SIMULATING A SYNTHETIC POPULATION OF ESTABLISHMENTS

Diem-Trinh Le, Giulia Cernicchiaro, Chris Zegras, Joseph Ferreira

mobil.TUM 2016

SimMobility-Long-Term Group
MIT: Joseph Ferreira, Chris Zegras, Roberto Ponce Lopez, Jingsi Shaw
SMART: Yi Zhu, Diem-Trinh Le, Chetan Rogbeer, Gishara Indeewarie
NUS: Mi Diao
Singapore
Area: 719.1 km²
Density: 7,697/km²
GDP (PPP): $82,762
SimMobility Framework

LONG-TERM
Land development and location choices

AGENTs’ POPULATIONS year (t-1)

Households/Individuals
Developers
Firms/Establishments

MID-TERM MODULE

AGENTs’ POPULATIONS year t

Real Estate Market
HH/Establishments
Locations

Labor Market
Workers
Jobs

Distribution Market
Suppliers
Service Sector

— Modular
Parallel and distributed
Publish/subscribe mechanism

MARKET TRANSACTION MODELS

LONG-TERM

High resolution travel behavior
Accessibility
Logistic performances
Tours
Trip chains
Fleet operations schedule

MID-TERM

Daily activity and mobility patterns

SHORT-TERM

Location of HH/Firms
Vehicle ownership
Supply chain structure

MARKET TRANSACTION MODELS

Real Estate Market
HH/Establishments
Locations

Labor Market
Workers
Jobs

Distribution Market
Suppliers
Service Sector
HOW TO SIMULATE A POPULATION OF FIRMS?

- How many firms?
- What kind of firm?
- How much floor area is occupied?
- How many workers?
Step 1: Data Collection

- List of business entities registered at ACRA
- Building data
- National statistics: total employment, total occupied area, etc.

- Establishments’ floor size and number of workers
Step 2: Estimate Establishments’ Size

- **REALIS**
  - Use property transactions data
  - Estimate a unit’s area based on its characteristics (regression model)

- **ACRA**
  - Apply results from REALIS to ACRA dataset to estimate establishment’s floor area

- **ACRA**
  - Convert floor area to employment size
  - Conversion factor: average floor space per worker in Singapore

Data collection → Estimate firm’s size → Adjusting the stats → Distribute numbers
Step 3: Adjust The Numbers

1. **Method: Iterative proportional fitting (IPF).**
2. **Marginal controls: Official statistics from different government agencies.**

<table>
<thead>
<tr>
<th>Planning Area</th>
<th>Industry type</th>
<th>Industry 1</th>
<th>Industry 2</th>
<th>Industry k</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMK</td>
<td>Floor type</td>
<td>No. of Jobs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Office</td>
<td>$N_{j_{pa,ft=office}}^i$</td>
<td>$N_{j_{pa,ft=office}}^i$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Retail</td>
<td>$N_{j_{pa,ft=retail}}^i$</td>
<td></td>
<td>$N_{j_{pa,ft=retail}}^i$</td>
</tr>
<tr>
<td></td>
<td>Warehouse</td>
<td>$N_{j_{pa,ft=warehouse}}^i$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Industrial</td>
<td>$N_{j_{pa,ft=industrial}}^i$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The adjusted number of jobs in each industry for each planning area $N_{j_{pa,k}}^i$
Step 4: Distribute Jobs & Establishments to Buildings

- Establishment \( e \)
- Building \( i \)
- Industry type \( k \)
- Number of employees \( j \)
- Occupied floor area \( f \)
- Floor type \( ft \)

\[ n_{i,k}^j = N_{pa,k} \]
\[ n_{i,k}^e = N_{ft} \]
\[ n_{e}^f = N_{i,ft} \]

- \( n_{ik}^c \): the number of estabs/jobs/ floor area in building \( i \) of industry \( k \).
- \( N_k \): the total numbers of estab. (\( c = e \)), jobs (\( c = j \)), and floor size (\( c = f \)) of a particular industry type \( k \) in Singapore.
- \( N_{i,ft} \): the total numbers of estab. for a particular building and floor type.
Step 4: Distribute Jobs & Establishments to Buildings

Data collection
Estimate firm’s size
Adjusting the stats
Distribute numbers

Office
Retail
Industrial
Warehouse

<table>
<thead>
<tr>
<th>Floor type</th>
<th>Establishments</th>
<th>Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pop. Syn</td>
<td>SingStats</td>
</tr>
<tr>
<td>Office</td>
<td>56,800</td>
<td>139,718</td>
</tr>
<tr>
<td>Retail</td>
<td>76,648</td>
<td>9,577</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>16,368</td>
<td>11,076</td>
</tr>
<tr>
<td>Warehouse</td>
<td>10,184</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td><strong>160,000</strong></td>
<td><strong>160,371</strong></td>
</tr>
</tbody>
</table>
Buildings and Estab. by Job Size

No. of Jobs in buildings

No. of Jobs in Estab.
Locations of estab. with retail floor type
Locations of estab. with industrial floor type
Locations of estab. with warehouse floor type
Industrial jobs at zonal level

- 1-200
- 201-1000
- 1001-3000
- 3001-5000
- >5000
Warehouse jobs at zonal level
Conclusions

Data needed

- A sample of establishments
- Aggregate data on employment, establishments by industry and by area
- Building data

Office
Retail
Industrial
Warehouse
• SMART website http://smart.mit.edu
• Future Urban Mobility Lab http://ares.lids.mit.edu/fm/
• **Number of ACRA establishments by role and type.**
• **Conversion factor**
• **Building data**
• **Location of ACRA establishments.**
<table>
<thead>
<tr>
<th>Classification</th>
<th>2012</th>
<th>2015</th>
<th>15/12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturer</td>
<td>28,281</td>
<td>33,290</td>
<td>1.18</td>
</tr>
<tr>
<td>Supplier/Wholesaler</td>
<td>67,489</td>
<td>90,020</td>
<td>1.33</td>
</tr>
<tr>
<td>Retailer</td>
<td>46,665</td>
<td>64,914</td>
<td>1.39</td>
</tr>
<tr>
<td>Carrier</td>
<td>2,695</td>
<td>3,575</td>
<td>1.33</td>
</tr>
<tr>
<td>Other</td>
<td>137,778</td>
<td>200,302</td>
<td>1.45</td>
</tr>
<tr>
<td><strong>Total by Role</strong></td>
<td>282,908</td>
<td>392,101</td>
<td>1.39</td>
</tr>
<tr>
<td>Entity without branches</td>
<td>264,996</td>
<td>373,540</td>
<td>1.41</td>
</tr>
<tr>
<td>Entity with branches</td>
<td>4,801</td>
<td>5028</td>
<td>1.05</td>
</tr>
<tr>
<td>Branches</td>
<td>13,111</td>
<td>13,533</td>
<td>1.03</td>
</tr>
<tr>
<td><strong>Total by Branch</strong></td>
<td>282,908</td>
<td>392,101</td>
<td>1.39</td>
</tr>
<tr>
<td>Business entities</td>
<td>84,200</td>
<td>124,527</td>
<td>1.48</td>
</tr>
<tr>
<td>Business branches</td>
<td>13,111</td>
<td>13,533</td>
<td>1.03</td>
</tr>
<tr>
<td>Company entities</td>
<td>178,436</td>
<td>242,496</td>
<td>1.36</td>
</tr>
<tr>
<td>LLP entities</td>
<td>7,124</td>
<td>11,428</td>
<td>1.60</td>
</tr>
<tr>
<td>LP entities</td>
<td>37</td>
<td>124</td>
<td>3.35</td>
</tr>
<tr>
<td><strong>Total by Type</strong></td>
<td>282,908</td>
<td>392,108</td>
<td>1.39</td>
</tr>
</tbody>
</table>
ACRA dataset of live entities in Feb. 2015 minus entities registered after 31/12/2012.

<table>
<thead>
<tr>
<th>Description</th>
<th>Unique values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of establishments</td>
<td>282,907</td>
</tr>
<tr>
<td>Unique Entity Number</td>
<td>269,796</td>
</tr>
<tr>
<td>SSIC1</td>
<td>858</td>
</tr>
<tr>
<td>Full address</td>
<td>142,642</td>
</tr>
<tr>
<td>Postcode</td>
<td>37,566</td>
</tr>
</tbody>
</table>

ACRA Sample

<table>
<thead>
<tr>
<th>Description</th>
<th>Unique values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of establishments</td>
<td>142,642</td>
</tr>
<tr>
<td>Unique Entity Number</td>
<td>134,894</td>
</tr>
<tr>
<td>SSIC1</td>
<td>826</td>
</tr>
<tr>
<td>Postcode</td>
<td>37,566</td>
</tr>
</tbody>
</table>

*SSIC: Singapore Standard Industrial Classification*
<table>
<thead>
<tr>
<th>Floor Type</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office</td>
<td>9.282070</td>
</tr>
<tr>
<td>Retail</td>
<td>4.013971</td>
</tr>
<tr>
<td>Industrial</td>
<td>58.57063</td>
</tr>
<tr>
<td>Warehouse</td>
<td>33.01842</td>
</tr>
</tbody>
</table>
## Industry Type and Floor Type

<table>
<thead>
<tr>
<th>SSIC</th>
<th>Section</th>
<th>Floor type</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Manufacturing</td>
<td>Industrial</td>
</tr>
<tr>
<td>H</td>
<td>Transportation and Storage</td>
<td>Warehouse</td>
</tr>
<tr>
<td>J</td>
<td>Information and Communications</td>
<td>Office</td>
</tr>
<tr>
<td>K</td>
<td>Financial and Insurance Activities</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>Real Estate Activities</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>Professional, Scientific and Technical Activities</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Administrative and Support Service Activities</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Wholesale and Retail Trade</td>
<td>Retail</td>
</tr>
<tr>
<td>I2</td>
<td>Food Service Activities</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>Health and Social Services</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>Arts, Entertainment and Recreation</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Other Service Activities</td>
<td></td>
</tr>
</tbody>
</table>
Regression Models

• Two regression models:
  – Floor type “office” and “retail”:
    • Commercial property transactions 1995-Oct 2015
    • 16383 observations
  – Floor type “industrial” and “warehouse”
    • Factory/warehouse property transactions 1995-Oct 2015
    • 23289 observations

• Predictors:
  – Location
  – Building type
  – Floor level
  – Floor type
Building Data

• Building dataset by Yi:
  – Number of buildings: 109,709.
  – Info: location, building type, est. total space, est. floor area for different floor types (*not for all buildings).

• What was added:
  – Occupied floor area for each floor type.
  – Number of jobs for each industry type (15 types).
  – Number of establishment for each industry type.
  – No. of buildings that were assigned with jobs: 45,814.
Establishment Population Synthesis

- **List of buildings**
  - Building ID
  - Location (planning area, postcode)
  - Number of jobs for each industry

- **List of establishments**
  - Establishment ID
  - Location (planning area, postcode)
  - Size (floor area occupied and number of jobs)
  - Floor type
  - Industry type (SSIC section)

- **List of jobs**
  - Job ID
  - Establishment ID
  - Industry type
  - Location (planning area, postcode)
SYN POP VS. ACRA
Locations of ACRA estab. with office floor type
Locations of syn. pop. estab. with office floor type
Locations of ACRA vs. syn. pop. estab. with office floor type
Locations of ACRA estab. with retail floor type
Locations of syn. pop. estab. with retail floor type
Locations of ACRA vs. syn. pop. estab. with retail floor type
Locations of ACRA estab. with industrial floor type
Locations of syn. pop. estab. with industrial floor type
Locations of ACRA vs. syn. pop. estab. with industrial floor type
Locations of ACRA estab. with warehouse floor type
Locations of syn. pop. firms with warehouse floor type
Locations of ACRA vs. syn. pop. estab. with warehouse floor type