

Chapter 2

Buyer-Supplier Networks and Aggregate Volatility

Takayuki Mizuno, Wataru Souma and Tsutomu Watanabe

Abstract This chapter investigates the structure and evolution of customer–supplier networks in Japan using a unique dataset that contains information on customer and supplier linkages for over 500,000 incorporated non-financial firms for the 5 years from 2008 to 2012. We find, first, that the number of customer links is unequal across firms: the customer link distribution has a power-law tail with an exponent of unity (i.e., it follows Zipf’s law). We interpret this as implying that competition among firms to acquire new customers yields winners that attract a large number of customers, as well as losers that end up with fewer customers. We also show that the shortest path length for any pair of firms is, on average, 4.3 links. Second, we find that link switching is relatively rare. Our estimates indicate that 92 % of customer links and 93 % of supplier links survive each year. Third and finally, we find that firm growth rates tend to be more highly correlated the closer two firms are to each other in a customer–supplier network (i.e., the smaller is the shortest path length for the

This paper is an extended version of Mizuno et al. (2014).

T. Mizuno (✉)

National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan
e-mail: mizuno@nii.ac.jp

Department of Informatics, The Graduate University for Advanced Studies, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

PRESTO, Japan Science and Technology Agency, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

The Canon Institute for Global Studies, 5-1 Marunouchi 1-chome, Chiyoda-ku, Tokyo, Japan

W. Souma

College of Science and Technology, Nihon University, 7-24-1 Narashinodai, Funabashi, Chiba, Japan
e-mail: soumawataru@nihon-u.ac.jp

T. Watanabe

Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
e-mail: watanabe@e.u-tokyo.ac.jp

The Canon Institute for Global Studies, 5-1 Marunouchi 1-chome, Chiyoda-ku, Tokyo, Japan

© Springer Japan 2015

T. Watanabe et al. (eds.), *The Economics of Interfirm Networks*, Advances in Japanese Business and Economics 4, DOI 10.1007/978-4-431-55390-8_2

two firms). This suggests that a non-negligible portion of firm growth fluctuations stem from the propagation of microeconomic shocks—shocks that affect a specific firm—through the customer–supplier chains.

Keywords Buyer-supplier network · Aggregate volatility · Input–output analysis · Power-law distribution · PageRank

2.1 Introduction

Firms in a modern economy tend to be closely interconnected, particularly those in the manufacturing sector. Firms typically rely on the delivery of materials or intermediate products from their suppliers to produce their own products that in turn are delivered to downstream firms. Two recent episodes vividly illustrate how closely firms are interconnected. The first is the earthquake and tsunami that hit the Tohoku region in the northeastern part of Japan on March 11, 2011, resulting in significant human and physical damage to that region. However, the economic damage was not restricted to the geographical region and spread in an unanticipated manner to other parts of Japan through the disruption of supply chains. For example, vehicle production by Japanese automakers that were physically removed from the affected areas was either halted or reduced because of a shortage of auto part supplies from firms located in the affected areas. This shock even spread across borders and led to a substantial decline in North American vehicle production.¹ The second episode is the recent financial turmoil triggered by the subprime mortgage crisis in the United States. This adverse shock originally stemmed from the so-called toxic assets on the balance sheets of U.S. financial institutions and led to the failure of these institutions. The shock was transmitted beyond entities that had direct business dealings with the collapsed financial institutions to those that seemed to have no relationship with them, resulting in an upheaval that affected financial institutions around the world.

These two episodes show that both national economies and the global economy are subject to the risk of a chain reaction in product disruptions through customer–supplier linkages. Such risk is particularly high when the linkage structure in the economy is dominated by a few hub firms whose products are supplied to many other firms as input. Importantly, supply-chain disruptions are more serious when there are no close substitutes to the hub firms, at least in the short run. Motivated at least partly by these episodes, some recent studies in macroeconomics have sought to develop theoretical production chain models that extend input-output analysis—which dates back to the seminal work by Wassily Leontief published in the 1930s

¹ For example, the U.S. Federal Reserve chairman Ben Bernanke stated in the aftermath of the earthquake: “U.S. economic growth so far this year looks to have been somewhat slower than expected. Aggregate output increased at only 1.8 % at an annual rate in the first quarter, and supply-chain disruptions associated with the earthquake and tsunami in Japan are hampering economic activity this quarter.” (Speech at the International Monetary Conference, Atlanta, Georgia, U.S. on June 7, 2011).

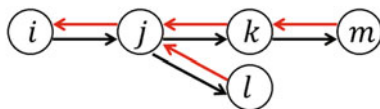


Fig. 2.1 An illustration of customer–supplier network consisting of five firms. The *red arrows* in the figure indicate the flow of money, while the *black arrows* indicate the flow of products each firm produces. For example, firm j purchased something from firm i and sells something to firms k and l

(Leontief 1936)—to identify conditions under which microeconomic shocks (idiosyncratic shocks to individual firms) can propagate to the rest of the economy through production chains, leading to fluctuations in production at the macro level (Acemoglu et al. 2012, 2013b; Carvalho 2010, 2012; Stella 2013; Kelly et al. 2013). Policymakers are also aware of the need to prepare for the propagation of adverse shocks through production chains.²

The present study seeks to provide empirical evidence on the structure and evolution of customer-supplier networks in Japan using a unique dataset that contains information on customer and supplier linkages for over 500,000 incorporated non-financial firms for the 5-year period from 2008–2012. This dataset provides the customer and supplier lists for each firm. We use these lists to produce a customer-supplier network. To illustrate this, Fig. 2.1 shows a simple example consisting of five firms. The red arrows in the figure indicate the flow of money, while the black arrows indicate the flow of products that each firm produces. Firm j purchases something from firm i (its supplier) and sells something to firms k and l (its customers). Moreover, firm k , which purchases something from firm j , sells to firm m . Note that firm k is both a customer (it buys from firm j) and a supplier (it sells to firm m).

The rest of the paper is organized as follows. Section 2.2 compares two important customer–supplier network models—Leontief and PageRank—showing that both models are equivalent under some assumptions. Section 2.3 provides a detailed description of the dataset used, while Section 2.4 examines the basic structure of customer–supplier networks, focusing particularly on how closely firms are interconnected. Section 2.5 investigates how customer–supplier networks evolve over time. Section 2.6 empirically evaluates the extent that firm sales and growth are affected by the propagation of idiosyncratic shocks through production chains. Finally, Section 2.7 concludes the paper.

² The study of networks as phenomena that deserve analysis goes back to the small-world network model by Watts (Watts and Strogatz 1998) and has gained popularity in a variety of scientific disciplines including statistical physics, computer science, biology, and sociology. The methodology developed in those disciplines has been introduced into economics only relatively recently (Jackson 2010; Goyal 2012). However, it has produced important contributions on bank-firm relationships (Souma 2003), cross shareholdings (Garlaschelli et al. 2005), supply chains (Atalay et al. 2011; Saito et al. 2007; Ohnishi et al. 2010; Takayasu et al. 2008; Fujiwara and Aoyama 2010; Watanabe et al. 2012), systemic risks in financial markets (Battiston et al. 2007; Acemoglu et al. 2013a), and international trade (Garlaschelli and Loffredo 2005; Garlaschelli and Loffredo 2004; Di Giovanni and Levchenko 2010).

2.2 Equivalence of Leontief and PageRank Models

The initial concept of interfirm networks originated in the 1930s (Leontief 1936), although the nodes used in the Leontief analysis at that time are economic sectors (i.e., industries) rather than individual firms. Let us denote the number of firms in an economy by N and the sales vector associated with those firms by x that is a column vector with the sales of firm i on the i th row. We normalize x so that $x' \mathbf{1} = 1$ holds. The input-output structure of the economy is represented by an $N \times N$ matrix A with a_{ij} as an (i, j) element. The element a_{ij} denotes the share of input j (i.e., the commodity produced by firm j) in the total intermediate input use of firm i . Market clearing conditions are given by

$$x = (1 - \alpha)A'x + f \quad (2.1)$$

where f is a column vector representing the final demand to individual firms and α is the share of value added to gross sales. The first and second terms on the right-hand side represent the intermediate and final demands. Two new assumptions are now introduced. The first assumption relates to the final demand vector that is given by

$$f = \frac{\alpha}{N} \mathbf{1} \quad (2.2)$$

That is, the final demand is equal across the firms. Second, we assume that the a_{ij} is equal to the reciprocal of the total number of suppliers to firm i if $a_{ij} > 0$ and zero otherwise. This means that the supply links to firm i are of the same thickness. A new input-output matrix defined in this way is denoted by \tilde{A} . Given these assumptions, Eq. (2.1) changes to

$$\tilde{x} = (1 - \alpha)\tilde{A}'\tilde{x} + \frac{\alpha}{N} \mathbf{1} \quad (2.3)$$

Notably, the column vector \tilde{x} in Eq. (2.3) is a PageRank vector (Brin and Page 1998). PageRank is an algorithm used by Google Search to rank websites in their search engine results. Equation (2.3) shows that the input-output model invented by Leontief in the 1930s is closely connected to the basic idea of PageRank.

Based on Eq. (2.3), Acemoglu et al. (2012) investigates how an economy's value added, the log of which is denoted by y , is affected by idiosyncratic shocks to individual firms in the economy. Denoting an idiosyncratic shock to firm i by v_i (which is assumed to be i.i.d. with mean zero and variance σ^2) and the corresponding column vector by v , we have $y = \tilde{x}'v$. If the distribution of PageRank across firms (i.e., \tilde{x}_i for $i = 1, \dots, N$) follows a uniform distribution (i.e., $\tilde{x}_i = 1/N$), then we have $\sqrt{\text{var}(y)} = \sigma/\sqrt{N}$, implying that the standard deviation of y converges to zero as $N \rightarrow \infty$ at the rate \sqrt{N} . Generally, the central limit theorem guarantees that the standard deviation of y decays at the rate \sqrt{N} if the PageRank distribution is sufficiently close to a uniform distribution. This implies that idiosyncratic shocks to individual firms would not translate into aggregate shocks because idiosyncratic shocks quickly cancel each other out as the number of firms increases (Dupor 1999).

However, this does not hold if PageRank is substantially unequal across firms (Acemoglu et al. 2012). Specifically, if the PageRank distribution has a power-law tail (i.e., $\Pr(x_i > x) \propto x^{-\zeta}$, where ζ is a power-law exponent with ζ between 1 and 2), then we have $\sqrt{\text{var}(y)} = \sigma/N^{1-1/\zeta}$. This means that the standard deviation of y decays at a rate slower than \sqrt{N} , implying that idiosyncratic shocks do not cancel each other out as quickly as implied by the central limit theorem. Hence idiosyncratic shocks to firms with very large PageRank may have a substantial impact on y . Typically, firms with large PageRank are hub firms that have a large number of trade partners. Idiosyncratic shocks to those hub firms spread to other firms through customer–supplier linkages, leading to a cascade phenomenon.

Notably, this argument is based on the two assumptions regarding final demand and the input-output matrix that may not actually hold in the data. However, one can make a similar argument by replacing $y = \tilde{x}'v$ with $y = x'v$. Gabaix (2011) shows that $\text{var}(y)$ converges to zero as $N \rightarrow \infty$ at the rate \sqrt{N} if the x_i 's are uniformly distributed (or close to it). Conversely, $\text{var}(y)$ decays at a slower rate if the distribution of x_i is heavy tailed, implying again that idiosyncratic shocks to individual firms would translate into macro shocks. This is referred to by Gabaix (2011) as granular hypothesis. It differs importantly from the cascade hypothesis proposed by Acemoglu et al. (2012) in that firms with large x_i may not necessarily be highly connected. For example, the large x_i of those firms may come from final demand rather than intermediate demand.

These two hypotheses have different implications on how policy makers should act to mitigate fluctuations in y . The granular hypothesis considers that fluctuations in y come from firms with large sales, so that it is important to mitigate idiosyncratic shocks to those firms. The “too big to fail” principle, often discussed in the context of preventing the failures of large financial institutions, is an example of such an action. The cascade hypothesis, however, implies that how closely a firm is connected to other firms through its customer–supplier linkages is vital rather than the firm size. This corresponds to the idea of “too interconnected to fail” that has been discussed by Markose et al. (2012) among others in the context of the recent financial crisis.

The cascade hypothesis has two testable implications. The first implication is that the number of trade links is highly unequal across firms. Particularly, the number of customer links for a firm, which is closely related to its PageRank, must be highly unequal and its distribution must have a heavy upper tail. Second, the cascade hypothesis implies that firm growth rates should be more highly correlated the closer two firms are to each other in a customer–supplier network. We test both implications using a dataset that contains information on customer and supplier linkages for over 500,000 incorporated non-financial firms.

2.3 Data

The dataset we use is jointly compiled by Teikoku Databank, Ltd. (TDB), one of the largest business database companies in Japan, and the HIT-TDB project of Hitotsubashi University. The dataset mainly provides information related to corporate bankruptcies and credit ratings and covers about 1.3 million incorporated

non-financial firms. Since the number of corporations in Japan in 2006 (as reported in the 2006 Establishment and Enterprise Census) was 1493 million, our dataset covers about 90 % of all incorporated firms in Japan. TDB collects various kinds of information from these firms, including annual or more frequent financial statement data.

Two types of information on customer–supplier relationships are recorded in this dataset. First, the dataset contains information on the number of three types of relationships a firm has with other firms, namely relationships with customers (i.e., firms to which a firm sells its products), suppliers (i.e., firms from which a firm purchases raw materials and intermediate products), and owners (i.e., firms by which a firm is owned). Since in this paper we focus on customer–supplier relationships, we mainly use information on customer and supplier linkages. We denote the total number of firm i 's customer links by N_i^C and the total number of supplier links by N_i^S . Second, the dataset lists the firms with which a firm has links (i.e., customers or suppliers to the firm) with their identification codes. However, the list is not exhaustive and its length cannot exceed 60 firms. This means that for smaller firms with fewer than 60 partners all of their partners are listed, but for large firms with more partners only the 60 most important ones are listed. In all cases, transaction partners are listed in descending order of importance based on the transaction volume.

Table 2.1 presents descriptive statistics on customer and supplier linkages. All statistics in the table are calculated using the total number of linkages, that is, N_i^C and N_i^S . Note that the table provides linkage information for five different years (i.e., 2008, 2009, 2010, 2011, and 2012), allowing us to investigate not only the structure of customer–supplier networks at a particular point in time but also their evolution. The sample mean for the number of customer links per firm is about 340 each year, and the median for the number of customer links per firm is 50, which is about one seventh of the mean, implying that the customer link distribution is not symmetric, but is substantially skewed to the right. In fact, the maximum number of customer links in 2012 was 95,512, which is far greater than the mean or the median, given that the standard deviation is only 2053. Turning to the number of supplier links, the sample mean is about 60 each year, which is much smaller than the number of customer links. A typical firm has six times as many customer links as supplier links. The median number of supplier links per firm is 20, implying again that the distribution for the number of supplier links is not symmetric but is skewed to the right. The maximum number of supplier links per firm is also much greater than the mean or the median.

To investigate the structure of customer–supplier networks and their evolution over time, we use the list of firms linked to a firm with their identification codes. As mentioned, the list is not exhaustive, so that, as far as large firms are concerned, links with less important partners are not recorded. The number of customers and suppliers in the list is 6.7 and 6.4 for a typical firm, which is much smaller than the means of the *total* number of customer and supplier links presented in Table 2.1. We augment the customer/supplier lists as follows. We first identify firm A as a supplier of firm B using the *customer* list of firm A, thereby producing an augmented supplier list of firm B. We add up the number of customer links originally shown in the customer list of a

Table 2.1 Number of customer and supplier links per firm

Customer links	2008	2009	2010	2011	2012
Number of firms	160,508	155,806	144,006	142,931	145,317
Number of links per firm					
Mean	339	343	341	340	339
Median	50	50	50	50	50
Std. Dev.	2107	2090	2015	2022	2053
Max.	90,200	90,504	90,000	90,000	95,512
Min.	0	0	0	0	0
Supplier links	2008	2009	2010	2011	2012
Number of firms	215,562	208,459	192,111	189,493	193,045
Number of links per firm					
Mean	56	58	61	62	61
Median	20	20	20	20	20
Std. Dev.	281	314	368	332	351
Max.	52,100	55,100	70,000	70,000	70,000
Min.	0	0	0	0	0

firm and the number of customer links identified in this way, and denote the sum by \tilde{N}_i^C . Similarly, we use the supplier lists of firms to produce augmented customer lists and define \tilde{N}_i^S . This kind of “reverse lookup” method has been applied to different datasets in previous studies on interfirm relationships, including Saito et al. (2007), Fujiwara and Aoyama (2010), Takayasu et al. (2008). Comparing N_i^C and \tilde{N}_i^C , we observe a relationship of the following form:

$$\langle \tilde{N}^C \mid N^C = n \rangle \propto n^{0.83} \quad \text{for} \quad 20 \leq n \leq 10000. \quad (2.4)$$

where $\langle \tilde{N}^C \mid N^C = n \rangle$ represents the mean of \tilde{N}_i^C across i given that the total (true) number of customer links, N^C , for those firms is equal to n . Interestingly, the power exponent of n is smaller than unity, implying that for firms with a large number of customers the augmented list still does not capture the true number of customers. The example of a firm leasing vending machines to other firms explains why. This firm has a very large number of customer firms, but because vending machines are not regarded as a key input to production by most customer firms, they do not include the leasing firm in their list of suppliers. In this case, \tilde{N}^C for the leasing firm is much smaller than N^C .

Turning to supplier lists, we have

$$\langle \tilde{N}^S \mid N^S = n \rangle \propto n^{1.19} \quad \text{for} \quad 10 \leq n \leq 1000. \quad (2.5)$$

indicating that the exponent of n is now greater than unity, which means that \tilde{N}^S more than doubles when N^S doubles, and in this sense \tilde{N}^S overestimates N^S . A likely

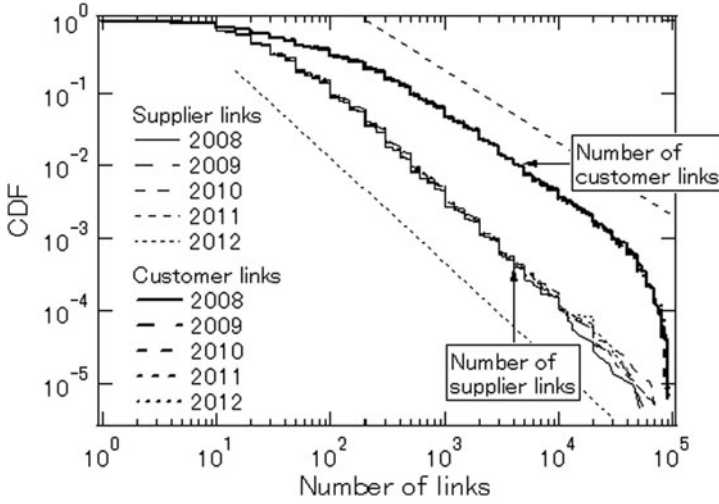


Fig. 2.2 Cumulative distributions of customer and supplier links in 2008–2012. The horizontal axis represents the *total* number of links, i.e., N^C and N^S , while the vertical axis represents the corresponding cumulative densities. The dotted straight lines are reference lines with a slope of -1 and -1.5 respectively. The number of firms used in this calculation is shown in Table 2.1

reason is that small suppliers to a prestigious firm with a large number of suppliers will include the prestigious firm in their customer list reported to TDB, since the prestigious firm is regarded as a key constituent of their customer base. However, this effect will be weak or absent if a customer firm is not that prestigious, which makes the exponent of n in Eq. (2.5) greater than unity.

2.4 The Structure of Customer–Supplier Networks

2.4.1 Unequal Links Across Firms

The number of links is unequal across firms with regard to both customer and supplier linkages, as we saw in Table 2.1. One may wonder how unequal it is across firms and whether the degree of inequality differs between customer and supplier linkages. To address these questions, we show in Fig. 2.2 the cumulative distribution functions (CDFs) of links across firms. The horizontal axis represents the number of links, while the vertical axis shows the corresponding cumulative densities. The horizontal and vertical axes are both in logarithm. For example, the number on the vertical axis corresponding to 10^2 on the horizontal axis is about 10^{-1} for supplier linkages, indicating that firms with more than 10^2 supplier links account for one tenth of all firms. The figure shows the CDFs for the customer and supplier linkages for each of our five observation years (2008, 2009, 2010, 2011, and 2012).

Given that the mean for the logarithm of the number of customer links is 1.72 and the corresponding standard deviation is 0.783, a number like 5000 links deviates from the mean by more than 2.52σ , and a number like 50,000 links deviates by more than 3.80σ . If the number of customer links is lognormally distributed, the cumulative probabilities corresponding to 5000 and 50,000 links are 0.0058 and 0.000072, which is much lower than the probabilities that we actually observe, indicating that the number of customer links has a heavier upper tail than a lognormal distribution.

The CDFs of customer links show a linear relationship between the log of the number of links and the log of the corresponding cumulative probability for the number of links within the range of 80–50,000. The slope is around -1 and is not significantly different from this value in each of the 5 years, that is,

$$P_{>}(N^C) \propto \frac{1}{N^C} \quad \text{for} \quad 80 \leq N^C \leq 50000 \quad (2.6)$$

where $P_{>}(N^C)$ represents the probability that the number of customer links exceeds a certain value. Equation (2.6) shows that N^C follows a power-law distribution and, more importantly, that its exponent is very close to unity. Power-law distributions with exponent 1 are found in various economic phenomena, including the distribution of city sizes, asset price changes, and firm sizes, a phenomenon referred to as Zipf's law. Most importantly, as shown by previous studies (Axtell 2001), firm sales follows Zipf's law, suggesting that the sales of a firm are related to the number of customers the firm has. We will come back to this issue in Section 2.5.

Turning to the number of supplier links, we again find a linear relationship between the log of the number of supplier links and the log of the corresponding cumulative density, indicating that the number of supplier links also follows a power-law distribution. However, the slope of the linear relationship is much larger than that in the case of customer links, implying that the tail part of the supply link distribution is less fat than that of the customer link distribution. The slope associated with supplier linkages is about -1.5 , so that the CDFs for the number of supplier links can be characterized by

$$P_{>}(N^S) \propto \left(\frac{1}{N^S}\right)^{1.5} \quad \text{for} \quad N^S \geq 30. \quad (2.7)$$

Since the power-law exponent in this case exceeds unity, Zipf's law does not hold. Note that the power-law exponent ζ is related to the Gini coefficient, G , in the form $G = 1/(2\zeta - 1)$. Therefore, the fact that the power-law exponent is larger for supplier linkages than for customer linkages implies that the Gini coefficient is smaller for supplier linkages and that, therefore, the number of supplier links across firms is less unequal than the number of customer links.

What explains this result? As emphasized in the recent literature on customer search models (Luttmer 2006; Gourio and Rudanko 2011), firms spend substantial resources on marketing to acquire as many customers as possible in order to increase their sales and profits. Such competition among firms produces winners with a large number of customers as well as losers with a small number of customers, resulting

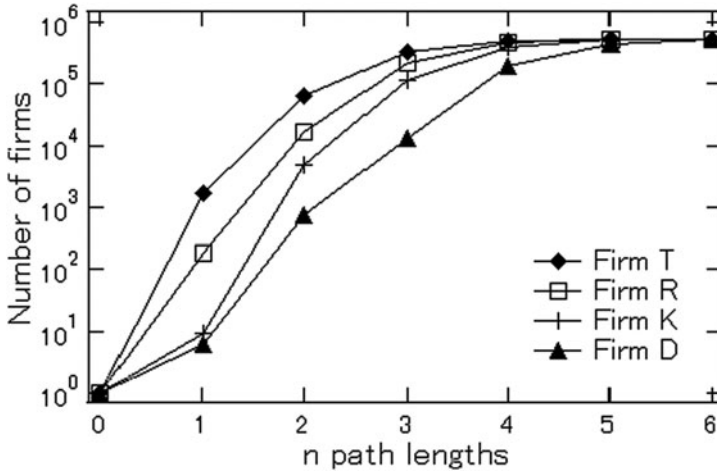


Fig. 2.3 Number of firms connected to a particular firm by n path lengths. Firms T , R , K , and D are randomly picked from the sample, which consists of all the firms on the augmented customer/supplier lists

in huge inequality in the number of customers. In contrast, with regard to supplier linkages, firms have little incentive to increase their number of suppliers because it is not necessarily profitable to buy materials and intermediate products from more suppliers. It may even be the case that purchases from more suppliers increase the associated costs (e.g., shipping costs) and therefore reduce profits. Therefore, because firms do not compete to have as many suppliers as possible, the extent of inequality is not as high as that with regard to the number of customers.

2.4.2 How Closely are Firms Interconnected?

To investigate how closely firms are interconnected, we use the augmented customer/supplier lists of partners mentioned in Section 2.2 for the set of firms whose identification codes are listed in the customer and/or supplier lists of the other firms. The number of firms that appear in the augmented lists is about 500,000.³ Specifically, we randomly pick four firms (Firms T , R , K , and D) to examine the number of firms connected to a particular firm by one, two, three, or more path lengths. The result is shown in Fig. 2.3. Firm T is connected to about 1700 firms by one path length, but it is connected to more than 60,000 firms by two path lengths. The corresponding number for four path lengths increases to 503,796, which is only slightly

³ The number of firms in the augmented lists is 552,145 for 2008, 541,816 for 2009, 518,565 for 2010, 520,087 for 2011, and 525,836 for 2012.

The Economics of Interfirm Networks

Watanabe, T.; Uesugi, I.; Ono, A. (Eds.)

2015, X, 263 p. 45 illus., 14 illus. in color., Hardcover

ISBN: 978-4-431-55389-2