathbb{T}_2$ to whom every retailer links—are unique.

From Assumption 1 it is clear that if some—but not all—retailers have no links, the corresponding network is not an equilibrium—by assumption, these retailers would earn positive profit by maintaining a single link having no out-neighborhood overlap with other retailers; in an arbitrary network (rather than a network with parallel links and sufficiently many suppliers), the overlap can only increase a retailer's expected payoff. Thus, we are only concerned with proving that networks where every retailer is active are not equilibria unless it is a cone network.

Now, assume a network, where every retailer is active, $\forall i \in \mathbb{T}_1 : d^+_{1,i} > 0$, yet the retailer degree sequence is non-uniform. We show that for a retailer i such that $d^+_{1,i} > 1$, one can always find a link in $\mathcal{N}^+_{1,i}$ to drop which would be strictly beneficial to i. Assume that retailer i is considering dropping a link to upstream supplier $k \in \mathcal{N}^+_{1,i}$; the payoff and the corresponding neighborhood overlap after the link is dropped are denoted by $\widetilde{\pi}_{1,i}$ and $\widetilde{\rho}^+_{1,i}$, respectively. Then,

$$\begin{aligned}
&\mathbb{E}[\widetilde{\pi}_{1,i} - \pi_{1,i}] \\
&\quad = \left[\lambda(1-\lambda)^2 D^2 \frac{1+(1+\lambda)\widetilde{\rho}^+_{1,i}}{d^+_{1,i}-1} - c(d^+_{1,i}-1)\right] - \left[\lambda(1-\lambda)^2 D^2 \frac{1+(1+\lambda)\rho^+_{1,i}}{d^+_{1,i}} - c d^+_{1,i}\right] \\
&\quad = \frac{\lambda(1-\lambda)^2 D^2}{d^+_{1,i}(d^+_{1,i}-1)} - \lambda(1-\lambda)(1-\lambda^2)D^2 \left(\frac{\widetilde{\rho}^+_{1,i}}{d^+_{1,i}-1} - \frac{\rho^+_{1,i}}{d^+_{1,i}}\right) + c \\
&\quad = \lambda(1-\lambda)^2 D^2 \frac{1+(1+\lambda)(d^+_{1,i}(\widetilde{\rho}^+_{1,i} - \rho^+_{1,i}) + \rho^+_{1,i})}{d^+_{1,i}(d^+_{1,i}-1)} + c.
\end{aligned}$$

Let us take a closer look at one component of the obtained expression:

$$d^+_{1,i}(\widetilde{\rho}^+_{1,i} - \rho^+_{1,i}) + \rho^+_{1,i} = (\text{from Lemma 1}) = d^+_{1,i}\Big(\sum_{j\in\mathcal{N}^+_{1,i}\setminus\{k\}} F^{-i}_{2,j} - \rho^+_{1,i}\Big) + \rho^+_{1,i}$$

$$= d^+_{1,i}((\rho^+_{1,i} - F^{-i}_{2,k}) - \rho^+_{1,i}) + \rho^+_{1,i} = \rho^+_{1,i} - F^{-i}_{2,k} d^+_{1,i} = d^+_{1,i}\left(\frac{\rho^+_{1,i}}{d^+_{1,i}} - F^{-i}_{2,k}\right)$$

$$= (\text{from Lemma 1}) = d^+_{1,i}\left(\frac{\sum_{j\in\mathcal{N}^+_{1,i}} F^{-i}_{2,j}}{d^+_{1,i}} - F^{-i}_{2,k}\right) = d^+_{1,i}(\langle F^{-i}_{2,j}\rangle_{j\in\mathcal{N}^+_{1,i}} - F^{-i}_{2,k}),$$

where $F^{-i}_{2,j} = \sum_{i'\in\mathcal{N}^-_{2,j}\setminus\{i\}} 1/d^+_{1,i'}$. Taking into account that $k \in \mathcal{N}^+_{1,i}$, in the obtained expression we are subtracting one $F^{-i}_{2,k}$ from its arithmetic average per retailer i. It is clear that we can always pick $k = \arg\min_j F^{-i}_{2,j}$ to make the obtained expression non-negative. Thus, $d^+_{1,i}(\widetilde{\rho}^+_{1,i} - \rho^+_{1,i}) + \rho^+_{1,i} \geq 0$, and consequently, $\mathbb{E}[\widetilde{\pi}_{1,i} - \pi_{1,i}] > 0$ for such k. In other words, an arbitrarily picked retailer with out-degree exceeding 1 has a strict incentive to drop a link to its "least useful" supplier. Hence, for any non-cone network there is a sequence of strictly improving unilateral deviations that terminate in a cone network. ■

Theorem 3 generalizes easily to the T-tier case and is summarized in Corollary 1.

Corollary 1 **(Non-Empty Equilibria in T-tier Supply Chain)**. *In a T-tier supply chain network formation game without congestion, the networks illustrated in Fig. 4—in which $\mathbb{T}_1$ retailers concentrate links, and in each subsequent tier $\mathbb{T}_t$, $t \in \{2, \ldots, T\}$, only one active supplier maintains a single link—are the unique non-empty pure strategy Nash equilibria.*

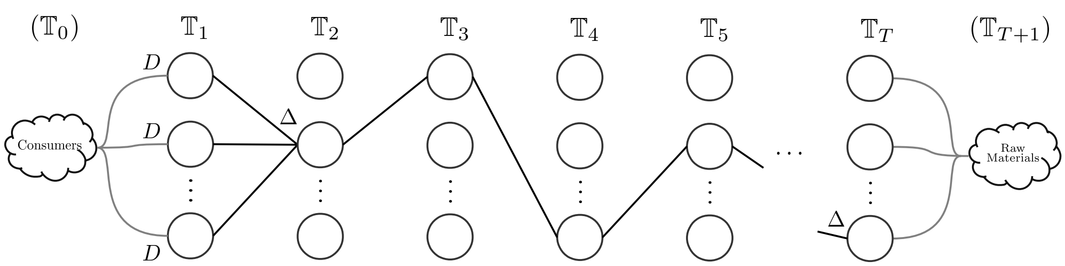


Figure 4 **Pure strategy Nash equilibrium in multi-tier supply chain network formation game without congestion.**

Thus, in the absence of congestion equilibrium supply chains are almost chains, up to the linkage between the first two tiers $\mathbb{T}_1$ and $\mathbb{T}_2$. From this corollary, it is clear why in Bimpikis et al. (2019), among the exogenously given complete k-partite supply chain networks, the inverted pyramid-shaped networks—with the number of agents per tier decreasing along the supply chain—appear as optimal.

Prior work considering production uncertainty, focuses on strategically setting order quantities and prices (Cachon and Lariviere 1999a, Cachon 2003, Cho and Tang 2013) rather than strategically choosing suppliers, as in our case. One may wonder whether Theorem 3 and Corollary 1 hold if we also allow the agents to strategically set order quantities, under the assumption that the quantity that an agent can sell downstream does not exceed its downstream demand[9]. The answer is *yes*: without congestion, retailers' linking and order quantity setting decisions are orthogonal, so retailers would form a cone network (or the chain-like network from Corollary 4) at equilibrium, with the equilibrium order quantities being above downstream demand to hedge against the supplier's under-delivery. Due to our focus on strategic linking, computing these quantities is beyond the scope of this paper.

4.3. Discussion of the Model Without Congestion

The analysis in Sec. 4.2 of the model of Sec. 4.1 produced cone-shaped non-empty equilibria networks, in which the retailers, having one link each, point to the same supplier upstream. The sparsity of such networks as well as the link concentration behavior are surprising. Intuitively, one would expect a resilient / efficient network would have some link redundancy. It is also surprising that too high production reliability, i.e., λ large, can actually hurt retailers; in particular, there is no way retailers can make a profit if production never fails. We discuss these observations below.

4.3.1. Sparsity *Equilibrium networks are sparse because there are no bounds on supplies.* Each upstream supplier can produce as much product as requested (conditional upon production success at that supplier, and, possibly, at other higher-level suppliers in a multi-tier model). The consequence of this is easy to see in a simplified model where we remove price formation and focus only on quantities: assume that a single retailer needs to satisfy a demand of D units, and has an option to source it from $d^+_{1,i}$ upstream suppliers ($d^+_{1,i}/D$ units from each) each of whom successfully delivers the requested quantity with a fixed probability λ. Ignoring the cost of link formation, the expected payoff of the retailer is λD. Hence, it does not matter through how many links to source product, even in the presence of failure. As soon as we introduce a positive linking cost, the retailer prefers to source product via a single link. If, instead, each upstream supplier had a hard cap on its production output strictly lower that a retailer's demand D, retailers would be forced to multi-source. Our model with congestion—described in Sec. 5—incorporates a soft cap via a congestion penalty.

4.3.2. Link Concentration *Retailers favor link concentration as sourcing from a single supplier allows them to buy at a low upstream price (conditional upon that supplier successfully producing).* The positive effect of link concentration by the retailers is best illustrated with a simple example. Consider a supply chain, with two retailers 1 and 2 and two suppliers A and B, in which we are concerned with the expected payoff of retailer 1. For simplicity, suppose $D=1$ and $c=0$ (the introduction of linking costs does not affect the conclusion). Now, let us compare the link concentration scenario (*cone network*) with the scenario when the retailers source from separate suppliers (*network with parallel links*).

- *Cone Network:* If both retailers source from the same supplier, say, A, the upstream price is always low (0 in this example) when upstream production succeeds, and the positive expected payoff of retailer 1 is obtained entirely from the case when both retailer 1 and the supplier A succeed, while retailer 2 fails, thereby, creating a 1 unit gap between upstream and downstream prices, generating a payoff $\pi_{1,1}=1$, with probability $\lambda^2(1-\lambda)$.

[9] This assumption is encoded in the payoff expression (14) implicitly. When an agent does not strategically set order quantities, her output cannot exceed her demand. If we allow order quantities to vary, we need to truncate the quantity an agent sells downstream, changing the payoff expression (14) to $\pi_{t,i}=\min\{D_{t,i},S_{t,i}\}\cdot p_t - R_{t,i}\cdot p_{t+1} - c\cdot d^+_{t,i}$.

- *Network with Parallel Links:* If the retailers source from different suppliers, then, the expected payoff of retailer 1 has two components. One when 1 succeeds, its peer 2 fails, and both upstream suppliers succeed establishing a low upstream price—which happens with probability $\lambda^3(1-\lambda)$, and in which retailer 1 has a payoff $\pi_{1,1}=1$. The other case is when 1 fails, yet, its supplier A succeeds (so 1 does not sell to consumers, yet has to buy from A), and B fails (so 1 buys at a high upstream price of 1), which happens with likelihood $\lambda(1-\lambda)^2$ and corresponds to 1's payoff $\pi_{1,1}=-1$. Hence, the expected payoff of retailer 1 in the network with parallel links is $\lambda^3(1-\lambda)-\lambda(1-\lambda)^2 = \lambda(1-\lambda)(\lambda^2+\lambda-1) < \lambda^2(1-\lambda)$ for all $\lambda \in (0,1)$.

The benefit that retailers enjoy from supply variance is related to Weitzman (1974) who compares the benefit of controlling a system through quantities rather than prices when production costs are uncertain. In the first case, quantities are fixed and prices adjust to clear the market. In the second case, prices are fixed and quantity adjusts to clear the market. Weitzman argued that control through quantities is superior to control via prices, and the advantage of such control for the system scales with the variance of the (component of the) production cost. Thus, higher production cost variance makes quantity control more advantageous for the producer. Our results provide a complementary perspective: If the buyer can choose to source from distinct supply chains that are quantity controlled, it may prefer to source from the chain having higher output uncertainty. Consequently, while, according to Weitzman (1974), higher cost variance encourages control through quantities, our results suggest that competition among quantity controlled supply chains will increase output uncertainty.

4.3.3. Retailers' Welfare vs. Production Failure *Retailers' welfare suffers when λ is close to 1 because a small number of failures among a retailer's peers cannot result in a large enough gap between upstream and downstream prices, to guarantee the retailer a positive expected payoff.* If we resort to the same simple supply chain example from the discussion of link concentration behavior, with two retailers and two suppliers, the above mentioned effect clearly manifests itself in Fig. 5 when we look at the dependency of a retailer's expected payoff upon productivity λ.

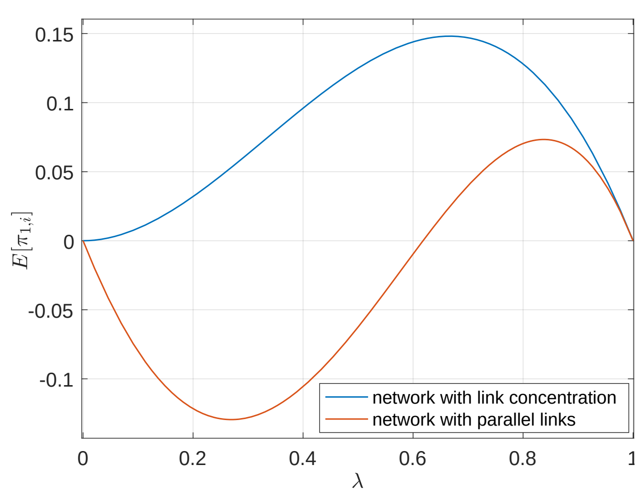


Figure 5 **Dependency of a retailer's expected payoff in a small supply chain in a network where both retailers source from the same supplier and a network where they source from two different suppliers.**

The non-monotonic dependence of a retailer's expected payoff upon λ is valid only when we vary reliability of *every* agent in the system. If we admit a heterogeneous environment, where every retailer and supplier $i \in \mathbb{T}_t$ had its own reliability parameter $\lambda_{t,i}$, then, in the model without congestion, from the point of view of retailer $i \in \mathbb{T}_1$, a perfect situation would be if every retailer including i were linking to the most reliable supplier, i itself would also have maximal λ_i, while its peers' $\lambda_{i'}$ were minimal—this way, retailer i could guarantee itself both a large product quantity to sell, and a large gap between upstream and downstream prices.

5. Strategic Formation of Supply Chain Networks With Congestion

In this section, we extend the previous model by incorporating a congestion penalty, modeling either limited supply or a delay in supply delivery—based upon the total product quantity being produced by a supplier—and show that the congestion effect changes the formed supply chains qualitatively.

5.1. Model

The payoff function (19) for each agent $i \in \mathbb{T}_t$ is an extension of the payoff function (14) that incorporates a congestion penalty term $L_{t,i}$:

$$\pi_{t,i} = \underbrace{S_{t,i} \cdot p_t}_{\substack{\text{selling} \\ \text{downstream}}} \underbrace{- R_{t,i} \cdot p_{t+1}}_{\substack{\text{buying} \\ \text{upstream}}} \underbrace{- c \cdot d^+_{t,i}}_{\substack{\text{linking} \\ \text{cost}}} \underbrace{- L_{t,i}}_{\substack{\text{congestion} \\ \text{penalty}}}, \tag{19}$$

$$\pi_{T,i} = S_{T,i} \cdot p_T - R_{T,i} \cdot p_{T+1},$$

where

$$L_{t,i} = \frac{1}{d^+_{t,i}} \sum_{j \in \mathcal{N}^+_{t,i}} \ell(S_{t+1,j}) = \frac{1}{d^+_{t,i}} \sum_{j \in \mathcal{N}^+_{t,i}} \frac{\gamma}{2}(S_{t+1,j})^2. \tag{20}$$

On can interpret the penalty as a soft constraint on supplies. Another way is to treat it as a delay or latency in product delivery—the larger the amount of product being in production at a supplier, the longer a retailer would wait, on average, for the delivery of goods by that supplier. The congestion function $\ell(x)$ depends upon the amount of product actually produced by a supplier, though demand-dependent ℓ may also be a viable option[10].

We use congestion function $\ell(x) = \frac{\gamma}{2}x^2$ due to its simplicity and strict convexity. The rationale for strict convexity is to model decreasing returns to scale in production. We assume that the production process is already well-established and, at best, scales linearly in the product quantity to be produced, yet, delivering higher product quantities may incur extra delays, potentially resulting from the need for sequential production due to

[10] Dependency of the congestion function $\ell(x)$ upon either the requested or produced amount of product is meaningful, depending on when a requesting party learns about the upstream failure. For example, if a failure occurs due to a natural disaster or a union strike—both of which are publicly observed—the congestion penalty would depend on the amount delivered; if, however, a supplier reaches the deadline having not managed to produced any output and having not timely informed its clients about it, then the congestion penalty's dependency upon (non-realized) demand may be more appropriate. We chose a supply-dependent congestion function to ensure that, if an upstream supplier fails, a retailer sourcing from that supplier does not incur additional penalties associated with the failed product delivery.

the limited output capacity of manufacturing equipment, or transportation of extra product from additional warehouses, or the need to service production equipment due to its amortization.

There is no consensus in literature on the form of the congestion function $\ell(x)$. For example, in Song et al. (2000), the "tardiness penalty" depends linearly on a product of excess wait time and product quantity, making the penalty function super-linear in its input; while in Wang and Tomlin (2009), the lead time-related penalty is a linear function of the order fulfilment excess wait time. Akan et al. (2012) use a convex-concave lead time cost function: it is initially convex, modeling customers' increasing impatience prior to a deadline, and concave past the deadline. One can imagine a purely concave congestion function, modeling a supplier's gradual learning of more efficient methods of production while executing large orders. However, in environments where the congestion function models delays—or, equivalently, lead time uncertainty—convexity is a natural choice (see Johari et al. (2010, Sec. 4) and references therein).

Next, we update the expression for the expected payoff.

Proposition 3 **(Retailer's Expected Payoff in the Model With Congestion)**. *For an active retailer $i \in \mathbb{T}_1^a$,*

$$\begin{aligned}
\mathbb{E}[\pi_{1,i}] &= \mathbb{E}[S_{1,i}\cdot p_1 - R_{1,i}\cdot p_2 - c\cdot d_{1,i}^+ - L_{1,i}] \\
&= \lambda(1-\lambda)D(\lambda D((1+\lambda)n_1^a - \lambda) - \Delta) + \frac{\lambda D^2}{d_{1,i}^+}\left((1-\lambda)^2 - \frac{\gamma}{2d_{1,i}^+}\right) \\
&\quad - cd_{1,i}^+ + \frac{\lambda D^2}{d_{1,i}^+}\sum_{j\in\mathcal{N}_{1,i}^+} F_{2,j}^{-i}\left((1-\lambda)(1-\lambda^2) - \frac{\gamma}{d_{1,i}^+} - \frac{\gamma}{2}F_{2,j}^{-i}\right) \qquad (21) \\
&= \lambda(1-\lambda)D(\lambda D((1+\lambda)n_1^a - \lambda) - \Delta) \\
&\quad + \frac{\lambda D^2}{d_{t,i}^+}\left((1-\lambda)^2 - \frac{\gamma}{2d_{1,i}^+} + \left((1-\lambda)(1-\lambda^2) - \frac{\gamma}{d_{1,i}^+}\right)\rho_{1,i}^+\right) \\
&\quad - cd_{1,i}^+ - \frac{\lambda D^2\gamma}{2d_{1,i}^+}\sum_{j\in\mathcal{N}_{1,i}^+}\left(F_{2,j}^{-i}\right)^2. \qquad (22)
\end{aligned}$$

Here $n_1^a \le n$ is the number of active retailers, having at least one out-link. $\rho_{1,i}^+ = \sum_{i'\in\mathbb{T}_1^a\setminus\{i\}} d_{1,i'\cap i}^+ \;/\; d_{1,i'}^+ = \sum_{i'\in\mathbb{T}_1^a\setminus\{i\}} |\mathcal{N}_{1,i}^+ \cap \mathcal{N}_{1,i'}^+| \;/\; d_{1,i'}^+$ is the aggregate relative overlap of out-neighborhoods of active retailers with the out-neighborhood of i, and $F_{2,j}^{-i} = \sum_{i'\in\mathcal{N}_{2,j}^-\setminus\{i\}} 1/d_{1,i'}^+$ is the congestion at supplier $j\in\mathbb{T}_2$ excluding the contribution of retailer $i\in\mathbb{T}_1$. If retailer i is inactive, then, $d_{1,i}^+ = 0$ and $\mathbb{E}[\pi_{1,i}] = 0$.

Proof of Proposition 3 The proof is easily obtained by substituting the congestion penalty (20), the expression for supply (12), and the expected payoff for the model without congestion given in Proposition 1 into the above expression for $\mathbb{E}[\pi_{1,i}]$. ■

5.2. Analysis of the Model With $n = 2$ and $m = 2$

In this section, we provide the analysis of the supply chain formation model with congestion, described in Sec. 5.1 in the case of two tiers, having two retailers and two suppliers, as illustrated in Fig. 6. More general results for the model with congestion appear in Sec. 5.3.

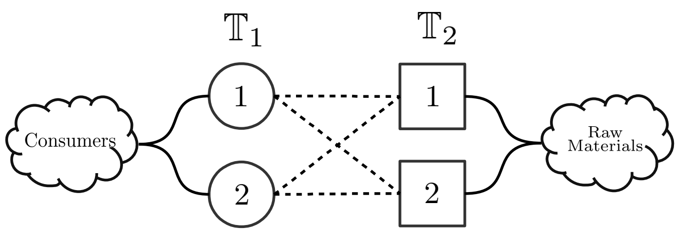
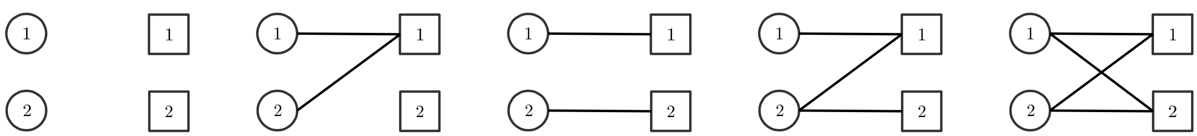


Figure 6 A small two-tier supply chain with two retailers $\mathbb{T}_1$, $|\mathbb{T}_1| = n = 2$, and two suppliers $\mathbb{T}_2$, $|\mathbb{T}_2| = m = 2$. The link that the retailers may create are displayed dashed.

We will be interested in which of the networks shown in Fig. 7 are equilibrium networks. The networks in Fig. 7 exhaust the set of equilibrium network candidates, up to agent



(a) empty (b) cone (c) parallel (d) zee (e) full

Figure 7 Equilibrium network candidates. The empty network is always an equilibrium; other networks may be equilibria (potentially, simultaneously) in different regions of the parameter space.

labeling: (i) an empty network is always an equilibrium for the same reason as in the case of the model without congestion (see Theorem 1)—even in the absence of the congestion penalty, it is strictly preferred to the networks that can be obtained from it via unilateral deviations, so (ii) the latter networks where only one retailer has links cannot be equilibria; (iii) the zee-shaped network is unique up to retailer labeling, and the cone network is unique up to supplier labeling.

5.2.1. Payoffs We use Proposition 3 to compute a retailer's expected payoff in the candidate equilibrium networks.

Proposition 4 **(Retailer's Expected Payoff in 2x2 Candidate Networks).** *Retailers have the following expected payoffs in each of the candidate networks from Fig. 7:*

$$\mathbb{E}[\pi^*_{1,i}\big|] = 0,$$

$$\mathbb{E}[\pi^*_{1,i}\big|] = \lambda(\lambda^3 - 2\lambda^2 + 1 - \tfrac{7}{2}\gamma)D^2 - c,$$

$$\mathbb{E}[\pi^*_{1,i}\big|] = \lambda(-\lambda^3 + 2\lambda - 1 - \tfrac{1}{2}\gamma)D^2 - c,$$

$$\mathbb{E}[\pi^*_{1,1}\big|] = \tfrac{1}{2}\lambda(-\lambda^3 - \lambda^2 + 3\lambda - 1 - \tfrac{9}{4}\gamma)D^2 - c,$$

$$\mathbb{E}[\pi^*_{1,2}\big|] = \tfrac{1}{2}\lambda(-\lambda^3 - 2\lambda^2 + 5\lambda - 2 - \tfrac{5}{4}\gamma)D^2 - 2c,$$

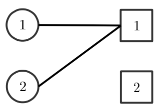
$$\mathbb{E}[\pi^*_{1,i}\big|] = \tfrac{1}{2}\lambda(-\lambda^3 - 2\lambda^2 + 5\lambda - 2 - \gamma)D^2 - 2c.$$

where $i \in \{1,2\}$, $\lambda \in (0,1)$ is the production success likelihood, D is the consumer demand per retailer, and c and γ are the linking and the congestion costs, respectively.

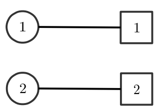
Proof of Proposition 4 The expressions for the payoffs as functions of consumer demand D per retailer are obtained directly by specializing expression (21) of a retailer's expected payoff from Proposition 3 to the case of 2 retailers and 2 suppliers. ■

5.2.2. Bounding Costs First, we determine the model parameters for which each of the candidate networks yields non-negative expected payoffs for each retailer.

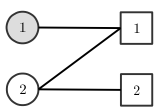
PROPOSITION 5 **(Retailer Network Feasibility)**. *For each of the candidate networks, retailer $i \in \{1,2\} = \mathbb{T}_1$ enjoys non-negative expected payoff only within the following model parameter ranges:*



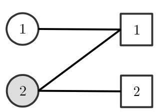
$$\begin{cases} \gamma \ < \ \gamma^{max}_{\text{cone}} = \frac{2}{7}(1-\lambda)(\frac{\sqrt{5}+1}{2}-\lambda)(\frac{\sqrt{5}-1}{2}+\lambda), \\ \lambda \ \in \ (0,1), \\ c \ \le \ c^{max}_{\text{cone}} = \lambda(\lambda^3 - 2\lambda^2 + 1 - \frac{7}{2}\gamma)D^2; \end{cases}$$



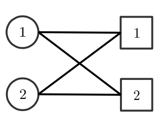
$$\begin{cases} \gamma \ < \ \gamma^{max}_{\text{para}} = 2(1-\lambda)(\lambda + \frac{\sqrt{5}+1}{2})(\lambda - \frac{\sqrt{5}-1}{2}), \\ \lambda \ \in \ (\frac{\sqrt{5}-1}{2}, 1) = (\lambda^{min}_{\text{para}}, 1), \\ c \ \le \ c^{max}_{\text{para}} = \lambda(-\lambda^3 + 2\lambda^2 - 1 - \frac{1}{2}\gamma)D^2; \end{cases}$$



$$\begin{cases} \gamma \ < \ \gamma^{max}_{\text{zee-1}} = \frac{4}{9}(1-\lambda)(\lambda + (\sqrt{2}+1))(\lambda - (\sqrt{2}-1)), \\ \lambda \ \in \ (\sqrt{2}-1, 1) = (\lambda^{min}_{\text{zee-1}}, 1), \\ c \ \le \ c^{max}_{\text{zee-1}} = \frac{1}{2}\lambda(-\lambda^3 - \lambda^2 + 3\lambda - 1 - \frac{9}{4}\gamma)D^2; \end{cases}$$



$$\begin{cases} \gamma \ < \ \gamma^{max}_{\text{zee-2}} = \frac{4}{5}(1-\lambda)(\lambda + \frac{\sqrt{17}+3}{2})(\lambda - \frac{\sqrt{17}-3}{2}), \\ \lambda \ \in \ (\frac{\sqrt{17}-3}{2}, 1) = (\lambda^{min}_{\text{full}}, 1), \\ c \ \le \ c^{max}_{\text{full}} = \frac{1}{4}\lambda(-\lambda^3 - 2\lambda^2 + 5\lambda - 2 - \frac{5}{4}\gamma)D^2; \end{cases}$$



$$\begin{cases} \gamma \ < \ \gamma^{max}_{\text{full}} = (1-\lambda)(\lambda + \frac{\sqrt{17}+3}{2})(\lambda - \frac{\sqrt{17}-3}{2}), \\ \lambda \ \in \ (\frac{\sqrt{17}-3}{2}, 1) = (\lambda^{min}_{\text{zee}}, 1), \\ c \ \le \ c^{max}_{\text{zee-2}} = \frac{1}{4}\lambda(-\lambda^3 - 2\lambda^2 + 5\lambda - 2 - \gamma)D^2; \end{cases}$$

Proof of Proposition 5 We verify the proposition for the cone network the proof for the other networks is similar.

From Proposition 4, we have

$$\mathbb{E}[\pi^*_{1,i}|] = \lambda(\lambda^3 - 2\lambda^2 + 1 - \tfrac{7}{2}\gamma)D^2 - c.$$

In this expression, if the coefficient of D^2 is non-positive, then $\mathbb{E}[\pi_{1,i}|] < 0$ because we assumed that the consumer demand D per retailer is positive. In the latter case, the cone network would not be a best response, as a retailer would prefer to drop its links, increasing its expected payoffs to zero. Hence, for a retailer to get non-negative expected payoff, that coefficient of D^2 must be positive, resulting in

$$\gamma < \tfrac{2}{7}(\lambda^3 - 2\lambda^2 - 2\lambda + 3) = \tfrac{2}{7}(1-\lambda)(\tfrac{\sqrt{5}+1}{2} - \lambda)(\tfrac{\sqrt{5}-1}{2} + \lambda).$$

(For the obtained upper bound to be well-defined, we do not need additional restrictions on $\lambda \in (0,1)$, though, in the proof for the other networks we must lower-bound λ to ensure that the upper bound of γ is non-negative.)

Additionally, for the cone network to be an equilibrium candidate, we require $\mathbb{E}[\pi^*_{1,i}\big|] \geq 0$, resulting in the upper bound for c

$$c \leq \lambda(\lambda^3 - 2\lambda^2 + 1 - \tfrac{7}{2}\gamma)D^2 = c^{max}_{\text{cone}}.$$

■

Similarly to how it was done for the model without congestion, we will assume that the parameters of the model with congestion are such that the network with parallel links is feasible, that is, the retailers are getting positive payoffs in such a network. The cost bounds in the following assumption follow directly from Proposition 5.

Assumption 2 **(Bounding Costs in 2x2 Model With Congestion)**. *Assume that the parameters of the supply chain network formation model with congestion are such, that, if the number of suppliers were at least as large as the number of retailers, then, in the network with parallel links, every retailer would have a positive expected payoff. From Proposition 5, this assumption holds if and only if*

$$\gamma < \gamma^{max}_{\text{para}} = 2(1-\lambda)(\lambda + \tfrac{\sqrt{5}+1}{2})(\lambda - \tfrac{\sqrt{5}-1}{2}) \qquad \textit{and} \qquad \lambda > \tfrac{\sqrt{5}-1}{2} = \lambda^{min}_{\text{para}}.$$

5.2.3. Nash Equilibria When $\gamma > 0$ and $c = 0$ In what follows, we analyze the small supply chain network formation model with congestion assuming a negligible linking cost c. For now, we focus on how different combinations of (λ, γ) affect retailers' behavior.

In the light of Proposition 5 and Assumption 2, the relevant space of parameters is depicted in Fig. 8. Due to Assumption 2, we are interested only in the part of the parameter space under the curve γ^{max}_{para}.

In order to reason about when each of the candidate networks is an equilibrium, let us outline the possible unilateral deviations in Fig. 9. From Theorem 1 we know that no unilateral deviation from an empty network can provide a non-negative expected payoff to a retailer, so the empty network is isolated in Fig. 9, and, hence, is always an equilibrium. Other candidate networks may or may not be equilibria depending on which of them are preferred by the retailers performing the corresponding unilateral deviations. These latter preferences vary across the parameter space, as the following proposition establishes.

Proposition 6 **(Retailers' Preference Over Equilibrium Network Candidates)**. *Let us assume that linking cost c is negligibly small, and define*

$$\widehat{\gamma}_{fz1} = \tfrac{4}{5}(1-\lambda)^2, \quad \widehat{\gamma}_{z2c} = \tfrac{4}{23}(1-\lambda)^2(3\lambda + 4), \quad \widehat{\gamma}_{pc} = \tfrac{2}{3}(1-\lambda)^2(\lambda+1), \quad \widehat{\gamma}_{pz2} = \lambda(1-\lambda)^2.$$

Then, for all $\lambda \in (\tfrac{\sqrt{5}-1}{2}, 1)$, $\widehat{\gamma}_{fz1}(\lambda) < \widehat{\gamma}_{z2c}(\lambda) < \widehat{\gamma}_{pc}(\lambda) < \widehat{\gamma}_{pz2}(\lambda)$, and for the different ranges of γ, retailers' preferences over networks are as shown in Fig. 10,

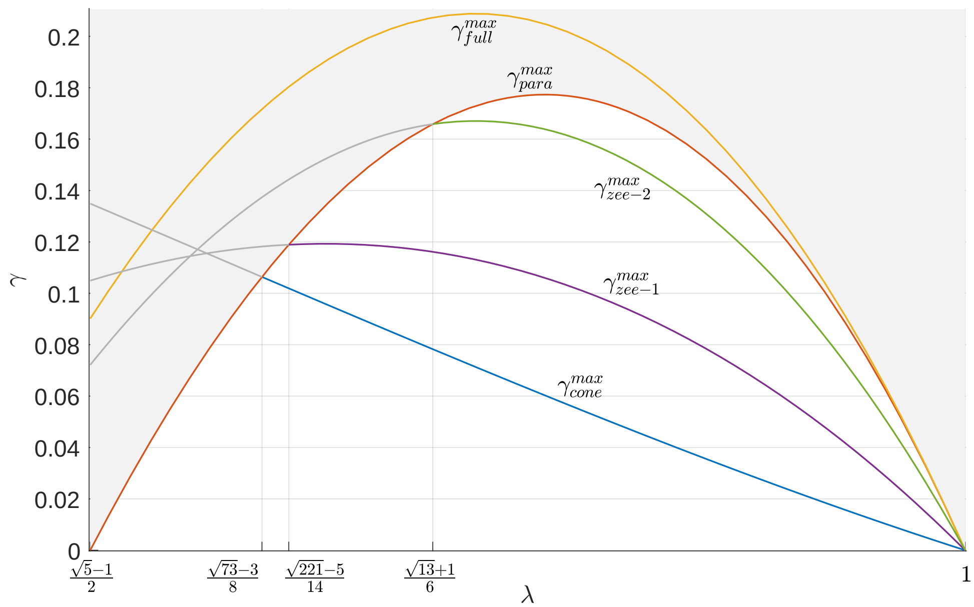


Figure 8 Parameter space for the model with congestion having 2 retailers and 2 suppliers. The feasibility regions of each equilibrium candidate network (for zee-shaped network—from the points of view of both retailers)—in which the corresponding retailers have positive expected payoffs—are enclosed between the horizontal axis, strictly below the curve γ^{max}_{para} and the curve for the corresponding candidate network.

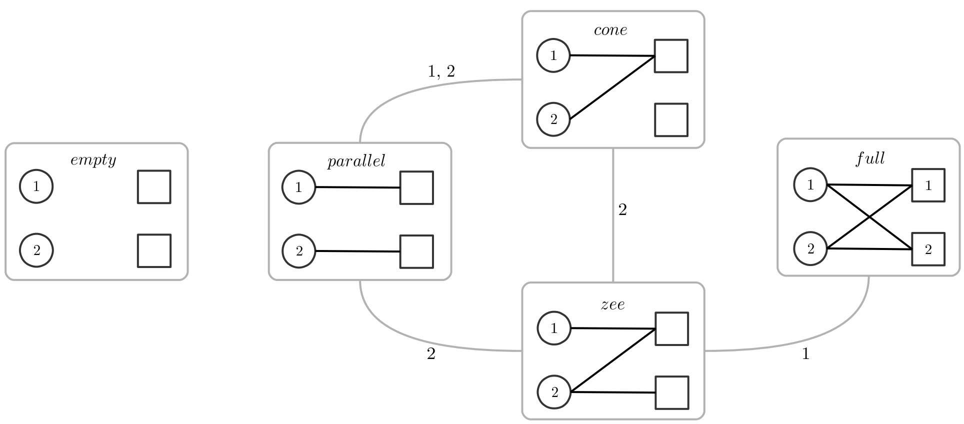


Figure 9 Possible unilateral deviations in a supply chain with 2 retailers and 2 suppliers. Links between networks indicate a possibility of a unilateral deviation by the retailers whose indices label that link.

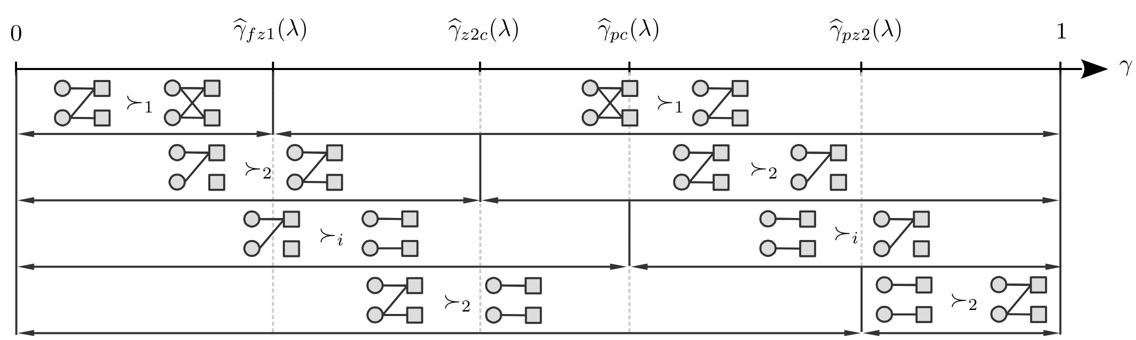


Figure 10 Retailers' preferences over equilibrium network candidates for different congestion costs.

where $A \succ_i B$ indicates that $\mathbb{E}[\pi_{1,i}\big|_A] > \mathbb{E}[\pi_{1,i}\big|_B]$, that is, retailer $i \in \{1,2\}$ strictly prefers network B to network A, and there is a unilateral deviation via which i can switch between A and B. Non-strict preference $\succeq_i$ is defined analogously.

Proof of Proposition 6 Expressions for $\widehat{\gamma}_{fz1}$, $\widehat{\gamma}_{z2c}$, $\widehat{\gamma}_{pc}$, and $\widehat{\gamma}_{pz2}$ are obtained by solving equations $\mathbb{E}[\pi_{1,i}\big|] = \mathbb{E}[\pi_{1,1}\big|]$, $\mathbb{E}[\pi_{1,2}\big|] = \mathbb{E}[\pi_{1,i}\big|]$, $\mathbb{E}[\pi_{1,i}\big|] = \mathbb{E}[\pi_{1,i}\big|]$, and $\mathbb{E}[\pi_{1,i}\big|] = \mathbb{E}[\pi_{1,2}\big|]$, respectively, under the assumption that linking cost c can be dropped. In these equations, the expected payoffs are given in Proposition 5. The rest is straightforward. ■

We, now, can augment the parameter space in Fig. 8 with the obtained thresholds $\widehat{\gamma}$. The result is shown in Fig. 11.

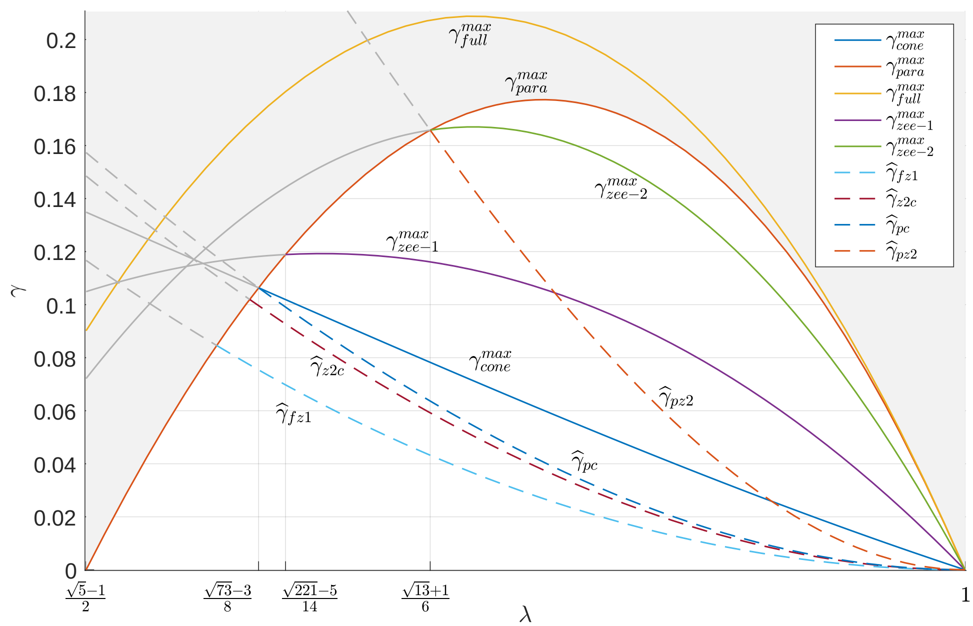


Figure 11 Parameter space for the model with congestion having 2 retailers and 2 suppliers. In addition to the upper bounds on γ necessary for candidate network feasibility, we show thresholds $\widehat{\gamma}$ that affect the retailers' preferences over equilibrium candidate networks.

Having information about both candidate network feasibility and retailers' preference over them in different regions of the parameter space, we can characterize equilibrium networks, as we do in Theorem 4 and Fig. 12.

THEOREM 4 **(Equilibria in 2x2 Model With Congestion and $c = 0$).** *Assume that the linking cost c is negligibly small, and γ^{max}_* and $\widehat{\gamma}_*$ are defined in Propositions 5 and Proposition 11, respectively. Then, in the supply chain network formation game with congestion (Definition 1) with 2 retailers and 2 suppliers, the following holds for the pure strategy Nash equilibria networks:*

1. *An empty network is always an equilibrium.*
2. *If $\frac{\sqrt{5}-1}{2} < \lambda < 1$, and $0 \le \gamma < \gamma^{max}_{para}$, and*
   (a) *$\gamma < \widehat{\gamma}_{fz1}(\lambda)$, then the* cone *network is the unique non-empty equilibrium;*
   (b) *$\widehat{\gamma}_{fz1}(\lambda) \le \gamma \le \widehat{\gamma}_{z2c}(\lambda)$, then the* cone *and* full *networks are the only non-empty equilibria;*
   (c) *$\widehat{\gamma}_{z2c}(\lambda) < \gamma < \widehat{\gamma}_{pz2}(\lambda)$, then the* full *networks is the unique non-empty equilibrium;*
   (d) *$\gamma \ge \widehat{\gamma}_{pz2}(\lambda)$, then the* parallel *and the* full *networks are the only non-empty equilibria.*

3. *If none of the above conditions is met, then the empty network is the only equilibrium.*

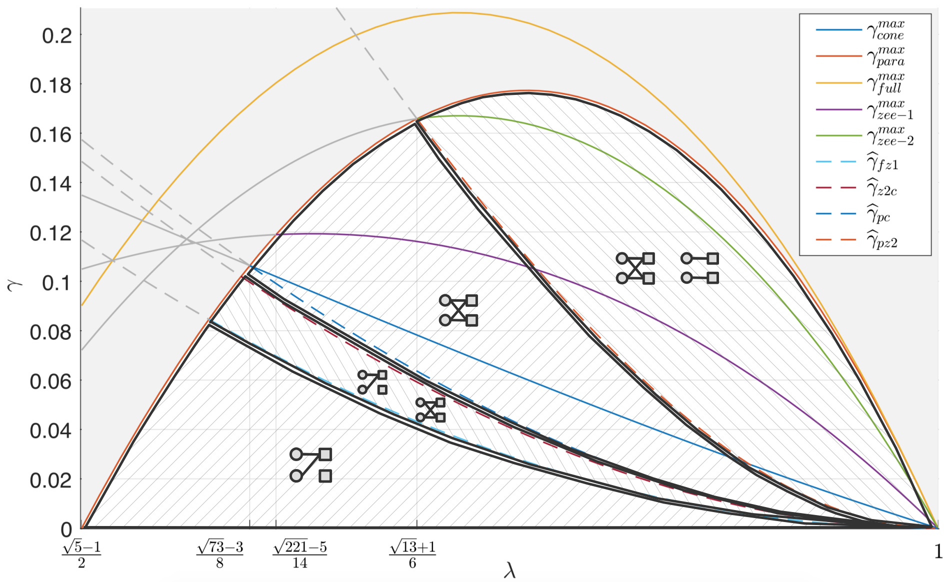


Figure 12 **Equilibria networks in different parts of the parameter space for the model with congestion with 2 retailers and 2 suppliers, and a negligible linking cost $c \approx 0$.**

Proof of Theorem 4 To characterize equilibria networks in different parts of the model parameter space—shown in Fig. 11—we will rely on Proposition 5 that provides us with retailers' expected payoffs in a best response as well as the conditions for when that payoff is non-negative, as well as on Proposition 6 that establishes the retailers' preference over equilibrium network candidates.

According to Assumption 2, and Proposition 5 characterizing the necessary condition for the assumption to hold, we are interested only in the region of the parameter space strictly[11] above the horizontal axis and strictly below curve $\gamma^{max}_{\text{para}}(\lambda)$, which also implies a lower bound on λ:

$$\begin{aligned} 0 < \ & \gamma < 2(1-\lambda)\left(\lambda + \tfrac{\sqrt{5}+1}{2}\right)\left(\lambda - \tfrac{\sqrt{5}-1}{2}\right), \\ \tfrac{\sqrt{5}-1}{2} < \ & \lambda < 1. \end{aligned}$$

In the above defined region of the parameter space, we will focus on 5 parts that curves $\widehat{\gamma}_{fz1}$, $\widehat{\gamma}_{z2c}$, $\widehat{\gamma}_{pc}$, and $\widehat{\gamma}_{pz2}$ slice the region into:

1) $\gamma < \widehat{\gamma}_{fz1}(\lambda)$: According to Proposition 5, all the non-empty candidate networks are feasible here (the retailers are getting a non-negative payoff), and, based on Proposition 6, the retailers' preferences over networks are as follows:

$$\succ_2 \ , \qquad \succ_1 \ , \qquad \succ_{1,2} \ , \qquad \succ_2$$

The same relationships summarized as a diagram in Fig. 13. Thus, the only non-empty

[11] The case $\gamma = 0$ corresponds to the model without congestion, which is the topic of study in Sec. 4.

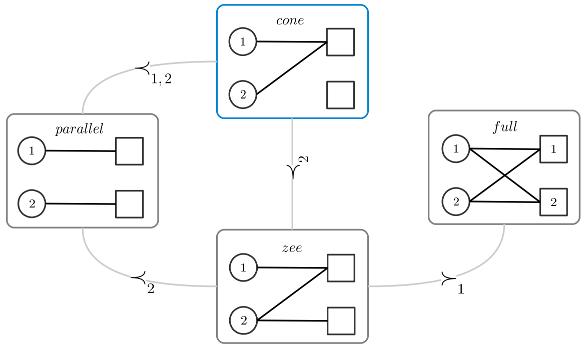


Figure 13 **Retailers' preference over networks when $\gamma < \widehat{\gamma}_{fz1}(\lambda)$. $A \succ_i B$ indicates that in network A, retailer i has a strictly larger expected payoff than in network B. Each network is feasible (retailers have non-negative expected payoffs in each of them). The cone network is an equilibrium.**

network from which no retailer wants to unilaterally deviate is the cone network, making it the unique non-empty equilibrium network up to supplier labeling.

2) $\widehat{\gamma}_{fz1} \le \gamma \le \widehat{\gamma}_{z2c}$*:* Proceeding similarly to the previous case, we end up with the relationships between the candidate networks shown in Fig. 14. While all networks are feasible,

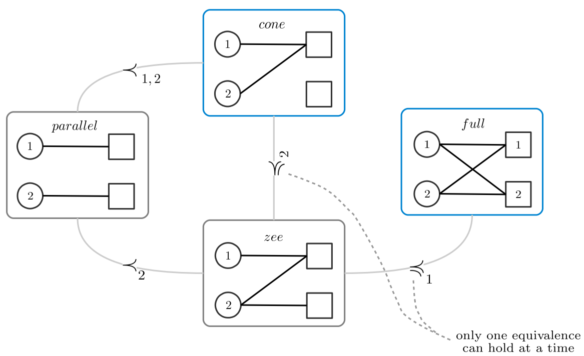


Figure 14 **Retailers' preference over networks when $\widehat{\gamma}_{fz1}(\lambda) \le \gamma \le \widehat{\gamma}_{z2c}(\lambda)$. $A \succ_i B$ indicates that in network A, retailer i has a strictly larger expected payoff than in network B, and $A \succcurlyeq_i B$ allows for expected payoff equality. Among the two $\succcurlyeq_i$ relationships, equality can hold only for one of them at a time. All networks are feasible, and the cone and full networks are equilibria.**

only in the cone and the full networks, retailers prefer not to unilaterally deviate. Notice that, out of two non-strict preference relations— $\succcurlyeq_2$ and $\succcurlyeq_1$ —equivalence can hold in one of them at a time (either when $\gamma = \widehat{\gamma}_{fz1}(\lambda)$ or when $\gamma = \widehat{\gamma}_{z2c}(\lambda)$, which cannot hold simultaneously in $\lambda \in (\frac{\sqrt{5}-1}{2}, 1)$), which is why zee-shaped network is not an equilibrium.

3) $\widehat{\gamma}_{z2c}(\lambda) < \gamma < \widehat{\gamma}_{pz2}(\lambda)$: The retailers' preferences over the candidate networks are shown in Fig. 15. Here, the feasibility of networks varies across the region, and the cone network's relationship with the parallel network also varies. However, this does not affect the full network's being the only non-empty equilibrium in this region of the parameter space.

4) $\widehat{\gamma}_{pz2}(\lambda) \le \gamma < \gamma_{para}^{max}(\lambda)$*:* In the last slice of the parameter space partition, the retailers' preferences over the candidate networks are shown in Fig. 16. Here, the parallel and the full network are the only non-empty equilibria.

■

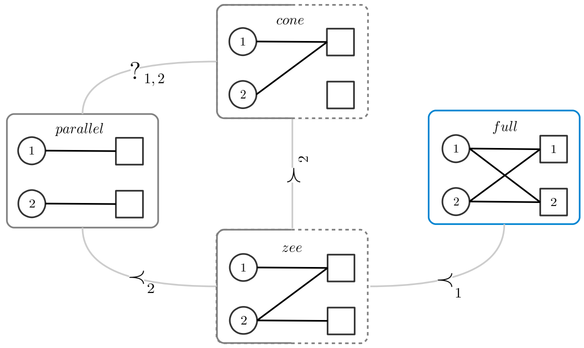


Figure 15 Retailers' preference over networks when $\widehat{\gamma}_{z2c}(\lambda) < \gamma < \widehat{\gamma}_{pz2}(\lambda)$. The retailers' expected payoffs are non-negative in the cone network only in the part of the region, and so is the expected payoff of retailer 1 in the zee-shaped network; also the relationship between the cone and the parallel networks change within the region.

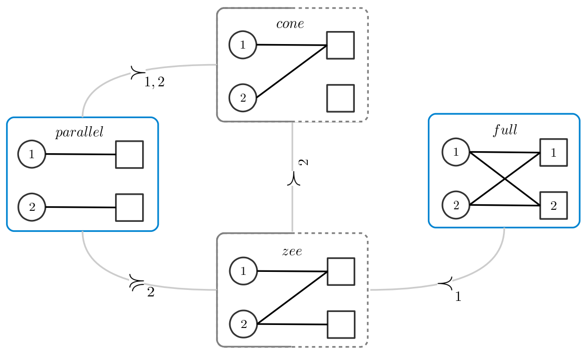
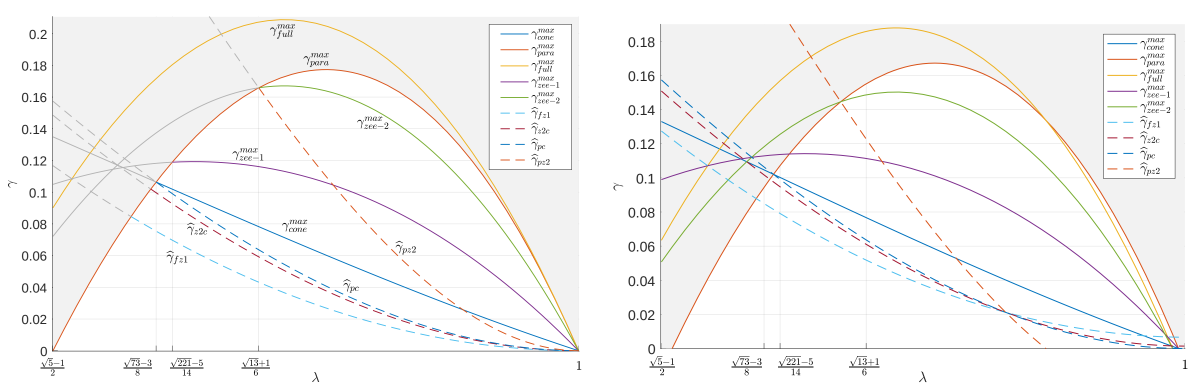


Figure 16 Retailers' preference over networks when $\widehat{\gamma}_{pz2}(\lambda) \leq \gamma < \gamma^{max}_{\textsf{para}}(\lambda)$. Feasibility of zee and cone networks varies across the region. Parallel and full networks are unique non-empty equilibria.

5.2.4. Nash Equilibria When $\gamma > 0$ and $c > 0$ In this section, we characterize equilibrium networks when $c > 0$. The qualitative changes in the parameter space partitioning are depicted in Fig. 17. It is easy to see that the introduction of positive linking cost c results



(a) $c = 0$ (b) $c > 0$

Figure 17 Parameter space of the model with congestion with two retailers and suppliers, with zero and positive linking cost.

in a shift of curves $\widehat{\gamma}_{z2c}(\lambda)$ and $\widehat{\gamma}_{pz2}(\lambda)$. However, there are two critical things to notice:

(i) all such curves intersect at the same point, whose location in the parameter space is described in Fig. 18; and

(ii) the order of the curves $\widehat{\gamma}$ on each side of that intersection point is the same, regardless of the value of c.

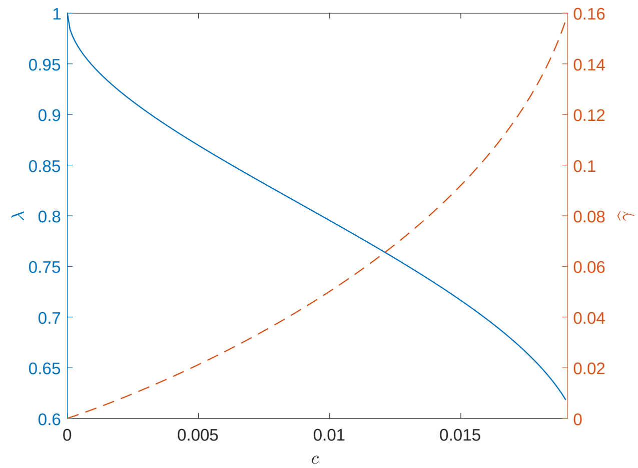


Figure 18 **The location of the intersection point of the curves $\widehat{\gamma}_{fz1}(\lambda, c)$, $\widehat{\gamma}_{z2c}(\lambda, c)$, $\widehat{\gamma}_{pc}(\lambda, c)$, and $\widehat{\gamma}_{pz2}(\lambda, c)$.**

Thus, following the reasoning from the proof of Theorem 4, it is easy to generalize the latter's statement to the case of positive linking cost c. To aid understanding, we provide this generalization here informally: Fig. 19 shows how equilibrium networks change when the linking cost c is positive but not large.

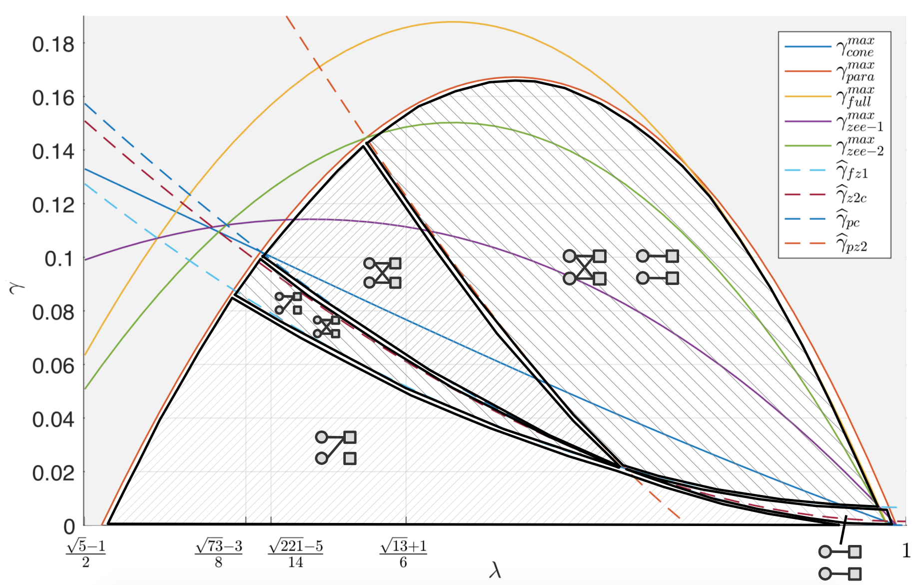


Figure 19 **Equilibrium networks in different parts of the parameter space for the model with congestion having 2 retailers and 2 suppliers, and a small positive linking cost c.**

Besides $\widehat{\gamma}$ curves' shifting with growing c, the network feasibility regions—outlined by curves γ^{max}—shrink, and the feasibility regions of denser networks (only the full network for the case of 2 retailers and suppliers) shrink faster than those of sparser networks. As a result, when c grows further, the equilibria network distribution over the parameter space changes as shown in Fig. 20.

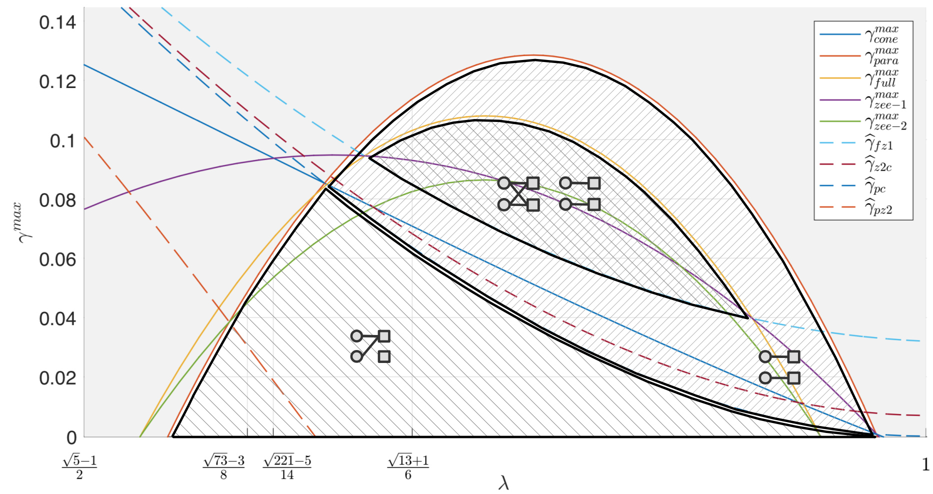


Figure 20 **Equilibria networks in different parts of the parameter space for the model with congestion having 2 retailers and 2 suppliers, and a positive linking cost c of larger magnitude.**

As c grows even larger, first, the island region in which the full network is an equilibrium gradually disappears; then the region where the parallel network is feasible shrinks to a point (it does not disappear because of Assumption 2).

We conclude the discussion of the model with 2 retailers and suppliers with Table 3 showing the values of expected payoff of a retailer at equilibrium for the candidate networks when the linking cost c grows. We see that the positive values of the expected payoffs outline the earlier defined regions within which different networks may be equilibria.

5.3. Analysis of the General Model

In this section, we will provide a limited set of theoretical results for the general model with congestion defined in Sec. 5.1, putting no constraints on the number of retailers and suppliers. These results will be proved under the following assumptions.

ASSUMPTION 3 **(Cost Magnitude in General Model)**. *Let us assume that the linking cost c and the congestion cost γ in the general model with congestion are not too high, so that*

1. *every retailer can be active—having $d_{1,i}^{+} > 0$ and $\mathbb{E}[\pi_{1,i}] > 0$; and*
2. *there is a positive value of consumer demand D per supplier, under which the retailers have non-negative expected payoffs $\mathbb{E}[\pi_{1,i}] \geq 0$.*

The second item in Assumption 3 states that there is a network in which retailers enjoy non-negative profit. The first item in the assumption restricts attention to those networks where every retailer links to some suppliers (there may be equilibria networks in which some retailers have no links).

First, let us define

$$\widehat{d}_{1,i}^{+} = \frac{\lambda(1-\lambda)D}{\sqrt{c}}, \tag{23}$$

$$\widehat{F}_{2,j}^{-i} = \max\left\{0, \frac{(1-\lambda)(1-\lambda^2)}{\gamma} - \frac{1}{d_{1,i}^{+}}\right\}, \tag{24}$$

and assume that $\widehat{d}_{1,i}^{+} \in \{1, \dots, m\}$.

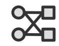
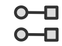
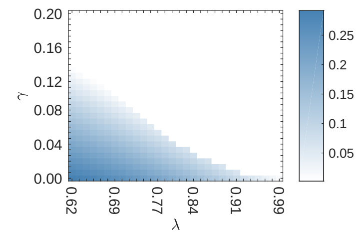
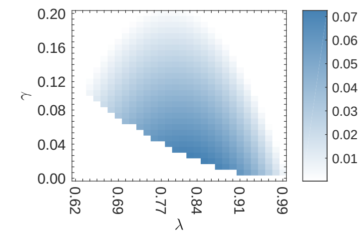
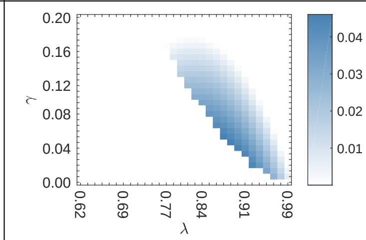
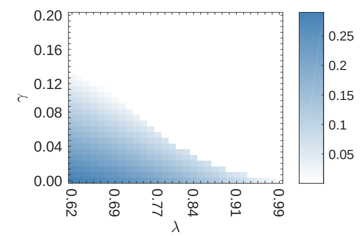
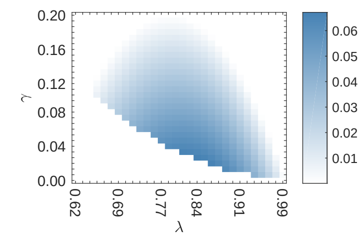
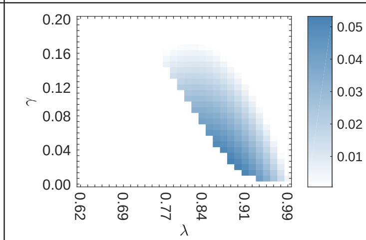
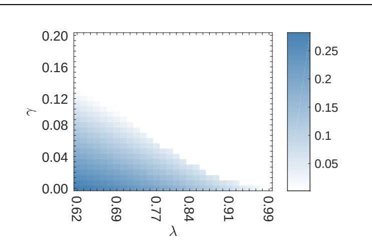
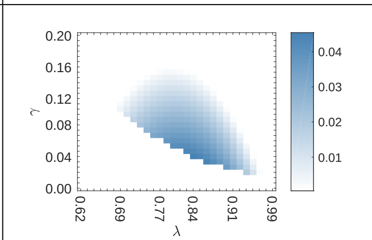
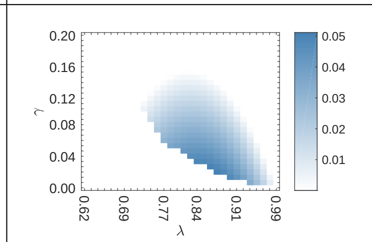
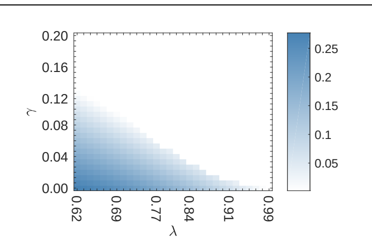
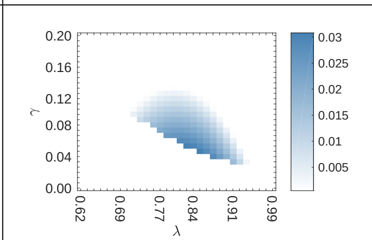
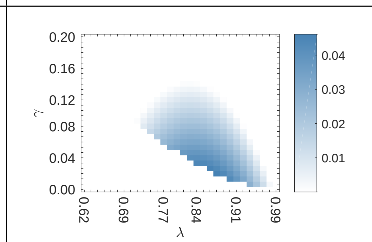
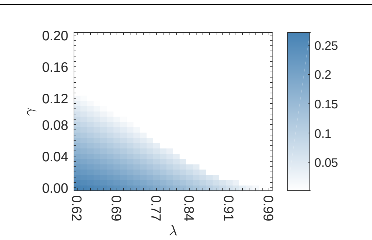
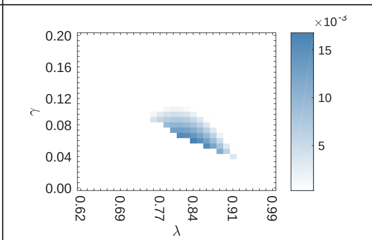
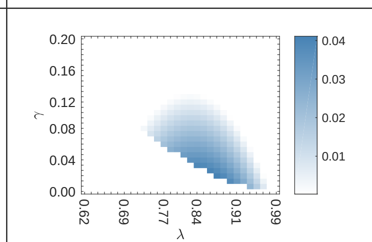
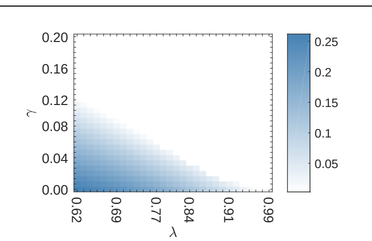
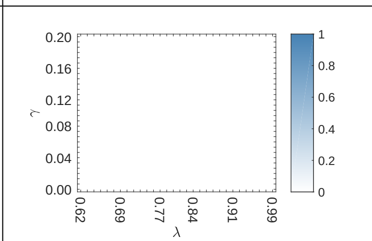
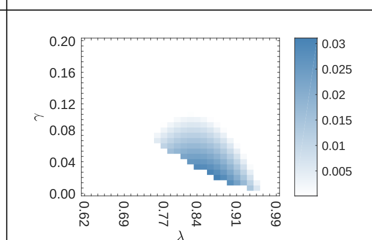
c			
0.000			
0.002			
0.010			
0.015			
0.020			
0.030			

Table 3 **Expected payoff $\mathbb{E}[\pi^*_{1,i}|\bullet]$ of a retailer at equilibrium when linking cost c varies, and the consumer demand per retailer is normalized to $D=1$. The 0 values in these figures indicate that the corresponding network is not an equilibrium for the given combination of model parameter values (λ, c, γ).**

THEOREM 5 **(Equilibria in Low Congestion Cost Regime)**. *Given that Assumption 3 holds, in a* low congestion cost regime*, when*

$$\gamma < \frac{(1-\lambda)(1-\lambda^2)}{n}$$

the behavior of the general model with congestion is qualitatively identical to that of the model without congestion—characterized earlier in Theorem 3—so there is only one unique (up to supplier labeling) non-empty equilibrium network where every retailer maintains a single link, and all the retailers source from the same supplier.

THEOREM 6 **(Symmetric Equilibrium in Higher Congestion Cost Regime)**. *Given that Assumption 3 holds, in a* high congestion cost regime*, where, from the perspective of every retailer i, the congestion at every supplier $F_{2,j}^{-i}$ can reach $\widehat{F}_{2,j}^{-i}$, let us ignore parity and divisibility issues and assume existence of a regular network, in which all retailers have out-degree $d_{1,i}^{+} = \widehat{d}_{1,i}^{+}$, all the congestion at every supplier is $\widehat{F}_{2,j}^{-i}|_{d_{1,i}^{+}=\widehat{d}_{1,i}^{+}}$. Then, this network is a pure strategy Nash equilibrium of the general model with congestion.*

Proof of Theorems 5 and 6 From (21), we have the expression for the expected payoff of a retailer in the general model with congestion

$$\begin{aligned}\mathbb{E}[\pi_{1,i}] =& \lambda(1-\lambda)D(\lambda D((1+\lambda)n_1^a - \lambda) - \Delta) + \frac{\lambda D^2}{d_{1,i}^{+}}\left((1-\lambda)^2 - \frac{\gamma}{2d_{1,i}^{+}}\right) \\ &- cd_{1,i}^{+} + \frac{\lambda D^2}{d_{1,i}^{+}} \sum_{j\in\mathcal{N}_{1,i}^{+}} F_{2,j}^{-i}\left((1-\lambda)(1-\lambda^2) - \frac{\gamma}{d_{1,i}^{+}} - \frac{\gamma}{2}F_{2,j}^{-i}\right).\end{aligned}$$

From Assumption 3, we know that at an equilibrium, $n_1^a = n$.

1) *Low congestion cost regime:* Based on the above expression for $\mathbb{E}[\pi_{1,i}]$, when a best-responding retailer i decides which $d_{1,i}^{+}$ suppliers to link to, it assesses each supplier with respect to the value of

$$f(F_{2,j}^{-i}) = F_{2,j}^{-i}\left((1-\lambda)(1-\lambda^2) - \frac{\gamma}{d_{1,i}^{+}} - \frac{\gamma}{2}F_{2,j}^{-i}\right),$$

which is retailer dependent. $f(F_{2,j}^{-i})$ is a quadratic function, reaching its maximal value on $F_{2,j}^{-i} \in \mathbb{R}_+$ at

$$F_{2,j}^{-i} = \widehat{F}_{2,j}^{-i} = \max\left\{0, \frac{(1-\lambda)(1-\lambda^2)}{\gamma} - \frac{1}{d_{1,i}^{+}}\right\}.$$

In a low congestion regime, when $\gamma < \frac{(1-\lambda)(1-\lambda^2)}{n}$, for any retailer i and supplier j, $\widehat{F}_{2,j}^{-i} > n - \frac{1}{d_{1,i}^{+}} \geq n-1$ is so large, that

$$F_{2,j}^{-i} < \widehat{F}_{2,j}^{-i}.$$

Taking into account that, under the same constraint imposed upon γ, the following term in the expected payoff expression is non-negative

$$(1-\lambda)^2 - \frac{\gamma}{2d_{1,i}^{+}} > (1-\lambda)^2 - \frac{(1-\lambda)(1-\lambda^2)}{2nd_{1,i}^{+}} = (1-\lambda)^2\left[1 - \frac{1+\lambda}{2nd_{1,i}^{+}}\right] \geq 0,$$

we see that congestion in the low congestion cost regime is unambiguously good from the perspective of a retailer, and the latter also has no incentive to create more than one link as the expected payoff is a decreasing function of $d^+_{1,i}$. As a result, the behavior of the model is the same as that of the model without congestion, as described in Theorem 3.

2) *Higher congestion cost regime:* In this regime, and unlike the previous case of low γ, congestion $F^{-i}_{2,j}$ at every supplier can actually reach its optimal value $\widehat{F}^{-i}_{2,j}$. Let us consider the network in which retailers have identical out-degrees $d^+_{1,i} = \tau \geq 1$, and suppliers have identical congestion values $F^{-i}_{2,j} = \widehat{F}^{-i}_{2,j}$. This network's existence is assumed in Theorem 6. In this network, in terms of retailers' link distribution over suppliers, the retailers are best-responding, as all the suppliers have identical value $f(F^{-i}_{2,j})$, and it does not matter which $d^+_{1,i}$ of them to link to. Thus, we are interested in the question of whether any retailer would like to unilaterally change its outdegree $d^+_{1,i}$ improving its expected payoff. Let us pick the best possible out-degree from which a retailer would not have an incentive to deviate.

$$\mathbb{E}[\pi_{1,i}]\big|_{F^{-i}_{2,j}=\widehat{F}^{-i}_{2,j}} \xrightarrow{d^+_{1,i}} \max\,.$$

The considered objective function, as a function of $d^+_{1,i}$, reaches its maximum at

$$d^+_{1,i} = \widehat{d}^+_{1,i} = \frac{\lambda(1-\lambda)D}{\sqrt{c}}.$$

If $\forall i \in \mathbb{T}_1 : d^+_{1,i} = \widehat{d}^+_{1,i}$ and $\forall i \in \mathbb{T}_1, j \in \mathbb{T}_2 : F^{-i}_{2,j} = \widehat{F}^{-i}_{2,j}$, does a retailer i have an incentive to unilaterally deviate? If it decides to reduce its outdegree $d^+_{1,i}$, then no changes happen to i's preferences on which suppliers to link to, as its $\widehat{F}^{-i}_{2,j}$ decreased, and by simply dropping any of existing links it is already doing its best at reducing congestion $F^{-i}_{2,j}$ at its suppliers. Hence, as $d^+_{1,i}$ is already at its optimal value, a link drop cannot improve i's expected payoff. Similarly, if we consider link addition, it does not matter whom to create an extra link to, as all the suppliers have the same value, and a retailer's out-degree is already at its optimal value. Thus, the considered network is an equilibrium.

■

We list several observations obtained from computing equilibria in specific instances of the general model with congestion:

- We conjecture that, under moderate linking and congestion costs, a non-empty equilibrium network always exists.
- There are irregular equilibrium networks (not left- or right-regular).
- In some equilibrium networks, a fraction of retailers have no links.
- While the symmetric equilibrium network of Theorem 6 need not exist in general, we often observe existence of an equilibrium network structurally similar to it.

5.4. Supplier Heterogeneity and Incentives for Reliability Improvement

The general model with congestion assumed a homogeneous set of suppliers, having identical production success likelihoods $\lambda_i = \lambda$ and congestion costs $\gamma_i = \gamma$. Here we are interested in whether suppliers are incentivized to invest in improving their reliability through either reducing their congestion costs γ_i or increasing their production success likelihood λ_i. Furthermore, if there are two options for a supplier—either to invest in production reliability, or to invest congestion reduction—which should they pick?

THEOREM 7 **(Investment in Quality by Heterogeneous Strategic Suppliers)**. *Assume that the linking and congestion costs are moderate, so that non-empty equilibria exist. Then, the following holds in our two-tier models:*

1. *Absent congestion, suppliers are always incentivized to maximize their production success likelihood λ_i.*
2. *With congestion, while a reduction in congestion cost γ_i is unambiguously good for supplier i, increasing its production success likelihood λ_i is* not *always profitable—retailers may prefer to source from suppliers that are* less *likely to succeed in production.*

Proof of Theorem 7 Let us, now, extend the general model with congestion with heterogeneous reliability parameters. Since we are mostly interested in the strategic behavior of suppliers competing for retailers, we will assume that retailers are all equally reliable, having the likelihood of production success $\lambda_{1,i} = \lambda_r$ for all $i \in \mathbb{T}_1$. Reliability of suppliers, however, varies: supplier $j \in \mathbb{T}_2$ has production success likelihood $\lambda_{2,j} = \lambda_j$, and congestion cost $\gamma_{2,j} = \gamma_j$.

First, we must modify the expected payoffs of the general model with congestion to this environment.

1) Payoffs: The general expressions (19) for agent payoff still holds in the heterogeneous model:

$$\pi_{1,i} = S_{1,i}p_1 - R_{1,i}p_2 - cd^+_{1,i} - L_{1,i},$$

$$\pi_{2,j} = S_{2,j}p_2 - R_{2,j}p_3,$$

where, as before, $S_{t,i}$ is the supply of $i \in \mathbb{T}_t$, $R_{t,i}$ is the same agent's realized demand, p_t is the price at which tier t trades product downstream, and

$$L_{1,i} = \frac{1}{d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \frac{\gamma}{2}(S_{2,j})^2$$

is the penalty incurred by agent $i \in \mathbb{T}_t$ due to congestion at the upstream suppliers $\mathcal{N}^+_{t,i}$ it is linked to. Notice that, as suppliers in $\mathbb{T}_2$ do not strategically create links—all of them are assumed to have access to the raw material market—they do not suffer penalties associated with linking or upstream congestion.

Now, we need to derive expected payoffs. This derivation will go along the lines of the proof of Proposition 3, with the difference that, suppliers are heterogeneous.

$$\begin{aligned}\mathbb{E}[\pi_{1,i}] &= \mathbb{E}[S_{1,i}p_1 - R_{1,i}p_2 - cd^+_{1,i} - L_{1,i}] = (\text{from proof of Proposition 3}) \\ &= \mathbb{E}\Big[R_{1,i}\big(\sum_{i' \in \mathbb{T}^a_1} (1 - \omega_{1,i}\omega_{1,i'})R_{1,i'} - (1 - \omega_{1,i}\Delta\big) - cd^+_{1,i} - L_{1,i}\Big].\end{aligned}$$

At first, focus on the terms other than the congestion penalty $L_{1,i}$; we deal with the later in the last part of the proof.

$$\begin{aligned}\mathbb{E}[R_{1,i}] &= (\text{from (10)}) = \mathbb{E}\left[D\frac{e^+_{1,i}}{d^+_{1,i}}\right] = \mathbb{E}\left[D\frac{\sum_{j \in \mathcal{N}^+_{1,i}} \omega_{2,j}}{d^+_{1,i}}\right] \\ &= \frac{D}{d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \mathbb{E}[\omega_{2,j}] = \frac{D}{d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \lambda_j = \frac{D}{d^+_{1,i}}\lambda^+_{1,i} = D\bar{\lambda}^+_{1,i},\end{aligned}$$

where
$$\lambda^+_{1,i} = \sum_{j\in\mathcal{N}^+_{1,i}} \lambda_j \quad \text{and} \quad \bar{\lambda}^+_{1,i} = \frac{\lambda^+_{1,i}}{d^+_{1,i}}.$$

Within the scope of this proof, we will be using some extra notation to handle supplier heterogeneity; in this notation, the plus superscript indicates relation to the out-neighborhood of an agent specified in the subscript, while the bar indicates averaging over that out-neighborhood.

$$\begin{aligned}
\mathbb{E}[R_{1,i}R_{1,i'}] &= (\text{from (10)}) = \mathbb{E}\left[\Big(D\frac{e^+_{1,i}}{d^+_{1,i}}\Big)\Big(D\frac{e^+_{1,i'}}{d^+_{1,i'}}\Big)\right] = \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\,\mathbb{E}\left[\Big(\sum_{j\in\mathcal{N}^+_{1,i}} \omega_{2,j}\Big)\Big(\sum_{j'\in\mathcal{N}^+_{1,i'}} \omega_{2,j'}\Big)\right] \\
&= \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\Big(\sum_{j\in\mathcal{N}^+_{1,i}}\sum_{j'\in\mathcal{N}^+_{1,i'}} \mathbb{E}[\omega_{2,j}]\,\mathbb{E}[\omega_{2,j'}] + \sum_{j\in\mathcal{N}^+_{1,i'}\cap\mathcal{N}^+_{1,i'}} \mathrm{Var}[\omega_{2,j}]\Big) \\
&= \frac{D^2}{d^+_{1,i}d^+_{1,i'}}\Big(\sum_{j\in\mathcal{N}^+_{1,i}}\sum_{j'\in\mathcal{N}^+_{1,i'}} \lambda_j\lambda_{j'} + \sum_{j\in\mathcal{N}^+_{1,i'}\cap\mathcal{N}^+_{1,i'}} \lambda_j(1-\lambda_j)\Big) \\
&= D^2\left(\frac{\sum_{j\in\mathcal{N}^+_{1,i}}\lambda_j}{d^+_{1,i}}\cdot\frac{\sum_{j'\in\mathcal{N}^+_{1,i'}}\lambda'_j}{d^+_{1,i'}} + \frac{\sum_{j\in\mathcal{N}^+_{1,i'}\cap\mathcal{N}^+_{1,i'}}\lambda_j(1-\lambda_j)}{d^+_{1,i}d^+_{1,i'}}\right) = D^2\Big(\bar{\lambda}^+_{1,i}\bar{\lambda}^+_{1,i'} + \sigma^+_{1,i\cap i'}\Big),
\end{aligned}$$

where
$$\sigma^+_{1,i\cap i'} = \frac{\sum_{j\in\mathcal{N}^+_{1,i'}\cap\mathcal{N}^+_{1,i'}}\lambda_j(1-\lambda_j)}{d^+_{1,i}d^+_{1,i'}}.$$

Substituting the obtained expressions involving realized demands into the expression for the expected payoff of a retailer, we end up with

$$\begin{aligned}
\mathbb{E}[\pi_{1,i}] &= \mathbb{E}\left[R_{1,i}\Big(\sum_{i'\in\mathbb{T}^a_1}(1-\omega_{1,i}\omega_{1,i'})R_{1,i'} - (1-\omega_{1,i}\Delta\Big) - cd^+_{1,i} - L_{1,i}\right] \\
&= \sum_{i'\in\mathbb{T}^a_1}(1-\mathbb{E}[\omega_{1,i}]\,\mathbb{E}[\omega_{1,i'}])\,\mathbb{E}[R_{1,i}R_{1,i'}] - \Delta(1-\mathbb{E}[\omega_{1,i}])\,\mathbb{E}[R_{1,i}] - cd^+_{1,i} - \mathbb{E}[L_{1,i}] \\
&= D^2(1-\lambda_r^2)\sum_{i'\in\mathbb{T}^a_1}(\bar{\lambda}^+_{1,i}\bar{\lambda}^+_{1,i'} + \sigma^+_{1,i\cap i'}) - \Delta(1-\lambda_r)D\bar{\lambda}^+_{1,i} - cd^+_{1,i} - \mathbb{E}[L_{1,i}].
\end{aligned}$$

The expected payoff $\mathbb{E}[\pi_{2,j}]$ of a supplier has the following form:

$$\begin{aligned}
\mathbb{E}[\pi_{2,j}] &= \mathbb{E}[S_{2,j}p_2 - R_{2,j}p_3] = \mathbb{E}[\omega_{2,j}R_{2,j}(\Delta - S_2) - R_{2,j}(\Delta - S_3)] \\
&= (\text{as raw material production never fails}) \\
&= \mathbb{E}[\omega_{2,j}D_{2,j}(\Delta - \sum_{j'\in\mathbb{T}^a_2} S_{2,j'}) - D_{2,j}(\Delta - n^a_1 D)] \\
&= D_{2,j}\,\mathbb{E}[\omega_{2,j}(\Delta - \sum_{j'\in\mathbb{T}^a_2}\omega_{2,j'}D_{2,j'}) - (\Delta - n^a_1 D)]
\end{aligned}$$

$$= D_{2,j}\big(\lambda_j(\Delta - D_{2,j} - \sum_{j' \in \mathbb{T}_2^a \setminus \{j\}} \lambda_{j'} D_{2,j'}) - (\Delta - n_1^a D)\big).$$

Now, we use the obtained expected payoffs to analyze the behavior of retailers and suppliers, first, in the model without congestion, and, then, with congestion.

2) Supplier Behavior Without Congestion: Expected payoff of a retailer without congestion is

$$\mathbb{E}[\pi_{1,i}] = D^2(1-\lambda_r^2) \sum_{i' \in \mathbb{T}_1^a} (\bar{\lambda}_{1,i}^+ \bar{\lambda}_{1,i'}^+ + \sigma_{1,i\cap i'}^+) - \Delta(1-\lambda_r) D \bar{\lambda}_{1,i}^+ - c d_{1,i}^+.$$

From the above expression and the definition of

$$\sigma_{1,i\cap i'}^+ = \frac{\sum_{j \in \mathcal{N}_{1,i'}^+ \cap \mathcal{N}_{1,i'}^+} \lambda_j (1-\lambda_j)}{d_{1,i}^+ d_{1,i'}^+},$$

it is clear that the link concentration behavior by retailers in the model without congestion transfers persists with heterogeneous suppliers. Thus, at an equilibrium, there will be only one supplier, say $s \in \mathbb{T}_2$, to which all the active retailers link. Thus, at an equilibrium,

$$\bar{\lambda}_{1,i}^+ = \frac{\lambda_{1,i}^+}{d_{1,i}^+} = \lambda_{1,i}^+ = \sum_{j \in \mathcal{N}_{1,i}^+} \lambda_j = \lambda_s \quad \text{and} \quad \sigma_{1,i'\cap i}^+ = \lambda_s(1-\lambda_s).$$

We can also repeat the reasoning from the analysis of equilibria for the model without congestion, and conclude that, at an equilibrium,

$$\begin{aligned}
\mathbb{E}[\pi_{1,i}] &= D^2(1-\lambda_r^2) \sum_{i' \in \mathbb{T}_1^a} (\bar{\lambda}_{1,i}^+ \bar{\lambda}_{1,i'}^+ + \sigma_{1,i\cap i'}^+) - \Delta(1-\lambda_r) D \bar{\lambda}_{1,i}^+ - c d_{1,i}^+ \\
&= D^2(1-\lambda_r^2) \sum_{i' \in \mathbb{T}_1^a} (\lambda_s^2 + \lambda_s(1-\lambda_s)) - \Delta(1-\lambda_r) D \lambda_s - c \\
&= D^2(1-\lambda_r^2) n_1^a \lambda_s - nD(1-\lambda_r) D \lambda_s - c \\
&= D^2(1-\lambda_r)(n_1^a(1+\lambda_r) - n)\lambda_s - c.
\end{aligned}$$

As we assumed that non-empty equilibria exist, the factor in front of λ_s in the obtained expression must be non-negative. Hence, a best-responding retailer maximizes λ_s.

At the same time, at the equilibrium where retailers concentrate links, supplier expected payoff is

$$\mathbb{E}[\pi_{1,j}] = D_{2,j}(\lambda_j(\Delta - D_{2,j}) - (\Delta - n_1^a D))$$

if $j = s$ and $\mathbb{E}[\pi_{1,j}] = 0$ otherwise. Notice that, generally, $\Delta > D_{2,j}$, so, unsurprisingly, suppliers are incentivized to attract more demand from retailers.

Consequently, if suppliers are strategic about choosing their production success likelihoods λ_j, and taking into account that, in the absence of congestion, only one supplier gets links, suppliers are unconditionally incentivized to maximize their reliability to get a positive expected payoff.

3) Supplier Behavior With Congestion: To analyze the model with congestion, let us, first, update the corresponding expression for the expected retailer payoff, and, in particular, get the expected value of the congestion penalty.

$$L_{1,i} = \frac{1}{d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \frac{\gamma}{2}(S_{2,j})^2, \qquad \mathbb{E}[L_{1,i}] = \frac{\gamma}{2d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \lambda_j (D_{2,j})^2 = \frac{\gamma D^2}{2d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \lambda_j (F_{2,j})^2,$$

where $F_{2,j} = \sum_{i' \in \mathcal{N}^-_{2,j}} 1/d^+_{1,i'}$ describes how congested supplier $j \in \mathbb{T}_2$ is. Consequently, a retailer's expected payoff in the model with congestion and heterogeneous suppliers is

$$\mathbb{E}[\pi_{1,i}] = D^2(1-\lambda_r^2) \sum_{i' \in \mathbb{T}^a_1} (\bar{\lambda}^+_{1,i}\bar{\lambda}^+_{1,i'} + \sigma^+_{1,i\cap i'}) - \Delta(1-\lambda_r)D\bar{\lambda}^+_{1,i} - cd^+_{1,i} - \frac{\gamma D^2}{2d^+_{1,i}} \sum_{j \in \mathcal{N}^+_{1,i}} \lambda_j (F_{2,j})^2.$$

Already from the expression above we can see that, now, a retailer, besides aiming at picking reliable suppliers due to the first two summands in the expected payoff expression, may avoid linking to highly reliable suppliers, as the corresponding large λ_j would increase the congestion penalty—the last term in the expected payoff.

To get a better feeling for why retailers may prefer lower-reliability suppliers, let us look at two specific equilibria that we have already encountered.

First, let us inspect the symmetric equilibrium of Theorem 6. In it,

$$d^+_{1,i} = \tau, \quad d^-_{2,j} = n\tau/m, \quad F_{2,j} = \sum_{i' \in \mathcal{N}^-_{2,j}} 1/d^+_{1,i'} = n/m, \quad \bar{\lambda}^+_{1,i} = \lambda^+_{1,i}/\tau.$$

Consequently, again, assuming that every retailer is active, that is, $n^a_1 = n$,

$$\mathbb{E}[\pi_{1,i}] = D^2 \sum_{i' \in \mathbb{T}_1} \left(\bar{\lambda}^+_{1,i}\left((1-\lambda_r^2)\bar{\lambda}^+_{1,i'} - (1-\lambda_r) - \frac{\gamma n}{2m^2} \right) + (1-\lambda_r^2)\sigma^+_{1,i\cap i'} \right) - c\tau.$$

Notice that, in the obtained expression, the factor next to $\bar{\lambda}^+_{1,i}$ can be positive or negative, depending on the balance between the sizes of supplier and retailer sets, as well as the congestion penalty value. If this is the case, the first summand under the sum would be such that a best-responding retailer would choose lower-reliability suppliers to lower its $\bar{\lambda}^+_{1,i}$—the average reliability over the suppliers retailer i is linking to. Another term under the sum, involving $\sigma^+_{1,i\cap i'}$ would still drive the retailers to link to "mid-reliability" suppliers. Thus, the retailers would be driven to link to a few suppliers having an intermediate value of production success likelihood λ_j.

Secondly, let us investigate a simpler equilibrium of Sec. 5.2, when $n = m = 2$, and in the network, every retailer maintains a single link, sourcing from a separate supplier (a network with parallel links). In that network, assuming, w.l.o.g., that retailer $i \in \mathbb{T}_1$ links to supplier $i \in \mathbb{T}_2$,

$$d^+_{1,i} = 1, \quad d^-_{2,j} = 1, \quad F_{2,j} = 1, \quad \bar{\lambda}^+_{1,i} = \lambda_i, \quad \sigma^+_{1,i} = \lambda_i(1-\lambda_i),$$

which gives

$$\mathbb{E}[\pi_{1,i}] = D^2\big[-(1-\lambda_r^2)\lambda_i^2 + ((1-\lambda_r^2)(\lambda_1+\lambda_2+1) - 2(1-\lambda_r) - \gamma/2)\lambda_i\big] - c.$$

Expected payoff is a quadratic function, whose maximum in λ_i is attained at

$$\widehat{\lambda}_i = \lambda_1 + \lambda_2 + 1 - \frac{2}{1+\lambda_r} - \frac{\gamma}{2(1-\lambda_r^2)}.$$

In general, it is possible that $\widehat{\lambda}_i \in (0,1)$, in which case, suppliers would compete to drive their reliability λ_i towards an intermediate value.

Hence, in the model with congestion, suppliers may not be incentivized to improve their production reliability λ_i. At the same time, however, it is clear from the expression for the expected payoff of a retailer that improvement of the congestion cost γ_i is unambiguously good and lets the corresponding supplier attract more demand (links).

■

Theorem 7 confirms something intuitive—there is no reason why reduction of congestion cost γ_j can hurt a supplier. Indeed, reduction of delays in order fulfillment unequivocally makes the corresponding supplier more attractive for retailers, boosting the supplier's demand and, consequently, payoff. Alternatively, if we interpret the congestion penalty as a soft cap on supply, then it is unsurprising that suppliers prefer higher supply caps.

Surprisingly, higher production reliability, can actually harm a supplier. The intuition for why this happens is as follows. In a model with congestion, there are two competing forces present. One, coming from the base model without congestion, drives the retailers towards link concentration to secure better upstream prices. Another force, present in the form of the explicit congestion penalty term in the retailer's payoff in the model with congestion, drives the retailers towards diversifying and spreading their supplier bases to avoid congestion or long waits for their order fulfillment. When both these forces are present, their balance results in some optimal value of congestion for the retailers, and to approach that optimal congestion, the retailers are incentivized to link to "medium congestion" suppliers. Such suppliers are characterized by lower demand or lower production success likelihood.

In the light of Theorem 7, we can conclude that, despite the seeming similarity between production failures and production delays, these two types of failures are qualitatively different.

5.5. Discussion of the Model With Congestion

The model with congestion of Sec. 4.1 was obtained by extending the model of Sec 5.1. Being an extension, the model with congestion inherited the retailers' drive towards creating sparse networks (as they are cheap from the linking cost perspective) are concentrating links (as it allows retailers to attain better upstream prices). However, the congestion penalty introduces a countervailing force, that drives the retailers to create redundant links and spread them to achieve a certain optimal supplier overlap with their peers.

While the equilibrium networks of the model without congestion—where agents concentrate links—are absent in the supply chain literature, the equilibrium networks of the model with congestion: sparse networks possessing a sufficient amount of redundancy—and, in particular, the symmetric equilibrium network of Theorem 6—resemble k-partite graph expanders, which have been argued to form resilient supply chains (Chou et al. 2011). However, unlike this latter work, where the network structure was exogenously imposed, our resilient networks are endogenously formed by the agents in an uncoordinated fashion.

It is also surprising that, in a heterogeneous environment, according to Theorem 7, suppliers may not want to improve their production reliability, while always being willing to improve their production delays via reducing congestion costs. This behavior also stems from the balance between two forces present in the system—network sparsification and link concentration versus redundancy creation.

Allowing the agents to strategically set order quantities in addition to strategic linking does not change equilibrium networks in the model without congestion. One may wonder whether it affects the structure of equilibrium networks when congestion is present. Unlike the case without congestion, creation of redundant links and boosting order quantities interact through the congestion function. Strategic order quantity setting may result in the shrinkage of the retailers' supplier bases at equilibrium. Characterization of the balance between the extent of supplier base diversification and over-ordering at equilibrium is *an open problem.*

6. Conclusion

This paper introduced a model of strategic formation of supply chain network in the presence of yield uncertainty and congestion. We use it to derive three conclusions. First, in the absence of congestion, retailers tend to create very sparse networks and concentrate links, which results in chain-like multi-tier networks. Sparsity is the result of unconstrained supplies, while link concentration lets retailers secure lower prices at the high yield upstream.

In the presence of congestion, retailers tend to form expander-like networks, which are sparse, yet possessing sufficient redundancy. Finally, we show the qualitative difference between yield uncertainty and congestion failures in an environment with heterogeneous strategic suppliers: reducing congestion costs is unambiguously beneficial. Improving production reliability, however, is beneficial only in the absence of congestion. In the presence of congestion making production process more reliable can actually hurt a supplier. That is, there can be too much production reliability in a supply chain.

In this paper, we focused on multi-sourcing as the primary device for mitigating production disruptions. Others focus on varying ordered quantities to tackle uncertainty. Combining the two in one model—would not change the equilibrium networks that emerge in the absence of congestion—is an important direction for future study of supply chain network models with congestion.

Acknowledgments

The authors are grateful to Ozan Candogan and Ben Golub for constructive suggestions that have led to tangible improvements of this work.

References

Adida E, DeMiguel V (2011) Supply chain competition with multiple manufacturers and retailers. *Operations Research* 59(1):156–172.

Akan M, Ata B, Olsen T (2012) Congestion-based lead-time quotation for heterogeneous customers with convex-concave delay costs: Optimality of a cost-balancing policy based on convex hull functions. *Operations Research* 60(6):1505–1519.

Ambrus A, Mobius M, Szeidl A (2014) Consumption risk-sharing in social networks. *American Economic Review* 104(1):149–82.

Amelkin V, Vohra R (2019) Yield uncertainty and strategic formation of supply chain networks. *arXiv:1907.09943 [cs.GT]* Available at `https://arxiv.org/abs/1907.09943`.

Ang E, Iancu DA, Swinney R (2016) Disruption risk and optimal sourcing in multi-tier supply networks. *Management Science* 63(8):2397–2419.

Anupindi R, Akella R (1993) Diversification under supply uncertainty. *Management Science* 39(8):944–963.

Babich V, Burnetas AN, Ritchken PH (2007) Competition and diversification effects in supply chains with supplier default risk. *Manufacturing & Service Operations Management* 9(2):123–146.

Bernstein F, Federgruen A (2005) Decentralized supply chains with competing retailers under demand uncertainty. *Management Science* 51(1):18–29.

Bimpikis K, Candogan O, Ehsani S (2019) Supply disruptions and optimal network structures. *Management Science* 65(12):5504–5517.

Bimpikis K, Fearing D, Tahbaz-Salehi A (2018) Multi-sourcing and miscoordination in supply chain networks. *Operations Research* 66(4):1023–1039.

Bohn RE, Terwiesch C (1999) The economics of yield-driven processes. *Journal of Operations Management* 18(1):41–59.

Cachon GP (2003) Supply chain coordination with contracts. *Supply Chain Management: Design, Coordination and Operation*, volume 11 of *Handbooks in Operations Research and Management Science*, 227–339 (Elsevier).

Cachon GP, Lariviere MA (1999a) Capacity choice and allocation: Strategic behavior and supply chain performance. *Management Science* 45(8):1091–1108.

Cachon GP, Lariviere MA (1999b) An equilibrium analysis of linear, proportional and uniform allocation of scarce capacity. *Institute of Industrial and Systems Engineers (IISE) Transactions* 31(9):835–849.

Cachon GP, Terwiesch C (2008) *Matching Supply with Demand* (McGraw-Hill Publishing).

Carr SM, Karmarkar US (2005) Competition in multiechelon assembly supply chains. *Management Science* 51(1):45–59.

Cho SH, Tang CS (2013) Advance selling in a supply chain under uncertain supply and demand. *Manufacturing & Service Operations Management* 15(2):305–319.

Chod J, Trichakis N, Tsoukalas G (2019) Supplier diversification under buyer risk. *Management Science* 65(7):3150–3173.

Chopra S, Meindl P (2016) *Supply Chain Management. Strategy, Planning & Operation* (Pearson).

Chou MC, Chua GA, Teo CP, Zheng H (2011) Process flexibility revisited: The graph expander and its applications. *Operations Research* 59(5):1090–1105.

Cormican K, Cunningham M (2007) Supplier performance evaluation: Lessons from a large multinational organisation. *Journal of Manufacturing Technology Management* 18(4):352–366.

Deo S, Corbett CJ (2009) Cournot competition under yield uncertainty: The case of the US influenza vaccine market. *Manufacturing & Service Operations Management* 11(4):563–576.

Fang Y, Shou B (2015) Managing supply uncertainty under supply chain Cournot competition. *European Journal of Operational Research* 243(1):156–176.

Hwang W, Bakshi N, DeMiguel V (2015) Simple contracts for reliable supply. *Working Paper* Available at `http://faculty.london.edu/avmiguel/HBD-2015-09-02.pdf`.

Johari R, Weintraub GY, Van Roy B (2010) Investment and market structure in industries with congestion. *Operations Research* 58(5):1303–1317.

Kotowski MH, Leister CM (2018) Trading networks and equilibrium intermediation. *HKS Working Paper No. RWP18-001* Available at `http://dx.doi.org/10.2139/ssrn.3104754`.

Kranton RE, Minehart DF (2001) A theory of buyer-seller networks. *American Economic Review* 91(3):485–508.

Mendelson H, Tunca TI (2007) Strategic spot trading in supply chains. *Management Science* 53(5):742–759.

Rong Y, Snyder LV, Shen ZJM (2017) Bullwhip and reverse bullwhip effects under the rationing game. *Naval Research Logistics (NRL)* 64(3):203–216.

Snyder LV, Atan Z, Peng P, Rong Y, Schmitt AJ, Sinsoysal B (2016) OR/MS models for supply chain disruptions: A review. *Institute of Industrial and Systems Engineers (IISE) Transactions* 48(2):89–109.

Song JS, Yano CA, Lerssrisuriya P (2000) Contract assembly: Dealing with combined supply lead time and demand quantity uncertainty. *Manufacturing & Service Operations Management* 2(3):287–296.

Tang SY, Kouvelis P (2011) Supplier diversification strategies in the presence of yield uncertainty and buyer competition. *Manufacturing & Service Operations Management* 13(4):439–451.

Tomlin B (2006) On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science* 52(5):639–657.

Wang Y, Tomlin B (2009) To wait or not to wait: Optimal ordering under lead time uncertainty and forecast updating. *Naval Research Logistics (NRL)* 56(8):766–779.

Weitzman ML (1974) Prices vs. quantities. *The Review of Economic Studies* 41(4):477–491.