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
Understanding China’s Growth (and Recent Slowdown)

• Why does China have a thriving private sector?
  • “The Fundamental Institutions of China's Reforms and Development” (Xu, Journal of Economic Literature, 2011)
  • “Institutional Foundations of China’s Growth and Slowdown” (Bai, Hsieh and Song, prepared for NBER Macro Annual 2019)

• Lack of evidence on the (in)formal institutions

• This paper: Conglomerate as a special institutional arrangement
  • Overcome (institutional) frictions
  • Strike special deals with many private businesses
Road Map

• Data and basic facts

• Identifying “conglomerates” and their main patterns

• A model of network formation, its empirical predictions and aggregate implications

• Future work
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

• 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)

Galaxy 2
(2M Firms, 15% of Total Registered Capital)
1995

Galaxy 1
(50K Firms, 1.6% of All firms and 31% of Total Registered Capital)

No Significant Galaxy 2
2000

Galaxy 1
(160K Firms, 3.8% of All firms, 46% of Total Registered Capital)

No Significant Galaxy 2
2005

Galaxy 1
(240K Firms, 4% of All Firms, 50% of Total Registered Capital)

Galaxy 2
(100K Firms, 2% of Total Registered Capital)
2010

Galaxy 1
(400K Firms, 4.3% of All Firms, 52% of Total Registered Capital)

Galaxy 2
(500K Firms, 7% of Total Registered Capital)
Galaxy 1
(700K Firms, 48% of Total Registered Capital)

Galaxy 2
(2M Firms, 15% of Total Registered Capital)
Relative Size of Galaxy 1

Dynamics of Galaxy Relative Size

Registered Capital  
Firm Number

Relative Size of Galaxy 2
Galaxies in Germany

• 2.9 million German firms in Orbis
  • 0.52 million with legal-person shareholders (20% > 6% in China)

• 0.26 million in Galaxy 1
  • $0.26/0.52 = 50.5% < 64.8\%$ in China

• 87 thousand in Galaxy 2
  • $0.087/2.9 = 3\% << 11\%$ in China
Some Stylized Facts

• Extensive and intensive margins: Firm size, age, ownership, YK ...

• Firms connecting the two galaxies: Peripheral in G1 and center in G2

• Center in 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≠i), standardized into [0,1].
Distance: distance to the set of core firms (central and provincial SOEs).
State-Centered ≠ State-Owned

<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>

<table>
<thead>
<tr>
<th>Firms Controlled by 772 SOEs</th>
<th>Firm Number</th>
<th>RC share in Galaxy 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.5%</td>
<td>772</td>
<td>10.5%</td>
</tr>
<tr>
<td>26.0%</td>
<td>40,234</td>
<td>26.6%</td>
</tr>
<tr>
<td>29.9%</td>
<td>52,446</td>
<td>30.9%</td>
</tr>
<tr>
<td>32.8%</td>
<td>63,678</td>
<td>34.0%</td>
</tr>
<tr>
<td>33.9%</td>
<td>69,385</td>
<td>35.7%</td>
</tr>
</tbody>
</table>
Connectivity and Firm Characteristics (Current Network)

### Legal-Person Firms

<table>
<thead>
<tr>
<th>Galaxy 1</th>
<th>log RC</th>
<th>Age</th>
<th>SOE</th>
<th>log Y</th>
<th>log YK</th>
<th>log YL</th>
<th>log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.295</td>
<td>3.359</td>
<td>0.204</td>
<td>0.616</td>
<td>-0.370</td>
<td>0.323</td>
<td>0.149</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Closeness</th>
<th>log RC</th>
<th>Age</th>
<th>SOE</th>
<th>log Y</th>
<th>log YK</th>
<th>log YL</th>
<th>log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.048</td>
<td>4.910</td>
<td>1.163</td>
<td>2.136</td>
<td>-0.757</td>
<td>1.148</td>
<td>0.665</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distance</th>
<th>log RC</th>
<th>Age</th>
<th>SOE</th>
<th>log Y</th>
<th>log YK</th>
<th>log YL</th>
<th>log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.230</td>
<td>-0.421</td>
<td>-0.083</td>
<td>-0.100</td>
<td>0.041</td>
<td>-0.052</td>
<td>-0.027</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

### Natural-Person Firms

<table>
<thead>
<tr>
<th>Galaxy 2</th>
<th>log RC</th>
<th>Age</th>
<th>SOE</th>
<th>log Y</th>
<th>log YK</th>
<th>log YL</th>
<th>log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.843</td>
<td>0.304</td>
<td>-0.045</td>
<td>0.031</td>
<td>-0.079</td>
<td>0.014</td>
<td>-0.009</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Closeness</th>
<th>log RC</th>
<th>Age</th>
<th>SOE</th>
<th>log Y</th>
<th>log YK</th>
<th>log YL</th>
<th>log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.195</td>
<td>4.241</td>
<td>0.011</td>
<td>0.603</td>
<td>-0.594</td>
<td>0.211</td>
<td>0.003</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distance</th>
<th>log RC</th>
<th>Age</th>
<th>SOE</th>
<th>log Y</th>
<th>log YK</th>
<th>log YL</th>
<th>log TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.010</td>
<td>0.029</td>
<td>0.007</td>
<td>0.035</td>
<td>0.042</td>
<td>0.035</td>
<td>0.034</td>
</tr>
<tr>
<td>Sample</td>
<td>Full</td>
<td>Full</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
<td>NBS</td>
</tr>
</tbody>
</table>

Closeness(i) = 1 / Sum(distance(i,j), j≠i), standardized into [0,1] (0, periphery; 1 center)
## Connectivity and Firm Characteristics (Current Network, Cont’d)

Using eccentricity and historical network data: Very similar results

<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
All Companies Owned by East Hope

• Red: State JV
• Grey: Solely Owned by East Hope
• Blue: Private JV
Agriculture
- Beijing Fangshan Hope Animal Feed Co.
- East Hope Changge Animal Nutrition Co.
- Beijing Beautiful Hope Animal Feed Co.
- Wuhu Hope Animal Feed Co.
- Funing Hope Animal Feed Co.
- Yangzhou Hope Animal Feed Co.
- Xinyang East Hope Animal Nutrition Co.
- Nanyang Hope Animal Feed Co.
- Xunxian Zhongyuan Hope Animal Feed Co.
- East Hope (Neiqiu) Livestock Co.

Heavy Industry
- East Hope (Sanmenxia) Alumina Co.
- CNMC Alumina Development Co.
- Chongqing Pengwei Petroleum and Chemical Co.
- Chongqing Liangjiang Energy Development Co.
- Wulong Hongneng Coal Co.
- Chongqing Zhengxin Alumina Co.
- Chongqing Zhengyang New Material Co.
- Sanmenxia Dachang Mining Co.
- Hulun Buir Shenbao Dongneng Coal Co.
- Xinjiang Tianlong Hope Energy Co.
- Baotou Horizon Environment Protection Technique Co.

Other
- Shanghai Yehai Trade Co.
- Bright Dairy & Food Co.
- Minsheng Life Insurance Co.
- Chengdu East Hope Tianxiang Real Estate Co.
- Chengdu Dongxiang Property Management Service Co.

Notes:
1. Joint ventures established by East Hope Group and SOEs are highlighted by *Red*, while joint ventures established by East Hope Group and other private firms, excepted for the other three Hope Groups held by Mr. LLIU Yongxing’s brothers, are highlighted by *Blue*.
2. Within each industry, joint ventures are sorted by dates of establishment.
East Hope’s Expansion through Ownership Network

• All State JVs are outside Sichuan (home province)

• State JVs are the oldest and private JVs are the youngest

• Agriculture: 116 companies (9 state JVs; 1 private JV)
  • State JVs are on average 9.7 year older than non-JVs (i.e., those solely owned by East Hope).
  • Among the 10 oldest companies, 7 are state JVs (recall 9 state JV in total).

• Heavy industry (aluminum, energy, etc.): 60 companies (9 state JVs; 1 private JV)
  • State JVs are on average 3.0 years older than non-JVs.

• Other industries: 39 companies (1 state JVs; 4 private JVs)
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)
Basic Information

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

Log (1-F) vs Log Conglomerate Size
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>
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>Proportion</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](image)
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.
Regional and Sectoral Concentrations
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

Average Log YK in a Conglomerate (NBS Sample)

The 1st Percentile Log YK in a Conglomerate (NBS Sample)

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.</th>
<th>Log 90&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 95&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 99&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log Ave.</th>
<th>Log 90&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 95&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 99&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log Ave.</th>
<th>Log 10&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 5&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 1&lt;sup&gt;st&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Log C Size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.158</td>
<td>0.296</td>
<td>0.332</td>
<td>0.414</td>
<td>0.082</td>
<td>0.202</td>
<td>0.256</td>
<td>0.479</td>
<td>0.022</td>
<td>-0.083</td>
<td>-0.114</td>
<td>-0.321</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.032)</td>
<td>(0.033)</td>
<td>(0.038)</td>
<td>(0.018)</td>
<td>(0.030)</td>
<td>(0.038)</td>
<td>(0.057)</td>
<td>(0.021)</td>
<td>(0.026)</td>
<td>(0.031)</td>
<td>(0.049)</td>
</tr>
<tr>
<td><strong>N</strong></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.</th>
<th>Log 90&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 95&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 99&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log Ave.</th>
<th>Log 90&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 95&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 99&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log Ave.</th>
<th>Log 10&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 5&lt;sup&gt;th&lt;/sup&gt;</th>
<th>Log 1&lt;sup&gt;st&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ave. Distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.209</td>
<td>-0.311</td>
<td>-0.363</td>
<td>-0.467</td>
<td>-0.076</td>
<td>-0.146</td>
<td>-0.197</td>
<td>-0.327</td>
<td>0.007</td>
<td>0.113</td>
<td>0.138</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.027)</td>
<td>(0.027)</td>
<td>(0.030)</td>
<td>(0.015)</td>
<td>(0.025)</td>
<td>(0.032)</td>
<td>(0.050)</td>
<td>(0.018)</td>
<td>(0.021)</td>
<td>(0.025)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>N</strong></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></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></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>-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>
<td>0.018</td>
<td>0.021</td>
</tr>
<tr>
<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>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td><strong>C. Distance * C. State Dummy</strong></td>
<td>-0.051</td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td></td>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td><strong>C. Distance * Log C. Size</strong></td>
<td>-0.069</td>
<td></td>
<td></td>
<td></td>
<td>-0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.003)</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>0.02</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>37,123</td>
<td>37,123</td>
</tr>
</tbody>
</table>
Summary of the Stylized Facts

• Network features: Firm size, age and SOE share (YK) negatively (positively) correlated with firm distance to the center of the network

• Conglomerate analysis: Firm size, age and SOE share (YK) negatively (positively) correlated with distance to the center of a conglomerate

• Conglomerate analysis: Strong correlation between conglomerate size and the top firm size or the bottom YK ratio in a conglomerate (weak correlation between conglomerate size and the average firm size or YK ratio in a conglomerate)
Take-Aways

• Growth of incumbent conglomerates is the key.

• Firms in the center of a conglomerates are important for the size (and growth?) of the conglomerate.

• No severe misallocation across conglomerates
A Simple Model

• 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 constant saving rate (s).
Firm Dynamics

• $N(t)$ incumbents and $1 - N(t)$ potential entrants

• Incumbents own capital and invest in matched entrants via random search.

• The new entrants will make production, return a fixed proportion of output to their investors and then become incumbents (capital owners) in the subsequent periods.

• Myopic: Both incumbents and potential entrants are run by one-period-lived entrepreneurs (can be relaxed).
Firm Characteristics and Distributions

• Incumbents: \((A, Z^c)\), where \(A \in \{A_1, \ldots, A_I\}\) is capital owned by incumbent 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 \(P^c(A, Z^c, 0), P^e(Z^e, 0)\) and \(N(0)\).
One-to-One Matching

- An incumbent randomly meets a potential entrant in each period and then form a joint venture under some conditions.

- Extension: Allowing joint venture between incumbents.
Capital Allocation in a Joint Venture

• In a matched pair, the incumbent will allocate capital by

\[ \pi^c(A, Z^c, Z^e) = \max_K Z^c(A - K)^\alpha + \beta Z^e K^\alpha \]

• Denote \( K(A, Z^c, Z^e) \) the capital allocation rule. Then, the entrant’s profit is

\[ \pi^e(A, Z^c, Z^e) = (1 - \beta)Z^e K(A, Z^c, Z^e)^\alpha \]

• \( \beta \) captures contractual frictions (efficient capital allocation by Nash Bargaining)
Capital Allocation in a Joint Venture

• Fixed cost $C$ for forming/maintaining a joint venture (paid by incumbent)

• Surplus for incumbent and entrant are

\[
\tilde{\pi}^c (A, Z^c, Z^e) = \max \{ \pi^c (A, Z^c, Z^e) - Z^c A^\alpha - C, 0 \}
\]

\[
\tilde{\pi}^e (A, Z^c, Z^e) = \pi^e (A, Z^c, Z^e)
\]

• Incumbent more capital + Entrant high TFP => More surplus of the joint venture
Random Search

• Incumbent and entrant’s search efforts are $\lambda^c(A, Z^c, t)$ and $\lambda^e(Z^e, t)$.

• Search cost is $SC(\lambda^i)$, where $SC(\cdot)$ is homogeneous of degree $1$, convex and satisfies $SC(0) = 0$, $SC'(\cdot) > 0$, $SC''(\cdot) > 0$, $\lim_{x \to \infty} SC(x) = \infty$.
  - Example: $SC(\lambda^i) = \frac{B}{\eta} (\lambda^i)^\eta$, where $\eta > 1$. 
Random Search

• “Market tightness” on incumbent and entrant sides:

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

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

\[
\lambda^c(A, Z^c, t) \cdot \theta^c(t) \cdot \Gamma^e(Z^e, t), \text{ where } \Gamma^e(Z^e, t) = \frac{\lambda^e(Z^e,t) p^e(Z^e,t)}{\sum_{Z^e} \lambda^e(Z^e,t) p^e(Z^e,t)}
\]

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

\[
\lambda^e(Z^e, t) \cdot \theta^e(t) \cdot \Gamma^c(A, Z^c, t), \text{ where } \Gamma^c(A, Z^c, t) = \frac{\lambda^c(A,Z^c,t) p^c(A,Z^c,t)}{\sum_{A,Z^c} \lambda^c(A,Z^c,t) p^c(A,Z^c,t)}
\]
Optimal Search Efforts

• The optimal search effort for incumbents and potential entrants:

\[
\max_{\lambda^c} \lambda^c \theta^c \sum_{Z^e} \pi^c (A, Z^c, Z^e) \cdot \Gamma^e (Z^e, t) - SC(\lambda^c)
\]

\[
\max_{\lambda^e} \lambda^e \theta^e \sum_{A, Z^c} \pi^e (A, Z^c, Z^e) \cdot \Gamma^c (A, Z^c, t) - SC(\lambda^e)
\]

• The FOCs solve \( \lambda^c = \lambda^c (A, Z^c, t) \) and \( \lambda^e = \lambda^e (Z^e, t) \).

• Denote the matching rate: \( M(A, Z^c, Z^e, t) = 1 \) if \( \pi^c (A, Z^c, Z^e) > 0 \) and 0 otherwise.
Evolution of Capital

• Incumbent’s next-period capital is $A - K(A, Z^c, Z^e) + s \cdot (A, Z^c, Z^e)$.

• Denote $\Phi^c(A'|A, Z^c, Z^e) = 1$ if $A' \neq A$ and 0 otherwise.

• The current entrants will become incumbents in the next period, with capital $K(A, Z^c, Z^e) + \pi^e (A, Z^c, Z^e)$.

• Likewise, Denote $\Phi^e(A'|A, Z^c, Z^e) = 1$ if $A' > 0$ and 0 otherwise.
Evolution of Incumbents and Entrants

- The measure of matched entrants with $A'$ and $Z^e$: $N^{e,c}(A', Z^e, t) =$
  
  $$P^e(Z^e, t) \cdot (1 - N(t)) \cdot \sum_{A,Z^c} \lambda^e(Z^e, t) \cdot \theta^e(t) \cdot \Gamma^c(A, Z^c, t) \cdot M(A, Z^c, Z^e, t) \cdot \Phi^e(A' | A, Z^c, Z^e)$$

- The measure of incumbents with $A'$ and $Z^c$: $N^{c,c}(A', Z^c, t) =$
  
  $$\sum_A P^c(A, Z^c, t) \cdot N(t) \cdot \sum_{Z^e} \lambda^c(A, Z^c, t) \cdot \theta^c(t) \cdot \Gamma^e(Z^e, t) \cdot M(A, Z^c, Z^e, t) \cdot \Phi^c(A' | A, Z^c, Z^e)$$

  $$+ P^c(A', Z^c, t) \cdot N(t) \sum_{Z^e} [1 - \lambda^c(A', Z^c, t) \cdot \theta^c(t) \cdot \Gamma^e(Z^e, t) \cdot M(A', Z^c, Z^e, t)]$$
Aggregate Dynamics

• The measure of entrants evolves by

\[ N(t + 1) = \sum_{A,Z^c} N_{c,c}^{c,c}(A, Z^c, t) + \sum_{A,Z^c} N_{e,c}^{e,c}(A, Z^c, t) \]

• The distributions evolve by

\[ P_c^c(A, Z^c, t + 1) = \frac{N_{c,c}^{c,c}(A, Z^c, t) + N_{e,c}^{e,c}(A, Z^c, t)}{N(t + 1)} \]

\[ P_e^e(Z^e, t + 1) = \frac{N_e(Z^e, t) - \sum_A N_{e,c}^{e,c}(A, Z^c, t)}{1 - N(t + 1)} \]
Outcomes

• Firm TFP, size and YK distributions

• A connected network (all initial incumbents are connected)

• Conglomerate Formation

• Distance and degree distributions
Ownership Layer Distribution

• Define $D(x|A, Z^c, t)$ the distribution of incumbents $(A, Z^c)$ on the ownership layers (i.e., the distance to the “ultimate” capital owner), $x = 1, 2, \ldots$.

• The measure of matched entrants with $A', Z^e$ and $x + 1$:

$$ND^{e,c}(x + 1|A', Z^e, t) = P^e(Z^e, t) \cdot (1 - N(t))$$

$$\sum_{A, Z^c} D(x|A, Z^c, t) \cdot \lambda^e(Z^e, t) \cdot \theta^e(t) \cdot \Gamma^c(A, Z^c, t) \cdot M(A, Z^c, Z^e, t) \cdot \Phi^e(A'|A, Z^c, Z^e)$$
Ownership Layer Distribution (cont’d)

• The measure of incumbents with $A'$, $Z$ and $x$:

$$ND^{c,c}(x|A', Z^c, t)$$

$$= \sum_A D(x|A, Z^c, t) \cdot P^c(A, Z^c, t) \cdot N(t) \cdot \sum_{Z^e} \lambda^c(A, Z^c, t) \cdot \theta^c(t) \cdot \Gamma^e(Z^e, t) \cdot M(A, Z^c, Z^e, t) \cdot \Phi^c(A'|A, Z^c, Z^e)$$

$$+ D(x|A', Z^c, t) \cdot P^c(A', Z^c, t) \cdot N(t) \cdot \sum_{Z^e} [1 - \lambda^c(A', Z^c, t) \cdot \theta^c(t) \cdot \Gamma^e(Z^e, t) \cdot M(A', Z^c, Z^e, t)]$$

• Evolution:

$$D(x|A, Z^c, t + 1) = \frac{ND^{c,c}(x|A, Z^c, t) + ND^{e,c}(x|A, Z^c, t)}{N(t + 1)}$$
Degree Distribution

• Define $Q(x|A, Z^c, t)$ the distribution of incumbents $(A, Z^c)$ on the degrees (i.e., the number of invested firms), $x = 0, 1, \ldots$.

• The measure of matched entrants with $A^\prime, Z^e$ and 0:

$$NQ^{e,c}(0|A^\prime, Z^e, t) = P^e(Z^e, t) \cdot (1 - N(t))$$

$$\sum_{A, Z^c} \lambda^e(Z^e, t) \cdot \theta^e(t) \cdot \Gamma^c(A, Z^c, t) \cdot M(A, Z^c, Z^e, t) \cdot \Phi^e(A'|A, Z^c, Z^e)$$
Degree Distribution (cont’d)

- The measure of incumbents with $A'$, $Z^c$ and $x + 1$:

$$ NQ^{c,c}(x + 1|A', Z^c, t) $$

$$ = \sum_A Q(x|A, Z^c, t) \cdot P^c(A, Z^c, t) \cdot N(t) \cdot \sum_{Z^e} P^e(Z^e, t) \cdot M(A, Z^c, Z^e, t) \cdot \Phi^c(A'|A, Z^c, Z^e) $$

$$ + D(x + 1|A', Z^c, t) \cdot P^c(A', Z^c, t) \cdot N(t) \cdot \sum_{Z^e} [1 - \lambda^e(A', Z^c, t) \cdot \theta^c(t) \cdot \Gamma^e(Z^e, t) \cdot M(A', Z^c, Z^e, t)] $$

- Evolution:

$$ Q(x|A, Z^c, t + 1) = \frac{NQ^{c,c}(x|A, Z^c, t) + NQ^{e,c}(x|A, Z^c, t)}{N(t + 1)} $$
Simulations (50 Initial Incumbents)
Three Largest Conglomerates among 50 Identified by Multilevel Algorithm
Structural Estimation (VERY PRELIMINARY)

• Targeted moments for firms in Galaxy 1 (six moments): Variance of firm size, TFP and YK and their covariance with distance to the center (central and provincial SOEs)

• Targeted moments for conglomerates in Galaxy 1 (seven moments): Variance of conglomerate size, variance of top/bottom firm size, TFP and YK in a conglomerate and their covariance with conglomerate size
Specification and Estimation

• Simulated method of moments: Estimating $\beta$, $s$, $\sigma_{\log K}$, $\sigma_{\log Z}$, $\eta$ and $C$

• Very preliminary: Identity matrix, Grid search on coarse parameter space, no standard errors ...
• Predetermined parameters: $\alpha = 0.5$ and $N(0) = 0.05$

• Estimated parameters: $\beta = 0.7$, $s = 0.7$, $\sigma_{logA} = 1.5$, $\sigma_{logZ} = 1.0$, $\eta = 0.2$ and $C = 0.05$

<table>
<thead>
<tr>
<th></th>
<th>Firms</th>
<th></th>
<th></th>
<th>Conglomerates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Model</td>
<td></td>
<td>Data</td>
<td>Model</td>
</tr>
<tr>
<td>$\text{var}(\log_{RC})$</td>
<td>2.626</td>
<td>1.681</td>
<td>$\text{var}(\log_{CSize})$</td>
<td>0.658</td>
<td>0.711</td>
</tr>
<tr>
<td>$\text{var}(\log_{tfp})$</td>
<td><strong>1.221</strong></td>
<td><strong>0.55</strong></td>
<td>$\text{var}(\log_{RC_99th})$</td>
<td>0.499</td>
<td>0.491</td>
</tr>
<tr>
<td>$\text{var}(\log_{yk})$</td>
<td>1.643</td>
<td>0.321</td>
<td>$\text{var}(\log_{tfp_99th})$</td>
<td><strong>1.023</strong></td>
<td><strong>0.114</strong></td>
</tr>
<tr>
<td>$\text{cov}(\text{dist},\log_{RC})$</td>
<td><strong>-0.955</strong></td>
<td><strong>-0.294</strong></td>
<td>$\text{var}(\log_{yk_1st})$</td>
<td>0.705</td>
<td>0.362</td>
</tr>
<tr>
<td>$\text{cov}(\text{dist},\log_{tfp})$</td>
<td>-0.055</td>
<td>-0.045</td>
<td>$\text{cov}(\log_{Csize},\log_{RC_99th})$</td>
<td>0.223</td>
<td>0.353</td>
</tr>
<tr>
<td>$\text{cov}(\text{dist},\log_{yk})$</td>
<td>0.187</td>
<td>0.204</td>
<td>$\text{cov}(\log_{Csize},\log_{tfp_99th})$</td>
<td>0.315</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\text{cov}(\log_{Csize},\log_{yk_1st})$</td>
<td><strong>-0.211</strong></td>
<td><strong>-0.26</strong></td>
</tr>
</tbody>
</table>
Simulated Conglomerate Size Distribution

Conglomerate Size Distribution (Data)

Conglomerate Size Distribution (Simulated)
Simulated Conglomerate Size and Average/Top Log Firm Size
Simulated Conglomerate Size and Average/Bottom Log Firm YK
Aggregate Results
Main Predictions

• Network structures: Firm size and YK negatively correlated with distance to the core

• Conglomerate analysis:
  • Strong correlation between conglomerate size and the top firm size or the bottom YK ratio in a conglomerate (weak correlation between conglomerate size and the average firm size or YK ratio in a conglomerate)
  • Few new conglomerates

• Aggregate implications: Network expansion, firm entry, economic growth and correction of misallocation
Extension I: Introducing Entrant-Specific Distortions

- Potential entrants face tax/subsidy $\tau$

- In a matched pair, the incumbent will allocate capital by

$$\pi^c(A, Z^c, Z^e) = \max_K Z^c (A - K)^\alpha + \beta \frac{Z^e}{1 + \tau} K^\alpha$$

- Denote $K(A, Z^c, Z^e)$ the capital allocation rule. Then, the entrant’s profit is

$$\pi^e(A, Z^c, Z^e) = (1 - \beta) \frac{Z^e}{1 + \tau} K(A, Z^c, Z^e)^\alpha$$
Extension II: Long-Lived Incumbents and Endogenous Matching

- Let $V^c(A, Z^c)$ be the value function of an incumbent $(A, Z^c)$ under the stationary distribution $P^c$ and $P^e$.

- In a meeting, the incumbent will allocate and save capital by

$$\max_{K,K'} (1 - \beta) \left( Z^c (A - K)^\alpha - K' + \frac{1}{1 + r} V^c(K', Z^c) \right)$$

$$+ \beta \left( Z^e K^\alpha + \frac{1}{1 + r} V^c(K, Z^e) \right)$$

- $K = 0$ (failed match) or $K > 0$ satisfying the entry condition.
- $1/(1 + r)$ is a combination of discount factor and exit rate.
Incumbent’s Value Function

- Denote $K(A, Z^c, Z^e)$ the capital allocation rule and $K'(A, Z^c, Z^e)$ the saving rule. Then,

$$V^c(A, Z^c) = \sum_{Z^c, Z^e} P^e(Z^e) \left( (1 - \beta) \left( \frac{Z^c(A - K(A, Z^c, Z^e))^\alpha}{1 + r} - K'(A, Z^c, Z^e) \right) + \beta \left( Z^e K(A, Z^c, Z^e)^\alpha + \frac{1}{1 + r} V^c(K(A, Z^c, Z^e), Z^e) \right) \right)$$

- To be done ...
Interpretations

• Network as financial intermediary?

• Network as informal institutions?

• Limitations of conglomerations
Regional and Sectoral Concentrations
Summary

• Chinese firms are deeply connected through shareholding.

• State-centered conglomerates played a key role in firm entry, resource reallocation and aggregate TFP growth.

• The China model has its strengths and weaknesses. More formal institutions would need to be established for future efficiency gains.
To-Do List

• Finishing structural estimation

• Allowing matching among incumbents

• Introducing financial intermediary

• ...