

# Appendix

## **Energy and resilience: The effects of endogenous interdependencies on trade network formation across space among major Japanese firms**

Petr Matous  
The University of Tokyo

Yasuyuki Todo  
RIETI and the University of Tokyo

This appendix explains the method of stochastic actor-based modeling for network evolution. The model conditions on the first observation and tests hypothetical drivers of the evolution of the network, which is treated as a continuous-time Markov chain of single trading link changes between observations.

Between the observations, each firm may receive one or more opportunities in a random order to change its suppliers that are represented by its outgoing ties. The model includes ‘rate effects’ that regulate how often actors receive an opportunity to modify their outgoing ties. These rate effects depend on the number of observed changes. Only one actor acts at any given time, and coordination is not allowed.

Each firm chooses its suppliers to maximize its utility. As in generalized linear models, utility is expressed as a combination of hypothetically relevant network features  $f_i(\beta, x) = \sum_k \beta_k s_{ki}(x)$ . The utility function quantifies the desirability of each possible next state of the network  $x$  among the fixed set of actors from the viewpoint of actor  $i$ . A random component with a standard Gumbel distribution is added to the evaluation function. This procedure is included to respect the stochastic character of network evolution, which results from influences that are unrepresented by nodal or dyadic variables and from measurement errors. Thus, the actor does not necessarily choose the state with the highest utility, although such a choice is most likely. When a firm receives an opportunity to change its suppliers, the options are to create one new tie, delete one existing tie, or do nothing.

Each effect  $s_{ki}$  in the model corresponds to possible reasons why an actor might wish to change a tie or a behavior. These effects express the firm’s supply chain management tendencies. The explanations and mathematical formulas of effects  $s_{ki}$  are presented in Table 2.

The goal of the simulation is to estimate the relative weights  $\beta_k$  for the statistics  $s_{ki}$ . Parameter estimates can be used to compare how attractive various supply chain configurations are while controlling for other exogenous and endogenous effects. The signs of  $\beta_k$  indicate the preferred directions of network change, and their relative magnitudes can be interpreted similarly to parameters of multinomial logistic regression models in terms of the log-probabilities of changes among which the actors can choose.

The estimation was executed in SIENA package version 4 in R (Ripley, Snijders et al. 2012). The method of moments, which depends on thousands of iterative computer simulations of the change process (Snijders 2001), is used to estimate the parameters  $\beta_k$  that enable the reproduction of trading network evolution between 2011 and 2012. There is one target statistic for each estimated effect (for example, the number of ties in the network corresponds to the outdegree effect, the number of reciprocated ties corresponds to the reciprocity effect, and the amount of change in the network corresponds to the rate function). The models all converged with  $T$ -ratios, which quantify the deviations between the simulated and the observed values of the target statistics, between -0.1 and 0.1, indicating an excellent model convergence (Ripley, Snijders et al. 2012). In the final stage of the simulation, the standard errors of the estimated parameters are computed

by the finite difference method, based on the sensitivity of the target statistics to  $\beta_k$ .

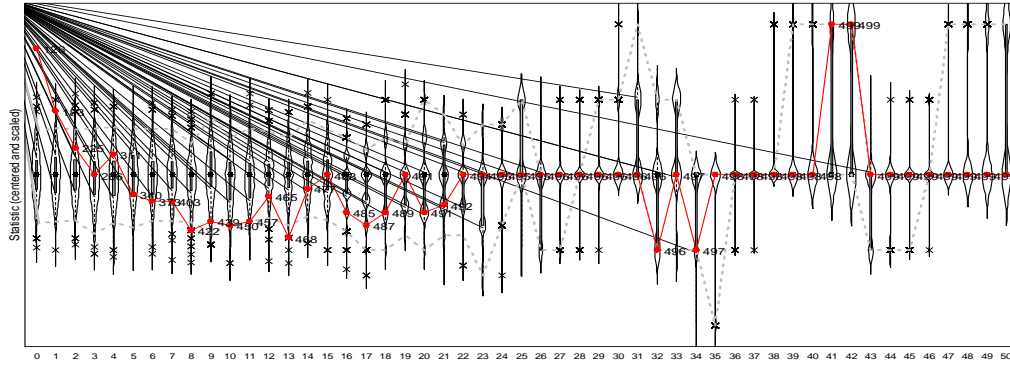
The diagrams below indicate the goodness of fit of the three presented models in terms of indegree distribution, outdegree distribution, geodesic distance distribution, and triadic census, using methods developed by Lospinoso and Snijders (2011).

The violin plots (Hintze and Nelson 1998) represent the kernel density distribution of the statistic and the red lines depict the cumulative distribution of the observed values. The violins are not smooth for less frequent higher degree nodes because the density plots approximate the distribution of a small number of discretely distributed values (Ripley, Snijders et al. 2013).

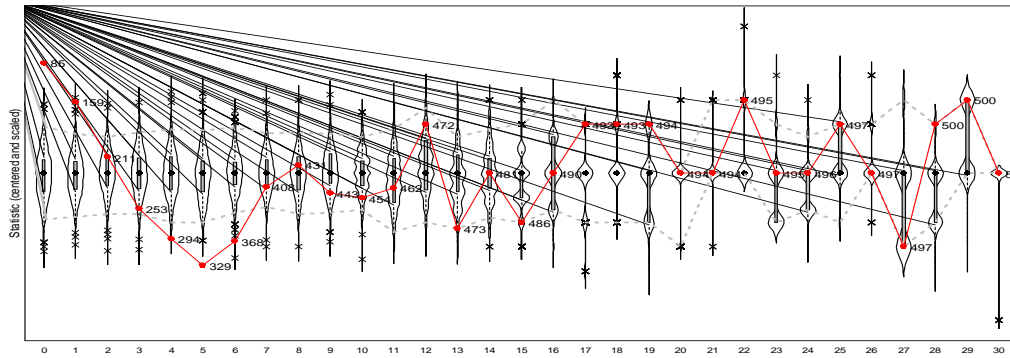
Because the values for different statistics within each plot vary widely, each violin is scaled and centered to maximize the visibility of the plot. The dotted grey lines designate a point-wise 90% relative frequency band for the simulated data. The fit is considered acceptable if the observed values (red lines) fall within this region. However, the goal is not necessarily to match the model exactly on every single statistic, which can be highly irregular. Such an approach would require over-fitting the model to all incidental lone observations or errors in the data and would require the addition of theoretically irrelevant effects.

Standard labeling is used for the classes of the triad census (Wasserman and Faust 1994).

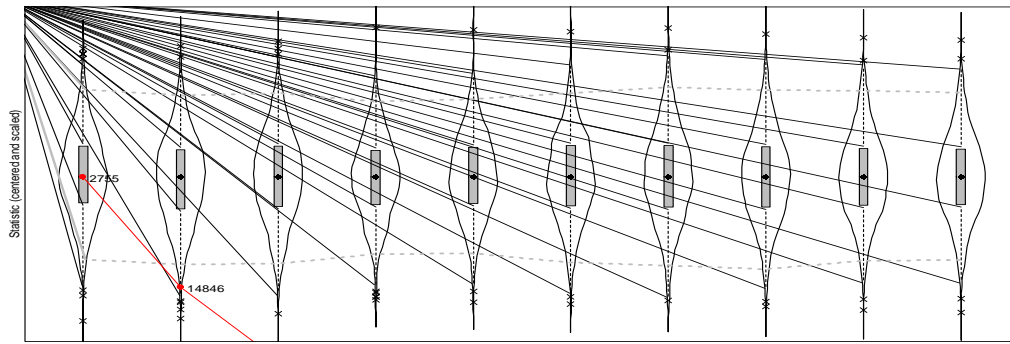
Model 1 – Independent links  
Indegree distribution



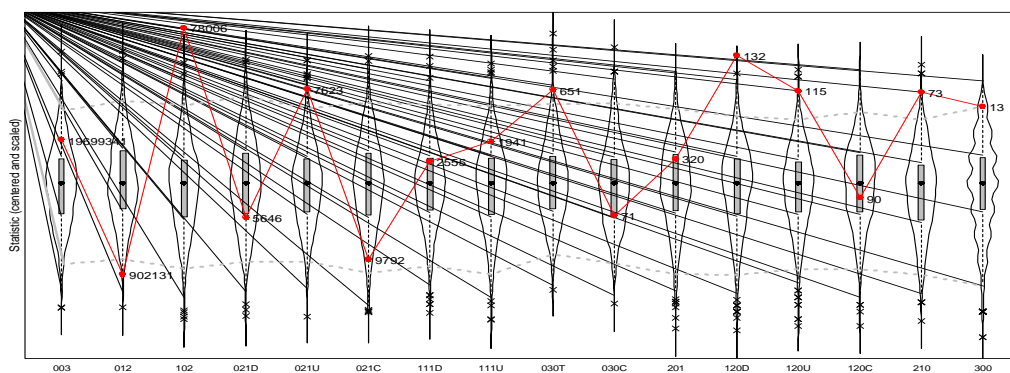
Outdegree distribution



Geodesic distribution

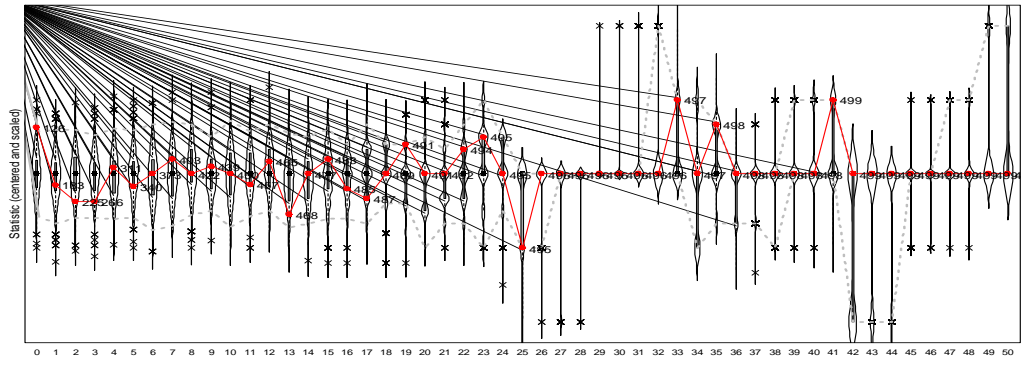


Triad census

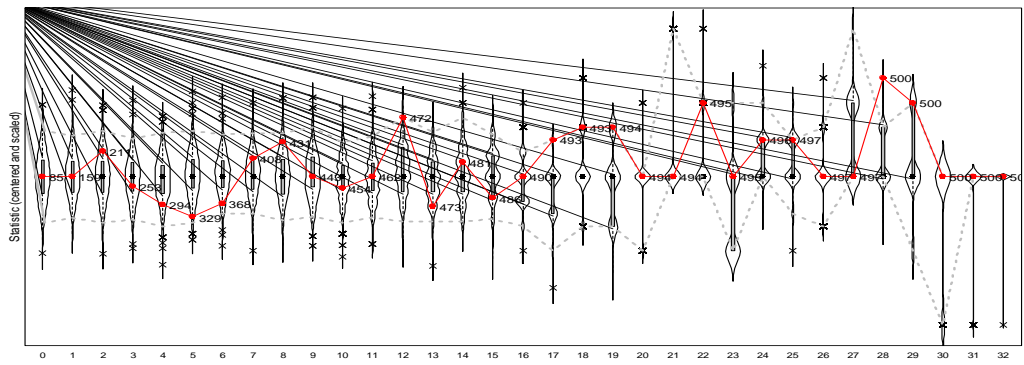


Model 2 – Interdependent links

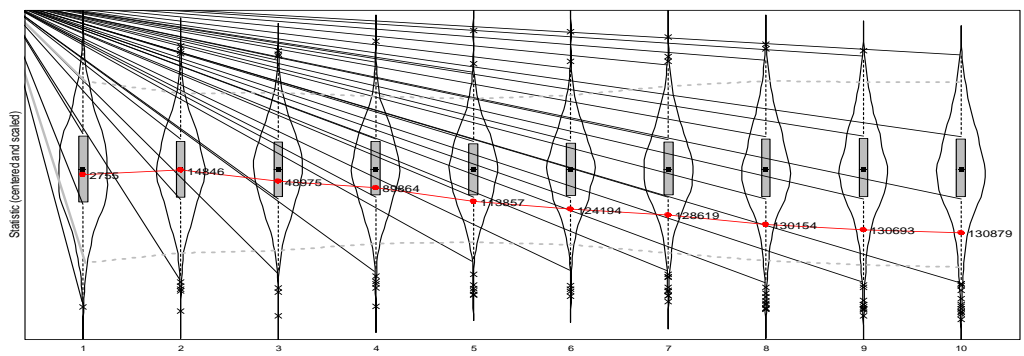
### Indegree distribution



### Outdegree distribution



### Geodesic distribution



### Triad census

