Constructing a Population of Singapore Establishments for Modeling Firm Behavior in SimMobility

Diem Trinh LE¹, Giulia CERNICCHIARO¹, Yi ZHU¹, Chris ZEGRAS², Joseph FERREIRA²
¹Singapore-MIT Alliance for Research and Technology (SMART), ²Massachusetts Institute of Technology (MIT)

1. OBJECTIVE
   • To construct a synthetic population of establishments in Singapore for incorporation into the SimMobility platform.
     - The population synthesis aims to reflect the reality of firms in Singapore in December 2012
     - Data to be generated include (1) establishments’ locations, (2) industry type, (3) employment size, and (4) occupied floor area

2. DATA
   • Building data from Singapore Land Authority (SLA)
   • Land use data from Urban Redevelopment Authority (URA)
   • Aggregate data on employment and businesses from Department of Statistics (SingStats) and Ministry of Manpower (MOM)
   • The study sample: Establishments registered at the Accounting and Corporate Regulatory Authority (ACRA) and SDirectory. Information included: (1) establishment’s name, (2) address, (3) industry type, (4) some sales and capital
     - Need to estimate establishment’s size

3. METHOD
   4-step procedure
   • Estimate the occupied floor space for each establishment: Regression models on observed property transactions (REALIS dataset)
     - A model for services and retail establishments (adj. R² = 0.39, NRMSE = 0.07)
     - A model for factory and warehouse establishments (adj. R² = 0.65, NRMSE = 0.06)
   • Estimate establishments’ job size based on their occupied floor area
     - Average floor space per employee for different floor types was used as a conversion factor between floor size and employment size
   • Adjust the number of jobs for establishments in each floor type at the planning area level using Interactive Proportional Fitting method (IPF)
   • Distribute the number of jobs among buildings proportionally to their approximated occupied floor area within each planning area and by floor type

4. RESULTS
   • A total number of 160,000 synthetic establishments were generated, each has the following attributes:
     - Location (postcode and planning area)
     - Floor area occupied and employment size
     - Industry type (SSIC1 code) and floor type (retail, office, industrial, or warehouse)
   • Jobs and establishments are relatively well distributed among the buildings according to the availability of suitable floor space.

5. CONCLUSIONS
   • A synthetic population of establishments in Singapore was constructed, which can be used for modeling firm behaviour in SimMobility.
   • The method can be applied in other contexts, if the following data are available: (1) a sample of establishments, (2) aggregate data on employment by industry and by area, and (3) building data.
   • Highlights:
     - The numbers of jobs in each industry type and in each planning area are close to real numbers.
     - The numbers of buildings that have jobs are similar to the number of buildings in the ACRA dataset.
   • Further improvements:
     - More detailed level of industry types
     - More realistic distribution of jobs by industry type

Acknowledgements The research is supported by the National Research Foundation, Prime Minister’s Office, Singapore, under its CREATE programme, Singapore-MIT Alliance for Research and Technology (SMART) Future Urban Mobility (FM) IRG.