Disruption Risk Evaluation on a Large-scale Production Network with Establishments and Products

Hiroyasu Inoue[1,2*] **and Yasuyuki Todo**[3,4]

[1]University of Hyogo, Graduate School of Information Science, Kobe, 6500047, Japan
[2]RIKEN, Center for Computational Science, Kobe, 6500047, Japan
[3]Waseda University, Graduate School of Economics, Tokyo, 1698050, Japan
[4]Research Institute of Economy, Trade and Industry, Tokyo, 1008901, Japan
[*]inoue@gsis.u-hyogo.ac.jp

ABSTRACT

This study investigates the propagation of economic shocks through production networks by comparing firm- and establishment-level networks. To achieve this, we integrate a comprehensive firm-level production network dataset for Japan with a dataset covering nearly all Japanese establishments and including detailed output information. This approach allows us to construct an establishment-level production network that identifies multiple inputs and outputs for each establishment, encompassing 183,951 manufacturing establishments within 157,537 firms and capturing 919,982 inter-establishment linkages. Applying the data to a probabilistic diffusion model, we simulate the transmission of supply chain disruptions and derive the following key findings: (1) The establishment-level network exhibits a higher degree of shock propagation than the firm-level network. (2) Incorporating product-level information enhances the estimated propagation effects relative to industry-level classifications alone. (3) Regional shocks propagate significantly differently across the two network levels, highlighting the misidentification of more micro supply chains in the firm-level network. These findings underscore the importance of granular data in accurately assessing economic shock transmission through production networks.

1 Introduction

Supply chain disruptions pose significant risks to the real economy, as they can propagate through production networks, amplifying their impact beyond the initially affected firms. When an exogenous shock,such as a natural disaster, pandemic, or geopolitical conflict, reduces the production capacity of certain firms, the effects extend to their suppliers, which experience a decline in demand, and to their customers, which face shortages of essential materials, parts, or components[1,2]. As a result, even relatively small initial shocks can generate substantial indirect effects, sometimes surpassing the direct impact of the original disruption[3,4]. These cascading effects can contribute to large-scale fluctuations across the economy[5–7].

A key limitation of prior research on shock propagation in supply chains is its reliance on inter-industry analysis based on input-output (IO) tables[7–9]. While IO tables provide a useful macroeconomic perspective, they fail to capture the granular complexity of actual production networks, which operate at the firm or even establishment level. The interactions in these networks are significantly more intricate than physical interactions, such as those between molecules, due to the heterogeneity of firms and their organizational structures. Furthermore, because IO tables aggregate firm-level transactions, they obscure individual

firm-specific circumstances and, as a result, tend to underestimate or misrepresent the true magnitude of shock propagation[10,11]. To address these limitations, recent research leveraging firm-level production network data has provided new insights into the complexities of supply chain disruptions, a development made possible by the increasing availability of detailed datasets[10,12–14].

Despite the progress made in firm-level analyses, several challenges remain. First, supply chains primarily operate at the establishment level, where intermediate goods are transacted between individual plants rather than between firms as a whole. Consequently, the structure of the production network at the establishment level differs from that at the firm level and is likely to be more intricate. Second, many firm-level datasets record the location of corporate headquarters but not the locations of individual establishments. Since headquarters tend to cluster in major metropolitan areas such as Tokyo, firm-level networks may fail to accurately represent inter-regional supply chain linkages. This limitation can lead to an underestimation of regional economic interdependence and an inability to analyze the propagation of region-specific shocks. Finally, standard firm-level datasets often classify firms by industry but do not provide detailed information about the specific products manufactured at each establishment. To account for input substitution and resilience within production networks, previous studies have often assumed that firms within the same industry produce homogeneous products. However, this assumption is overly simplistic and neglects potentially significant variations in product characteristics within a given industry.

To address these challenges, this study constructs a large-scale production network in Japan that explicitly identifies the linkages between establishments and the products they manufacture. This analysis is based on two key datasets. The first is the Company Information Database and Company Linkage Database compiled by Tokyo Shoko Research (TSR), which, as of 2020, contains data on 1,520,605 firms and 5,860,726 supplier-client relationships. Our analysis focuses on a subset of manufacturing firms. The second dataset is the Economic Census for Business Activity, conducted by the Ministry of Economy, Trade, and Industry in 2021, which includes data on 512,401 manufacturing establishments. Crucially, the Census data provide information on the specific products manufactured at each establishment, enabling us to decompose inter-firm linkages into establishment-level connections by inferring the "recipe" of input combinations required to produce a given output. From this process, we construct a production network comprising 183,951 manufacturing establishments affiliated with 157,537 firms and 919,982 supply chain linkages between these establishments. We then apply a probabilistic diffusion model to simulate the propagation of supply chain disruptions across networks of varying levels of granularity.

Our main findings are as follows. First, simulations comparing the TSR firm-level network with the newly constructed establishment-level network reveal that the latter exhibits greater shock propagation effects. Second, we analyze different levels of input substitutability, considering scenarios with no substitution, substitution between identical products as classified in the Census data, substitution between products within the same industry, and complete substitution. Our results indicate that incorporating detailed product-level information significantly amplifies shock propagation relative to models relying solely on industry classifications. Finally, when disruptions originate in Tokyo, where corporate headquarters are concentrated, firm-level analyses tend to overestimate the extent of propagation compared to shocks originating from other regions. In contrast, establishment-level analysis, which accounts for the precise locations of production facilities, does not exhibit such overestimation, highlighting the importance of accurately mapping network structures. By integrating establishment-level production network data and detailed product information, this study provides a more comprehensive framework for assessing the resilience and vulnerability of supply chains. Our findings underscore the necessity of granular data in capturing the true dynamics of economic disruptions and offer valuable insights for policymakers and businesses seeking to enhance supply chain robustness.

2 Data

2.1 Primary data

One of the primary data sources for this study is the Company Information Database and Company Linkage Database, collected by Tokyo Shoko Research (TSR). The Company Information Database provides key attributes for each firm, including its name, address, industry classification, sales, and corporate identification number. The Company Linkage Database details domestic supplier-client relationships. We utilize data from 2020 to ensure consistency with the Economic Census for Business Activity, conducted in 2021, as described below. Among the available years, 2020 is the closest to 2021, making it the most suitable for our analysis. The TSR dataset contains 1,520,605 firms and 5,860,726 trade links, encompassing a broad range of firms in Japan, and capturing the majority of significant production relationships. As outlined later, we extract manufacturing firms from this dataset for our analysis.

While the TSR dataset provides valuable information to construct the firm-level production network of most firms in Japan, a key limitation is its lack of information on individual establishments and the specific products they manufacture. To address this gap, we incorporate data from the Economic Census for Business Activity (hereafter referred to as the Census). The Census is conducted every five years by the Ministry of Internal Affairs and Communications and the Ministry of Economy, Trade and Industry in Japan[15]. It covers all establishments across all industries, including those of micro, small, and medium-sized enterprises. The Census dataset provides comprehensive details for each establishment, including its name, address, phone number, number and type of employees, business activities, enterprise type, breakdown of sales by products, breakdown of expenses, and purchased items. For this study, we specifically use the 2021 Census data, which includes 5,156,063 establishments.

To construct a more detailed production network, we integrate the Census data with the TSR dataset using corporate identification numbers. This process adds 1,014,673 non-headquarter establishments to the dataset, supplementing the TSR dataset's information on firm headquarters. Although the merged dataset encompasses all firms and establishments in Japan, our analysis focuses on the manufacturing sector for three primary reasons. First, identifying a network of establishments requires inferring the "recipe" for producing each product based on a combination of specific inputs. This inference is more feasible for the manufacturing sector compared to other industries. Second, including additional sectors, particularly the wholesale sector, complicates the analysis of production networks between establishments due to the nature of trade relationships. Finally, we utilize purchase data from the Basic Survey of Japanese Business Structure and Activities (BSJ) (Kigyo Katsudo Kihon Chosa), an annual survey conducted by the Ministry of Economy, Trade, and Industry. This survey primarily targets the manufacturing sector, making it a suitable complement to our dataset. As a result of this data integration and selection process, our final sample consists of 157,537 firms and 183,951 establishments, producing 2,396 distinct products, classified into 595 industries according to the Japanese Standard Industrial Classification (JSIC)[16].

2.2 Creating the production network with establishments and products

Using the two data sources, i.e., the TSR dataset containing firm-level links and the Census dataset with establishment-level attributes but no explicit link information, we infer plausible supplier-client relationships between establishments. To achieve this, we adopt a two-step methodology. First, we filter links based on standard input-output relationships. Second, we refine these connections using a set cover optimization approach.

In the first step, we establish initial links by assuming that every establishment of a supplier firm is potentially connected to every establishment of its client firms. We then refine these connections by identifying "standard recipes"—the typical input-output relationships observed in the economy. Specifically, we

consider an input essential for a product if it appears in at least 50% of the maximum number of observed connections for that product. For instance, if all 1,000 establishments producing automobiles are linked with machinery producers but only 300 have links to food producers, we infer that machinery is a valid input for automobile production, whereas food is not. While the 50% threshold is an arbitrary cutoff, we verify the robustness of our results using alternative values.

After filtering links based on these standard recipes, some client establishments may still have redundant supplier connections. For example, if a supplier firm operates multiple establishments producing the same product, a client establishment may not need to maintain links with all of them. To resolve this, we apply a set cover problem formulation to determine the minimal set of supplier establishments required to satisfy the input demands of each client establishment.

The set cover problem is a well-known optimization problem. It involves a universal set containing all elements and a collection of subsets each of which contains some elements from the universal set. The objective is to find the smallest number of these subsets that together include every element in the universal set. Although this problem is usually NP-hard, it is solvable in this study due to the relatively small size of each instance.

In our context, the set cover problem is to select the minimal combination of supplier establishments for each client establishment. The universal set consists of the collection of inputs identified by the cutoff recipes. The subsets represent supplier establishments, and their elements are the products produced by these establishments.

Additionally, we incorporate groups into the set cover problem. A group represents a collection of subsets, corresponding to a firm (as a group of establishments) in this study. The group constraint requires that at least one subset (establishment) from each group (firm) be chosen, ensuring consistency with firm-level transactions observed in the TSR data.

Furthermore, we introduce a priority into the set cover problem. If a subset (an establishment) is chosen, the priority of other subsets (establishments) within the same group (firm) is reduced by one. Therefore, if multiple subsets (establishments) contain the same elements (output products) but belong to different groups (firms), the selection process prioritizes those with lower priority values.

By solving this optimization problem, we construct a minimal yet functional establishment-level network that aligns with firm-level production relationships and product information for each establishment. However, it is important to note that this inferred network represents a lower-bound approximation of actual supplier-client relationships. In practice, additional supplier links may exist beyond those predicted by the set cover solution, but without direct establishment-level trade data, validation against the true network remains challenging.

2.3 Structure of establishment-level network

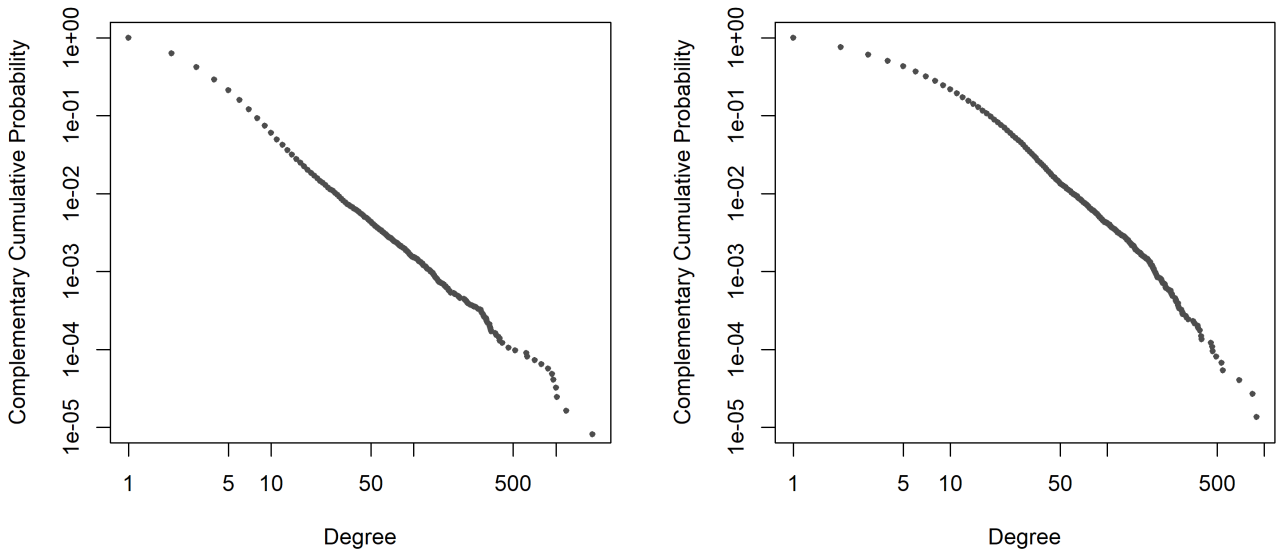
This subsection describes the establishment-level network obtained from the two-step procedure outlined above. In the first step, the total number of all possible product combinations based on the firm-level network and establishment-level production information is 1,182,274. After applying the cutoff, the standard recipes retain 43,340 input-output product combinations. The complete set of trade connections between supplier and client establishments, encompassing all establishments of supplier and client firms, amounts to 18,872,233. After applying the cutoff recipes and solving the set cover problem, we identify 919,982 plausible connections between establishments.

The firm-level network extracted from the TSR data consists of 157,537 nodes (firms) and 480,507 links, whereas the inferred establishment-level network includes 183,951 nodes (establishments) and 919,982 links, as summarized in Table 1. As expected, the establishment-level network exhibits a higher average number of links per node than the firm-level network.

Table 1. Number of nodes and links for firm- and establishment-level networks.

	node	Link	Average link per Node
Firm-level	157,537	480,507	6.1
Establishment-level	183,951	919,982	10.0

The degree distributions for both networks are illustrated in Figure 1, disregarding link directionality. Although the establishment-level network (panel (b)) is constructed using more granular data, its degree distribution retains the overall form of the firm-level network (panel (a)) and appears to follow a power-law distribution. Notably, networks with power-law degree distributions tend to have short average path lengths, implying that shocks can propagate rapidly throughout the network[17].



(a) Firm level **(b)** Establishment level

Figure 1. Degree of production networks

However, there are notable differences in the lower frequency of low-degree nodes. For example, 90% of the nodes have fewer than 7 links in the firm-level network, while 90% have fewer than 17 links in the establishment-level network. This difference may be due to our inference process where every establishment is assumed to find supplier establishments to fulfill its recipe, often resulting in multiple connections with establishments within the same supplier firm. These connections are simply counted as one in the firm-level network. This reasoning is supported by the difference observed in the distribution of the complete connections between establishments (SI Fig. 1). In this complete connection network, 90% of the nodes have fewer than 56 links, a significant deviation from the firm-level network.

One key advantage of constructing an establishment-level network, as opposed to a firm-level network, is the ability to analyze production linkages at a more granular geographic scale. Figure 2 compares the geographic distribution of firms and manufacturing establishments across Japan's prefectures. As expected, Tokyo (Code 13), the capital city, has a significantly higher concentration of corporate headquarters—approximately double the number of establishments. By contrast, major industrial prefectures

such as Saitama (Code 11), Shizuoka (Code 22), and Aichi (Code 23) exhibit the opposite pattern, with a greater number of establishments than firm headquarters.

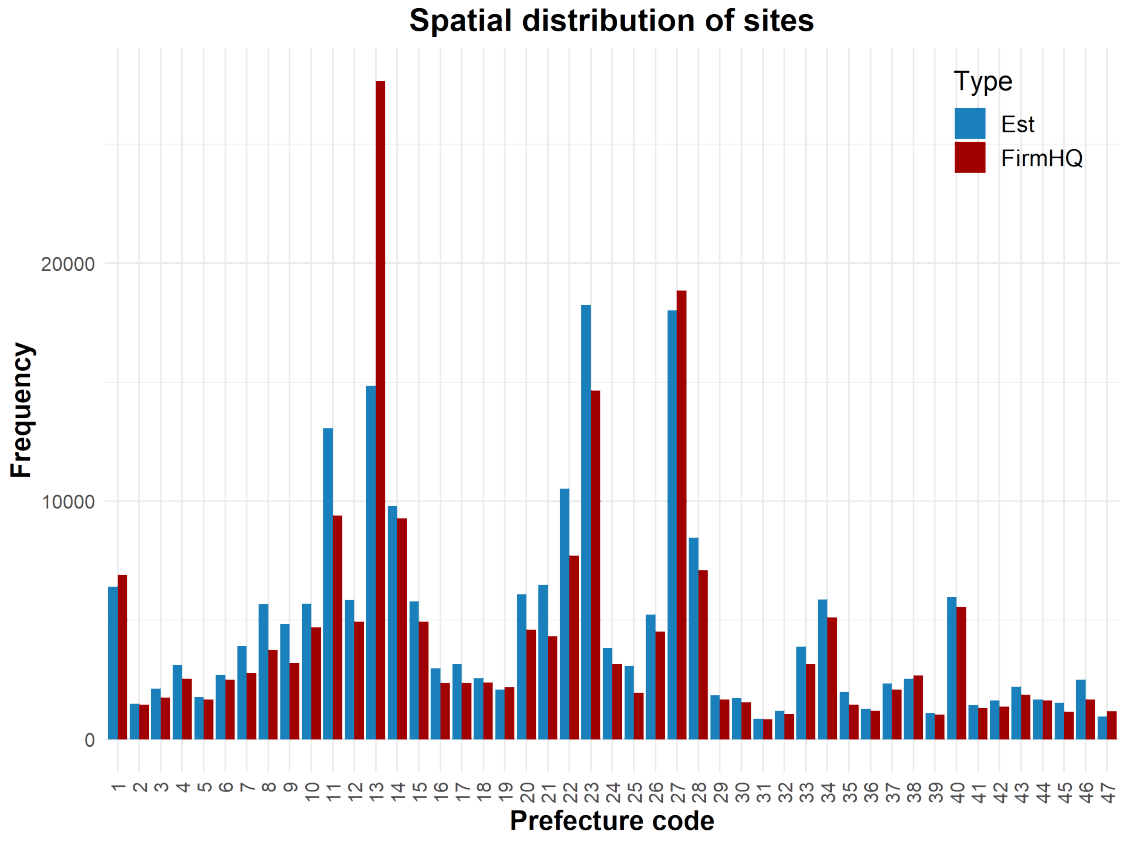


Figure 2. Degree of production networks

3 Method

3.1 Model

To examine propagation of the effect of supply chain disruptions, we propose a model in which infeasibility of production propagate probabilistically through a production network. While many existing models account for complex behaviors within and between entities to simulate shock propagation[10,18,19], this paper introduces a substantially simpler model, similar to the one in Zhao et al.'s work[20], allowing for straightforward comparison between networks at firm and establishment levels. A key feature of the present model is our assumption that each establishment produces its products from the same combination of inputs determined by the current network. In other words, in case of supply chain disruptions, substitution of the disrupted input for another type of product that is not currently used as an input is not possible. However, when an input product is provided by multiple suppliers, substitution between suppliers can mitigate the propagation of supply chain disruptions.

The formal description of the model is as follows. First, we define notations in the model. Consider a production network of entities where entities are either firms or establishments. The adjacency matrix between entities i and j is denoted by A_{ij} that is one if entity i supplies any of its products to j and zero otherwise. The set of output products of each entity i are denoted as P_i^{out}. Then, the input products of

entity i can be described as $P_i^{\text{in}} = \bigcup_j A_{ji} \otimes P_j^{\text{out}}$, where $\otimes$ indicates that if the left operand is equal to one, the right operand is valid, and vice versa.

Next, we describe the propagation process. The state of entity i at time t, $s_i(t)$, is one if i is active and produces outputs and 0 if i is inactive and does not produce any product. Once an entity becomes inactive due to the propagation of economic shocks, it cannot become active again during the simulation process. Let U be the set of entities that became newly inactive at time $t-1$. The set of active client entities that have at least one supplier in U is denoted as V. For an entity v in V, let U_v be the subset of U that supplies to v. Then, we assume that the probability that entity v becomes inactive at time t ($s_v(t) = 0$), defined as ρ_v, is given by:

$$\rho_v(t) = 1 - \Pi_{q \in \bigcup_{i \in U_v} P_i^{\text{out}}} (1 - p_{vqt}). \tag{1}$$

p_{vqt} is the probability of shock propagation to entity v at time t through product q. Eqn. (1) indicates that only when there is no shock propagation for all supply products that newly inactive suppliers have, v remains active and vice versa. Further, we assume that p_{vqt} is expressed by:

$$p_{vqt} = p \left(1 - \frac{\sum_j s_j(t) A_{jv} \mathbf{1}_{P^{\text{out}}{}_j}(q)}{\sum_j A_{jv} \mathbf{1}_{P^{\text{out}}{}_j}(q)} \right), \tag{2}$$

where p is a parameter that determine the benchmark propagation probability of becoming inactive and assumed to be either 0.2, 0.4, 0.6, 0.8, or 1, and $\mathbf{1}$ is an indicator function defined by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases} \tag{3}$$

Eqn. (2) shows that if client entity v has more than one suppliers producing product q and some of the suppliers become inactive, v can reduce the probability of becoming inactive by the share of its active suppliers of q. In other words, client firms can mitigate propagation of shocks through the production network by substitution disrupted suppliers for other existing suppliers producing the same product.

In each simulation run, all entities are active initially: $s_i(0)$ for any i. At time 1, one entity is uniformly and randomly selected to become inactive from all entities (or within specific candidates depending on the setting). Then, the inactive state propagates through the production network, following the assumptions defined by Eqns. (1)-(3), until any entity becomes newly inactive. At the conclusion of each simulation run, we measure the share of inactive entities in the economy, which serves as an indicator of the extent of shock propagation. To ensure robustness, we conduct 100,000 independent simulation runs for each scenario, varying the initially selected entity. We then compute the mean inactive ratio and its standard deviation to assess the overall impact and variability of disruptions under different network structures and settings.

3.2 Simulation frameworks

Using the data and model described above, we will address three issues by comparing simulation results using different networks, assumptions, and parameter values.

First, we analyze the differences in shock propagation between the firm-level network and the establishment-level network. To ensure comparability, we define "products" represented by q in Eqn. (2) using 545 industry classifications within the manufacturing sector, following previous studies[10,12–14]. This approach assumes that inputs from suppliers in the same industry are substitutable. Notably, we do not employ the product classifications for the establishment-level network in this step, as our primary goal is

to isolate the differences in propagation dynamics between the two networks under a consistent definition of q. In cases where a firm or establishment is classified under multiple industries, we use the primary industry classification. Accordingly, the industry of a firm and that of an establishment belonging to the firm are not necessarily identical.

Second, we assess the impact of different levels of input substitutability on shock propagation within the establishment-level network. We consider four scenarios. First, we assume no substitution between inputs even if inputs are defined by the same product classification. In other words, q in Eqn. (2) is defined by each establishment. This assumption implies that each establishment produces unique outputs. Second, we relax this assumption slightly and assume substitution within product classifications, defining q by the 2,396 product types in the census data. In other words, inputs are substitutable if they belong to the same product type. Third, substitution within industry classifications is allowed for. Inputs are substitutable if they belong to the same industry, even if they are classified under different product types. This assumption aligns with the industry-based classification used in the first analysis. Finally, we assume that all establishments produce the same product, allowing for universal substitution of inputs regardless of product or industry classification. Under this assumption, the probability of shock propagation, p_{vqt} in Eqn. (2), is equal to q times the share of inactive suppliers.

Finally, we examine how the geographic origin of shocks influences propagation dynamics. This analysis is motivated by the differences in the spatial distribution of firms (headquarters) and establishments, as illustrated in Fig. 2. Importantly, geographic distribution alone does not inherently affect the magnitude of shock propagation, as (i) only a single node is initially disrupted in each simulation and (ii) network topology, rather than physical location, determines propagation patterns. To test the hypothesis that firm-level data overstates the economic importance of Tokyo—where corporate headquarters are concentrated—we simulate shock propagation originating from Tokyo and compare it to propagation from other regions using both the firm-level and establishment-level networks. Regardless of the network type, we assume that each entity (firm or establishment) may produce multiple products to ensure comparability across scenarios. Specifically, in the establishment-level analysis, each establishment produces multiple products as in the second analysis, while in the firm-level analysis, each firm is assigned the combined products of its establishments. By incorporating actual product information rather than relying solely on industry proxies, we aim to provide a more realistic assessment of shock propagation. We hypothesize that simulations using the firm-level network will overestimate the impact of shocks originating from Tokyo, whereas the establishment-level network will offer a more accurate representation by incorporating the actual locations of production facilities.

4 Simulation Results and Discussions

4.1 Comparisons between firm- and establishment-level networks

We begin by comparing the extent of shock propagation between the firm- and establishment-level networks. Fig. 3 shows that the establishment-level network exhibits a significantly greater propagation effect, as measured by the mean share of inactive entities at the end of each simulation. The difference is statistically significant based on t-tests. In particular, when the baseline probability of propagation p is low, such as 0.2 and 0.4, the difference is substantial. As the baseline probability increases, economic shocks are more likely to propagate across both networks, reducing the relative difference between them.

The discrepancy is notable, because the links in the firm-level network are maintained in the establishment-level network and because the number of entities, 157,537 in the former and 183,951 in the latter, is not very different (Section 2.2). Furthermore, the variation in shock propagation is not driven by differences in input classifications, as both networks define products based on industries, ensuring comparability.

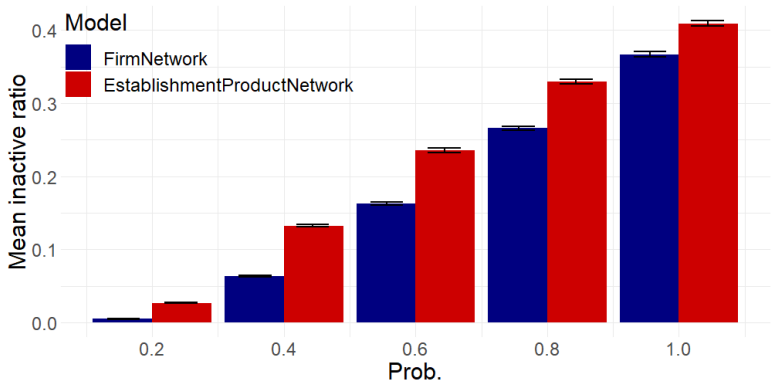


Figure 3. Comparison in the propagation effect between firm- and establishment-level networks: This figure shows the mean and standard error of the share of inactive entities at the end of each of 100,000 simulations, using either firm- (blue bars) or establishment-level (red bars) network and the baseline probability of propagation q of 0.2, 0.4, 0.6, 0.8, or 1. In both networks, industries are used as the definition of products q in the propagation model.

The primary explanation for this difference lies in the greater diversity of input linkages in the establishment-level network. As indicated in Table 1, the average number of suppliers per node is higher in the establishment-level network than in the firm-level network. While one might expect that a higher number of suppliers per entity would lead to more substitutable inputs, reducing the impact of shocks, our results indicate the opposite. Instead, the establishment-level network reveals a wider variety of input types with less overlap, implying that firms classified under a particular industry often operate establishments involved in different industries to support their diverse production activities. This structural complexity leads to a higher degree of shock propagation in the establishment-level network.

This distinction between firm- and establishment-level networks is conceptually similar to the well-documented differences between analysis using industry-level IO tables and firm-level production networks. Previous studies have shown that IO table analyses tend to underestimate supply-side shocks compared to firm-level analyses[21,22]. This discrepancy is often attributed to the difference in network structures: IO tables depict dense and nearly completely interconnected networks, whereas firm-level networks are sparse with each firm connecting to distinct partners. This sparsity and heterogeneity have been identified as a primary factor contributing to the observed differences. However, in our comparison between firm- and establishment-level network reveals that even within an already heterogeneous network, incorporating denser connections among establishments to reflect real-world interactions can amplify the propagation of shocks. In other words, this study demonstrates that disaggregation from firms to establishments has a significant impact on shaping differences in shock propagation.

4.2 Input substitutability in the establishment-level network

Next, we focus on the establishment-level network and compare shock propagation between different assumptions of the input substitutability. The results presented by Fig. 4 show that the mean ratio of inactive establishments, our measure of the propagation effect, is the lowest under the assumption of complete substitution between inputs, followed by that of substitution within identical products and then between products within the same industry, and the highest when no input substitution is assumed, for any value of p, the baseline probability of propagation. All differences between scenarios are statistically significant based on t-test.

These findings align with the structural properties of a power-law degree distribution network, where

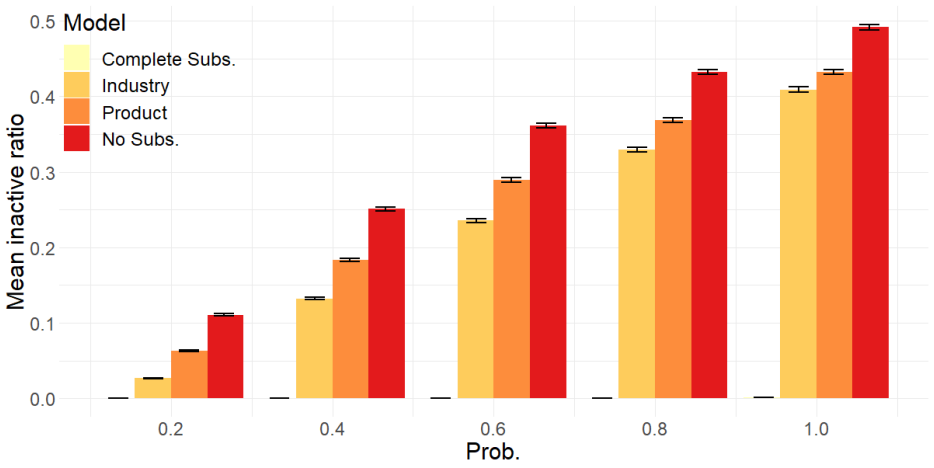


Figure 4. Comparison between different levels of input substitutability: This figure shows the mean ratios of inactive establishments and their standard errors from simulations using the establishment-level network under four different assumptions of input substitutability: complete substitution (labeled as "Complete Subs." and represented by yellow bars), substitution at the industry level ("Industry" and light orange), at the product level ("Product" and orange), and no substitution ("No Subs." and red). "Prob." indicates the baseline probability of propagation, p, in the model.

hubs (highly connected establishments) play a critical role in propagation dynamics. Large-scale shock transmission is only possible if the shock successfully spreads through these hubs. Under perfect substitution, highly connected hubs can replace inactive suppliers with alternative sources, preventing further propagation, as reflected in the (almost invisible) yellow bars in Figure 4. Conversely, when no substitution is possible, even hubs are vulnerable to shock propagation, leading to cascading failures across their dependent establishments, as illustrated by the red bars in the figure. These results underscore the crucial role of input substitutability in determining the extent of shock transmission, consistent with the findings from the firm- versus establishment-level comparison in the previous subsection.

However, it is important to emphasize that, conceptually, the assumption of substitution within identical products can either increase or decrease the propagation effect compared to the assumption of substitution between products within the same industry. On one hand, for instance, consider a scenario where establishment A_1 sources supplies from establishment B_1 of firm B and establishment C_1 of firm C. Suppose further that firms B and C belong to the same industry. Then, under the industry-level substitution assumption, suppliers B_1 and C_1 are considered substitutable. However, if B_1 and C_1 produce different products according to our product-level classification, they would not be substitutable under the product-level substitution assumption. Consequently, the assumption of substitution at the product level implies lower input substitutability and, therefore, predicts a greater propagation effect compared to substitution at the industry level.

On the other hand, our establishment-level analysis, which incorporates detailed product information, allows each establishment to produce multiple products. For example, if firms B and C belong different industries, but establishments B_1 and C_1 produce multiple products that are overlapped with each other, the two suppliers would be substitutable under the product-level substitution, but not under the industry-level substitution. In this case, shock propagation is greater in the firm-level network than in the establishment-level network. Our findings suggest that the reduction in substitutability due to finer product classifications outweighs the increase in substitutability that arise from establishments producing multiple products.

In conclusion, the approach of this paper—using establishments instead of firms, using products instead

of industries, and allowing for multiple products per establishment—offers a more realistic representation of shock propagation through production networks. These results emphasize the importance of granular data in accurately modeling economic disruptions.

4.3 Comparison between regional shocks

In the final set of simulations, we compare the propagation of shocks originating from Tokyo versus other regions, using both the firm-level network and the establishment-level network. Figure 5 illustrates the differences between these network structures.

When using the firm-level network, shock propagation is significantly greater when the shock originates in Tokyo (dark blue bars) compared to when it originates in other regions (lighter blue bars), across all values of p, the baseline probability of propagation. These differences are statistically significant according to t-tests. However, when using the establishment-level network, this pattern disappears. There is no statistically significant difference in propagation effects between Tokyo-origin and other regional shocks (red and orange bars).

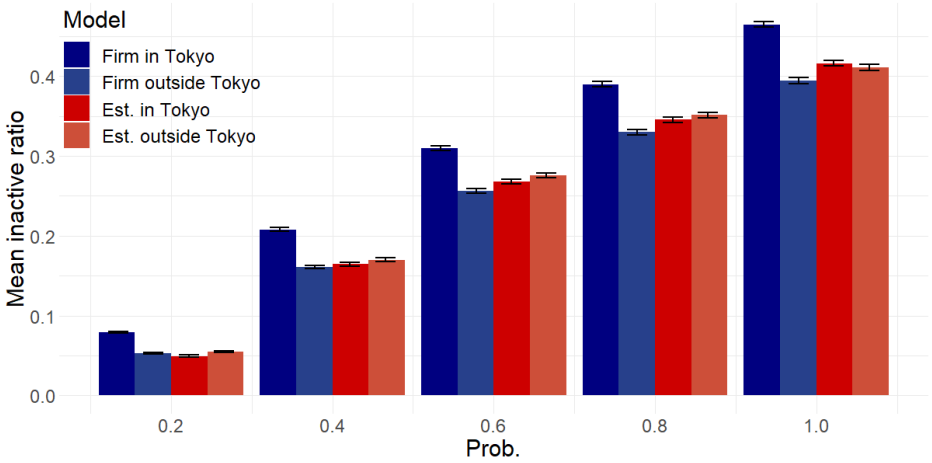


Figure 5. Comparison of firm- and establishment-level networks in response to different regional shocks: This figure shows the mean ratios of inactive establishments and their standard errors from simulations using the firm- and establishment-level networks assuming shocks originating in and outside Tokyo. "Prob." indicates the baseline probability of propagation, p, in the model.

These findings suggest that relying on a firm-level network may lead to incorrect conclusions due to the misrepresentation of the geographic distribution of production sites within supply chains. For example, the firm-level simulation results might imply that firms headquartered in Tokyo should be prioritized for protection against supply chain disruptions to mitigate widespread economic consequences. However, the establishment-level analysis challenges this implication, revealing that the actual production sites, which are more dispersed across regions, do not support such a policy recommendation.

Several factors contribute to this discrepancy between firm- and establishment-level analyses. A key factor is the concentration of corporate headquarters in Tokyo, while production activities are often spread across multiple regions. Firms headquartered in Tokyo tend to operate multiple establishments, often located outside Tokyo. Examining the distribution of out-degree (i.e., the number of client connections), we find that at the firm level, the average out-degree for firms in Tokyo is 7.24 and larger than that for firms in other regions, 3.66. However, at the establishment level, the average out-degree for Tokyo-based establishments is 5.61 and smaller than for establishments outside Tokyo, 7.79. Since out-degree is

more critical for shock diffusion than in-degree (the number of suppliers), these differences highlight the structural variation in networks at different levels of aggregation. However, shock diffusion is not solely determined by out-degree distributions; network topology and input substitutability also play crucial roles, as demonstrated in our previous analyses. This complexity underscores why our simulation-based approach, which utilizes large-scale firm and establishment data, is more effective than relying solely on conceptual reasoning or descriptive statistics when evaluating supply chain resilience and economic shock propagation.

5 Conclusion and future works

In this study, we constructed an establishment-level production network with detailed product information, covering 183,951 manufacturing establishments of 157,537 firms and 919,982 inter-establishment linkages, by integrating firm-level supply chain data with establishment-level economic census data. Using the firm- and establishment-level networks, we applied a probabilistic diffusion model to simulate the propagation of economic shocks resulting from supply chain disruptions. Specifically, we examined the differences in shock propagation: (1) Between firm- and establishment-level networks; (2) Under different assumptions of input substitutability; (3) Between shocks originating in Tokyo and those in other regions

Our key findings are as follows. (1) The establishment-level network exhibits greater shock propagation than the firm-level network. (2) Shock propagation is more extensive when input substitutability is defined at the product level rather than at the industry level. (3) Firm-level networks overstate the propagation of shocks originating from Tokyo, whereas establishment-level networks show no significant difference between Tokyo-origin and other regional shocks.

These results consistently indicate that greater network granularity amplifies shock propagation effects. That is, propagation is stronger when using establishment-level rather than firm-level data and when defining products at a more disaggregated level rather than by industry classification. These findings suggest that incorporating more detailed and realistic production network information can significantly alter the assessment of economic shock propagation.

Several avenues exist for extending this research. First, the network can be expanded to include non-manufacturing establishments. This study focuses on the manufacturing sector, utilizing 157,537 out of 1,520,605 firms and 183,951 out of 5,156,063 establishments in the census data. However, similar methodologies could be applied to other sectors, thereby expanding the sample. A key challenge is that wholesalers do not transform products but rather facilitate distribution between sectors. By establishing comprehensive input-output relationships for non-manufacturing sectors, future research could develop a more holistic model of inter-industry shock propagation.

Second, alternative algorithms for constructing establishment-level networks can be explored. In this study, we inferred a minimal connection network using a set cover optimization approach (Section 2.2). However, alternative algorithms could introduce additional structural complexity. For example, empirical evidence suggests that geographic proximity influences supplier-client relationships[23]. Future studies could incorporate spatial constraints, as well as additional firm-level factors such as sales relationships, shared financing banks, or cross-shareholding ties[24]. Comparing these alternative network structures in terms of their estimated shock propagation effects would be an important extension of this work.

Finally, a significant portion of firms operate with only one establishment, and this information could be used to refine the network further. For instance, if both a supplier and a client are single-establishment firms, their connection could be treated as a definitive trade relationship, eliminating the need for inference in those cases. Such an approach could improve the accuracy of network construction, particularly in sectors where firm-establishment correspondence is more straightforward.

Data availability

The data utilized in this study were obtained from Tokyo Shoko Research, Ltd., the Ministry of Internal Affairs and Communications, and the Ministry of Economy, Trade, and Industry. However, access to these data is restricted, and they are not publicly available without the appropriate licenses.

References

1. Carvalho, V. M., Nirei, M., Saito, Y. U. & Tahbaz-Salehi, A. Supply chain disruptions: Evidence from the great east japan earthquake. Tech. Rep. No. 17-5, Columbia Business School Research Paper (2016).
2. Barrot, J.-N. & Sauvagnat, J. Input specificity and the propagation of idiosyncratic shocks in production networks. *The Q. J. Econ.* **131**, 1543–1592 (2016).
3. Tierney, K. J. Business impacts of the northridge earthquake. *J. Contingencies crisis management* **5**, 87–97 (1997).
4. Pelling, M., Özerdem, A. & Barakat, S. The macro-economic impact of disasters. *Prog. Dev. Stud.* **2**, 283–305 (2002).
5. Bak, P., Chen, K., Scheinkman, J. & Woodford, M. Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics. *Ricerche economiche* **47**, 3–30 (1993).
6. Gatti, D. D. *et al.* A new approach to business fluctuations: heterogeneous interacting agents, scaling laws and financial fragility. *J. Econ. behavior & organization* **56**, 489–512 (2005).
7. Acemoglu, D., Carvalho, V. M., Ozdaglar, A. & Tahbaz-Salehi, A. The network origins of aggregate fluctuations. *Econometrica* **80**, 1977–2016 (2012).
8. Haimes, Y. Y. & Jiang, P. Leontief-based model of risk in complex interconnected infrastructures. *J. Infrastructure systems* **7**, 1–12 (2001).
9. Okuyama, Y., Hewings, G. J. & Sonis, M. Measuring economic impacts of disasters: interregional input-output analysis using sequential interindustry model. In *Modeling spatial and economic impacts of disasters*, 77–101 (Springer, 2004).
10. Inoue, H. & Todo, Y. Firm-level propagation of shocks through supply-chain networks. *Nat. Sustain.* **2**, 841–847, DOI: 10.1038/s41893-019-0351-x (2019). Number: 9 Publisher: Nature Publishing Group.
11. Diem, C., Borsos, A., Reisch, T., Kertész, J. & Thurner, S. Estimating the loss of economic predictability from aggregating firm-level production networks. *PNAS nexus* **3**, pgae064 (2024).
12. Poledna, S. *et al.* When does a disaster become a systemic event? estimating indirect economic losses from natural disasters. *arXiv preprint arXiv:1801.09740* (2018).
13. Diem, C., Borsos, A., Reisch, T., Kertész, J. & Thurner, S. Quantifying firm-level economic systemic risk from nation-wide supply networks. *Sci. reports* **12**, 7719 (2022).
14. Pichler, A. *et al.* Building an alliance to map global supply networks. *Science* **382**, 270–272, DOI: 10.1126/science.adi7521 (2023). Publisher: American Association for the Advancement of Science.
15. Statistics Bureau of Japan. Economic Census for Business Activity. https://www.stat.go.jp/english/data/e-census/2021/index.html (2021). Accessed: 08.02.2024.

16. Ministry of Internal Affairs and Communications, Japan. Japan Standard Industrial Classification (Rev. 13, October 2013). https://www.soumu.go.jp/english/dgpp_ss/seido/sangyo/index13.htm (2013). Accessed: 08.16.2024.

17. Barabási, A.-L. *Network science* (Cambridge University Press, 2016).

18. Battiston, S., Puliga, M., Kaushik, R., Tasca, P. & Caldarelli, G. Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Sci. reports* **2**, 1–6 (2012).

19. Fujiwara, Y., Terai, M., Fujita, Y. & Souma, W. Debtrank analysis of financial distress propagation on a production network in japan. *RIETI Discuss. Pap. Ser.* **15** (2016).

20. Zhao, K., Zuo, Z. & Blackhurst, J. V. Modelling supply chain adaptation for disruptions: An empirically grounded complex adaptive systems approach. *J. operations Manag.* **65**, 190–212 (2019).

21. Inoue, H. & Todo, Y. Propagation of negative shocks across nation-wide firm networks. *PLOS ONE* **14**, e0213648, DOI: 10.1371/journal.pone.0213648 (2019).

22. Diem, C., Borsos, A., Reisch, T., Kertész, J. & Thurner, S. Estimating the loss of economic predictability from aggregating firm-level production networks, DOI: 10.48550/arXiv.2302.11451 (2023). ArXiv:2302.11451 [physics, q-fin].

23. Nakajima, K., Saito, Y. U. & Uesugi, I. Measuring economic localization: Evidence from japanese firm-level data. *J. Jpn. Int. Econ.* **26**, 201–220 (2012).

24. Krichene, H. *et al.* The emergence of properties of the Japanese production network: How do listed firms choose their partners? *Soc. Networks* **59**, 1–9, DOI: 10.1016/j.socnet.2019.05.002 (2019).

Acknowledgements

This research was conducted as part of a project entitled “Research on Relationships between Economic Networks and National Security” undertaken at the Research Institute of Economy, Trade, and Industry (RIETI). The authors thank the seminar participants at RIETI.

Author contributions statement

H.I conceived the study. H.I and Y.T conducted the analyses and wrote and reviewed the manuscript.