Economics Bulletin

Volume 32, Issue 4

How links are formed according to the distance in transaction network of 800,000 firms

Tomohiko Konno
Princeton University

Abstract

We studied how links are formed according to the distance in transaction network of 800,000 firms. We investigated the inter-firm transaction network of $800,000$ firms in Japan. They are almost all firms in Japan. As far as we know, scale-freeness, hierarchy, and degree-degree correlation were discovered as fundamental characteristics in the network before the present paper. We discovered another fundamental characteristic as to distance. Distance plays an important role in the network. The number of transaction links between two regions is proportional to the product of the sales (degrees) of the two regions divided by the distance between the two regions. This is of a form of the gravity equation. It implies that the further the firms are, the less likely it is that they are linked in the network.

Citation: Tomohiko Konno, (2012) "How links are formed according to the distance in transaction network of 800,000 firms", *Economics Bulletin*, Vol. 32 No. 4 pp. 3279-3286.
Contact: Tomohiko Konno - tomohiko@princeton.edu.
Submitted: September 28, 2012.   **Published:** November 29, 2012.

1. Introduction

The investigation of the fundamental characteristics of inter-firm transaction networks is significant, because it has been well recognized that an underlying network changes the outcome of a model(see reviews by Goyal (2007); Jackson (2008); Vega-Redondo (2007); S.N.Dorogovtesev and J.F.F.Mendes (2003); Newman (2003); Boccaletti et al. (2006)). The same models on different networks can have different outcomes. This fact has motivated us to study networks. Therefore, when we attempt to build a model on a transaction network, we must know the fundamental characteristics of the real transaction network.

Since large data set was not available before, we could not know the characteristics previously. For example, we need a network with a lot of nodes to find a fat tail in degree distribution. The fat tail is one thing which determines the outcome of a model in many cases. However, we obtain such a large data containing almost all the firms in Japan recently. The data contains the transaction information of $800,000$ firms. The data opened the door to investigating the fundamental characteristics of the network. That the transaction network is scale-free network, has hierarchical structure, and has degree-degree correlation were discovered so far (Konno (2009); Saito et al. (2007)). We discovered another fundamental characteristic in the network. We showed that the distance also plays an important role in the transaction network. Why distance still matters in transaction networking with the advancement of IT? It is possibly because the transportation cost and face-to-face communication cost. Even though information technology made it easier to communicate with someone in distant than ever, face-to-face communication is most efficient and still much important.

Two main topics in network are as follows. One is how the outcome changes depending on the underlying network, which is already mentioned, and second is network formation mechanism. When we study a network formation mechanism, the discovery that distance also plays an important role in the network is significant. For example, a famous scale-free network formation mechanism called BA preferential attachment (Barabasi and Albert (1999)) does not make the scale-free network which has the same characteristics as to distance observed in the real transaction network. Gathering the fundamental characteristics of the network will help us study a network formation mechanism and the present paper contributes it to some extent.

A typical form of the gravity equation is as follows.

$$\mathrm{Trade_{ij}} = G \times \frac{\mathrm{GDP_i} \times \mathrm{GDP_j}}{\mathrm{Distance_{ij}}}, \tag{1}$$

where $\mathrm{GDP_i}$ is the GDP of country-i, $\mathrm{Distance_{ij}}$ is the physical distance between country-i and country-j, and G is a constant. Equation (1) explains the total amount of trade between two countries $\mathrm{Trade_{ij}}$ well. The theoretical basis of this gravity equation was studied by Anderson (1979), Helpman and Krugman (1985), Bergstrand (1990), Markusen and Wigle (1990), Eaton and Kortum (1997), and Deardoff (1998). They derived the equation from very different models such as Ricardian, Heckscher-Ohlin, increasing returns to scale model, differentiated goods model, homogeneous goods model, and so on. See also

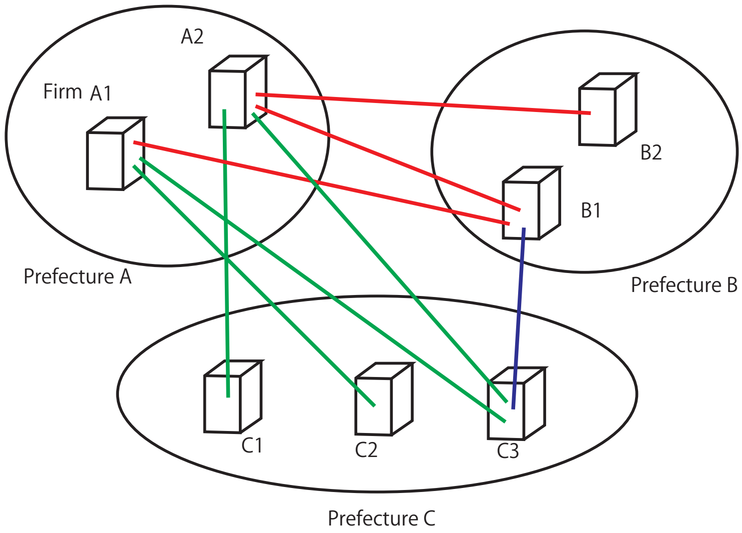

Figure 1: Transaction Network

the survey by Helpman (1998). In applied works, gravity equation often includes other variables in order to account for, language relationships, tariffs, income level, price level, contiguity. Gravity equation has been applied not only to the amount of trade but also to FDI, telephone traffic, merger and acquisition, immigration, hospital patient flow, and so on (Wong (2008); Cieslik and Ryan (2004); Lewer and Berg (2008); Lowe and Sen (1996)).

2. Gravity Equation of the Transaction Network

We found that the following equations as to distance hold well in the inter-firm transaction network of 800,000 firms in Japan.

$$\mathrm{Weight_{ij}} = G_1 \times \frac{\mathrm{Sales_i} \times \mathrm{Sales_j}}{\mathrm{Distance_{ij}}} \tag{2}$$

$$\mathrm{Weight_{ij}} = G_2 \times \frac{\mathrm{Degree_i} \times \mathrm{Degree_j}}{\mathrm{Distance_{ij}}}, \tag{3}$$

where G_1 and G_2 are constants and the other terms will be explained below. We call eqs. (2,3) network gravity equations. There are 47 prefectures in Japan. $\mathrm{Sales_i}$ is the total sales of all the firms in prefecture i. $\mathrm{Degree_i}$ is the total number of inter-prefecture links of all the firms in prefecture i in the network. $\mathrm{Distance_{ij}}$ is the physical distance between two prefectures. It is not a network distance. The distance is measured between capitals of the prefectures. The weight $\mathrm{Weight_{ij}}$ is the number of links between two prefecture i and j in the network. Hence, $\mathrm{Degree_i} = \sum_j \mathrm{Weight_{ij}}$ holds true and $\mathrm{Weight_{ii}} = 0, \forall i$. Let us explain $\mathrm{Weight_{ij}}$ in detail with the example illustrated in Fig. 1.

There are there prefectures, A, B, and C, and seven firms in Fig. 1. Two firms are in prefecture A and the names of them are A1 and A2. In the same way, let us label other five firms as B1, B2, C1, C2, and C3. A link between two firms means that the transaction between two firms exists. A link in the figure is defined between two firms. In a similar fashion we can define a link between two prefectures and a weight of transactions between two prefectures. For example, look at prefecture A and prefecture B, firm-A1 is linked to B1 and B2, firm-A2 is linked to B1. The three links are between the two prefectures, so that we can regard that prefecture A is linked to prefecture B with weight 3. In this way, we have the weight matrix Weight of the transaction network as

$$\text{Weight} = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 3 & 4 \\ 3 & 0 & 1 \\ 4 & 1 & 0 \end{pmatrix} \end{array}. \tag{4}$$

In this example, we have $\text{weight}_{\text{AB}} = 3$, $\text{weight}_{\text{AC}} = 4$, and $\text{weight}_{\text{BC}} = 1$ as illustrated in Fig. 1. The implications of eqs. (2,3) are mainly twofold. First, the more the product of the sales of two firms is, the more likely it is for the firms to be connected in the network. In other words, the more the sales of a firm is, the more links the firm attracts from other firms. The same two statements hold for the degree instead of sales. Second, the closer two firms are, the more likely it is for the two firms to do transaction and to be connected in the network.

2.1. The Data

The data was made by Tokyo Shoko Research Ltd (TSR). It contains the financial data and the transaction relations such as buy and sell of $800,000$ firms in Japan. It contains almost all firms in Japan. For example, we are able to know from the data that firm A is a supplier to firm B, firm C is a customer for firm D, and so on. If there is a transaction regardless of buy or sell between two firms, we regard that a link is between the two firms in the transaction network. We cannot know the volume and the frequency of a transaction from the data. We can see urban agglomeration in the transaction network. $100,000$ out of $800,000$ firms are in Tokyo. There are about $2,880,000$ links in the inter-prefecture transaction network and 27% of them are connected to Tokyo.

2.2. Gravity Equation for Sales

Now, we regress the following equation, which is summarized in table 1.

$$\text{Weight}_{\text{ij}} = G_1 \times \frac{\text{Sales}_{\text{i}} \times \text{Sales}_{\text{j}}}{\text{Distance}_{\text{ij}}}, \tag{5}$$

where G_1 is a constant. R^2 is 0.47, which tells that the equation holds well. In Fig. 2, $\text{Weight}_{\text{ij}}$ between prefecture i and j is plotted as a function of $\text{Sales}_{\text{i}} \times \text{Sales}_{\text{j}}/\text{Distance}_{\text{ij}}$. The logarithmic plot is illustrated in Fig. 3.

Next, we obtain the exponents in the following equation to check whether the exponents are close to $+1$ and -1. Although β_2 is redundant from the symmetry, we still remain β_2 for comprehensiveness.

$$\text{Weight}_{\text{ij}} = G_1 \times {\text{Sales}_{\text{i}}}^{\beta_1} \times {\text{Sales}_{\text{j}}}^{\beta_2} \times {\text{Distance}_{\text{ij}}}^{\beta_3} \tag{6}$$

Table 1: Estimation: eq. (5)

Variable	Coefficient	(Std. Err.)
$\text{Sales}_\text{i} \cdot \text{Sales}_\text{j}/\text{Distance}_\text{ij}$	82.6e-18	(1.9e-18)
Intercept	968.5	(87.1)

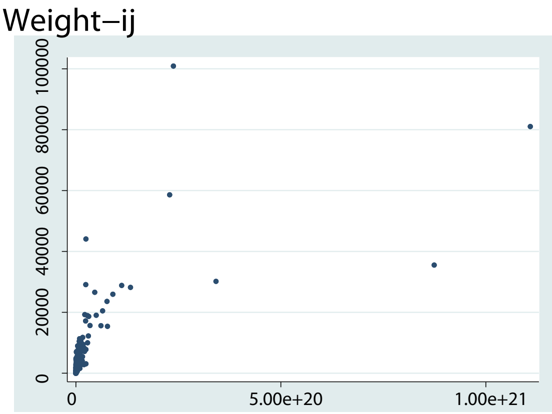

Figure 2: Scatter plot for eq. (5)

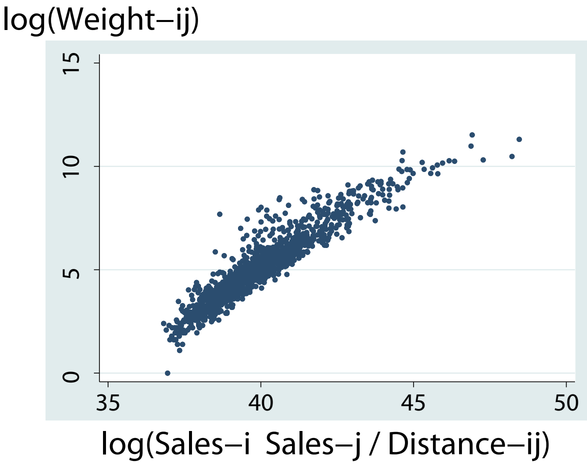

Figure 3: Logarithmic plot for eq. (5)

After taking the logarithm on both sides, we have

$$\log(\text{Weight}_\text{ij}) = \beta_1 \log(\text{Sales}_\text{i}) + \beta_2 \log(\text{Sales}_\text{j}) + \beta_3 \log(\text{Distance}_\text{ij}) + \log(G_1). \qquad (7)$$

We regress eq. (7). The result is summarized in table 2.

Table 2: Estimation results: eq. (7)

Variable	Coefficient	(Std. Err.)
$\log(\text{sales}_\text{i})$	0.969	(0.013)
$\log(\text{sales}_\text{j})$	0.969	(0.013)
$\log(\text{distance}_\text{ij})$	-0.938	(0.017)
Intercept	-33.8	(0.44)

The exponents β_1 and β_2 are close to $+1$ and β_3 is close to -1, which is desired. We have $R^2 = 0.88$. The result shows that the network gravity equation for sales holds for very well.

2.3. Gravity Equation for Degree

We now study gravity equation for degree. Instead of sales, we consider the relation among degree, distance, and weight as follows

$$\text{Weight}_\text{ij} = G_2 \times \frac{\text{Degree}_\text{i} \times \text{Degree}_\text{j}}{\text{Distance}_\text{ij}}. \qquad (8)$$

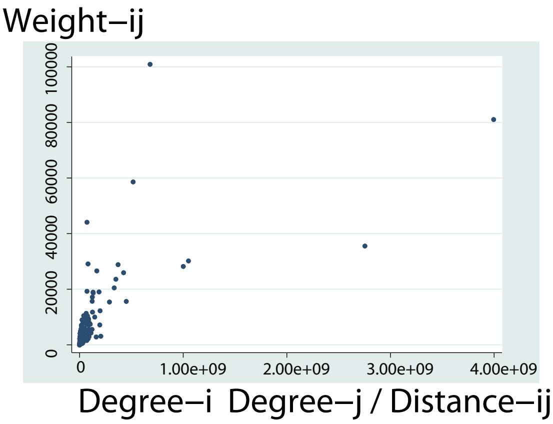

Figure 4: Scatter plot for eq. (8)

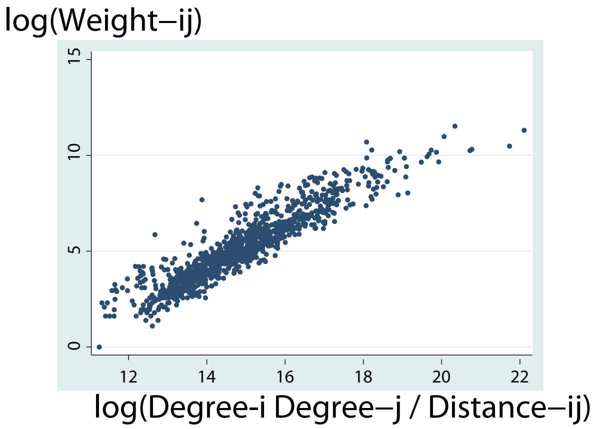

Figure 5: Logarithmic plot for eq. (8)

In Fig. 4, Weight_{ij} is plotted as a function of $\text{Degree}_i \times \text{Degree}_j/\text{Distance}_{ij}$. The logarithmic plot is illustrated in Fig. 5.

Table 3: Estimation: eq. (8)

Variable	**Coefficient**	**(Std. Err.)**
$\text{Degree}_i \times \text{Degree}_j/\text{Distance}_{ij}$	2.4e-5	(0.05e-5)
Intercept	844	(87.1)

Table 3 is the regression result of eq. (8). We have $R^2 = 0.55$. This gravity equation for degree also holds well. Now, we regress the following equation in order to check whether the exponents are close to +1 an −1,

$$\log(\text{Weight}_{ij}) = \beta_1 \log(\text{Degree}_i) + \beta_2 \log(\text{Degree}_j) + \beta_3 \log(\text{Distance}_{ij}) + \log(G_2). \qquad (9)$$

Table 4: Estimation eq. (9)

Variable	**Coefficient**	**(Std. Err.)**
$\log(\text{degree}_i)$	1.17	(0.016)
$\log(\text{degree}_j)$	1.17	(0.016)
$\log(\text{distance}_{ij})$	-0.88	(0.018)
Intercept	-13.9	(0.280)

Table 4 summarizes the regression result of eq. (9). We have $R^2 = 0.86$. As seen from the table 4, the exponents are close to +1 and −1, which is desired. This result shows that the network gravity equation for degree also holds very well.

3. Conclusion

We analyzed the inter-firm transaction network consisting of 800,000 firms in Japan. They are almost all firms in Japan. Other fundamental characteristics of inter-firm transaction network is that the network is scale-free network, has hierarchical structure, and has degree-degree correlation, which is summarized in Konno (2009). In the present paper, We discovered another fundamental characteristic as to distance in the transaction network. We found that the following two gravity-type equations

$$\mathrm{Weight_{ij}} = G_1 \times \frac{\mathrm{Sales_i} \times \mathrm{Sales_j}}{\mathrm{Distance_{ij}}} \tag{10}$$

$$\mathrm{Weight_{ij}} = G_2 \times \frac{\mathrm{Degree_i} \times \mathrm{Degree_j}}{\mathrm{Distance_{ij}}} \tag{11}$$

hold well. $\mathrm{Sales_i}$ is the total sales of all the firms in prefecture i. $\mathrm{Distance_{ij}}$ is the physical distance between two prefectures. The distance is measured between capitals in the prefectures. The weight $\mathrm{Weight_{ij}}$ is the number of links between firms in prefecture i and prefecture j in the network. $\mathrm{Degree_i}$ is the total number of inter-prefecture transaction links of all the firms in prefecture i in the network. Hence, $\mathrm{Degree_i} = \sum_j \mathrm{Weight_{ij}}$ holds. The other three characteristics, that the network is scale-free network, has hierarchical structure, and has degree-degree correlation, are as equally important as “distance” we discovered in this paper when we build a model on transaction network and study a network formation mechanism of the transaction one. True network formation mechanism of the transaction one must have the characteristics as to “distance” and other ones. An underlying network changes the outcome of a model. We can assume a realistic underlying network with the characteristic as to “distance” owing to the discovery in the present paper and other characteristics.

Anderson, JE (1979) “Theoretical Foundation For The Gravity Equation,” *American Economic Review*, Vol. 69, No. 1, pp. 106–116.

Barabasi, Albert-Laszlo and Reka Albert (1999) “Emergence of Scaling in Random Networks,” *Science*, Vol. 286, No. 5439, pp. 509–512.

Bergstrand, JH (1990) “The Heckscher-Ohlin-Samuelson Model, The Linder Hypothesis And The Determinants Of Bilateral Intraindustry Trade,” *Economic Journal*, Vol. 100, No. 403, pp. 1216–1229, DEC.

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

Cieslik, A and M Ryan (2004) “Explaining Japanese direct investment flows into an enlarged Europe: A comparison of gravity and economic potential approaches,” *Journal Of The Japanese And International Economies*, Vol. 18, No. 1, pp. 12–37, MAR.

Deardoff, A (1998) “Determinants of Bilateral Trade: Does Gravity Model Work in a Classical World.”

Eaton, J and SS Kortum (1997) "Technology and Bilateral Trade," *NBER Working Paper*.

Goyal, S. (2007) *Connections: An Introduction to the Economics of Networks*: Princeton University Press.

Helpman, E (1998) "The Structure of Foreign Trade," *NBER Working Paper*.

Helpman, E and PR Krugman (1985) *Market structure and foreign trade*: MIT Press Cambridge, Mass.

Jackson, MO (2008) *Social and economic networks*: Princeton University Press.

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*, Vol. 3, No. 2009-31.

Lewer, Joshua J. and Hendrik Van den Berg (2008) "A gravity model of immigration," *Economics Letters*, Vol. 99, No. 1, pp. 164–167.

Lowe, JM and A Sen (1996) "Gravity model applications in health planning: Analysis of an urban hospital market," *Journal Of Regional Science*, Vol. 36, No. 3, pp. 437–461, AUG.

Markusen, Jr and Rm Wigle (1990) "Explaining The Volume Of North-South Trade," *Economic Journal*, Vol. 100, No. 403, pp. 1206–1215, DEC.

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

Saito, Yukiko Umeno, Tsutomu Watanabe, and Mitsuru Iwamura (2007) "Do larger firms have more interfirm relationships?" *Physica A*, Vol. 383, No. 1, pp. 158–163.

S.N.Dorogovtesev and J.F.F.Mendes (2003) *Evolution of Networks*: Oxford University Press.

Vega-Redondo, F. (2007) *Complex social networks*: Cambridge University Press.

Wong, Wei-Kang (2008) "Comparing the fit of the gravity model for different cross-border flows," *Economics Letters*, Vol. 99, No. 3, pp. 474–477, JUN.