Modelling Supply Chain Adaptation for Disruptions: An Empirically Grounded Complex Adaptive Systems Approach

Kang Zhao[1], Zhiya Zuo[2], and Jennifer V. Blackhurst[1+]

[1]Department of Business Analytics, The University of Iowa, Iowa City, IA 52242, USA. kang-zhao@uiowa.edu, jennifer-blackhurst@uiowa.edu.

[2]Department of Information Systems, City University of Hong Kong, Kowloon Tong, Hong Kong. zhiyazuo@gmail.com.

[+]Corresponding author.

Abstract

Through the development and usage of an agent-based model, this paper investigates firms' adaptive strategies against disruptions in a supply chain network. Viewing supply chain networks as complex adaptive systems (CAS), we first construct and analyze a real-world supply chain network among 2,971 firms spanning 90 industry sectors. We then develop an agent-based simulation to show how the model of firms' adaptive behaviors can leverage competition relationships within a supply chain network. The simulation also models how disruptions propagate in the supply chain network through cascading failures. With the simulation, we seek to understand if firms' adaptive behaviors can reduce the impact of disruptions in these networks. Therefore, we propose, evaluate, and analyze two types of adaptive strategies a firm can leverage to reduce the negative effects of supply chain network disruptions. First, we deploy in our model a reactive strategy, which restructures the network in response to a disruption event among first-tier suppliers. Next, we develop and propose proactive strategies which are used when a distant disruption is observed but has not yet hit the focal firm. We discuss the implications related to how and when firms can improve their resilience against supply disruptions by leveraging adaptive strategies.

Keywords: Supply Chain Disruptions, Complex Networks, Agent-based Models, Resilience.

Citation:

Zhao, K., Zuo, Z., & Blackhurst, J. V. (2019). Modelling supply chain adaptation for disruptions: An empirically grounded complex adaptive systems approach. *Journal of Operations Management*, *65*(2), 190–212. https://doi.org/10.1002/joom.1009

1. Introduction

Due to the complexity, uncertainty and interdependence of today's supply chains, there is an increased risk of loss in the supply chain network due to a disruption event (Bode et al., 2011; Bode and Wagner, 2015; Kamalahmadi and Parast, 2016). A disruption in a supply chain network is defined as an event that disrupts the flow of goods or services (Craighead et al., 2007). Losses stemming from supply chain network disruptions may manifest as financial loss, a loss in operational performance and even a loss of market position (Hendrick and Singhal 2003; Hendricks and Singhal, 2005; Wagner and Bode, 2008). Moreover, because of the interconnected nature of supply chain networks, a disruption may propagate and cascade through the supply chain (Hearnshaw et al., 2013; Fiksel et al., 2015) with increasing magnitude or severity of impact (Van der Vegt et al., 2015). In other words, a disruption may not originate from the focal firm's immediate suppliers but rather elsewhere in the network (Blackhurst et al., 2005; Kim et al., 2015). A lack of understanding of how the supply chain network is structured may exacerbate the impact of disruptions and inadvertently allow disruptions to propagate (Kim et al., 2015). Managers of real-world supply chains find the cascading effect or propagation of a disruption difficult to understand (Fiksel et al., 2015). The ability to restructure the supply chain in the face of changing conditions is critical to maintain continuity of supply chain performance (Hearnshaw, et al. 2013). Flows of materials within the supply chain network need to be redirected and structures need to be adapted to allow for continuity in operations. As such, there have been calls to examine the structure of supply chain networks and determine the ability of the network to adapt in the face of supply chain disruptions (Hearnshaw et al., 2013; Kim et al., 2015; Van der Vegt et al., 2015).

In this study, we view a supply chain network as a complex adaptive system (CAS) (Choi et al., 2001) where, in the face of a disruption, firms connected in a complex network have the ability to adapt and restructure their connections. The CAS framework provides a useful theoretical foundation for this study (Thomson, 1967; Anderson, 1999; Choi et al., 2001) as firms in a supply chain operate as an interconnected network in a dynamic environment (Blackhurst et al., 2011; Bode et al., 2011; Kim et al., 2011). Therefore, even a small change at one node in the chain can cause a disruption to spread, impacting other nodes in the chain (Craighead et al., 2007). We posit that firms in a supply chain constitute self-organizing networks. In addition, some supply chains can be adaptive or resilient. When hit with a disruption, they can adapt or restructure themselves to reach a desirable state (back to the original state, an equivalent state, or better) (Ambulkar et al., 2015). In viewing supply chain networks as adaptive systems, the ability to adapt and restructure is critical for minimizing losses from disruptions (Hearnshaw et al, 2013; Ambulkar et al., 2015). The effectiveness of adaptive restructuring strategies in improving network resilience after node removal has been illustrated in other complex systems, such as food webs (Staniczenko et al. 2010). In addition, Nair and Vidal (2011) noted that network topology is an important factor with regards to spreading disruptions. However, recent research on resilience to supply chain disruptions has not fully incorporated the role of network structures (Kim et al., 2011) and lacks a clear understanding of disruptions and their impact at a network level (Kim et al., 2015). In other words, understanding how disruptions impact multiple tiers in a supply chain and how the structure of the network may play a role in this impact is lacking. In order to address these gaps in the research, we seek to answer the following research question:

> *How can firms leverage different types of adaptive strategies in the supply network to improve resilience against supply disruptions?*

Inspired by both supply chain management and network science literatures on rewiring edges (Watts and Strogatz, 1998; Zhao et al., 2011b), our study presents and examines two types of adaptive strategies to restructure a supply chain network: 1) a *reactive strategy*, which restructures the network in response to a disruption event among first-tier suppliers. In other words, reactive strategies are used when an immediate supplier of a focal firm fails. Next, we develop and propose 2) *proactive strategies*. These strategies focus on restructuring the network after observing a distant firm failure (beyond first tier) in order to avoid possible disruptions to the focal firm. Representing a forward-looking approach, proactive strategies are in anticipation of a disruption (which has already occurred in another part of the network) hitting the focal firm and will identify the weakest spot specific to the disrupted distant firm in the network.

In order to study firms' adaptive strategies that improve their resilience to supply chain disruptions, this study develops agent-based simulations based on large-scale real-world supply networks. Our modeling of adaptive behaviors incorporates the structure of both supply chain networks (which connects partner firms in the supply chain,) and competition networks (which connect competing firms in the supply chain) so that we can investigate how competition relationships among firms in a supply chain network can be exploited to develop resilience against disruptions (in Sections 4 and 5.1). The two networks are again used to model and analyze firms' proactive strategies (in Section 5.2) including factors related to the effectiveness of proactive strategies (in Section 5.3).

This research proceeds in four steps: First, we collect data of 2,971 firms from 90 industries to construct a large-scale supply chain network among these firms, along with an accompanying competition network. The data was collected through scraping a database for information on firms including their financial data as well as relationship data among firms. We reveal the complex structural properties of these networks and show a firm's partnership and competition with others are interweaved. Second, we design agent-based simulation models for firms' reactive strategies in this complex system, and the propagation of disruption impact. Third, we use the models to evaluate the impact of disruptions and illustrate the effectiveness of reactive behaviors in reducing the impact of disruptions. Fourth, we propose, evaluate and analyze proactive strategies that firms can use to improve their supply chain resilience against distant disruptions.

This study makes a number of important contributions to the understanding of supply chain networks. First, our agent-based model leverages structures of both real-world supply chain and competition networks as well as firm attributes, to realistically model key components of complexity in supply chain networks, namely the propagation of a disruption in the supply chain and firms' adaptive behaviors to manage disruption risk. The use of competition networks opens interesting possibilities to not only handle disruptions more effectively, but also to gain advantage in the market by leveraging visibility of relationships and structures within the network. Second, we illustrate how the insights gained in this study can be used by a focal firm to restructure its supply chain network so that it becomes more resilient against supply chain network disruptions in a real-world setting. By using both the supply chain and competition networks, this research helps to better understand the effectiveness of adaptive strategies within complex supply chain networks in the face of supply chain disruptions.

The remainder of this paper is organized as follows. After covering the theoretical foundation and related studies for this research in Section 2, we introduce how we collect empirical data to construct and analyze large-scale supply chain and competition networks in Section 3. Section 4 describes the agent-based model we develop for this research, and Section 5 shows results from our simulations and related experiments. The paper concludes with a discussion of the results, future work and limitations in Section 6.

2. Literature

This section covers the literature related to supply chain networks as complex adaptive systems as well as disruptions in the supply chain network.

2.1 Supply Chain Networks as Complex Adaptive Systems

Based on the seminal work of Choi et al. (2001), a CAS is defined as an "interconnected network of multiple entities (or agents) that exhibit adaptive action in response to both the environment and the system of entities itself" (Pathak et al., 2007, pg. 550). A CAS is a self-organizing system and it reconfigures its internal and external linkages to continually evolve over time (Anderson, 1999; Choi et al., 2001). Kim et al. (2015) and Nair et al. (2009) note that CAS is a useful theory in describing supply chain network structures. Pathak et al. (2007) term supply networks as a typical case of CAS because a supply chain will adapt via interactions of nodes within the network and evolve over time. In applying CAS to supply chain networks, Pathak (2007, pg. 562) states that such a network consists of "interconnected autonomous entities that make choices to survive and, as a collective, the system evolves and self-organizes over time". This is particularly applicable in looking at disruption propagation in supply chains. In a supply chain network, a disruption such as a supplier failing will cause the agent (focal firm) to seek an alternative supplier (using schema defined as a plan or decision-making logic) leading to a change in the network structure. Interestingly, Choi et al. (2001) note that supply networks are complex and dynamic, and changes that occur within the network (such as at a second or third tier supplier) are often outside of a focal firm's awareness. In a supply chain network disruption context, this means that a disruption can occur without the focal firm knowing that it will be affected. However, because of the interconnected nature of the network, a disruption may propagate and worsen and eventually have a severe impact to the focal firm (Hearnshaw et al., 2013; Fiksel et al., 2015). In this paper, a CAS lens allows us to view a supply network as a complex system where individual firms can adapt and restructure their networks in the face of a supply chain disruption.

2.2 Disruptions to Supply Chain Networks

While researchers have studied the resilience of supply chain networks from a complex network perspective (Thadakamalla et al. 2004; Zhao et al. 2011b; Kim et al., 2015; Zhao et al., forthcoming), structures of large-scale supply chain networks have often been synthetically created based on the assumption that supply chain networks follow certain network topologies, such as ER random, small-world, or scale-free networks (Pathak et al., 2007). While these topologies have been observed in various complex social and physical networks, their applicability for supply chain networks are not clear in the literature. For example, empirical studies on supply chain network structures have found that some supply chain networks have truncated power-law degree distributions (Saavedra et al., 2008), while some do not (Atalay et al., 2009; Kito et al., 2014). Another disadvantage of using synthetically created supply chain networks is that the connection between firm/node attributes (such as industry and size) and their network positions is often ignored, even though both firm attributes and network positions are important when measuring resilience against disruptions. For example, a large retailer (e.g., Walmart) may have a great number of suppliers but few customer firms in the supply chain network, while a semiconductor firm with a medium size (e.g., ARM Holdings PLC) can have the opposite network position: many customers and few suppliers.

There have been studies based on real-world supply chain networks, but they are often limited to networks for a specific product, such as automobiles (Kim et al., 2011). Although such an approach can help to understand a focal firm's operation at a very fine-grained level (e.g., different parts for a product), it can be challenging to obtain such detailed information of product-level flows among firms and how each firm uses parts it procures from suppliers to produce its own products, especially when the network goes beyond a couple of hundred nodes.

A recent study (Brintrup et al., 2016) made a notable effort to address this problem by integrating product flows among more than 18,000 inter-connected firms in the automobile industry and studied disruptions to this network from a topological perspective. However, at such a large scale, product information is only available at the level of generic product categories (e.g., air conditioner and gearbox) rather than for specific product models, and relationships among different products (e.g., what parts are needed to produce an air conditioner) are very difficult to capture. Another limitation of these product-specific supply networks is that ties among firms mainly reflect how parts or materials related to the product would flow from various suppliers to the focal firm. Other connections among the focal firm's suppliers are often missing or incomplete if they are not directly related to the product. For possible disruptions to a specific product (e.g., automobiles), such networks are sufficient and provide accurate information on the flow of related goods. However, firms often feature different lines or types of products for different types of customers (e.g., Samsung Electronics manufactures smartphones, TVs, and home appliances for consumers, along with semiconductors, LED panels, and network infrastructures for other firms). Thus, studying disruptions at the firm level requires the collection and aggregation of such product-specific supply networks for different products a firm provides, which is a daunting task.

Meanwhile, while some have examined the resilience of large-scale real-world supply chain networks, studies focused on the propagation of disruptions and the effect of adaptive behaviors during propagation are lacking in the literature. For example, Brintrup et al. (2011) focus on the resilience of Toyota's supply chain network from a topological perspective. Saavedra et al. (2008) analyze how a supply-chain network in the garment industry shrank over years. They propose a preferential-attachment model for how a node replaces a lost partner with a new one that is already well connected in the network. However, these studies do not incorporate the cascading failures of nodes nor the effects of firms' adaptive strategies. Therefore, this study seeks to understand how firms can restructure supply chain networks to improve their resilience against cascading disruptions.

3. Network Construction and Analyses

In this section, we will describe how we collect data, construct the supply and competition networks, and illustrate characteristics of the two networks.

3.1 Data Collection and Network Construction

Our supply chain network is created from a secondary data source using Mergent Horizon (http://www.mergenthorizon.com). Mergent Horizon lists information about global firms including company reports, financial data, competitors, customers and suppliers. We scrape the database using a snow-ball sampling approach, a technique for collecting large-scale network data (Carrington et al. 2005), with the Boeing Company as the seed node or anchor for the entire network in order to give perspective to the network. Raw HTML files from each firm's web page in the database are collected by our scrapers (a computer program that retrieves and parse files from the Web), and parsed to retrieve not only its attributes, but also its suppliers, customers, and competitors, for subsequent scraping. Firms in our final dataset include Boeing's tier-1 suppliers, tier-2[1] suppliers, and tier-3 suppliers, as well as Boeing's customers, and customers of Boeing's tier-1 and tier-2 suppliers. In addition to firms, the final dataset also includes supplier-customer

[1] Note that when going beyond first tier suppliers, we cannot guarantee a supplier of a firm's tier-1 suppliers is a tier-2 supplier of the firm. This is because we do not have data on exact product flows among firms and how a firm uses its incoming materials to generate outgoing products. However, for the purpose of simplicity and naming convention, we still call these firms as tier-2 suppliers of the focal firm. Similar naming schemes apply to tier-3 and tier-4 suppliers.

relationship among all the firms in the dataset. The unshaded portion of Figure 1 shows the coverage of firms in our dataset, where the direction of arrows corresponds to the direction of data collection. Also, even though our data collection starts with Boeing, the network is not centered on Boeing, because for each firm, we also capture its relationships with all other firms that appear in the dataset.

Our data collection (see Figure 1) yields 2,971 firms that are headquartered in 63 countries from North America, Europe, Asia, Latin America and Africa. As earlier mentioned, firms in our dataset span beyond the Aerospace industry, and cover 90 different sectors, with Internet & Software (8.65%), Semiconductors (5.39%), and Industrial Machinery & Equipment (4.54%) being the top 3 most-represented sectors. Such a rich dataset enables us to build a large-scale global supply chain network that spans multiple industries[2].

In addition to building a supply chain network based on relationships of a firm's suppliers and customers, our collection of competitor data for these 2,971 firms, along with the product overlap between two competitors (provided by Mergent), makes it possible to construct a competition network among these 2,971 firms. Competitors beyond these 2,971 firms are not included in the competition network. The supply and the competition networks among the same set of firms are essentially a multi-relational network among these firms (Yan et al, 2012: Zhao et al., 2016), but we treat them as two networks to simplify implementations.

Constructions of the two networks are shown side by side in Figure 1 with the shaded portion being the competition network. The original supply chain network is denoted by $G_S(V,E_S)$ and the competition network by $G_P(V,E_P)$. The two networks share the same set of nodes V, with

[2] To show the dataset we collected with Boeing as the seed firm is a representative sample, we also retrieve another set of data, and find that the new data and the network based on it feature similar characteristics with the ones used in this paper. More details are in Appendix 1.

|V|=2,971. Each node $v_i \in V$ corresponds to a firm. However, the edge sets are different for the two networks. $e_{i,j} \in E_S$ in the supply chain network represents a directed and unweighted supply edge from v_i to v_j, which means v_i is a supplier of v_j. By contrast, $e'_{i,j} \in E_P$ in the competition network is a directed and weighted edge between v_i and v_j, indicating that v_i is a competitor of v_j. The weight of $e'_{i,j}$ is proportional to the product overlap between v_i and v_j. Take three firms, Boeing (ID 1048), Curtiss-Wright (ID 287) and Airbus Group (ID 103714), as an example. Curtiss-Wright is a supplier of Boeing and $e_{287,1048} \in E_S$ is an edge in the supply chain network. Airbus Group is a competitor of Boeing. Among Boeing's areas of business, 10 out of 17 are also within Airbus Group's areas of business. Thus, there is an edge $e'_{1048,\,103714} \in E_P$ in the competition network with weight 10/17=0.59.

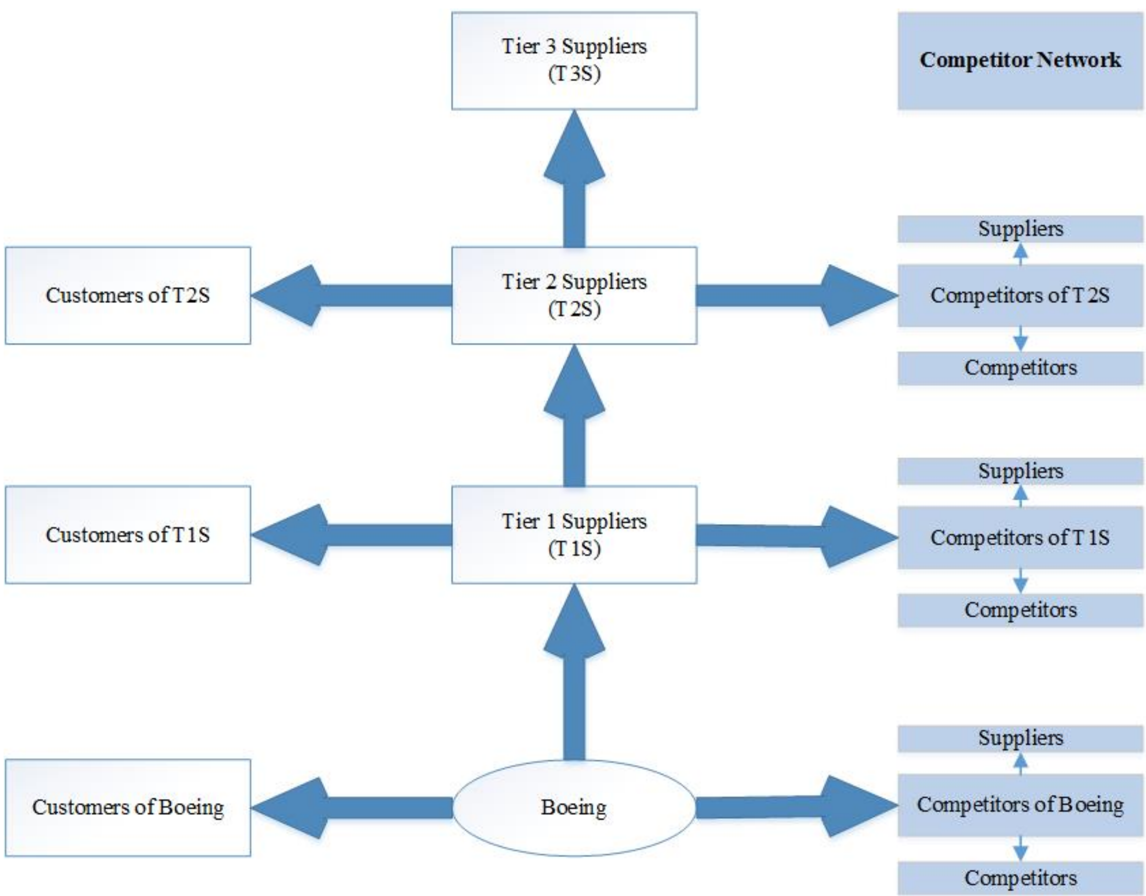

Figure 1. Data collection flow for our supply chain network and competition network.

3.2 Topological Analyses

Figure 2 visualizes the supply chain network. Table 1 provides a summary of the basic statistics of the supply and competition networks including number of nodes, number of edges, characteristic path length (the average shortest path length between a pair of nodes), diameters (the maximum of shortest path length between any two nodes), and clustering coefficients (the probability that a node's two neighbors are connected to each other). For example, the average distance between two nodes in the supply network is only 4.7 hops, and any two nodes are no more than 13 hops away from each other. These characteristics are similar to many real-world complex networks (Newman, 2003).

Both the supply and the competition networks feature complex topologies with highly-skewed degree distributions (Barabasi and Albert, 1999; Albert and Barabasi, 2002). In Table 2, we compare four common degree distributions for complex networks—Power-law (PL), Exponential (EXP), Truncated Power-law (Truncated PL), and Log-Normal (LN). We find that Truncated PL offers the best fit for the supply network's degree distribution. As for the competition network, Truncated PL still fits its degree distribution better than the other three, although differences are not statistically significant when comparing Truncated PL with LN and PL. Figure 3 shows both networks' degree distributions with fitted Truncated PLs estimated using the approach by (Clauset et al., 2009). In other words, most nodes in the network have few neighbors, while there are few nodes with many neighbors. According to previous studies (Thadakamalla et al. 2004; Zhao et al. 2011a), such a supply network with highly skewed degree distributions is usually robust against random failures but is more fragile when important nodes with high degrees are removed. We will evaluate this later in our simulation analysis.

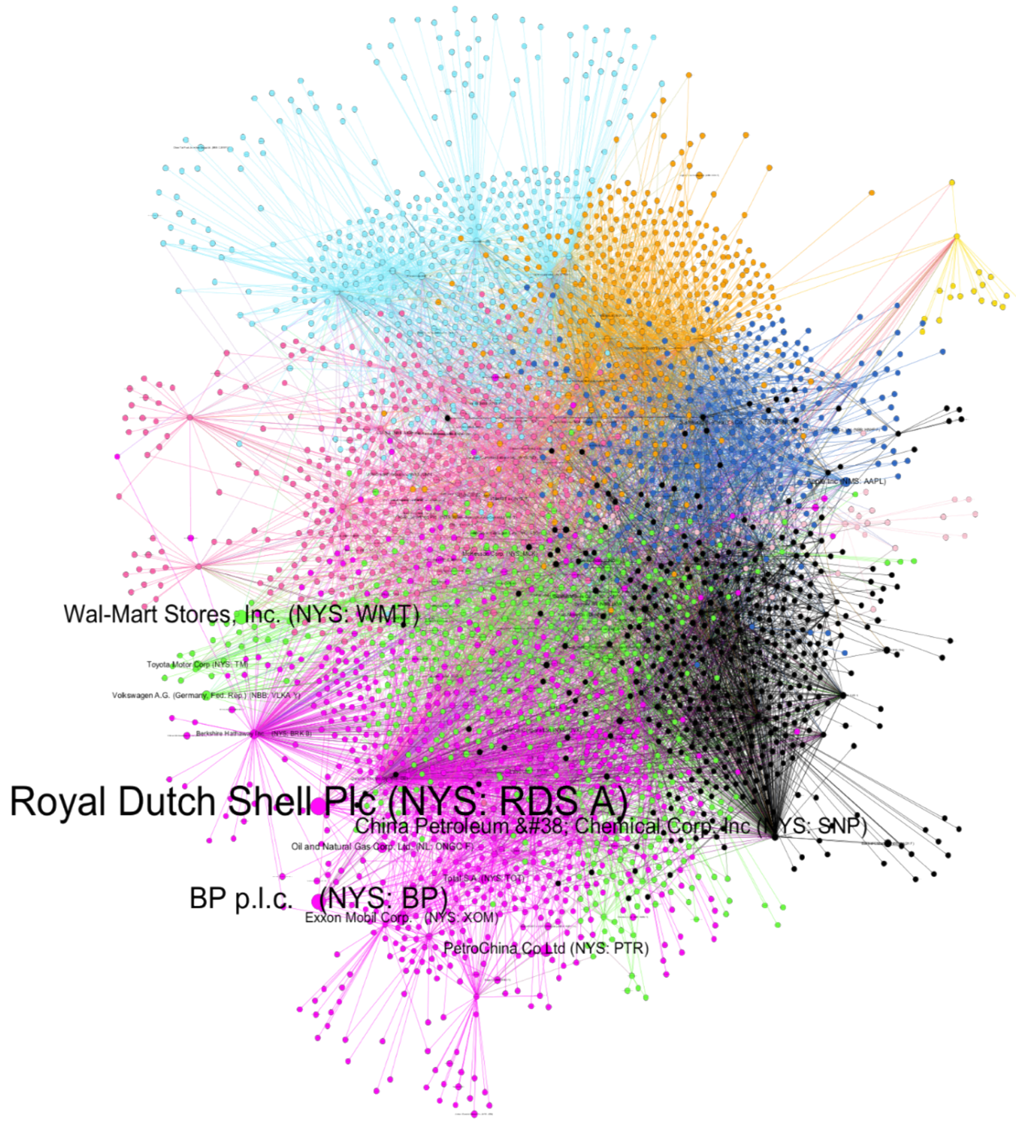

Figure 2. Visualization of the supply chain network. The size of a node is proportional to the corresponding firm's size, measured by log(revenue); the color of a node represents the network cluster (generated with modularity maximization) the node belongs to.

Table 1. Basic network statistics

	# of nodes	*# of edges*	*characteristic path length*	*diameter*	*clustering coefficient*
Supply Network	2,971	9,535	4.704	13	0.179
Competition Network	2,971	6,372	5.229	13	0.401

Table 2. Degree distributions of the two networks.

	Supply Network	Competition Network
Truncated PL vs. PL	5.461***	0.406
Truncated PL vs. EXP	10.001***	2.460**
Truncated PL vs. LN	2.452*	0.745
LN vs. EXP	9.795***	2.458*
LN vs. PL	4.745***	0.216
$*: p < 0.05$; $**: p < 0.01$; $***: p < 0.001$		
Note: The table shows log-likelihood ratios for pairwise comparisons between candidate distributions. If the value is positive, the first distribution is better. Otherwise, the second one fits the data better. For example, 6.168 indicates that, for the supply network, Truncated Power-law fits the degree distribution better than Power-law.		

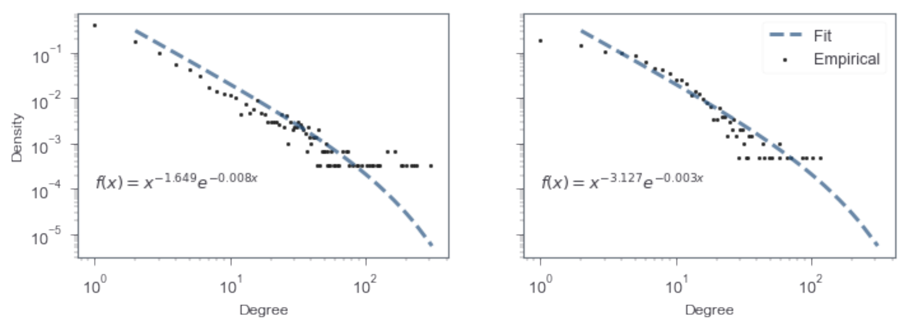

Figure 3. Log scale degree distributions of the supply network (left), and the competition network (right). The mathematical equations are fitted Truncated Power-law for each degree distribution.

In addition, we also find a surprising relationship between the supply chain and competition networks. As shown in Figure 4, it is possible that a firm's competitor is also among the same firm's upstream suppliers. For example, there is a probability of 7% that a competitor of a focal firm also serves as the focal firm's upstream supplier that is 3 hops away in the supply chain network. Such a probability is as high as 20% for a competitor to be an upstream supplier that is 4 or more hops away.

All these findings further highlight the complexity embedded in supply chain networks: two randomly chosen firms are, on average, within only 5 hops away from each other in the supply chain network, and a firm's competitor can also serve as an upstream supplier. They also illustrate the necessity to build a large-scale firm-level supply chain network using real-world data, because synthesized networks or product-specific networks can hardly reveal such structural complexity.

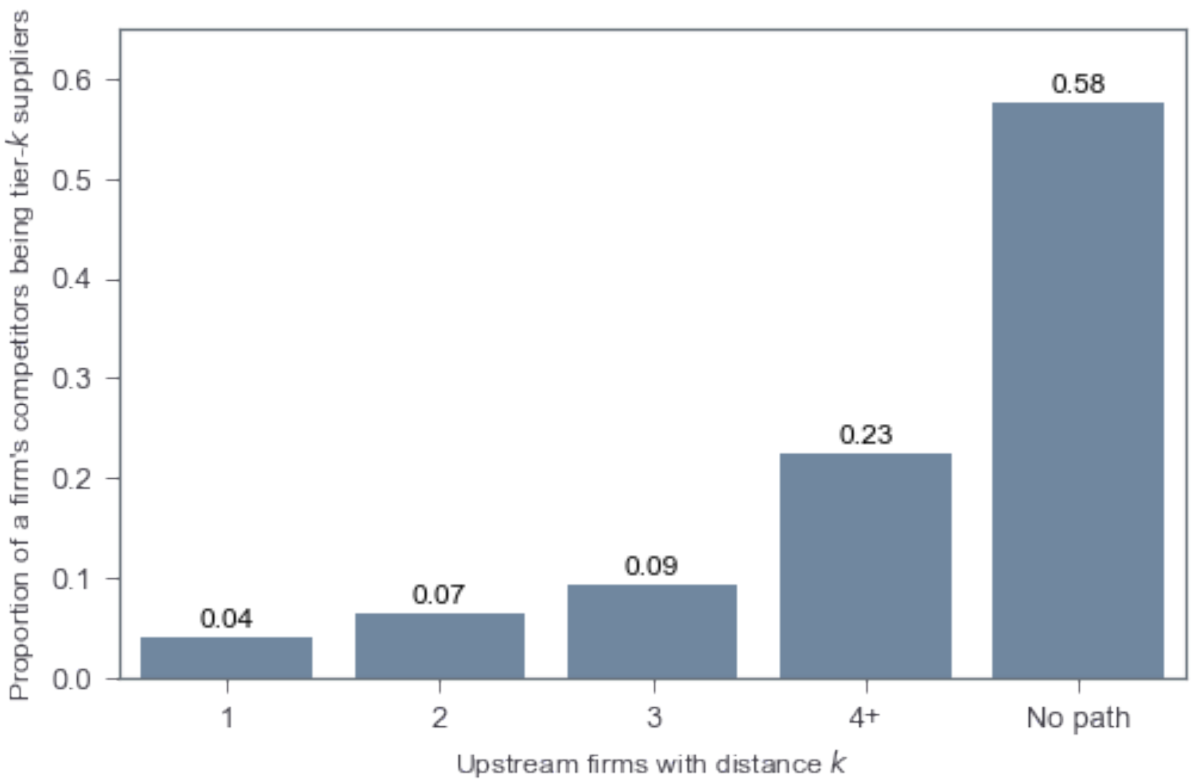

Figure 4. Probability of a firm's competitor is among its upstream firm at different distance.

4. The Agent-based Model

Agent-based modelling (ABM) is a powerful tool for the study of CAS (Axelrod 1997; Wilensky and Rand, 2015; He et al., 2016). ABM can capture phenomena in CAS by simulating how each individual (i.e., agent) makes decisions based on its interactions with the environment and other agents (Wu et al., 2013). Agents can also adapt and evolve (He et al., 2015). Pathak et al. (2007) and Nair and Vidal (2011) discuss how interconnected entities may adapt in response to a change in the system. This adaptation could lead to a restructuring of the network. Following this logic, we develop an agent-based model to study a supply chain network as a CAS. Specifically, the model simulates how firms respond to disruptions via reactive behaviors and how the impact of disruptions propagates in large-scale supply chain networks. In our ABM, each firm/node is represented as an autonomous agent. In reality, a firm whose supplier ceases to operate may not simply wait for the supplier to recover. Instead, the firm will try to find alternative sources of supplies and could request new connections with one of these alternative suppliers in order to resume its own normal operations. When receiving such requests to build new connections, these alternative suppliers will also decide whether to accept such requests.

Therefore, the first key component of our ABM is to model firms' adaptive behaviors when facing disruptions in the supply chain network $G_s(V, E_s)$ by leveraging the competition network G_p (V, E_p) among firms, where V is the set of all firms, and E_s and E_p are sets of edges in the two networks respectively. In other words, this adaptive strategy of network restructuring is reactive, as it occurs in response to a first-tier supplier failure. To implement this strategy, each run of our ABM consists of multiple iterations of inter-agent interactions, and each iteration has two steps: Step 1 and Step 2. Assume an initial disruption occurs at agent v_i and forces it to cease operations (i.e., removing v_i from V, and its edges from E_s and E_p) at time *t*. After that initial disruption, time ticks *t+odd_number (e.g., t+1, t+3, t+5,…)* in our model are for customers of v_i to find and send requests to alternative suppliers (described in Step 1.1 and Step 1.2), and time ticks *t+even_number* *(e.g., t+2, t+4, t+6…)* are for alternative suppliers to decide which requests to accept (described in Step 2).

Another key component of our ABM is to model how the impact of a disruption propagates in a supply chain network. Our modeling of such propagations was based on agents' reactive behaviors in seeking alternative suppliers. If an agent, who needs to find an alternative supplier due to the failure of one of its original suppliers, cannot secure one such supplier, then the agent's operation will be disrupted, and may even cease all of its operations with certain probabilities. If the agent does stop operating due to the lack of alternative suppliers, then it will be removed from the supply chain network and as a result, all its customers will need to seek alternative suppliers (described in Step 1.3). Such consecutive removal of nodes from the supply chain network after the initial node removal will constitute cascading failures, and model the propagation of disruptions across the whole supply chain network.

The complete model is specified as follows (Appendix 2 lists pseudo-code of the model). After the initial node removal (Step 0), the model includes two major components: firms seeking alternative suppliers and send requests (Step 1), and alternative suppliers deciding which requests to accept (Step 2).

Step 0: The Initial Node Removal. At the very beginning, the model will remove one firm from the supply chain network. Edges attached to the node are also removed. Users of our model can decide which node is removed initially. Such a single node removal at the beginning may cause cascading failures of other nodes later.

Step 1: Seeking Alternative Suppliers. After the removal of a node v_i from G_s (it could be the initially removed node, or a node removed by cascading failures), each customer of v_i (denoted as $v_m \in C_i$, where $\exists e_{i,m} \in E_S$) will try to find alternative suppliers in the following way:

Step 1.1. Identify potential alternative suppliers. v_m considers all direct competitors of v_i from G_p (denoted as $v_n \in P_i$, where $\exists e'_{i,n} \in E_P$) as its list of candidates. Each agent v_n in the candidate list of v_m will be approached by v_m with probability $p_{n,m} = \frac{k * w_{i,n}}{\sum_{P_i} k * w_{i,n}}$, where $w_{i,n}$ is the edge weight between v_i and $v_{i,n}$ in the competition network G_p. Recall that edge weight between two firms in the competition network represents the two firms' overlap in products. The definition of $p_{n,m}$ reflects the intuition that a competitor that is more similar to v_i in terms of products is more likely to provide what v_i supplies to its customers previously. k is a weighting factor to give higher preference to existing partners (customers or suppliers). The preference for and benefits of using existing suppliers is well documented in the literature, including potential liabilities and risk exposure in using new and unproven partners where capabilities and trust have not been established (Dyer and Singh, 1998; Wagner and Friedl, 2007), "switching inertia" in using new suppliers (Li et al., 2006), and the ability to leverage specific assets of the relationship (Dwyer et al., 1987; Azadegan et al. 2011). In fact, some literature has noted the "liability of newness" can make ventures susceptible to risk events (Azadegan et al., 2013).

In our model, if $\exists e_{m,n} \in E_S$ or $\exists e_{n,m} \in E_S$, we try two different ways to sample the value of k: (1) $k \sim N(1.5, 0.1)$, a normal distribution with a mean of 1.5 and a standard deviation of 0.1 with a minimum of 1; (2) $k \sim N(1.5, 0.2)$, a normal distribution with a mean of 1.5 and a standard deviation of 0.2 with a minimum of 1. Otherwise, k is set to 1 for non-partners.

Step 1.2. Stop approaching alternative suppliers. If an alternative supplier approached by v_m does not accept its request, then v_m will remove that supplier from its candidate list and decide which one to approach in subsequent trials by recalculating $p_{n,m}$ for each remaining candidate. v_m will approach one alternative supplier at each time tick for Step 1 (*t+1, t+3, t+5,* ….) and will stop approaching alternative suppliers when any one of the three following conditions is met:

(1) A request for an alternative supplier is accepted.

(2) It has reached the maximum number of trials allowed for an agent but has not secured an alternative supplier. The maximum number of trials is set to 10 (i.e., one iteration of a simulation will stop after *t+20*)[3].

(3) v_m already exhausts all of its alternative supplier candidates (e.g., v_i only has 5 competitors) before reaching its maximum number of trials.

If the 2nd or the 3rd condition is met, that means v_m cannot secure an accepted request from an alternative supplier by the end of the iteration, and v_m will be marked as *"disrupted"*.

Step 1.3. Possible cascading failures. To simulate cascading failures, we remove a "disrupted" agent v_i from the network with mean probability of P_i^R, which depends on two factors about v_i.

First, P_i^R is proportional to the percentage of lost suppliers. A lost supplier for v_i refers to a supplier that meets both of the following conditions: (1) it has ceased operation and been removed from the supply chain network; and (2) v_i cannot find an alternative

[3] The degree distribution of the competition network is highly skewed. Only 10% of the firms in the network have more than 10 competitors. In other words, it is unlikely for a firm to have more than 10 alternative suppliers. Meanwhile, simulation data also shows that the conditional probability that a firm's seeking of alternative suppliers was stopped because of the threshold of 10 trials given that the firm needs an alternative supplier is below 0.03%. Thus, we believe that increasing this threshold will have minimal impact on simulation results.

supplier to replace the removed suppliers after going through Step 1.1 and Step 1.2. The percentage of lost suppliers refers to the ratio between a firm's number of lost suppliers and the firm's number of original suppliers prior to disruptions. For example, a firm that has 10 original suppliers but loses 1 supplier will have 10% of lost suppliers. The idea behind this factor is that the more suppliers a firm loses, the more likely this firm will fail to operate. For instance, all other things being equal, a firm that has lost 60% of its suppliers is more likely to fail than another one that lost only 10% of its suppliers. This is supported by supply chain risk research such as Trkman and McCormack (2009) who note that supplier failure is a key driver of risk in the supply chain where the loss of a supplier increases the riskiness to the firm.

Second, P_i^R is inversely proportional to the firm size of v_i. In addition to the status of a node's neighbors, as in the first condition, the size of a company also matters in cascading failures – other things being equal, larger firms are less likely to cease operations while smaller firms are more likely to fail due to limited resources and relationships (Azadegan et al., 2013).

Specifically, P_i^R, the mean probability that v_i is removed, is a function of agent v_i's size and percentage of lost suppliers. It is defined in Equation 1, where $Z_i = \log(Revenue_i)$ is the size of v_i, and L_i represent the percentage of lost suppliers for agent v_i. According to this formula, if all the suppliers of a firm are lost ($L_i = 1$), then a firm will be removed from the network with a mean probability of 1. As for firm sizes, the biggest firm serves as the baseline and does not get any size-based discount on its ability to survive. Meanwhile, the smallest firm would still have a non-zero probability to survive as long as it does not lose all the suppliers. Given the same L_i, the probability is linearly and inversely proportional to firm size Z_i. This logic is supported in the literature, where the size of the firm had been shown to be an advantage in managing risk and disruptions. For example, Chopra and Sodhi (2014) state that large firms have the ability to build resilience relatively inexpensively to better manage supply chain disruptions. Related to this in a supply chain security context, Park et al. (2016, pg. 126) found that larger firms are better able to handle security and safety issues due to "greater affordability for needed resource commitments". Finally, Azadegan et al. (2013) note that in comparison, larger and older firms are better at managing risk and are less susceptible to damage from disruptions than newer and small firms in a new venture context.

$$P_i^R = \begin{cases} 0, & if\ L_i = 0 \\ 1 - \frac{Z_i - \min(Z) + 1}{\max(Z) - \min(Z) + 1} \times (1 - L_i), & Otherwise \end{cases} \qquad \text{(Equation 1)}$$

Because P_i^R represents the mean probability that v_i is removed, we also add some variations to such a mean value. To ensure the robustness of our simulation results, we tried to randomly sample the value of node removal from two distributions with P_i^R as the mean: the first one is a normal distribution $N(P_i^R, 0.1)$ with a standard deviation of 0.1; the second one is a uniform distribution $Unif(P_i^R - 0.1, P_i^R + 0.1)$. Both distributions are truncated to make sure values sampled from them are in the range of [0,1].

Note that cascading failures may cause an agent to have more than one of its original suppliers removed. In this case, the agent will need to identify an alternative supplier for each of its original suppliers that are removed. Also, if a disrupted node is not removed in one iteration, it may still be removed in a subsequent iteration with a mean probability of P_i^R, although its P_i^R may increase if it continues to lose suppliers.

Step 2: Decisions by Alternative Suppliers. After receiving new requests from v_m to provide supplies, alternative supplier v_n needs to decide whether to accept new requests, and if so, which one(s) to accept. Note that there is an upper limit (U_k) on how many new requests each agent k can accept. The limit is proportional to the agent's firm size (represented by revenue) because we assume that larger firms are often able to accommodate more requests as they have more resources to review and accommodate requests.

Specifically, we evaluate two different ways to sample U_k. The first way is to divide all firms in our study into three categories based on their revenues, and sample U_k from three different normal distributions. Small firms are those whose revenues rank in the bottom 1/3 of all firms; large firms have revenues that rank within the top 1/3 among all; and the rest are medium firms. For small firms, $U_k \sim N(2, 0.5)$, where $N(2, 0.5)$ represents a normal distribution with a mean of 2 and a standard deviation of 0.5; for medium firms $U_k \sim N(4,1)$; for large firms $U_k \sim N(6,2)$. The actual values of U_k are rounded to the nearest integer with a minimum value of 0. The second way is to sample U_k from one discrete uniform distribution $Unif(\frac{Z_i'}{2} - 2, \frac{Z_i'}{2} + 2)$, where Z_i' is the integer closest to $Z_i = \log(Revenue_i)$. Such a distribution ensures that the smallest company would still have certain probabilities of offering extra capacities.

Once an agent has exhausted its capacity (i.e., reached its U_k), it will not accept any new request. If an agent loses a customer, who is removed from the network due to disruptions, then the agent can accommodate one more supply request beyond its original U_k. It is also worth noting that this upper limit of accepting requests for each agent is set at the beginning of a simulation run. In other words, for a given agent, its U_k will stay the same for one simulation run but may vary from one run to another.

If at any given time *t+t'* (after the initial node removal at *t*), v_n has not reached its upper limit U_n, it will accept new requests. If there is only one request to v_n, then v_n accepts it. If there is more than one request received by v_n, it creates a list of candidate requesters and decide which one(s) to accept based on the following rules:

Step 2.1. Preference to requests from existing partners. Among agents on the list, v_n will first consider requests from those who are already connected with v_n (either as a supplier or a customer). This assumption is similar to the one about picking alternative suppliers--existing partners are more attractive than new partnership. If there are multiple such requesters, then the probability for v_n to pick an existing neighbor v_m would be proportional to the product of the requester's firm size (following logic from Azadegan et al., 2013; Chopra and Sodhi, 2014; Park et al., 2016) and the similarity between v_n and the supplier v_m is trying to replace.

Step 2.2. Preference to requests from larger firms. If $v_{i,n}$ has not reached its upper limit after accepting requests from existing network neighbors at *t+t'*, it will try to accommodate requests from non-network-neighbors. Similar to Step 2.1, if there are multiple such requesters, then the probability for v_n to pick a new customer v_m would be proportional to the product of v_m's firm size and the similarity between v_n and the supplier v_m is trying to replace. If a request from a non-partner is accepted, a new edge will be added between v_n and v_m in the supply chain network.

After each iteration of a simulation run, our model checks if any agent(s) was removed from the network due to cascading failures during the iteration. The model will repeat Steps 1 and 2 for customers of newly removed agent(s) in a new iteration of the same simulation (all candidate lists and request lists will be cleared for the new iteration). When receiving a new request, those who have reached their upper limit U_k in previous iterations will no longer accept it, until a new opening becomes available after one of their customers gets removed. A simulation will stop when the simulation has finished its 13th iteration. 13 is the diameter of the supply chain network. Therefore, 13 iterations should be long enough in most simulations to spread the effect of the initial disruption to all nodes in the network[4].

In all, we try two different settings for each of the three parameters in the model: the preference to existing partners when approaching alternative suppliers, the probability to remove a node from the supply network, and the extra capacity for a firm to accept new requests. That leads to 2^3=8 different settings for our simulation (listed in Table 3).

5. Results

The ABM is developed using Python and simulations are run on a high-performance computing cluster with each of the 8 settings in Table 3 for the following analyses: (1) showing the impact of disruptions with and without adopting reactive strategies, (2) evaluating the effectiveness of proactive strategies for firms to improve their resilience against an ongoing distant disruptions, and (3) analyzing factors related to the performance of proactive strategies. In each simulation, we simulate the impact of one node removal (i.e., the initial disruption), although more nodes may be removed due to cascading failures.

5.1. The impact of high and low-degree disruptions.

This set of experiments compares the effects of disruptions to high-degree and low-degree firms by simulating the removal of high or low total degree nodes from the supply chain network.

---

[4] Simulation results also show that most of the disruptions occur during the first four iterations. After the 13th iteration, the number of newly disrupted firms only increases by less than 2% for high-degree initial disruptions. Thus we believe that extending the maximum value of iterations beyond 13 will have minimal impact on the simulation results.

Specifically, we rank each node based on their degrees in a descending order. Those ranked within top 10% nodes are considered as high degree nodes (with degrees ranging from 15 to 312), whereas the low-degree nodes are those with degree 1, 40% out of 2,971 firms in the network. For each simulation setting in Table 3 and each type of node removal (i.e., high and low degree), repetitive simulations are conducted 1,000 times. Overall, we run 8*2*1000=16,000 simulations for this experiment.

Table 3. Different simulation settings.

Setting	Preference to existing partners as alternative suppliers	Extra capacity a firm can accommodate	The probability of removing a disrupted node.
1	k~N(1.5, 0.1)	Normal distribution based on firm size.	$N(P_i^R, 0.1)$
2	k~N(1.5, 0.1)	Normal distribution based on firm size.	$Unif(P_i^R - 0.1, P_i^R + 0.1)$
3	k~N(1.5, 0.1)	$Unif(\frac{Z_i'}{2} - 2, \frac{Z_i'}{2} + 2)$	$N(P_i^R, 0.1)$
4	k~N(1.5, 0.1)	$Unif(\frac{Z_i'}{2} - 2, \frac{Z_i'}{2} + 2)$	$Unif(P_i^R - 0.1, P_i^R + 0.1)$
5	k~N(1.5, 0.2)	Normal distribution based on firm size.	$N(P_i^R, 0.1)$
6	k~N(1.5, 0.2)	Normal distribution based on firm size.	$Unif(P_i^R - 0.1, P_i^R + 0.1)$
7	k~N(1.5, 0.2)	$Unif(\frac{Z_i'}{2} - 2, \frac{Z_i'}{2} + 2)$	$N(P_i^R, 0.1)$
8	k~N(1.5, 0.2)	$Unif(\frac{Z_i'}{2} - 2, \frac{Z_i'}{2} + 2)$	$Unif(P_i^R - 0.1, P_i^R + 0.1)$

As mentioned earlier, after an initial node removal, other nodes can be in one of the following states at a given time: undisrupted, disrupted, and removed. Undisrupted nodes operate as usual, although they may have to request alternative suppliers or accept new customers. Undisrupted nodes become disrupted when they fail to secure alternative suppliers to replace original suppliers that are removed from the network. A disrupted firm can be removed from the network with a probability that takes into consideration both the size of the firm and the percentage of lost suppliers (as described in Step 1.3 and Equation 1). We measure the impact of an initial disruption by the total number of disrupted firms, because the operations of these firms are negatively affected by the disruption. While the number of removed firms can also indicate the severity of a disruption, the removal of a firm indicates the cease of operation for the whole company, which is rare in reality.

Figure 5 (top) shows that an initial node removal can indeed disrupt multiple other firms. As we would expect, high-degree node removal in such a network with highly skewed degree distributions cause more firms to be disrupted and are more damaging to the whole network than low-degree node removals. Simulation results are consistent no matter which simulation setting is used: Pearson correlation coefficients among the average numbers of disrupted firms from simulations with the 8 different settings for simulations range from 0.86 to 1.00 (all with p-value<0.001).

As a comparison, we also showed results from another set of 16,000 simulations, which have the same settings, except that agents in this new set of simulations will not adapt to a disruption by reactively connecting to alternative suppliers (results in Figure 5 bottom). Specifically, the new sets of simulations ignore Step 1.1, Step 1.2 and Step 2 in our original ABM, and only keep the consecutive removal of nodes to model cascading failures. Comparing simulation results where agents have reactive behaviors versus not, we can see that ignoring agent's adaptive behaviors from the model greatly increases the impact of the initial disruption. With a reactive strategy, the maximum numbers of disrupted firms are lower than 30, whereas the minimum numbers of disrupted firms after an initial high-degree node removal are higher than 1,400 without a reactive strategy. In other words, the inclusion of an agent's reactive strategy in face of disruptions into our ABM plays an important role in reducing the negative impact of supply chain disruptions. Without considering such adaptive behaviors, many previous studies of complex network resilience may have over-estimated the negative impact of a disruption.

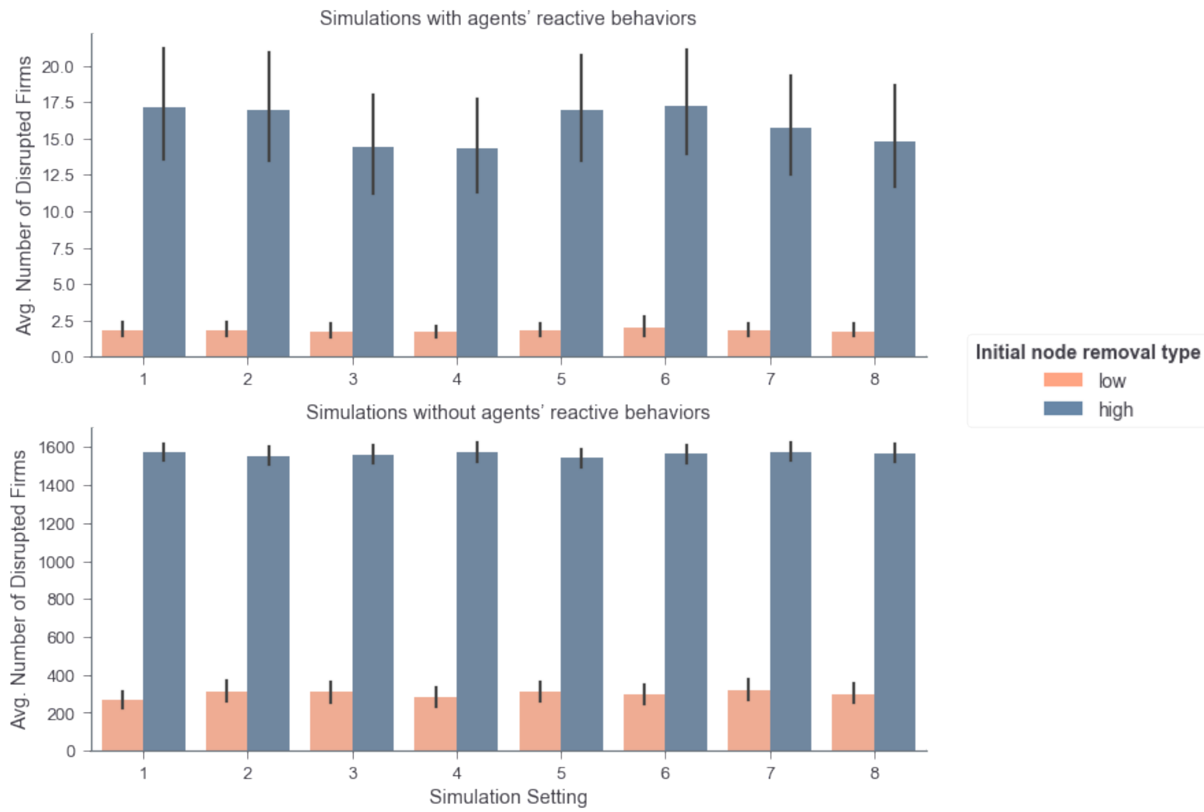

Figure 5. Total numbers of disrupted firms in the supply chain. Results for high and low degree node removal with 8 different simulation settings.

5.2. Proactive strategies after distant disruptions.

Although our ABM incorporates firms' adaptive behaviors in seeking alternative suppliers when their original suppliers fail, a reactive strategy only occurs when a firm is forced to deal with the impact of a failed tier-1 supplier. In many cases, a firm may want to be more proactive in preparing for the impact of a disruption at a distant firm in the supply network (i.e., at least two hops away) and even before any of its immediate suppliers fails. Intuitively, more proactive identifications of which supplier is the riskiest spot given a distant disruption allows a focal firm to prepare an alternative supplier to the riskiest spot. By doing so, the focal firm gains two advantages: First, it faces less competition for alternative suppliers with other firms that only react to the failure of a direct supplier. Second, because such proactive behavior occurs before the distant disruption actually hits the focal firm, it has more time to streamline supplies from the alternative supplier than when it only adopts the reactive strategy after a direct supplier has failed.

Therefore, we propose and compare two proactive strategies for a firm to reduce its supply network risk after a distant disruption is observed. Using 181 firms from 4 sectors[5] (referred to as the focal firm v_f), we simulate the removal of high-degree distant firms that are at least 2 hops away from v_f and compare the probabilities that they are disrupted with and without using proactive strategies.

Based on which distant firm is removed, our proactive strategies identify the riskiest spot among the focal firm's suppliers corresponding to the distant removal. Then the proactive strategies add a new supplier as a backup or alternative for the riskiest spot. The first proactive strategy we propose (S1) identifies riskiest spots based on simulation results. Given a focal firm v_f, for each high-degree firms that are at least two hops away from the focal firm, we remove the high-degree firm, and observe if the initial node removal will eventually disrupt v_f in 100 simulations. If v_f is disrupted, it must have one or more suppliers that are removed during the propagation of the impact from the initial removal. With the simulation results, for each pair of focal firm v_f and the removal of distant node v_r, we can obtain a Disruption Causing Probability (DCP) distribution over the suppliers of v_f: $< p^r_{f,1}, p^r_{f,2}, \ldots, p^r_{f,i}, \ldots, p^r_{f,|S_f|} >$, where S_f is the set of tier-1 suppliers for v_f, and $p^r_{f,i}$ represents the probability that the initial removal of distant node v_r

[5] The four sectors are Computer Hardware & Equipment, Aerospace, Beverage and Food, and Retail (general retail and specialty retail). The four sectors are the most represented among top 27 firms in Gartner's supply chain performance ranking in 2015 (Gartner, 2015).

will disrupt v_f by removing $v_i \in S_f$. Then naturally, supplier v_w with $p^r_{f,w} = argmax_{v_i \in S_f}(p^r_{f,1}$, $p^r_{f,2}$,..., $p^r_{f,|S_f|}$) is the riskiest spot among suppliers of v_f if v_r is removed. When more than one supplier has the same maximum DCP, strategy S1 randomly selects one of them as the riskiest spot.

Proactive strategy S1 requires a large number of simulations to obtain DCP for each pair of focal firm and distant node removal. Therefore, we also propose a heuristic proactive strategy (S2) to help managers approximate the riskiest spots without running many simulations. S2 is developed based on a heuristic measure called Disruption Risk Score (DRS), which considers firm sizes and topologies of both the supply and competition networks. Specifically, once a distant node v_r is removed from the supply chain network, S2 will evaluate focal firm $v_f's$ suppliers $v_i \in S_f$ based on their sizes, degrees in the competition network (i.e. the number of competitors), and Node-to-Node (N2N) betweenness in the supply chain network. Built on the concept of network betweenness, which measures the probability that a node appears on the shortest paths between all possible pairs of nodes, N2N betweenness of a node is the probability for the node to appear on shortest paths between two given nodes. Take Boeing as an example: the shortest path length from firm A to Boeing is 4, and there are 5 such paths with this length. Among Boeing's Tier-1 suppliers, firm B is on 3 of the 5 shortest paths, and firm C is on 2. Then for Boeing and firm A, $N2N$ (A, Boeing, B) $= 3/5$ and $N2N$(A,Boeing, C) $= 2/5$.

Specifically, the DRS of a supplier $v_i \in S_f$ to focal firm v_f after the removal of distant node v_r is defined in Equation 2, where P_i is the set of v_i's competitors in the competition network. $|P_i|$ is v_i's degree centrality in the competition network; Z_i is the firm size of v_i; and $N2N(v_r$,v_f, $v_i)$ is the Node-to-Node betweenness of v_i on paths from v_r to v_f. The logic of using the DRS measure is that among a focal firm's direct suppliers, those who have fewer competitors (i.e., lower degree in the competition network), smaller sizes, and higher probabilities to be on the shortest path between the removed distant node to the focal firm (i.e., higher N2N betweenness in the supply network) are riskier for the focal firm.

$$DRS_{f,i}^{r} = \frac{N2N(v_r, v_f, v_i)}{|P_i| \times Z_i} \qquad \text{(Equation 2)}$$

After the removal of distant node v_r, proactive strategy S2 identifies a firm $v_w \in S_f$ as the riskiest spot for focal firm v_f when $v_w = argmax_{v_i \in S_f} DRS_{f,i}^{r}$. Similar to S1, ties are broken randomly. In other words, riskier firms have higher probabilities to cause an impact to the focal firm, because (1) they have relatively fewer backup or replacement options, (2) they are more susceptible to disruptions due to their smaller sizes, and (3) they are more likely to spread the disruption to the focal firm.

After finding the riskiest spot v_w, both strategies randomly pick one of v_w's competitors, which is not a current Tier-1 supplier of the focal firm, $v_k \in P_w \cap \overline{S_f}$, and add a directed link from v_k to the focal firm v_f. If v_w has no competitors ($|P_w| = 0$) or all of its competitors are already the focal firm's suppliers ($P_w \subseteq S_f$), we exclude v_w from the candidate list and move to the 2nd riskiest spot. For both strategies, we simply add one new supplier for the focal firm right after a distant node removal is observed, assuming the new supplier will accept the request. We make such an assumption for two reasons: First, doing so can simplify the evaluation of proactive strategies. By contrast, allowing a new supplier to reject a proactive request will prolong the simulation and could lead to no new tie formed during proactive restructuring, which makes our evaluation difficult. Second, because such a connection is built by a focal firm proactively to avoid possible disruptions, the focal firm's urgency for supplies is lower than in reactive strategies. Therefore, compared to handling urgent requests sent via reactive strategies, an alternative supplier has a higher probability to adjust its capacities to accommodate such a proactive request.

To evaluate the two proactive strategies, we also add our original reactive strategy from our ABM (described in Section 4) as a baseline adaptive approach for comparison where the focal firm will passively wait till one of its tier-1 suppliers gets removed, and then try to find alternative supplier(s). In the experiments, we only simulate the removal of high-degree nodes, whose degrees rank within top 10% in the whole network, as removing these nodes is the most damaging. The removal is also limited to distant nodes that are at least two hops away from a focal firm, because proactive strategies take place when a focal firm's tier-1 suppliers are not yet affected. After the initial removal of high-degree distant node, we compare the probabilities that focal firms get disrupted when no proactive strategy is used (baseline), proactive strategy S1 is used, and proactive strategy S2 is used. We run 500 simulations for each pair of distant high-degree node removal and focal firm.

Figure 6 uses a scatter plot to show the decrease in focal firms' probabilities of being disrupted after using the two proactive strategies S1 and S2. A positive decrease means a proactive strategy helps to improve a focal firm's resilience against high-degree distant node removals. Compared to the baseline with only reactive behaviors, proactive strategies can reduce the probabilities of disruptions for 150 (82.87%) of the 181 focal firms in our simulation. The maximum decrease is 0.36 for Inventec Corp with S2.

Meanwhile, the performance of the two strategies is highly correlated (r=0.99, p-value<0.0001), with S1 performing slightly better. For example, both strategies lead to similar average disruption probabilities: 3.3% for both S1 and S2, whereas such probability is 7.8% on average without proactive strategies. Besides 12 firms where both strategies have the same performance, S1 outperforms S2 for 90 focal firms, while S2 performs better than S1 for 79 focal firms. At the same time, the performance of S2 compared to S1 also illustrates the effectiveness of our heuristic measure DRS for riskiest spot identifications without running a large number of simulations.

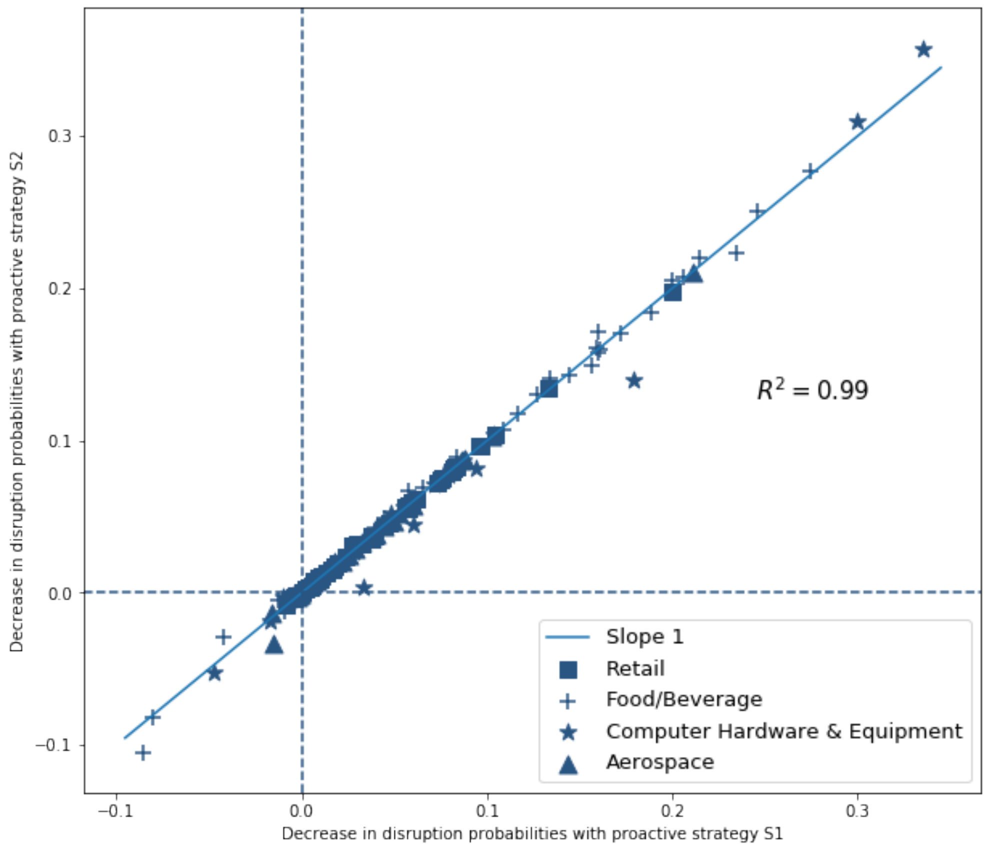

Figure 6. Decrease in focal firms' disruption probabilities for 181 focal firms with proactive strategies S1 and S2.

5.3. Factors impacting the effectiveness of proactive strategies

As Figure 6 shows, the effectiveness of proactive strategies varies from one focal firm to another. A better understanding of which factors make such strategies more or less effective can help a firm better decide if it should adopt proactive strategies. Therefore, we hypothesize two factors that can influence the effectiveness of proactive strategies for a focal firm and run OLS regressions to evaluate the impact of these factors. Because the two proactive strategies have very similar performance with S1 slightly outperforming S2, our analysis on the effectiveness of proactive strategies focuses on S1.

5.3.1. Unevenness of Risk among Suppliers in the Supply Chain Network

With proactive strategy S1, we can obtain the distribution of average risk from all suppliers to a focal firm by averaging a focal firm's DCP for each distant node removal. We hypothesize that the more even the risk distribution is, the less effective a proactive strategy becomes. This is because a more uneven risk distribution means some suppliers are much riskier than others for the focal firm, which can then use proactive strategies to address such vulnerability. This indeed reflects the reality of today's supply chain networks where some suppliers are riskier partners or more likely to cause a disruption. Each supplier has characteristics or dynamic factors that impact riskiness or resilience to disruptions (Blackhurst et al., 2011; Ho et al., 2015). This results in different suppliers having different levels of risk to a focal firm. Hence, risk exposure is uneven across the network. In fact, recent research has noted the lack of research monitoring and understanding supplier risk levels and its impact on the network (Ho et al., 2015).

We will use two extreme examples to illustrate the idea behind this hypothesis related to the unevenness of risk amongst supplier. When the risk distribution follows a uniform distribution, every supplier shares the same probabilities of disrupting the focal firm, but a proactive approach is limited to taking care of one of these firms. On the other end of the spectrum, if the risk distribution follows a Dirac Delta Distribution with one firm having a probability of 1 and the others being 0, it is obvious which supplier is the riskiest spot. After a proactive approach handles the riskiest spot, other suppliers have no chance to disrupt the focal firm anymore. To measure the unevenness of the risk distribution, we calculate Gini coefficients (Gini, 1912) of the average DCP distribution for S1. Higher Gini coefficients mean more unevenly distributed risk among a focal firm's suppliers. Therefore, we hypothesize:

H1: *In the presence of a remote supply chain disruption, the effectiveness of a proactive strategy is positively associated with the unevenness of risk among a focal firm's suppliers.*

In our models, the baseline disruption probability (BDP) is the probability of a focal firm being disrupted in our baseline setting without using a proactive strategy. Intuitively, such a probability will impact the effectiveness of a proactive strategy. In other words, a focal firm starting with a higher risk of disruption can benefit more from being proactive in approaching others, while a proactive strategy will not help as much when a focal firm already has lower disruption probabilities. Therefore, we also hypothesize:

> ***H1a****: The positive effect of risk unevenness on the effectiveness of proactive strategies is moderated by BDP.*

We contend that when BDP is higher, the positive effect of risk unevenness becomes stronger. This is because when a focal firm is more susceptible to remote disruptions, the overall risk caused by the focal firm's suppliers is higher. Meanwhile, given the same unevenly distributed DCP, with a higher overall risk of disruptions, the potential risk from the riskiest spot of the focal firm will increase. Therefore, proactive strategies can be more effective after finding an alternative to the riskiest spot.

5.3.2. Multi-Sourcing Ratio among Suppliers in the Supply Chain Network

A common practice for a firm to improve its supply chain network resilience against disruptions is to add back-up suppliers by procuring the same product from more than one supplier (Sawik, 2014a; Sawik, 2014b). Multi-sourcing approach can also be leveraged to maintain competitiveness amongst suppliers (Heese, 2015). We have seen an increasing use of competing suppliers in a multi-sourcing strategy in real world supply chains. For example, Apple sources displays from multiple suppliers which maintains competitiveness between suppliers and reduce the risk of supply disruptions (Li and Debo, 2009; Hu et al., 2017). In this paper, we measure the level of multi-sourcing (using the supply chain network and its corresponding competition network) with a new network-based measure called multi-sourcing ratio (MR). To calculate MR, we examine all suppliers of a focal firm, and find the percentage of suppliers that also compete with another supplier of the same focal firm. A higher ratio means a higher level of multi-sourcing for a focal firm and more competition amongst its suppliers. For example, firm A has 4 suppliers B, C, D, and E. Among the 4 suppliers, B competes with C (where B and C are multiple sources to the focal firm A and they are connected in the competition network), and C competes with E. Then the MR for firm A is 75%, because 3 out of 4 suppliers for A have competitor(s) among A's suppliers.

If a focal firm has a supplier and the competitors of that supplier are also already serving the focal firm (in a multi sourcing situation), then the proactive strategy will be less effective. This is because the proactive strategy works in the following way: for the riskiest supplier of a focal firm (depending on which remote firm is disrupted), proactive strategies pick one firm from the riskiest supplier's competitors and add the firm as a supplier. If a competitor of the riskiest supplier is already a supplier of a focal firm, then adding another competitor of the riskiest supplier as a supplier becomes redundant.

Back to the example of Apple's display suppliers. Assume A is a display supplier for Apple and is identified as the riskiest spot after a remote disruption. If Apple only uses supplier A as its display supplier, then adding supplier B (which is a competitor to supplier A) as another supplier (multi-sourcing) during proactive restructuring can be effective. However, if Apple is already buying displays from both supplier A and supplier B, Apple already has a back-up supplier in place in the event of supplier A failing. In this case, adding another display manufacturer C using proactive strategies may still help, but the improvement will be less than in the case where there was no backup supplier in place. Essentially, when a focal firm has no backup to its riskiest supplier, proactive strategies help more. If a focal firm already has backup to its riskiest supplier (via multi-sourcing), then adding another backup via proactive strategies is less effective. Therefore, we hypothesize:

> ***H2**: In the presence of a remote supply chain disruption, the effectiveness of a proactive strategy is negatively associated with a focal firm's multi-sourcing ratio.*

5.3.3. Models and results

To test our hypotheses, we run a multiple regression model on simulation results of proactive strategy S1 for the 181 firms in Section 5.2. The dependent variable (DV), $\Delta DisProb$, is the decrease in focal firm's disruption probability after restructuring with strategy S1, compared to baseline with no proactive restructuring. In other words, the DV shows how much the disruption probability decreases after a focal firm uses proactive strategy S1.

As for covariates, control variables include a focal firm's sector (3 dummy variables for 4 sectors), its out-degree centrality in the supply chain network[6] (i.e., number of customers, *OutDgrSupply*), its degree centrality in the competition network (i.e., number of competitors, *DgrComp*), and its BDP. The two independent variables are (1) the focal firm's Gini coefficient of average DCP (*Risk_Gini*), and (2) the focal firm's multi-sourcing ratio (*MR*). We also added an interaction term of *Risk_Gini*BDP* to test Hypothesis 1a. The full model is specified in Equation 3. There is no strong correlation between any pair of covariates (Figure 7). Because this is a linear regression model, we tested assumptions for such a model and included results in Appendix 3. Note that all covariates except dummies for sectors are log-transformed to address assumptions of linear regression models.

$$\Delta DisProb_i = \beta_0 + \beta_1 * Sector_i + \beta_2 * OutDgrSupply_i + \beta_3 * DgrComp_i + \beta_4 * BDP_i + \beta_5 * Risk_{Gini_i} + \beta_6 * MR_i + \beta_7 * Risk_{Gini_i} * BDP_i + \epsilon_i,\ i = 1,2,\ldots,181. \qquad \text{(Equation 3)}$$

[6] In-degree centrality in the supply chain network is not included, because it is highly correlated with many other covariates (e.g., 0.68 with MR and 0.65 with Risk.)

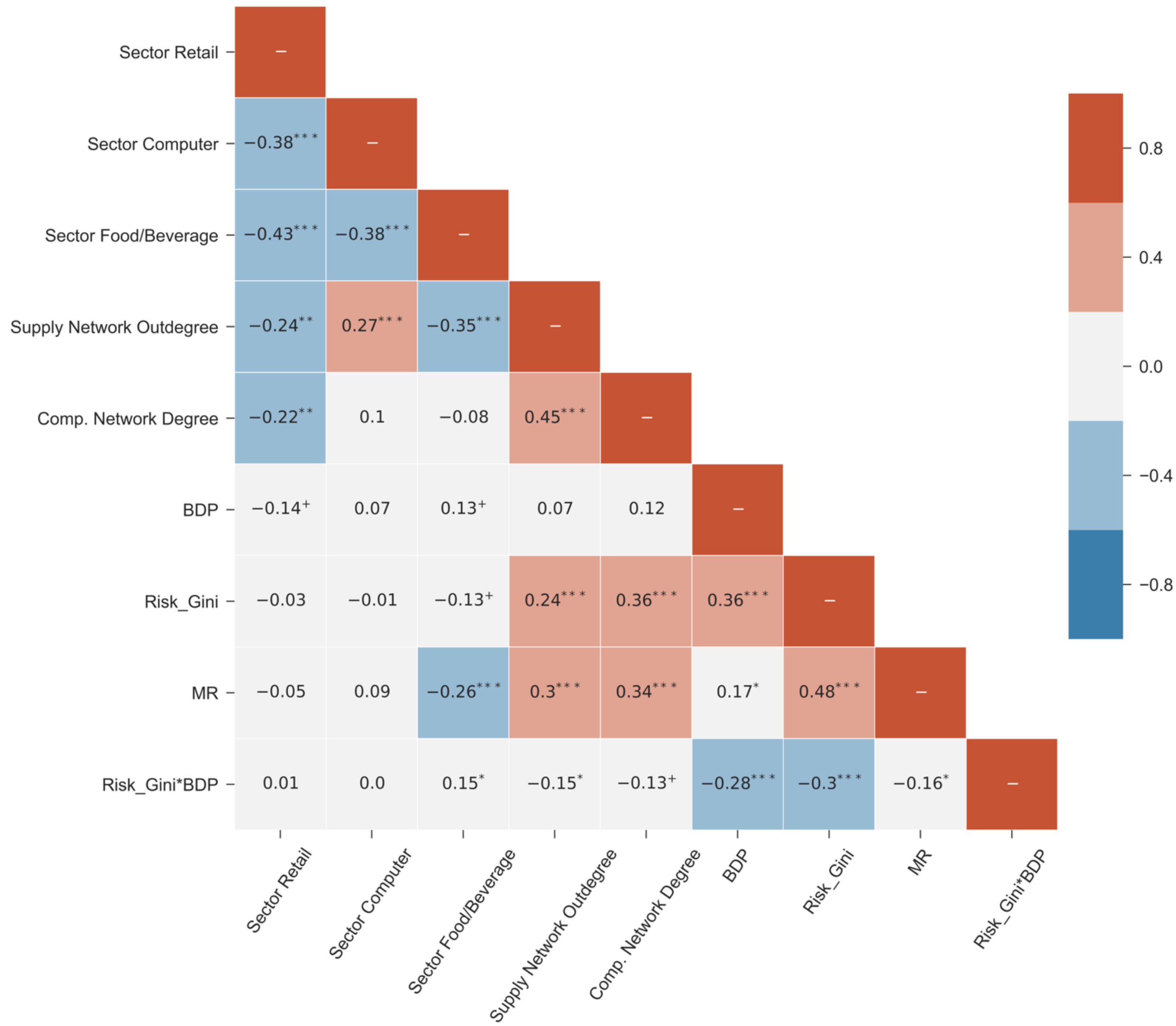

Figure 7. Pair-wise Pearson correlation between covariates:
$*$:p < 0.05; $**$:p < 0.01; $***$:p < 0.001

Table 4 summarizes results of our regression models, each with a different set of covariates. Confidence intervals and statistical significance are based on robust standard errors (Arellano, 1987). First, among control variables, only BDP is a significant predictor for the effectiveness of proactive strategies. As expected, its sign is consistently positive confirming that firms that suffer from higher disruption probabilities without proactive restructuring benefit more from the proactive strategy. Second, Hypotheses 1 and 2 are supported. *Risk_Gini* is a positive and significant predictor, while *MR* is a negative and significant predictor of the DV. In other words, a firm with more unevenly distributed risk among its suppliers and a lower ratio of supply multi-sourcing would benefit more from a proactive strategy. Last, Hypothesis 1a is also supported as the interaction term *Gini_Risk*BDP* has a positive and significant coefficient (Model 3).

Table 4. OLS regression model for the effectiveness of proactive strategy S1.

	Variable	Coefficient (Std. Err.)		
		Model 1	Model 2	Model 3
Control Variables	Sector Retail	-0.05 (0.18)	-0.09 (0.20)	-0.17 (0.21)
	Sector Computer	-0.05 (0.19)	-0.03 (0.20)	-0.12 (0.21)
	Sector Food/Beverage	0.18 (0.23)	0.14 (0.25)	0.01 (0.26)
	Supply Network Outdegree	-0.02 (0.08)	-0.02 (0.08)	-0.03 (0.08)
	Competition Network Degree	-0.06 (0.07)	-0.07 (0.07)	-0.07 (0.06)
	BDP	0.62*** (0.08)	0.58*** (0.09)	0.62*** (0.10)
Independent variables	Risk_Gini	-	0.19* (0.08)	0.22* (0.09)
	MR	-	-0.18** (0.06)	-0.19** (0.06)
	Risk_Gini*BDP	-	-	0.16* (0.08)
Adjusted R^2		0.38	0.41	0.43
F		11.38***	10.28***	13.98***
$+: p < 0.1$; $*: p < 0.05$; $**: p < 0.01$; $***: p < 0.001$				

6. Discussion

In this section, we discuss implications for researchers and practitioners. Future research directions extending from this study as well as limitations are also presented.

6.1. Implications for Theory and Practice

This paper builds upon recent work analyzing supply chain networks as complex adaptive systems. We have developed a method to model and understand strategies for supply chain adaptation in the face of disruptions. Not only does recent research tell us that the network level implication of cascading disruption is difficult to understand (Fiksel et al., 2015), but there is a need to reconfigure and restructure supply chain networks in response to disruptions (Hearnshaw et al., 2013). In fact, the more that can be done to identify and “shore up” risky spots in the supply chain, the better the supply chain performance (Blackhurst et al., 2018). As such, we have answer calls to build adaptive capabilities into a model for supply chain disruptions (Hearnshaw et al., 2013; Kim et al., 2015; Van der Vegt et al., 2015) by leveraging inter-firm competition relationships. We model the supply chain network as a CAS (Choi et al., 2001; Nair et. al., 2009) and leverage the adaptive capabilities to reconfigure connections and structures (Anderson, 1999; Choi et al., 2001). From a CAS perspective, we can model both close and distant disruptions and examine strategies to mitigate their impacts. This is important as disruptions may occur outside of the direct purview of a focal firm and propagate to it with intensifying and devastating effects. Our approach of modeling supply chain networks as an agent-based system examines two types of strategies: reactive and proactive to determine how firms can leverage such adaptive strategies to improve their resilience against supply chain network disruptions.

We present our results in three stages. First, we model and analyze the impact of disruptions on a real-world large-scale supply network and demonstrate the use and effectiveness of reactive strategies. Next, we develop and evaluate the effectiveness of proactive strategies for firms to improve their resilience against a distant disruption. Third, we analyze factors related to the performance of proactive strategies. Our results have implications to both managers and researchers alike. In our first stage, we measure the impact of reactive strategies where disruptions occurred at high-degree and low-degree nodes. Not surprisingly, but now empirically validated through this study, we show that high-degree node removal is more damaging to the supply network as compared to low-degree node removal. We also illustrate the spread of a supply disruption with and without a reactive strategy. A reactive strategy was shown to reduce the number of nodes impacted by almost 50-fold (from over 1,400 to less than 30). This provides a baseline for our model and also demonstrates the importance of considering firms' adaptive behaviors when evaluating a supply chain network's performance against disruptions.

However, based on recent calls in the research, such as Blackhurst et al. (2018), to focus more on proactively managing risk, we develop and model proactive strategies. Such strategies are used when a disruption occurs at a distant firm (beyond first tier) but has not yet impacted the focal firm. Proactive strategies identify the weakest spot specific to the disrupted distant firm in the network. A replacement supplier is identified from the competition network and the supply network is restructured in a proactive manner. The first proactive strategy (S1) requires a large number of simulations, while the second proactive strategy (S2) uses a heuristic approach to identify the riskiest spots in the supply network. We empirically model and validate the superiority of proactive strategies over reactive strategies. We illustrate that even though a disruption may not originate from a focal firm's immediate neighbors, it can propagate to the focal firm (Kim et al., 2015; Blackhurst et al., 2005). Our strategies demonstrate a way to mitigate the impact of these potential damaging disruptions and develop resilience to allow firms to continue adding value to customers as called for by Ambulkar et al. (2015).

Next, in order to better understand factors impacting the effectiveness of proactive strategies, we ran regression analyses. We proposed that two factors specifically could affect the effectiveness of proactive strategies: the evenness of risk among suppliers of the focal firm and the ratio of multi-sourcing among suppliers of the focal firm. We find that with higher evenness of risk, proactive strategies become more effective. However, the more multi-sourcing exists in the supply base, proactive strategies become less effective. Such findings greatly improve the practical value of our proactive strategies because they can better inform managers on whether their firms should adopt the proactive strategy when a distant disruption occurs. This is important for managers, because they are charged with the delicate balancing act of risk versus reward in the supply chain. The findings can also help managers address the question of where and how to invest valuable and limited resources (Chopra and Sodhi, 2004; Tomlin, 2006; Chopra and Sodhi, 2014).

6.2. Directions for Future Research and Limitations

There are also many exciting possibilities to extend this research. For example, in our agent-based models, a firms' decision-making strategies, especially how an alternative supplier decides which request to accept, can incorporate more factors, such as geographical proximity, contract negotiation, the competition relationship between the requester and the alternative supplier. The way we design alternative suppliers' upper limit of accommodating new requests can be improved to be more realistic as well. In addition, this research may have interesting implications for supply base management policies. For example, Shao (2017) notes that a supplier may (on its own) subcontract out to competition in order to win a bid. In this case, the focal firm views the supplier as a single source but that supplier may be subcontracting out to other suppliers. In this case the multi-sourcing nature of the supply base in not controlled by the focal firm but rather a supplier. Such nuances would be interesting extensions.

Also, our model focuses on how a disruption propagates from disrupted suppliers to their customers. However, losing a customer may negatively affect a supplier's supply chain operation as well. Adding such upstream propagation to our model will be helpful to better capture the impact of disruptions. In addition, a model that considers the recovery of a firm after it is removed from the supply chain network will be interesting, although our current model does not incorporate such recovery due to increased complexity. Comparing to removing one firm and examining how the disruption propagates, removing multiple firms simultaneously at the beginning of the simulation can also be valuable, because this can reflect to disruptions caused by disasters in a larger geographical region or political/military turmoil in a country. Alternatively, we can also specify the type of disruptions for the initial node removal because disruptions caused by exogenous shocks, such as natural disasters and industry-wide decline, and endogenous processes, such as competitive dynamics, may affect other firms in the network in different ways.

In terms of adaptive strategies, the reactive and proactive strategies we investigated in this paper restructure a focal firm's network after a disruption occurs among the focal firm's immediate suppliers or distant nodes respectively. It would be interesting to develop a preemptive strategy, which guides a focal firm to more strategically restructure its networks for possible disruptions in the future. Last but not least, analyzing how large-scale supply chain networks and the competition networks co-evolve over time would also be an interesting undertaking, because we may be able to identify major disruptions that actually happen, and track how firms react to these disruptions from real-world data over time. Such longitudinal data can potentially help us validate our model of firms' adaptive behaviors.

Finally, this study is not without limitations. First, our supply chain network model is constructed from Mergent Horizon. The data is verified to be accurate and enables us to build a large-scale multi-tier supply chain network along with a competition network that adds another layer of relationship among these firms. However, this dataset may not capture all the relationships and entities in the network. For example, Boeing's customers include government agencies, which may also be customers to Microsoft and thus constitute hidden connections between Boeing and Microsoft. As we focus on firms only, government agencies are not included in our dataset, although a political or military event can disrupt such entities and may affect supply chain operations of Boeing and Microsoft. Also, our sampling of the dataset relies on snow-ball sampling starting from one seed node. This method helps to yield a multi-sector multi-country supply networks, but it may also introduce bias as more connected firms are more likely to be sampled. We try another sample with 27 seed nodes and find that this sample yields a network similar to the one used in this study (see Appendix 1 for details). However, we still acknowledge that potential bias may exist in the supply network used in this study. Second, while our verified ABM is based on a decision-making logic that reflects our knowledge of real-world supply chain operations, it is extremely difficult to obtain empirical data on firms' moves after supply chain disruptions to validate such a schema. That is also a common challenge for many ABMs, and we hope we can address this issue in the future. Third, our model focuses on firms' short-term reactions to a disruption, because we remove a firm from the supply network and does not consider if and when the firm will come back to normal operations. To include med- or long-term reactions to a disruption, the model will have to incorporate a node availability check component, as well as how firms deal with alternative suppliers when their original suppliers resume normal operations. Last, when evaluating the effectiveness of different adaptive strategies, we do not consider the cost of adding new suppliers. Therefore, our adaptive strategies provide the best-case scenario. Firms need to consider costs of such strategies when deciding which alternative supplier(s) to approach.

References

Agarwal, N. C., 1979. On the interchangeability of size measures. The Academy of Management Journal 22 (2), 404–409.

Albert, R., Barabasi, A.L., 2002. Statistical mechanics of complex networks. Reviews of Modern Physics 74, 47–97.

Ambulkar, S., Blackhurst, J. Grawe, S., 2015. Firm Resilience to Supply Chain Disruptions: Scale Development and Empirical Examination. Journal of Operations Management 33-34, 111-122.

Anderson, P., 1999. Complexity theory and organization science. Organization Science 10 (3), 216-232.

Arellano, M., 1987. Computing Robust Standard Errors for Within-groups Estimators. Oxford Bulletin of Economics and Statistics 49 (4), 431–434.

Atalay, E., Hortaçsu, A., Roberts, J., Syverson, C., 2009. Network structure of production, Proceedings of the National Academy of Sciences of the Unites States of America (PNAS) 108(13), 5199–5202.

Axelrod, R., 1997. The complexity of cooperation: Agent-based models of competition and collaboration. Princeton University Press, Princeton, NJ.

Axtell, R. L., 2001. Zipf distribution of U.S. firm sizes. Science 293 (5536), 1818–1820.

Azadegan, A., 2011. Benefiting from supplier operational innovativeness: The influence of supplier evaluations and absorptive capacity. Journal of Supply Chain Management 47 (2), 49-64.

Azadegan, A., Patel, P. C., Parida, V., 2013. Operational slack and venture survival. Production and Operations Management 22 (1), 1-18.

Barabasi, A. L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509-512.

Blackhurst, J., Craighead, C. W., Elkins, D., Handfield, R., 2005. An empirically derived agenda of critical research issues for managing supply-chain disruptions. International Journal of Production Research 43 (19), 4067-4081.

Blackhurst, J., Rungtusanatham., M., Scheibe, K., Ambulkar, S., 2018. Supply chain vulnerbaility: A network based visualizationa nd clustering analysis approach. Journal of Purchasing and Supply Management 24 (1), 21-30.

Blackhurst, J., Dunn, K., Craighead, C. W., 2011. An empirically derived framework of global supply resiliency. Journal of Business Logistics 32 (4), 374-391.

Bode, C., Wagner, S. M., Petersen, K. J., Ellram, L. M., 2011. Understanding responses to supply chain disruptions: Insights from information processing and resource dependence perspectives. Academy of Management Journal 54 (4), 833-856.

Bode, C., Wagner, S. M., 2015. Structural drivers of upstream supply chain complexity and the frequency of supply chain disruptions. Journal of Operations Management 36, 215-228.

Brintrup, A., Kito, T., Lopez, E., New, S., Reed-Tsochas, F., 2011. Mapping the Toyota supply network: Implications for resilience. Retrieved from http://www.cabdyn.ox.ac.uk/complexity_PDFs/Working%20Papers%202011/ToyotaRobustness.pdf

Brintrup, A., Ledwoch, A., & Barros, J., 2016. Topological robustness of the global automotive industry. Logistics Research 9 (1), 1. https://doi.org/10.1007/s12159-015-0128-1

Carrington, P. J., Scott, J., Wasserman, S., 2005. Models and methods in social network analysis. Cambridge University Press.
Choi, T.Y., Dooley, K.J., Rungtusanatham, M., 2001. Supply networks and complex adaptive systems: Control versus emergence. Journal of Operations Management 19 (3), 351-366.
Chopra, S., Sodhi, M.S., 2004. Managing Risk to Avoid Supply-Chain Breakdown. MIT Sloan Management Review 46 (1) 53-61.
Chopra, S., Sodhi, M.S., 2014. Reducing the Risk of Supply Chain Disruptions. MIT Sloan Management Review 55 (3), 72-80.
Clauset, A., Shalizi, C.R., Newman, M.E., 2009. Power-law distributions in empirical data. SIAM Review 51 (4), 661-703.
Craighead, C., Blackhurst, J., Rungtusanatham, M., Handfield, R., 2007. The Severity of Supply Chain Disruptions: Design Characteristics and Mitigation Capabilities. Decision Sciences Journal 38 (1), 131-156.
Dwyer, F. R., Schurr, P. H., Oh, S., 1987. Developing buyer-seller relationships. The Journal of Marketing 51 (2), 11-27.
Dyer, J. H., Singh, H., 1998. The relational view: Cooperative strategy and sources of interorganizational competitive advantage. Academy of Management Review 23 (4), 660-679.
Fiksel, J., Polyviou, M., Croxton, K.L., Pettit, T.J., 2015. From risk to resilience: learning to deal with disruption. MIT Sloan Management Review 56 (2), 79-86.
Gini, C., 1912. Variabilità e mutabilità. Reprinted in Pizetti, E.; Salvemini, T., eds. (1955). Memorie di metodologica statistica. Rome: Libreria Eredi Virgilio Veschi.
He, Z., Cheng, T.C.E., Dong J., Wang, S., 2016. Evolutionary location and pricing strategies for service merchants in competitive O2O markets. European Journal of Operational Research 254 (2), 595-609.
Heese, H., S., 2015. Single vs. Multiple Sourcing and The Evolution of Bargaining Positions. Omega 54, 125-133.
Hearnshaw, E. J., Wilson, M.M., 2013. A complex network approach to supply chain network theory. International Journal of Operations & Production Management 33, 442-469.
Hendricks, K., Singhal, V., 2003. The effect of supply chain glitches on shareholder wealth. Journal of Operations Management 21 (5), 501–22.
Hendricks, K., Singhal, V., 2005. Association between supply chain glitches and operating performance. Management Science 51 (5), 695–711.
Ho, W., Zheng, T., Yildiz, H., Talluri, S., 2015. Supply chain risk management: A literature review. International Journal of Production Research 53 (16), 5031-5069.
Hu, S., Wan, Z., Ye, Q., Chi, W., 2017. Supplier behavior in capacity investment competition: An experimental study. Production and Operations Management 26 (2), 273-291.
Kamalahmadi, M., Parast M. M., 2016. A review of the literature on the priciples of enterprise and supply chain resilince: Major findings and directions for future research. International Journal of Production Research 171, 116-133.
Kim, Y., Choi, T. Y., Yan, T., Dooley, K., 2011. Structural investigation of supply networks: A social network analysis approach. Journal of Operations Management 29 (3), 194-211.
Kim, Y., Chen Y.-S., Linderman, K., 2015. Supply network disruption and resilince: A network structural perspective. Journal of Operations Management 33-34, 43-59.

Kimberly, J. R., Evanisko, M. J., 1981. Organizational innovation: The influence of individual, organizational, and contextual factors on hospital adoption of technological and administrative innovations. Academy of Management Journal 24 (4), 689–713.

Kito, T., Brintrup, A., New, S., Reed-Tsochas, F., 2014. The Structure of the Toyota Supply Network: An Empirical Analysis (SSRN Scholarly Paper No. ID 2412512). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2412512

Li, C., Debo, L., 2009. Second Sourcing vs. Sole Sourcing with Capacity Investment and Asymmetric Information. Manufacturing and Service Operations Management 11 (3), 448-470.

Li, S., Madhok, A., Plaschka, G., Verma, R., 2006. Supplier switching inertia and competitive asymmetry: a demand side perspective. Decision Sciences 37 (4), 547-576.

Nair, A., Narasimhan, R., Choi, T., 2009. Supply networks as a complex adaptive systems: Toward simulation-based theory building on evolutionary decision making. Decision Sciences Journal 40 (4), 783-815.

Nair, A., Vidal, J.M., 2011. Supply network topology and robustness against disruptions–an investigation using multi-agent model, International Journal of Production Research 49, 1391-1404.

Newman, M. E. J., 2003. The structure and function of complex networks. SIAM Review 45 (2), 167–256.

Park, K., Min, H., Min, S., 2016. Inter-relationship among risk taking propensity, supply chain security practices and supply chain disruption occurrence. Journal of Purchasing and Supply Management 22, 120-130.

Pathak, S., Day, J., Nair, A., Sawaya, W., Kristal, M., 2007. Complexity and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective. Decision Sciences Journal 38 (4), 547-580.

Saavedra, S., Reed-Tsochas, F., Uzzi, B., 2008. Asymmetric disassembly and robustness in declining networks. Proceedings of the National Academy of Sciences of the United States of America 105, 16466–16471.

Saavedra, S., Stouffer, D.B., Uzzi, B., Bascompte, J., 2011. Strong contributors to network persistence are the most vulnerable to extinction. Nature, 478, 233-235.

Sawik, T., 2014a. Joint Supplier Selection and Scheduling of Customer Orders Under Disruption Risks: Single vs. Dual Sourcing. Omega 43, 83-95.

Sawik, T., 2014b. Optimization of Cost and Service Level in the Presence of Supply Chain Disruption Risks: Single vs. Multiple Sourcing. Computers and Operations Reseach 51, 11-20.

Shao, L., 2018. Competition under diseconomies of scale: The role of subcontracting and single-sourcing commitment. Decision Sciences Journal 49 (1), 154-179.

Staniczenko, P., Lewis, O.T., Jones, N.S. Reed-Tsochas, F., 2010. Structural dynamics and robustness of food webs. Ecology Letters 13 (7), 891-899.

Thadakamalla, H. P., Raghavan, U. N., Kumara, S., Albert, R., 2004. Survivability of multiagent-based supply networks: A topological perspective. IEEE Intelligent Systems 9, 24-31.

Thompson, D., 1967. Organizations in Action. McGraw-Hill, New York.

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

Trkman, P., McCormack, K., 2009. Supply chain risk in turbulent environments--A conceptual model for managing supply chain network risk. International Journal of Production Economics 119 (2), 247-258.
Van der Vegt, G., Essens, P., Wahlstrom, M., George, G., 2015. Managing risk and resilience. Academy of Management Journal 58 (4), 971-980.
Wagner, S. M., Bode, C., 2008. An empirical examination of supply chain performance along several dimensions of risk. Journal of Business Logistics 29, 307–325.
Wagner, S. M., Friedl, G., 2007. Supplier switching decisions. European Journal of Operational Research 183 (2), 700-717.
Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of "small-world" networks. Nature 393, 440–442.
Wilensky, U., Rand, W., 2015. An introduction to agent-based modeling. MIT Press, Cambridge, MA.
Wu, T., Huang, S., Blackhurst, J., Zhang, X. Wang., S., 2013. Supply Chain Risk Management: An Agent-Based Simulation to Study the Impact of Retail Stockouts. IEEE Transactions on Engineering Management 60(4), 676-686.
Yang, Y., Chawla, N., Sun, Y., Han, J. , 2012. Predicting Links in Multi-relational and Heterogeneous Networks. 2012 IEEE 12th International Conference on Data Mining (ICDM), 755–764. https://doi.org/10.1109/ICDM.2012.144
Zhao, K., Kumar, A., Harrison, T. P., Yen, J., 2011a. Analyzing the resilience of complex supply network topologies against random and targeted disruptions. IEEE Systems Journal 5, 28-39.
Zhao, K., Kumar, A., Yen, J., 2011b. Achieving high robustness in supply distribution networks by rewiring. IEEE Transactions on Engineering Management 58, 347-362.
Zhao, K., Scheibe, K., Blackhurst, J., Kumar, A., forthcoming. Supply Chain Network Robustness Against Disruptions: Topological Analysis, Measurement, and Optimization. IEEE Transactions on Engineering Management.
Zhao, K., Wang, X., Cha, S., Cohn, A. M., Papandonatos, G. D., Amato, M. S., Pearson, J. L., Graham, A. L., 2016. A Multirelational Social Network Analysis of an Online Health Community for Smoking Cessation. Journal of Medical Internet Research 18 (8), e233. https://doi.org/10.2196/jmir.5985

Appendix 1. Comparison between two datasets and supply chain networks

This appendix compares the dataset (D_1) and the supply chain network based on it (referred to as G_s) used in the paper with another dataset (D_2) and the supply chain network $G_s{}'$ based on D_2.

D_1 was collected using Boeing as the seed firm. By contrast, D_2 used 27 seed firms from various sectors. According to Gartner's supply chain performance ranking in 2015 (Gartner, 2015), 25 of these 27 firms were ranked with in top 25, and the other two were named "supply chain masters". The 27 firms are listed in Table A1-1. Note that Boeing is not one of the 27 firms. The collection of D_2 followed the same procedure as D_1. Both datasets were collected from the Mergent Horizon database.

Table A1-1. List of seed firms for D_2. ("Supply chain masters" are marked with *)

Amazon.com	The Coca-Cola Company	Johnson & Johnson
McDonald's	Starbucks	L'Oréal
Unilever	Wal-Mart Stores	Cummins
Intel	3M	Toyota Motor
Inditex	PepsiCo	Home Depot
Cisco Systems	Seagate Technology	Apple*
H&M	Nestlé	P&G*
Samsung Electronics	Lenovo	
Colgate-Palmolive	Qualcomm	
Nike	Kimberly-Clark	

Compared to the 2,791 firms in D_1, D_2 has 4,406 firms. However, the sector distributions (i.e., the percentage of firms from each sector) are very similar. The Pearson correlation coefficient between sector distributions in the two datasets is 0.88 (p-value<0.001). When we fit a linear regression between sector percentages from D_2 vs sector percentages from D_1, we got a fitted straight line with a slope of 0.86 and R-squared of 0.78. The results are statistically significant with a p-value < 0.001. This indicates a strong linear and positive relationship between the industry sector distributions in the two datasets.

Using dataset D_2, we also built a supply chain network $G_s{}'$ using the same approach we built G_s with D_1. Figure A1-1 shows the degree distribution $G_s{}'$ in log-log scale. Similar to G_s, the degree distribution is also best represented by a Truncated Power-law: log-likelihood ratios for pairwise comparisons between Power-law (PL), Truncated Power-law (Truncated PL), Exponential (EXP), and Log-Normal (LN) are 1.682 (p-value <0.1) for PL vs Truncated PL, 6.805 (p-value <0.001) for Truncated PL vs EXP, 3.194 (p-value <0.05) for Truncated PL vs LN, 6.388 (p-value <0.001) for PL vs. EXP, and 6.679 (p-value < 0.001) for LN vs. EXP.

$Table$ A1-2 compares the four network statistics of the two networks, namely network density, average path length, network diameter, and average clustering coefficient. First, the two networks share similar densities, even though $G_s{}'$ is much bigger. Second, $G_s{}'$ has slightly longer average shortest path length (a.k.a., characteristics path length), but lower clustering coefficient. This is expected for $G_s{}'$ that has more nodes, because for networks with power-law-like degree distributions, characteristics path length scales with *log(N)/log(log(N))*, and clustering coefficient scales with *log(N)/N,* where *N* is the number of nodes (Thadakamalla, Raghavan, Kumara, & Albert, 2004).

All the comparisons above showed that dataset D_1 and supply chain network G_S used in our study represent a reasonable sample of the Mergent database.

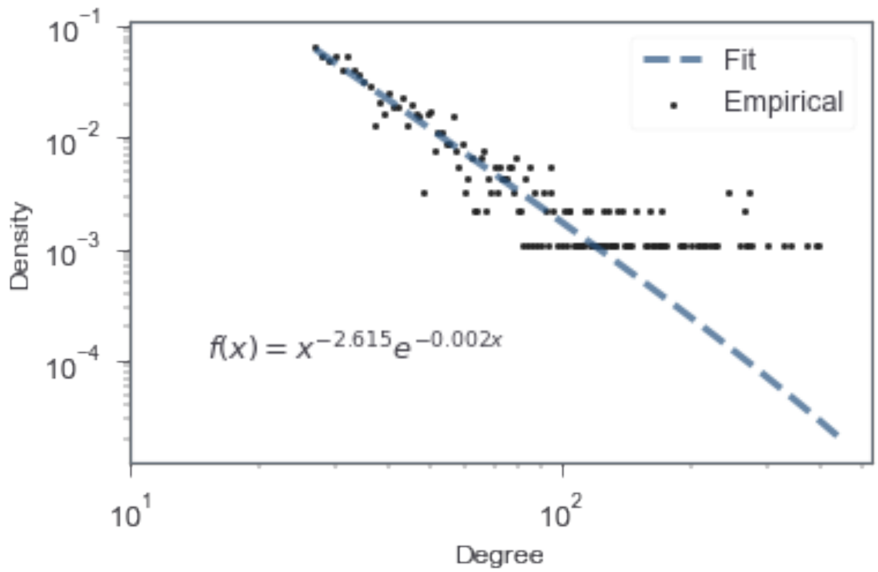

Figure A1-1. Degree distributions of the new supply network $G_S{}'$.

Table A1-2. Network Statistics of G_S and G_S'

	DENSITY	CHAR. PATH LEN.	AVG. CLUSTERING COEF.
G_S	0.001	4.704	0.057
$G_S{}'$	0.001	5.007	0.048

References:

Gartner. (2015). Gartner Supply Chain University Top 25. Retrieved September 26, 2016, from https://www.gartner.com/technology/supply-chain/top25.jsp

Thadakamalla, H. P., Raghavan, U. N., Kumara, S., & Albert, R. (2004). Survivability of Multiagent-Based Supply Networks: A Topological Perspective. IEEE Intelligent Systems, 19, 24–31. https://doi.org/10.1109/MIS.2004.49

Appendix-2. Flow diagrams and pseudo code for the agent-based model

Figure A2-1 shows the flow diagram of our agent-based model. Figure A2-2 summarizes agents' behaviors in the model.

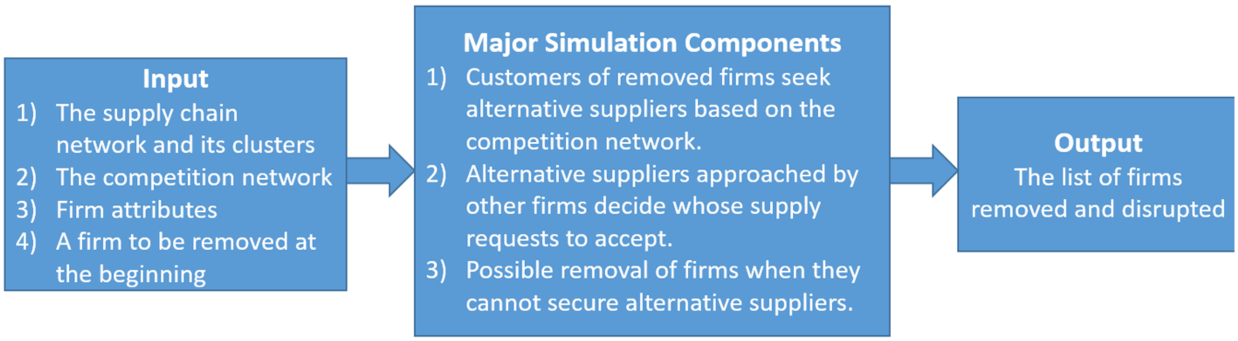

Figure A2-1. The flow diagram of the agent-based model.

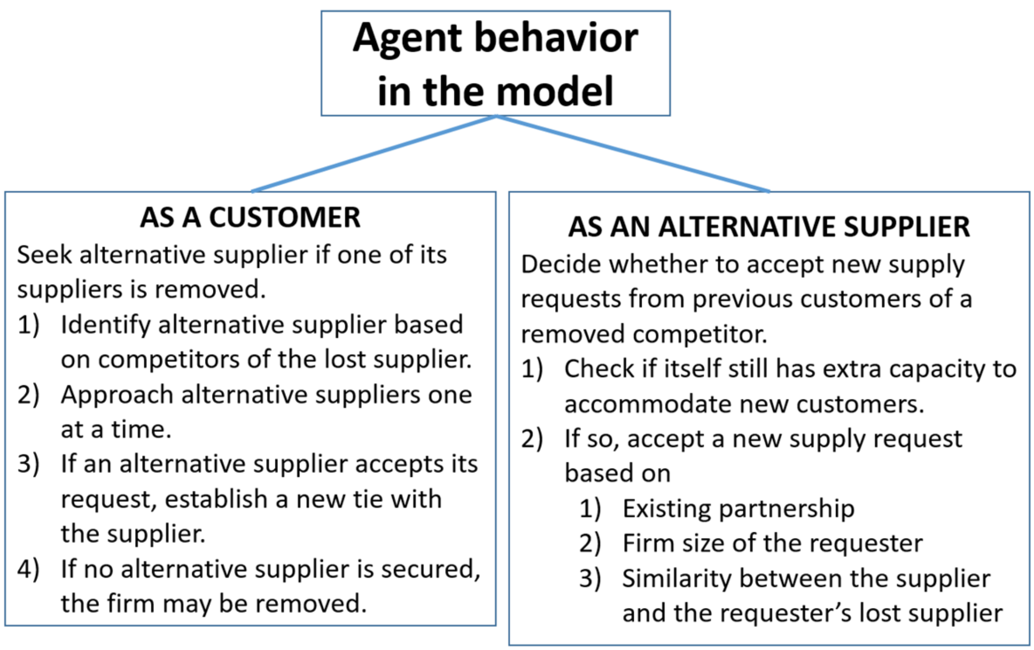

Figure A2-2. A summary of agent behaviors in the model.

The following pseudocode summarizes the major steps of our simulation framework.

VARIABLES :

- $G_s(V, E_s)$ *--Supply chain network*
- $G_p(V, E_p)$ *--Competition network*
- V*--A full set of firms (i.e. nodes)*
- E_s*—A full set of edges in the supply chain network*
- E_p*—A full set of edges in the competition network*
- v_x*--A node in V*
- V_d*--A list of disrupted nodes*
- V_r*--A list of nodes to be removed*
- v_i*--A node in* V_{iter}
- S_x *--Suppliers of* v_x *(preceding nodes in* G_s*)*
- s_x*--A node in* S_x
- C_x*--Customers of* v_x *(succeeding nodes in* G_s*)*
- v_m*--A node in* C_x
- P_x*--Competitors of* v_x *(neighbors in* G_p*)*
- v_n*--A node in* P_x
- U_x*--The upper limit of* v_x *to accept requests as alternative suppliers*
- LS_x*--A list of lost suppliers of* v_x
- ls_x*--A node in* LS_x
- q_x*--The alternative supplier that node* v_x *sends request to*
- Req_x*--A list of nodes that send requests to node* v_x
- req_x*--A node in* Req_x
- Acc_x*--A list of nodes whose requests accepted by node* v_x
- acc_x*--A node in* Acc_x
- Rej_x*--A list of nodes whose requests rejected by node* v_x
- rej_x*--A node in* Rej_x
- $trial_x$*--Number of requests sent by node* v_x
- $iter$*--Current iteration of a simulation cycle*
- t*--Current time tick in one iteration of a simulation cycle*
- $p_{m,n}$*--Probability that* v_m*, a customer of* v_i *, sends a request to candidate supplier* v_n
- E_a*--A list of edges to be added to* G_s

CONSTANTS:

- $maxIter = 13$ *// The maximum iteration of one simulation cycle.*
- $maxTrial = 10$ *// The maximum number a node can send requests for alternative suppliers in one simulation cycle.*
- $Revenue_x$ *// Revenue level (firm size) of node* v_x*: small, medium, and large*

MAIN PROGRAM

```
START simulation

FOR iter = 1 to $maxIter$
   SET $E_a = \emptyset$ // Empty the set of edges added in the previous iteration.
   IF $iter$==1
         INITIALIZE $V_d = \emptyset$, $V_r = \emptyset$
         FOR each $v_x \in V$
               INITIALIZE $U_x$ based on simulation settings in Table 4.
               INITIALIZE $Req_x$, $Acc_x$, $Rej_x$
               SET $trial_x = 0$
         END FOR
   ELSE
         FOR each $v_x \in V$
               RESET $Req_x$, $Acc_x$, $Rej_x$
         END FOR
   END IF

   FOR each $v_i \in V_r$  // Remove nodes
         OBTAIN $C_i$ and $P_i$ // Find customers and competitors for nodes to be removed
         FOR each $v_m \in C_i$ and each $v_n \in P_i$
               OBTAIN  $p_{m,n}$
         END FOR
         $G_s = G_s - v_i$; $G_p = G_p - v_i$;
         Remove $v_i$'s edges from $E_s$ and $E_p$
   END FOR
   SET $V_r = \emptyset$ // Empty the set of nodes to be removed

   FOR $t$ = 1 to $maxTrial$
         // Send requests to alternative suppliers
         FOR each $v_x \in V$
               IF $v_x$ has no lost suppliers OR $trial_x == maxTrial$
                     CONTINUE
               END IF
               // Send out requests for each of the lost suppliers of node $v_x$
               FOR each $ls_x \in LS_x$
                     DRAW $q_x$ according to $p_{m,n}$
                     Send request to $q_x$
                     $trial_x = trial_x$ +1 // $v_x'$s number of trials increase by 1
               END FOR
         END FOR

         // Alternative suppliers handle requests
         FOR each $v_x \in V$
               IF $U_x < 1$ OR $Req_x = \emptyset$  // This node has no extra capacity
                     CONTINUE
               END IF
```

```
            OBTAIN  Acc_x,Rej_x // Decide which requests to accept and reject

            FOR each acc_x ∈ Acc_x // Accept requests and add edges
                IF edge V_x→acc_x NOT in G_s
                    APPEND (V_x,acc_x) to E_a
                END IF
                U_x = U_x −1  // The extra capacity decrease by 1
            END FOR

            FOR each rej_x ∈ Rej_x
                REMOVE v_x from rej_x's candidate list
            END FOR
        END FOR
    END FOR

    FOR each v_i ∈ V_r // Increase the capacity for suppliers for removed nodes
        FOR each s_i ∈ S_i
            U_i = U_i +1
        END FOR
    END FOR

    ADD edges in E_a into G_S

    FOR each v_x ∈ V  //Check which node is disrupted
        IF v_x cannot secure suppliers compared to what is has in iter − 1
            V_d = V_d + v_x  // Set this node as disrupted
        END IF
    END FOR

    FOR each v_x ∈ V_d//Check which disrupted node(s) should be removed
        IF v_x is to be removed based on a normal/uniform distribution with mean in Eq.-1.
            V_r = V_r + v_x  // Set this node as to be removed
            V_d = V_d − v_x
        END IF
    END FOR

END FOR
STOP simulation
```

Appendix 3: Examining assumptions of linear regressions

The validity of linear regression results depends on three assumptions. In this appendix, we test these assumptions for our regression model on factors impacting the effectiveness of proactive strategies.

1. Linearity

Linear regressions assume the existence of linear relationships between independent and dependent variables. We first apply Harvey Collier test (Harvey & Collier, 1977) for linearity validation. The p value is less than 0.001 – there is sufficient evidence that we reject the null hypothesis of linearity.

To address such non-linearity issue, we apply logarithm transformations to all covariates except the three dummy codes for sectors. Applying the same test on the linear regression model with transformed data, the p value becomes 0.1 where we fail to reject the linearity assumption. In the following two sections, we test the model with log-transformed data against normality and homoscedasticity assumptions. Regression models we used in the main paper are also based on log-transformed data.

2. Normality

We apply quantile-quantile (Q-Q) plot to examine normality of the residual score. As Figure A.3.1 shows, there is a strong linear relationship ($r^2 = 0.89, p < 0.001$) between sample and theoretical (in this case, our reference distribution is normal distribution) quantiles. At the same time, the scatter points do not fall on a straight line, indicating some degree of non-normality in the residual.

Nonetheless, we argue that, first of all, normality is not a necessary assumption for linear regression models. Specifically, according to Lumley et al. (2002):

> "… [Normality] is not necessary for the least-squares fitting of the regression model but it is required in general for inference making … only extreme departures of the distribution of Y from normality yield spurious results.
>
> This is consistent with the fact that the Central Limit Theorem is more sensitive to extreme distributions in small samples, as most textbook analyses are of relatively small sets of data…"

Further, our sample size of 181 is large enough for the statistical inference to be effective. Past studies have shown that sample sizes of 40 (Barrett & Goldsmith, 1976) or 80 (Ratcliffe, 1968) are large enough to diminish the departure from normality for inference.

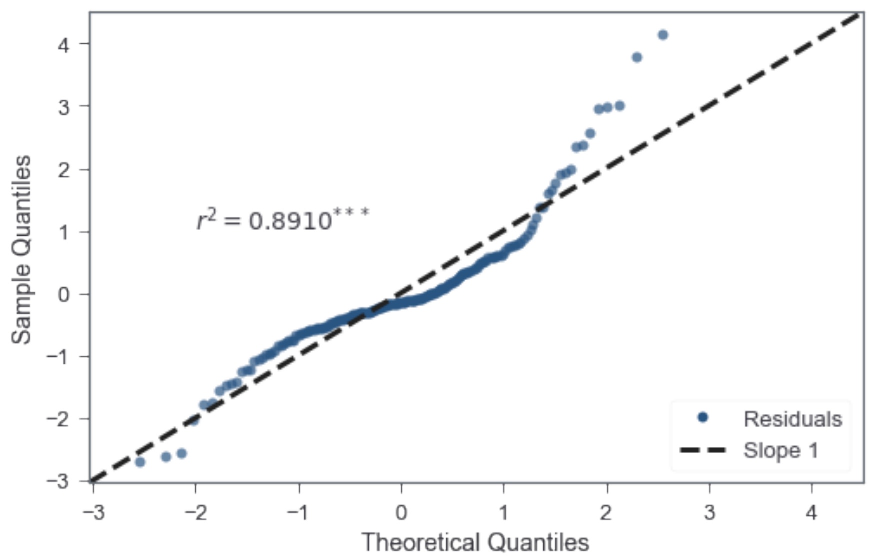

Figure A.3.1. Quantile – Quantile plot for residuals.

3. Homoscedasticity

Homoscedasticity requires that the variance of error terms (i.e. residuals) stay constant across different values of independent variables. We apply the Breusch-Pagan test (Breusch & Pagan, 1979) and obtain a p-value less than 0.001, which indicates heteroscedasticity. To address this problem, we used robust standard errors [a.k.a, White standard errors; (Arellano, 1987)] in the model, which is a common way for dealing with heteroscedastic data in linear regression models.

References

Arellano, M., 1987. Computing Robust Standard Errors for Within-groups Estimators. Oxford Bulletin of Economics and Statistics 49 (4), 431–434.

Barrett, J. P., & Goldsmith, L. (1976). When is n sufficiently large? *American Statistician*, *30*(2), 67–70.

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, *47*(5), 1287. The Econometric Society.

Harvey, A. C., & Collier, P. (1977). Testing for functional misspecification in regression analysis. *Journal of Econometrics*, *6*(1), 103–119. North-Holland.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annu. Rev. Public Health*, *23*, 151–69.

Ratcliffe, J. F. (1968). The Effect on the t Distribution of Non-Normality in the Sampled Population Author. *Journal of the Royal Statistical Society*, *17*(1), 42–48. WileyRoyal Statistical Society.