RIETI Discussion Paper Series 19-E-101

# Communities of Investors and Start-up Companies: An Analysis Using Bipartite Stochastic Block Model

**KOJAKU, Sadamori**
Kobe University

**MIYAKAWA, Daisuke**
Hitotsubashi University

RIETI
Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 19-E-101
**December 2019**

# Communities of Investors and Start-up Companies: An Analysis Using Bipartite Stochastic Block Model[1]

SADAMORI KOJAKU[2]
Kobe University

DAISUKE MIYAKAWA[3]
Hitotsubashi University

Abstract

Using unique investment-level data accounting for around 15,000 individual investments done by various types of investors in start-up companies over the periods from 2000 to 2014, we empirically detect community structures consisting of investors and start-up companies through the bipartite stochastic block model and examine their implications on investment performance. The detected community structure represented by multiple groups of investors and start-up companies suggests, first, large heterogeneity of each community in terms of clustered investor types but less so in terms of start-up companies' industry composition. Second, we observe investment performance is higher when the communities are populated by clusters of VCs or non-financial companies. These results jointly imply the systematic concentration of specific types of investors associated with investment performance.

*Keywords*: Start-up firms, investors, investment-level data, stochastic block model
*JEL classification*: M13, G11, G24



[1] This study is conducted as a part of the Project "Creation and Development of High-tech Startups" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The author is grateful for helpful comments and suggestions by Yuji Honjo (Chuo Univ.), Kenta Ikeuchi (RIETI), Makoto Yano (RIETI), Atsushi Nakajima (RIETI), Masayuki Morikawa (RIETI), Fabio Bertoni (Emlyon Business School), and Discussion Paper seminar participants at RIETI. Miyakawa gratefully acknowledges financial support from the Grant-in-Aid for Scientific Research No. 16K03736 and from the grant-in-aid from Zengin Foundation for Studies on Economics and Finance.

2 Research Associate (Postdoc) Kobe University, Research Institute for Economics and Business Administration, skojaku@rieb.kobe-u.ac.jp.

3 Associate professor, Hitotsubashi University Business School (HUB), dmiyakawa@hub.hit-u.ac.jp.

## 1. Introduction

Start-up companies have been considered as a major source for economic growth. Thus, it has been also considered as an especially important policy target for the governments of developed countries to facilitate active business activities of those start-up companies. Specifically because financial constraint is one of the most crucial obstacles those start-up companies are facing, a large number of policy discussions and extant academic studies have been devoted to the issue on how to facilitate finance to start-up companies. In this context, extant literature has been intensively examining how the attributes of venture capital funds (VCs) affect the successfulness of the start-up companies and their investments.

Although our knowledge about the mechanism and the role of such specialized financial institution as VCs have been largely accumulated, there still exist two important open questions we would like to examine. First, investors are essentially diverse. Most of start-up companies are typically founded by individuals and obtain financial resources from various sources including banks, security companies, insurance company as well as non-financial business enterprises, governmental institutions, and universities. In addition to those entities, individuals including angels and Employee Stock Ownership Plan (ESOP) are also important investors for start-up companies. Thus, it is important to take into account those diverse investor types for analyzing the successfulness of the start-up companies and their investments.

Second, start-up companies are also diverse. While we observe many start-up companies in the field of information technology and communication, there also exist a large number of firms starting their businesses in variety of industries such as financial, business service, biotechnology, medical, energy, environment, and consumer products. Thus, it is quite natural to include not only investor heterogeneity but also the heterogeneity associated with start-up companies simultaneously into analyses. As far as we notice, mainly due to the limitation of data, there have been only few studies taking into account the abovementioned two points (i.e., heterogeneity of investor types and start-up companies' industry) at the same time.

Against this background, first, we construct a unique investment-level microdata accounting for around 15,000 individual investments from various types of investors such as venture capital firms, financial institutions, non-financial companies, founders, universities, etc., to start-up companies in various industries over the periods of 2000 to 2014 in Japan. Although the data coverage is not necessarily exhaustive for the universe of start-up companies simply because it is impossible to measure the activities of all the start-up companies and the investments to those companies through privately-placed equity investments, our dataset still accounts for a substantial number of firms establishing IPOs over the periods from 2004 to 2014 in Japan.

Using this dataset, second, we employ the method used in network science to detect multiple communities consisting of start-up companies and investors. To be more specific, we use a type of stochastic block model, which model the interaction among multiple "communities" consisting of nodes represented for by either investors or start-up firms. The model estimates both the set of nodes (i.e., investors or start-up firms) and the probability of interaction (i.e., investments) from a set of investor nodes to a set of start-up firms by fitting the model to the actually observed investment pattern. Taking look at such estimated communities, we can detect what kind of investors, which include various types of investors, tend to invest on what kind of start-up firms in various industries.

Third, in order to extract empirical implications associated with those communities, we test whether each community has specific features represented for by start-up companies' industry, investors' type, investment returns, and post-IPO dynamics (i.e., upgrading the listed market and delisting). The results of these statistical tests allow us to characterize each community in terms of its composition and performance.

Our estimation for the community detection provides 324 communities consisting of 18 groups of start-up companies and 18 groups of investors. First, our statistical test exclusively focusing on the top 10 communities in terms of the total number of investments done in each community suggest that all the 10 communities exhibit statistically distinct heterogeneity in terms of investor types. This suggests the existence of clustering of specific types of investors. As an illustration, three communities out of the top 10 communities are detected to systematically consist of a large number of VCs but a small number of non-financial investors for the two among such three communities. Other three communities consist of a small number of VCs, and the last four communities consist of average number of VCs while other industries such as non-financial investors, ESOP, and/or founder are either over- or under-representing in those four communities.

Regarding the industry composition of start-up companies, however, we identify only two out of the top 10 communities systematically consisting of a large number of companies in specific industries. To be more specific, one community is populated by a large number of biotechnology companies while another one consists of a large number of semi-conductor and energy companies (and a small number of start-up companies in financial industry). All the other eight communities do not exhibit any statistically significant features in terms of start-up firms' industry composition.

These findings imply that it is not apparent for each investor type to be associated with high industry specialization while types of investors are still clustered[1]. On interpretation for these two patterns is that investors in a same type tend to invest together but the target of their investments (i.e., start-up companies) are well diversified. We should note that there is still a chance for each individual investors such as a specific VC tends to focus on some specific industries, and thus there could be industry-specialization in each investor-level. Thus, our results do not preclude the

[1] As we will mention later, there is one community systematically consisting of a large number of biotechnology companies and VCs (with a small number of non-financial investors).

possibility of each investor's industry specialization and the accumulation of those industry expertise. Nonetheless, we believe that it is still informative to confirm the independence of start-up firms' industry composition and investor types.

Second, for those communities characterized above, we examine the association between each community and observed investment performance measured by either benchmark-adjusted investment return, probability of upgrading the listed market (e.g., moving up from emerging market to the first or second section of the Tokyo Stock Exchange (TSE)), and probability of delisting after IPO. Our statistical test suggests that those performance measured for a community tends to become better when investors in the community consist of a large number of VCs or non-financial investors. It is also confirmed that the community consisting of a small number of VCs tend to be associated with low investment return. As an exception, we also find one community consisting of a large number VCs and yet showing low investment return. Notably, such a community systematically consists of a large number of start-up companies in biotechnology industry. Given start-up companies in biotechnology industry are supposed to need large R&D investments and take a longer period for going public, it could be natural for those companies to exhibit lower return.

The results reported above jointly imply the systematic emergence of investor communities in terms of investor type, and the cluster is associated with investment performance. A type of investors (e.g., VCs and non-financial investors) are likely to invest together and exhibit high performance. Such a high performance is not necessarily an artifact of start-up companies' industry characteristics as there are only weak patterns of industry composition in each community. Our results suggest that the interaction among investors in a same type could be a source of highly performing investment through, for example, sourcing activities among them.

The rest of the present paper proceeds as follows. Section 2 explains the empirical approach we take in the present paper. Section 3 details the data we use for our analysis. In Section 4, we show empirical findings. Section 5 concludes and discusses the future research issues.

## 2. Empirical methodology

In this section, we briefly go over our empirical methodology. After introducing a standard notation used in network science, we present the method for detecting community structure and testing statistical patterns associated with features of each community.

### 2.1 Notation

Suppose there are a set of start-up firms $V_c \subseteq \mathbb{Z}$ and a set of investors $V_b \subseteq \mathbb{Z}$. The number of investment from investor-$i$ to start-up firm-$j$ is denoted by $w_{i,j}$. Note that in the latter section, we employ the volume of investment instead of the number of investment to measure $w_{i,j}$. Using $w_{i,j}$, we can define $b_{i,j}$ as $\mathbf{1}\{w_{i,j} > 0\}$, which are the elements of adjacent matrix. Each start-up firm or investor is denoted by a node $v$, which is an element in union of $V_c$ and $V_b$. An edge between investor-$i$ and start-up firm-$j$ is denoted by $e_{i,j}$. Weighted adjacent matrix $W$ is defined as $W_{i,j} = w_{i,j}$. Adjacent matrix $B$ is defined as $B_{i,j} = b_{i,j}$. $E$ is a set of edges between start-up firms and investors. Then, we can define $G$ as a bipartite graph $(V_c \cup V_b, E)$ of investment relation.

### 2.2 Stochastic block model (SBM)

In the present paper, we use Stochastic block model (SBM) proposed by Karrer and Newman (2011). SBM is developed to model the probability of graph $\boldsymbol{G}$ given $\boldsymbol{\omega}$ and $\boldsymbol{g}$ where $\boldsymbol{g} \equiv \{g_k\}$ is group assignment for vertex $k$ and $\boldsymbol{\omega} \equiv \{\omega_{r,s}\}$ denotes $\mathrm{E}[W_{i,j}]$ for $i, j$ lying in group $r, s$. Using the notation we introduced, such a probability of graph $\boldsymbol{G}$ given $\boldsymbol{\omega}$ and $\boldsymbol{g}$ is constructed as follows. We can estimate the $\boldsymbol{\omega}$ and $\boldsymbol{g}$ by maximizing the likelihood based on this probability.

$$P(\boldsymbol{G}|\ \boldsymbol{\omega}\ \ ,\ \boldsymbol{g}\ ) = \prod_{i<j}\frac{\left(\omega_{g_i g_j}\right)^{W_{i,j}} exp\left(-\omega_{g_i g_j}\right)}{W_{ij}!} \times \prod_{i}\frac{\left(\frac{1}{2}\omega_{g_i g_i}\right)^{W_{i,i}/2} exp\left(-\frac{1}{2}\omega_{g_i g_i}\right)}{(W_{ii}/2)!}$$

In order to apply SBM to the bipartite-graph, we rely on the method proposed in Larremore et al. (2014). In our estimation, $w_{i,j}$ is considered by using degree (i.e., $W_{i,j}$ in our data) corrected SBM. We set the number of groups for start-up companies and investors as 18 each, which is based on the method proposed by Peixoto (2015).

2.3 Test for features

To test whether each community has specific features represented by start-up companies' industry composition, investors' type, and post-IPO dynamics (i.e., upgrading and delisting), we employ a following set-up. First, $x_j \in \{0,1\}$ and $x_i \in \{0,1\}$ denotes a feature (i.e., industry, investor type, and delist) for start-up firm-*j* and investor-*i*. Using Kronecker delta $\delta$, we can construct a number of start-up firm-*j* or investor-*i* in group $k$ as follows:

$$B_k = \sum_{i=1}^{N_2} \delta(k, g_i)$$

Then, the number of start-up firm-*j* or investor-*i* in group $k$ with $x_j = 1$ or $x_i = 1$ can be expressed as follows:

$$B_k^{(1)} = \sum_{i=1}^{N_2} x_i \delta(k, g_i)$$

We can generate $b_i \in \{0,1\}$ indepedent of $g_i$ by reshuffling the label and use this as a null model so that we can test if $B_k^{(1)}$ is larger or smaller than $B_k^{(1)}$ under the null model. Under the null model, we assume that $B_k^{(1)}$ follows the binomial distribution:

$$Pr\left(B_k^{(1)} = b\right) = \binom{B_k}{b} p^b (1-p)^{B_k - b} \text{ where } p = \sum_{i=1}^{N} \frac{x_i}{N}$$

Given we have already detected communities consisting for start-up firms and investors, we can apply this method to establish an empirical association between a community and a specific feature.

To test how each community is associated with investment returns, we regress the benchmark adjusted-return measured for each investment on dummy variables representing for each community so that we can see the difference of investment return among multiple communities. In this estimation, we employ quantile regression instead of a standard OLS so that we can explicitly examine the association the upper tail behavior of return distribution of each community.

## 3. Data

In this section, we will go over the data we use in the present study. All the data are obtained from Japan Venture Research Inc. (JVR), through an explicit contract between Hitotsubashi University and JVR for an academic use. The dataset contains detailed investment information measured for each triplet consisting of start-up companies, investors, and investment round. In the present paper, we use the data associated with the investments on start-up companies that eventually accomplished IPO.[2] Each data entry is identified by the abovementioned triplets and associated with the date of investment, the share price of start-up companies paid by investors for purchase at each investment,

[2] Although it is possible to use the additional three datasets corresponding to VBs that either (i) acquired by other companies, (ii) liquidated, or (iii) have not faced IPO, M&A, or liquidation, we are not using this as the issue associated with the comprehensiveness of the dataset we mentioned above is supposed to be severer.

the date of IPO, the initial share price for the start-up companies as of IPO, the initial market the start-up company is listed as of IPO, whether the start-up company changed the listed market after IPO or not (i.e., to the first and second sections of TSE), and whether the VB was delisted or not. The frequency of the data is monthly.

Figure 1 depicts the number of IPOs recorded in our dataset, the actual number of IPOs occurring, and the ratio of the former to the latter (line chart) over the period from 2004 to 2014. We can immediately notice that the dynamics of the number is consistent with the actual number of IPOs in Japanese stock market. As the bar chart (A) in Figure 1, which is smaller than the actual IPO numbers denoted by the bar chart (B), suggests, our dataset is not comprehensive and missing some of the IPO records. Figure 2 depicts the number of investments observed in each year. Table 1 summarizes the IPO markets each start-up company was originally listed as well as the average, median and standard deviation of the number of investors investing to those companies. Among the various features associated with start-up firm-*j* and investor-*i*, Table 2 and Table 3 list the start-up companies' industry classification and investor type, respectively. Table 4 further summarizes the average, median and standard deviation of the number of investor types.

One unique feature of our data is that we can measure investment amounts, the individual investment return that each investor obtained from their investments, and the post-IPO dynamics (e.g., delist) of each VB. Table 5 summarizes the average, median and standard deviation of the invested amounts (thousand JPY) to each start-up company. Such information is not necessarily available in the extant studies. For example, Hochberg et al. (2007) employ the number of investment rounds for measuring VCs' performance. The investment return measure we employ in the present paper is more suitable for characterizing investment performance. As we mentioned in the previous section, to measure the return, we compute the annualized return from each investment by using the abovementioned data and subtract the market return over the same period. To be more precise, we compute the annualized return of the investment from investor-$i$ to start-up company-$j$ implemented at $t$ as $r_{ij}^{(t_{inv}, t_s(j))}$ as in the following form:

$$r_{ij}^{(t_{inv}, t_s(j))} = \left(\frac{s_j}{w_{ij}^{(t)}}\right)^{\frac{t_s(j)-t+1}{365}}$$

In this expression, $t_s(j)$ denotes the time when start-up company-$j$ accomplished IPO. $w_{ij}^{(t)}$ accounts for the price per share paid by investor-$i$ when it invested on start-up company-$j$ at time $t$. $s_j$ accounts for either the initial price or offer price of start-up company-$j$'s share as of IPO. Note that this return is measured for each investment from investor-$i$ to start-up company-$j$. We compute the benchmark-adjusted return by subtracting the market return based on TOPIX over the same period from the return above.

As another performance measure of investment, we track the market movement of each start-up company-$j$ invested by investor-$i$ at $t$. Here, the market movement denotes the delisting from market, staying in the originally listed market, and market upgrade. Table 6 summarizes number of post-IPO dynamics consisting of DELIST_a, DELIST_b, STAY, TSR1, and TSE2. DELIST_a denotes a dummy variable taking value of one if the start-up company delist due to M&A while DELIST_b accounts for the delisting due to other reasons mainly accounted for by violation of listed criteria. STAY accounts for the case that start-up companies stay in the same market as they were originally listed while TSE1 and TSE2 account for the cases that the start-up companies move to another (larger) stock markets.

Using the investment data of investors to start-up companies over multiple investment rounds, we construct bipartite graph which incorporate both investors and start-up companies in the data. As it is bipartite, each node of the network accounts for either start-up companies or investors, and the former (the latter) can be connected only to the latter (the former) through edge.

## 4. Empirical findings

### 4.1 Community detection

Figure 3 illustrates a result of our bipartite degree-corrected SBM. Each cell corresponds to a potential community and the color chart indicates how "active" those communities are. Based on the level of the activity measured by the degree (i.e., the number of investments done within the community), we pick up the top 10 communities from the matrix.

Table 7a summarizes, in those top 10 communities, how many start-up companies and investors are incorporated. For example, the community-1, which is associated with the highest degree (i.e., the number of investments), consist of 60 start-up firms and 204 investors. Table 7b also summarizes the median of the returns based on either initial or offer price as of IPO as well as the initial price decline rate associated with those top 10 communities.

### 4.2 Testing features

As the first test extracting an association between detected communities and a specific feature, we test whether start-up companies included in each group belong to a specific industry or not. InTable 8, we show the results of our test on whether start-up companies in a specific industry (e.g., Communication) is more or less likely to be included in each community. The shaded area denotes the case where a specific industry denoted in the column over-represent in the community denoted in the row. We can confirm that only the 6$^{th}$ and 8$^{th}$ community are associated with some distinct features in terms of start-up firms' industry composition. To be more specific, the 6$^{th}$ community is populated by a large number of biotechnology companies while the 8$^{th}$ community consists of a large number of semi-conductor and energy companies but a small number of start-up companies in finance industry. All the other eight communities do not exhibit any statistically significant features in terms of start-up firms’ industry composition. These findings imply that at least for the top 10 active communities where a large number of investments, there is small industry concentration.

As the second test, we repeat the same test for investor types. Table 9 summarizes the results. First of all, we can immediately confirm that most of the communities show some distinctive feature associated with investor types. 1st, 6th, and 7th communities are more likely to include investors categorized as VCs while the 2nd, 3rd, and 4th communities are less likely to include VCs. This suggests the existence of clustering of specific types of investors. Regarding the former groups (1st, 6th, and 7th), over-representing investor type is only VCs in 1st and 6th groups. Among the latter groups (i.e., 2nd, 3rd, and 4th), there is a single over-representing investor type such as non-financial investors, real estate, and bank. Notably, out of the top 10 communities, only the three groups (i.e., 7th, 8th, and 9th) exhibit multiple over-representing investor types. In fact, the 8th group is overpopulated by Angel and ESOP are majority and 9th group by Founder and ESOP. As the Angel investors are not well specified in our dataset and could be relatives of the founders and some related parties, these two groups can be also considered as the groups with only one type of over-representing investor type. Although VCs invest together with other types of investors such as Shosha (i.e., trading companies, bank, real estate, insurance) in the 7th group, the abovementioned results suggest the large heterogeneity of each community in terms of clustered investor types.

### 4.3 Performance

How should we interpret the detected community structure and the associated features? Table 10 summarizes the performance measured by various metrics for each community. The first two columns show either the investment returns based on the initial price as of IPO is larger in its 75 percent and 90 percent quantile. The third and fourth columns repeat the same exercise for the return based on the offer price as of IPO. The fifth and sixth columns show whether each community is more likely to be associated with market upgrade to TSE first section or second section. The last column shows whether each community is more likely to be delisted or not.

These results suggest that those performance measured by investment return for a community tends to become better when investors in the community consist of a large number of VCs (i.e., 1st community) or non-financial investors (i.e., 2nd and 5th communities). It is also confirmed that the community consisting of a small number of VCs tend to be associated with low investment return (i.e., 3rd community)). As an exception, we also find one community consisting of a large number VCs and yet showing low investment return (i.e., 6th community). Notably, such a community systematically consists of a large number of start-up companies in biotechnology industry. Given start-up companies in biotechnology industry are supposed to need large R&D investments and take a longer period for going public, it could be natural for those companies to exhibit lower return. We should also notice that the likelihood of delisting is higher in the 5th community, which is associated with higher investment returns and populated by a large number of non-financial investors. Interestingly, the 4th community is associated with high likelihood of market upgrading to TSR first section. Given this community is overpopulated by banks as investors, it might be the case that banks aim at long-term growth of their investment targets.

In order to see the robustness of our result, we repeat all the estimation from community detection to performance test by referring the investment amounts instead of investment numbers as the degree and confirm the robustness of our results. As an additional information, when we repeat the same exercise by using only the information at the first round investment, we can confirm only few features associated with each community.

## 5. Conclusion

In the present paper, using unique investment-level data accounting for around 15,000 individual investments from various types of investors to start-up companies over the periods from 2000 to 2014, we empirically detect community structures consisting of investors and start-up companies through the bipartite stochastic block model and examine its implication on investment performance. The detected community structure represented by multiple groups of investors and start-up companies suggests, first, large heterogeneity of each community in terms of clustered investor types but less so in terms of start-up companies' industry composition. Second, we observe investment performance is higher when the communities are populated by clusters of VCs or non-financial companies. These results jointly imply the systematic concentration of specific types of investors associated with investment performance.

As our analysis is focusing only on the top 10 communities detected by stochastic block model, one promising way is to extend the discussion to a larger number of communities and test its statistically association with the performance. Also, it could be informative to focus on more detailed investor types (e.g., VCs backed by different financial sources) and start-up firms' industry classification so that we can see the pattern of clustering and its association with investment performance in more detail.

**Reference**


Larremore, D. B., Clauset, A., and Jacobs, A. Z. 2014. Efficiently Inferring Community Structure in Bipartite Networks. Phys. Rev. E 90, 012805.

Hamao, Y., Packer, F., and J. R. Ritter, 2000. Institutional Affiliation and the Role of Venture Capital: Evidence from Initial Public Offerings in Japan. Pacific-Basing Finance Journal 8, 529–558.

Hellmann, T., L. Lindsey, and M. Puri, 2008. Building Relationships Early: Banks in Venture Capital. Review of Financial Studies 21, pp. 513–541.

Hochberg, Y., Lindsey, L., Westerfield, M. 2015 "Resource Accumulation through Economic Ties: Evidence from Venture Capital. Journal of Financial Economics 118: 245-267.

Hochberg, Y. V., Ljungqvist, A. Lu, Y. 2007. Whom You Know Matters: Venture Capital Networks and Investment Performance. Journal of Finance 62, pp. 251-301.

Hochberg, Y. V., Ljungqvist, A. Lu, Y. 2010. Networks as a Barrier to Entry and the Competitive Supply of Venture Capital. Journal of Finance 65, pp. 829-859.

Karrer, B., and Newman, M. E. J. 2011. Stochastic Block Models and Community Structure in Networks. Phys. Rev. E 83, 016107.

Miyakawa, D., M. Takizawa. 2015. Capital Supply Channel through Venture Capitals: Evidence from Matched Data. RIETI Discussion Paper Series 15-E-141.

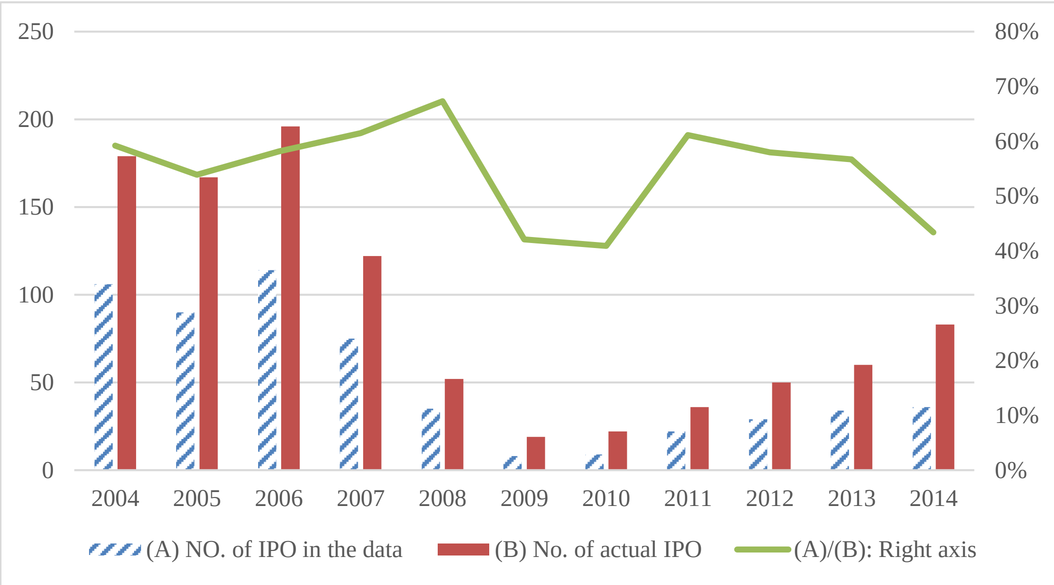
Peixoto, T. P. 2017. Nonparametric Bayesian Inference of the Microcanonical Stochastic Block Model. Phys. Rev. E 95, 012317.

Peixoto, T. P. 2015. Model Selection and Hypothesis Testing for Large-Scale Network Models with Overlapping Groups. Phys. Rev. X 5, 11033 (2015).

Rin, M. D., Hellmann, T., Puri, M. 2013. A Survey of Venture Capital Research. In Constan-tinides, G., Harris, M., Stulz, R. (eds.) Handbook of the Economics of Finance, vol 2, Am-sterdam, North Holland

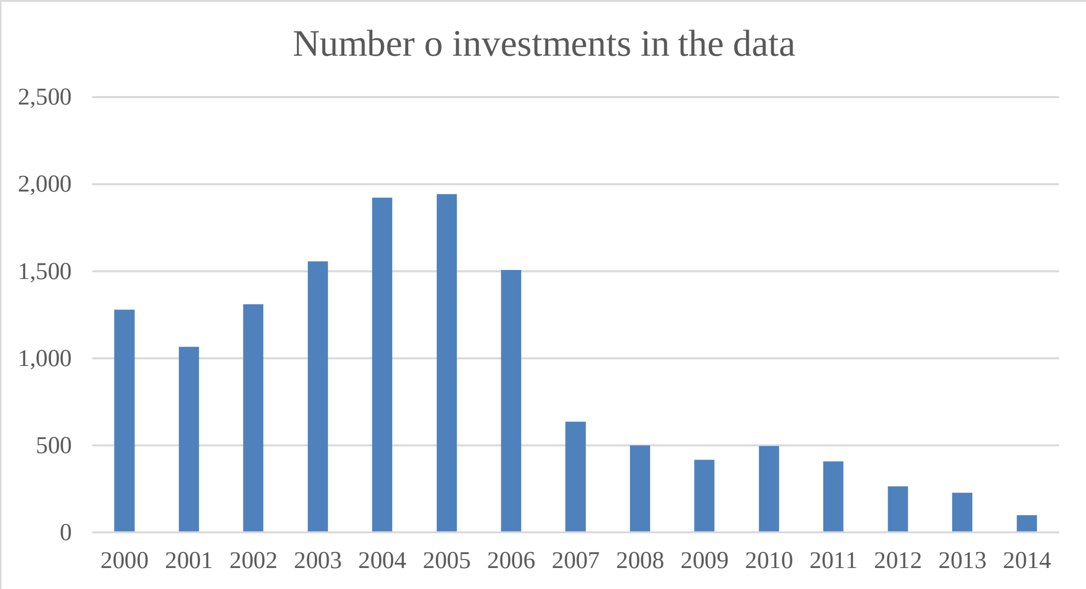
Snijders, T. A. B., 2002. Markov Chain Monte Carlo Estimation of Exponential Random Graph Models. Journal of Social Structure 3.

Figure 1 Evolution of IPO numbers in the data



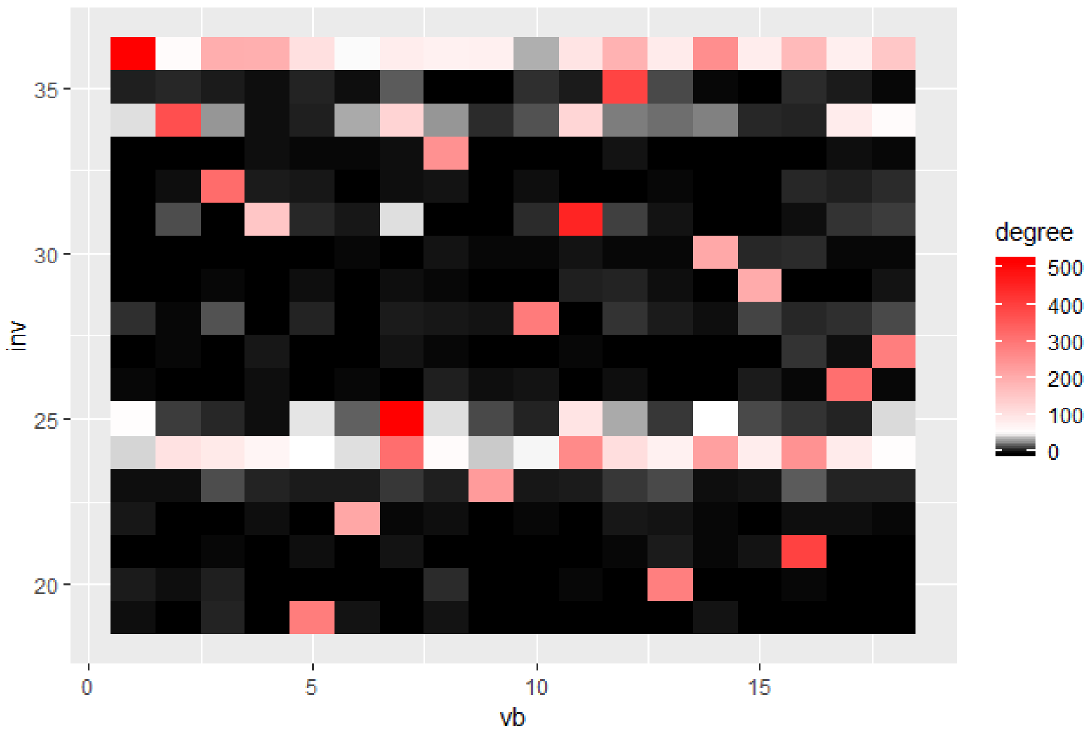

Note: Each bar accounts for the number of IPOs in each year, which is calculated by the data we use for our empirical analysis.

Figure 2 Evolution of investment number (degree)




Note: Each bar accounts for the number of investments to IPO companies in each year, which is calculated by the data we use for our empirical analysis.

Figure 3 Detected communities




Note: Each bar accounts for the number of investments to IPO companies in each year, which is calculated by the data we use for our empirical analysis.

Table 1 IPO markets

| IPO market | #(start-up companies) | #(investors) average | #(investors) median | #(investors) standard deviation |
|---|---|---|---|---|
| JASDAQ | 275 | 10.9 | 8 | 9.5 |
| LOCAL | 48 | 11.5 | 10 | 7.5 |
| MOTHERS | 210 | 12.8 | 11 | 8.3 |
| PRO | 2 | 14.0 | 14 | 8.5 |

Note: The table summarizes the number of start-up companies and the number of investors for each IPO market. All the numbers are calculated by the data we use for our empirical analysis.

Table 2 Industry

| Start-up company industry classification |
|---|
| Communication |
| Finanial |
| Financial-ICT |
| Business service |
| IT hardware |
| IT software |
| IT service |
| Semiconductor & products |
| Biotechnology |
| Medical |
| Medical-ICT |
| Energy |
| Energy-ICT |
| Environment |
| Consumer |

Note: The table summarizes the industry classification corresponding to the star-up companies going to IPO markets.

Table 3 Investor types

| Investor type |
|---|
| Angel |
| Green sheet |
| Individual |
| Founder |
| Unknown |
| VC |
| Shosha |
| Foreign |
| Bank |
| Cloud funding |
| Non-financial |
| ESOP |
| Security companies |
| Government |
| Others |
| Other financial |
| University |
| Real estate |
| Insurance |
| Foreign VC |

Note: The table summarizes the types of investors

Table 4 Investor types

| IPO market | #(start-up companies) | #(investor type) average | #(investor type) median | #(investor type) standard deviation |
|---|---|---|---|---|
| JASDAQ | 275 | 4.3 | 4 | 1.9 |
| LOCAL | 48 | 4.4 | 4 | 1.7 |
| MOTHERS | 210 | 4.9 | 5 | 1.7 |
| PRO | 2 | 5.5 | 6 | 0.7 |

Note: The table summarizes the number of investor types corresponding to the star-up companies going to IPO markets. All the numbers are calculated by the data we use for our empirical analysis.

Table 5 Investment amounts

| IPO market | #(start-up companies) | Investment amounts average | Investment amounts median | Investment amounts standard deviation |
|---|---|---|---|---|
| JASDAQ | 275 | 1,458,221 | 385,374 | 5,583,584 |
| LOCAL | 48 | 524,144 | 383,868 | 462,421 |
| MOTHERS | 210 | 1,419,183 | 587,168 | 2,668,533 |
| PRO | 2 | 463,777 | 463,777 | 317,383 |

Note: The table summarizes the amounts of investment for each start-up companies going to IPO markets. All the numbers are calculated by the data we use for our empirical analysis.

Table 6 Post-IPO dynamics

| Post-IPO dyanmics | #(start-up companies) | #(investors) average | #(investors) median | #(investors) standard deviation |
|---|---|---|---|---|
| DELIST_a | 71 | 12.0 | 11 | 8.1 |
| DELIST_b | 34 | 13.9 | 11 | 8.6 |
| STAY | 312 | 11.2 | 9 | 8.9 |
| TSE1 | 85 | 12.8 | 10 | 10.4 |
| TSE2 | 33 | 10.7 | 11 | 5.3 |

Note: The table summarizes the post-IPO dynamics for each start-up companies going to IPO markets. All the numbers are calculated by the data we use for our empirical analysis.

Table 7a Detected communities (top 10)

| Top 10 community (based on degree) | Group ID for start-up companies | Group ID for investors | #(start-up companies) | #(investors) |
|---|---|---|---|---|
| 1 | 14 | 21 | 60 | 204 |
| 2 | 2 | 31 | 29 | 291 |
| 3 | 17 | 29 | 23 | 241 |
| 4 | 13 | 20 | 48 | 260 |
| 5 | 3 | 34 | 35 | 284 |
| 6 | 11 | 28 | 30 | 139 |
| 7 | 6 | 27 | 26 | 126 |
| 8 | 12 | 32 | 41 | 108 |
| 9 | 1 | 24 | 29 | 137 |
| 10 | 10 | 23 | 23 | 138 |

Note: The table summarizes the detected communities. We list top 10 communities in terms of the number of investment (i.e., degree).

Table 7b Detected communities (top 10)

| Top 10 community (based on degree) | Item | Group ID for start-up companies | Group ID for investors | Median |
|---|---|---|---|---|
| 1 | Initial price/Investment price - 1 | 14 | 21 | 0.83 |
| | Offer price/Investment price -1 | 14 | 21 | 0.38 |
| | Initial price decline rate (%) | 14 | 21 | 105.00 |
| 2 | Initial price/Investment price - 1 | 2 | 31 | 0.22 |
| | Offer price/Investment price -1 | 2 | 31 | -0.08 |
| | Initial price decline rate (%) | 2 | 31 | 84.61 |
| 3 | Initial price/Investment price - 1 | 17 | 29 | 0.06 |
| | Offer price/Investment price -1 | 17 | 29 | -0.06 |
| | Initial price decline rate (%) | 17 | 29 | 75.84 |
| 4 | Initial price/Investment price - 1 | 13 | 20 | 0.06 |
| | Offer price/Investment price -1 | 13 | 20 | -0.06 |
| | Initial price decline rate (%) | 13 | 20 | 30.00 |
| 5 | Initial price/Investment price - 1 | 3 | 34 | 0.78 |
| | Offer price/Investment price -1 | 3 | 34 | 0.23 |
| | Initial price decline rate (%) | 3 | 34 | 90.47 |
| 6 | Initial price/Investment price - 1 | 11 | 28 | -0.08 |
| | Offer price/Investment price -1 | 11 | 28 | -0.21 |
| | Initial price decline rate (%) | 11 | 28 | 56.92 |
| 7 | Initial price/Investment price - 1 | 6 | 27 | 0.39 |
| | Offer price/Investment price -1 | 6 | 27 | 0.01 |
| | Initial price decline rate (%) | 6 | 27 | 100.00 |
| 8 | Initial price/Investment price - 1 | 12 | 32 | -0.10 |
| | Offer price/Investment price -1 | 12 | 32 | -0.13 |
| | Initial price decline rate (%) | 12 | 32 | 40.00 |
| 9 | Initial price/Investment price - 1 | 1 | 24 | 0.12 |
| | Offer price/Investment price -1 | 1 | 24 | -0.14 |
| | Initial price decline rate (%) | 1 | 24 | 123.68 |
| 10 | Initial price/Investment price - 1 | 10 | 23 | -0.31 |
| | Offer price/Investment price -1 | 10 | 23 | -0.37 |
| | Initial price decline rate (%) | 10 | 23 | 51.81 |

Note: The table summarizes the detected communities. We list top 10 communities in terms of the number of investment (i.e., degree).

Table 8 Test for start-up companies' industry

| | Communication | Finanial | Financial-ICT | Business service | IT hardware | IT software | IT service | Semiconductor & products |
|---|---|---|---|---|---|---|---|---|
| 1 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 2 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 3 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 4 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 5 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 6 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 7 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 8 | FALSE | smaller | FALSE | FALSE | FALSE | FALSE | FALSE | larger |
| 9 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 10 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |

| | Biotechnology | Medical | Medical-ICT | Energy | Energy-ICT | Environment | Consumer |
|---|---|---|---|---|---|---|---|
| 1 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 2 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 3 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 4 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 5 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 6 | larger | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 7 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 8 | FALSE | FALSE | FALSE | larger | FALSE | FALSE | FALSE |
| 9 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 10 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |

Note: The table summarizes the results of the test for start-up companies industry.

Table 9 Test for investor type

| | 1 | 3 | 4 | 6 | 7 | 8 | 9 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| | Angel | Individual | Founder | VC | Shosha | Foreign | Bank | Non-financial | ESOP |
| 1 | FALSE | FALSE | smaller | larger | FALSE | FALSE | FALSE | FALSE | FALSE |
| 2 | FALSE | FALSE | FALSE | smaller | FALSE | FALSE | FALSE | larger | FALSE |
| 3 | FALSE | FALSE | FALSE | smaller | FALSE | FALSE | FALSE | FALSE | FALSE |
| 4 | FALSE | FALSE | FALSE | smaller | FALSE | FALSE | larger | FALSE | FALSE |
| 5 | FALSE | FALSE | smaller | FALSE | FALSE | FALSE | FALSE | larger | smaller |
| 6 | FALSE | FALSE | FALSE | larger | FALSE | FALSE | FALSE | smaller | FALSE |
| 7 | FALSE | FALSE | FALSE | larger | larger | FALSE | larger | smaller | FALSE |
| 8 | larger | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | smaller | larger |
| 9 | FALSE | FALSE | larger | FALSE | FALSE | FALSE | FALSE | FALSE | larger |
| 10 | FALSE | FALSE | larger | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |

| | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 21 |
|---|---|---|---|---|---|---|---|---|
| | Security companies | Government | Others | Other financial | University | Real estate | Insurance | Foreign VC |
| 1 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 2 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 3 | FALSE | FALSE | FALSE | FALSE | FALSE | larger | FALSE | FALSE |
| 4 | FALSE | FALSE | larger | FALSE | FALSE | FALSE | FALSE | FALSE |
| 5 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 6 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 7 | FALSE | FALSE | FALSE | FALSE | FALSE | larger | larger | FALSE |
| 8 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 9 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 10 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |

Note: The table summarizes the results of the test for investor type.

Table 10 Investment performance

| | 75_fp | 90_fp | 75_op | 90_op | TSE1 | TSE2 | Delist |
|---|---|---|---|---|---|---|---|
| 1 | larger | larger | larger | larger | FALSE | FALSE | FALSE |
| 2 | larger | larger | larger | larger | FALSE | FALSE | FALSE |
| 3 | smaller | smaller | smaller | smaller | FALSE | FALSE | FALSE |
| 4 | FALSE | FALSE | FALSE | FALSE | larger | FALSE | FALSE |
| 5 | larger | larger | larger | larger | FALSE | FALSE | larger |
| 6 | smaller | smaller | smaller | smaller | FALSE | FALSE | FALSE |
| 7 | FALSE | FALSE | FALSE | smaller | FALSE | FALSE | FALSE |
| 8 | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE | FALSE |
| 9 | FALSE | smaller | FALSE | FALSE | FALSE | FALSE | FALSE |
| 10 | smaller | FALSE | smaller | FALSE | FALSE | FALSE | FALSE |

Note: The table summarizes the results of the test for post-IPO dynamics (i.e., DELIST_b).