J Econ Interac Coord (2008) 3:73–80
DOI 10.1007/s11403-008-0035-5

REGULAR ARTICLE

Correlated performance of firms in a transaction network

Yuichi Ikeda · Hideaki Aoyama · Hiroshi Iyetomi · Yoshi Fujiwara · Wataru Souma

Published online: 21 March 2008

Abstract The correlation of firms' performance on a transaction network is studied by analyzing financial and transaction data. Statistically significant correlation coefficients are obtained as evidence for the firm interactions. The firm interactions are taken into account in the basic equation of firm activity. Forty percent of residuals are explained by considering the firm interactions. The overall structure of the transaction network, i.e., the connectivity of industry sectors, is analyzed.

Keywords Complex network · Firm activity · Network growth · Connectivity · Econophysics

1 Introduction

Heterogeneous interacting firms are among the fundamental constituents of an industrial economy. The heterogeneity of firms can be characterized by their size, ratio, and industry sector. It is known that the performance of firms is strongly correlated as a consequence of their direct interactions in transactions. Although we intuitively expect

Y. Ikeda (✉)
Hitachi Research Institute, 4-14-1 Soto-kanda, Chiyoda-ku, Tokyo 101-8010, Japan
e-mail: yuichi.ikeda.ze@hitachi.com

H. Aoyama
Department of Physics, Kyoto University, Kyoto 606-8501, Japan

H. Iyetomi
Department of Physics, Niigata University, Ikarashi, Niigata 950-2181, Japan

Y. Fujiwara · W. Souma
NiCT/ATR CIS, Applied Network Science Lab,
2-2-2 Hikari-dai, Seika-chou, Kyoto 619-0288, Japan

the performance of firms to be correlated, few studies have addressed a fundamental understanding of this correlation. In this paper, we study the correlation of firms' performance in a transaction network by analyzing a large set of financial and transaction data for Japanese industry.

This paper is organized as follows. First, the evidence for correlated performance obtained from an analysis of transaction and financial data is shown. Then, the basic equation of firm activity is proposed by taking into account direct interaction between firms. Finally, the overall structure of the transaction network, such as connections between industry sectors, is analyzed.

2 Evidence for correlated performance

Following is a brief review of current models for firm activity. Basic equations without interaction, based on a stochastic multiplicative process (SMP), have been widely used. One example is SMP with additive noise (Kesten 1973):

$$R_i(t+1) = X_i(t)R_i(t) + A_i(t).$$

Here R is the revenue, X is the growth rate of revenue, and A is additive noise. Another example is stochastic multiplicative process with reset events (Manrubia and Zanette 1999):

$$R_i(t+1) = \begin{cases} X_i(t)R_i(t) & \text{with } 1-p \\ R_i(0) & \text{with } p. \end{cases}$$

Although these models reproduce the power law distribution for firm size and the Laplace distribution for growth rate (Fujiwara et al. 2004, Pammolli et al. 2007), they do not pay attention to the interaction between firms. Thus, if firms' performances are strongly correlated, the models without interaction are not satisfactory.

Next we describe attributes of the analyzed data. The transaction database contains suppliers and customers for the 1,405 firms which were listed in the first section of the Tokyo Stock Exchange at Japanese Fiacal Year (JFY) 2003. The data is sold by Teikoku Databank Ltd. The financial database contains income statements and balance sheets for the period JFY1993–JFY2003. Note that this is annual data, thus the length of the time series is very short. The statistical reliability of such short dataset is examined by a statistical test, as explained later. The data is sold by Nikkei NEEDS. Figure 1 shows the observed distribution of growth rate. On the left is the growth rate of cost $Y_i(t) = C_i(t+1)/C_i(t)$ and on the right is the growth rate of revenue $X_i(t) = R_i(t+1)/R_i(t)$. The central parts of both distributions are characterized by the Laplace distribution and outer parts exhibit fat-tail. The QQ-plots of normalized growth rate are shown in Fig. 2. The left-hand figure is for the normalized growth rate of cost $\delta Y_i(t) = (Y_i(t) - \bar{Y}_i)/\sigma_i^Y$ and that on the right is for the normalized growth rate of revenue $\delta X_i(t) = (X_i(t) - \bar{X}_i)/\sigma_i^X$. Here $\bar{Y}_i$, σ_i^Y, $\bar{X}_i$, σ_i^X are the average growth rate of cost, the standard deviation of the growth rate of cost, the average growth rate of revenue, and the standard deviation of the growth rate of revenue, respectively.

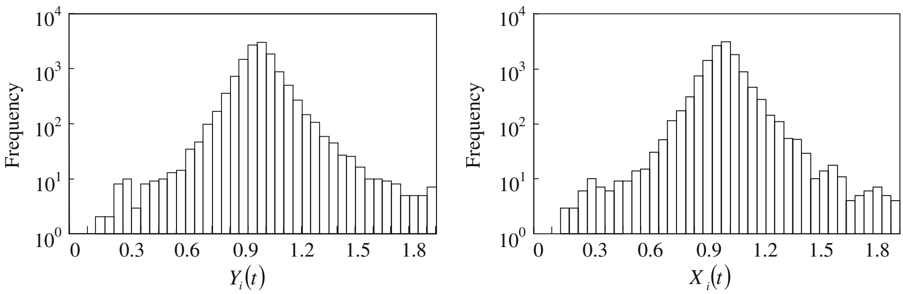

Fig. 1 Observed distributions of growth rate. The figure on the *left* shows the growth rate of cost $Y_i(t)$ and on the *right* is the growth rate of revenue $X_i(t)$

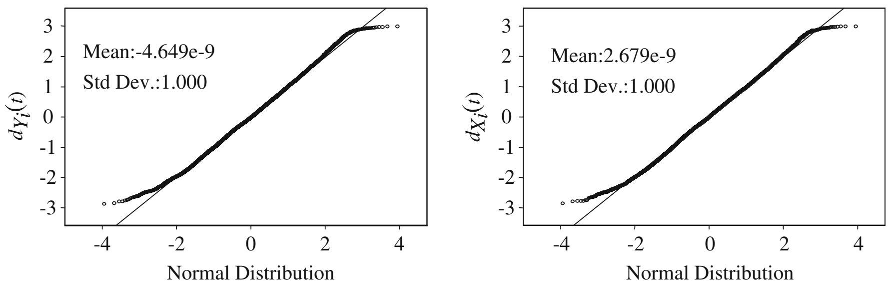

Fig. 2 QQ-plots of normalized growth rate. The figure on the left is the normalized growth rate of cost $\delta Y_i(t)$ and the right is the normalized growth rate of revenue $\delta X_i(t)$

The central parts of both distributions are characterized by the normal distribution and outer parts exhibit fat-tail. This implies the Laplace distribution is obtained as a superposition of the normal distributions with different standard deviations.

Then, mean values of the correlation coefficients of the growth rate are calculated. The correlation coefficient of growth rates is defined by $\rho_{ij}^{\alpha\beta} = \langle \delta\alpha_i(t)\delta\beta_i(t)\rangle(\alpha = \{X, Y\}, \beta = \{X, Y\})$. In Table 1, the correlation coefficients between X, Y, YX, and XY are shown for pairs without and with transactions. From this comparisons, we find that these coefficients are almost the same magnitude: $Mean[\rho_{ij}^{(XX)}] \approx Mean[\rho_{ij}^{(YX)}] = Mean[\rho_{ij}^{(XY)}]$; correlation coefficients with transactions are larger than those without. The distribution of the correlation coefficients between Y and X is shown in Fig. 3. The distribution of pairs with transactions is right-skewed, compared with the distribution of pairs without transaction. We presume that large correlations without transactions are caused by second neighbor links, because average correlation coefficient for second neighbor links is as large as 60% of coefficient for transaction pairs [see Fig. 2c of (Souma et al. 2006)]. A portion of negative correlations corresponds to insignificant transactions due to small transaction value. If the network contains just two firms, the correlation coefficient is always positive. It is however possible that a firm pair with small transaction value shows negative correlation in the network consisting many firms. Thus we consider the negative correlations as a kind of noise.

Table 1 Correlation coefficients

Type	$Mean[\rho_{ij}^{(XX)}]$	$Mean[\rho_{ij}^{(YY)}]$	$Mean[\rho_{ij}^{(YX)}]$	$Mean[\rho_{ij}^{(XY)}]$
Without transaction	0.1475	0.1381	0.1432	0.1432
With transaction	0.2435	0.2194	0.2352	0.2348

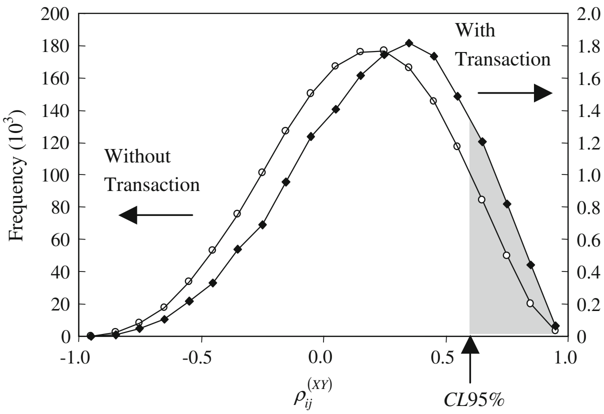

Fig. 3 Distributions of the correlation coefficients between Y and X. The distribution of pairs with transactions is right-skewed, compared with the distribution without transactions. The shaded portion of the distribution indicates statistically significant correlation coefficient at $CL95\%$

Noise filtering techniques such as the random matrix theory can not be applied to the correlation coefficients because the number of firms N is much larger than the length of time series L. Thus, we use the test for zero correlation (Kendall and Stuart 1973). In this test the null hypothesis, two variables are mutually uncorrelated, is used. If $\rho = 0$, variables $T = r\sqrt{L-2}/\sqrt{1-r^2}$ calculated from sample correlation coefficients r are governed by the t-distribution with $L-2$ degrees of freedom, where L is the number of time series data. The null hypothesis is rejected with a $1 - P\%$ confidence level (CL). The shaded portion of the distribution in Fig. 3 indicates statistically significant correlation coefficient at $CL95\%$. Statistically significant correlation coefficients give evidence of correlated performance. As a result, we show that the interactions between firms must be taken into account in the basic equation of firm activity.

In summary, the widely used equations based on a stochastic multiplicative process do not consider firm interaction due to transactions. Statistically significant correlation coefficients were obtained as evidence for firm interaction, showing that firm interactions must be taken into account in the basic equation of firm activity.

3 Basic equation of firm activity

A filtered transaction network, where 940 of the 1405 firms are connected, was obtained by filtering correlation coefficients on the original transaction network with $CL95\%$. We estimated the cumulative degree distributions for the filtered network, shown in

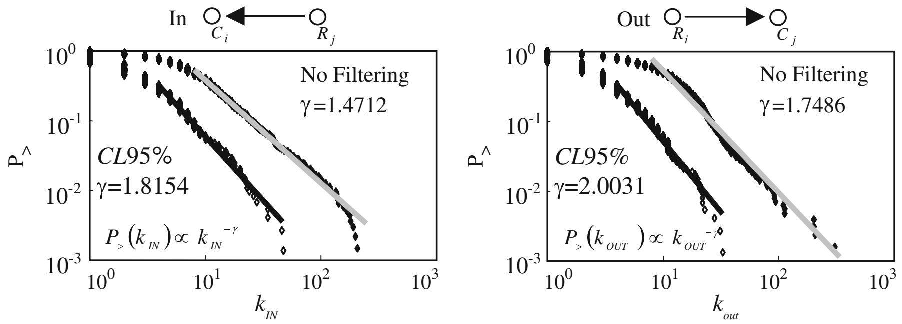

Fig. 4 Cumulative degree distributions for the filtered network. The figure on the *left* is the in-degree distribution and on the *right* is the out-degree distribution. Both filtered and non-filtered degree distributions are characterized by the power law distribution

Fig. 4. On the left is the in-degree distribution and on the right is the out-degree distribution. Both filtered and non-filtered degree distributions are characterized by the power law distribution. The exponents γ for the filtered distributions are close to two. Roughly speaking, the number of filtered degree is smaller than that of non-filtered degree by one order of magnitude.

Now we introduce firm interaction into the basic equation of firm activity:

$$\delta Y_i(t+1) = \sum_{j=1}^{L-1} k_{ij}\delta X_j(t) + \epsilon_i(t), \tag{1}$$

and

$$\delta X_i(t+1) = \sum_{j=1}^{L-1} k_{ij}\delta Y_j(t) + \epsilon_i(t). \tag{2}$$

Equation (1) is for the growth rate of cost and Eq. (2) is for the growth rate of revenue. The first term of the RHS is the interaction term. Parameters k_{ij} are estimated using regression analysis with residual ϵ. These equations are the foundation of our agent-based model of firms (Ikeda et al. 2007).

The QQ-plots in Fig. 5 show that the distribution of the residual is a normal distribution. The standard deviations are about 0.8 for both cases, and less than those of δY and δX: $StdDev[\epsilon_i(t)] < StdDev[\delta Y_i(t)]$ and $StdDev[\epsilon_i(t)] < StdDev[\delta X_i(t)]$. This means the firm interactions of Eqs. (1) and (2) were successfully introduced. Then, we lowered the CL of the filtering in order to take into account more numbers of interaction pairs in the basic equation. If the number of interaction pairs exceeded $L-1$ by lowering the CL of the filtering, then interaction pairs with larger correlations were selected. Figure 6 shows that, at $CL50\%$, forty percent of residuals are explained by considering firm interactions.

In summary, 940 linked firms among 1,405 firms were obtained by filtering the transaction network with $CL95\%$. In-degree and out-degree distributions are characterized by the power law with the exponents $\gamma = 2$. Interactions between firms were

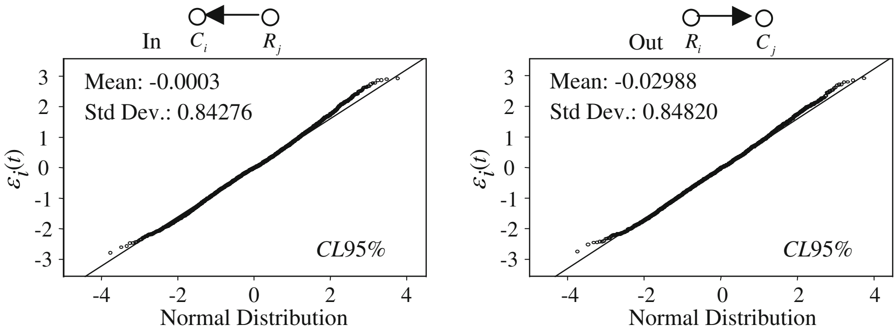

Fig. 5 The QQ-plots show that the distribution of the residual is a normal distribution. The standard deviations are about 0.8 which is less than those of δY and δX. This means the firm interactions of Eqs. (1) and (2) were successfully introduced

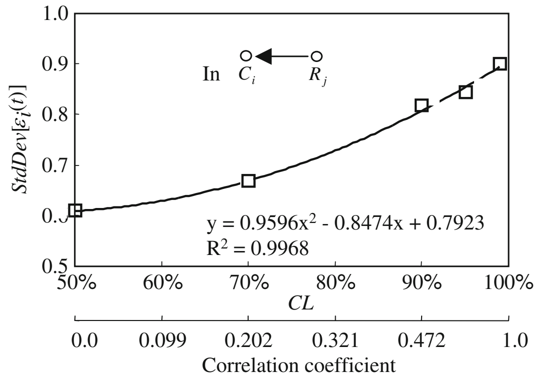

Fig. 6 The CL of the filtering was lowered in order to take into account larger numbers of interaction pairs in the basic equation. Forty percent of residuals are explained by considering the firm interactions at $CL50\%$

taken into account in the basic equation of firm activity. Forty percent of residuals are explained by considering firm interactions.

4 Overall structure of transaction network

Finally, we discuss the overall structure of the transaction network. The degree for an industry sector gives an insight into the overall structure of the network, i.e., the connectivity of the industry sector. The idea of the connectivity for network had been introduced in Broder et al. (2000). The in-degree for sector k^S_{IN} and the out-degree for sector k^S_{OUT} are defined by $k^S_{IN} = \sum_{i\in S}\sum_j f(\rho^{YX}_{ij} - \rho_{CL})$ and $k^S_{OUT} = \sum_{i\in S}\sum_j f(\rho^{XY}_{ij} - \rho_{CL})$, respectively. Here ij is a linked pair, ρ_{CL} is the threshold for statistically significant correlation coefficients at the given CL, and

$$f(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0). \end{cases}$$

Table 2 Strongly connected components with $CL95\%$

In-degree		Out-degree	
Industry sector	k^S_{IN}	Industry sector	k^S_{OUT}
Electrical machinery	581	Electrical machinery	527
Mercantile houses	453	Mercantile houses	472
Chemicals	243	Chemicals	287
Construction	232	Machinery	239
Machinery	221	Construction	188

Table 3 Upstream components with $CL95\%$

Industry sector	k^S_{IN}	k^S_{OUT}	$(k^S_{OUT} - k^S_{IN})/k^S_{IN}$
Mining	0	5	–
Warehousing	3	32	9.667
Land transportation	16	51	2.188
Shipbuilding	4	12	2.000
Air transportation	1	3	2.000

Table 4 Downstream components with $CL95\%$

Industry sector	k^S_{IN}	k^S_{OUT}	$(k^S_{IN} - k^S_{OUT})/k^S_{OUT}$
Retail	88	4	21.00
Gas	19	3	5.333
Electric power	99	24	3.125
Railroads	21	7	2.000
Telecommunication	24	9	1.667

The cumulative distributions of the in-degrees and out-degrees for each sector are obtained by analyzing the transaction data. The industry sectors with the five highest degrees are shown in Table 2. These sectors are recognized as the main part of the strongly connected component. In Tables 3 and 4, the upstream and the downstream components are specified by the difference between in-degree and out-degree, respectively. The upstream sectors consist of raw materials and infrastructures, and the downstream sectors are mainly for consumers. The specified components are shown as a connectivity graph of industry sectors in Fig. 7. This figure indicates the overall structure of the transaction network.

In summary, the degrees for industry sectors provide insights into the overall structure of the network. The upstream, downstream, and strongly connected components are specified by the in-degree and out-degree for each sector, and are summarized as a connectivity graph of industry sectors.

5 Conclusions

We studied the correlation of firms' performance in a transaction network by analyzing financial and transaction data. Statistically significant correlation coefficients were

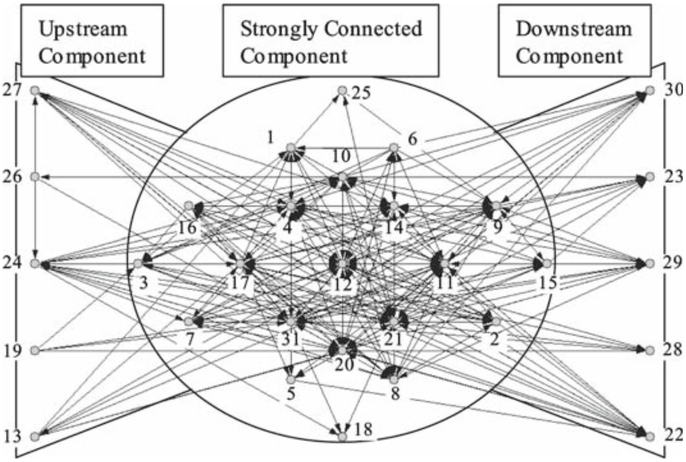

Fig. 7 The specified components are shown as a connectivity graph of industry sectors. Each number indicates an industry sector: *1* Food products, *2* Textiles, *3* Paper manufacturing, *4* Chemicals, *5* Pharmaceuticals, *6* Oil, *7* Rubber, *8* Ceramics, *9* Steel, *10* Nonferrous metal, *11* Machinery, *12* Electrical machinery, *13* Shipbuilding, *14* Auto-manufacturing, *15* Transport machines, *16* Precision apparatus, *17* Other manufacturing, *18* Fisheries, *19* Mining, *20* Construction, *21* Mercantile houses, *22* Retail, *23* Railroads, *24* Land transportation, *25* Marine transportation, *26* Air transportation, *27* Warehousing, *28* Telecommunication, *29* Electric power, *30* Gas, and *31* Service companies

obtained as evidence for the firm interaction. Then, interactions between firms were taken into account in the basic equation of firm activity. Forty percent of residuals were explained by considering the firm interactions. Finally, the overall structure of the transaction network was analyzed, and a connectivity graph of industry sectors was obtained.

References

Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, Wiener JL (2000) In: Proceedings of the 9th international world wide web conference. Elsevier, Amsterdam p 309

Fujiwara Y, Di Guilmi C, Aoyama H, Gallegati M, Souma W (2004) Do Pareto-Zipf and Gibrat laws hold true? An analysis with European firms. Physica A 335(1–2):197–216

Ikeda Y, Souma W, Aoyama H, Iyetomi H, Fujiwara Y, Kaizoji T (2007) Quantitative agent-based firm dynamics simulation with parameters estimated by financial and transaction data analysis. Physica A 375(2):651–667

Kendall MG, Stuart A (1973) The advanced theory of statistics, 3rd edn, vol 1. Griffin, London, p 415

Kesten H (1973) Random difference equations and renewal theory for products of random matrices. Acta Math 131:207–248

Manrubia SC, Zanette DH (1999) Stochastic multiplicative processes with reset events. Phys Rev E 59:4945–4948

Pammolli F, Fu D, Buldyrev SV, Riccaboni M, Matia K, Yamasaki K, Stanley HE (2007) A generalized preferential attachment model for business firms growth rates I. Empirical evidence. Eur Phys J B 57(2):127–130

Souma W, Aoyama H, Fujiwara Y, Ikeda Y, Iyetomi H, Kaizoji T (2006) Correlation in business networks. Physica A 370(1):A151–155