Konno, Tomohiko

**Article**

# Network structure of Japanese firms. Scale-free, hierarchy, and degree correlation: analysis from 800,000 firms

Economics: The Open-Access, Open-Assessment E-Journal

**Provided in Cooperation with:**
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

*Suggested Citation:* Konno, Tomohiko (2009) : Network structure of Japanese firms. Scale-free, hierarchy, and degree correlation: analysis from 800,000 firms, Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 3, Iss. 2009-31, pp. 1-13,
https://doi.org/10.5018/economics-ejournal.ja.2009-31

This Version is available at:
https://hdl.handle.net/10419/27725

Economics
The Open-Access, Open-Assessment E-Journal

Vol. 3, 2009-31 | July 1, 2009 | http://www.economics-ejournal.org/economics/journalarticles/2009-31

# Network Structure of Japanese Firms. Scale-Free, Hierarchy, and Degree Correlation: Analysis from 800,000 Firms

*Tomohiko Konno*

*Graduate School of Economics, University of Tokyo*
*Research Fellow of the Japan Society for the Promotion of Science*

**Abstract**

We analyze fundamental characteristics of the inter-firm transaction network through the data of 800,000 Japanese firms. We find that there exists a hierarchical structure and a negative degree correlation in this transaction network. We also find that this undirected network is a scale-free network. We bring to light these characteristics of the network and discuss why there is an important need to conduct research work on the actual network structure.

Special issue Reconstructing Macroeconomics

**JEL:** L16

**Keywords:** Network of firms; scale-free network; complex network; hierarchy; degree correlation

**Correspondence**

Tomohiko Konno, Bunkyo-ku, Hongo 7-3-1 Tokyo, Japan,
E-mail: tomo.konno@gmail.com

*The author acknowledges Hiroshi Yoshikawa and The Research Institute of Economy, Trade and Industry (RIETI) for their help, especially for valuable data set. This research was, in part, financially supported by JSPS.*

# 1 Introduction

In this paper, we study the inter-firm transaction network that exists in Japanese firms. We find that in this network, there is a hierarchy and a degree correlation. We discover the hierarchy by analyzing the clustering coefficient. We also discover that an undirected network has scale-free degree distribution

Garlaschelli et al. (2005), Souma et al. (2006) studied the existing shareholders' network in Japan, Italy, and the US and found that this network contained scale-free network structures. However, their data comprised only a small number of firms. In contrast, our data contain 800,000 firms. Saito et al. (2007) used the same data as ours and their study revealed that the directed network had a scale-free distribution. We believe that some important network characteristics remain undiscovered and feel that we will be able to discover some significant ones. Our motivation originates from the development of the complex network theory and the belief that the network is as significant in the field of economics as it is in the fields of other sciences. In fact, we believe that a complex network view can benefit Economics greatly. Some reviews of complex networks are detailed below Dorogovtsev and Mendes (2003), Vega-Redondo (2007), Barabási and Albert (2002), Newman (2003), Boccaletti et al. (2006).

The purpose of the present paper is to mainly provide information on the inter-firm transaction network in order to facilitate further studies on models on networks in the future; for example, competing firms on network and so on. Unless we have information about real network structures, we will have no idea about the kind of network that we can assume. As has been already been discovered in connection with models on networks, the network structure in evolutionary games affects the outcome of the model. For example, Abramson and Kuperman (2001) contend that the Prisoner's Dilemma evolutionary games, referred to as PD games hereafter, differ across network structures that range from regular lattices to random networks. Santos and Pacheco (2005) contend that the cooperative behavior in a PD game is enhanced by a scale-free network. In his paper,Ohtsuki et al. (2006) discusses the rule for enhancing cooperation in evolutionary games on networks. Hence, it is likely that in Economics, the outcomes of the models on network are dependent on the underlying network structures. This paper will facilitate further such research in the near future.

## 1.1 Models on Network

In the complex network theory, as is mentioned above, it has been discovered that there is a significant relationship between the behavior of agents and their underlying network structures, for example in evolutionary games Ohtsuki et al. (2006), Santos and Pacheco (2005), Abramson and Kuperman (2001). Such a relationship between the games and networks is one of the motivations that stimulate us to study real network structures. With regard to the hierarchical structure, Vukov and Szabó (2005) contend that in PD games on hierarchical regular lattices, if there are a sufficient number of layers, the highest frequency of cooperation occurs in the middle layers. In their study, they discuss the optimum number of layers that the commu-

nity should have. They also contend that there is diminished cooperation among the PD games on the hierarchical scale-free networks. Aoki and Yoshikawa (2006) discuss the importance of a hierarchical structure in Finance and Economics.

Moving on to the topic of degree correlation, the co-star and co-author relationship has a positive degree correlation. On the other hand, the concepts of gene network, protein network, nerve circuit, and food chain appearing in Biology have a negative correlation. Artificial networks like power grids and the Internet have a negative but weak correlation. Degree correlation differs across different networks. Positive correlation tends to lead to a lower percolation transition point Newman (2002), Callaway et al. (2001). In particular, it is well known that degree correlation affects the synchronization of oscillators on networks Motter et al. (2005), Di Bernardo et al. (2007), Sorrentino et al. (2006), Di Bernardo et al. (2005). In this sense, degree correlation is also an important network characteristic.

## 1.2 Outline of this Paper

The outline of this paper is as follows. Section 2 consists of three subsections. The first subsection shows the approach to identifying hierarchical structures and degree correlation. It comprises four blocks. We begin by explaining the figure that we will be using many times thereafter. Then, we briefly touch upon the random and scale-free networks. Next, we describe the hierarchical structure of a network by comparing it with other types of networks that are devoid of hierarchy and also describe the approach to identifying a hierarchical structure. Finally, we explain what is meant by degree correlation. The second subsection has three blocks. Here, we first describe the data. Thereafter, we explain the method through which degree distribution is calculated. Finally, we introduce clustering coefficients. In the last subsection, we show the results of the study and discuss them in three blocks. First, we review the degree distribution of the Japanese inter-firm undirected network. Next, we show that the network has a hierarchical structure by analyzing the clustering coefficient. Finally, we show that the network has a negative degree correlation. Section 3 contains the conclusion.

# 2 Analysis

## 2.1 Identifying Hierarchical Structure and Degree Correlation

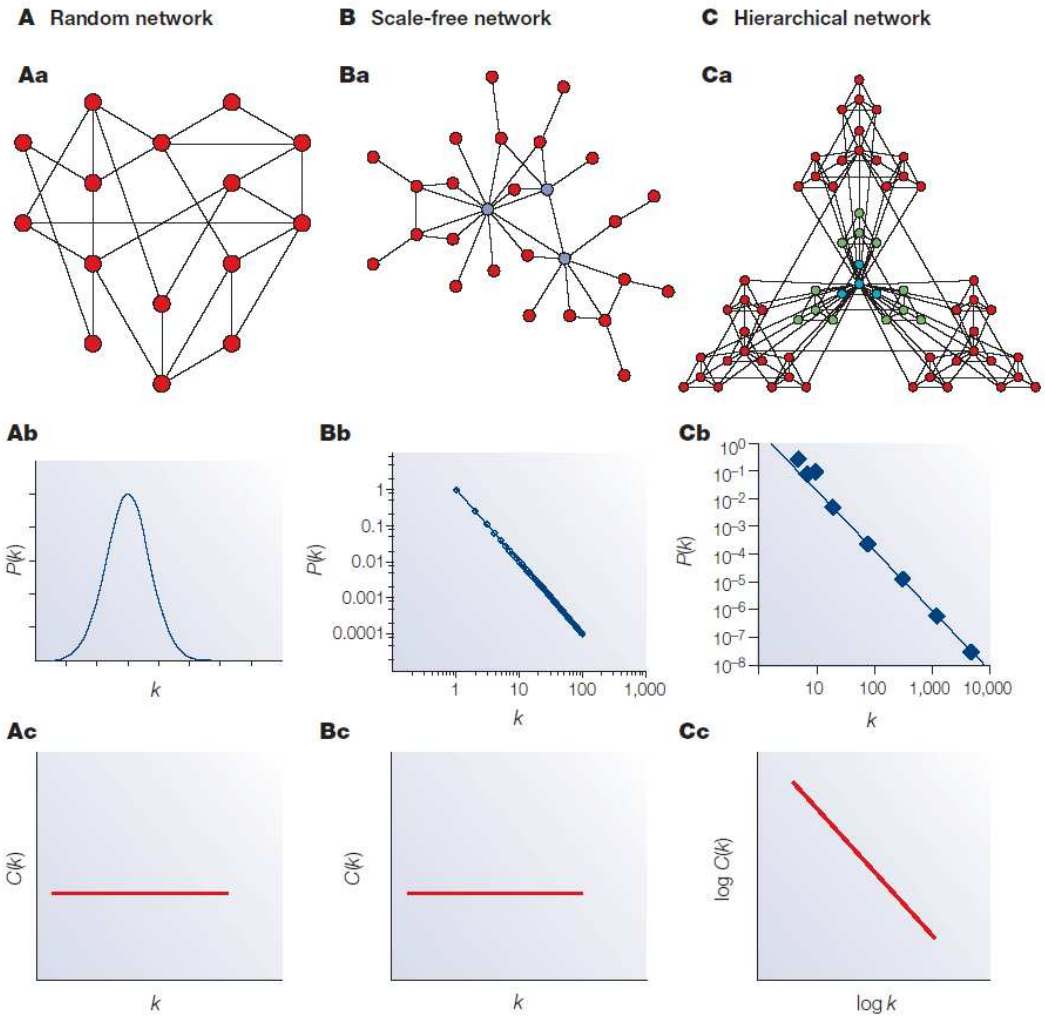


Figure 1: Network Structures

First , we explain Fig.1[1] briefly. In this figure, there are three types of networks: random networks, scale-free networks, and hierarchical networks. We will explain the networks in this given order. In the figure, $k$ stands for the degree, which is defined by the number of links that the vertex has. The first row (a) illustrates the examples of each network, the second row (b) exhibits degree distribution $P(k)$, and the third row (c) shows the clustering coefficient $C(k)$. The latter two terms will be explained later.

[1]Reprinted by permission from Macmillan Publishers Ltd: [Nature Reviews Genetics] Barabási and Oltvai (2004), copyright(2004)

### 2.1.1 Scale-Free Network and Random Network

Firstly, let us explain scale-free networks. A scale-free network is a network whose degree distribution is as follows.

$$P(k) \sim k^{-\gamma} \tag{1}$$

The degree is the number of links that a vertex has and is denoted by $k$. A scale-free network is quite different from the "Random Graph" that was defined by Erdos and Renyi (1959). A random graph is constructed in the following way. Choose two vertices and either link them with uniform probability $p$ or do not link them with probability $1-p$. Complete this procedure for all pairs of vertices. If the whole network has $n$ vertices, the degree distribution is binomial.

$$\begin{aligned} p(k) &= {}_{n-1}C_k \, p^k (1-p)^{n-1-k} \\ &\sim {}_nC_k \, p^k (1-p)^{n-k} \quad (\because n-1 \sim n). \end{aligned} \tag{2}$$

In the limit $n \to \infty, p \to 0$, with keeping $np = \lambda$, the Eq.(2) becomes Poisson distribution

$$P(k) = \frac{e^{-\lambda}\lambda^k}{k!}. \tag{3}$$

Here, $\lambda$ is the mean degree of the network. In Fig.1, Aa and Ab illustrate a random graph and its degree distribution.

On the other hand, Ba and Bb in Fig.1 show a scale-free network and its degree distribution in a log-log plot. In a random graph, there are no vertices with very large degrees. In a scale-free network, in contrast, there are a small number of vertices with very large degrees that are called "Hubs." Roughly speaking, the two networks differ from each other in that hubs are present in one while being absent in the other.

### 2.1.2 Identifying Hierarchical Structure

The third network (C) is a hierarchical network and has a scale-free degree distribution $P(k) \sim k^{-\gamma}$. The difference between a scale-free network (B) and a hierarchical network is depicted in the third row (Bc, Cc), which illustrates the clustering coefficients of these networks. The clustering coefficient, which will be explained later, of the scale-free network in the figure is constant $C(k) = \text{Const}$. On the other hand, the clustering coefficient of a hierarchical network is dependent on $k$ as $C(k) \sim k^{-1}$. When we compare this hierarchical network with other two types of networks,the structure of the hierarchical network becomes clear. In many real networks, the clustering coefficient and the degree have the special relation of $C(k) \sim k^{-1}$. Examples of this relation are the Co-Actor network, the Language network, the World Wide Web, and the Internet at the autonomous system level. Ravasz and Barabási (2003) show that this relation, $C(k) \sim k^{-1}$, implies a hierarchical structure. The authors term networks like (Cc) in Fig.1 "hierarchical structure" networks. Barabási and Oltvai (2004) have also explained this relation and hierarchical structure.

As is explained in graph (Cc), $\log c(k)$ and $\log k$ are linearly proportional and the proportionality coefficient is $-1$, while the clustering coefficient $C(k)$ is constant in the other two networks. Therefore, in order to detect the relation $C(k) \sim k^{-1}$ that implies a hierarchical structure as in (Ca)[2] in Fig.1, we study the clustering coefficient.

### 2.1.3 Degree Correlation

We study the degree correlation of the network. Degree correlation is defined by the following equation:

$$k_{nn}(k) \equiv \sum_{k'} k' \Pr(k' \mid k). \tag{4}$$

$\Pr(k' \mid k)$ is the conditional probability that the vertex with degree $k$ is adjacent to the vertex with degree $k'$. In a nutshell, $k_{nn}(k)$ implies that the vertex with degree $k$ is adjacent to the vertex with degree $k_{nn}(k)$. This is also an important network characteristic.

## 2.2 Data, Degree Distribution, and Clustering Coefficient

### 2.2.1 The Data

The data for this study have been supplied by Tokyo Shoko Research, Ltd. (TSR) via RIETI and consist of the financial data and network relationships of 800,000 firms: in other words, the data comprise their buying, selling, and shareholder relationships. The data contain reports on 4,000,000 of the firms' relations and include information on gross sales, region, year of establishment, number of employees, number of offices, number of factories, industrial classification, and so on. We did not use the data of any firm that does not report its gross sales; thus, the number of firms we used reduced to 800,000. We used the relationship between the firms' buying and selling. This data set was created by asking the firms to report their business partners. Thus, the data set suffers from one limitation: it does not include all the transactional relationships. In this paper, we construct an undirected network in which we do not discriminate the fact that firm A sells to firm B from the fact that firm B sells to firm A. This is because we believe that when analyzing the clustering coefficient, it is appropriate to study an undirected network. Furthermore, we believe that studying the clustering coefficient of an undirected network is more essential than studying that of a directed one. Thus, this paper only studies the relationship in which there exists a transaction between two firms. We built an adjacency matrix, which is the common method through which networks are analyzed. In the adjacency matrix, we set element ij as 1 if there is any transaction, regardless of whether it involves buying or selling . We set element ij as 0 if no transaction takes place between firm i and firm j. Hence, the number of transactions between firms is not considered either; our only concern is whether any transaction has taken place. Actually, we

[2]Similar figure appears in Ravasz and Barabási (2003)

are unable to discern the number of transactions between specific firms from the data. Thus, the new data that include the number of transactions will be able to reveal new information.

### 2.2.2 Degree Distribution

Now, we present the method of calculating degree distribution in detail. First, we count all the links from the vertex with degree $k$. Let this be denoted by "All degree(k)." Remember that the degree indicates the number of links that the vertex has. Second, we calculate $P(k)$ as follows:

$$P(k) = \frac{\text{All degree}(k)}{\sum_{k'} \text{All degree}(k')} \tag{5}$$

In order to detect scale-free distribution, it is better to draw the CDF rather than the PDF, the reason for which is given in Dorogovtsev and Mendes (2003). Hence, we illustrate the CDF.

### 2.2.3 Clustering Coefficient

Now, we provide the definition of the clustering coefficient. The clustering coefficient can be defined for every vertex whose degree is larger than 1. For example, the clustering coefficient of vertex j is defined as follows:

$$C_j = \frac{\text{The number of triangles around vertex-j}}{\text{The maximum number of possible triangles around vertex-j}} \tag{6}$$

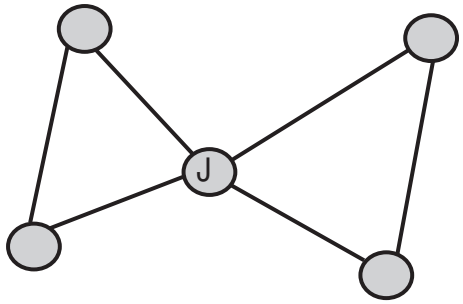


Figure 2: explanation for clustering coefficient

Fig.2 explains the clustering coefficient. The number of triangles around vertex j is 2. The maximum number of triangles we could make around vertex j is ${}_4C_2 = 6$ because there are 4 vertices around vertex j. Thus, the clustering coefficient of vertex j is $\frac{2}{6} = \frac{1}{3}$. If the degree is 1, we cannot define the clustering coefficient because we cannot make any triangle around the vertex. Recall that the degree is defined as the number of edges that the vertex has. For example, in Fig.2, the degree of vertex j is 4. In a nutshell, the clustering coefficient helps us measure how densely the vertices are connected locally among their neighbors. We use a friends' network as an example. If the clustering coefficient is large, it is likely that a friend of a

friend is also your friend. The same in unlikely if the clustering coefficient is small. It is worth noting that many real networks have a larger clustering coefficient and smaller mean path length as compared to random networks. The mean path length is defined as the average of the path lengths over all the pairs of vertices.

## 2.3 Results and Discussions

### 2.3.1 Degree Distribution of the Network

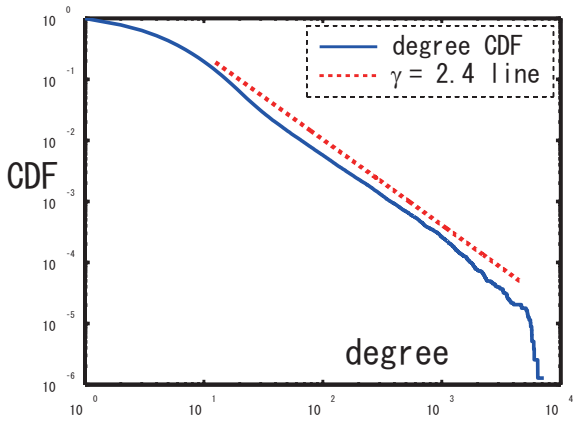


Figure 3: Degree distribution of Japanese firms' transaction network

We study the transaction network of Japanese firms.
The solid line in Fig.3 illustrates the $1-\text{CDF}$ of degree distribution in log-log plots. Saito et al. (2007) shows that the directed network of the transactions of Japanese firms in which the buying and selling transactions are distinguished from each other has a scale-free distribution. On the other hand, we show that the degree distribution of an undirected network,[3] in which we consider that the firms are linked if there exists either transaction, follows a scale-free distribution of the form $P(k) \sim k^{-2.4}$. $1-\text{CDF}$ of $P(k) \sim k^{-2.4}$ is illustrated through a dashed line in in the same figure. The drop of the actual data line in the right of the figure stems from the fact that there are only finite number of vertices in the network. If we want to see a perfect scale-free distribution, we will need a network with infinite number of vertices.

### 2.3.2 Hierarchical Structure of Japanese Firms

In this section, we discuss the hierarchical structure of the transaction network of Japanese firms that is implied by a clustering coefficient.

We draw the clustering coefficient of the transaction network of 800,000 Japanese firms. Fig.4 depicts the scatter plots and estimated line. The x-axis is $\log(k)$ and y-axis is log(clustering coefficient). Table 1 shows the estimation result.

[3]In undirected network adjacency matrix is symmetric, while in directed network adjacency matrix is not generally symmetric.

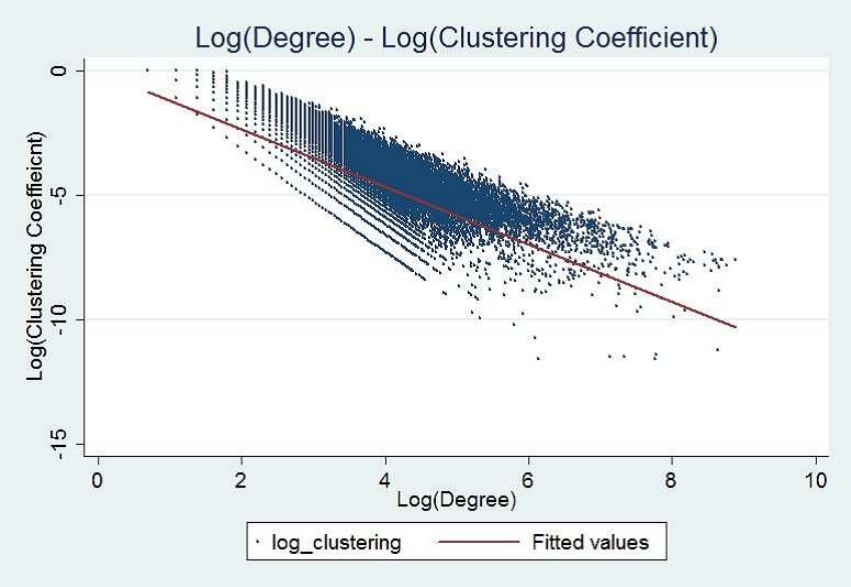


Figure 4: Log(Degree) - Log(Clustering Coefficient)

Table 1: Estimation results: Clustering coefficient

| | **Coefficient** | **(Std. Err.)** |
|---|---|---|
| Log (degree) | -1.159 | (0.002) |
| Intercept | -0.049 | (0.004) |

The estimated relation is

$$\log C(k_j) = -1.159 \log k_j - 0.049 \tag{7}$$

$R^2$ is 0.66. Remember that $C(k)$ stands for the clustering coefficient of the vertex with degree $k$. Eq.(7) strongly demonstrates that the coefficient of $\log(k)$ is very close to $-1$; this relation is equivalent to $C(k) \sim k^{-1}$, which is desired. As we have discussed previously, this relation implies that the transaction network of Japanese firms not only has a scale- free structure, it also possesses a hierarchical structure. This fact is clearly exemplified (Ca) in Fig.1.

Fig.4 seems to comprise less than 4 structures of dots aligning on lines with a negative slope. We need to explain this observation. Remember that the clustering coefficient is defined as

$$\begin{aligned} C(k) &\equiv \frac{\text{The number of triangles}}{k(k-1)/2} \\ &\sim \frac{\text{The number of triangles}}{k^2/2}. \end{aligned} \tag{8}$$

The number of triangles in Eq.(8) is discrete, i.e., 1, 2, 3, and so on. Hence, the bottommost structure consists of points that comprise 1 triangle, and thus, the clustering coefficients are $2 \times 1/k^2$. Similarly, the second structure from the bottom

consists of points that comprise 2 triangles. Subsequently, the clustering coefficients are $2 \times 2/k^2$. The clustering coefficients of the third structure from the bottom are $2 \times 3/k^2$ and so on. Since we take their logarithm, the slope of these structures is $-2$, so that in the region where the number of triangles is small, they appear to be aligned.

We would like to mention the following. Barabási and Albert (1999) introduced the famous scale-free network generating mechanism known as preferential attachment. In a nutshell, the higher the vertex's degrees, the more likely it attracts links from other vertices. However, it is well known that the network generated by this preferential attachment mechanism does not have a hierarchical structure. In this network, the relation $C(k) \sim k^{-1}$ cannot be observed. Since the transaction network has a hierarchical structure, another mechanism must be governing the firms' transaction network.

### 2.3.3 Degree Correlation

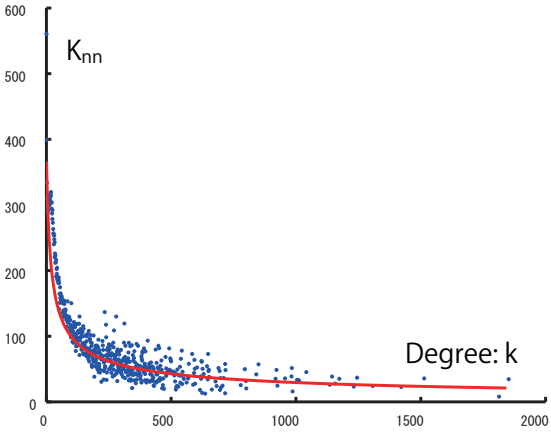


Figure 5: Degree Correlation

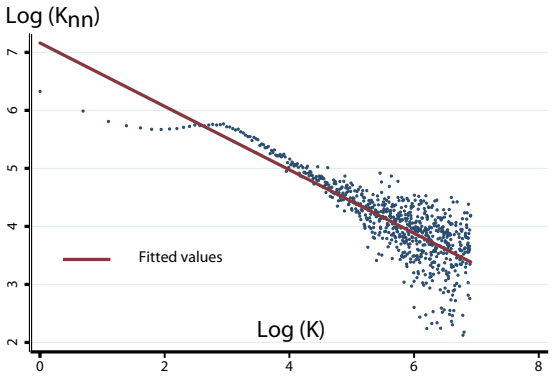


Figure 6: In Log Log Plots

Another important discovery is the existence of a degree-degree correlation. The degree and the mean of the next neighbor degrees have the following relation in the network.

$$k_{nn} \sim k^{-0.5} \tag{9}$$

Here, $k_{nn}$ stands for the mean of the next neighbor degrees. The result is $k_{nn} = 1289\ k^{-0.546}$. Table 2 shows the regression results.
We thus obtain $\log(k_{nn}) = -0.546\ \log(k) + 7.162$, which is almost same as Eq.(9). Further, $R^2$ is 0.681. Fig.5 illustrates the scatter plot and fitted curve while Fig.6 demonstrates the log-log plot and fitted curve.

Table 2: Estimation results: Degree Correlation

| | **Coefficient** | **(Std. Err.)** |
|---|---|---|
| Log degree | -0.546 | (0.014) |
| Intercept | 7.162 | (0.080) |

# 3 Conclusion

We study the existing transaction network in Japanese firms by analyzing the firms' degree distribution, clustering coefficient, and degree correlation. We discover the following three important characteristics.

First, we find that an undirected network is a scale-free network. Second, we discover that the network has a hierarchical structure. Third, we realize that there exists a degree correlation.

As mentioned earlier, we believe that the study of a real network will lead to further research that will reveal the hidden relation between the underlying network structure and the economy. We need information on the real network structure when we build models on networks. In many other fields, it has been discovered that the micro and macro properties are dependent on the network structures. They depend on, for example, whether the network structures are random or scale-free networks, hierarchical structure, clustering coefficients, degree correlation, and so on. We expect that a similar relation will be discovered in Economics as well. In this sense, the study of a real network toward obtaining the above information is crucial.

# References

Abramson, G., and Kuperman, M. (2001). Social Games In a Social Network. *Physical Review E*, Vol. 6303, No. 3, p. 4.

Aoki, M., and Yoshikawa, H. (2006). *Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes*. Cambridge University Press.

Barabási, A.L., and Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, Vol. 286, No. 5439, pp. 509–512.

Barabási, A.L., and Albert, R. (2002). Statistical Mechanics of Complex Networks. *Rev. Mod. Phys.*, Vol. 74, No. 1, pp. 47–97, Jan.

Barabási, A.L., and Oltvai, Z.N. (2004). Network Biology: Understanding the Cell's Functional Organization. *Nature Reviews Genetics*, Vol. 5, No. 2, pp. 101–113. 10.1038/nrg1272.

Di Bernardo, M., Garofalo, F., and Sorrentino, F. (2005). Synchronizability of Degree Correlated Networks. Arxiv preprint cond-mat/0504335.

Di Bernardo, M., Garofalo, F., and Sorrentino, F. (2007). Effects of Degree Correlation on the Synchronization of Networks of Oscillators. *International Journal of Bifurcation and Chaos*, Vol. 17, No. 10, pp. 3499–3506.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.U. (2006). Complex Networks: Structure and Dynamics. *Physics Reports*, Vol. 424, No. 4-5, pp. 175–308.

Callaway, D.S., Hopcroft, J.E., Kleinberg, J.M., Newman, M.E.J., and Strogatz, S.H. (2001). Are Randomly Grown Graphs Really Random? *Physical Review E*, Vol. 64, No. 4, p. 41902.

Dorogovtsev, S.N., and Mendes, J.F.F. (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW.* Oxford University Press.

Erdos, P. and Renyi, A. (1959). On Random Graphs. *Publ. Math. Debrecen*, Vol. 6, pp. 290–297.

Garlaschelli, D., Battiston, S., Castri, M., Servedio, V.D.P., and Caldarelli, G. (2005). The Scale Free Nature of Market Investment Network. *Physica A*, Vol. 350, No. 2-4, pp. 0310503–491.

Motter, A.E., Zhou, C.S., and Kurths, J. (2005). Enhancing Complex-Network Synchronization. *Europhysics Letters*, Vol. 69, No. 3, pp. 334–340.

Newman, M.E.J. (2002). Assortative Mixing in Networks. *Physical Review Letters*, Vol. 89, No. 20, p. 208701.

Newman, M.E.J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, Vol. 45, No. 2, pp. 167–256.

Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M.A. (2006). A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks. *Nature*, Vol. 441, No. 7092, pp. 502–505.

Ravasz, E. and Barabási, A.L. (2003). Hierarchical Organization in Complex Networks. *Physical Review E*, Vol. 67, No. 2, p. 026112.

Saito, Y., Watanabe, T., and Iwamura, M. (2007). Do Larger Firms Have More Interfirm Relationships? *Physica A*, Vol. 383, No. 1, pp. 158–163.

Santos, F.C., and Pacheco, J.M. (2005). Scale-Free Networks Provide a Unifying Framework for the Emergence Of Cooperation. *Physical Review Letters*, Vol. 95, No. 9, p. 098104.

Sorrentino, F., Di Bernardo, M., Cuellar, G.H., and Boccaletti, S. (2006). Synchronization in Weighted Scale-Free Networks with Degree–Degree Correlation. *Physica D: Nonlinear Phenomena*, Vol. 224, No. 1-2, pp. 123–129.

Souma, W., Fujiwara, Y., and Aoyama, H. (2006). Change of Ownership Networks in Japan. *Practical Fruits of Econophysics: Proceedings of the Third Nikkei Econophysics Symposium.*

Vega-Redondo, F. (2007). *Complex Social Networks.* Cambridge University Press.

Vukov, J. and Szabó, G. (2005). Evolutionary Prisoner's Dilemma Game on Hierarchical Lattices. *Physical Review E*, Vol. 71, No. 3, p. 036133.