Conglomerate Formation in China

Chong-En Bai, Tsinghua University
Chang-Tai Hsieh, Chicago Booth
Zheng (Michael) Song, Chinese University of Hong Kong
Xin Wang, Chinese University of Hong Kong
China’s Growth Puzzle

• Why does China have a thriving private sector?
  • “Regionally Decentralized Authoritarianism” (Xu, Journal of Economic Literature, 2011)
  • “Crony Capitalism with Chinese Characteristics” (Bai, Hsieh and Song, prepared for NBER Macro Annual 2019)

• This paper: Conglomerate as a special institutional arrangement
  • Overcome economic and institutional frictions
  • Strike special deals with many private businesses
Summary of the Main Findings

• Facts from firm ownership networks in China’s economic universe with 17 million firms
  • Fast-growing “giant component”
  • Mainly driven by expansion of incumbent conglomerates
  • Firm size and TFP distributions across and within conglomerate

• A model of network formation that overcomes financial constraints (or entry barriers)
  • Giant component and conglomerates
  • Firm size and TFP distributions
  • Structural estimation and aggregate implications (very preliminary)
China’s Firm Ownership Network Data

• Firm Registration Data of State Administration for Industry and Commerce

• All firms, including holding companies (28 million firms, 11 million exit)

• Owner can be individual or legal persons (firms or holding companies)
China’s Firm Ownership Network Data (cont’d)

• Owners in 2015 (or exit date)
  • Name and ID of legal person (including holding companies) and individuals
  • Equity share in 2015 (or exit date) of each owner
  • Change in legal person owners from 2004 to 2014 for 11 provinces (no equity information, robustness check)

• Registered capital, year of establishment, exit year

• Matched with NBS data on industrial firms
Economic Activity Dominated by Two Largest Connected Networks (Giant Components)

• Connected by Legal Person Ownership
  • 4% by firms and 48% by registered capital
  • 0.75% reduction by removing listed firms

• Connected by Individual Person Ownership
  • 11% by firms and 15% by registered capital
  • 0.35% reduction by removing listed firms

• Firms connecting the two networks (i.e., firms in both the two networks)
  • 0.8% by firms and 4.7% by registered capital
Galaxy 1
(700K Firms, 48% of Total Registered Capital)

2015

Galaxy 2
(2M Firms, 15% of Total Registered Capital)
Historical Ownership Networks

• Approach 1: Inferred from 2015 ownership data + Firm entry and exit year (our main approach)
  • Strength: Full coverage
  • Weakness: Strong assumption of initial firm registration information (annual updating rate of 5% for firm ownership information by Annual Survey of Industrial Firms)

• Approach 2: Using records on changes in firm registration information in 11 provinces since 2004 (robustness check)
  • Strength: Real data
  • Weakness: 11 provinces only
1995

Galaxy 1
(50K Firms, 31% of Total Registered Capital)

No Significant Galaxy 2
2005

Galaxy 1
(240K Firms, 50% of Total Registered Capital)

Galaxy 2
(100K Firms, 2% of Total Registered Capital)
Galaxy 1
(700K Firms, 48% of Total Registered Capital)

2015

Galaxy 2
(2M Firms, 15% of Total Registered Capital)
Size of Galaxy 1 and 2 Relative to the Universe
Size of Galaxy 1 Relative to Firms with Legal-Person Connections

Robustness Check: Historical Data with 11 Provinces
Galaxies in Germany

- 2.9 million German firms in Orbis
  - 0.79 million with legal-person connections (27% > 9% in China)

- 0.26 million in Galaxy 1
  - 0.26/0.79 = 33% < 44% in China

- 87 thousand in Galaxy 2
  - 0.087/0.57 = 15% < 27% in China
Some Stylized Facts

• Fast-growing giant components (G1 and G2)

• Correlations with firm size, age, ownership, YK and its distance to the center of the galaxy

• Center of G1 (by closeness): Big SOEs
  • Germany: Financial institutions (DEUTSCHE BANK AG; COMMERZBANK AG; UNICREDIT BANK AG; DZ BANK AG; DEUTSCHE ZENTRAL-GENOSSENSCHAFTSBANK ...
772 Central and Provincial SOEs as the Center of Galaxy 1

Correlation between closeness and distance: -0.86
Closeness\(i\) = 1 / Sum(distance\(i,j\), \(j\neq i\)), standardized into \([0,1]\).
Distance: distance to the set of core firms (central and provincial SOEs).
### State-Centered ≠ State-Controlled

<table>
<thead>
<tr>
<th>Firms Controlled by 772 SOEs</th>
<th>Direct + Indirect Equity Shares</th>
<th>Controlling Shareholder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Firm Number</td>
<td>RC share in Galaxy 1</td>
</tr>
<tr>
<td>772 SOEs</td>
<td>Threshold</td>
<td>772</td>
</tr>
<tr>
<td>50%</td>
<td>40,234</td>
<td>26.0%</td>
</tr>
<tr>
<td>25%</td>
<td>52,446</td>
<td>29.9%</td>
</tr>
<tr>
<td>10%</td>
<td>63,678</td>
<td>32.8%</td>
</tr>
<tr>
<td>5%</td>
<td>69,385</td>
<td>33.9%</td>
</tr>
</tbody>
</table>
Connectedness Firm Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Legal-Person Firms</th>
<th></th>
<th>Natural-Person Firms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log RC</td>
<td>Age</td>
<td>SOE</td>
<td>log Y</td>
</tr>
<tr>
<td>Galaxy 1</td>
<td>2.295</td>
<td>3.359</td>
<td>0.204</td>
<td>0.616</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.006)</td>
<td>(0.001)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Firms in Galaxy 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log RC</td>
<td>Age</td>
<td>SOE</td>
<td>log Y</td>
</tr>
<tr>
<td>Closeness</td>
<td>5.048</td>
<td>4.910</td>
<td>1.163</td>
<td>2.136</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.054)</td>
<td>(0.007)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.230</td>
<td>-0.421</td>
<td>-0.083</td>
<td>-0.100</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.004)</td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

Closeness\((i) = 1 / \text{Sum(distance}(i,j), j\neq i),\) standardized into \([0,1]\) (0, periphery; 1 center)
## Connectedness and Firm Characteristics

<table>
<thead>
<tr>
<th></th>
<th>G1 = 1</th>
<th>Closeness in G1</th>
<th>Distance in G1</th>
<th>G2 = 1</th>
<th>Closeness in G2</th>
</tr>
</thead>
<tbody>
<tr>
<td>log TFP</td>
<td>0.117</td>
<td>0.029</td>
<td>-0.367</td>
<td>0.034</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.031)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>log YK</td>
<td>-0.075</td>
<td>-0.018</td>
<td>0.232</td>
<td>-0.020</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.013)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>log YL</td>
<td>-0.044</td>
<td>-0.009</td>
<td>0.108</td>
<td>-0.023</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.024)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Age</td>
<td>0.005</td>
<td>0.001</td>
<td>-0.013</td>
<td>-0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>SOE</td>
<td>0.581</td>
<td>0.089</td>
<td>-1.874</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>301,246</td>
<td>55,418</td>
<td>55,418</td>
<td>301,246</td>
<td>51,590</td>
</tr>
</tbody>
</table>

Using eccentricity and historical network data: Very similar results
Identifying “Conglomerates” in the Firm Network

• Identify communities (conglomerates) in the first galaxy (700k firms)

• Method: Multilevel algorithm (“Fast Unfolding of Communities in Large Networks”, Blondel et al., 2008)

• Results: 413 communities (or conglomerates) identified, with modularity = 0.94 (from -1 to 1, a measure of performance in network partition)

• Other algorithms: informap (m=0.86), walktrap (m=0.83) and label propagation (m=0.82)
Modularity Maximization

• Basic idea: Firms are more likely to be connected in the conglomerate (higher edge densities within conglomerate than between conglomerate).

• The modularity quantifies the deviation of within-conglomerate edge density from the expected density in a random network:

\[
Q = \frac{1}{2M} \sum_{ij} \left( A_{ij} - \frac{k_i k_j}{2M} \right) \delta(C_i, C_j)
\]

• It can also be rewritten as

\[
Q = \sum_c \left[ \frac{k_{in}^c}{2M} - \left( \frac{k_{tot}^c}{2M} \right)^2 \right]
\]
Basic Results

Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm Number</td>
<td>1702.8</td>
<td>1187.0</td>
<td>1544.0</td>
<td>31.0</td>
<td>11014.0</td>
</tr>
<tr>
<td>RC (0.1 Billion Yuan)</td>
<td>1544.4</td>
<td>686.8</td>
<td>2491.1</td>
<td>0.9</td>
<td>23098.5</td>
</tr>
</tbody>
</table>

Conglomerate Size Distribution
Growth of Conglomerates

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Conglomerate</th>
<th>Mean of Conglomerate Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>212</td>
<td>245.4</td>
</tr>
<tr>
<td>2000</td>
<td>308</td>
<td>512.5</td>
</tr>
<tr>
<td>2005</td>
<td>334</td>
<td>731.5</td>
</tr>
<tr>
<td>2010</td>
<td>372</td>
<td>1088.1</td>
</tr>
<tr>
<td>2015</td>
<td>413</td>
<td>1702.8</td>
</tr>
</tbody>
</table>
## Growth of Conglomerates – 11 Provinces

<table>
<thead>
<tr>
<th>Year</th>
<th>Using Historical True Shareholders (Method 2)</th>
<th>Using Most Recent Shareholders (Method 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Conglomerate</td>
<td>Mean of Conglomerate Size</td>
</tr>
<tr>
<td>2004</td>
<td>216</td>
<td>396.1</td>
</tr>
<tr>
<td>2007</td>
<td>231</td>
<td>438.7</td>
</tr>
<tr>
<td>2011</td>
<td>273</td>
<td>565.2</td>
</tr>
<tr>
<td>2013</td>
<td>292</td>
<td>686.8</td>
</tr>
</tbody>
</table>
Conglomerate Stability

- Conglomerate ID: Conglomerate $j$ in period $t$ is identical to conglomerate $i$ in period $t - 1$ if $(j, t)$ has the maximum number of incumbent firms in $(i, t - 1)$

Proportion of Incumbent Firms Staying in the Same Conglomerate

<table>
<thead>
<tr>
<th>Period</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>95-00</td>
<td>44.97%</td>
</tr>
<tr>
<td>00-05</td>
<td>40.60%</td>
</tr>
<tr>
<td>05-10</td>
<td>42.20%</td>
</tr>
<tr>
<td>10-15</td>
<td>42.40%</td>
</tr>
</tbody>
</table>

Annual Rate of Incumbent Firms Exiting Conglomerate
State-Centered Conglomerates

- 772 central and provincial SOEs in 210 conglomerates

- These SOEs are in the center of their conglomerate (with much higher closeness, larger size and lower YK).

- Closeness within conglomerate is highly correlated with distance to the 772 SOEs (correlation: -0.72).

- The 210 “state-centered” conglomerates account for two-thirds of Galaxy 1 by firm number and registered capital.
Conglomerate Size vs. Average/Top Firms

Average Log RC in a Conglomerate

The 99th Percentile Log RC in a Conglomerate

Robustness check: Using Log RC of the core firms in the center of conglomerate
Conglomerate Size vs. Average/Top Firms

Robustness check: Using Log YK of the core firms in the center of conglomerate
## Cross-Conglomerate Variations

<table>
<thead>
<tr>
<th></th>
<th>Log Ave. RC</th>
<th>Log 90th RC</th>
<th>Log 95th RC</th>
<th>Log 99th RC</th>
<th>Log Ave. TFP</th>
<th>Log 90th TFP</th>
<th>Log 95th TFP</th>
<th>Log 99th TFP</th>
<th>Log Ave. YK</th>
<th>Log 10th YK</th>
<th>Log 5th YK</th>
<th>Log 1st YK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log C Size</td>
<td>0.158 (0.025)</td>
<td>0.296 (0.032)</td>
<td>0.332 (0.033)</td>
<td>0.414 (0.038)</td>
<td>0.082 (0.018)</td>
<td>0.202 (0.030)</td>
<td>0.256 (0.038)</td>
<td>0.479 (0.057)</td>
<td>0.022 (0.021)</td>
<td>-0.083 (0.026)</td>
<td>-0.114 (0.031)</td>
<td>-0.321 (0.049)</td>
</tr>
<tr>
<td>N</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Log Ave. RC</th>
<th>Log 90th RC</th>
<th>Log 95th RC</th>
<th>Log 99th RC</th>
<th>Log Ave. TFP</th>
<th>Log 90th TFP</th>
<th>Log 95th TFP</th>
<th>Log 99th TFP</th>
<th>Log Ave. YK</th>
<th>Log 10th YK</th>
<th>Log 5th YK</th>
<th>Log 1st YK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ave. Distance</td>
<td>-0.209 (0.020)</td>
<td>-0.311 (0.027)</td>
<td>-0.363 (0.027)</td>
<td>-0.467 (0.030)</td>
<td>-0.076 (0.015)</td>
<td>-0.146 (0.025)</td>
<td>-0.197 (0.032)</td>
<td>-0.327 (0.050)</td>
<td>0.007 (0.018)</td>
<td>0.113 (0.021)</td>
<td>0.138 (0.025)</td>
<td>0.239 (0.042)</td>
</tr>
<tr>
<td>N</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>413</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
<td>406</td>
</tr>
</tbody>
</table>
## Within-Conglomerate Variations

<table>
<thead>
<tr>
<th></th>
<th>Log RC</th>
<th>Log RC</th>
<th>Log RC</th>
<th>Log TFP</th>
<th>Log TFP</th>
<th>Log TFP</th>
<th>Log YK</th>
<th>Log YK</th>
<th>Log YK</th>
<th>Log YK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distance</strong></td>
<td>-0.232</td>
<td></td>
<td></td>
<td>-0.030</td>
<td></td>
<td></td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C. Distance</strong></td>
<td></td>
<td></td>
<td></td>
<td>-0.077</td>
<td>-0.060</td>
<td>-0.088</td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.006</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td><strong>C. Distance * C. State Dummy</strong></td>
<td></td>
<td>-0.051</td>
<td>-0.088</td>
<td></td>
<td>-0.004</td>
<td></td>
<td>0.002</td>
<td></td>
<td></td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td>(0.002)</td>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td><strong>C. Distance * Log C. Size</strong></td>
<td></td>
<td>-0.069</td>
<td></td>
<td>-0.14</td>
<td></td>
<td></td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C. FE</strong></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>R2</strong></td>
<td>0.08</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>703,128</td>
<td>471,849</td>
<td>471,849</td>
<td>471,849</td>
<td>55,443</td>
<td>37,123</td>
<td>37,123</td>
<td>37,123</td>
<td>56,040</td>
<td>37,548</td>
</tr>
</tbody>
</table>
Summary of the Stylized Facts

• Universe: Fast-growing Galaxy 1 and 2

• Galaxy 1: Growth mainly driven by the expansion of incumbent conglomerates.
  • Firm size, age and SOE share (YK) negatively (positively) correlated with distance to the center of the galaxy

• Conglomerates:
  • Between: Strong correlation between conglomerate size and the top firm size or the bottom YK ratio in a conglomerate
  • Within: Firm size, age and SOE share (YK) negatively (positively) correlated with distance to the center of a conglomerate
A Model of Conglomerate Formation

• Consider an economy without financial intermediary.

• All firms produce the same goods by a decreasing-return-to-scale production technology: \( Y_i = Z_i K_i^\alpha \), where firm TFP, \( Z_i \), is constant and exogenous.

• Assume all firms run by two-period-lived entrepreneurs with log preferences.
  • Constant saving rate: \( \beta / (1 + \beta) \).
  • Robustness check: Infinitively-lived entrepreneurs
Firm Dynamics

- $N^C(t)$ incumbents (capital owner) and $N^e(t)$ potential entrants (no capital).

- An incumbent may invest in a potential entrant or another incumbent via random search.

- The invested firms will pay a fixed proportion of surplus to their investors (frictions for misallocation between investor and investee).

- The invested entrants will become incumbents (capital owners) in the subsequent periods.

- Incumbents exit at the rate of $q$, while $N^e(t)$ is constant.
Firm Characteristics and Distributions

• Incumbents: \((A, Z^c)\), where \(A \in \{A_1, \ldots, A_f\}\) is capital owned by incumbents at the beginning of each period and \(Z^c \in \{Z_1, \ldots, Z_I\}\) is the incumbent’s TFP.

• Potential entrants: \(Z^e\), where \(Z^e \in \{Z_1, \ldots, Z_J\}\) is the entrant’s TFP.

• The initial distributions are \(P^c(A, Z^c, 0)\), \(P^e(Z^e, 0)\), \(N^c(t)\) and \(N^e(t)\).
Capital Allocation in a Joint Venture

- Capital allocation decided by the investor

- For a joint venture between an incumbent and an entrant: \( \max_K Z^c (A - K)^\alpha + \gamma K^\alpha \)

  \[
  \Rightarrow K^{c,e} (A, Z^c, Z^e) = \frac{(\gamma Z^e)^{\frac{1}{1-\alpha}}}{(\gamma Z^e)^{\frac{1}{1-\alpha}} + (Z^c)^{\frac{1}{1-\alpha}}} A
  \]

- For a joint venture between two incumbents: \( \max_K Z^c (A - K)^\alpha + \gamma \hat{Z}^c \left( (\hat{A} + K)^\alpha - \hat{A}^\alpha \right) \)

  \[
  \Rightarrow K^{c,c} (A, Z^c, \hat{A}, \hat{Z}^c) = \max \left\{ 0, \frac{(\gamma \hat{Z}^c)^{\frac{1}{1-\alpha}}}{(\gamma \hat{Z}^c)^{\frac{1}{1-\alpha}} + (Z^c)^{\frac{1}{1-\alpha}}} A - \frac{(Z^c)^{\frac{1}{1-\alpha}}}{(\gamma \hat{Z}^c)^{\frac{1}{1-\alpha}} + (Z^c)^{\frac{1}{1-\alpha}}} \hat{A} \right\}
  \]

- Capital misallocation between investors and investees
Decision of Forming a Joint Venture

• Investors pay the cost of forming joint venture, $C$.

• Conditional on the meeting of $(A, Z^c, Z^e)$, the joint venture will be formed if $I^{c,e}(A, Z^c, Z^e) = 1$ or

\[ Z^c (A - K^{c,e}(A, Z^c, Z^e))^\alpha + \gamma Z^e K^{c,e}(A, Z^c, Z^e)^\alpha - Z^c A^\alpha > C \]

• Conditional on the meeting of $(A, Z^c, \hat{A}, \hat{Z}^c)$, the joint venture will be formed if $I^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) = 1$ or

\[ Z^c (A - K^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c))^\alpha + \gamma \hat{Z}^c \left( (\hat{A} + K^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c))^\alpha - \hat{A}^\alpha \right) - Z^c A^\alpha > C \]
Investor’s Surplus

• For investor of a joint venture \((A, Z^c, Z^e)\):

\[
\pi^{c,e}(A, Z^c, Z^e) = Z^c \left( A - K^{c,e}(A, Z^c, Z^e) \right) + \gamma Z^e K^{c,e}(A, Z^c, Z^e) - Z^c A^\alpha - C
\]

• For investor of a joint venture \((A, Z^c, \hat{A}, \hat{Z}^c)\):

\[
\pi^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) = Z^c \left( A - K^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) \right) + \gamma \hat{Z}^c \left( \hat{A} + K^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) \right) - \hat{A}^\alpha - Z^c A^\alpha - C
\]
Investee’s Surplus

• For investee of a joint venture \((A, Z^c, Z^e)\):

\[
\hat{\pi}^{c,e}(A, Z^c, Z^e) = (1 - \gamma)Z^eK^{c,e}(A, Z^c, Z^e)^\alpha
\]

• For investee of a joint venture \((A, Z^c, \hat{A}, \hat{Z}^c)\):

\[
\hat{\pi}^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) = (1 - \gamma)\hat{Z}^c\left(\hat{A} + K^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c)\right)^\alpha + \gamma\hat{Z}^c\hat{A}^\alpha
\]
Random Search

• Incumbent’s search effort to find an investee: $\lambda^c (A, Z^c, t)$

• Incumbent and entrant’s search effort to find an investor: $\lambda^c (A, Z^c, t)$ and $\lambda^e (Z^e, t)$.

• Search cost is $SC(x)$, $x = \lambda^c, \hat{\lambda}^c$ or $\lambda^e$.
  
  • $SC(\cdot)$ is homogeneous of degree 1, convex and satisfies $SC(0) = 0, SC'(\cdot) > 0, SC''(\cdot) > 0, \lim_{x \to \infty} SC(x) = \infty$.
  
  • For simulation and estimation: $SC(x) = \frac{B}{\eta} (x)^\eta$, where $\eta > 1$. 
• Two sides of the market: Investors (incumbents) and investees (incumbents and potential entrants)

• The share of entrants on the side of investees:

\[
ES(t) = \frac{\sum_{Z^e} \mu^e(Z^e, t)P^e(Z^e, t)}{\sum_{Z^e} \mu^e(Z^e, t)P^e(Z^e, t) + \sum_{A,Z^c} \mu^c(A,Z^c,t)P^c(A,Z^c,t)}
\]

• Market tightness:

\[
\theta(t) = \min \left\{ 1, \frac{\sum_{Z^e} \mu^e(Z^e, t)P^e(Z^e, t) + \sum_{A,Z^c} \mu^c(A,Z^c,t)P^c(A,Z^c,t)}{\sum_{A,Z^c} \lambda(A,Z^c,t)P^c(A,Z^c,t)} \right\}
\]

\[
\hat{\theta}(t) = \min \left\{ 1, \frac{\sum_{A,Z^c} \lambda(A,Z^c,t)P^c(A,Z^c,t)}{\sum_{Z^e} \mu^e(Z^e, t)P^e(Z^e, t) + \sum_{A,Z^c} \mu^c(A,Z^c,t)P^c(A,Z^c,t)} \right\}
\]
• The “natural” probability for an incumbent \((A, Z^c)\) to meet a potential entrant \(Z^e\):

\[
Prob^{c,e}(A, Z^c, Z^e, t) = \theta(t) \cdot \frac{\mu^e(Z^e, t) P^e(Z^e, t)}{\sum_{Z^e} \mu^e(Z^e, t) P^e(Z^e, t)}
\]

• The probability for an incumbent \((A, Z^c)\) to meet an incumbent \((\hat{A}, \hat{Z}^c)\):

\[
Prob^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c, t) = \theta(t) \cdot \frac{\mu^c(\hat{A}, \hat{Z}^c, t) P^c(\hat{A}, \hat{Z}^c, t)}{\sum_{\hat{A}, \hat{Z}^c} \mu^c(\hat{A}, \hat{Z}^c, t) P^c(\hat{A}, \hat{Z}^c, t)}
\]

• The probability for an entrant \(Z^e\) to meet an incumbent \((A, Z^c)\):

\[
\overline{Prob}^{c,e}(A, Z^c, Z^e, t) = \hat{\theta}(t) \cdot \frac{\lambda(A, Z^c, t) P^c(A, Z^c, t)}{\sum_{A, Z^c} \lambda(A, Z^c, t) P^c(A, Z^c, t)}
\]

• The probability for an incumbent investee \((\hat{A}, \hat{Z}^c)\) to meet an incumbent investor \((A, Z^c)\):

\[
\overline{Prob}^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c, t) = \hat{\theta}(t) \cdot \frac{\lambda(A, Z^c, t) P^c(A, Z^c, t)}{\sum_{A, Z^c} \lambda(A, Z^c, t) P^c(A, Z^c, t)}
\]
Optimal Searching Efforts

• As an investor:

$$\lambda^c(A, Z^c, t) = \operatorname{argmax}_{\lambda^c} \lambda^c \cdot \left( \sum_{Z^e} ES(t) \cdot \operatorname{Prob}^{c,e}(A, Z^c, Z^e, t) \cdot \pi^{c,e}(A, Z^c, Z^e) \cdot I^{c,e}(A, Z^c, Z^e) + \sum_{\hat{A}, \hat{Z}^c} (1 - ES(t)) \cdot \operatorname{Prob}^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c, t) \cdot \pi^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) \cdot I^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) \right) - SC(\lambda^c)$$

• As an entrant investee:

$$\hat{\lambda}^c(Z^e, t) = \operatorname{argmax}_{\hat{\lambda}^c} \hat{\lambda}^c \cdot \sum_{A, Z^c} \operatorname{Prob}^{c,e}(A, Z^c, Z^e, t) \cdot \hat{\pi}^{c,e}(A, Z^c, Z^e) \cdot I^{c,e}(A, Z^c, Z^e) - SC(\hat{\lambda}^c)$$

• As an incumbent investee:

$$\lambda^e(A, Z^c, t) = \operatorname{argmax}_{\lambda^e} \lambda^e \cdot \sum_{\hat{A}, \hat{Z}^c} \operatorname{Prob}^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c, t) \cdot \hat{\pi}^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) \cdot I^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) - SC(\lambda^e)$$
Evolution of Capital

\[
A'(A, Z^c, Z^e) = \frac{\beta}{1 + \beta} (\pi^{c,e}(A, Z^c, Z^e) + Z^c A^\alpha + (1 - \delta)A),
\]

\[
A'(A, Z^c, \hat{A}, \hat{Z}^c) = \frac{\beta}{1 + \beta} (\pi^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) + Z^c A^\alpha + (1 - \delta)A),
\]

\[
\hat{A}'(A, Z^c, Z^e) = \frac{\beta}{1 + \beta} \hat{\pi}^{c,e}(A, Z^c, Z^e),
\]

\[
\hat{A}'(A, Z^c, \hat{A}, \hat{Z}^c) = \frac{\beta}{1 + \beta} (\hat{\pi}^{c,c}(A, Z^c, \hat{A}, \hat{Z}^c) + (1 - \delta)\hat{A}).
\]
Outcomes

• Dynamic firm TFP and size distributions

• Dynamic ownership network

• Many connected networks and the giant component

• Conglomerate Formation in the giant component
Simulations

• $N^c(0) = 50$, $q = 2\%$ and $N^e(t) = 1000$ for all $t$ (i.e., $g = 0$)

• Special case: Incumbents can be investors only

• Benchmark: Incumbents can be both investors and investees

• Connectedness by new incumbents is the key to generate a big giant component.
Special Case vs Benchmark
Structural Estimation (VERY PRELIMINARY)

• Between-Group Moments: Cross-conglomerate correlations with conglomerate size and top/bottom firm size, TFP and YK in the conglomerate

• Within-Group Moments: (i) Within-conglomerate variance of firm size, TFP and YK relative to the variance in the giant component; (ii) Within-conglomerate correlation between firm distance to the center and firm size, TFP and YK
• Predetermined parameters: $\alpha = 0.5$ and saving rate $= 0.3$

• Estimated parameters: $\gamma = 0.2$, $\sigma_{logA} = 1$, $\sigma_{logZ} = 0.2$, $\eta = 3$ and $C = 0.04$

<table>
<thead>
<tr>
<th>Cross-Conglomerate Moments</th>
<th>Data</th>
<th>Model</th>
<th>Within-Conglomerate Moments</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross_C_corr(C_Size,log_RC_99th)</td>
<td>0.39</td>
<td>0.37</td>
<td>Within_C_corr(dist,log_RC)</td>
<td>-0.70</td>
<td>-0.23</td>
</tr>
<tr>
<td>Cross_C_corr(C_Size,log_tfp_99th)</td>
<td>0.38</td>
<td>0.11</td>
<td>Within_C_corr(dist,log_TFP)</td>
<td>-0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Cross_C_corr(C_Size,log_yk_1st)</td>
<td>-0.31</td>
<td>-0.56</td>
<td>Within_C_corr(dist,log_YK)</td>
<td>0.20</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within_C_var(log_RC) / var(log_RC)</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within_C_var(log_TFP) / var(log_TFP)</td>
<td>0.98</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within_C_var(log_YK) / var(log_YK)</td>
<td>0.98</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Simulated Conglomerate Size and Average/Top Log Firm Size

![Graphs showing simulated conglomerate size and average/top log firm size](image)

- \( \ln\text{groupsize} \)
- \( m_{\text{avg_logoutput}} \)
- Fitted values
- \( m_{\text{upper95_logoutput}} \)
- Fitted values
Simulated Conglomerate Size and Average/Bottom Log Firm YK
Non-Targeted Moment: Size of Galaxy 1

- The simulated capital share of the giant component: 61%

- Total registered capital in Galaxy 1 / Total registered capital in the universe: 48%

- Total registered capital in Galaxy 1 / Total registered capital in the firms with legal-person connections: 80%
Non-Targeted Moment: Conglomerate Size Distribution

Conglomerate Size Distribution (Data)

Log (1-F) vs. Log Conglomerate Size

Conglomerate Size Distribution (Simulated)

Log Conglomerate Size vs. lnF
Aggregate Results
The Explosion of the Giant Component
Main Insights from the Model

• Consider an economy with no financial intermediary and high entry barriers

• Conglomeration helps entry and resource reallocation

• Limitations of conglomerations as informal institution
  • Misallocation, Inefficient Selection, Systemic Risks ...
Interpretations

- Network as financial intermediary?
- Network as informal institutions?
- Limitations of conglomeration
Regional and Sectoral Concentrations
All Companies Owned by East Hope

- Red: State JV
- Grey: Solely Owned by East Hope
- Blue: Private JV
Conclusion

• Conglomeration as an informal institution that helps economic growth by reallocating resources and removing entry barriers

• The China model has its strengths and weaknesses. More formal institutions would need to be established to sustain long-run growth.