USING SPATIAL MICROSIMULATION FOR INTEGRATED LAND USE AND TRANSPORT MODELING IN METRO MANILA

IMPROVING TRANSPORT STATISTICS FOR POLICY ANALYSIS

By

Noriel Christopher C. Tiglao and Mark Angelo Y. Tacderas

Presented by

Noriel Christopher C. Tiglao

National College of Public Administration and Governance, UP Diliman

nctiglao@up.edu.ph

14th NCS National Convention on Statistics

1-3 October 2019 | Crowne Plaza Manila Galleria

Organized by the Philippine Statistical System

Spearheaded by the Philippine Statistics Authority

Outline

- Introduction
- Key Urban Issues
- Considerations for Urban Modeling
- Spatial Microsimulation Approach
- Applications
  - Location Choice
  - Car Ownership
- Concluding Remarks

Key Urban Issues in Metro Manila

- Rapid population growth
- Urban primacy
- Urban housing problem
- Urban poverty

Urban Modeling Issues

McGee (1971)

- Urbanization in the third world is happening at a compressed time-scale, greater magnitude and complex socio-economic conditions.
- Urbanization studies in the third world should be undertaken in the broader investigation of the ‘forces influencing the society and country as a whole’

Urban Modeling Issues

Lakshmanan (1981)

- The modeler must be in touch with the policy maker and the policy maker should understand the attributes of the models
- A mistake to make isomorphic transfers of urban development models from developed cities to the developing world

Mohan (1979)

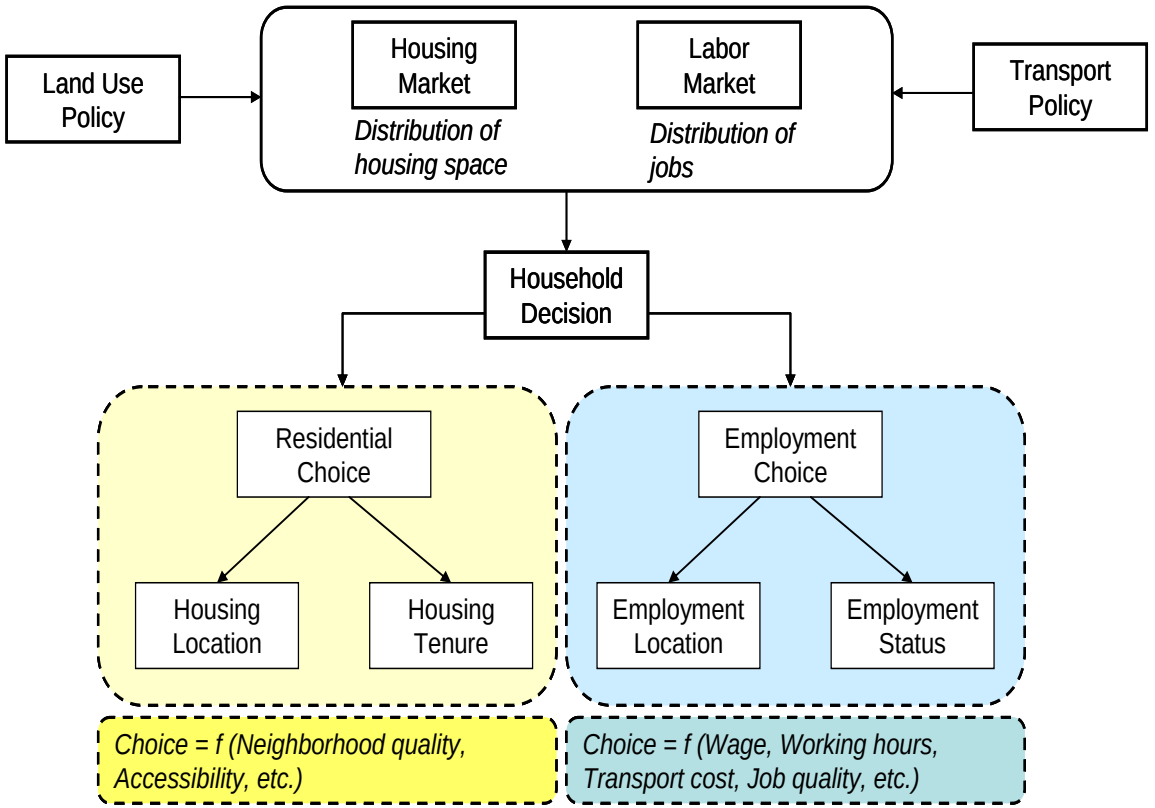
- Attention should be given to the particular institutional structure of the country concerned
- Clear and reliable information is required in areas such as transport, housing, and the informal sector in developing countries.

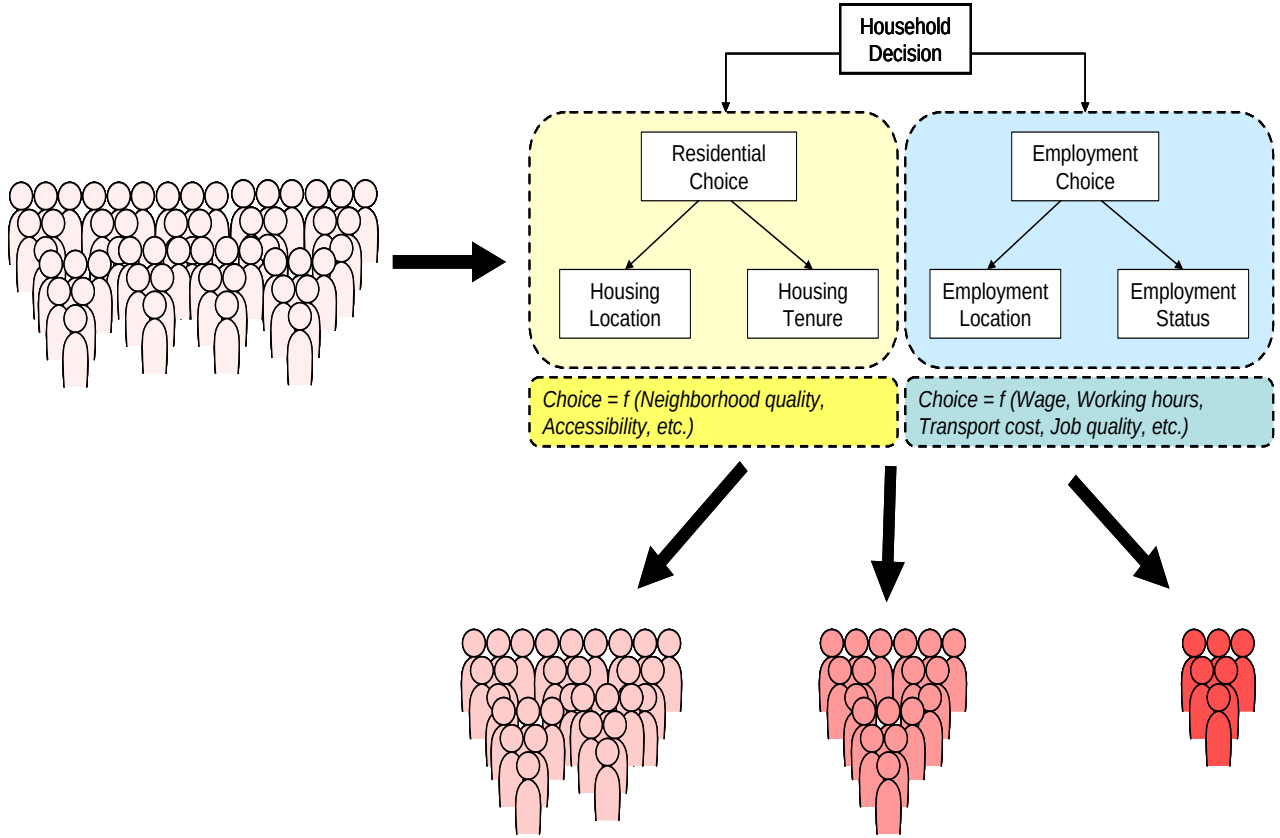
Urban Modeling Issues

Tiglao and Tsutsumi (2003)

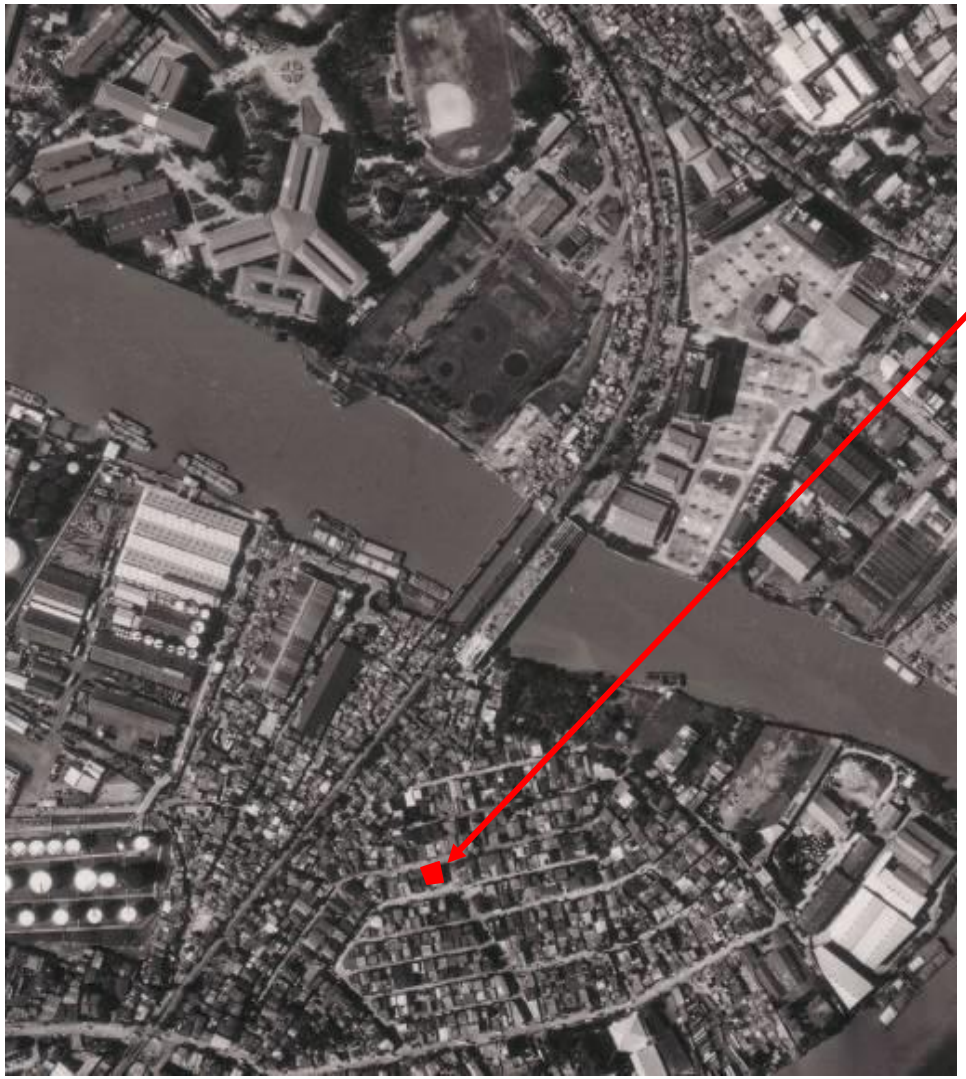
- Highlighted key modeling issues that need to be tackled in the development of urban models for developing countries
- Rapid growth in population is also coupled by the presence of severe economic inequality among individuals and households. The large gap between the rich and the poor is very much evident in the housing and labor sectors of the urban economy.
- Need to effectively distinguish the various income and social groups. The current modeling practice of defining ‘representative households’ needs to be refined in order to capture the household structure in developing countries at a disaggregated level

Spatial Analysis Framework





Spatial attribute of household microdata



HOUSEHOLD
Variables
Province-ID District-ID Barangay-ID Household-ID Household size Age of hh head Sex of hh head Marital status of hh head Education of hh head (Economic activity of hh head) (Occupation of hh head) (Employment sector of hh head) (Employment status of hh head) Members [Vector] Building type Roof type Wall type State of repair Year built (Household income) (Housing status) (Housing value)
Methods
GetEconomicActivityofHead GetOccupationofHead GetEmploymentSectorofHead GetEmploymentStatusofHead GetHouseholdIncome GetHousingStatus GetHousingValue ...

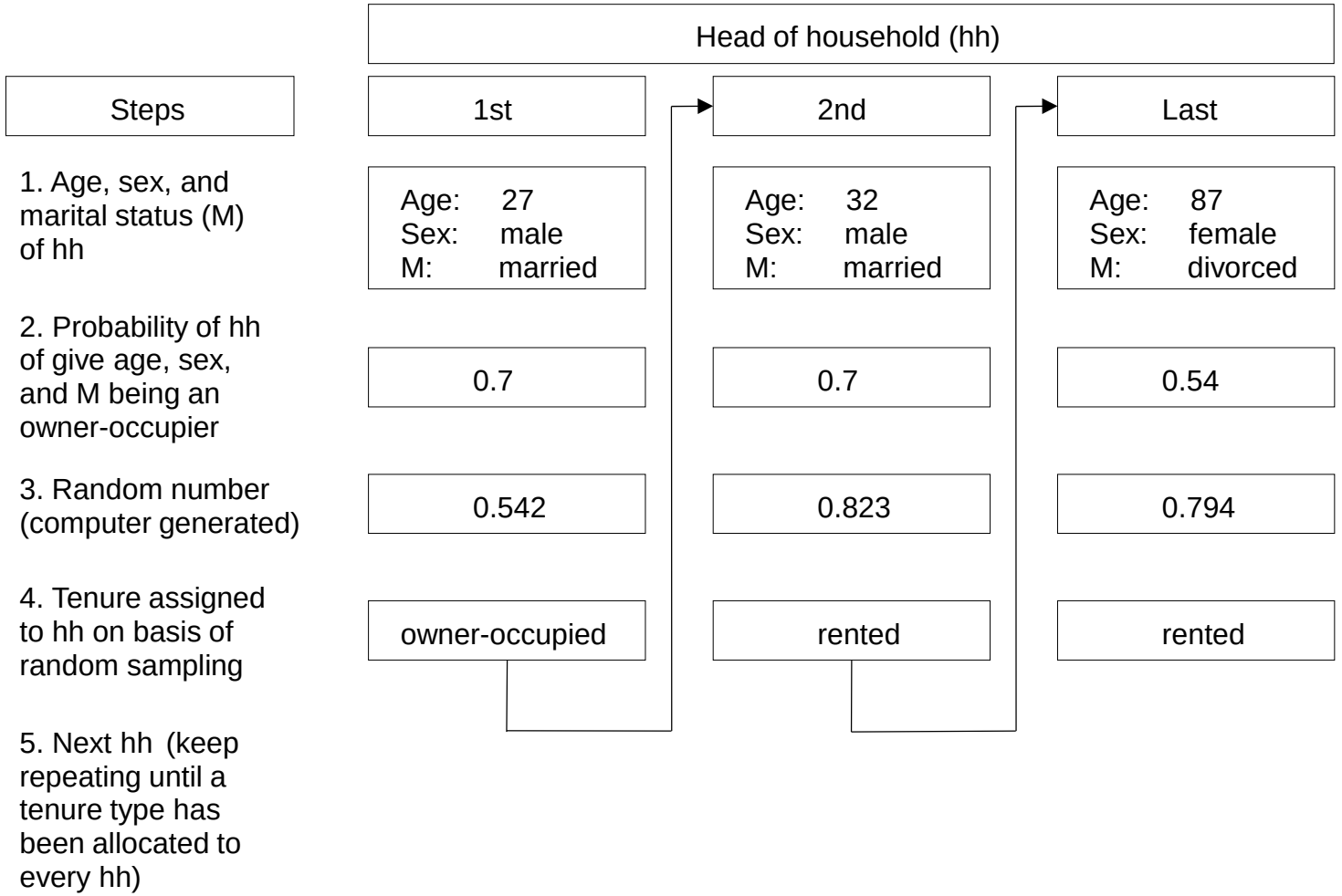
Spatial Microsimulation Approach

- Developed by Guy Orcutt in 1957; ‘A new kind of socio-economic system’
- Directly concerned with microunits such as persons, households, or firms
- Models lifecycle by the use of conditional probabilities
- One major objective in spatial microsimulation is the estimation of microdata

Spatial Microsimulation (cont.)

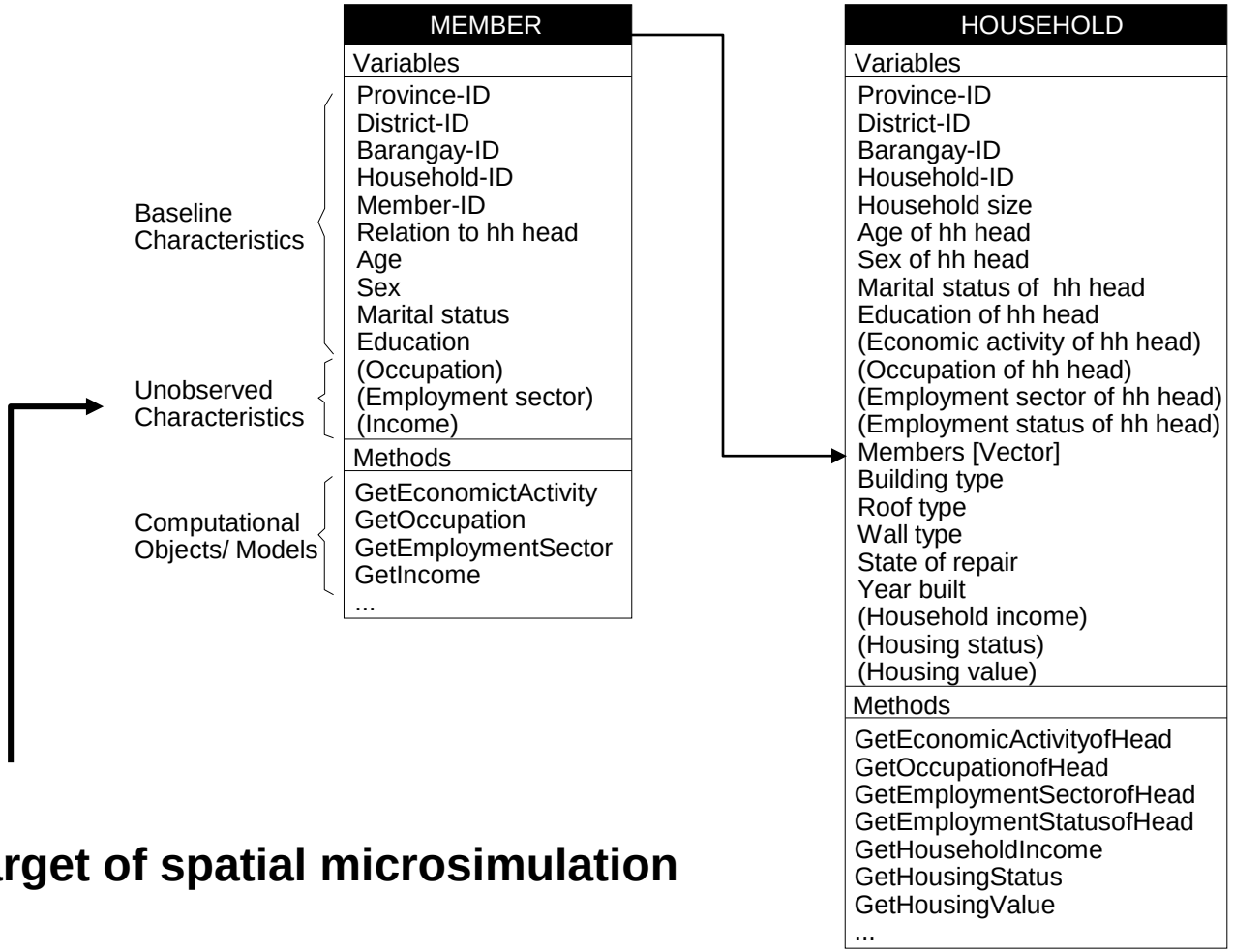
- Spatial microsimulation is increasingly applied in the quantitative analysis of economic and social policy problems (Clarke, 1996)
  - Tax benefit incidence
  - Income
  - Housing
  - Water consumption
  - Transportation
  - Health

Example of spatial microsimulation process



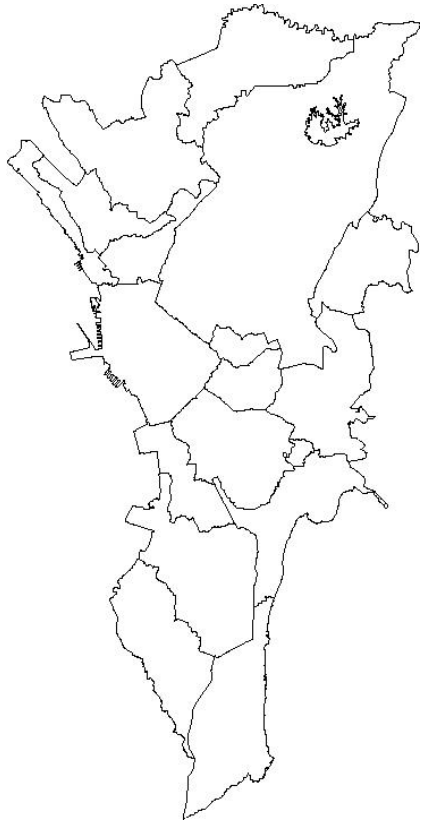
Source: Clarke (1996)

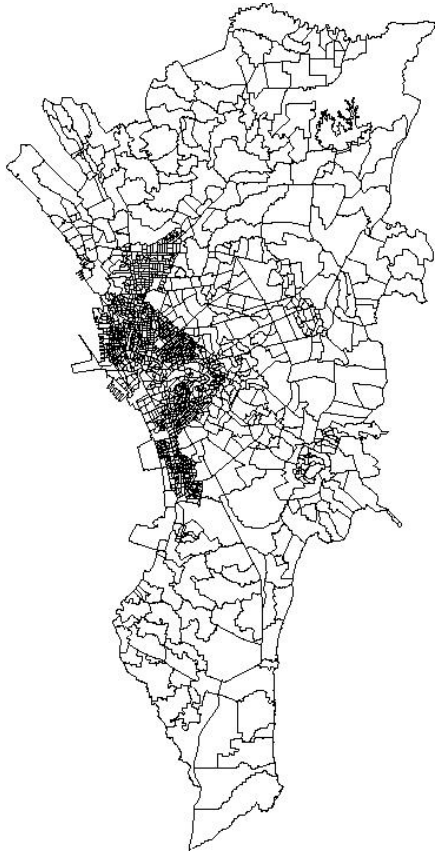
Object representation of household microdata



Target of spatial microsimulation

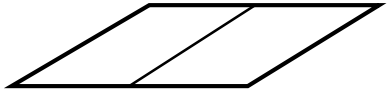
Disparate data sets, Different zoning systems



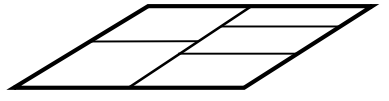


City

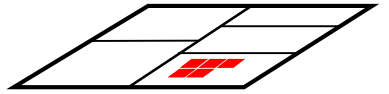
Traffic Zone



Barangay



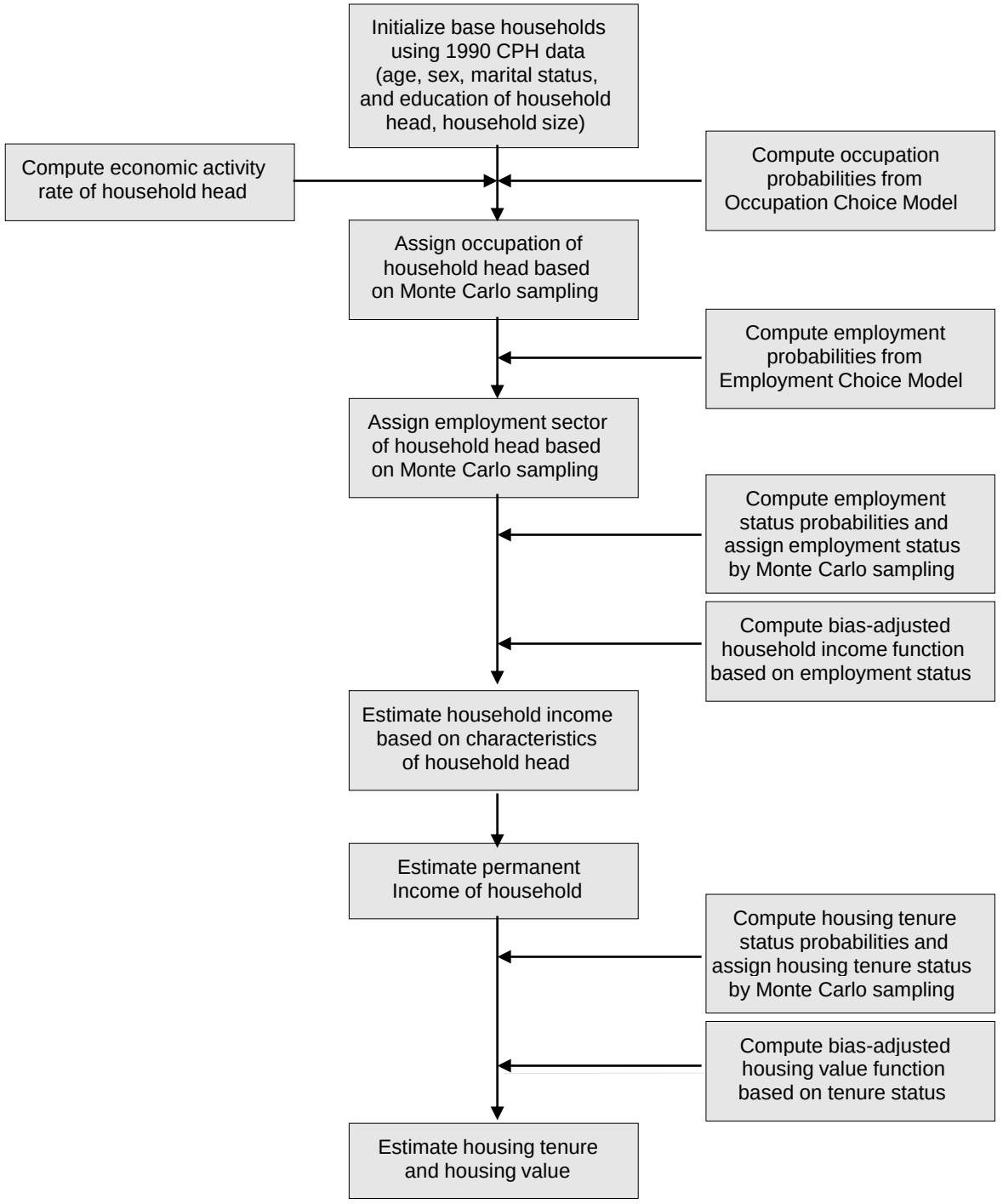
Households



Available data sets

Zone System	Data Set	Description/ Coverage
City	1997 Family Income and Expenditure Survey (FIES)	• Household demographics, some housing variables • Detailed household incomes and expenditures • 4,030 samples for Metro Manila
Traffic Zone	1996 Metro Manila Urban Transportation Integration Study (MMUTIS)	• Selected household demographics • Member/ household income • 50,000 samples for Metro Manila
Barangay	1990 Census of Population and Housing (CPH)	• Detailed household and housing characteristics • No income/employment variable • Non-response on housing variables • All households in 1990 (1,567,665 households)
GIS	1996 MMUTIS Land Use GIS 1997 Building Footprint Data	• Urban land use zoning map for entire Metro Manila • Building footprints for most cities

Spatial microsimulation for estimating household characteristics



Source: Tiglao (2002)

Spatial Microsimulation Modules

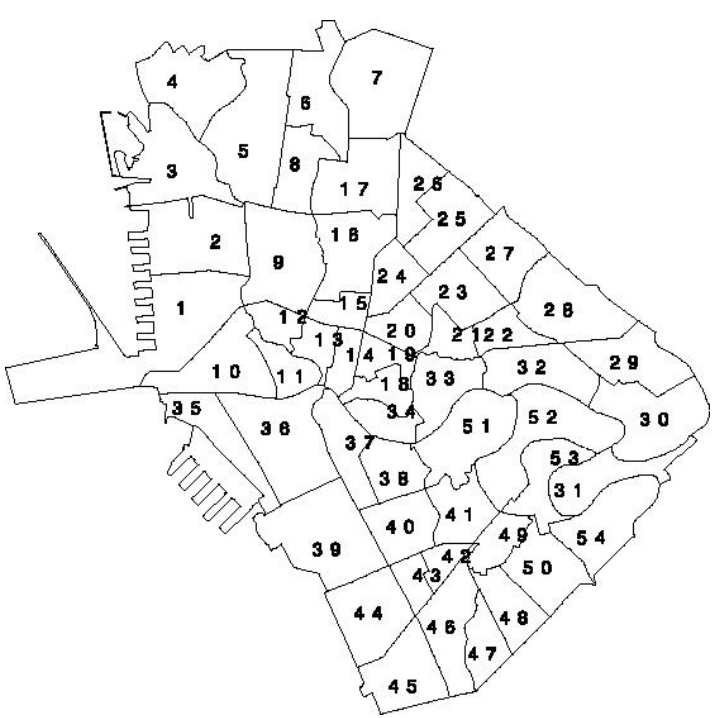
- **Economic Activity Module**. Economic activity rates are computed as conditional probability of an individual being economically active given age, sex and location using the 1996 MMUTIS data. The estimated rates for each zone are applied to all households in the barangays that are located with each particular zone. The assignment of whether a particular household head or member is economically active or not is determined using Monte Carlo sampling. The process involves drawings of random numbers and comparing it with the conditional probabilities.
- **Occupational Choice Module**. The occupational choice is formulated as a multinomial logit model with the actual occupation type as observed choices of the household head. The model includes education level and age (proxy for experience) as explanatory variables. Separate models are estimated for male and female household heads. There are seven occupation groups, namely: Professional, Administrative, Clerical, Sales, Services, Agriculture, and Production.
- **Employment Sector Choice Module**. The employment sector choice is also formulated as a multinomial logit model. It includes education level and age as explanatory variables for observed employment sector. Similarly, separate models are estimated for male and female household heads. There are six employment sectors, namely: Agriculture, Manufacturing, Wholesale & Retail, Transportation, Financing, and Community Services.
- **Employment Status Module**. The employment status model determines whether the household head works in the formal or informal sector. The household head works in the formal sector when he/she is employed by a firm, whether government or in the private sector. On the other hand, a household head who is self-employed or works for another household is also considered to be in the informal sector. Employment status is formulated as a probit model with the following explanatory variables: sex, age, age squared, marital status, education level, household size, occupation type, and employment sector. This model provides a reduced-form probit equation in a three-stage model of household income with selectivity on employment status.

Spatial Microsimulation Modules

- **Household Income Module**. The household income model estimates the household income for a household head and taking into account the employment status of that particular head. Rather than simply calibrating regression models by ordinary least squares (OLS), the household income models incorporate bias corrections for selectivity. Separate household income functions were estimated for the formal and informal sector. Moreover, correction terms were found to be statistically different from zero.
- **Permanent Income Module**. The permanent income model estimates the permanent income of the household given human and non-human wealth characteristics of the household. The explanatory variables include age, age squared, education level, education level squared, household type, and household income.
- **Housing Tenure Module**. The housing tenure sub-model determines whether the household belongs to the formal or informal housing. Formal tenure consists of owners, renters, and those who own land while informal tenure refers to those who may own house but does not own the land. Housing tenure status is formulated as a probit model with the following explanatory variables: education level of household head, household size, and permanent income. This model provides a reduced-form probit equation in a three-stage model of housing value with selectivity on housing tenure status.
- **Housing Value Module**. The housing value model estimates the imputed value of housing for each household which incorporates bias corrections for selectivity on housing tenure status. Separate housing value functions were estimated for the formal and informal housing tenures. Correction terms were found to be statistically different from zero.
- **Inequality Measures Module**. The module takes the full array of incomes in the household microdata and generates three measures of inequality, namely: Gini coefficient, Theil index, and Coefficient of Variation (CV). It is possible to incorporate other measures on inequality based on human capital.
- **Mapping and Visualization Module**. The mapping and visualization module provides the graphical interface for the internal data in the modeling system.

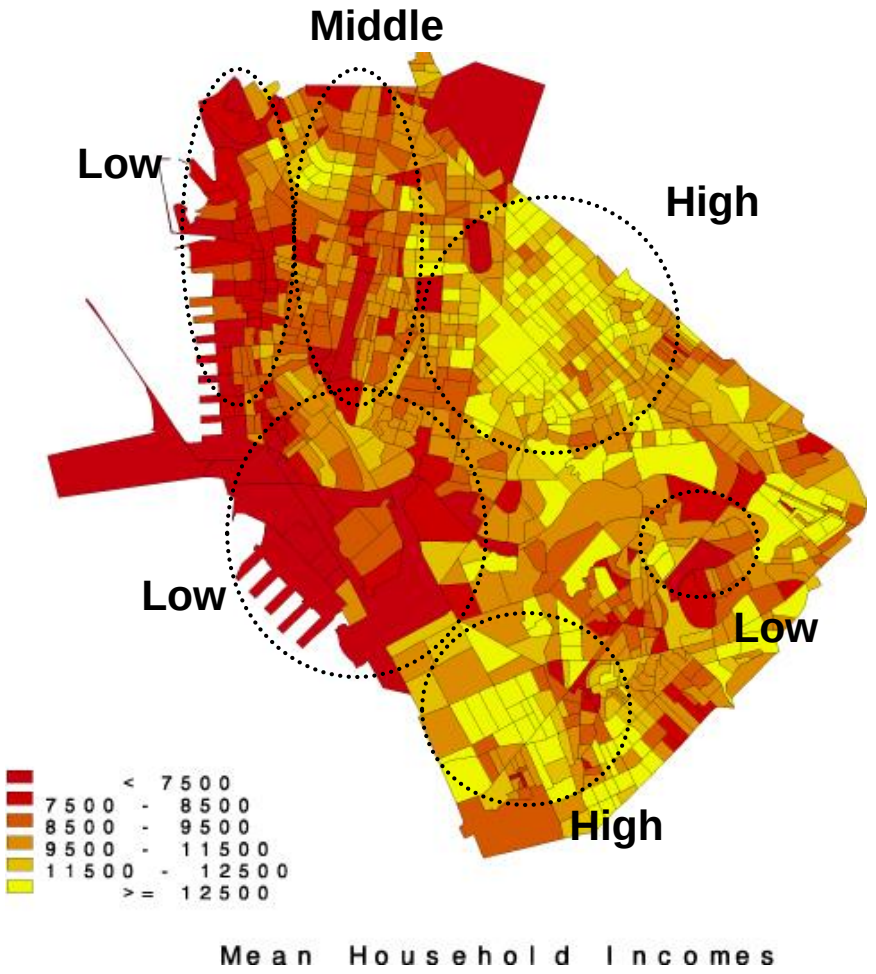
Spatial microsimulation of informal households in Manila City

- Manila City (54 traffic zones, 900 barangays, 1.59 million pop. in 1990, 308,874 households)

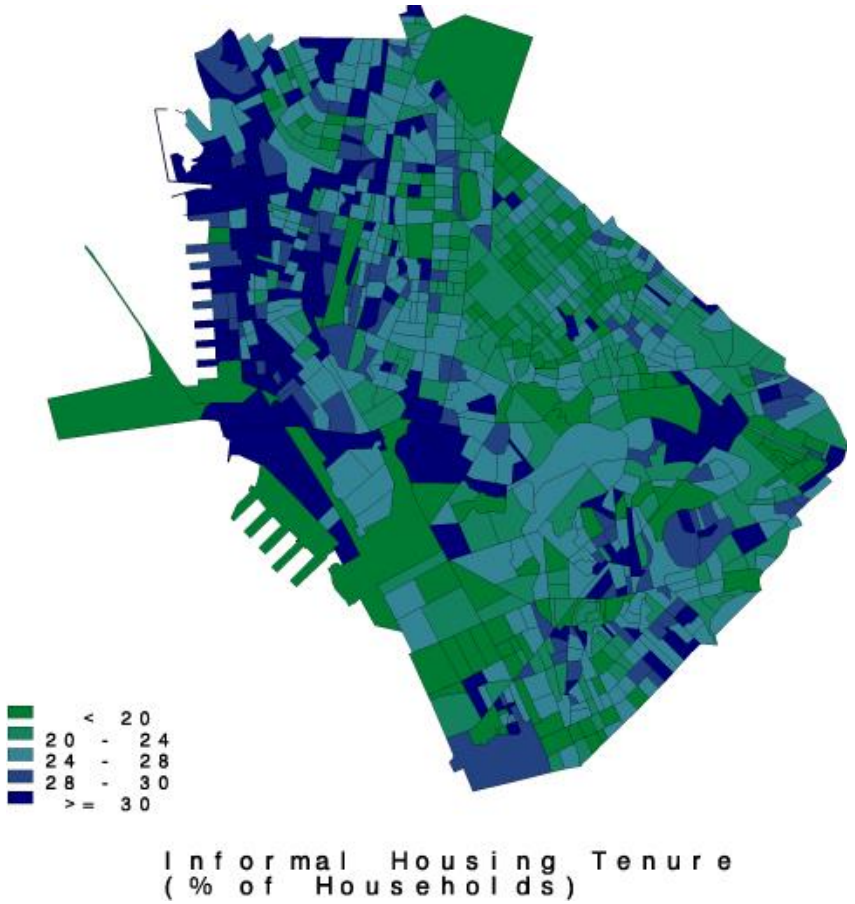


Source: Tiglao (2002)

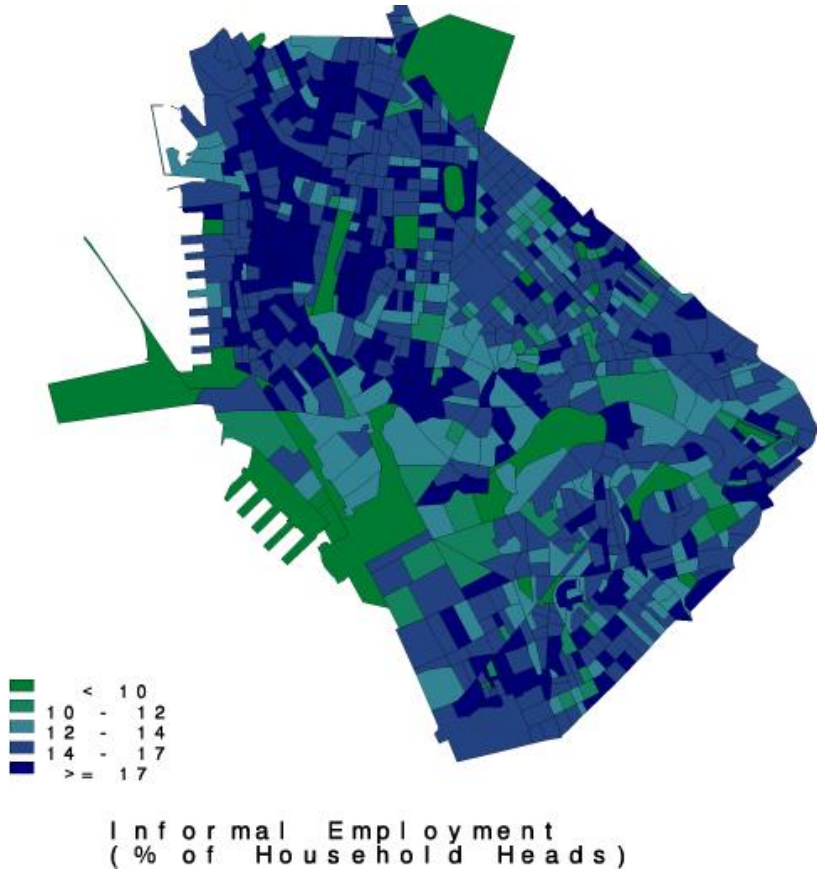
Simulated mean household incomes



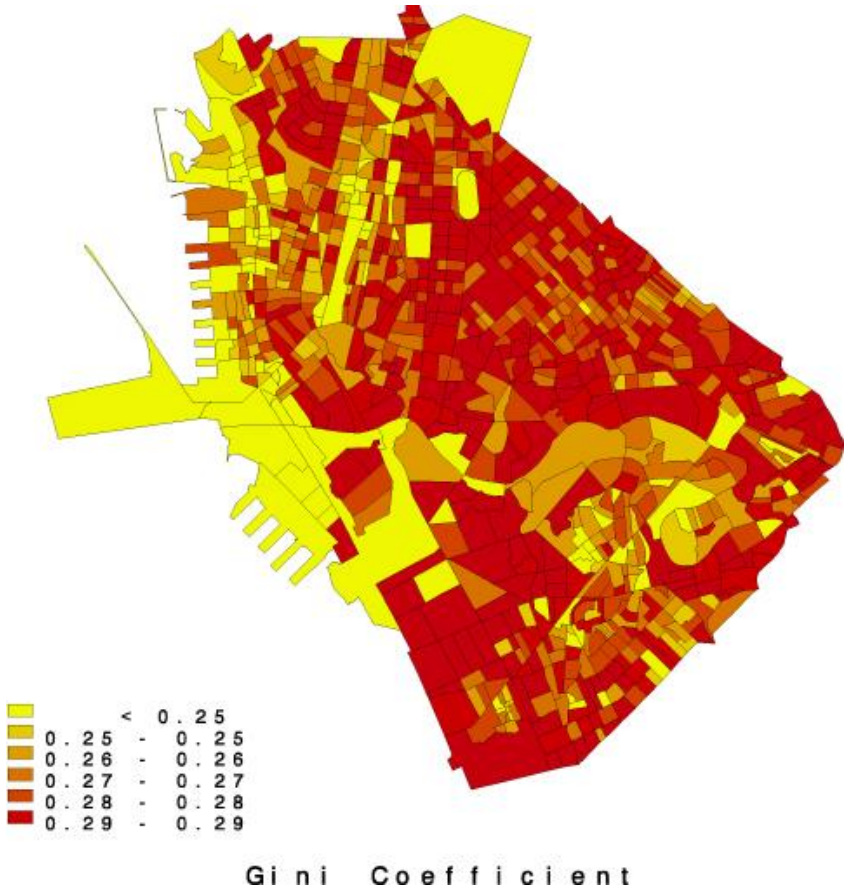
Simulated housing tenure



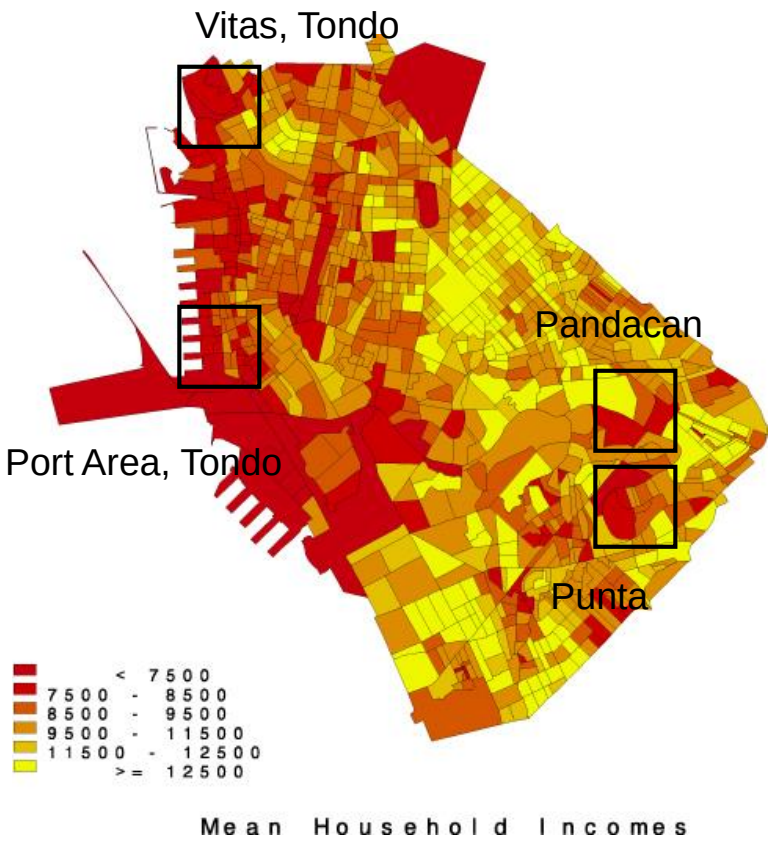
Simulated informal employment



Inequality measures



Ground truths



Modeling Location Choice

Location Choice Theory

- Lancaster's (1966) theory of consumer behavior, views housing as a bundle of services, and households as utility maximizing consumers based on some function of these underlying attributes of housing, including locational characteristics
- Rosen (1974) developed the hedonic theory of housing markets, in which households choose housing so as to maximize a utility function subject to a budget constraint
- McFadden (1978), among others, on the use of random utility theory to develop multinomial logit models of residential location opened a significant direction for research in this area. This body of work was applied to assess and highlight the importance of accessibility and travel mode on residential location

Bid Rent Theory

- Ellickson (1981) that develops a logit model of the property auction process using the bid rent function rather than the utility function. Essentially, this approach focused on the landowner's problem of selling to the highest bidder, which is the consumer making the highest bid.
- It differs from the majority of logit models of residential choice, which focus on the consumer's problem of choosing among properties based on maximizing their utility function. Essentially, this approach represents the two sides of the auction: the buyer's perspective and the seller
- Martinez (1992) extended Ellickson's work by developing a 'bid-choice' model that dealt with both sides of the auction simultaneously, through a nested logit formulation in which the higher level of the model represented the consumer's choice among properties, and the lower level represented the landowner's choice among bidders

Bid Rent Theory

- Martinez (1992) derives a multinomial logit model predicting the probability that a consumer h will choose lot i:

$$P_{i|h} = \frac{e^{\mu(\Theta_{hi} - p_i)}}{\sum_j e^{\mu(\Theta_{hj} - p_j)}} \quad (1)$$

where:

Θ_{hi} is the willingness of consumer h to pay for lot i; and

p_i is the market price of lot i

- The probability of choosing alternative i then is a function of the relative consumer surplus of the alternative:

$$CS_{hi} = \Theta_{hi} - p_i \quad (2)$$

Spatial Microsimulation of Residential Location Choice

- Waddell (2003) provides an approach to deal with aggregation of alternatives to the zone since the model does not explicitly deal with elemental housing or lots as the level of choice.
- This is done by including the size of the choice set represented by each of the aggregate choices. Substituting equation (2) into (1) and incorporating a size term yields

$$P_{i|h} = \frac{e^{\mu(CS_{hi} - \ln S_i)}}{\sum_j e^{\mu(CS_{hj} - \ln S_j)}} \quad (3)$$

Spatial Microsimulation of Residential Location Choice

- The first step in the development of the location choice model is the estimation of the bid functions.
- Bid prices are considered to be the successful bids that households make that match the market price for the alternative. The structure of the bid functions takes the following generic form:

$$BP_{hi} = \beta_0 + \sum \beta_j X_j + \sum \beta_k Z_k \qquad (4)$$

where:

BP_{hi} is the bid price of household h on dwelling unit i

X_j are dwelling attributes

Z_k are zone or neighborhood attributes

β are parameters to be estimated

Residential Location Choice Model

Household classification categories

Household income	Household size	Tenure
Under P9,000	Less than 5	Formal
P9,000 – P14,999	5 or more	Informal
P15,000 – P29,999		
P30,000 or more		

Household bid price variables

Variable	Definition
Occpd1, Occpd2, Occpd3, Occpd4, Occpd5, Occpd5	Dummy variable for occupation type of the household head: Professional (Occpd1=0), Administrative (Occpd1=1), Clerical (Occpd2=1), Sales (Occpd3=1), Services (Occpd4=0), Agriculture (Occpd5=0), Production (Occpd6=1),
Flrarea29, Flrarea30, Flrarea50	Percent of dwelling units with area less than or equal to 29 sq. m, 30 sq. m to less than 50 sq. m and 50 sq. m or more, respectively
Yrbuilt80, Yrbuilt81, Yrbuilt86	Percent of dwelling units that are built in 1980 and earlier, between 1981 and 1985, and after 1986, respectively
Rooftype	Percent of dwelling units with durable roof quality
Walltype	Percent of dwelling units with durable wall quality
Repair	Percent of dwelling units not needing repair
Lowinc, Midinc, Highinc	Percent of households with low income (less than P9,000), middle income (between P9,000 and P14,999), and high income (more than P15,000)
Formal	Percent of households with formal tenure
Single, Duplex, Multi	Percent of dwelling units under single, duplex and multi-unit types, respectively
Landval	Average land value
Access, Distmkti, Timemkti	Accessibility measure, Distance and travel time to the Makati CBD area
Density	Population density of the zone
Res, Educ, Ind, Comm	Percent of land classified as residential, educational, industrial, and commercial, respectively

Residential location choice model estimation results

Variable	Formal households	Informal households
Consumer Surplus	0.31685	0.56290
	(1.690)	(2.024)
Nunits	-0.31719	-0.30316
	(-1.383)	(-1.343)
Log-Likelihood	-8956.3508	-11510.0356

Modeling Employment Location Choice

- The employment location choice model may be specified as a multinomial logit model that includes the accessibility variables (e.g. access to population areas, distance or travel time to the CBD), agglomeration variables in the sense that similar employment tend to cluster in a zone (e.g. percent of employment or occupation type), and land use characteristics.
- In order to estimate the model, a new microsimulation module needs to be developed.
  - The aim of the module is to assign the workplace zone for each of the household heads in the microdata.
  - Once the workplace zones for each household head has been assigned, a sampling-of-alternatives approach can be done in order to generate a set of alternatives for which the employment location choice model can be estimated.

Modeling Employment Location Choice

- The process will be as follows:
  1) Generate conditional probability of a household head in each zone having a particular workplace zone given its age, sex, marital status, education level, economic activity rate, occupation type, employment type, and employment status (whether formal or informal);
  2) Assign the workplace zone for each household head using monte carlo sampling;
  3) Stratify the household heads according to employment status (formal and informal);
  4) Generate alternative workplace zoning by random sampling; and
  5) Estimate the employment location choice model for household heads in the formal and informal sector, respectively.

Modeling Car Ownership

- There are many factors affecting Car Ownership. Generally speaking, these factors can be divided into three main types:
  1) Household characteristics (e.g. income, number of members, age, etc.)
  2) Relative location of the household
  3) Cost and Service level (e.g. purchase price, repairs, fuel costs, etc.)

- Rubite and Tiglao (2003) presents a model for predicting the probability of the household's choice of owning a car or not as a function of the household's and house head's characteristics

Car Ownership Model

Car ownership model parameters

Variable	Parameter Estimate	Standard Error	*t*-statistic	Prob>X^2
Constant	-2.617	0.1095	-12.494	0.000
HHinc	0.1366	0.0409	3.268	0.000
Nwork	0.3876	0.717	5.331	0.000
EDSA-in	-0.2827	0.1458	-1.823	0.034

- The probability of a household to own a car can be expressed as:

$$Prob(own) = \frac{e^{logit(P)}}{1 + e^{logit(P)}}$$

where:

$logit(P)$ = −2.617 + 0.1366·HHinc + 0.3876·Nwork − (0.2827·EDSA-in)

Concluding Remarks

- Spatial Microsimulation methods can overcome data problems in ‘data-poor’environments
- Resulting microdata enables analyst to make full use of existing but disparate data sets and produce reliable and spatially-disaggregate information
- Further research work should be pursued in developing the methods as practical tools for improving integrated land use and transport policy