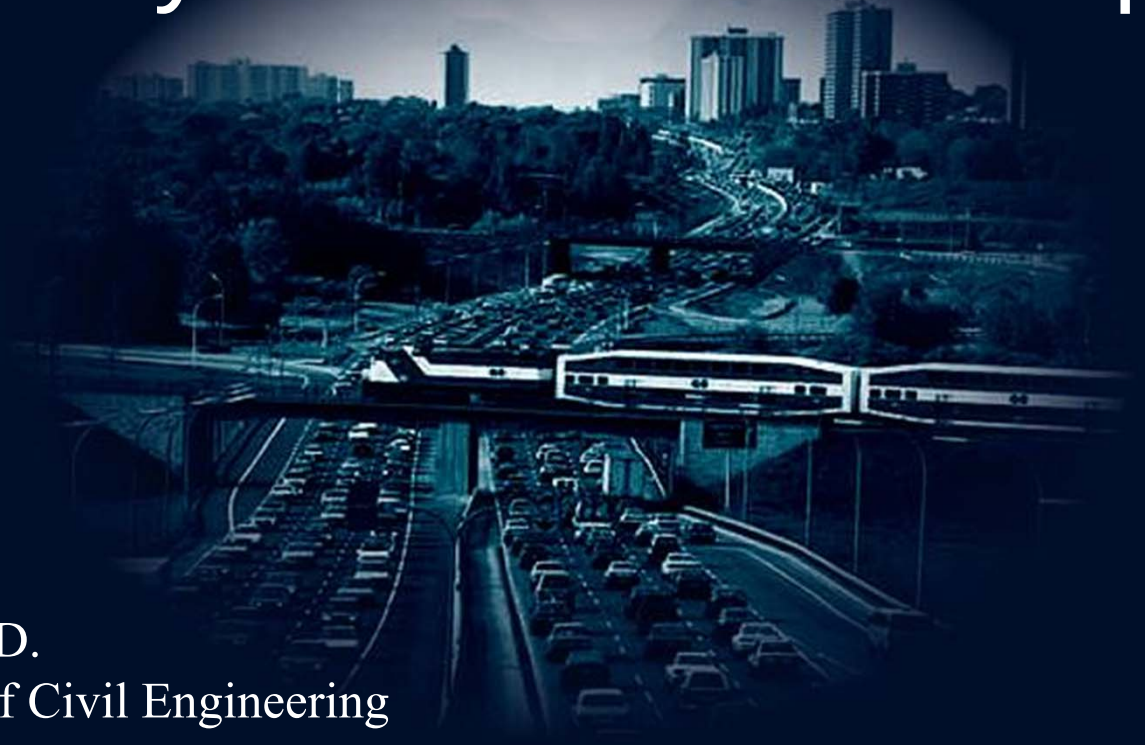
Modelling Information & Learning in Agent-Based Models of Activity/Travel: Issues & Options



Eric J. Miller, Ph.D.
Professor, Dept. of Civil Engineering
University of Toronto

Frontiers in Transportation 2013 Workshop
Herrsching am Ammersee, Germany
August 1, 2013

ILUTE

Presentation Outline

- Issues in location choice modelling
- The location choice set formation problem
- Some propositions
- Next steps

This talk builds upon J. Wang and E.J. Miller. *A prism- and gap-based approach to shopping location choice*. Forthcoming, Transportation Research B, 2014.

Spatial Choice

- Defined most broadly, a spatial choice is any discrete choice from a set of possible spatial locations or destinations.
- Examples include:
  - Residential location choice
  - Job location choice
  - Firm location choice
  - School location choice
  - Shopping destination choice
  - Vacation location choice
  - …

Accuracy of Spatial Choice Models

- Despite a long history of usage, spatial choice models have progressed relatively little (Timmermans, 2003) and are often found to be inaccurate at the traffic zone level at which travel demand models generate operate, e.g.:
  - Work trip commuting distributions (Hutchinson & Smith, 1979).
  - Small office location choice (Elgar, et al., 2009).
  - Shopping destination choice (Wang & Miller, 2014).

Accuracy, cont'd

- This poor performance is worrisome, since it calls into question the validity of travel demand forecasting models in general.
- Surprising little attention has been paid to this problem in operational practice:
  - Model parameters are estimated to fit observed trip length frequency distributions, not to O-D cell values.
  - Models are calibrated to fit aggregate screen-line counts.
- The shift to activity-based models has arguably not improved operational performance to date.

Location Choice in Activity/Travel Models (ATMs)

- Work & school locations can be assumed to be determined by longer-run market processes.
- Non-work/school activity episode location (trip destination) choices are, arguably, the "weak link" in ATMs.
- Will focus on shopping episode location in this presentation.

“Classic” Problems of Spatial Choice Models

- Extremely large (universal) choice sets.
- General lack of good rules for determining choice sets.
- Very static; lack of a dynamic representation of choice set evolution over time
- Many latent elements

Latent Elements

- History of shopping experience generally not observed ("left-censoring" in both data & models)
- Aggregate activity episode categories (e.g., what kind of shopping?)
- Detailed attributes of shopping locations not known.
- Usual inter-personal heterogeneity
- Information channels & their impacts (media, social networks, personal experience,…)

Dynamics

- Again, history – what are the "initial conditions"?
- We will never simulate every hour/day in the life.
- When is an episode location chosen? Prior to scheduling? During scheduling? Prior to mode choice?

Information

- Full social networks do not (yet) exist in practical models; so how to model social influence?
- How to model media information dissemination?
- What does “information” mean in a computer-based model? How is it represented and stored?

Information, cont'd

- "Aware" of a location (yes/no).
- Perceptions of location attributes
  - Another potential level of latency/modelling
  - Computational/memory issues (maintaining personalized "mental maps" of an urban region for each agent a potentially challenging task)

Agent-Based Microsimulation (ABM)

- ABM holds the promise for providing a computational framework for addressing many of these issues:
  - Can track/store “history”
  - Implement “learning”
  - Potentially computationally efficient (depending on implementation)

ABM (cont'd)

Even in agent-based microsimulations, in any operational model we are inevitably trying to simulate an arbitrary day/week in the life of our agents, without knowing the history of what they did the day/week/month before. This is equally true for integrated urban simulation models that step agents through multiple years. This works well (or at least potentially can) for housing, labour force participation, demographics, auto ownership, etc., all of which evolve on a (more or less) yearly (or, worst case, monthly) basis and for which a "full history" can be simulated and maintained. But activity/travel goes on day-by-day and we are not yet willing to simulate every day in everyone's life! (And do we need to?)

ABM, cont'd

- Can ABM be used to "get around" some of the computational & behavioural issues of "traditional", "analytically-based" models?
- Arguably we do not yet always exploit the ABM approach to its fullest in the way we formulate and apply our models.

Model Components

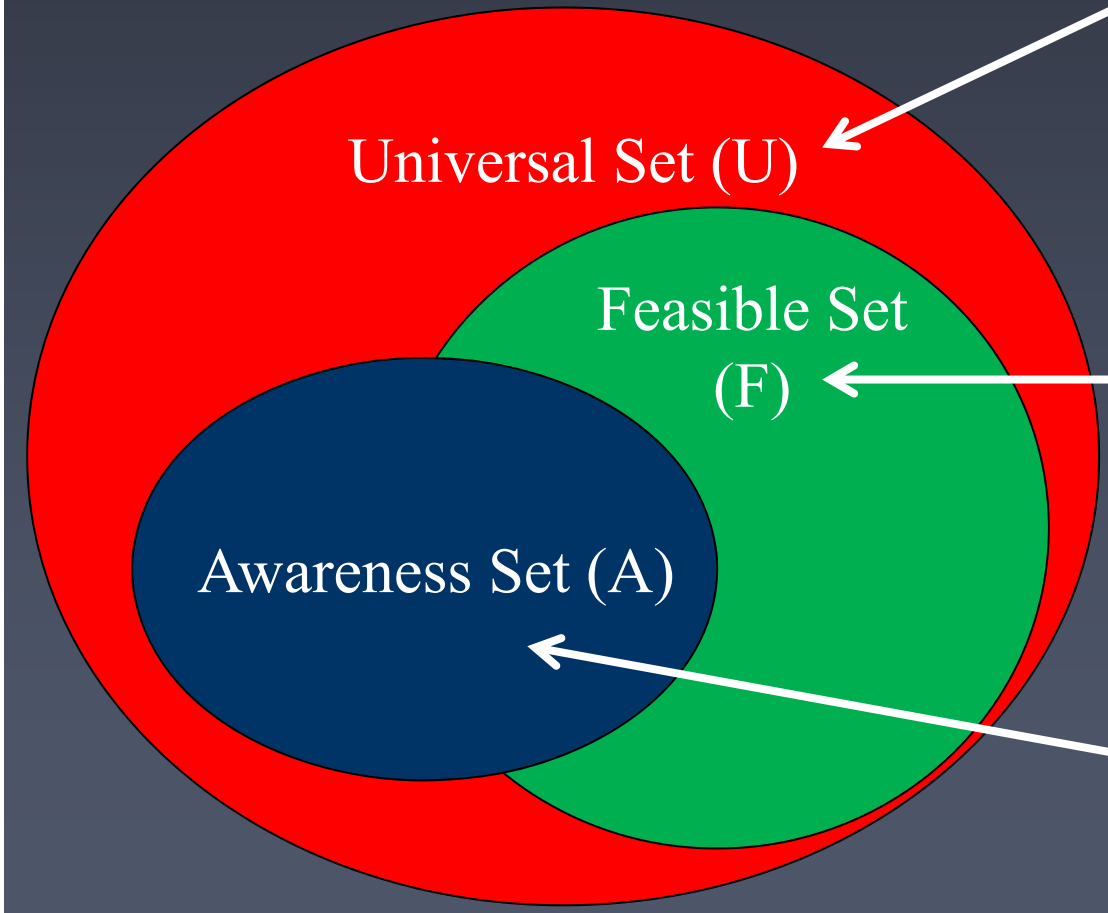
- As with market-based processes any destination choice model needs:
  - A mechanism for determining the choice set.
  - A preference function.
  - A decision rule.

Model Components

- As with market-based processes any destination choice model needs:
  - A mechanism for determining the choice set.
  - A preference function.
  - A decision rule.

Focus of this discussion.

Choice Sets: The "800-lb Gorilla"



In principle, U is "known", but data problems may exist in terms of adequately characterizing it.

F is defined by constraints (time-space, etc.), but is the context defining these constraints known?

A depends on the agent's perceptions, history, learning.

Choice set at time t, $C(t) = A(t) \cap F(t)$

Constraints: Determining Feasibility

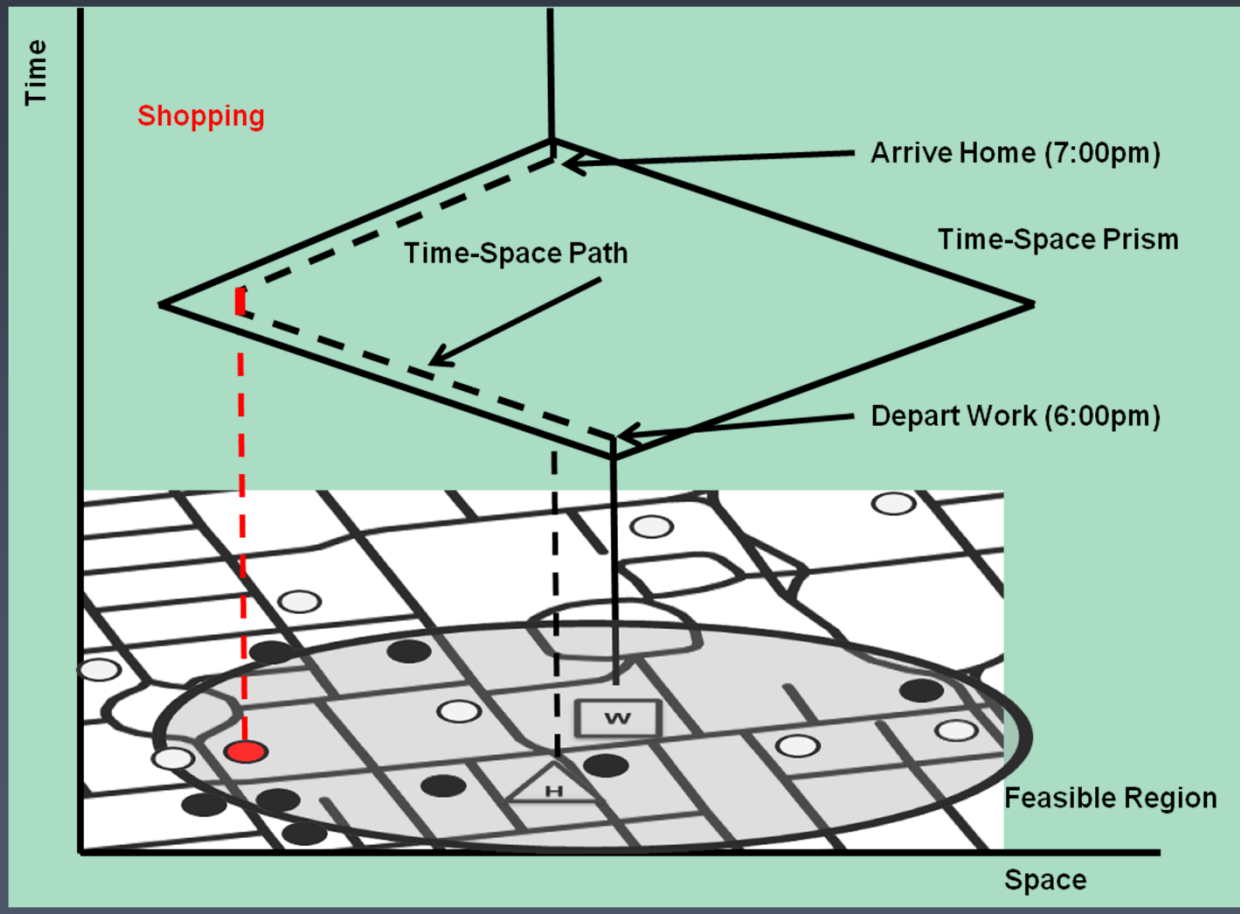
Operative constraints are person-specific:

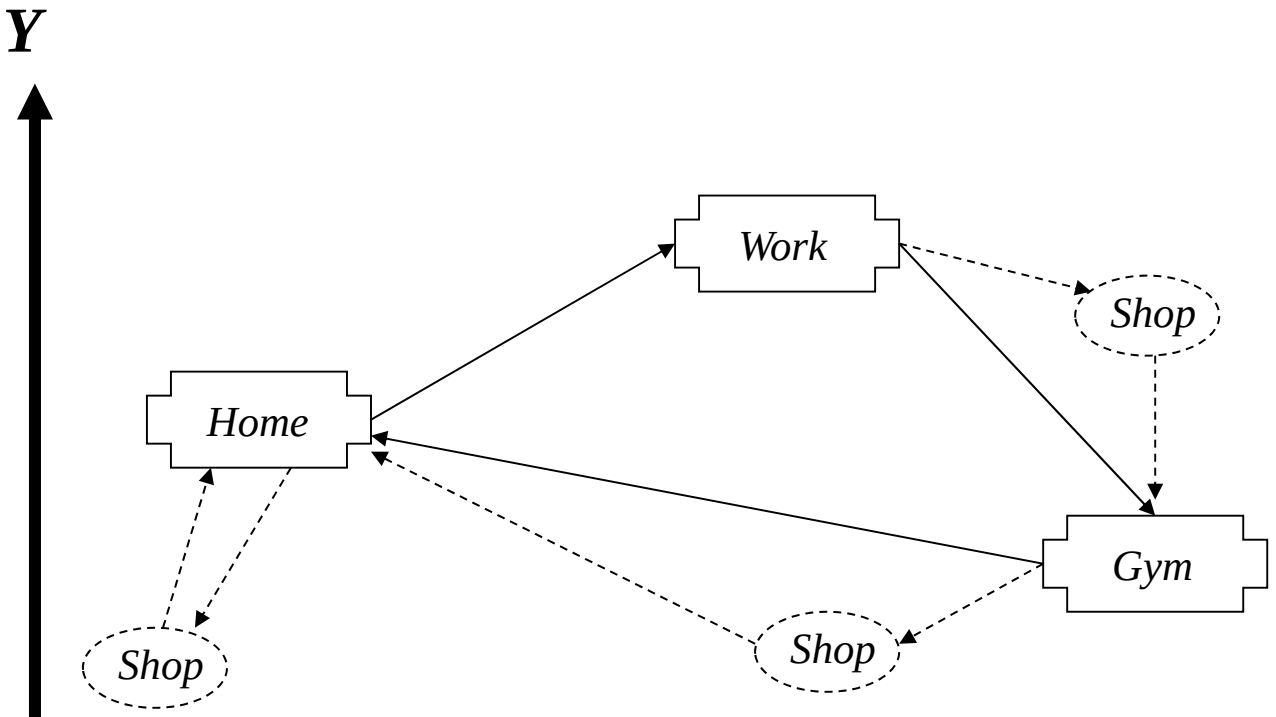
- Time
- Space
- Income
- Capability
- Household
- …

(Hagerstrand, 1970)

Time-space prisms depend on:

- Mode
- "Gap" (prior & posterior episode locations & end/start times





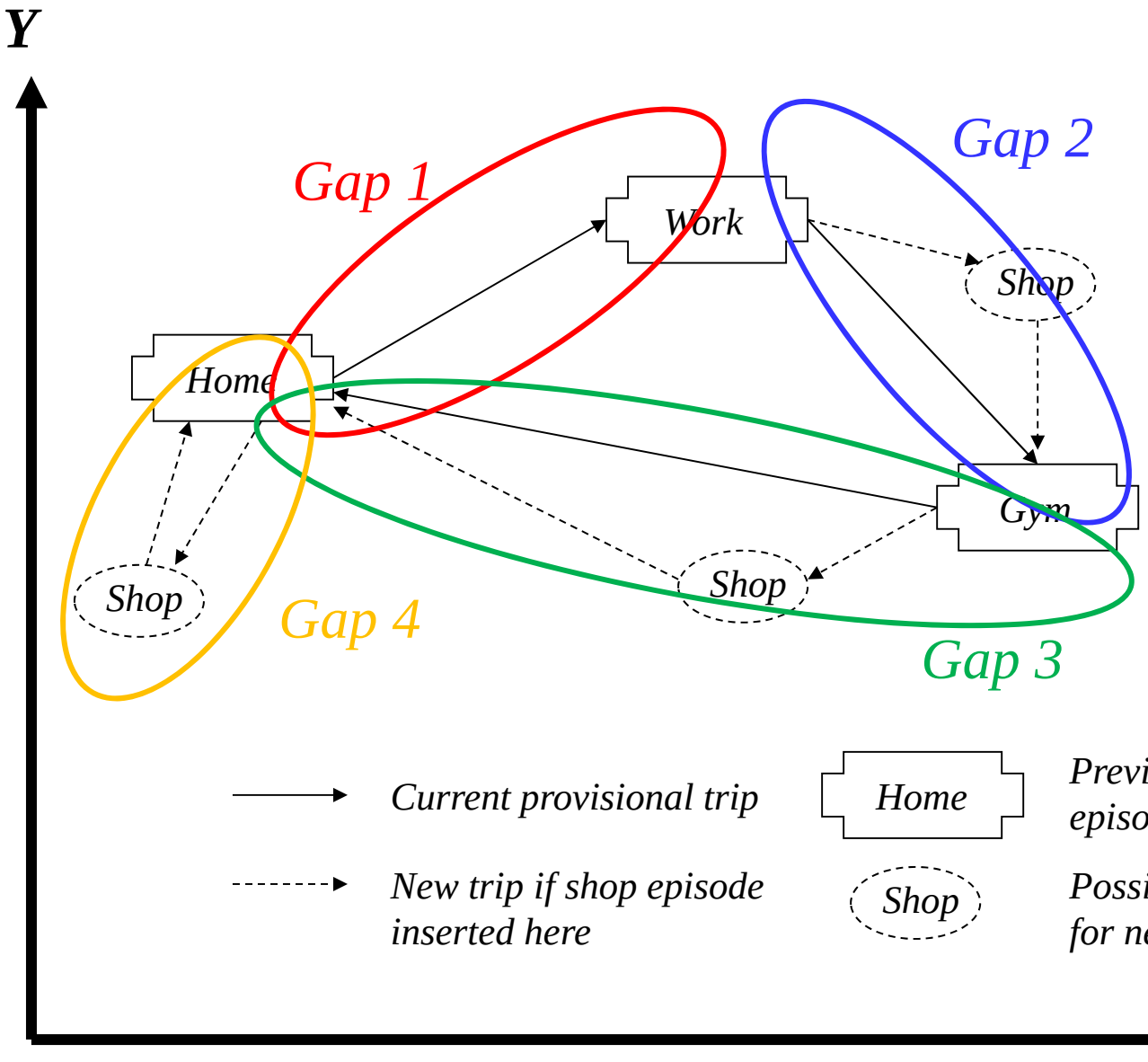
Feasible choice sets and the "utility" of locations depends on when the episode is to occur, the pre-existing "provisional schedule", the mode of travel, etc.

Current provisional trip

New trip if shop episode inserted here

Home — *Previously inserted episode*

Shop — *Possible insertion point for new shop episode*



Shopping location & timing (schedule gap) choice arguably should be treated as a joint choice over the given planning period.

Current Practice

- For logit model estimation random choice sets are often used (McFadden, 1978).
- For model application/forecasting two common approaches (both of which are wrong, at least in the absence of time-space constraints):
  - Use random choice sets (the random number generator dominates the choices).
  - Use the universal choice set (behaviourally incorrect and computationally burdensome; also generates too many low probability events).

Manski Model

A classic approach to the choice set problem is to treat the choice set as probabilistic:

$$P(j) = \Sigma_k \, P(j|C_k)P(C_k|C_k \, \varepsilon \, U) \qquad [1]$$

P(j) = Probability of choosing location j
U = Universal choice set
C_k = k^{th} feasible subset

[1], however, is not computationally tractable for even medium-sized universal choice sets.

A Proposition: Awareness Probability

$A_{ia}(t)$ = set of locations for activity a that person i is aware of at time t.

$P(j \in A_{ia}(t))$ = the probability that location j is in $A_{ia}(t)$

Arguably:

$$P(j \in A_{ia}(t)) = f(X_{ija}(t), D_{ija}, \beta_{ia})$$

$\boldsymbol{X}_{ija}(t)$ = Vector of explanatory variables for location j for person i for activity a at time t

$\boldsymbol{D}_{ij}$ = Vector of distances for location j relevant for person i

β_{ia} = Vector of parameters for person i for activity a

Awareness Probability, cont'd

- Rules may exist to define locations for which

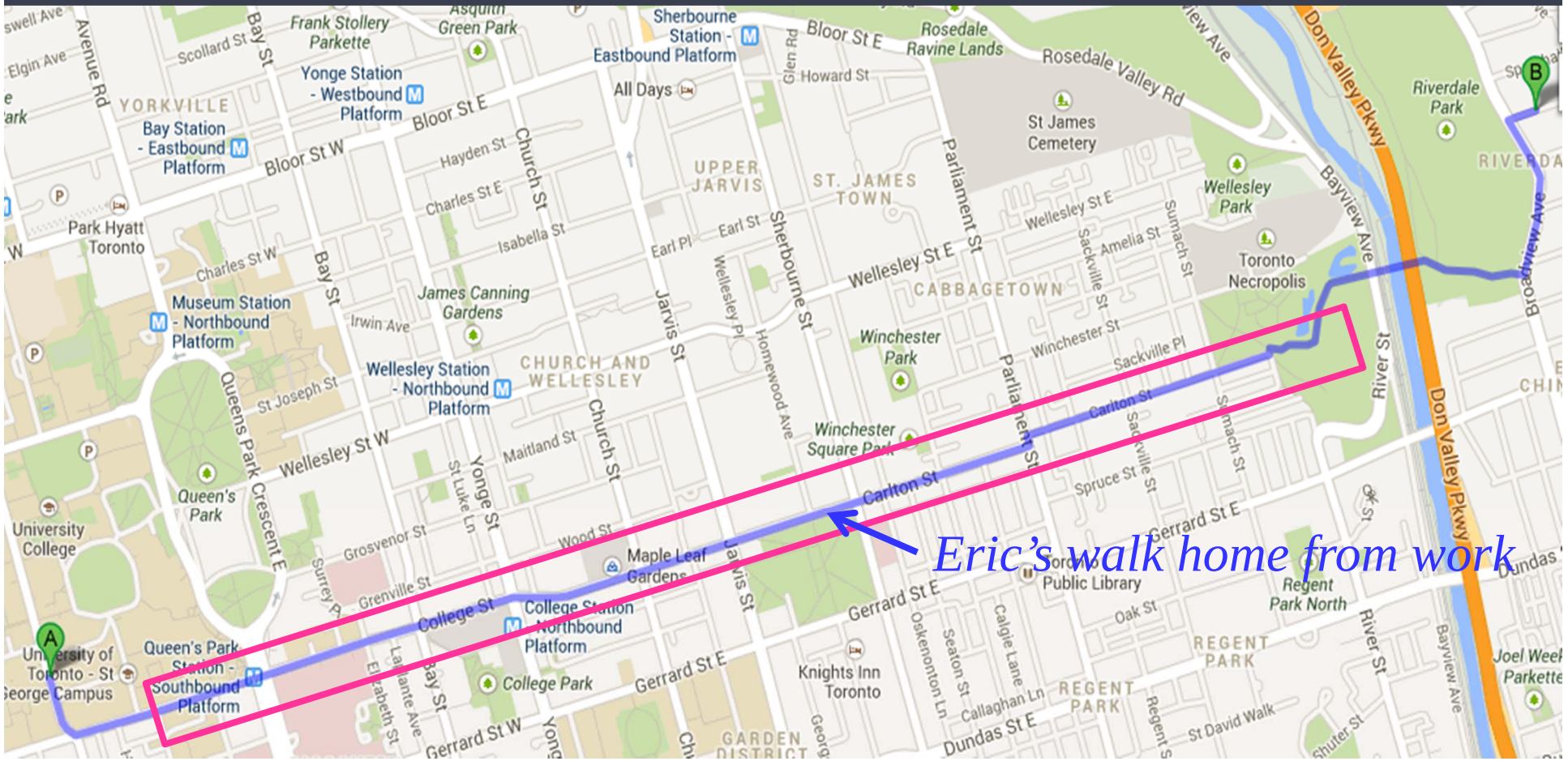
  $P(j \varepsilon A_{ia}(t)) = 1$

- E.g.:
  - Locations within radius R of one's home, place of work, place of school, etc.
  - Locations that have been previously visited
  - Locations discovered while travelling
- Maybe also $P(j \varepsilon A_{ia}(t)) = 0$?

ILUTE

"Discovering locations while travelling"

Agent-based network "assignment" models (e.g., MATSim, among others) potentially allow us to track agents' "exposure" to locations along their paths.



Awareness Probabilities, cont'd

- $P(j \varepsilon A_{ia}(t))$ might also be affected by:
  - Social influence (neighbours, family, friends)
  - Media (advertising, etc.)
- Membership in $A_{ia}(t)$ might decline over time if a given location is not visited for a long period of time.

A Proposed Microsimulation Framework for Awareness Set Evolution & Episode Location Choice

Generate a shopping episode to be scheduled with a nominal duration.

For each gap g in the provisional schedule for the planning period:
- Determine the feasible location set, F_g
  - For each location $j \in F_g$:
    - If $P(j \in A_{ia}(t)) = 1$ then add j to C
    - Else:
      - Generate a RV $u \sim U[0,1]$
      - If $u \leq P(j \in A_{ia}(t))$ then add j to C

Evaluate $P(j|C)$; draw RV & choose j* as the location.
Set $P(j^* \in A_{ia}(t)) = 1$

Proposed Microsimulation Model, cont'd

- The proposed model is a "straightforward" extension of current practice with $P(j \varepsilon A_{ia}(t))$ replacing uniform random draws.
- Builds memory into the system.
- Provides the opportunity for information from a variety of sources to affect choice sets.
- If the model runs fast enough, can run many "replications" (aka "days/weeks") per model time step (aka "years"), thereby getting closer to a quasi-continuous adaptive model.

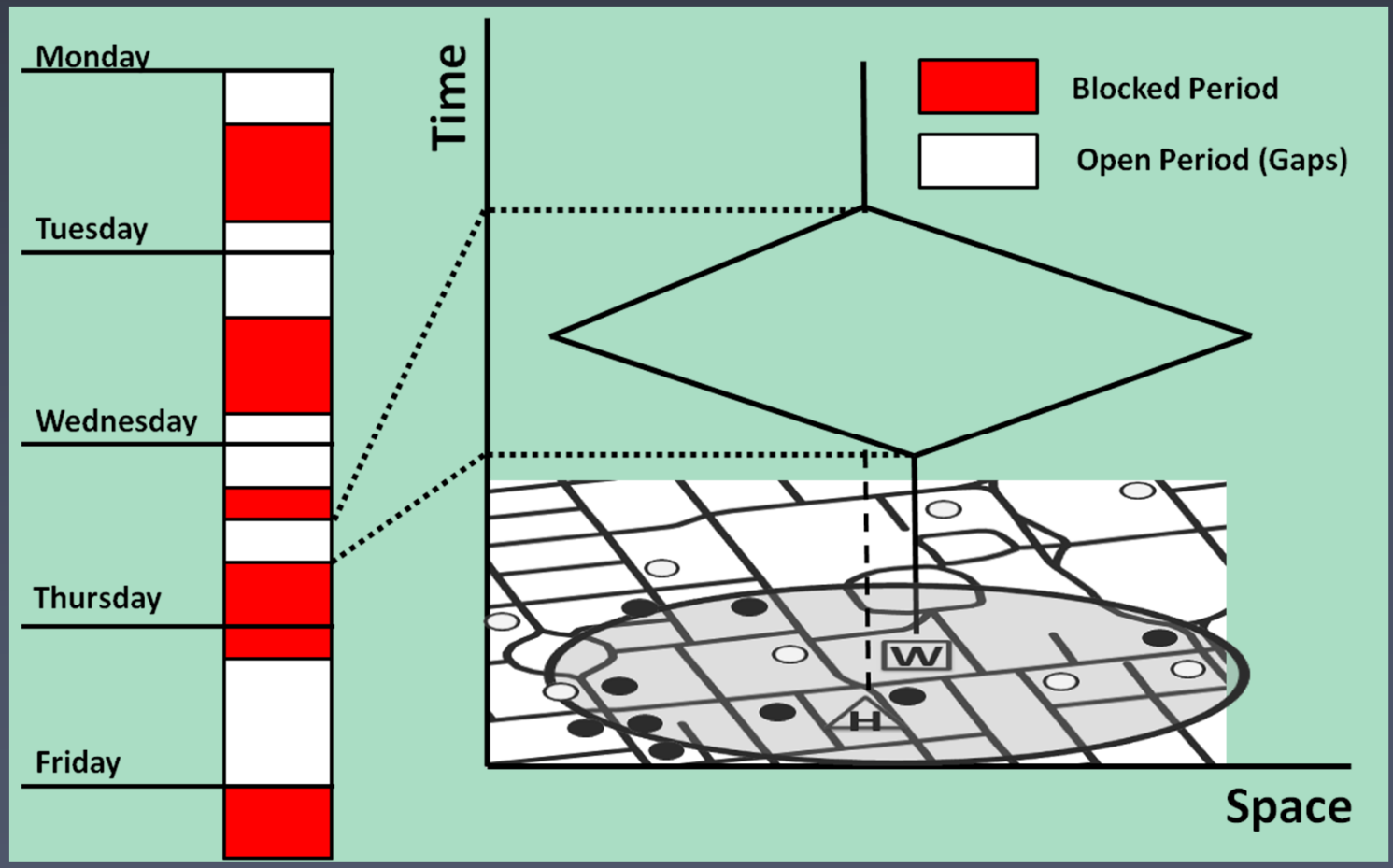
Key Questions

- Functional form for $A_{ia}(t)$?
- Can $A_{ia}(t)$ be estimated using only RP (observed location choice) data?
- If not, what data are required to build this model?
- Does it work better than random draws (if not it's not of much use!).

Next Steps

- Answer the key questions!
- A new Smart Phone based survey is under development in Toronto that (hopefully) will help address (some of) these questions.

Thank you! Questions?



References

Elgar, B. Farooq and E.J. Miller. Modeling location decisions of office firms: Introducing anchor points and constructing choice sets in the model system. Transportation Research Record, Journal of the Transportation Research Board 2133:56-63, 2009.

Hägerstrand, T. What about people in regional science? Papers of the Regional Science Association, 24, 7-21, 1970.

B.G. Hutchinson and D.P.Smith. Empirical studies of the journey to work in urban Canada", Canadian Journal of Civil Engineering 6(2):308 318, 1979.

McFadden, D. Modelling the Choice of Residential Location, In Spatial Interaction Theory and Residential Location, pp. 75-96, Ed. A. Korlquist, North Holland, Amsterdam, 1978.

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

H. Timmermans. The saga of integrated land use-transport modeling: how many more dreams before we wake up? Paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, 2003.