RIETI Discussion Paper Series 16-E-057

Shock Propagations in Granular Networks

FUJII Daisuke
RIETI

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

Shock Propagations in Granular Networks

Daisuke Fujii*

USC Dornsife INET

March 31, 2016

Abstract

This paper studies a number of features of transaction networks, firm sales growth, and buyer-supplier comovements of sales using a large-scale dataset on the Japanese interfirm transaction network. Larger firms have higher sales growth rates and smaller growth dispersion. Well-connected firms also exhibit higher growth rates, but there is no systematic relationship between the number of partners (degree) and sales growth dispersion. Using a statistical test for spatial interdependence, it is confirmed that there exists a significant network interdependence of sales growth. By employing spatial autoregressive models, various propagation factors are estimated. In the baseline specification, the elasticity of average sales growth of suppliers is estimated to be 0.153 while that of customers is 0.257 for year 2012. In all years, the upstream propagation factor is larger than the downstream factor implying a difficulty of replacing an existing customer or adjusting to a demand shock. The manufacturing sector is characterized by a large degree of propagation. For both downstream and upstream propagations, manufacturing and wholesale sectors exhibit higher propagations factors while retail and service sectors exhibit lower propagation factors. The interdependence of intermediate physical inputs produced by other firms may generate an additional margin for the buyer-supplier comovements. It was also found that larger firms have higher propagation factors. Larger firms have more partners, and their degree of propagation is also higher. This result stresses an even larger impact of big firms for aggregate fluctuations in a granular production network.

Keywords: networks, shock propagation, firm growth, aggregate volatility

JEL classifications: D22, D57, D85, L14

*Email: dfujii@usc.edu

1 Introduction

Modern economies are characterized by a complex inter-firm network structure. Firms, the fundamental units of production, are interconnected through financial linkages or supply chains. The financial crisis of 2008-2009, which triggered the Great Recession, underscored the role of banking networks in the context of macroeconomic volatility. The subprime mortgage risk drove a few U.S. financial institutions into bankruptcy. Shortly after, the failure spread to other institutions and connected firms due to credit crunch. The domino effect of adverse shocks was not limited in the U.S. causing the global recession. Firms are also connected in a supply chain network. Many firms rely on the use of intermediate inputs in the form of both physical goods and services produced by other firms. Some manufacturers such as automakers require thousands of parts delivered by their subcontractors. In this production network, supply chain disruptions create the ripples of negative shocks which may spread to many other firms in the economy. As can be learned from these lessons, it is critical to elucidate the properties of an inter-firm network and how shocks spread through the network in order to understand the sources of aggregate volatility. Though there has been a surge of research on economic and social networks in the past decade, there is little empirical evidence on the structure of inter-firm transaction networks and shock propagation at a large scale mainly due to the lack of data. The purpose of the current work is to fill this gap by employing a novel dataset on a buyer-supplier transaction network in Japan.

This paper investigates a number of features of transaction networks, their relationship with firm growth, and buyer-supplier comovements of sales using a comprehensive dataset on the Japanese inter-firm transaction networks. The unique data provide information on buyer, supplier, and ownership links of Japanese firms as well as their sales, profit, number of employees, industry classification, location, and so forth. For each firm, we can identify its suppliers, buyers and owners up to 24 firms. Although the truncation limit of 24 seems restrictive, it is possible to capture the entire production network quite well by merging self-reported and other-reported data. In this manner, we can see that some firms have thousands of transaction partners becoming a hub of the production network. The comprehensive nature of the data enables us to examine the detailed mechanism of shock propagations.

In the data, larger firms have higher sales growth rates and smaller growth dispersion. Well-connected firms also exhibit higher growth rates, but there is no systematic relationship between the number of partners (degree) and sales growth dispersion. Using a statistical test for spatial interdependence, it is confirmed that there exists a significant network interdependence of sales growth. By employing spatial autoregressive models, various propagation factors are estimated. In the baseline specification, the elasticity of average sales growth of suppliers is estimated to be 0.153 while that of customers is 0.257 for year 2012. In all years, the upstream propagation factor is larger than the downstream factor implying a difficulty of replacing an existing customer or adjusting to a demand shock. The manufacturing sector is characterized by a large degree of propagation. For both downstream and upstream propagations, manufacturing and wholesale sectors exhibit higher propagations factors while retail and service sectors exhibit lower propagation factors. The interdependence of intermediate physical inputs produced by other firms may generate an additional margin for the buyer-supplier comovements. It was also found that larger firms have higher propagation factors. Larger firms have more partners, and their degree of propagation is also higher. This result stresses an even larger impact of big firms for aggregate fluctuations in a granular production network.

This paper is related to several strands of literature. Some authors study the structure of the Japanese production networks. Saito et al. (2007) show the strong positive relationship between size and the number of links. Bernard et al. (2014, 2015) investigate the geography and firm performance using the same dataset. Carvalho et al. (2014) quantify the spillover effect of earthquakes on other firms through the supply chain network. Unlike their work, this paper focuses on buyer-supplier sales comovements rather than the effect of some exogenous shocks. The main purpose of the current work is to investigate more general patterns of shock propagations for different groups in the economy. Recent studies demonstrate that idiosyncratic shocks to heterogeneous firms may give rise to a sizable impact in aggregate fluctuations. Gabaix (2011) argues that a "granular" effect may be important when idiosyncratic volatilities are fixed and a firm size distribution is fat-tailed. Acemoglu et al. (2012) and others examine the propagation mechanism of idiosyncratic shocks through a network structure of an economy. The aggregate effect of firm-level shocks depends on the size distribution of shocks as well as a buyer-supplier network structure. Kelly et al. (2013) also consider customer-supplier connectedness to study the link between firm size distribution and firm volatility distribution. They show that the sales network structure is an important determinant of firm-level volatility. Mizuno et al. (2015) document a number of stylized facts about Japanese transaction networks, and report that firms' sales growth rates are more correlated if they locate closer in the transaction network.

The rest of the paper is organized as follows. Next section elaborates the data on the Japanese production networks in detail and presents figures to show the relationship between network statistics and firm sales growth. Section 3 contains the results of main empirical analyses on the shock propagation factors and how shocks are propagated through the production network in various specifications, and Section 4 concludes.

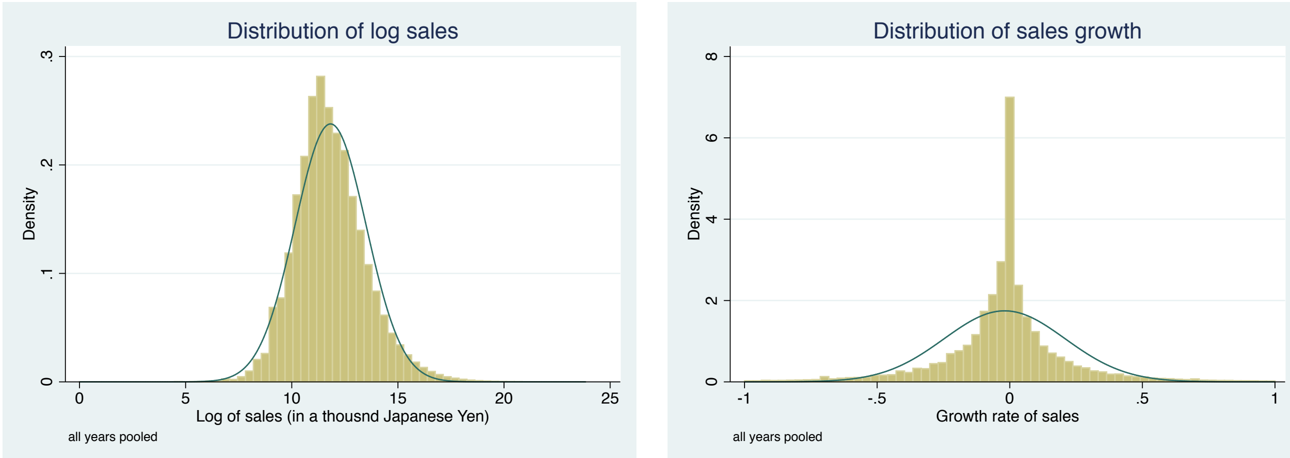
2 Data

2.1 Interfirm Network Data

The data on firm demographics and buyer-supplier networks come from Tokyo Shoko Research (henceforth TSR). TSR is a credit reporting company, which collects detailed information on Japanese firms to assess their credit scores. Firms provide their information in the course of obtaining credit reports on potential suppliers and customers or when attempting to qualify as a supplier. The information is updated at an annual frequency, and the datasets compiled in 2006, 2011, and 2012 are provided to the author by RIETI. On average, the TSR dataset covers about a million firms from all sectors. Carvalho et al. (2014) compare the firm size distributions of TSR and census data, and report that TSR data is under-sampled in very small firms. Nonetheless, the TSR data captures almost all economic activities in Japan. This comprehensive nature of the data, unlike the data of listed firms or only manufacturing sectors, allows us to directly investigate the propagation of shocks to small firms and its macroeconomic implications.

The firm-level data are retrieved from the TSR Company Information Database. For each firm, we have the information on its name, company code, address of headquarters, four-digit Japanese Standard Industrial Classification (JSIC) code, year of establishment, credit scores, number of employees, sales and profits of the most two recent periods available. Since the information on sales and profits for some firms is outdated, I discard firms whose most recent fiscal term is more than two years old for the analysis in this section. I also drop firms whose fiscal duration is not 12 months.

The unique feature of this dataset is the large-scale inter-firm network data, which is contained in the TSR Company Linkage Database. Firms report their suppliers, customers, and major shareholders up to 24 firms. Despite this truncation threshold, we can capture the interfirm network quite well by merging self- and other-reported data. For instance, many firms may identify one large firm as their customer. From the large firm's perspective, those firms are other-reported suppliers. In this way, we record more than 24 partners for some firms. Indeed, a small number of hub firms have several thousand links. For instance, the data in 2012 reveal that the top supplier has more than 12,000 customers. Following the literature on networks and graph theory, I use a term "link" for a transaction relationship and "node" for a firm interchangeably. The transaction network is a directed graph since



(a) Sales (b) Sales growth rate

Figure 1: Histograms of sales and sales growth

links are distinguished by suppliers and customers. For each firm, in-degree (out-degree) is defined as the number of suppliers (customers). In the data for 2006, there are 2 million supplier links and 1.9 million customer links. Table 9 displays the number of observations for each item by year. Note that the link information is only binary (whether a link exists or not) and does not give the dollar amount of the transaction. Nevertheless, the 2006 data contain the ranking of partners by their importance. This information can be used to weight the transaction links from a reporter's perspective. We can also put different weights on self- and other-reported links depending on the specifications.

2.2 Sales and Degree Distributions

The distributions of both log of sales and sales growth rates are well-approximated by a normal distribution as shown in Figure 1. Both histograms are fitted by a normal density based on the sample mean and variance. The fat-tailed distribution of firm size is observed in many countries, and indicates that a small number of large firms are responsible for a majority of economic activities. There is a disproportionately large mass at zero in the sales growth distribution. This may be due to coarse rounding of sales figures. For some firms, a small change in sales is rounded up, and hence, the growth rate is exactly zero. If we have finer data, the sales growth distribution is expected to resemble a normal distribution. Table 1 summarizes the sample statistics of sales growth.

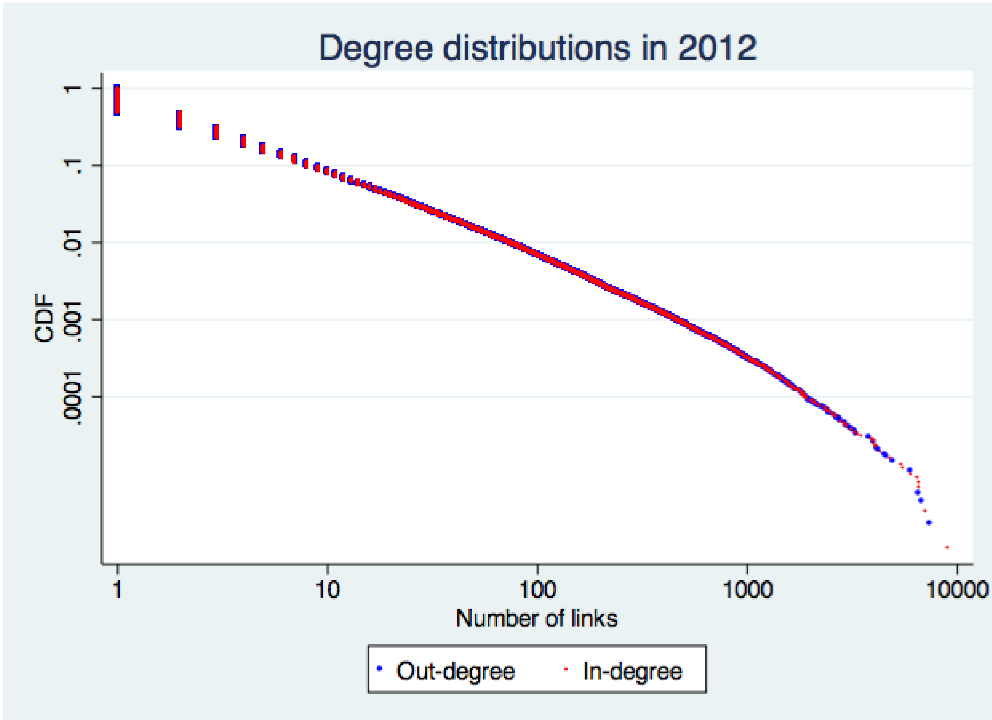


Figure 2: Degree distributions in 2012

year	mean	sd	p1	p25	p50	p75	p99
2006	0.0009935	0.3023747	-0.8350906	-0.0669117	0	0.064539	0.8785505
2011	-0.0497141	0.3512851	-1.098612	-0.1355257	-0.0171247	0.0339022	0.9346428
2012	-0.0087587	0.3422133	-0.9981375	-0.0901642	0	0.0689926	0.9757528
All pooled	-0.0203978	0.3353348	-0.9947567	-0.0988455	0	0.0557003	0.9328775

Table 1: Summary statistics of sales growth by year

For all years pooled, the mean is slightly negative, the median is zero, and the standard deviation is 0.335.

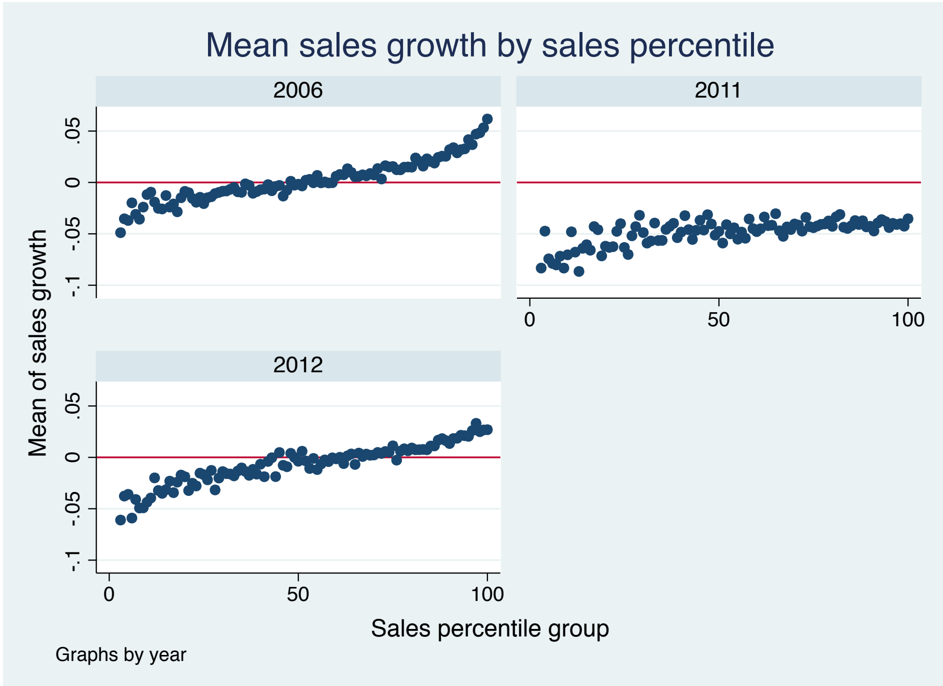
Figure 2 plots out-degree (number of customers) and in-degree (number of suppliers) distributions for 2012. Like many other social and economic networks, the log-log plot of the degree and its cdf exhibits a linear relationship implying that the degree distribution can be well approximated by a Pareto distribution. Bernard et al. (2014) report that the estimated Pareto shape parameter is -1.50 for out-degree and -1.32 for in-degree for the 2006 data. Some firms have thousands of transaction partners while large majority of firms have only one or two partners. As documented by Ohnishi et al. (2010) and others, the Japanese production network is characterized by a number of features such as a "small world" property, a small clustering coefficient, and negative assortativity.

2.3 Size, Degree and Firm Volatility

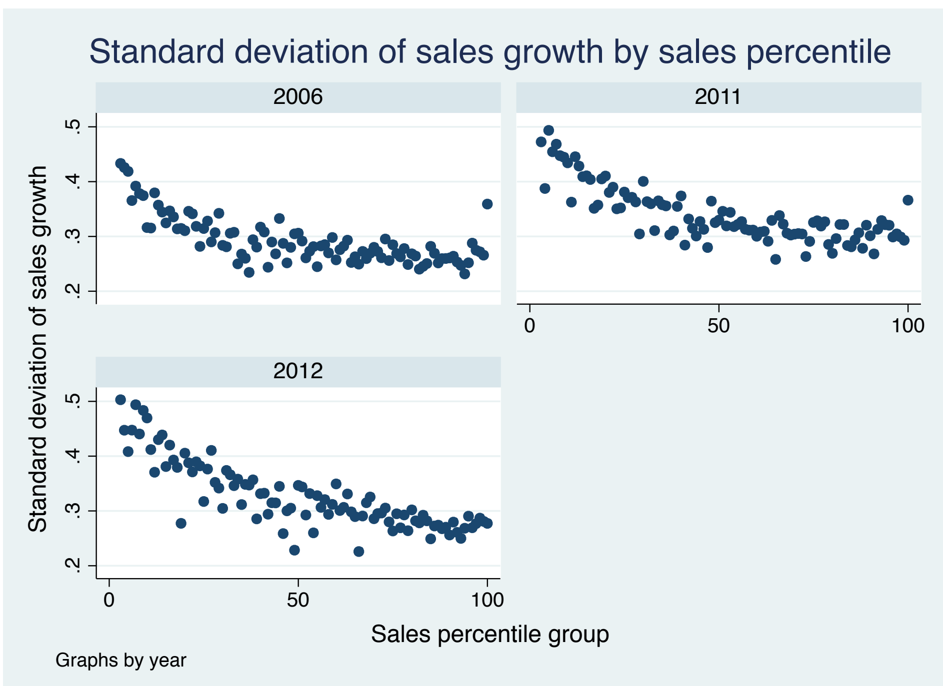
Is there any systematic pattern between firm growth and size? To answer this question, I divide firms into percentile groups based on the average sales between two consecutive years (Group 100 being the largest firms) for each year, calculate the within-group mean and standard deviation of sales growth, and examine their relationships with size. Figure 3 presents the scatter plots of the mean and standard deviation of firm growth against size groups for each year. [1] There is a striking pattern between sales growth and size. As size increases, the mean of sales growth becomes higher whereas the growth dispersion decreases, but both at a decreasing rate. Larger firms have a higher growth rate and smaller volatility, but this pattern is stark among the bottom half group of sales size. For firms in the top half group, the slope coefficient is small or insignificant. Larger firms achieve higher growth rates maybe because they have more market power, better access to credit, and other economic advantages. Smaller firms include those who received a sequence of adverse shocks, so their average growth rates are lower. Also, larger firms are expected to have a smaller volatility due to risk diversification on their subunits of economic activities. If a firm has many factories, the factory-level idiosyncratic shocks are averaged out.

Next, I conduct the same exercise for degree percentile groups. Firms are divided into 100 bins based on the number of customers (out-degree) and suppliers (in-degree). Figure 4 shows the two scatter plots for mean and volatility. Since there is a large number of firms with only one transaction partner, the first percentile groups contain disproportionately large mass of firms, and the next group is the 49th percentile. Similar to the previous exercise, there is a positive relationship between the mean of sales growth and the number of transaction partners. We cannot confirm a clear pattern between the growth dispersion and degree as displayed in the bottom plot. In-degrees of 2011 and 2012 exhibit a negative relationship at a 5 % significance level, but there is no significant relationship for 2006 and out-degrees. In a model of shock propagation where a firm's growth rate in part depends on its transaction partners' growth rates, a well-connected firm is expected to have a small volatility due to the law of large numbers. This pattern is not confirmed in the Japanese production network. It is interesting that the scatter dots of in-degree percentile groups exhibit systematically higher volatility than those of out-degree groups. On average, a firm's sales volatility tends to be

[1] The two smallest groups exhibit very low mean and large standard deviation. In the figure, these two outliers are removed to display an appropriate scale for the rest of the data.



(a) Mean of sales growth and size



(b) Standard deviation of sales growth and size

Figure 3: Firm growth and size

	mean	sd	p25	p75
supplier's relative size	0.209845	2.84482	-1.480199	1.912449
customer's relative size	0.2224562	2.842383	-1.391742	1.860825
pooled	0.2160606	2.843625	-1.436194	1.886274

Table 2: Summary statistics of the relative size of partners $r_{i,k}$

smaller if it has many customers, not suppliers, conditional on the number of transaction partners.[2] This may imply a stronger influence of downstream propagation.

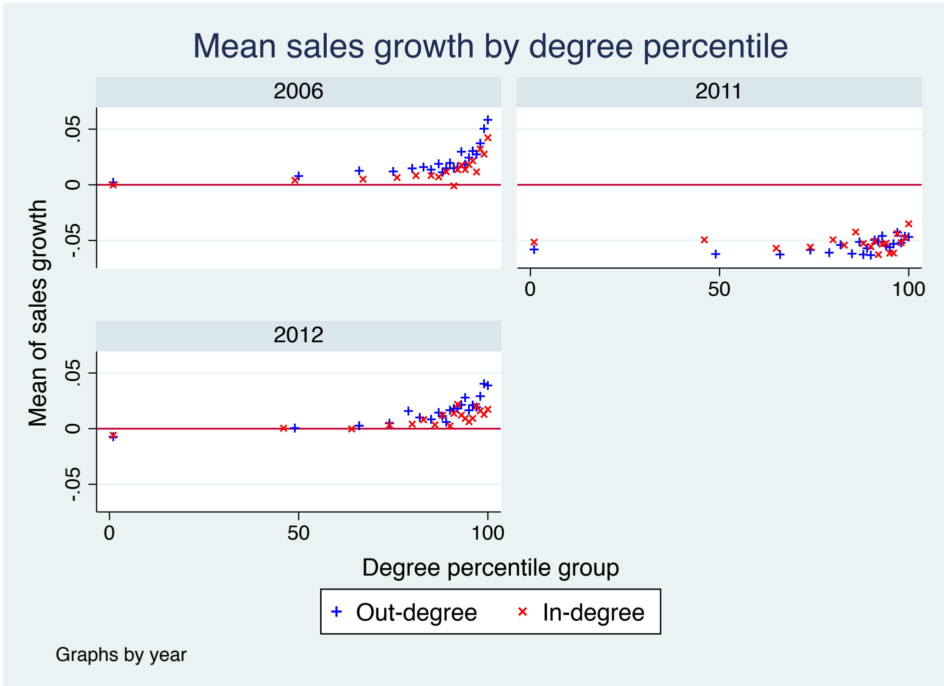
Figure 5 illustrates the relationship between degree and size for all years pooled. Both out- and in-degrees are positively correlated with size. The elasticity of the number of customers (out-degree) with respect to sales is 0.36 whereas that of the number of suppliers (in-degree) is 0.35. There is a positive and significant correlation between sales and the number of transaction partners, but it's not perfect (R-squared is 0.32 and 0.37 respectively). In general, well-connected firms are very large in size, but the reverse does not hold necessarily. There are some very large firms with less than 10 links.

2.4 Size and Link Weight

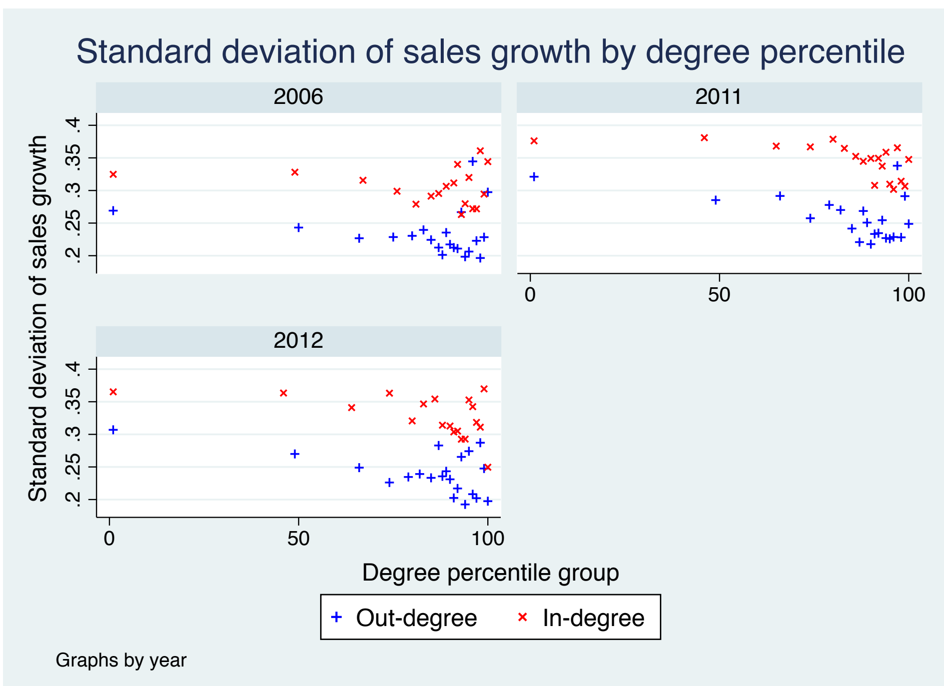
Although the data do not report the dollar amount of transactions, the data for 2006 report the ranking of transaction partners. Firms that report more than one partner order their partners by importance (one being the most important). This information of relative importance represents ordinal link weights (not cardinal since we only have rankings) for each transaction. I now examine the relationship between size and link weights. For firm i, denote the sales of the k-th most important partner by $S_{i,k}$. Define the k-th relative size of partners as $r_{i,k} = \ln\left(\frac{S_{i,k}}{S_{i,k+1}}\right)$. [3] Table 2 summarizes the statistics of $r_{i,k}$ for all available i and k. On average, a transaction partner is 21% larger in size than the next important partner. The magnitude is similar for suppliers and customers, but the standard deviation is large for both. There are many cases where more important partners are smaller in size. Figure 6 plots the average relative size for each available rank. In general, more important partners are larger since most of them lie above zero. The magnitude of the relative size is decreasing as we go down the ranking. After the 15th rank, there some cases the average size

[2] The degree distribution is almost identical for in- and out-degrees, so this systematic difference is not due to the average number of links for each percentile group.

[3] The sales data of some firms are not available. If $S_{i,k+1}$ is missing, it is replaced by the next important partner such as $S_{i,k+2}$ or $S_{i,k+3}$ to compute $r_{i,k}$.



(a) Mean of sales growth and degree



(b) Standard deviation of sales growth and degree

Figure 4: Firm growth and degree

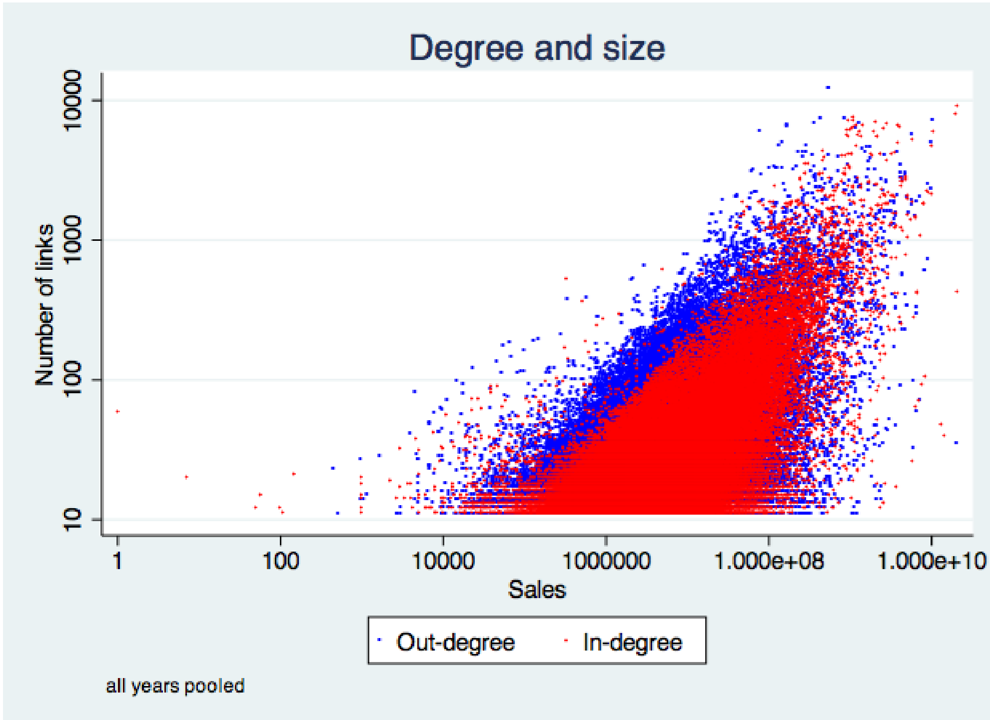


Figure 5: Degree and size

of the more important partner is smaller.

3 Buyer-Supplier Comovements in a Production Network

This section investigates buyer-supplier sales comovements in the Japanese production network. It is important to quantify the magnitude of the comovements and the pattern of shock propagation, which determines the importance of idiosyncratic shocks at a macro level as shown in Acemoglu et al. (2012). There are many studies that estimate the degree of shock propagation using industry-level I-O matrix or international I-O table for intermediate goods, but little research has quantified the buyer-supplier comovements using a large-scale production network data.

Due to the network feature of the model, estimation techniques developed in spatial econometrics literature are applied. To guide empirical analyses, consider a simple model of firm growth which captures spillover effects from transaction partners. A firm's sales growth rate is determined by an aggregate shock (common to all firms in a given year), firm-specific idiosyncratic shock, and the growth rates of its suppliers and customers. Both upstream and downstream propagations are examined. These shock propagations are the sources of an

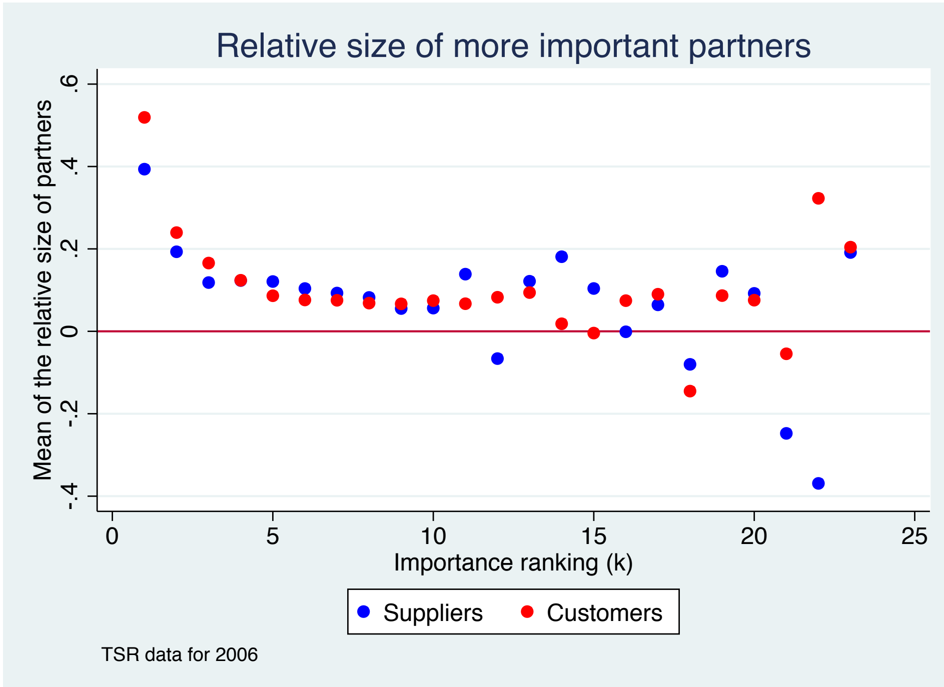


Figure 6: Size and link weight

endogenous systemic risk. Let $y_{it} = \ln\left(\frac{sales_{it}}{sales_{it-1}}\right)$ be the sales growth rate of firm i in year t. Since only cross-sectional data are available, the time subscript will be dropped. Denote the set of firm i's customers by Ω_i^C and the number of customers by N_i^C. Similarly, define Ω_i^S and N_i^S for suppliers.The sales growth rate is modeled as

$$y_i = \mu + \lambda \sum_{j=1}^{N} w_{ij} y_j + \epsilon_i$$

where N is the number of all firms, μ is the common shock, and λ is a spillover parameter which controls the rate of propagation through the network. A firm-specific sales shock is denoted by ϵ_i, and is assumed to be a white noise with $E(\epsilon) = 0$. The network weight w_{ij} governs the degree of influence of firm j on firm i. The full matrix of w_{ij} is denoted by $\boldsymbol{W}$. There are multiple ways to define these weights. A customer network is characterized by the weights $w_{ij} > 0$ if j is a customer of i ($j \in \Omega_i^C$), and $w_{ij} = 0$ otherwise. This network measures upstream propagations. By convention, diagonal elements are assumed to be zero ($w_{ii} = 0$ for all i), and the weights are row-normalized as $w_{ij} = \frac{1}{N_i^C}$ so that $\sum_j w_{ij} = 1$ for all firms. This implies that the intensive margin of the upstream propagation is the same across all customers since the TSR data only report the extensive margin of the transaction network (whether firms are connected or not). For now, I don't distinguish self-reported and other-reported customers as explained in Section 2. Later, more general specifications are fitted to accommodate different parameters for these two types of partners. A supplier network is defined analogously. We obtain this reduced form expression by building a multi-sector general equilibrium model with Cobb-Douglas technology and the I-O matrix of intermediate inputs. The share of intermediate goods is reflected in λ and the Cobb-Douglas expenditure share parameters show up as w_{ij}. The parameters λ and $w_{i,j}$ determine the magnitude and path of shock propagation.

As discussed later, one cannot estimate the parameters in this model consistently by the standard ordinary least squares (OLS) estimation due to the dependence of y_i on $\sum_j w_{ij} y_j$. Hence, I employ spatial econometric tools, which are originally developed to study the interdependence of geographic units such as counties in the U.S. In those models, the spatial weights w_{ij} are given by the inverse distance between i and j, or a contiguity matrix of sharing borders. This concept of spatial weights can be extended to our transaction network of firms. The terms "spatial" and "network" are used interchangeably. Due to the size of the matrix, an instrumental variable (IV) approach coupled with generalized method of moments (GMM) is applied rather than maximum likelihood estimation (MLE).

For the analyses in this section, firms whose information was updated in the latest year are selected and studied. For example, in the 2006 data, I extracted firms whose most recent fiscal ending date is between April 2005 and March 2006. The same criterion was applied for the years 2011 and 2012. Firms whose fiscal duration is not 12 months or whose sales data in the previous year is not available are dropped from the sample. This gives 300518, 578712, and 708288 firms for the years 2006, 2011, and 2012 respectively.

3.1 Network Interdependence of Sales Growth

Before conducting more solid analyses on buyer-supplier sales comovements, we need to test whether firm-level sales growth rates exhibit any network interdependence. For this purpose, Moran's I is computed. This statistic is defined as

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} \left(y_i - \bar{y}\right)\left(y_j - \bar{y}\right)}{\sum_i \left(y_i - \bar{y}\right)^2}$$

where N is the number of observations, y_i is the sales growth of firm i, $\bar{y}$ is the mean of y, and w_{ij} is the network weight of firm j for firm i. This statistic was proposed in Moran (1950), and has been widely used to test the existence of spatial dependence. Under the null

Year	Moran's I	z score	p-value	N
2006	0.0361	16.48	$< 10^{-6}$	300518
2011	0.0327	25.52	$< 10^{-6}$	578712
2012	0.0271	25.90	$< 10^{-6}$	708288

Table 3: Moran's I for the supplier network

hypothesis of no spatial interdependence, the statistic has the following expectation

$$E(I) = -\frac{1}{N-1}$$

Its variance $V(I)$ can be computed (see Appendix A.1), and the statistic $z = \frac{I-E(I)}{\sqrt{V(I)}}$ follows a standard normal distribution asymptotically. Thus, we can compute a p-value to test the null hypothesis. Let $\boldsymbol{y}$ be the column vector of y_i and define the detrended vector of y as $\boldsymbol{x} = \boldsymbol{y} - \bar{y}$. With the row normalization of the network matrix $\boldsymbol{W}$, we can simplify the above expression as $I = \frac{\boldsymbol{x'Wx}}{\boldsymbol{x'x}}$.

Table 3 shows the statistics for the supplier network. The p-values are smaller than 10^{-6} in all years implying the significant network interdependence of sales growth. Similar results are obtained for the customer network. The network sales growth is far from random, and we confirm that a firm's sales growth is correlated with the sales growth rates of its suppliers and customers.

3.2 Estimation of the Propagation Factor

Let $\boldsymbol{y}$ be the vector of detrended sales growth, and consider the following simple model of shock propagation

$$\boldsymbol{y} = \lambda \boldsymbol{W}\boldsymbol{y} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon}$ is a vector of i.i.d idiosyncratic shocks with zero mean and a finite variance σ^2. This is one of the most widely used models of spatial autocorrelation, and is referred to as a spatial autoregressive (SAR) model. To consistently estimate the parameters λ and σ^2, one cannot run a simple OLS regression of $\boldsymbol{y}$on $\boldsymbol{W}\boldsymbol{y}$.[4] Typically, MLE with normally distributed ϵ is employed to consistently estimate the parameters in a SAR model as shown

[4]The OLS estimator of λ is computed as $\hat{\lambda} = (\boldsymbol{y'W'Wy})^{-1}\boldsymbol{y'W'y}$. Its expectation is

$$\begin{aligned} E\left(\hat{\lambda}\right) &= E\left[(\boldsymbol{y'W'Wy})^{-1}\boldsymbol{y'W'}(\rho\boldsymbol{W}\boldsymbol{y}+\boldsymbol{\epsilon})\right] \\ &= \lambda + E\left[(\boldsymbol{y'W'Wy})^{-1}\boldsymbol{y'W'}\epsilon\right] \end{aligned}$$

	$\hat{\lambda}$			$\hat{\sigma^2}$		
	supplier	customer	both	supplier	customer	both
2006	0.043	0.060	0.043	0.077	0.080	0.078
2011	0.050	0.073	0.054	0.104	0.108	0.106
2012	0.038	0.059	0.047	0.103	0.108	0.105

Table 4: Estimated propagation factors

in Cliff and Ord (1973, 1981) and many others. One of the limitations of the MLE is its computational infeasibility when the sample size is large. The log-likelihood function requires the determinant of $I - \lambda \boldsymbol{W}$ whose operation counts grow at a rate proportional to N^3.[5] Obviously, the MLE is not practical for our problem where N is more than half a million. Thus, I take the GMM approach suggested by Kelejian and Prucha (1998). The moment conditions and the definition of the GMM estimators is described in Appendix A.2. The propagation factor λ is estimated for three types of network matrix $\boldsymbol{W}$: supplier, customer, and both ($w_{ij} = 1$ if j is either i's supplier or customer). Throughout this section, all network matrices are row-normalized, so the vector $\boldsymbol{W}\boldsymbol{y}$ corresponds to the average sales growth rate of partners. Table 4 displays the estimated values of λ and σ^2. In all years, the upstream propagation factor is higher than the downstream propagation factor. A firm's sales growth is correlated with the average sales growth of both its suppliers and customers, and the effect is stronger for customers.

Now, I consider a more general specification

$$\boldsymbol{y} = \lambda \boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\beta + \boldsymbol{\epsilon}$$

where $\boldsymbol{y}$ and $\boldsymbol{W}$ are the same as before, and $\boldsymbol{X}$ is the matrix of exogenous variables that

The second term is not zero, and hence, the OLS estimator is biased. See LeSage (1999) for more details. Anselin (1988) also shows

$$plim \frac{1}{N} \left(\boldsymbol{y}'\boldsymbol{W}'\boldsymbol{\epsilon} \right) = plim \frac{1}{N} \boldsymbol{\epsilon}'\boldsymbol{W} \left(\boldsymbol{I} - \boldsymbol{\lambda} \boldsymbol{W} \right)^{-\mathbf{1}} \boldsymbol{\epsilon}$$

Thus, the OLS estimator is not consistent.

[5]When ϵ is normally distributed with zero mean and σ^2, the log likelihood function is

$$\ln \mathcal{L} = -\frac{N}{2} \left[\ln 2\pi\sigma^2 \right] - \frac{1}{2\sigma^2} \boldsymbol{y}' \left(I - \lambda \boldsymbol{W} \right)' \left(I - \lambda \boldsymbol{W} \right) \boldsymbol{y} + \ln \| I - \lambda \boldsymbol{W} \|$$

where the last term is the log determinant. Ord (1975) proposes the use of eigenvalues to compute the determinant. Pace and Berry (1997) suggest quick computation of spatial autoregressive estimators by exploiting sparsity and rearranging the rows of $\boldsymbol{W}$. While these methods can accelerate the estimation time significantly for a mid-sized matrix, the computational burden is still a serious problem for the size of our matrix.

	2006		2011		2012	
	supplier	customer	supplier	customer	supplier	customer
$\hat{\lambda}$	0.180***	0.258***	0.134***	0.155***	0.153***	0.257***
	(0.0207)	(0.0211)	(0.0115)	(0.0121)	(0.0116)	(0.0139)
employment	0.008***	0.009***	0.004***	0.002***	0.005***	0.004***
	(0.0009)	(0.001)	(0.0006)	(0.0006)	(0.0005)	(0.0005)
credit score	0.170***	0.175***	0.347***	0.352***	0.310***	0.314***
	(0.008)	(0.009)	(0.006)	(0.006)	(0.005)	(0.005)
age	-0.050***	-0.048***	-0.040***	-0.043***	-0.044***	-0.045***
	(0.0014)	(0.0015)	(0.0009)	(0.0010)	(0.0008)	(0.0008)
constant	-0.478***	-0.507***	-1.241***	-1.242***	-1.052***	-1.057***
	(0.0323)	(0.0344)	(0.0230)	(0.0237)	(0.0191)	(0.0201)
2-digit JSIC FE	Yes	Yes	Yes	Yes	Yes	Yes
prefecture FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	126,983	104,330	320,535	286,971	430,995	375,540
R-squared	0.023	0.004	0.030	0.0316	0.028	0.013

Table 5: Mixed SAR Model

explain the variation of $\boldsymbol{y}$. This is called a mixed spatial autoregressive (SAR) model. The exogenous variables include the number of employees, credit score, age (all in logs), prefecture fixed effects (FE), and 2-digit JSIC industry FE. The IV estimation approach described in Kelejian and Prucha (1998) and Anselin (1988) is employed.[6] Unlike their model, I do not consider the spatial autocorrelation of the error term to reduce computational burdens, so ϵ_i is assumed to be independent with mean zero. The first and second order network matrices $\boldsymbol{W}$ and $\boldsymbol{W}^2$ are used to instrument the endogenous vector $\boldsymbol{W}\boldsymbol{y}$. Ideally, all higher order matrices should be included to approximate $(\boldsymbol{I}-\lambda\boldsymbol{W})^{-1}$, but given the estimated magnitude of λ, the marginal gain in precision of including more terms is very small. In the estimation, linearly independent columns of $[\boldsymbol{X},\boldsymbol{W}\boldsymbol{X},\boldsymbol{W}^{\boldsymbol{2}}\boldsymbol{X}]$ are used as instruments. Again, both types of the network matrix $\boldsymbol{W}$(supplier and customer) are considered. The results are shown in Table 5. In all specifications, employment and credit score have positive effects on sales growth whereas age is negatively correlated with it. In all years, the upstream propagation factor is larger than the downstream factor. In 2012, the elasticity of average sales growth of suppliers is 0.153 while that of customers is 0.257. The higher degree of upstream propagation implies that it is more difficult to replace an existing customer or adjust to a demand shock.

[6] I used the MATLAB functions developed in Spatial Econometrics Toolbox by LeSage (1999).

3.3 Group-Specific Propagation Factors

So far, we assumed the common propagation factor λ for all firms in the economy, but this factor can be heterogeneous across different groups of firms such as industries. For instance, consider two types of firms: manufacturing and non-manufacturing firms. Denote the manufacturing firms by group 1 and non-manufacturing firms by group 2. Consider the following system of SAR equations

$$\begin{aligned} \boldsymbol{y}_{1,m} &= \lambda_{11}\boldsymbol{W}_{11}\boldsymbol{y}_{1,m} + \lambda_{12}\boldsymbol{W}_{12}\boldsymbol{y}_{2,n} + \boldsymbol{X}_{1,m}\beta_1 + \epsilon_{1,m} \\ \ni \boldsymbol{y}_{2,n} &= \lambda_{22}\boldsymbol{W}_{22}\boldsymbol{y}_{2,n} + \lambda_{21}\boldsymbol{W}_{21}\boldsymbol{y}_{1,m} + \boldsymbol{X}_{2,n}\beta_2 + \epsilon_{2,n} \end{aligned}$$

where $\boldsymbol{y}_{1,m}$ represents the vector of sales growth of manufacturing firms and $\boldsymbol{y}_{2,n}$ represents the vector of sales growth of non-manufacturing firms. $\boldsymbol{W}_{11,m}$and $\boldsymbol{W}_{22,n}$ are the (square) network matrices for manufacturing and non-manufacturing firms respectively. The $m \times n$ matrix $\boldsymbol{W}_{12}$ shows the connection of manufacturing firms to non-manufacturing firms, and the $n \times m$ matrix $\boldsymbol{W}_{21}$ is defined analogously. In addition, $\boldsymbol{X}_{1,m}$ and $\boldsymbol{X}_{2,n}$ are the matrices of common exogenous variables. Since the sales growth vector of each type depends on both types of networks, the parameters must be estimated simultaneously. As shown in Bao (2010), the above system can be rewritten as the following

$$\begin{aligned} \begin{pmatrix} \boldsymbol{y}_{1,m} \\ \boldsymbol{y}_{2,n} \end{pmatrix} &= \lambda_{11} \begin{pmatrix} \boldsymbol{W}_{11} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_{1,m} \\ \boldsymbol{y}_{2,n} \end{pmatrix} + \lambda_{12} \begin{pmatrix} 0 & \boldsymbol{W}_{12} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_{1,m} \\ \boldsymbol{y}_{2,n} \end{pmatrix} \\ &+ \lambda_{21} \begin{pmatrix} 0 & 0 \\ \boldsymbol{W}_{21} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_{1,m} \\ \boldsymbol{y}_{2,n} \end{pmatrix} + \lambda_{22} \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{W}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_{1,m} \\ \boldsymbol{y}_{2,n} \end{pmatrix} \\ &+ \begin{pmatrix} \boldsymbol{X}_{1,m} & 0 \\ 0 & \boldsymbol{X}_{2,n} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,m} \\ \epsilon_{2,n} \end{pmatrix} \end{aligned}$$

By changing notations, this can be simplified to

$$y_N = \lambda_{11} W_1 y_N + \lambda_{12} W_2 y_N + \lambda_{21} W_3 y_N + \lambda_{22} W_4 y_N + X_N \beta + \epsilon_N$$

This equation is in the form of a high order SAR model. This model, sometimes referred to as SARMA model, has been studied by several researchers including Bloommestein (1983), Huang (1984), Anselin (2006), Lee and Liu (2010) and Gupta and Robinson (2015), and various estimators have been developed. Considering the number of parameters to be estimated, a simple OLS specification is employed in this subsection to minimize computational complexities. Kolympiris et al. (2011) list a number of justifications for the use of an OLS estimator in a similar context. Lee (2002) proved that OLS is consistent if the number of partners can become infinitely large as the sample size increases. Anselin (2006) shows that the OLS estimator is relatively robust to various model assumptions compared to the ML estimator. Based on Monte Carlo simulations, Franzese and Hays (2007) argue that the finite sample bias of the OLS estimator is reasonably small with at least 50 observations that have relatively small propagation factors ($\lambda < 0.3$). Given the magnitude of $\hat{\lambda}$ from the previous subsections, the potential bias associated with the OLS estimator is not large. Moreover, the focus of this subsection is to show differential degrees of propagation across different subgroups, and not the level of propagation factors per se. For expositional purpose, I only present the results for the year 2012, but similar results are obtained for other years as well.

First, I consider two types of groups: manufacturing and non-manufacturing firms.[7] In the 2012 data, about 16% of firms are in the manufacturing sector. The same set of exogenous variables as in the previous section is included. The estimated propagation factors are shown in Table 6, and the results for other variables are shown in the first two columns of Table 10 in Appendix. Denote manufacturing and non-manufacturing by M and N respectively. For both downstream and upstream propagations, the propagation factor of M-to-M firms is larger compared to the case of N-to-N. In the upper panel, the value of M-to-M is more than three times larger than that of M-to-N while the values of N-to-M and N-to-N are quite close. Manufacturing firms as suppliers do not give much impact on their non-manufacturing customers whereas non-manufacturing suppliers may have larger impact on their manufacturing customers. Cravino and Levchenko (2014) also found stronger sales comovements between multinational firms and their foreign affiliates in the manufacturing sector compared to the service sector. In manufacturing sectors, the interdependence of intermediate physical inputs produced by other firms may generate an additional margin for the buyer-supplier comovements.

The above model can be extended to incorporate more groups. If we divide firms into m groups, the total of m^2 propagation factors for all pairwise combinations are estimated. To further investigate the sectoral heterogeneity in the degree of propagation, firms are now divided into five sectors: 1) manufacturing, 2) construction, 3) wholesale, 4) retail, and 5) services. For the definition and shares of each sector, please see Table 11 in Appendix.[8] The

[7]Manufacturing firms are identified by their 2-digit JSIC code being between 09 and 32.

[8]Agriculture, food, and textiles are included in the manufacturing sector. Since their share is very small (only 1% of total number of firms), this is not an issue even though their nature of production is different from that of manufacturers.

From \ To	manufacturing	non-manufacturing
manufacturing	0.105*** (0.008)	0.028*** (0.004)
non-manufacturing	0.056*** (0.007)	0.051*** (0.003)

(a) downstream

From \ To	manufacturing	non-manufacturing
manufacturing	0.096*** (0.007)	0.055*** (0.005)
non-manufacturing	0.043*** (0.005)	0.038*** (0.002)

(b) upstream

Table 6: Propagation factors of manufacturing and non-manufacturing sectors

estimated propagation factors for each sectoral pair are displayed in Table 7and other results are shown in Appendix. Figure 7 presents the 3-D bar plots of the estimated propagation factors that are significant at least 10% level. For both types of propagations, the diagonal elements are all positive and significant implying the existence of within-sector propagations. The manufacturing-to-manufacturing pair has the highest propagation factor in both cases. For the downstream propagation, the manufacturing-to-wholesale pair has the second highest value. The elements of the third column (wholesale) of Table 7 (a) are all positive and significant. This means that the wholesale sector tends to be affected by suppliers, and the impact is largest if the supplier is a manufacturer, and smallest if the supplier is in the service sector. In the same table, the values of retail and service sectors (both rows and columns) are smaller compared to other sectors. Firms in these two sectors do not receive nor propagate downstream shocks to their transaction partners. The same pattern can be found in the upstream propagation. Manufacturing and wholesale sectors exhibit higher propagations factors while retail and service sectors exhibit lower propagation factors. The wholesale-to-manufacturing pair shows a relatively high value, which underscores the demand effect by wholesalers to manufacturers.

We can also estimate the propagation factors for different size groups. For this analysis, firms are divided into five groups based on their sales quintiles: smallest, small, medium, large and largest. With the same set of exogenous variables, the estimated propagation factors are shown in Table 8, and Figure 8 displays the graphical representation. In both types of propagations, the largest size group has higher propagation factors with he largest-to-largest

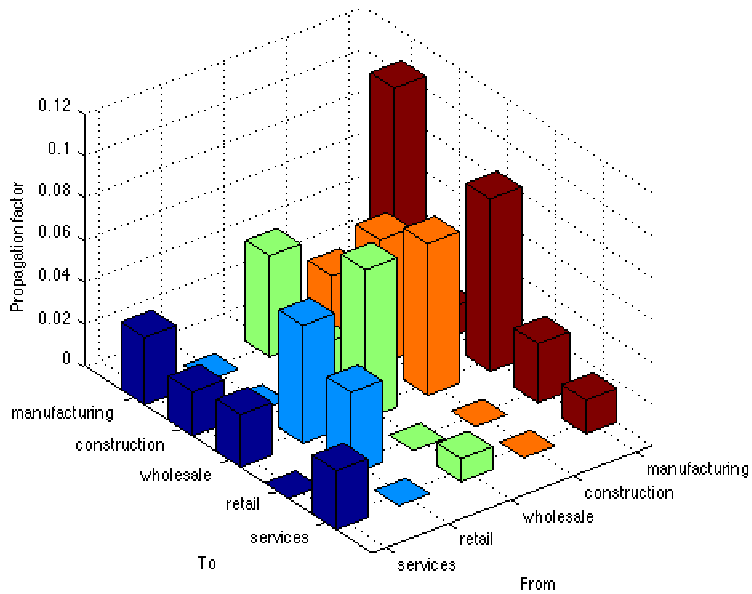
From \ To	manufacturing	construction	wholesale	retail	services
manufacturing	0.105*** (0.008)	0.014* (0.008)	0.082*** (0.008)	0.028* (0.016)	0.016*** (0.006)
construction	0.027*** (0.007)	0.059*** (0.004)	0.072*** (0.007)	0.006 (0.012)	0.008 (0.005)
wholesale	0.048*** (0.007)	0.020* (0.011)	0.071*** (0.008)	0.025 (0.017)	0.011* (0.006)
retail	0.004 (0.007)	0.004 (0.015)	0.056*** (0.010)	0.039** (0.018)	-0.004 (0.009)
services	0.032*** (0.009)	0.021*** (0.008)	0.025** (0.010)	0.008 (0.011)	0.028*** (0.006)

(a) downstream

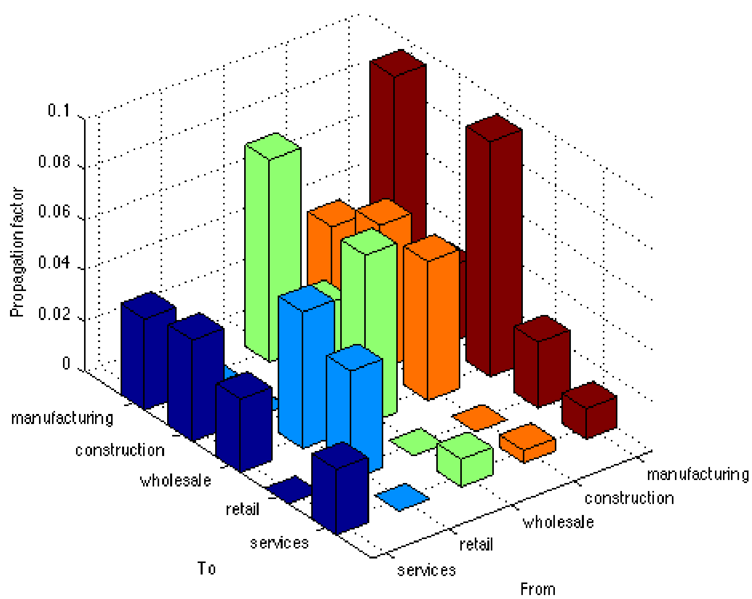
From \ To	manufacturing	construction	wholesale	retail	services
manufacturing	0.094*** (0.007)	0.030*** (0.010)	0.093*** (0.009)	0.026** (0.012)	0.012*** (0.004)
construction	0.044*** (0.009)	0.057*** (0.003)	0.055*** (0.011)	0.014 (0.017)	0.005* (0.003)
wholesale	0.080*** (0.010)	0.034*** (0.009)	0.067*** (0.010)	0.015 (0.012)	0.011** (0.005)
retail	0.027 (0.020)	-0.004 (0.015)	0.054** (0.026)	0.043* (0.025)	0.011 (0.013)
services	0.036*** (0.010)	0.040*** (0.011)	0.029** (0.014)	0.003 (0.018)	0.026*** (0.006)

(b) upstream

Table 7: Propagation factors of five sectors



(a) downstream



(b) upstream

Figure 7: Propagation factors of five sectors

From \ To	smallest	small	medium	large	largest
smallest	0.018* (0.009)	0.026*** (0.008)	0.032*** (0.007)	0.016** (0.006)	0.023*** (0.005)
small	0.025** (0.013)	0.065*** (0.010)	0.049*** (0.008)	0.038*** (0.007)	0.034*** (0.005)
medium	0.017 (0.013)	0.025** (0.011)	0.062*** (0.009)	0.058*** (0.008)	0.043*** (0.006)
large	0.028** (0.013)	0.030*** (0.011)	0.061*** (0.010)	0.058*** (0.008)	0.055*** (0.006)
largest	0.020*** (0.007)	0.035*** (0.006)	0.047*** (0.007)	0.062*** (0.006)	0.083*** (0.006)

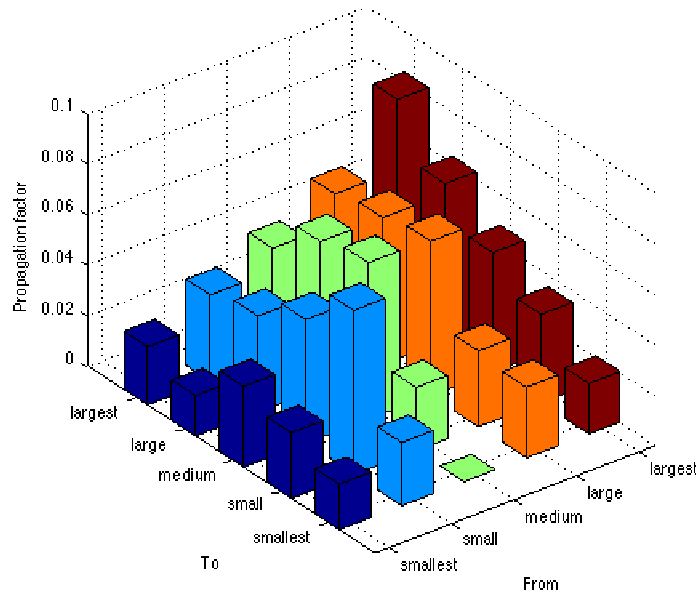
(a) downstream

From \ To	smallest	small	medium	large	largest
smallest	0.002 (0.006)	0.008 (0.007)	0.002 (0.007)	0.007 (0.006)	0.019*** (0.005)
small	0.014** (0.007)	0.035*** (0.007)	0.015* (0.008)	0.016** (0.007)	0.012** (0.006)
medium	0.041*** (0.008)	0.035*** (0.008)	0.040*** (0.008)	0.033*** (0.007)	0.029*** (0.007)
large	0.028*** (0.006)	0.016*** (0.005)	0.030*** (0.005)	0.025*** (0.005)	0.027*** (0.005)
largest	0.032*** (0.006)	0.049*** (0.006)	0.051*** (0.006)	0.071*** (0.006)	0.106*** (0.007)

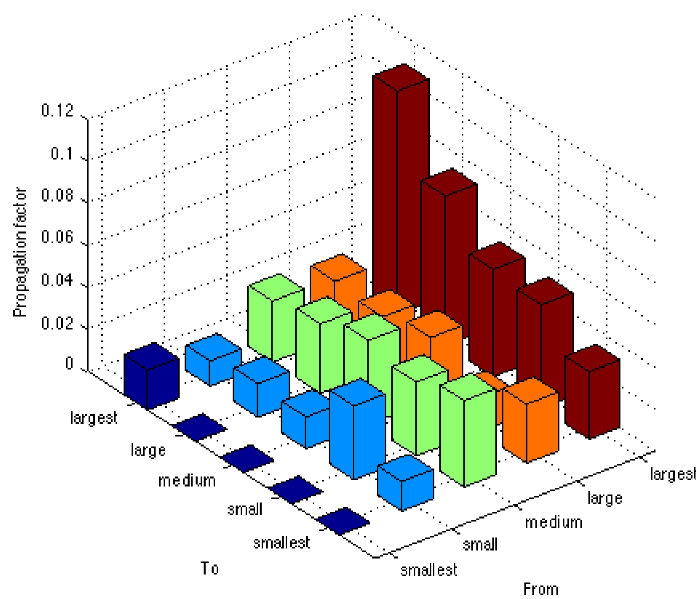
(b) upstream

Table 8: Propagation factors for different size groups

pair being the highest. If we focus on the row of the largest group, we see a decreasing pattern from the largest to the smallest. In general, larger firms have a relatively more impact on their partners, and the effect is stronger if the partners are also large. For the medium, small, and smallest groups, there is no such a monotonic relationship. From the rows of the smallest group in Table 8, it is apparent that small firms do not affect their partners both upstream and downstream directions. Large firms do not receive shocks if they come from small partners.



(a) downstream



(b) upstream

Figure 8: Propagation factors for different size groups

4 Conclusion

This paper investigates a number of features of transaction networks, their relationship with firm sales growth, and buyer-supplier comovements of sales using a large-scale dataset on the Japanese inter-firm transaction networks. The distributions of both log of sales and sales growth rates are well-approximated by a normal distribution, and the degree distribution can be well approximated by a Pareto distribution. Larger firms have a higher growth rate and smaller volatility, but this pattern is stark in the bottom half group of sales size. There is a positive relationship between the mean of sales growth and the number of transaction partners, especially among the top sales percentile groups. We cannot confirm a clear pattern between the growth dispersion and degree.

From the Moran's I statistic, a significant network interdependence of sales growth is confirmed for both supplier and buyer networks. After controlling various firm characteristics, industry and geographical fixed effects, the elasticity of average sales growth of suppliers is 0.153 while that of customers is 0.257 for 2012. In all years, upstream propagation is estimated to be stronger than downstream propagation. A firm receives more adverse shock when its customer's sales declines. This result implies that it is more difficult to replace an existing customer or adjust to a demand shock. It also suggests that we need to take into account the role of a well-connected firm as a customer when discussing a bailout policy. If a large customer exits from the market, the negative impact might be even larger for other firms and the economy as a whole at least in the short run.

From the estimation of group-specific propagation factors, it is clear that the manufacturing sector is characterized by a large degree of propagation. Because manufacturers trade physical inputs that are often times essential for production process, the network interdependence of intermediate goods may generate an additional margin for the buyer-supplier comovements. This results is consistent with other research such as Cravino and Levchenko (2015). For both downstream and upstream propagations, manufacturing and wholesale sectors exhibit higher propagations factors while retail and service sectors exhibit lower propagation factors. This result emphasizes a stronger propagation effect of government policies such as tax cuts, subsidies or bailout for manufacturing firms. It was also found that larger firms have higher propagation factors. Larger firms have more partners, and their degree of propagation is also higher. This result stresses an even larger impact of big firms for aggregate fluctuations in a granular production network.

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012) "The Network Origins of Aggregate Fluctuations," *Econometrica*, Vol. 80, No. 5, pp. 1977–2016.

Anselin, Luc (2006) *Spatial Econometrics, Palgrave Handbook of Econometrics*: Palgrave MacMilan, New York.

Bao, Yan (2010) "CMM Estimation of the Spatial Autoregressive Model in a System of Interrelated Networks," *Working Paper*.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito (2014) "Geography and Firm Performance in the Japanese Production Network," *RIETI Discussion Paper Series 14-E-034*.

——— (2015) "Production Networks, Geography and Firm Performance," *NBER Working Paper 21082*.

Blommestein, Hans J. (1983) "Specification and estimation of spatial econometric models: A discussion of alternative strategies for spatial economic modelling," *Regional Science and Urban Economics*, Vol. 13, No. 2, pp. 251 – 270.

Carvalho, Vasco M., Makoto Nirei, and Yukiko U. Saito (2014) "Supply Chain Disruptions: Evidence frmo the Great East Japan Earthquake," *RIETI Discussion Paper Series 14-E-035*.

Cliff, Andrew D. and Keith Ord (1973) *Spatial Autocorrelation*, London: Pion.

——— (1981) *Spatial Process: Models and Applications*, London: Pion.

Cravino, Javier and Andrei A. Levchenko (2015) "Multinational Firms and International Business Cycle Transmission," *Working Paper*.

Franzese, Robert J. and Jude C. Hays (2007) "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data," *Political Analysis*, Vol. 15, No. 2, pp. 140–164.

Gabaix, Xavier (2011) "The Granular Origins of Aggregate Fluctuations," *Econometrica*, Vol. 79, No. 3, pp. 733–772.

Gupta, Abhimanyu and Peter M. Robinson (2015) "Inference on Higher-Order Spatial Autoregressive Models with Increasingly Many Parameters," *Journal of Econometrics*, Vol. 186, No. 19 - 31.

Huang, Jun S. (1984) "The Autoregressive Moving Average Model for Spatial Analysis," *Australian Journal of Statistics*, Vol. 26, No. 2, pp. 169–178.

Kelly, Bryan, Hanno Lustig, and Stijn Van Nieuwerburgh (2013) "Firm Volatility in Granular Networks," *Working Paper*.

Kolympiris, Christos, Nicholas Kalaitzandonakes, and Douglas Miller (2011) "Spatial Collocation and Venture Capital in the US Biotechnology Industry," *Research Policy*, Vol. 40, No. 9, pp. 1188 – 1199.

Lee, Lung-fei (2002) "Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregressive Models," *Econometric Theory*, Vol. 18, No. 2, pp. 252–277, cited By 67.

Lee, Lung-fei and Xiaodong Liu (2010) "Efficient GMM Estimation of High Order Spatial Autoregressive Models with Autoregressive Disturbances," *Econometric Theory*, Vol. 26, pp. 187–230.

Mizuno, Takayuki, Wataru Souma, and Tsutomu Watanabe (2015) "Buyer-Supplier Networks and Aggregate Volatility," *RIETI Discussion Paper Series 15-E-056*.

Saito, Yukiko U., Tsutomu Watanabe, and Mitsuru Iwamura (2007) "Do Larger Firms Have More Interfirm Relationship?" *RIETI Discussion Paper Series 07-E-028*.

Appendix

A Mathematical Formula

A.1 The Variance of Moran's I

The variance of Moran's I has the following expression

$$V(I) = \frac{NS_4 - S_3 S_5}{(N-1)(N-2)(N-3)\left(\sum_i \sum_j w_{ij}\right)^2} - \left(\frac{1}{N-1}\right)^2$$

where

$$\begin{aligned}
S_1 &= \frac{1}{2}\sum_i \sum_j (w_{ij} + w_{ji})^2 \\
S_2 &= \sum_i \left(\sum_j w_{ij} + \sum_j w_{ji}\right)^2 \\
S_3 &= \frac{N^{-1}\sum_i (y_i - \bar{y})^4}{\left(N^{-1}\sum_i (y_i - \bar{y})^2\right)^2} \\
S_4 &= \left(N^2 - 3N + 3\right) S_1 - NS_2 + 3\left(\sum_i \sum_j w_{ij}\right)^2 \\
S_5 &= \left(N^2 - N\right) S_1 - 2NS_2 + 6\left(\sum_i \sum_j w_{ij}\right)^2
\end{aligned}$$

A.2 Moment Conditions of the FAR Model

Consider the matrix and vector

$$G = \frac{1}{N}\begin{bmatrix} 2y'\bar{y}' & -\bar{y}'\bar{y} & 1 \\ 2\bar{\bar{y}}'\bar{y} & -\bar{\bar{y}}'\bar{\bar{y}} & Tr\left(W'W\right) \\ (\bar{y}'\bar{\bar{y}} + \bar{y}'\bar{y}) & -\bar{y}'\bar{\bar{y}} & 0 \end{bmatrix} \text{ and } g = \frac{1}{N}\begin{bmatrix} y'y \\ \bar{y}'\bar{y} \\ y'\bar{y} \end{bmatrix}$$

where $\bar{y} = Wy$ and $\bar{\bar{y}} = W^2 y$. The estimated coefficients $\tilde{\lambda}$ and $\tilde{\sigma}^2$ are defined as the minimizers of

$$\left[g - G\begin{bmatrix} \lambda \\ \lambda^2 \\ \sigma^2 \end{bmatrix}\right]' \left[g - G\begin{bmatrix} \lambda \\ \lambda^2 \\ \sigma^2 \end{bmatrix}\right]$$

More details can be found in Kelejian and Prucha (1998).

B Tables

	2006	2011	2012
Total number of firms	807,727	1,161,096	1,193,283
Sales	802,584	1,061,706	1,089,596
Profits	594,542	724,546	736,800
Number of employees	801,729	1,078,992	1,106,667
# of supplier links	2,005,948	2,454,035	2,508,972
# of customer links	1,898,432	2,707,173	2,799,553

Table 9: Number of observations (firms or transactions) by year

	manu vs. non-manu		five sectors		five size groups	
	supplier	customer	supplier	customer	supplier	customer
λs	shown in text	shown in text	shown in text	shown in text	shown in text	shown in text
employment	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
credit score	0.345***	0.346***	0.345***	0.345***	0.344***	0.345***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
age	-0.047***	-0.047***	-0.047***	-0.047***	-0.047***	-0.047***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
constant	-1.177***	-1.179***	-1.173***	-1.176***	-1.170***	-1.176***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Observations	615,998	615,998	615,998	615,998	615,998	615,998
R-squared	0.030	0.030	0.030	0.030	0.030	0.030
2-digit JSIC FE	Yes	Yes	Yes	Yes	Yes	Yes
prefecture FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 10: Estimation results of the high order SAR models in subsection 3.3

	share	definition
manufacturing	17.2%	2-digit JSIC between 09 and 32, or between 01 and 05
construction	34.3%	2-digit JSIC between 06 and 08
wholesale	14.1%	2-digit JSIC between 50 and 55
retail	11.9%	2-digit JSIC between 56 and 61
services	22.5%	all other firms

Table 11: Share and definition of each sector