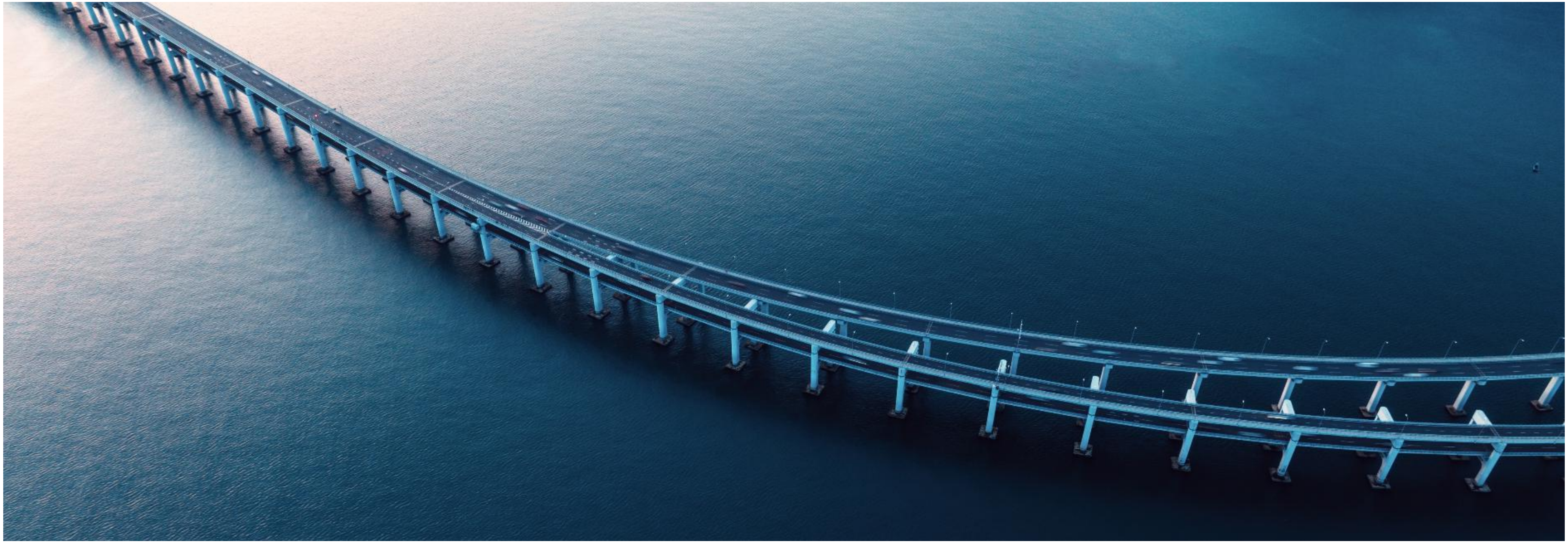
Extending the toolbox of travel modeling in practice

Modeling Mobility 2025



Peter Vovsha, Gaurav Vyas, Jim Hicks, *Bentley Systems*

MARICOPA ASSOCIATION of GOVERNMENTS

Bentley®

Intro

- Activity-Based Models (ABMs) in practice mostly rely on traditional logit models to generate individual travel choices

- It becomes important to extend the toolbox:
  - Growing number of choice dimensions
  - Realistic representation of constraints and interactions associated with individual travel behavior
- Multiple promising directions in academic research and some stand-alone examples in practice

- Mainstream of ABM industry quite reluctant to adopt new methods

 |

General factors for acceptance in practice

Acceptance in practice is function of many factors including:

- ***Simple*** estimation and calibration
- Computationally ***efficient*** application for a real-size population/network
- ***Compatibility*** with the ABM framework
- Realistic ***data*** input requirements
- Controllable ***elasticities*** that ensure logical policy sensitivity and transparency

Logit models:

- Satisfy all requirements
- However, when the model framework is extended to reflect travel behavior at a finer level, limitations of logit models manifest themselves:
  - For example, the activity/travel scheduling process is inherently better represented in continuous time

Selected extensions for discussion

- 15-min framework:
  - Optimization methods such as Liner Programming (LP)
  - Machine Learning (ML) / Artificial Intelligence (AI)
- 30-min framework w/bonus slides:
  - Multiple Discrete-Continuous Extreme Value (MDCEV)
  - Max entropy
  - Mixed logit models with random coefficients
  - Latent class discrete choice models
  - Incorporating travel time reliability

What value these methods can bring and what hampers their immediate adoption?

Factors affecting acceptance in practice

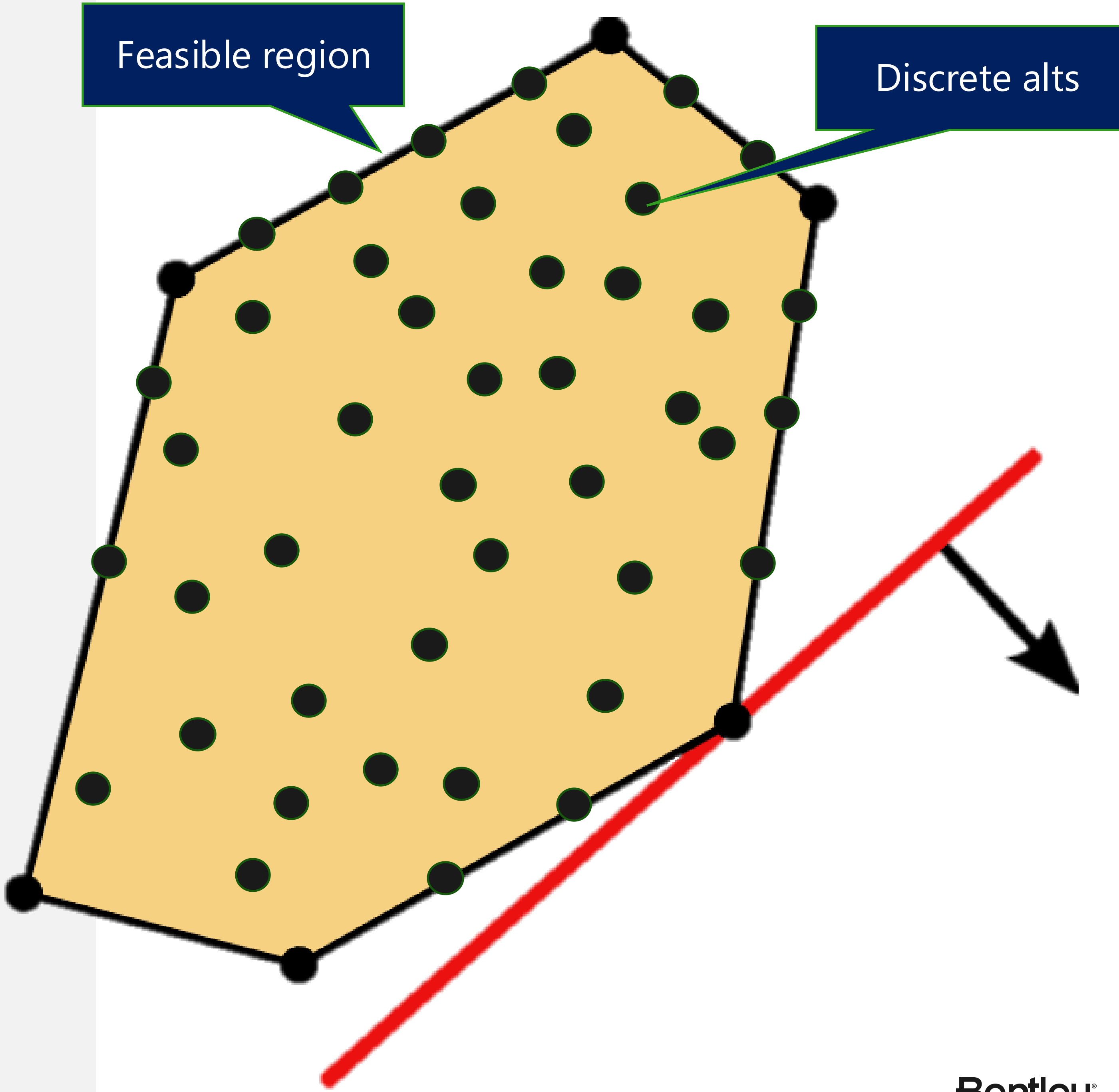
Tool	Simplicity	Efficiency	Compatibility	Data	Elasticity
Optimization	?				
ML					?
MDCEV	?		?		
Entropy					
Mixed logit	?				
Latent class				?	?
Reliability			?	?	

MoMo, 2025

AGENT

Optimization (LP)

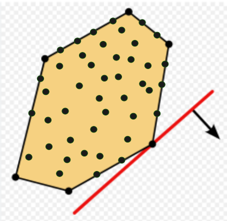
- Handling complex joint choices (avoiding enumeration of alternatives)
- Practical for:
  - Coordinated scheduling of individual activities and travel in continuous time
  - Resource allocation problems such as intrahousehold car allocation.
- Infeasible choice set size for conventional discrete choice:
  - Trip-by-trip departure time choice is meaningless due to cross-impacts
  - Millions of alternatives for daily schedule as joint choice w/1-min resolution
- Require special methods of calibration:
  - H. Mahmassani (joint publications with Bentley)
  - M. Bierlaire (OASIS, attempts of econometric estimation)
  - J. Chow (ML “inverse” optimization for calibration)



Bentley

Discrete choice model

$$\max_{i \in C} \left\{ U(i) = \left(\sum_k c_{ik} x_{ik} \right) + \varepsilon_i \right\}$$

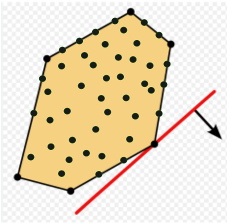


- $i \in C$ = discrete alternatives from the set
- k = explanatory variables / attributes
- x_{ik} = values of attributes
- c_{ik} = coefficients
- ε_i = random terms
- $P(i)$ = choice probability
- i^* = chosen alternative
- $x_{i^*,k} = x_k^*$ = chosen attributes

MoMo, 2025

Discrete choice model

$$\max_{i \in C} \left\{ U(i) = \left(\sum_k c_{ik} x_{ik} \right) + \varepsilon_i \right\}$$



LP

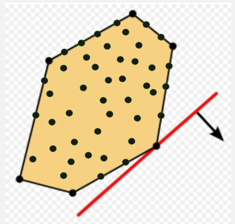
$$\max \left\{ F = \sum_k c_k x_k \right\}$$

$$\underline{A}_r \le \sum_k a_{kr} x_k \le \overline{A}_r$$

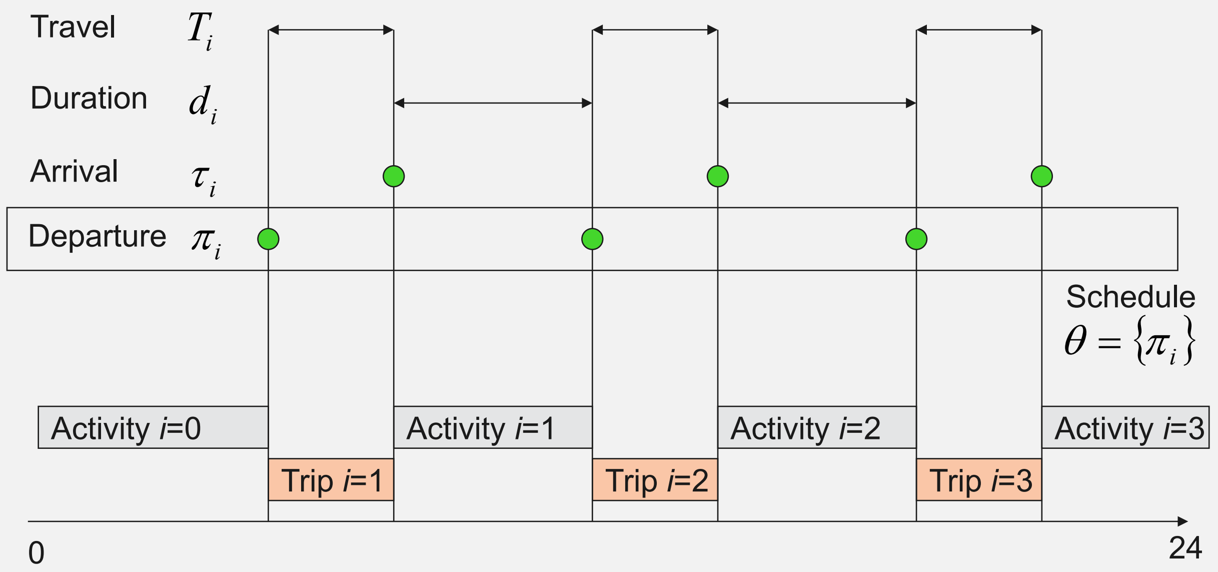
- r = constraints
- $\underline{A}_r, \overline{A}_r$ = lower and upper bounds
- x_k^* = optimal solution (chosen attributes)

Bentley

Schedule coordinator

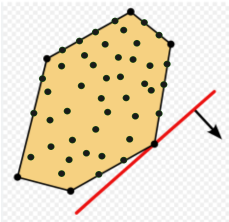


- Creates a consistent individual schedule of activities and trips in continuous time given:
  - Sequence of activities and trips for each individual
  - Travel times from the time sensitive network model
  - Original crude / inconsistent schedule from prior ABM choices
  - Joint trips of HH members
  - Seamless ABM-DTA integration



Bentley

Schedule coordinator interface



Add Model Step

Create new Import from specification

Type

Select a type...

- Generic Choice Model
- Location Choice Model
- Temporal Choice Model
- Aggregate Table
- Insert Into Table
- Delete From Table
- Table Calculator
- Schedule Coordinator

Decision-maker | Schedule attributes | Time-space constraints | Penalty expressions | Result attributes

Decision-maker: What agent(s) the schedule coordinator is applied for?

Schedule attributes: Preliminary schedule variables, departure/arrival time of trips

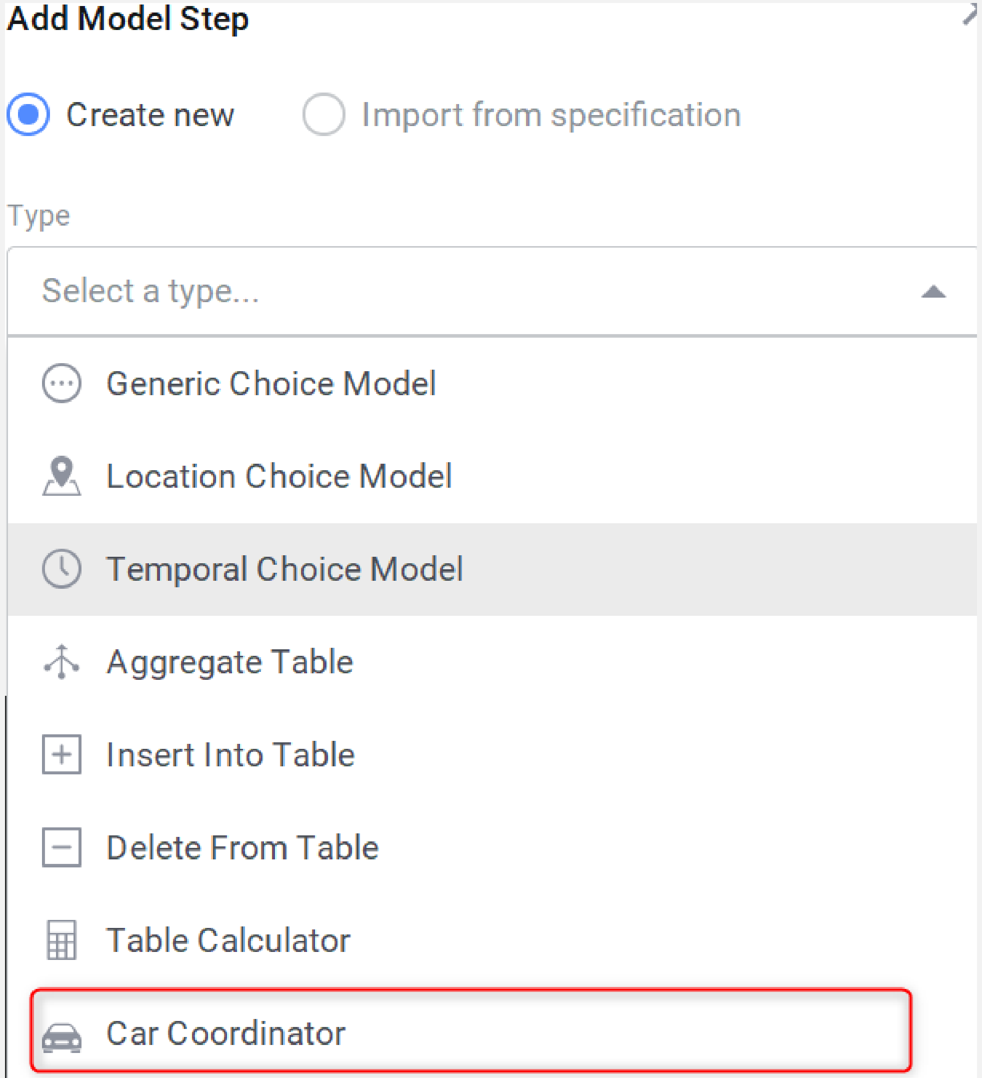
Time-space constraints: Activity duration and travel time

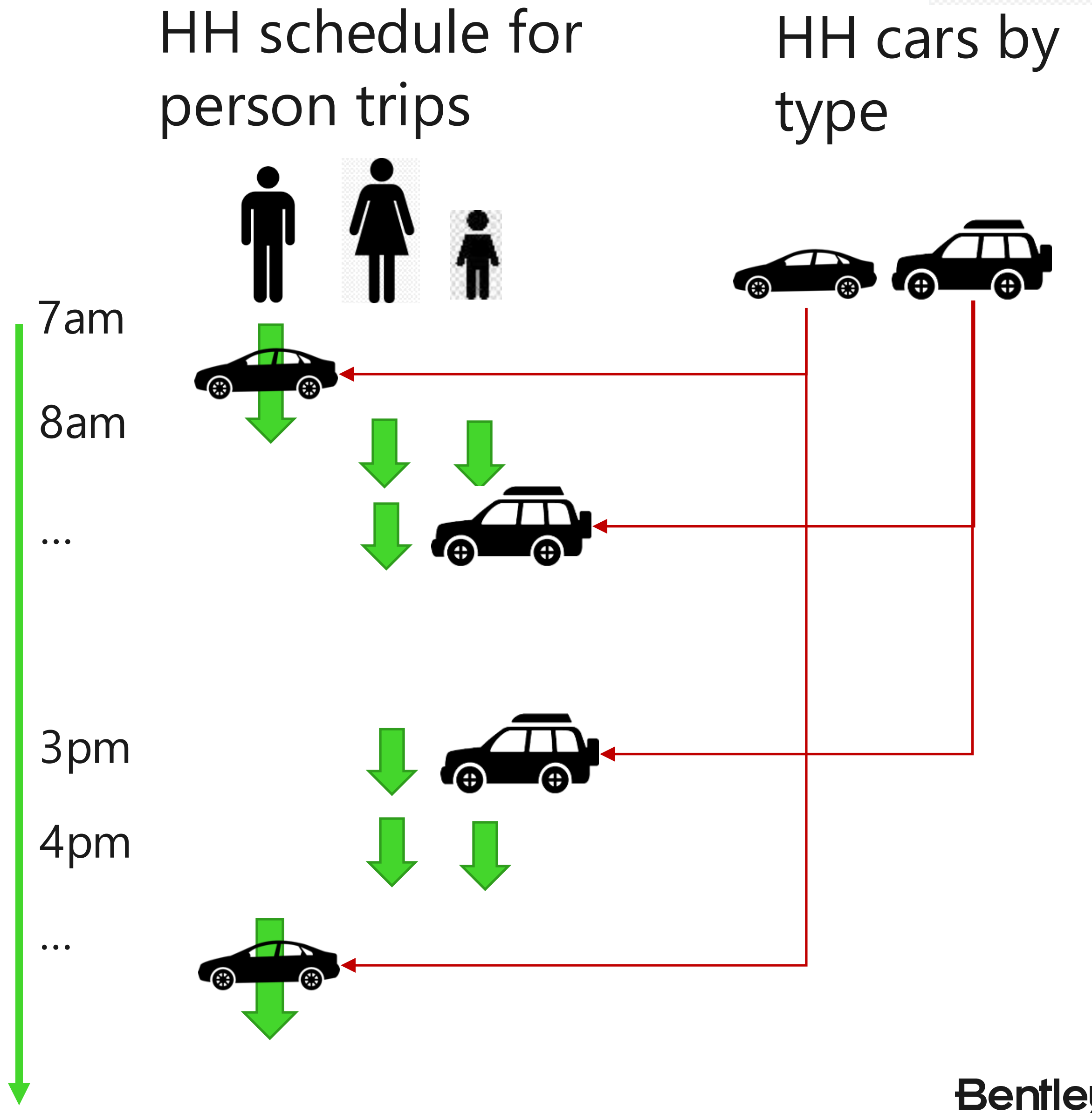
Penalty expressions: Penalties for changing schedule by tour/trip/person characteristics

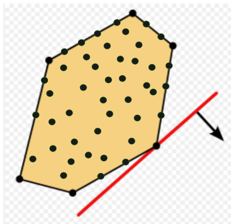
Result attributes: Result attributes to save adjusted schedule

Bentley

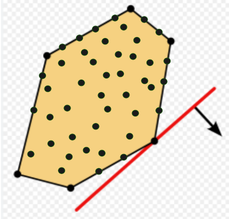
- Intrahousehold car allocation to person trips (“Car Coordinator”)
- Addresses different car types by body, fuel, level of automation, age, etc.







Solved as Mixed-Integer linear Program (MILP)



- Boolean decision variables:
  - Car ID for each auto trip
- Objective function is sum across all HH auto trips:
  - Cost for loaded trips
  - Cost for empty trips (AVs only)
  - Car allocation preferences
  - Penalty for unsatisfied demand
  - Bonus for usual driver
  - Penalty for trip departure time adjustment (if allowed)
- Constraints:
  - Car size matching party size
  - Car trips cannot overlap in time
  - Close chain of consecutive trips for each car

- LP integrates car choices for all trips: C×T Booleans
- Infeasible dimensionality if treated as joint discrete choice: C^T alts

MoMo, 2025

Bentley

Specification of car type dimensions of interest: Example

3 fuel types

- Gas/diesel
- Hybrid
- Electric (EV)

×

6 body types

- Motorcycle
- Mini
- Regular
- SUV
- Van
- Pickup

×

2 automation levels

- Regular (RV)
- Autonomous (AV)

18 car types for individual car segmentation

HH segmentation

ML/AI

- Represent a potential replacement for logit models, but this is hampered by ***complex elasticities*** and non-transparency
- Some of these methods proved to be very useful as a behavioral analysis tool
- Appealing practical direction is borrowing the ML backward propagation methods for travel model system ***calibration*** to match traffic/transit counts and/or ‘big data”.

Single choice example to compare different models

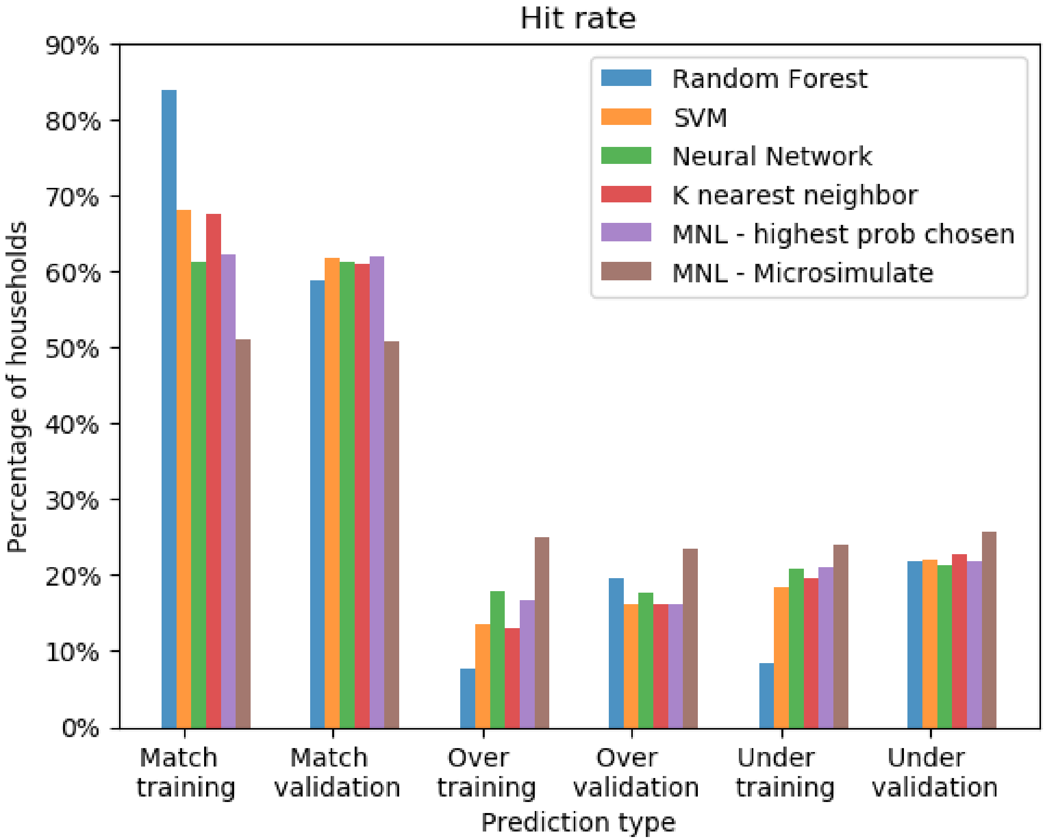
Auto ownership model: 0, 1, 2, 3+ cars

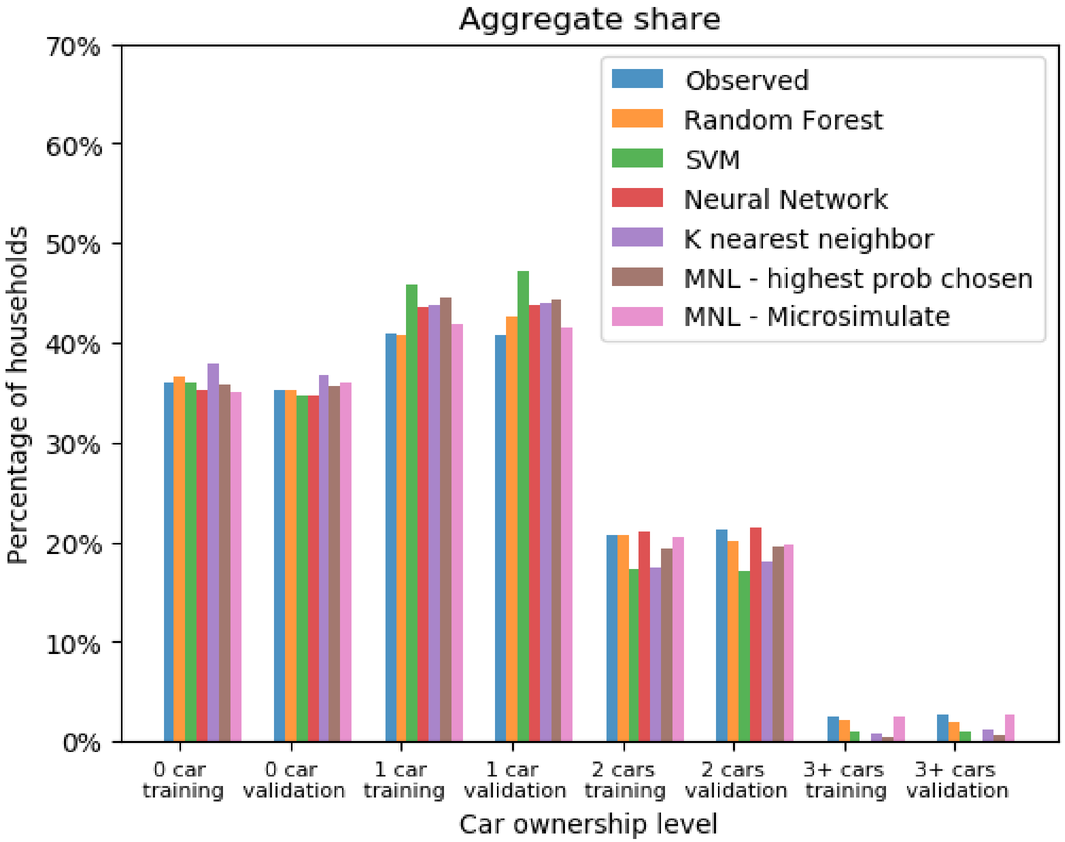
6 model applications: 4 ML + 2 MNL (highest probability chosen and microsimulation)

Model assessment criteria:

- Predictive power (individual hit & aggregate shares)
- Behavioral insights
- Policy sensitivity

Single model performance / slightly in favor of ML





MoMo, 2025

Bentley

Sensitivity tests / ML pitfalls

Transit travel time changed for the entire region

- Scenario 1 → Better transit service: all IVTs halved
- Expectation: car ownership would decrease

- Scenario 2 → Worse transit service: all IVTs doubled
- Expectation: car ownership would increase

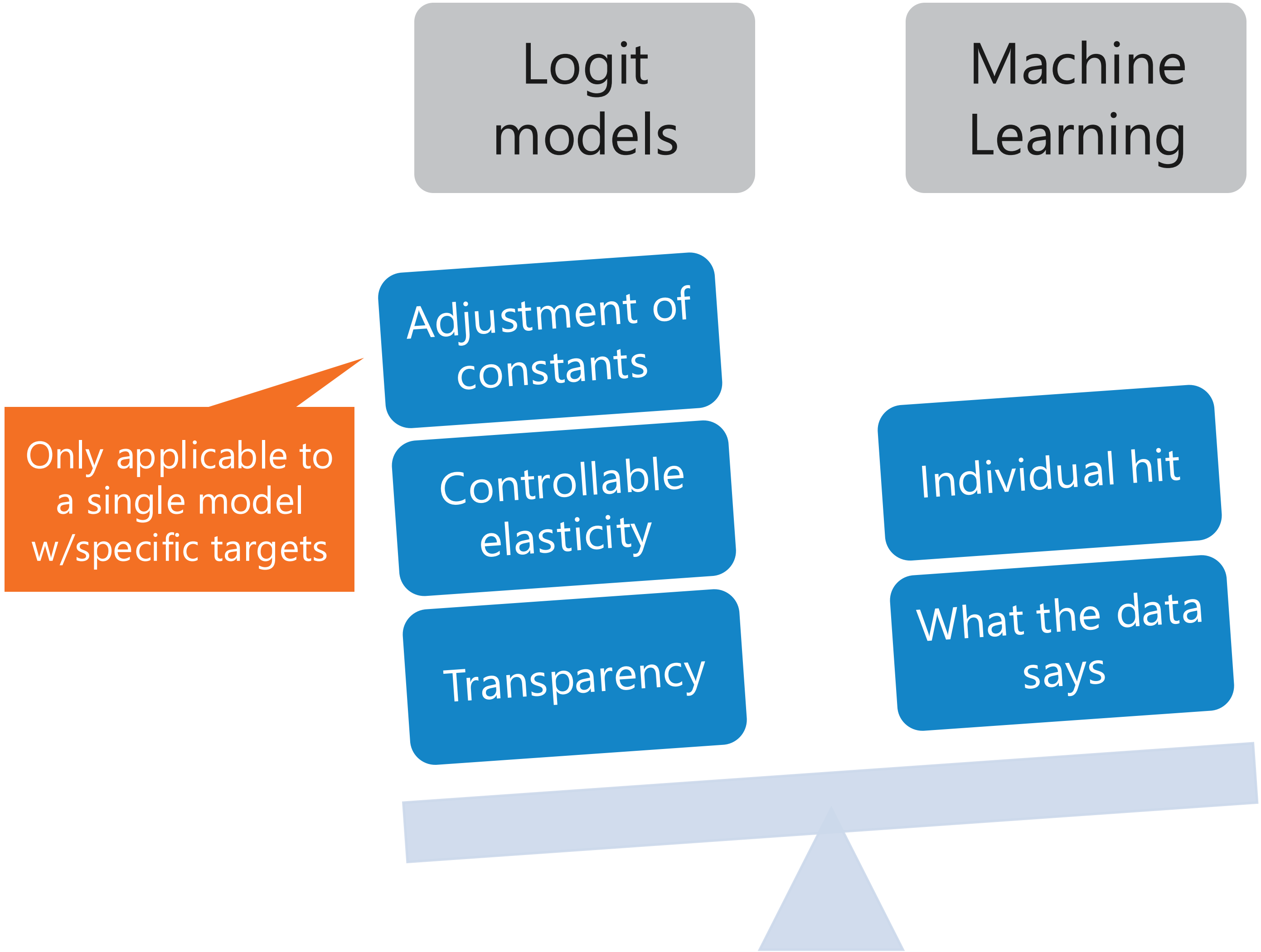
Only MNL passed these tests

- Desired elasticities guaranteed by the model structure

Practical pros and cons of logit vs. ML for single model

Hybridization is attractive direction:

- Mixed MNL-NN – the best of both worlds?



More details:

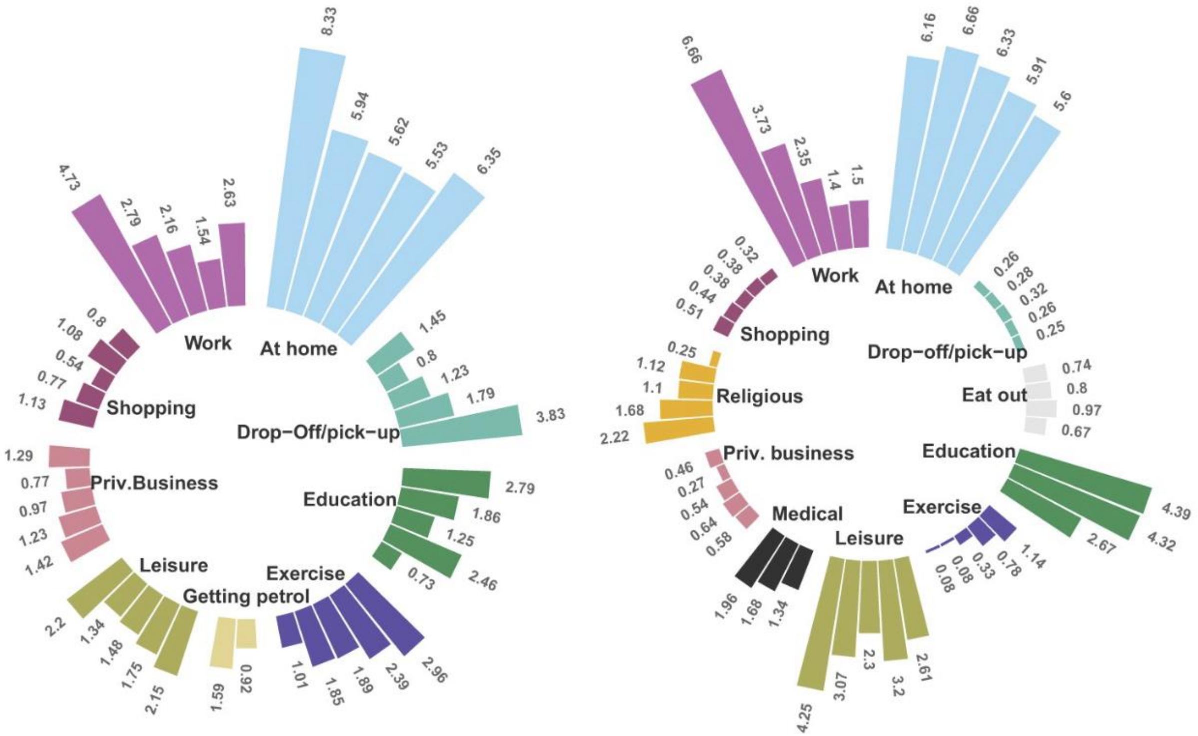
Vyas, G., P. Vovsha (2020) Assessment of Machine Learning methods versus logit models for travel modelling in practice.

Presented at the 99th TRB Annual Meeting, Washington, D.C.

Start of bonus slides for 30-min presentation

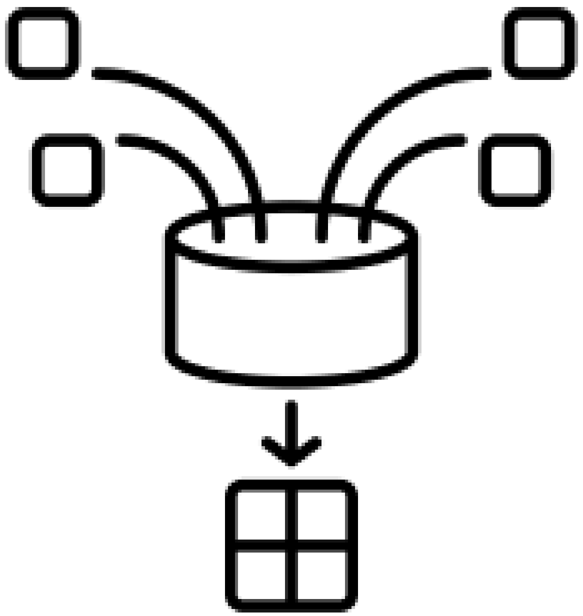
MDCEV

- MDCEV class of models allows for an ***integrated generation of individual activity participation under the time/budget constraints*** without a combinatorial explosion pertinent to discrete choice models
- Adoption of MDCEV in mainstream ABM until recently was hampered by a ***complex iterative application algorithm*** and somewhat ***incompatible ABM framework*** that operates with discrete activity/travel episodes rather than continuous time allocation
- ***Simple application algorithm*** was developed to overcome the first problem, but the second one still remains with the existing ABM structure
- Promising future direction can be restructuring the ABM as a ***time use*** model



Entropy max

- Recognition of imperfect / inconsistent data by allowing ***relaxations***:
  - spatial trip distribution: doubly constrained and singly constrained gravity models are particular cases
  - population synthesis
  - survey expansion
  - consolidation of different data sources



General formulation of balancing

- Inputs:
  - Seed structure to preserve to the max extent
  - Controls to match
- Method:
  - Iterative balancing
- Naive IPF may not work:
  - Inconsistent controls infeasible to match exactly
  - Non-convergent process depending on when it is stopped
  - ***Extended balancing formulation w/relaxations*** is solution
  - Simple and efficient implementation

General balancing formulation (population synthesis example)

Variables:
x_{nz} **= synthetic population (HHs from the sample expanded and allocated to TAZs)**
$\gamma_{iz}, \lambda_{jd}, \rho_k$ **= relaxation factors for controls**
Objective function:
$\min_{\{x_{nz},\gamma_{iz},\lambda_{jd},\rho_k\}} \left\{ \begin{array}{l} G = \sum_{nz} x_{nz} \times \left(\ln \frac{x_{nz}}{w_{nz}} - 1\right) + \sum_{iz} \mu_i \times \gamma_{iz} \times (\ln \gamma_{iz} - 1) \\ + \sum_{jd} \nu_j \times \lambda_{jd} \times (\ln \lambda_{jd} - 1) + \sum_k \eta_k \times \rho_k \times (\ln \rho_k - 1) \end{array} \right\}$
Subject to (relaxed constraints):
$\sum_n a_{ni} \times x_{nz} = A_{iz} \times \gamma_{iz}$
$\sum_{n,z \in Z_d} b_{nj} \times x_{nz} = B_{jd} \times \lambda_{jd}$
$\sum_{n,z \in Z} c_{nk} \times x_{nz} = C_k \times \rho_k$
$x_{nz}, \gamma_{iz}, \lambda_{jd}, \rho_k \geq 0$

MoMo, 2025

Mixed logit

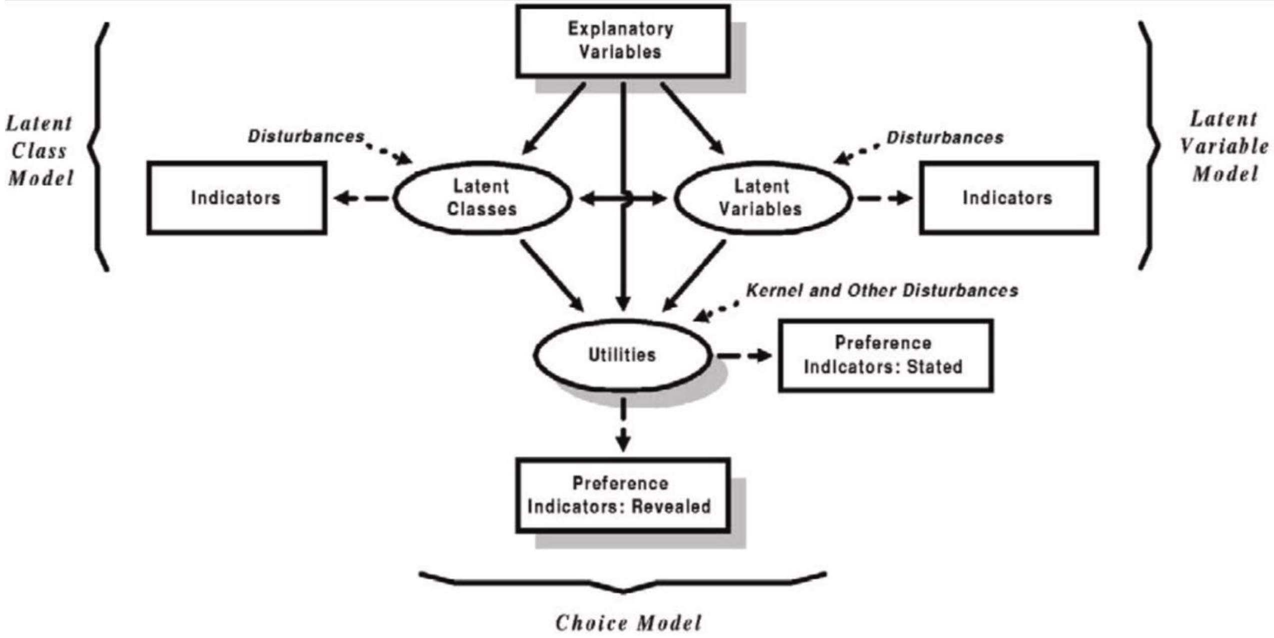
- Behaviorally appealing randomized coefficients ("taste variation").
- Very difficult in estimation with real-world data due to the ***statistical identification*** problems
- Important advantage for accounting of ***distributed Value of Time (VOT)*** for toll road forecasting. Replacing a distributed individual VOT with crude average values results in systematic biases in traffic and revenue forecasting (presentation)

$$P\left(y_{nt} = j \middle| \boldsymbol{X}_{ntj}, \boldsymbol{\beta}_{nt}, \right) = \frac{e^{V\left(\boldsymbol{X}_{ntj}, \boldsymbol{\beta}_{nt}\right)}}{\sum_{j' \in \mathscr{C}} e^{V\left(\boldsymbol{X}_{ntj'}, \boldsymbol{\beta}_{nt}\right)}},$$

$$\boldsymbol{\beta}_{nt} \sim \mathrm{N}\left(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_W\right)$$

Latent class

- Quite complex in model estimation these models offer unique insights into the ***hidden segmentation*** of individuals:
  - Segmentation becomes a part of the model
- The main stumbling block in practice proved to be ***policy sensitivity of the class membership component*** that in many cases resulted in a very constrained model response.



From Walker and Ben-Akiva, 2003

Bentley

Travel time reliability

- For quantification of travel time reliability, behaviorally appealing methods were suggested such as "mean-variance" and "schedule delay" concept
- Adoption in forecasting models proved to be problematic because of the ***required input*** such as preferred arrival time that is unobserved

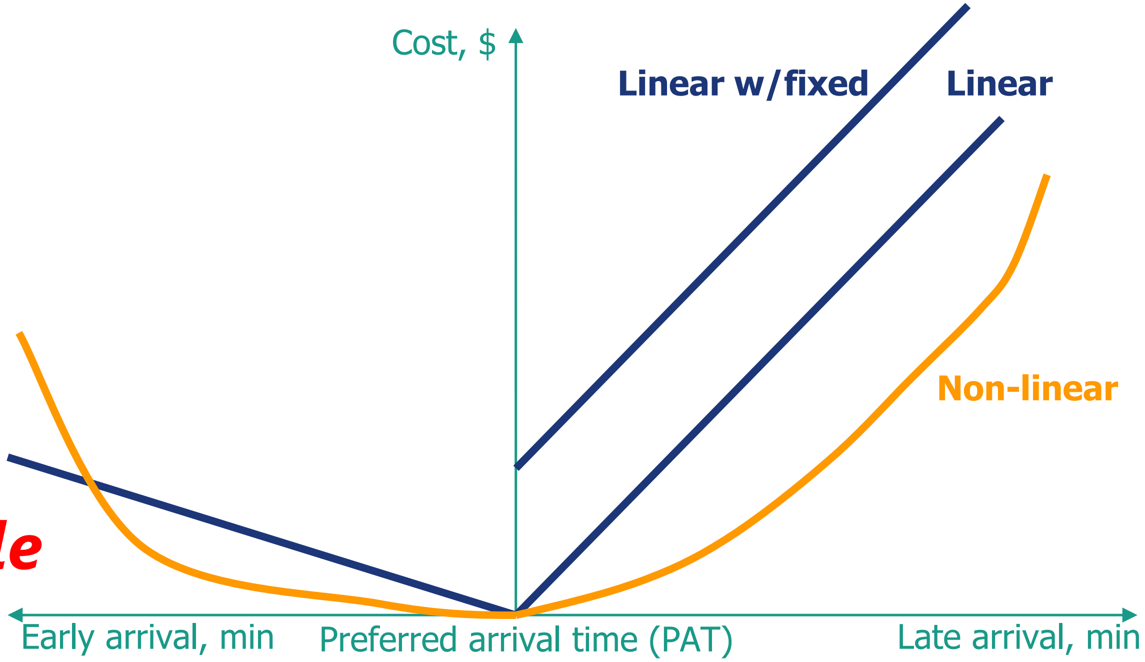
Bentley

Method 1: fundamental challenge resolved by simplification

- Random Utility Model (RUM): $U = \alpha \times \mathrm{Time} + \beta \times \mathrm{Cost} + \ldots + \varepsilon$
- Random coefficients RUM: $U = \alpha \times \mathrm{Time} + \beta \times \mathrm{Cost} + \ldots + \varepsilon$
- Random travel time is different:
  - $U = \alpha \times \mathrm{Time} + \beta \times \mathrm{Cost} + \ldots$ \\ difficult to handle
  - $U = \alpha \times \mathrm{Time} + \beta \times \mathrm{Cost} + \gamma \times \mathrm{SD} \ldots + \varepsilon$ \\ back to familiar RUM
  - $U = f(\mathrm{Time}) + \beta \times \mathrm{Cost} + \ldots$ \\ Risk or prospect theory

Method 2: schedule delay cost

- $\boldsymbol{U = \alpha \times T + \beta \times SDE + \gamma \times SDL + \delta \times L}$
- In presence of random travel times:
  - *f(T)* – travel time distribution
  - *E(U)* – expected utility dependent on *f(T)* and departure time/*PAT*
  - Improvement of reliability in *f(T)* is evaluated by *E(U)*
- Considerable body of literature:
  - SP estimates: $\gamma >> \alpha$ & $\gamma >> \beta$
- Behaviorally appealing but ***not compatible with the travel forecasting framework***



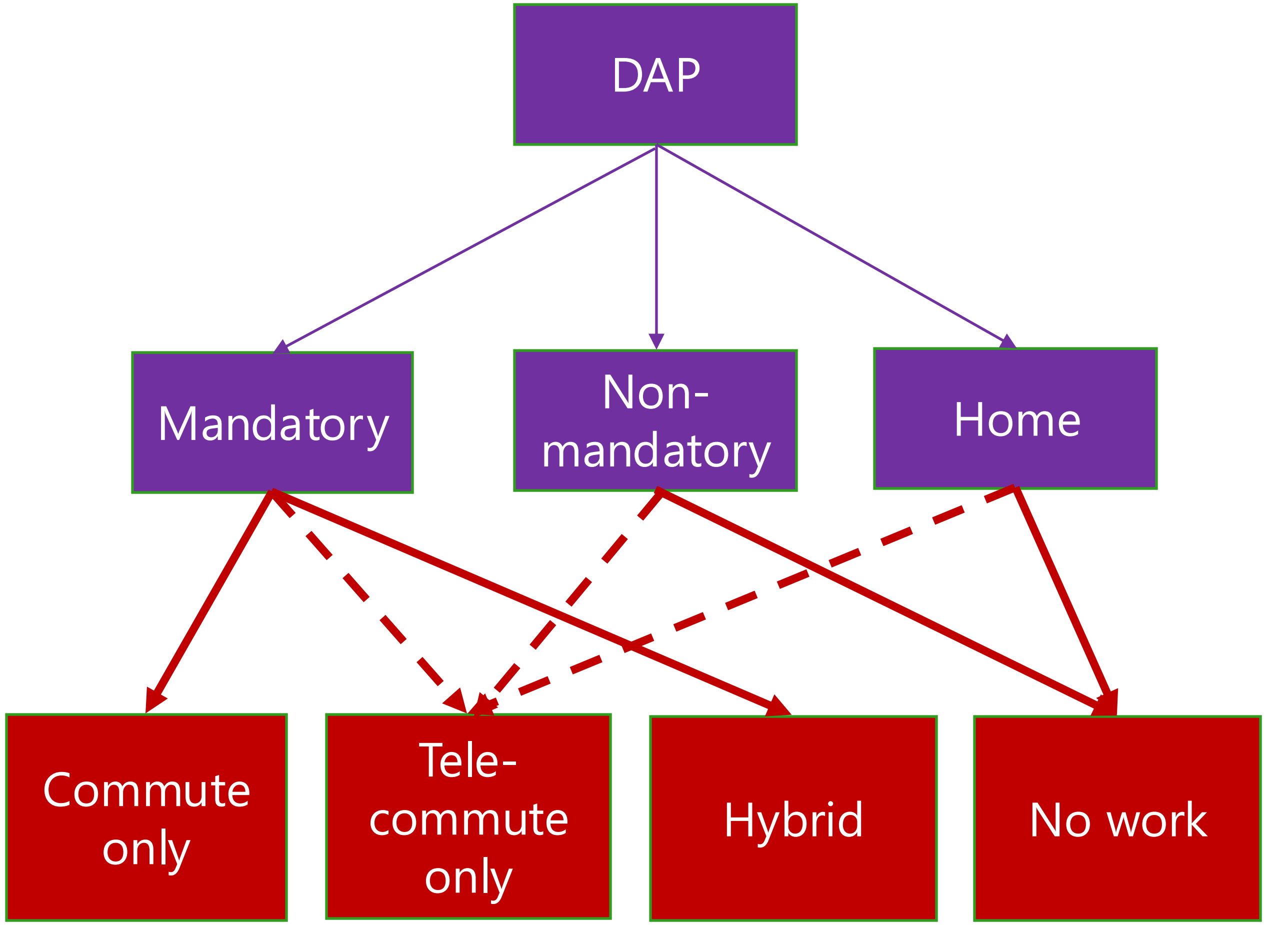
 |

Bentley

End of bonus slides for 30-min presentation

Typical cycle of innovation adoption (CDAP)

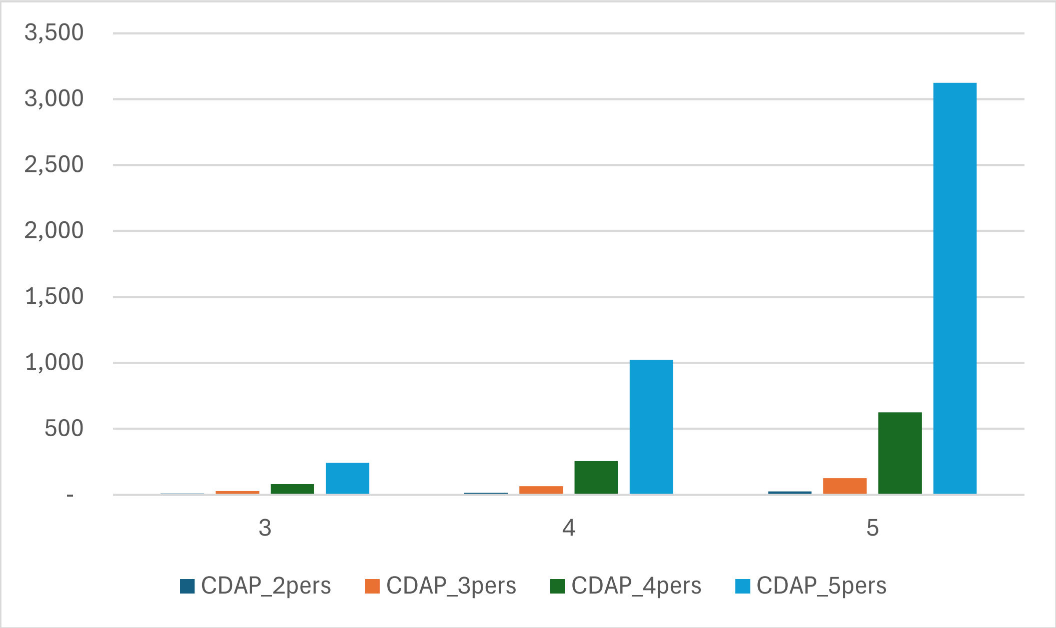
- First introduced in CT-RAMP1 MORPC ABM in **2005**:
  - Explodes with number of persons but still practical up to 5 persons in HH 3^5=243 alts
  - Accounts for intrahousehold interactions
  - Captures major features of individual daily activity pattern
  - Streamlines the subsequent sequence of sub-models
- Included in almost all ABMs in practice in **2005-2025** (CT-RAMP1/2, ActivitySim, TourCast)
- From **2025** becoming a limitation with the growing share of work from home:
  - Need an extension of DAP types



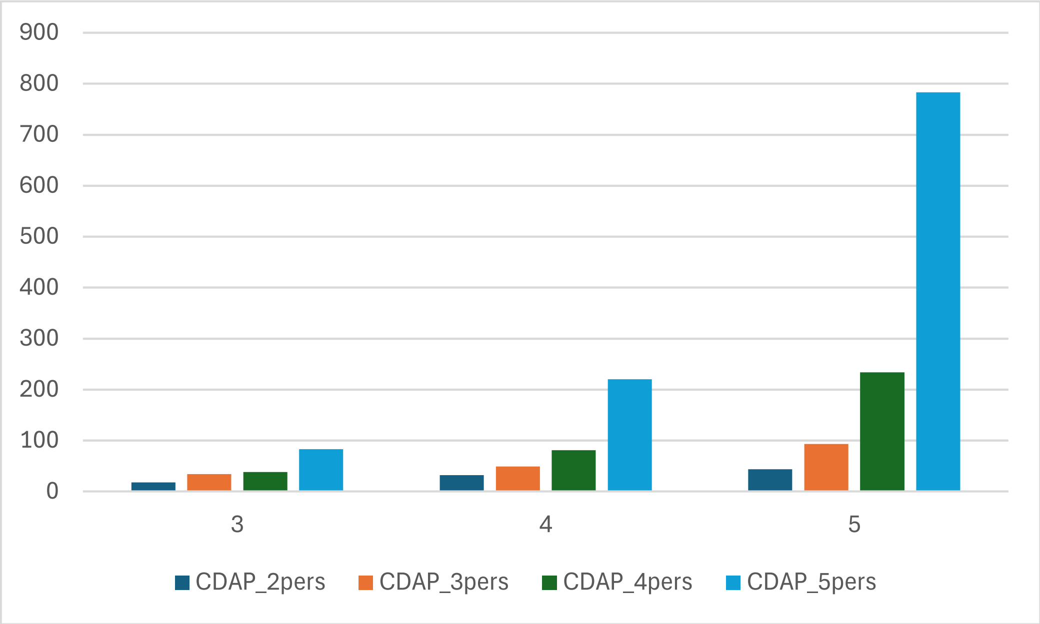
 |

Efficient implementation of CDAP for explicit modeling of telecommuting details

- Number of CDAP choice alts as function of number of DAP alts and HH size



- Runtime, sec (765K HHs, Perth, Western Australia; Intel(R) Xeon(R) Platinum 8370C CPU @ 2.80GHz 2.79 GHz, 128GB RAM)



Runtime increases but still feasible

General conclusions

Adoption of innovations rarely happens automatically, it requires an effort

Practice dictates many constraints that need to be recognized

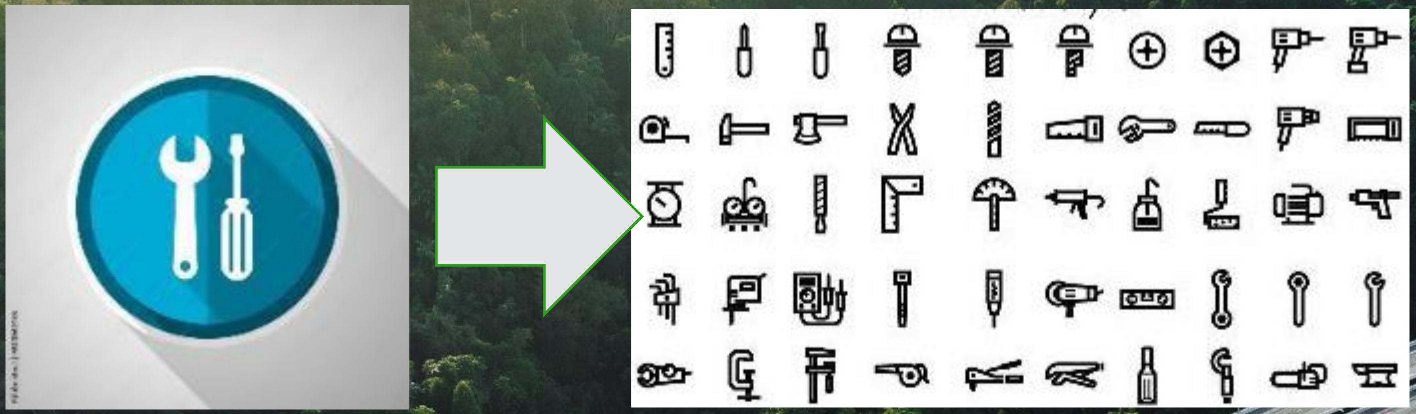
Not the best models win, but the most compatible with the ABM structure and supporting the business process at agency

Encourage R&D that addresses adoption in practice; bridging academic research and practical applications, "mixed" funding

Questions?
Thank you!

Peter.Vovsha@Bentley.com

Bentley®



Extending the toolbox of travel modeling in practice

Peter Vovsha, Gaurav Vyas, Jim Hicks, *Bentley Systems*

Bentley®