HISTORICAL VALIDATION OF AN INTEGRATED TRANSPORT – LAND USE MODEL SYSTEM

Eric J. Miller, Director, Cities Centre, University of Toronto, miller@ecf.utoronto.ca

Bilal Farooq, Ph.D. Candidate, Dept. of Civil Engineering, University of Toronto, bilal.farooq@gmail.com

Franco Chingcuanco, Division of Engineering Science, University of Toronto, franco.chingcuanco@utoronto.ca

David Wang, Division of Engineering Science, University of Toronto, David Wang dz.wang@utoronto.ca

ABSTRACT

The ILUTE (Integrated Land Use, Transportation, Environment) model system is an agent-based microsimulation model for the Greater Toronto-Hamilton Area (GTHA) in which disaggregate, process-based models of spatial socio-economic processes are used to evolve the GTHA system state from a known base case to a predicted future year end state in one-year time steps. ILUTE has reached a state of operational implementation in which historical validation runs are being undertaken. A 100% GTA population of persons, families, households and dwelling units has been synthesized for the year 1986. Twenty-year historical simulations (1986-2006) have been run, with model outputs being compared to Canadian Census data and Transportation Tomorrow Survey (TTS) data for 1991, 1996, 2001 and 2006. This paper presents recent findings from these historical validation tests, with particular emphasis on the system's modelling of demographic evolution of the population and of the region's housing market.

Keywords: integrated modelling, land use, microsimulation, validation

INTRODUCTION

The ILUTE (Integrated Land Use, Transportation, Environment) model system is an agent-based, microsimulation model that dynamically evolves urban spatial form, demographics, travel behaviour and environmental impacts over time for the Greater Toronto-Hamilton Area (GTHA). ILUTE has been under development for quite some time (Miller and Salvini, 1998, 2001; Miller, *et al*., 2004; Salvini and Miller, 2005; Miller, 2008). Currently, it has reached the point where it is being tested within a twenty-year historical (1986-2006) time period. The primary purposes of this paper are to:

- Provide an update on the current state of ILUTE, with an emphasis on its dynamic properties.
- Present our most recent historical test results.

Section 2 of the paper provides a high-level description of ILUTE, along with references to more detailed documentation of the model system. Following from this general description of ILUTE, two key aspects of ILUTE design of specific interest within this paper, system dynamics and market processes, are discussed in further detail in Sections 3 and 4. The primary focus of this paper is on two key components of the model system that currently are undergoing extensive testing: demographics and the housing market. These two components and their current test results are discussed in some detail in Sections 5 and 6 of the paper, respectively. Section 7, then provides a brief summary of the paper and a discussion of on-going and future work.

OVEVIEW DESCRIPTION OF ILUTE

ILUTE is a comprehensive, integrated modelling system designed to project the evolution of demographics, land use and travel within an urban region over time. It is an object- and agent-based, microsimulation modelling system. ILUTE is a classical time-driven simulation model in which the system state is evolved from a known initial or base case to some future end state one time-step at a time. Given that ILUTE is a microsimulation model, the system state is defined in terms of the individual persons, households, dwelling units, firms, etc. that collectively define the urban region being modelled. That is, the evolutionary engine operates upon lists of persons (and their attributes), households (and their attributes), and so on, simulating the behaviour of each of these agents (changes in residential location, labour force participation, activity/travel, etc.) over time. Figure 1 summarizes key elements of the current implementation of the model system.

As shown in this figure, key processes modelled within ILUTE include:

- The model system is initialized with a set of agents/objects which are synthesized from base year Census (and perhaps other) data. A 100% population of persons, families, households and dwelling units for each census tract in the study area has been constructed for 1986 using a modified iterative proportional fitting (IPF) procedure (Pritchard and Miller, 2008, 2009) that:
    - Simultaneously generates these four objects in a fully consistent manner.
    - Permits a large number of object attributes to be included in the synthesis.
    - Is computationally efficient.
    - Makes full use of multiple multivariate tables of observed data.
    - Is extendable to include additional elements (e.g., household auto ownership, which is not yet included in the synthesis procedure).

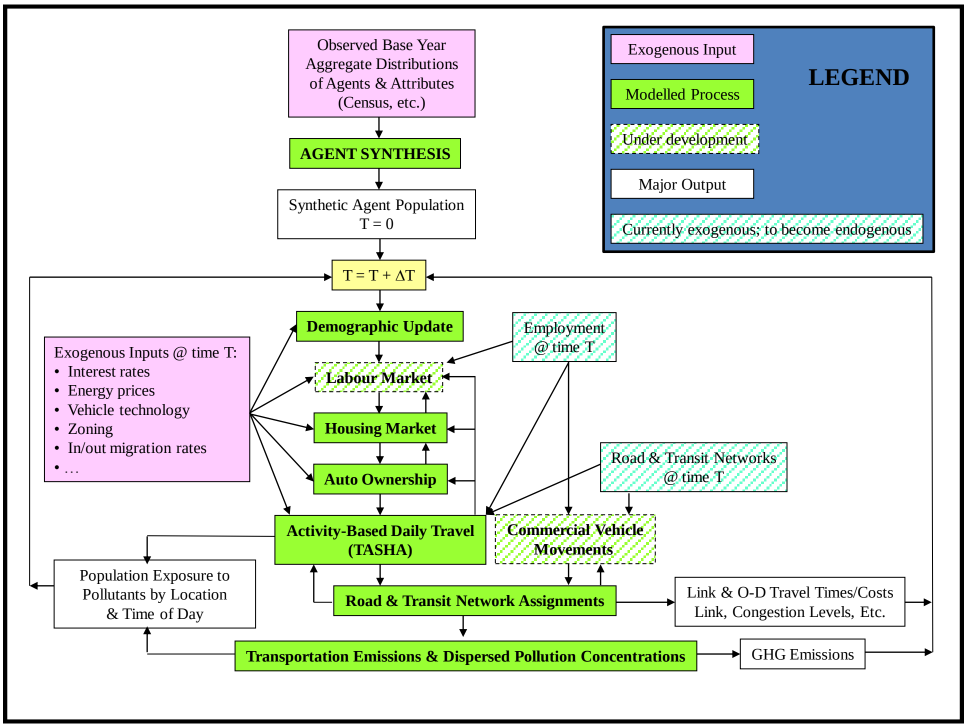


Figure 1 – High-level flowchart of ILUTE processes

For model testing purposes, either the full 100% population can be used, or a smaller subset, randomly drawn from the full population, can be used to speed up run times, with all other model elements and processes (building supply, etc.) being appropriately scaled.

- The current implementation models all processes using a standard one-year time step. The ILUTE software, however, permits individual processes to occur at finer time steps, down to a one month resolution level, if so desired.
- Resident population demographics are updated each time step. This includes dealing with in- and out-migration processes, which are very significant in the current application, given that the GTA has been growing (and is projected to continue to grow) by approximately 100,000 people per year. The linking of marriage partners together within the "marriage market" is also of particular concern. The demographic model is discussed in detail in Section 5.
- The labour market component evolves the labour force over time in terms of:
  - Entry and exit of persons to/from the labour market over time.
  - Mobility of workers within the labour market from one job to another.
  - Allocation of workers actively seeking employment to currently available jobs in the market.
  - The determination of worker wages/salaries by occupation, industry and location over time.
- The housing market component similarly evolves the residential location of households over time. It includes the endogenous supply of housing by type and location, as well as the endogenous determination of sales prices and rents. The housing market model is discussed in detail in Section 6.
- Household auto ownership is dynamically evolved using the models of household vehicle transactions and vehicle type/vintage choice developed by Mohammadian and Miller (2002a,b; 2003a,b).

- Once household demographics, labour market characteristics, residential location and auto ownership levels have been determined, the activity/travel patterns for each person within each household for a typical weekday are estimated using the agent-based microsimulation model TASHA (Travel/Activity Scheduler for Household Agents) developed by the ILUTE team. For full documentation of TASHA see Miller and Roorda (2003); Roorda and Miller (2006), Roorda, *et al*. (2006), Roorda, Miller and Habib (2008), Miller, *et al.*, (2006, 2008).
- TASHA is designed so that it readily interfaces with a variety of network assignment models. Currently it can be used with either EMME (either EMME/2 or EMME/3) for road and transit assignments or MATSim for road assignments (Gao, *et al*., 2010). Advantages of using a microsimulation model such as MATSim relative to an aggregate model such as EMME include:
    - The ability to retain an agent-based representation throughout the modelling process; i.e., the identity of individual agents and their travel behaviour is retained in MATSim, but is lost in aggregate, zone-based assignment procedures such as EMME, which require aggregation of individual agent trips into origin-destination flow tables.
    - MATSim (and other dynamic assignment procedures) deal explicitly with network dynamics and provide an explicit temporal representation, in contrast to the static representation in conventional deterministic user equilibrium methods such as EMME.
    - For full 24-hour network modelling, we have found that MATSim is at least as computationally efficient as running 24 one-hour static assignments in EMME, while providing much higher level behavioural fidelity and enhanced representation of network performance (Gao, *et al*., 2010).
- Considerable work in recent yours has focussed on developing an environmental modelling component within ILUTE (Hatzopoulou, *et al*, 2007; Hatzopoulou and Miller, 2008b; Miller, *et al*., 2009; Hao, *et al*., 2010). As illustrated in Figure 1 this involves:
    - Modelling both link-based and zone-based vehicle emissions. The dynamic, disaggregate nature of both TASHA and MATSim permit both the running and stationary emissions of each vehicle to be dynamically computed.
    - These emissions can then be fed into an atmospheric dispersion model (in our case, CALPUFF) so that pollutant concentrations over time and space can be computed.
    - At the same time, TASHA generates a dynamic population of where each person is over time and space.
    - Putting people and concentrations together at each location in each time period results in being able to compute persons' exposures to pollutants over time.
- It is the intent to implement some form of firmographic model within ILUTE, perhaps in the spirit of Maoh (2005) or Moeckel (2005). This has not yet been accomplished, and so for current historical model system testing purposes, observed employment levels by occupation and industry for each census track in the study area are exogenous inputs to the simulation.

- Similarly, a much longer-term project is to implement a microsimulation-based commercial vehicle movements (CVM) model to ILUTE (Roorda, Cavalcante, McCabe and Kwan, 2008). Currently, in the absence of such a model, typical GTHA four-step modelling practice of reducing major roadway capacities to "account for" trucks is followed, albeit with the recognition that this is a very crude approximation of the effects of CVM within the road system.
- Finally, another long-term project within ILUTE is to develop the capability to endogenously evolve at least the "routine" components of the road and transit network (e.g., streets, bus routes) over time in response to land use development, so that changes in these important components of the network do not need to be anticipated by the modeller and pre-defined as exogenous inputs to a given simulation run. "Major", policy-based network components such as new freeways, subways, etc. probably can and should remain as scenario-based inputs to the model. But endogenous evolution of the "base supply" of transportation services ideally should be part of integrated urban model systems, in the same way that endogenous evolution of housing and commercial floorspace supply currently is.

From the brief comments above, it is clear that two key elements of ILUTE are its dynamic elements and its focus on market-based processes. These two elements of ILUTE are discussed further in this paper in the following two sections. Much more extensive documentation of the ILUTE system as a whole can be found in a 10-volume series of technical reports that can be downloaded from: http://www.ecf.utoronto.ca/~miller/ILUTE.zip.

ILUTE DYNAMICS

As has already been briefly discussed, ILUTE is a time-driven microsimulation model in which several time units are used in the current implementation. These time units are:

- All demographic and spatial processes are modelled in one-year time steps. ILUTE is capable of modelling using arbitrary time steps that can vary from one process to another, down to the resolution of a single month. For present purposes, it was felt that a one-year time step was sufficient, although it is recognized that this introduces a certain amount of temporal aggregation that might be significant in some cases. As is discussed in Section 4, this one-year time step is implicitly relaxed in the case of ILUTE market processes.
- Activity/travel/environmental processes are modelled for a single, "typical" 24-hour weekday time period for each year in the simulation run. Specific time steps within these components of ILUTE currently consist of the following:
    - TASHA models activity and trip start times and activity durations at the level of 5 minute intervals (i.e., there are 288 5-minute time intervals within TASHA).
    - If EMME is used to model road and transit network flows, a representative hour for each of the morning peak period, the afternoon peak period and off-peak periods is modelled. A procedure for interpolating travel times for other hours of the day is then used to assign travel times to each trip based on the hour of its start time. (Guan, *et al.*, 2003). Alternatively, 24 one-hour road

  assignments can be run. For EMME-based emissions calculations, the 24 one-hour assignment approach is required.
  - If MATSim is used for road assignment, then the entire 24-hour period is simulated in continuous (1-second interval) time. Emissions are computed per vehicle per link during the simulation run.
  - Outputs of pollutant concentrations from the CALPUFF dispersion are collected for each traffic zone in the system at a temporal resolution of each hour within the 24-hour day being modelled.

As is also implicit in Figure 1 and in the discussion of this figure in the previous section, ILUTE is a disequilibrium model. In particular, a basic assumption within ILUTE is that urban areas are open, dissipative, path-dependent systems which are never in equilibrium, but rather are constantly responding to the constantly varying endogenous and exogenous forces that “drive” the evolution of the urban system state and that do not permit that system state to ever stabilize to an equilibrium. Hence, the adoption of a simulation framework, in which future states are explicitly evolved in a path-dependent fashion from (assumed) known base conditions.

Within individual model components both rule-based and utility-maximizing models are used, depending on the process being modelled. Assumptions concerning within-component equilibrium/stabilization/optimization also vary from one process to another. In general, our preference is for “myopic” processes in which individual agents are seeking to maximize their utility within individual decisions (e.g., what mode of travel to take to work) but not “globally” across multiple decisions (e.g., overall optimization of daily activity pattern). Similarly, our preference is for modelling market transactions that leave both individual buyers and sellers “satisfied” with their exchange of a good or service, but which does not involve imposing strict global equilibrium constraints. Assumptions, however, do vary from one model component to another, depending on both behavioural concerns and practical considerations of available modelling methods, computational efficiency, etc. More detailed discussion of this issue with respect to our modeling of markets is presented in the next section. Detailed discussion of these issues with respect to all other components of ILUTE is well beyond the scope of this paper, but a few key points to note in this regard include:

- TASHA assumes that persons myopically adjust their schedules to accommodate as best as possible newly generated activity episodes, but it does not attempt to optimize the overall daily schedule by rescheduling the entire day each time a new episode is added. Random utility maximization, however, is used to select travel modes, the allocation of household vehicles for usage by household, and the determination of ridesharing behaviour among household members so as to stochastically maximize household travel utility (given the utility-based preferences and travel needs of each individual household member).
- Both EMME and MATSim effectively equilibrate flows on the transportation network, given known trip inputs. EMME imposes a static equilibrium assumption; MATSim iteratively “stabilizes” the assigned flow patterns.
- ILUTE explicitly assumes disequilibrium between transportation and land use, in that network travel times from time period t (which are based on travel demand for time

period t) are used to determine travel demand (and influence other spatial processes) in time period (t+1).

SIMULATING SPATIAL MARKETS IN ILUTE

Spatial markets play a central role in ILUTE in that it is through market demand-supply interactions that all spatial processes of interest within ILUTE play out. All markets in ILUTE involve a demand process, a supply process, and a market clearing process that mediates between demand and supply by determining the exchange of goods and services between demanders/buyers/consumers and suppliers/sellers/producers and the prices at which these exchanges occur. All such market processes are modelled at the level of the individual agent (buyer, seller) and individual transactions between buyers and sellers (e.g., dwellings are “auctioned off” one unit at a time within the housing market model.

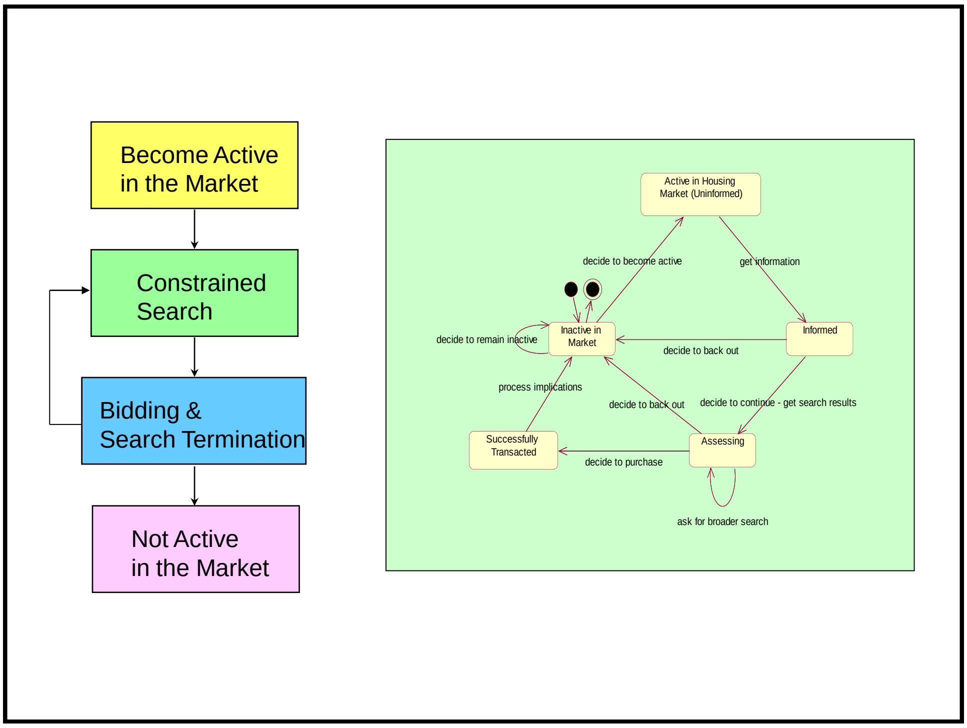


Figure 2 – Representation of market processes in ILUTE

In terms of demand processes, no agent is continuously active in any given market. Households do not daily search for new residential locations; workers do not continuously switch jobs; etc. Rather, market participation is characterized by protracted periods of inactivity (i.e., being in a passive state). At any point in time, however, an agent may decide to become active in the market, in response to potentially a wide variety of “push” and “pull” factors. That is, a decision to become active in a given market (look for a new home or job) is triggered by some combination of factors. Once active in a market, the agent engages in a search process, looking for a new house, job, etc. that represents an improvement over his/her current situation. The agent will remain active in the market until either a satisfactory new alternative is found and successfully obtained (e.g., a successful bid on a new house or success in receiving a job offer), or until a point in time is reached in which the agent decides that an improved alternative relative to the status quo cannot be feasibly obtained within the current market (no better house that he/she can afford, no job that is better / pays more, etc.) and so the agent decides to remain in the *status quo* state. In either case, the agent returns to the passive state, in either a new or the old location, depending on the active state termination criterion. Figure 2 provides two illustrations of this general market participation process, one using a conventional flowchart representation and one in Unified Modelling Language (UML) – a graphical "language" for representing object-oriented entities and processes.

Supply side processes within ILUTE currently are generally modelled in a more aggregate fashion (e.g., the "development industry" produces a certain number of new dwellings by type and location each year of the simulation). The outputs from these processes, however, are a list of new individual dwellings, jobs, etc. (and their associated individual attributes) which are then matched with demanders on a one-to-one basis.

Two generic market clearing processes are currently employed in ILUTE:

- A variable-price process, in which prices are endogenously determined on a transaction-by-transaction basis within each market clearing episode, using an auction-type process.
- A fixed-price process, in which prices are fixed within each market clearing episode and then are "globally updated" between market clearing episodes in response to general demand-supply characteristics. This process also applies to in which prices do not exist but a market-like matching of agents is required.

Thus, prices are endogenous within the market in both processes; what differs is the way in which prices are determined by the market interactions.

The owner-occupied housing market is an important example of the variable-price, auction-based process. It is discussed in greater detail in Section 6 of this paper. Several fixed-price market processes exist within ILUTE. Currently these consist of:

- The rental housing market, within which rents are either fixed by policy (rent control, assisted housing) or through an aggregate market adjustment process.
- The labour market, in which it is assumed that workers are in the short-run salary-takers, with salaries adjusting over time in response to aggregate market adjustments.
- The "marriage market" in which single males and females are matched to form married couples. Clearly no "price" exists in this "market", but it otherwise can be modelled in a fashion similar to other fixed-price markets.

The rental housing and labour market models are currently in final development stages and are not discussed further in this paper. The "marriage market" model is described in somewhat greater detail in the next section.

THE ILUTE DEMOGRAPHIC MODEL & SELECTED TEST RESULTS

The demographic updating component of ILUTE is concerned with evolving the person, family and household agents and their associated attributes over time. Families are explicitly maintained within the model system so that family relationships and interactions can be tracked over time and used to help explain family related behaviours. In ILUTE, nuclear families (i.e., families living as a household) are defined as husband-wife couples, with or without children, or single-parents living with one or more children. Extended family relationships, however, are also maintained throughout the simulation, in terms of children knowing where their parents live (and *vice versa*; and hence, grandparents knowing where their children live) and divorced people knowing where their ex-spouses live. Thus, family social networks are endogenously evolved over time.

A household is, as is usual in integrated models, defined as one or more persons living within the same dwelling unit. Hence, there is a one-to-one mapping between households and occupied dwelling units. A household can consist of any of the following combinations of persons and families:

- Single person.
- Multi-person, non-family.
- Single family.
- Multi-family.
- Single family with individual (i.e. non-family) persons.
- Multi-family with individual persons.

The ILUTE demographic component is updated each year. A bottom-up approach is employed in which the demographic changes of a region emerge through the sequential updating of each household. Likewise, the families and/or individuals within each household are updated consecutively. For each person in a household (whether part of a family or an individual unit), a series of demographic events are evaluated. Afterwards, these events are processed in order to reflect the changes or manage their effects. Finally, in-migration events are evaluated which introduce new persons, families and households into the ILUTE study region. Out-migration is also handled as part of the demographic processes.

Figure 3 outlines how the agents' attributes are updated each year in the simulation. The right column shows the sequence in which demographic events are evaluated for each person in a household. The left column of Figure 3 depicts the input variables that are currently used in these calculations. These variables can be *generic* (e.g. age) that are used to evaluate a multitude of events (e.g. death), or *specific* variables that are only applicable to certain events. For example, the length of a marriage is the main input in divorce decisions. While the previous two types of inputs derive directly from the agent being updated, *zonal* variables are also used, which are determined from the agent's area of residence. These aggregate inputs are employed to reflect the similarity in behaviour within a spatial region. For example, people in the suburbs tend to have higher driver's license ownership rates than people who live in the central business district. Finally, the current simulation year is used as a *temporal* variable which help captures the dynamically changing nature of demographic attributes.

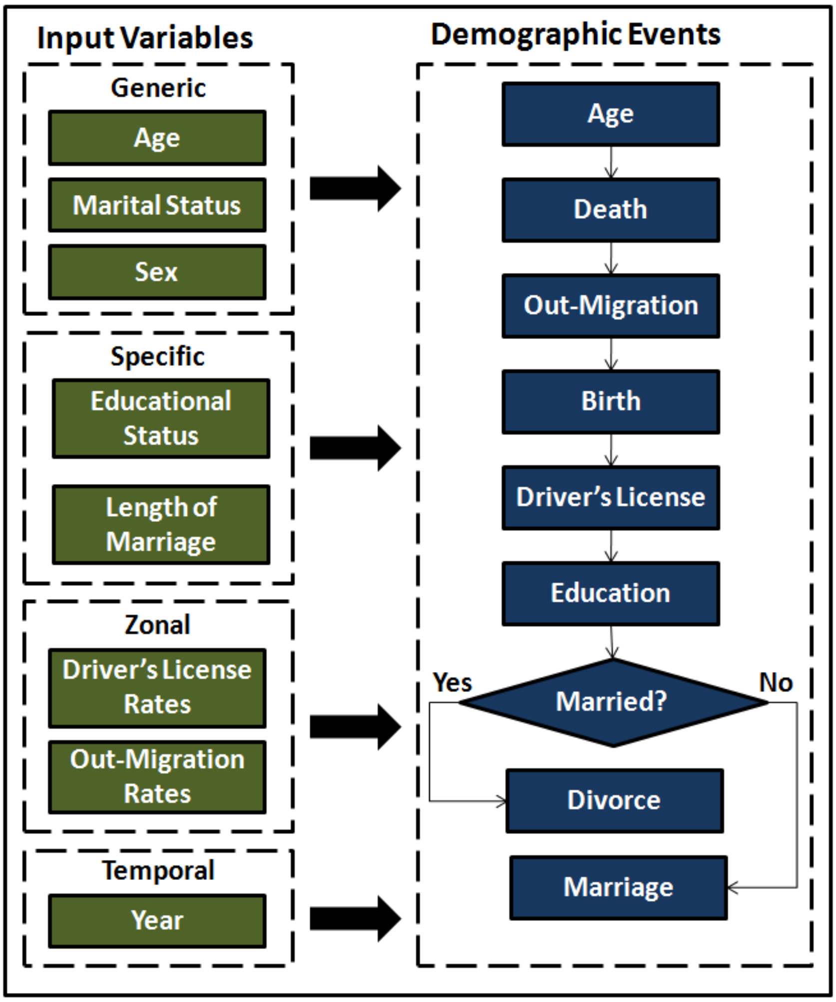


Figure 3 – Demographic updating of person attributes

With the exception of aging, each demographic event uses simple transition probabilities, conditioned on a person's current state, to determine a change in a demographic attribute. Cross-tabulations of the input variables in Table 1, derived from empirical data, were used to compute these conditional probabilities. Monte Carlo simulation is employed to determine the outcomes of all demographic events. Further details concerning demographic updating procedures can be found in Miller, Chingcuanco, Farooq, Habib and Habib (2008) and Chingcuanco (2010).

Figure 4 compares the GTHA age distributions by year as predicted with a 100,000 initial household run (hereafter referred to as a 100K run) of ILUTE, compared to Census data. Note that the 1986 distribution represents the synthesized input distribution used to initialize the ILUTE run. In general, the generated distribution tracks the Census distributions well, especially given that the predicted distribution is the net outcome of numerous processes -- births, deaths, in-migration and out-migration – in addition to the aging process *per se.*

Overall predicted and observed births, deaths and out-migrations are shown by year in Figure 5. ILUTE is doing a very good job in predicting births and deaths, but needs to improve its predictions of out-migrations somewhat. The out-migration model is currently hampered by the lack of good local data concerning out-migration rates by household/family type for the region. In-migrations are an exogenous input to the model, and so comparisons between predicted and observed values are of little value to show.

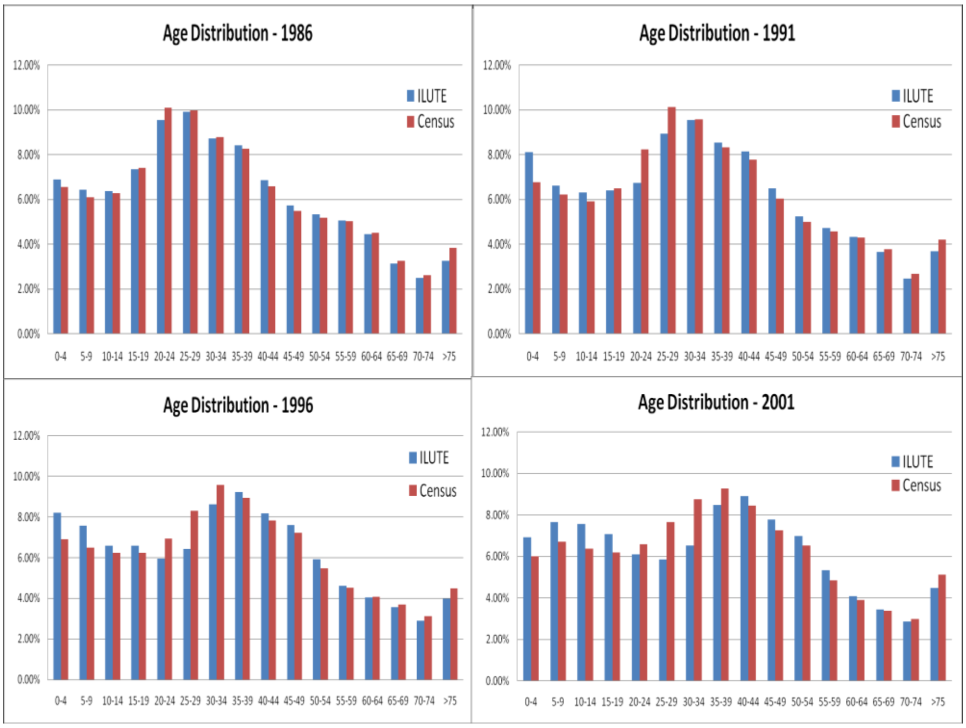


Figure 4 – Predicted & observed GTHA age distributions by year

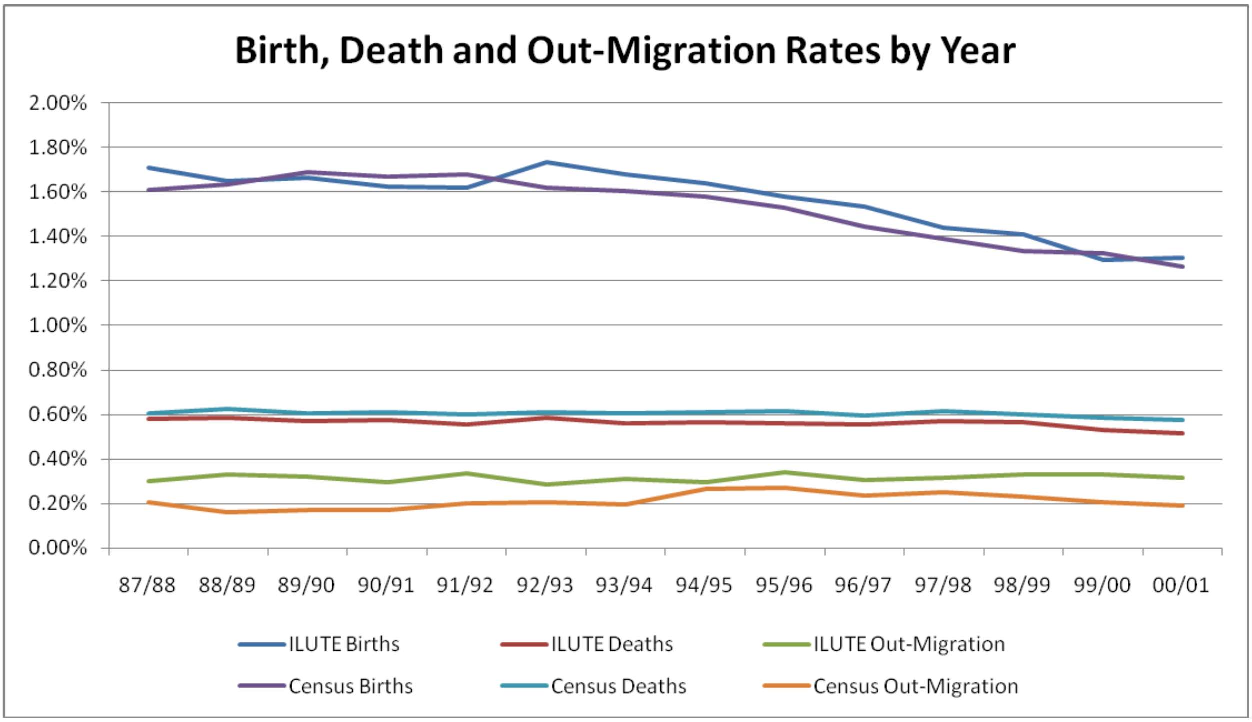


Figure 5 – Predicted & observed births, deaths & out-migrations by year

Marriage Market Model

Of particular interest within this paper is the "marriage market", which matches prospective husbands and wives together within a utility maximization framework, both as an example of an important demographic process and as an example of a fixed-price market process. The problem can be defined as follows. A decision-maker (either a prospective husband or wife) has *N* potential matches within a pre-determined choice set. His/her objective is to find a match that maximizes his/her utility. There are no constraints in this search process. However, constraints (age and spatial distance) do enter during the choice set generation.

In the ILUTE marriage market, potential marriages have an associated utility that is based on the characteristics of the prospective match. These utilities are adapted from (Choo, *et al.*, 2008) and are based on the potential couple's income(s) and the male/female ratios in their respective geographic areas:

- For the male participant, the utility of marrying a potential spouse is:

$U_m = f(combined\ income, sex\ ratio, single\ male\ income, wife's\ income)$.

- For the female participant, the utility of marrying a potential spouse is:

$U_f = f(combined\ income, sex\ ratio, single\ female\ income, husband's\ income).$

The average of the male and female utilities for a given potential pair can be interpreted as the weight of the arc between the two nodes of a graph. The solution to this problem is determined by finding the assignment of each element $m \in Males$ to each $f \in Females$ that maximizes the utility of the entire market (i.e. maximizes the sum of the utilities of all the participants), given the "choice sets" for each male and female. At present, choice set definition follows a simple rule-based approach based on age difference and geographic constraints.

This problem is equivalent to finding the matching that maximizes the sum of the graph's edges, or, in other words, solving a maximum weighted bipartite matching problem. This problem can be expressed as the following mathematical program:

$$MAX \quad \sum_{m}^{M} \sum_{f}^{F} x_{mf} U_{mf}$$

$$s.t. \quad \sum_{f}^{F} x_{mf} \leq 1 \; \forall \, m$$

$$\sum_{m}^{M} x_{mf} \leq 1 \; \forall \, f$$

where U_{mf} is the utility associated with the marriage of *m* and *f*, and x_{mf} is 1 if *m* & *f* are a match, and 0 otherwise. The *Hungarian algorithm* (also known as *Munkres' assignment algorithm*) can be used to solve this combinatorial optimization problem and is currently being investigated. At present, however, a simpler algorithm is used in ILUTE in which randomness is introduced. It is believed that such randomness provides a closer representation to reality. It also relaxes the rather strong assumptions made regarding overall market maximization and individual utility maximization. Equally important, the implemented procedure saves significant computational time. The simple algorithm is as follows:

1. If there are more wives than husbands, pick a wife randomly, else pick a husband.
2. If a wife (husband) is chosen, find a prospective husband (wife) that maximizes their marriage utility, given the wife's (husband's) choice set.
3. If the combined utility of the prospective match is higher than a threshold utility (at present, this is set at 0), marry the new couple.
4. Repeat until either the wife or husband pool is empty.

Table 1 presents results from the most recent ILUTE test run, in which the run was initialized with a 100,000 1986 random sample of GTHA households. The table compares the age distributions of married couples in 2001 (15 years into the simulation) as observed in the Canadian census and as predicted by ILUTE. In general, the predicted results correspond well with observed values.

Table 1 – Observed and predicted 2001 age distributions for married couples

CENSUS 2001 MARRIED COUPLES		**Age Male**							
		18 - 24	**25 - 34**	**35 - 44**	**45 - 54**	**55 - 64**	**65 – 74**	**75 - 84**	**85 & over**
Age Female	**18 - 24**	**0.28%**	**1.00%**	**0.14%**	**0.03%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**
	25 - 34	**0.18%**	**10.94%**	**7.10%**	**0.39%**	**0.06%**	**0.00%**	**0.00%**	**0.00%**
	35 - 44	**0.02%**	**1.57%**	**19.11%**	**7.84%**	**0.55%**	**0.08%**	**0.00%**	**0.00%**
	45 - 54	**0.01%**	**0.08%**	**1.59%**	**15.21%**	**6.19%**	**0.46%**	**0.03%**	**0.00%**
	55 - 64	**0.00%**	**0.01%**	**0.05%**	**0.95%**	**8.58%**	**4.40%**	**0.24%**	**0.02%**
	65 - 74	**0.00%**	**0.00%**	**0.01%**	**0.04%**	**0.51%**	**5.98%**	**2.39%**	**0.08%**
	75 - 84	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.03%**	**0.43%**	**2.56%**	**0.51%**
	85 & over	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.01%**	**0.11%**	**0.24%**

ILUTE 2001 MARRIED COUPLES		**Age Male**							
		18 - 24	**25 - 34**	**35 - 44**	**45 - 54**	**55 - 64**	**65 – 74**	**75 - 84**	**85 & over**
Age Female	**18 - 24**	**1.21%**	**0.71%**	**0.17%**	**0.00%**	**0.00%**	**0.01%**	**0.00%**	**0.00%**
	25 - 34	**0.05%**	**11.40%**	**3.78%**	**1.00%**	**0.03%**	**0.03%**	**0.00%**	**0.00%**
	35 - 44	**0.02%**	**0.97%**	**18.74%**	**8.73%**	**2.38%**	**0.12%**	**0.03%**	**0.00%**
	45 - 54	**0.00%**	**0.40%**	**4.62%**	**12.28%**	**6.32%**	**1.69%**	**0.07%**	**0.00%**
	55 - 64	**0.00%**	**0.01%**	**0.72%**	**3.44%**	**5.96%**	**3.33%**	**0.61%**	**0.02%**
	65 - 74	**0.00%**	**0.01%**	**0.02%**	**0.40%**	**1.83%**	**3.60%**	**1.54%**	**0.23%**
	75 - 84	**0.00%**	**0.00%**	**0.00%**	**0.01%**	**0.10%**	**1.02%**	**1.25%**	**0.47%**
	85 & over	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.01%**	**0.05%**	**0.28%**	**0.32%**

Figure 6 similarly compares the 2001 observed (Census) and predicted (100K ILUTE run) income differences for married couples. In this run, the model tends to over-predict matches with very small income differences; this will require some adjustment of the matching algorithm utility parameters in subsequent runs. Otherwise, however, good correspondence between observed and predicted age differences is obtained.

Figure 6: Observed and Predicted 2001 Income Differences for Married Couples

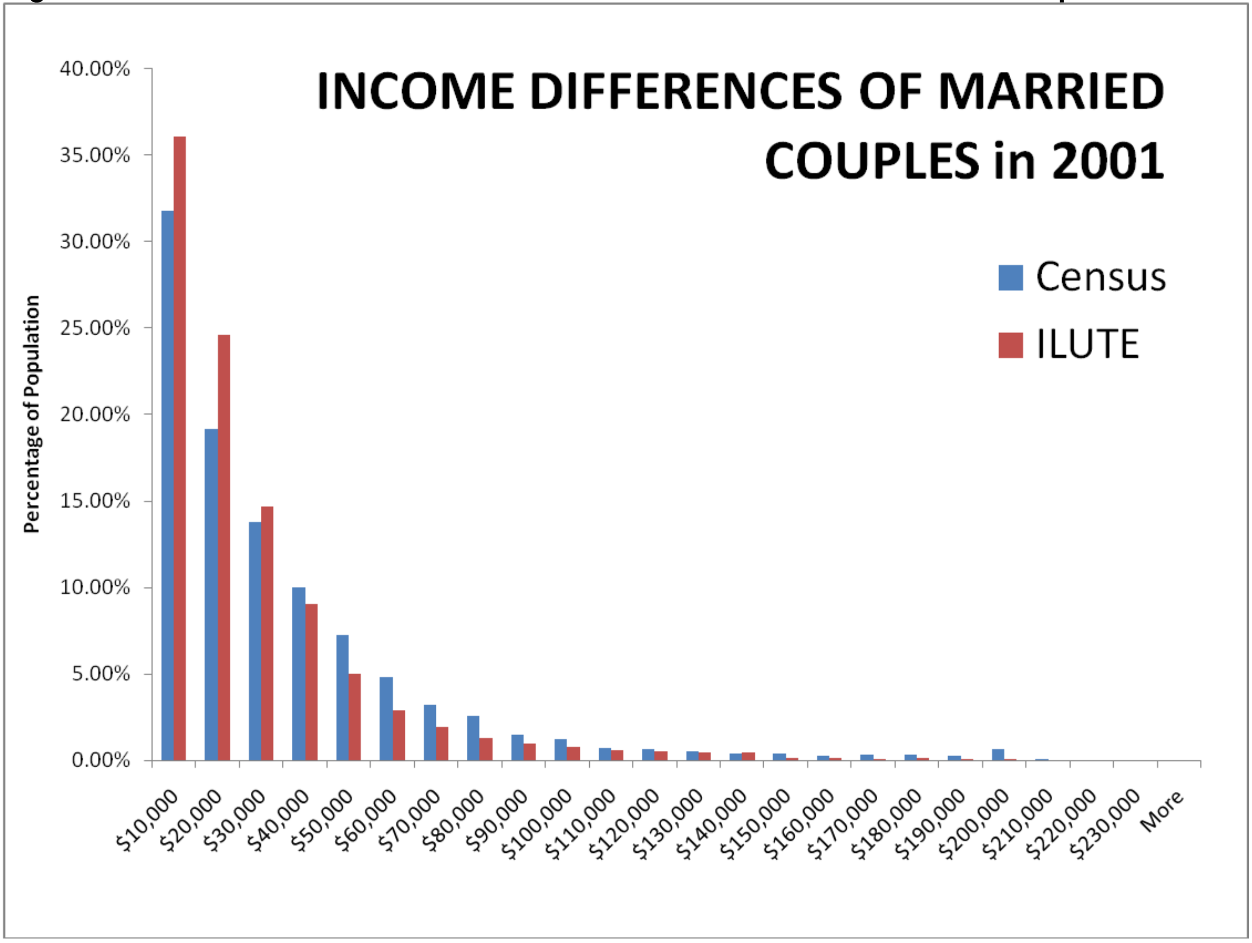


Figure 6 – Observed & predicted 2001 income differences for married couples

THE OWNER-OCCUPIED HOUSING MARKET: MODEL STRUCTURE AND SELECTED TEST RESULTS

Figure 7 elaborates the abstract market process discussed in Section 4 for the case of the owner-occupied housing market. As illustrated in this figure, at any point in time in an urban region there is a set (list) of active demanders/consumers of housing and a set (list) of active dwelling units available to be purchased or rented. Demanders and suppliers come together within the market and exchanges occur when a given supplier agrees to sell/rent a dwelling unit to a given demander at a mutually agreed-upon price. Thus, three types of agents exist within this market process: households, developers/landlords, and a *market agent* that is a "virtual agent" used within the simulation model to keep track of vacancies and prices, manage the lists of active demanders and suppliers, and manage the buying/selling market clearing process. Three key points to note about this process are:

- Prices are endogenously determined through the market clearing process; i.e., through the "bid-auction" (Martinez and Donoso, 2001) process of buyers and sellers interacting and coming to mutually agreeable prices for the exchange of the dwelling unit. These market prices can influence decisions by agents in subsequent time periods to become active in the market.

- The supply of housing comes from two very distinct processes. The first is the construction of new housing by developers/builders. The second (and quantitatively more important at any one instant in time) is the decision of owner-occupants to become active in the housing market. This decision results in the household becoming active both as a demander for a new dwelling unit and as a supplier of its old dwelling unit, which becomes active in the market on the "supply side" at the same time that the household becomes active on the "demand side"
- Until such time as a household sells its current dwelling it has the option of exiting from the housing market and staying in its current dwelling; i.e., it need not sell if it does wish to. Similarly developers of new housing stock do not need to sell in any given time period if they deem it economically unattractive to do so.

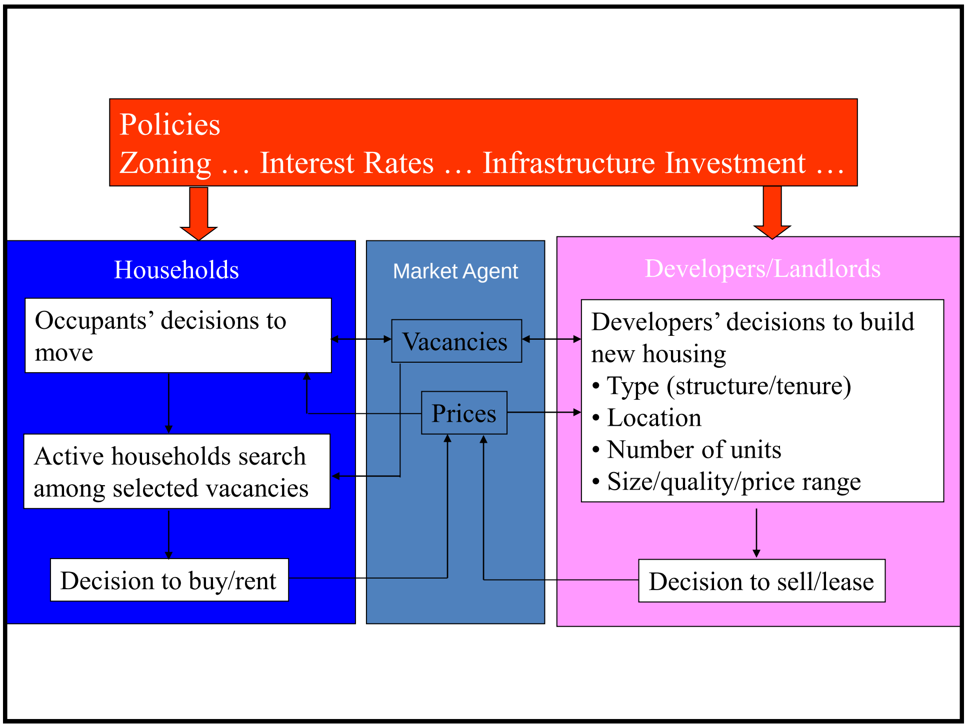


Figure 7 – Simulating owner-occupied housing demand, supply & market-clearling

Following on from this last point, while we speak of "market clearing", there is, in fact, no behavioural need nor mechanism within ILUTE to force the market to completely clear or to "equilibrate" within any one time period. Dwellings can remain vacant from one time period to another, and households can continue to be active (or drop out without transacting) from one time period to another as they deem fit. It is expected that in times of excess supply prices will tend to fall (or at least not increase) and rates of new supply provisions will tend to fall; while in times of excess demand the converse will be true (with appropriate response lags in both cases), but there is no reason to insist that demand match supply at any point in time or that prices systematically adjust so that the market equilibrates or stabilizes in any formal sense.

Figure 8 summarizes the overall housing market model system implemented within the current version of ILUTE, showing how the "demand" and "supply" components discussed in the previous chapters come together within the market clearing procedure to allocate individual households to individual dwelling units.

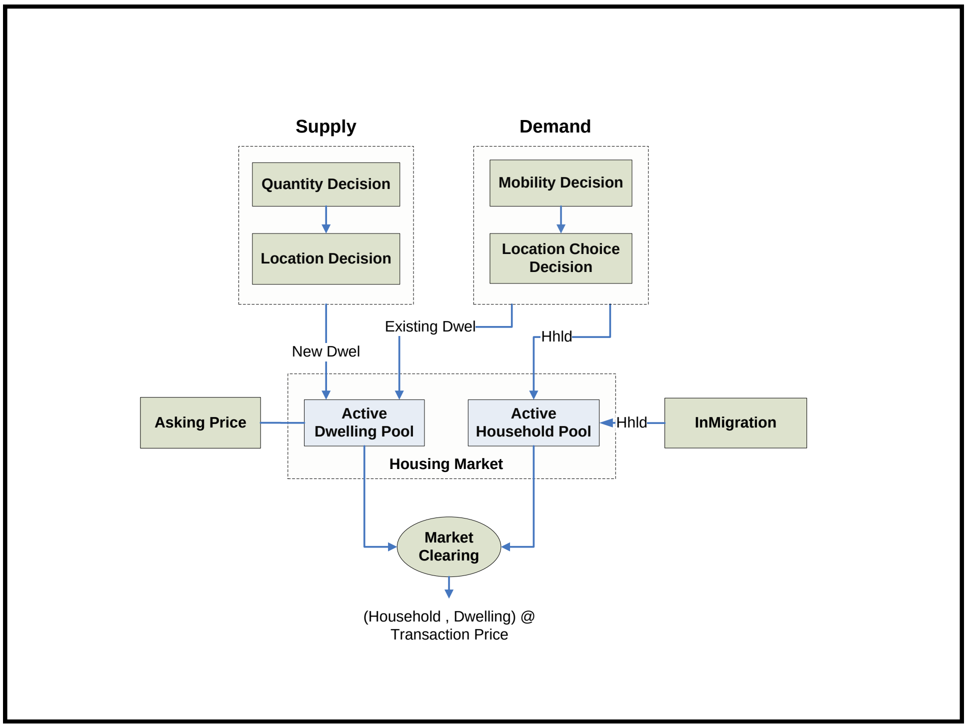
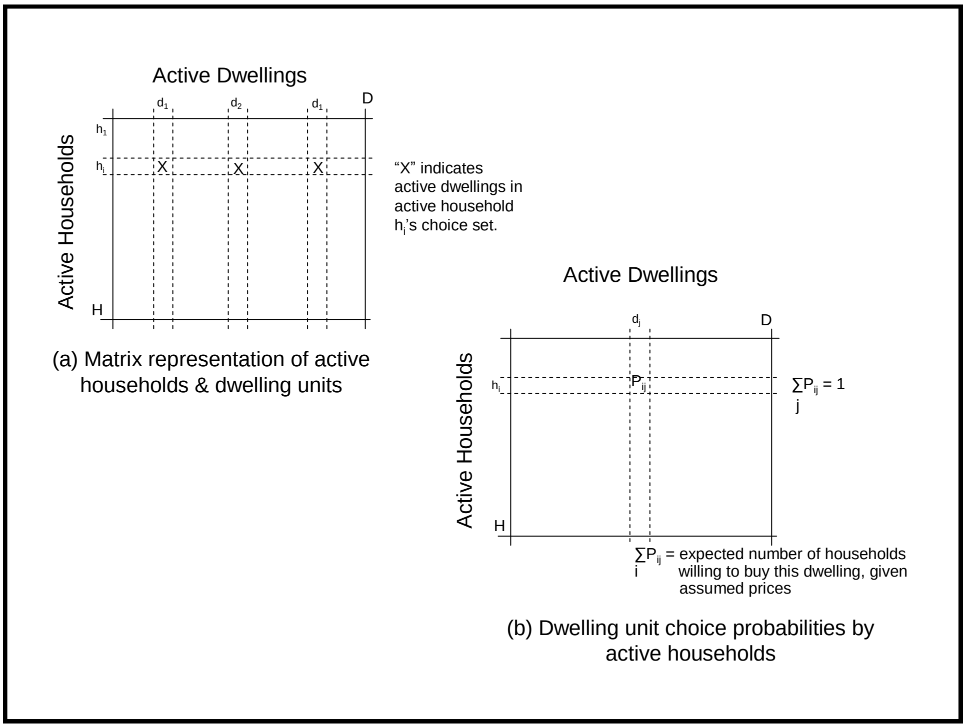


Figure 8 – ILUTE owner-occupied housing market model

At any point of time, the output from the demand and supply components of ILUTE generate a list of H active households looking for a new dwelling and a list of D active dwellings. As illustrated in Figure 9, these can be arranged in a matrix format, in which each row corresponds to a specific household and each column corresponds to a specific dwelling unit. Each household also has a choice set, C_h, of active dwellings that it is actively considering to purchase. This choice set is indicated by the "X's" in Figure 9(a).

As discussed in detail in Miller, Farooq and Habib (2008), a residential location choice model is used to compute location choice probabilities for each household, given an assumed set of prices. These prices are initialized using the asking price model presented in Miller, Farooq and Habib (2008). That is, as each household places its current dwelling on the market, it determines the initial asking price for this dwelling. This result is a matrix of household choice probabilities, as illustrated in Figure 9(b), in which the choice probabilities in each row sum to one by definition.

The column sums in Figure 9(b), however, can sum to any number greater or less than one. This column sum represents the expected number of households for whom the given dwelling would be their maximum utility alternative, given a set of prices, **p**, and, hence, the number of households who would be willing to purchase the dwelling at the specified price. If the price vector **p** where adjusted so that all columns sum to one (while maintaining all the row sums equal to one), then this would represent a classic Walrusian system-wide market equilibrium solution in which prices are set so as to “match demand to supply” and households could be allocated to dwellings based on the probabilities defined by this “equilibrated” table (Martinez and Donoso, 2001).



“X” indicates active dwellings in active household h_i’s choice set.

(a) Matrix representation of active households & dwelling units

$\sum_j P_{ij} = 1$

$\sum_i P_{ij}$ = expected number of households willing to buy this dwelling, given assumed prices

(b) Dwelling unit choice probabilities by active households

Figure 9 – Matrix representation of active households & dwellings

In ILUTE, however, market equilibrium in a given time period is not expected. Indeed, it is assumed that each buyer and seller in the market is myopic in that:

- Buyers are not aware of the actions of other buyers and only have detailed information concerning dwellings that are in their choice set.
- Suppliers are not aware of the actions of other sellers, and only have detailed information concerning the buyers that are in their prospect set.
- Dwellings are auctioned one at a time, and so prices are myopically determined one sale at a time, not by some global “balancing” process.

Given these assumptions, in ILUTE dwellings are sold one at a time by randomly selecting an active dwelling, determining its sales price (while holding the prices of all other active dwellings fixed for the moment), and then randomly selecting the household from within this dwellings prospect set that will purchase the dwelling at the sales price. Once this dwelling has been sold to the selected household, the following updating steps are taken:

- The household is moved to its new location and returned to the passive state with respect to the housing market.
- The dwelling is removed from the list of active dwellings in the housing market.
- The choice sets of all households in the dwelling’s prospect set are updated by deleting the sold dwelling from their choice sets. Their location choice probabilities are updated accordingly. The unsuccessful households can also decide to exit the housing market by becoming passive at their current location, providing that they have not yet sold this dwelling.

Figure 10 illustrates this process, which continues until all dwellings have been auctioned off for this time period.

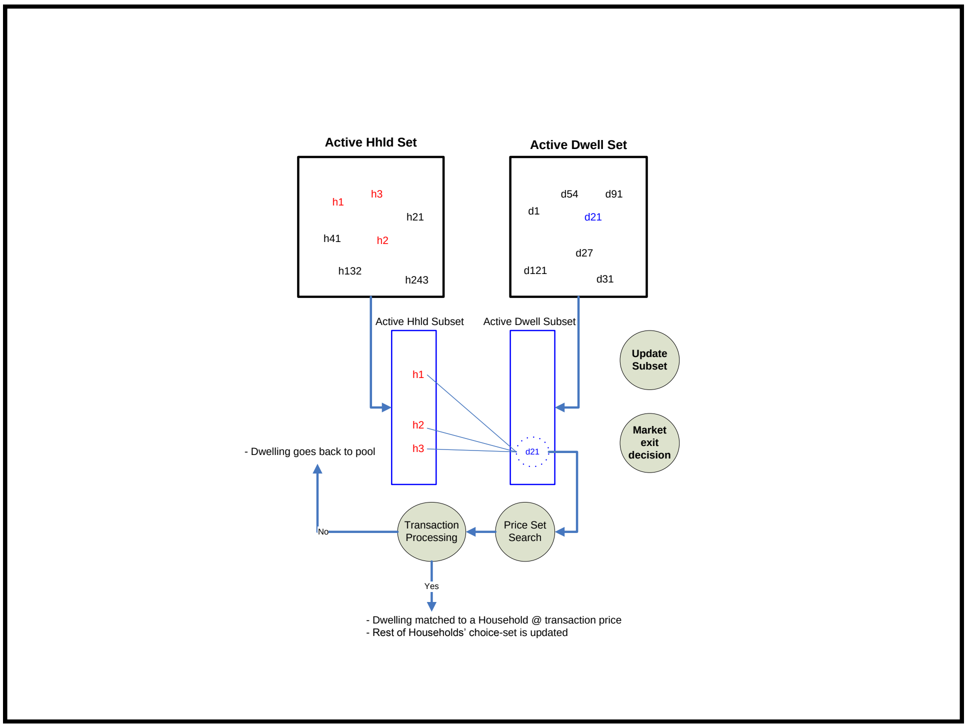


Figure 10: Owner-occupied housing market clearing process

The critical issue in implementing this procedure is how to determine the selling price for a given dwelling, given its prospect set of potential buyers and their preferences for this unit (relative to the other available dwellings in their choice sets), which will go up and down as the dwelling price is lowered or raised (relative to the temporarily fixed prices of the competing dwellings).

Given that the dwelling's column sum of the households' location choice probabilities equals the expected number of bidders for this dwelling, a sum of less than one can reasonably be assumed to indicate that the dwelling is over-priced relative to the current market, while a sum of greater than one implies that the dwelling is under-priced. Thus, it is assumed in this model that the dwelling owner will adjust the price of the unit until column sum equals one, at which point the dwelling is assumed to be neither over- nor under-priced. This is the "micro" equivalent of the system-wide equilibrium assumption, applied to a single dwelling. It is, indeed, equivalent to the system-wide method if all other prices are assumed as fixed. Again, it is argued that the "fixed price" assumption is appropriate for the individual dwelling auction case, since:

- It is assumed that dwellings are only auctioned one at a time, and so the "current" prices for all other active dwellings are, indeed, the asking price.
- The seller cannot influence these other prices, and, indeed, does not have detailed knowledge of them.
- The buyers' information consists only of the asking prices of the other dwellings in their choice sets. Given that they have not yet participated in auctions for these

dwellings, their "best guesses" at what price they could buy these dwellings for are the asking prices.

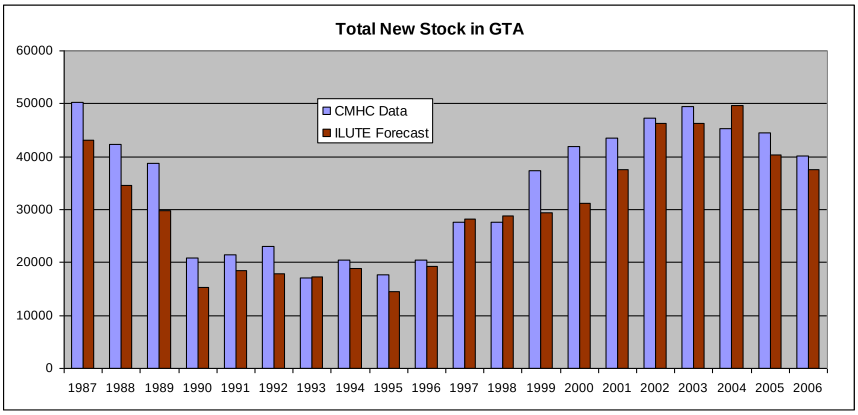
Solving for the price that causes the dwelling's prospect set (column) probabilities to sum to one is a non-trivial calculation. Details of the solution procedure used are provided in Miller, *et* al., (2008). Once the price is determined that causes the dwelling's column to sum to one, the non-zero elements of this column are interpreted as the probabilities that each household in the prospect set will be the successful bidder. Effectively this represents the assumption that households for whom this dwelling generates a higher systematic utility will be more likely to select this dwelling as their "best choice" and hence more likely to pay the sales price. Given these probabilities, the dwelling is sold to the household with the highest choice probability.

Finally with respect to the market clearing process, note that the supply and demand models generate the total number of households and dwellings that are predicted to be active in an entire year within the simulated time period. In reality, of course, only a fraction of these agents are active in the market at one point in time. In order to reflect this, without resorting to sub-year time steps, the process is initiated in Year 1 by randomly drawing 1/m of the total number of households and 1/m of the total number of dwellings from the Year 1 totals and this sub-sample is first processed. To avoid artificially creating a situation of "scarcity" in which very few dwellings and/or households are active in the market, once 1/(2m) (i.e., one-half of the sub-sample) of the dwellings have been allocated, another 1/(2m) random sample of dwellings and households is drawn. This process continues until only 1/(2m) dwellings for this year remain unsold. At this point the remaining active dwellings and households are "rolled over" as active agents for Year 2 processing, along with the first 1/(2m) sample from the new Year 2 active agents. Thus, the set of active agents in the market that are being processed always fluctuates between 1/(2m) and 1/(m) of the total pool of active agents. m is a model system parameter that is currently set at 12.

Sample Test Run Results, Owner-Occupied Housing Market

Testing of the housing market model is still in a very preliminary stage. Figure 12 compares total predicted and observed new housing by year from an early model run. In general, ILUTE is currently under-predicting the supply of new housing, but the overall shape of the trend curve is quite good.

Figure 13 plots the distribution of GTHA 2001 asking prices, averaged over ten 10% sample runs. All prices are in 2001 Canadian dollars. We do not have observed asking prices for 2001, but we do have average transaction prices. The average Toronto Real Estate Board (TREB) transaction price for 2001 is $222,000, whereas the ILUTE average asking price is $380,000. As is discussed further below, ILUTE transaction prices are higher than observed transaction prices. Given the asking prices shown in Figure 14, it is clear that the current asking price model is initializing the market with prices that are too high, leading to transaction prices that are also too high. A next step in the model development process will be to recalibrate the asking price model so that lower, more representative, asking prices are generated.



CMHC: Canadian Mortgage and Housing Corporation

Figure 12 – Predicted & observed GTHA supply of new housing

Table 2 – Predicted & observed transaction prices by dwelling structure type

	ILUTE		TREB	DELTA
Dwelling Type	**Average**	**St. Dev.**	**Average**	
Detached	480,000	200,000	307,000	173,000
Semi-Detached	280,000	130,000	230,000	50,000
Attached	260,000	110,000	212,000	48,000
Apartment	226,000	96,400	182,000	44,000
Total	392,000	180,000	222,000	170,000

Table 2 summarizes the means and standard deviations for 2001 predicted transaction prices, categorized by dwelling structure type, compared with average TREB values. In comparing ILUTE predictions to TREB values it should be noted that the TREB values are based on a single average value per "TREB zone", where these zones are geographically large. Thus, a considerable amount of variance has been squeezed out of the TREB data. In all cases, ILUTE is over-predicting average transaction prices relative to the observed average TREB values, although the TREB averages do all fall within one standard deviation of the ILUTE averages. Further, the predicted averages for 3 of the 4 dwelling types are within $50,000 of the observed averages. Prices for detached dwellings (which constitute a significant majority of the transacted dwellings), however, are clearly significantly over-predicted. Again, the most logical source of this error is the asking price model, which clearly needs to be revisited in subsequent testing.

Finally, Figure 14 displays the distributions of transactions prices generated by ILUTE by dwelling structural type, averaged over the ten 10% sample runs.

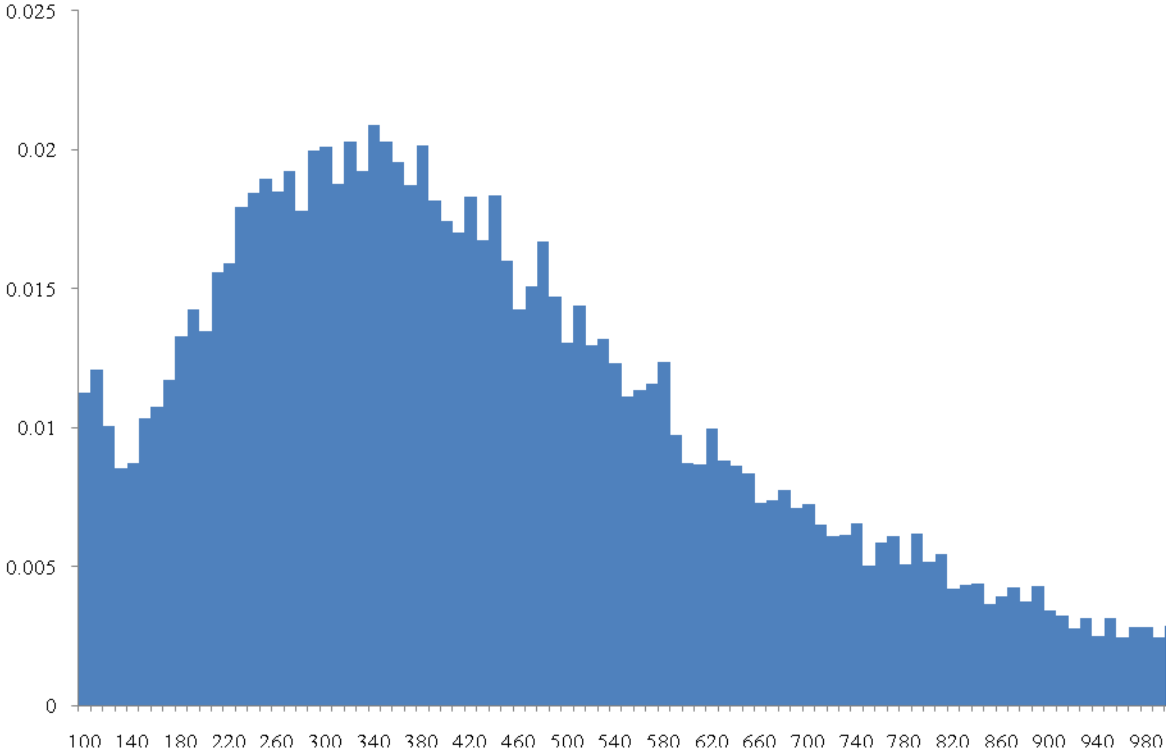


Figure 13 – 2001 predicted asking prices

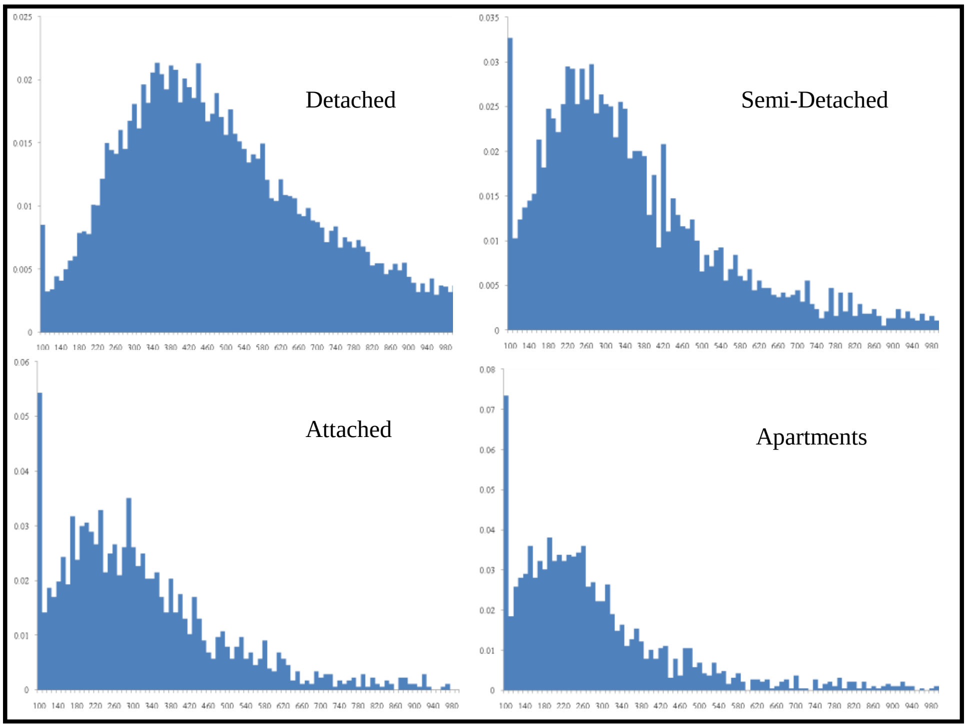


Figure 14 – 2001 predicted transaction prices by dwelling structure type

SUMMARY AND FUTURE WORK

This paper has summarized the current state of the ILUTE agent-based ,icrosimulation integrated urban model system. In addition to providing a general overview, the paper has focussed on the demographic and housing market components of the model system, which are two of the key components which currently undergoing extensive testing. This testing consists of running 20-year historical simulations (1986-2006) so that model outputs can be compared with observed data for this time period.

Much work remains to be done. With respect to the demographic model, improvements in both the in-migration and out-migration models are possible, providing that improved historical data for the Toronto region can be obtained. The marriage market model also requires further testing and calibration.

With respect to the owner-occupied housing market model, the asking price model, particularly for detached dwellings urgently requires recalibration/improvement. The new dwelling supply model also requires further calibration. Once this model has been improved, further detailed testing of the owner-occupied housing market model is required, with respect to temporal mobility rates, transaction prices, and the spatial distribution of location choices.

Several components of ILUTE are still currently under development. These include:

- The rental housing market (currently being developed as a Master's thesis).
- The labour market model (currently being developed as a Master's thesis).
- A firmographic model.
- Improved feedback from the transportation (TASHA-MATSim) model sub-system and the location choice processes modelled.

In addition to these components, a general concern that cuts across all of the spatial choice models within ILUTE is choice set modelling. As has often been observed in the literature (Timmermans, 2003; Miller, 2009), choice set modelling is all too often the weak link in integrated urban models. ILUTE is no exception in this regard. It is anticipated that a major research thrust by the ILUTE group in the near future will be to develop improved spatial choice models for location (residential and commercial), labour and travel (non-work/school destination) choice set determination. One recent attempt in this regard is the Ph.D. thesis of Elgar (Elgar, *et al*., 2008), but much remains to be done to generalize and operationalize these findings.

REFERENCES

Chingcuanco, F. (2010). "ILUTE Demographic Updating Module: Current Operational Status", unpublished working paper, Toronto: Urban Transportation Research and Advancement Centre, University of Toronto.

Choo, E., Seitz, S., & Siow, A. (2008). "The Collective Marriage Matching Model: Identification, Estimation and Testing", Univ. of Toronto Working Paper 340, 1-65.

Elgar, I, E.J. Miller and B. Farooq (2008). Development of an Operational Integrated Urban Model System, Volume VI: Modelling Firm Mobility & Location Choice, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Farooq, B., P.A. Salvini and E.J. Miller (2008). Development of an Operational Integrated Urban Model System, Volume X: ILUTE Software Documentation, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Gao, W., M. Balmer and E.J. Miller (2010). "Comparisons between MATSim and EMME/2 on the Greater Toronto and Hamilton Area Network", forthcoming, Transportation Research Records, Journal of the Transportation Research Board.

Guan, J., M.J. Roorda and E.J. Miller (2003). "Approximation of 24 Hour Travel Times in the Greater Toronto Area", pre-print CD, 82nd Annual Meeting of the Transportation Research Board, Washington, D.C., January.

Hao, J., M. Hatzopoulou and E.J. Miller (2010). "Integrating an Activity-Based Travel Demand Model with Dynamic Traffic Assignment and Emission Models: An Implementation in the Greater Toronto Area", forthcoming, Transportation Research Records, Journal of the Transportation Research Board.

Hatzopoulou, M. and E.J. Miller (2008a). Development of an Operational Integrated Urban Model System, Volume II: A Survey of Canadian Modelling Practice And Needs, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Hatzopoulou, M. and E.J. Miller (2008b). Development of an Operational Integrated Urban Model System, Volume VIII: Environmental Modelling, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Hatzopoulou, M., E.J. Miller and B. Santos (2007). "Integrating Vehicle Emission Modeling with Activity-based Travel Demand Modeling: A Case Study of the Greater Toronto Area (GTA)", Transportation Research Records, Journal of the Transportation Research Board, No. 2011, pp. 29-39.

Maoh, H. F. (2005). The Location of Business Establishments in the City of Hamilton, Canada: A Micro-Analytical Model Approach, Ph.D. thesis, Hamilton: McMaster University.

Martínez, F and P. Donoso (2001). "MUSSA: A Land Use Equilibrium Model with Location Externalities, Planning Regulations and Pricing Policies", 7th International Conference on Computers in Urban Planning and Urban Management (CUPUM 2001), Hawaii, July.

Miller, E.J. (2008). Development of an Operational Integrated Urban Model System, Volume I: Project Final Report, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Miller, E.J. (2009) "Integrated Urban Models: Theoretical Prospects", invited resource paper, Chapter 14 in Kitamura, R., T. Yoshii and T. Yamamoto (eds.), The Expanding Sphere of Travel Behaviour Research: Selected Papers from the 11th International Conference on Travel Behaviour Research, Bingley, UK: Emerald, pp. 351-384.

Miller, E.J., J.A. Carrasco, K.M.N.Habib and M.J. Roorda (2008). Development of an Operational Integrated Urban Model System, Volume VII: Activity Travel Modelling, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Miller, E.J., F. Chingcuanco, B. Farooq, K.M.N. Habib and M.A. Habib (2008). Development of an Operational Integrated Urban Model System, Volume IV: Demographic & Labour Market, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Miller, E.J., B. Farooq and M.A. Habib (2008). Development of an Operational Integrated Urban Model System, Volume V: Microsimulating Spatial Markets: Conceptual Design & Implementation of A Housing Market Model, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Miller, E.J., M. Hatzopoulou, J.Y. Hao and J. Vaughan (2009). Modelling of Transport Emissions in the Greater Toronto and Hamilton Area, final project report to Transportation Services Division, City of Toronto, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Miller, E.J., J.D. Hunt, J.E. Abraham and P.A. Salvini (2004). “Microsimulating Urban Systems”, Computers, Environment and Urban Systems, special issue, “Geosimulation: Object-Based Modeling of Urban Phenomena”, 28, pp. 9-44.

Miller, E.J. and M.J. Roorda (2003). “A Prototype Model of Household Activity/Travel Scheduling”, Transportation Research Record, Journal of the Transportation Research Board, No. 1831, 2003, pp. 114-121.

Miller, E.J., M.J. Roorda, C.A. Kennedy, A.S. Shalaby and H. MacLean (2006). Activity-Based, Multi-Modal Modelling of Travel Behaviour for Urban Design, final project report to Transport Canada Transportation Planning and Modal Integration Initiatives, Toronto: Joint Program in Transportation, University of Toronto.

Miller, E.J. and P.A. Salvini (1998). "The Integrated Land Use, Transportation, Environment (ILUTE) Modeling System: a Framework", presented at the 77th Annual Meeting of the Transportation Research Board, Washington, D.C., January 12-15.

Miller, E.J. and P.A. Salvini (2001). "The Integrated Land Use, Transportation, Environment (ILUTE) Microsimulation Modelling system: Description & Current Status", Chapter 41, in D. Hensher (ed.) The Leading Edge in Travel Behaviour Research, Amsterdam: Pergamon, pp. 711-724

Moeckel, R. (2005). “Microsimulation of firm location decisions”, in: Proceedings of the 9th International Conference on Computers in Urban Planning and Urban Management (CUPUM). London: Centre for Advanced Spatial Analysis (CASA UCL).

Mohammadian, A. and E.J. Miller (2002a). “Nested Logit Models and Artificial Neural Networks for Predicting Household Automobile Choices: Comparison of Performance”, Transportation Research Record, Journal of the Transportation Research Board, No. 1807, pp. 92-100.

Mohammadian, A. and E.J. Miller (2002b). “Estimating Expected Price of Vehicles in a Transportation Microsimulation Modeling System”, Journal of Transportation, pp. 537-541.

Mohammadian, A. and E.J. Miller (2003a). “An Empirical Investigation of Household Vehicle Type Choice Decisions”, Journal of Transportation Research Record, No. 1854, pp. 99-106.

Mohammadian, A. and E.J. Miller (2003b). “Dynamic Modeling of Household Automobile Transactions”, Transportation Research Record, Journal of the Transportation Research Board, No. 1831 2003, pp. 98-105.

Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (2007). "Root Finding and Nonlienar Sets of Equations", chapter 9 in Numerical Recipes: The Art of Scientific Computing, 3rd Ediiton, Cambridge: Cambridge University Press.

Pritchard, D.R. and E.J. Miller (2008). Development of an Operational Integrated Urban Model System, Volume III: Synthesizing Base Year Agents, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto.

Pritchard, D.R. and E.J. Miller (2009). "Advances in Agent Population Synthesis and Application in an Integrated Land Use / Transportation Model", pre-print CD, 88th Annual Meeting of the Transportation Research Board, Washington, D.C., January, 2009.

Roorda, M.J., R.A. Cavalcante, S. McCabe and H. Kwan (2008). Development of an Operational Integrated Urban Model System, Volume IX: A Conceptual Framework for Modelling Goods Movement And The Regional Economy, Toronto: Urban Transportation Research & Advancement Centre, University of Toronto

Roorda, M.J. and E.J. Miller (2006). "Assessing Transportation Policy Using an Activity-Based Microsimulation Model of Travel Demand", winner, ITE Past Presidents' Award 2006, ITE Journal, November, 2006, pp. 16-21.

Roorda, M.J., E.J.Miller and N. Kruchten (2006). "Incorporating Within-Household Interactions into a Mode Choice Model with Genetic Algorithm for Parameter Estimation", Transportation Research Records, Journal of the Transportation Research Board, No. 1985, 2006, pp. 171-179.

Roorda, M.J., E.J. Miller and K.M.N. Habib (2008). "Validation of TASHA: a 24-Hour Activity Scheduling Microsimulation Model", Transportation Research A, 42, 2008, pp. 360-375.

Salvini, P.A. and E.J. Miller (2005). "ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems", Networks and Spatial Economics, 5, 2005, pp. 217-234.

Timmermans, H. (2003). "The saga of integrated land use-transport modeling: how many more dreams before we wake up?" Conference keynote paper, 10th International Conference on Travel Behaviour Research, Lucerne.

Wahba, M. and A. Shalaby (2009a). "MILATRAS: A New Modelling Framework for the Transit Assignment Problem", in Schedule-Based Modeling of Transportation Networks: Theory and Applications, Chapter 9, pp. 171-194, Nigel Wilson and Agostino Nuzzolo (Eds.), Springer-Verlag (ISBN: 978-0-387-84812-9)

.Wahba, M. and A. Shalaby (2009b). "Learning-based Departure Time and Path Choice Modelling for Transit Assignment under Information Provision: A Theoretical Framework", presented at the 12th International Conference on Travel Behaviour Research, Jaipur, India, December 13-18.

Wang, J., M. Wahba and E.J. Miller (2010). "A Comparison of an Agent-Based Transit Assignment Procedure (MILATRAS) with Traditional Approaches ",forthcoming, Transportation Research Records, Journal of the Transportation Research Board.