Evidence-Based Transportation Demand Analysis

Disaggregate Demand Modelling

Khandker Nurul Habib, PhD. PEng

Outline

- Disaggregate Demand: Choice versus Demand
- Choice Theory
- Discrete Choice Model

Choice versus Demands

- Travel demand is derived from individuals' choices → influenced by behaviour

- Demand model considering travel behaviour should be:
  - **Descriptive (not prescriptive)**: how human beings behave, not how they should/ought to behave.
  - **Abstract**: that can be formalized in general cases not specific to particular circumstances.
  - **Operational**: can be applied to develop models with variables and parameters that can be observed and estimated

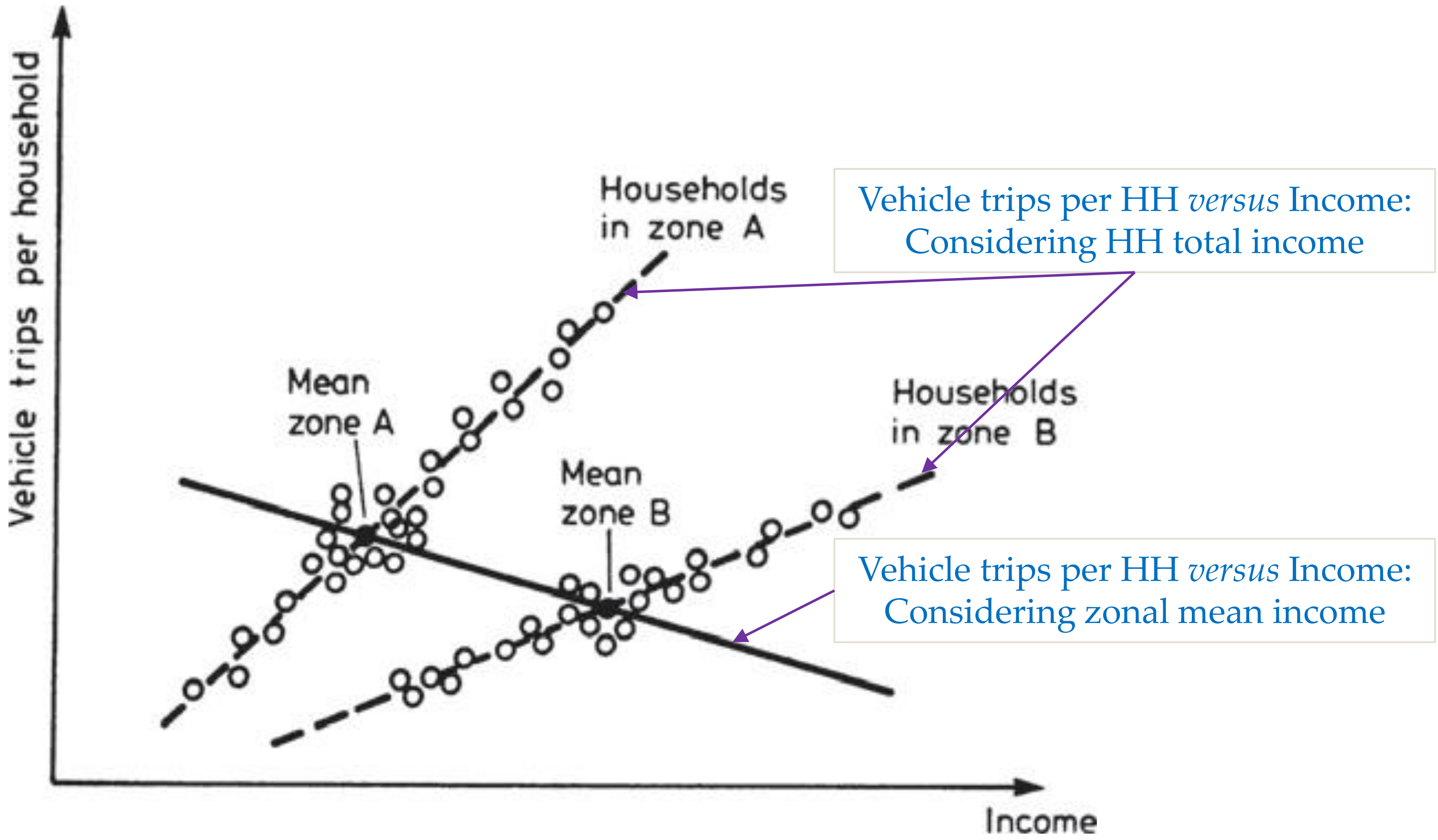
Framework of Choice Theory

- A choice as outcome of sequential decision-making process that includes:
  1. Definition of the choice problem
  2. Generation of alternatives
  3. Evaluation of attributes of alternatives
  4. Choice making
  5. Implementation
- A specific theory of choice is a collection of procedure that defines **the elements**:
  1. Decision maker
  2. Alternatives
  3. Attributes of alternatives
  4. Decision rules

Decision Maker in Demand Model

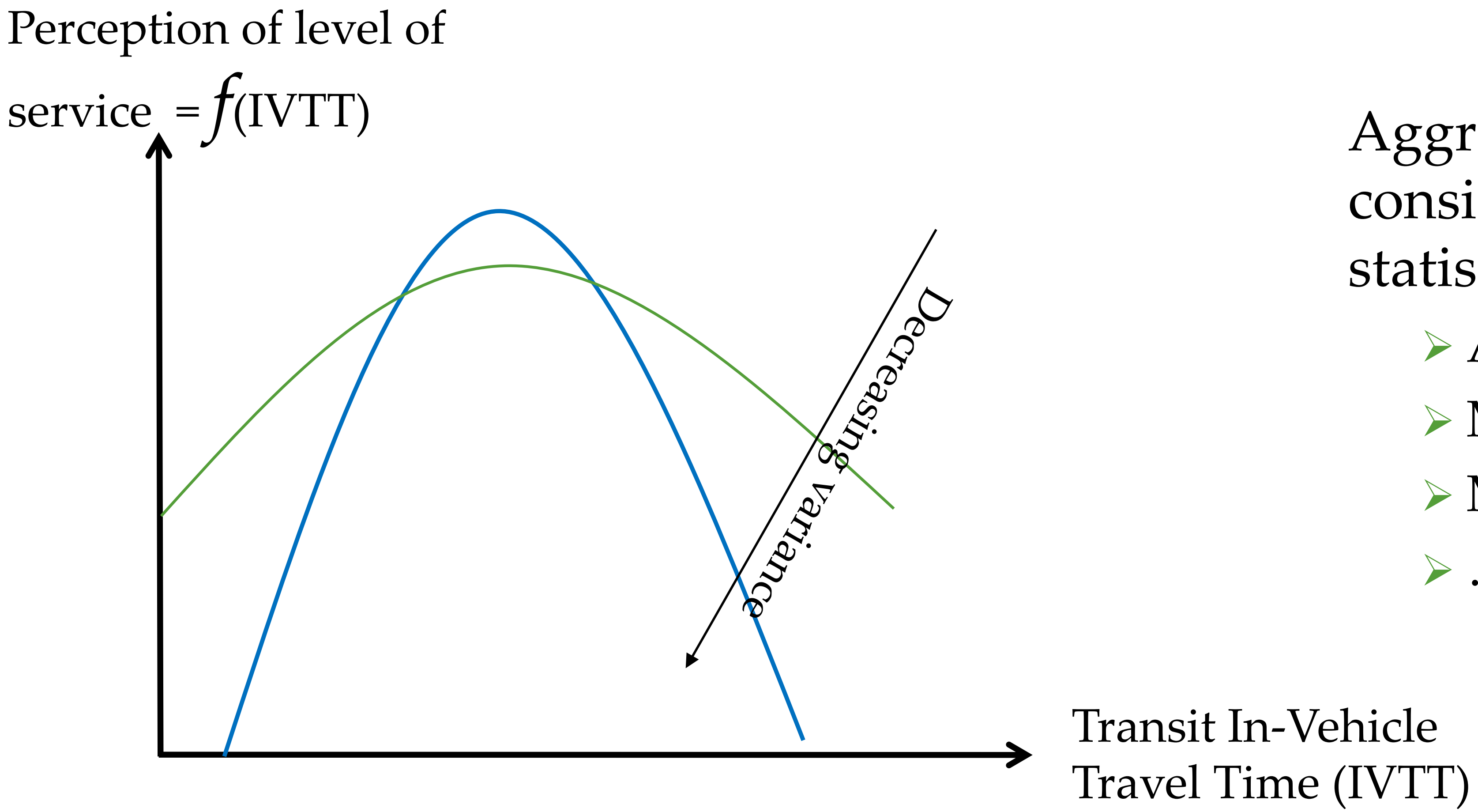
- **Aggregate**: Collective information
  - Number of trips generated by a zone
  - Number of people using bus
  - Number of people choosing a specific route
- **Disaggregate**: Individual decision maker level information
  - What is Mr x choosing to do.
- **Heterogeneous Decision maker**: Wide variety in choice behaviour across the population:
  - Induced from different individual attributes: socio-economic conditions, etc.
- **Aggregation bias**: biased induced for overlooking decision makers' heterogeneity and non-linearity of response:
  - Response function is non-linear with attributes

Aggregation Bias



- **Aggregation bias**: biased induced for overlooking decision makers' heterogeneity and non-linearity of response:
- Ecological Fallacy: Relationship changes between 'zonal mean income' *and* 'household total income'

Aggregation Subsides Variance/Heterogeneity



Aggregation requires considering sample statistics:

- Average
- Median
- Mode
- …..

Disaggregate Travel Demand

- Modelling travel demands:
  - Disaggregate: better capture systematic heterogeneity
  - Discrete: better define choice process and choice making behaviour
  - Random: better capture unobserved heterogeneity
  - Behavioural Theory: Relationship without underlying theory is meaningless
- An individual decision maker '*i*' who must chose one alternative '*j*' from a set of feasible alternative set 'C_i' (alternatives are mutually exclusive):
  - $P_i(j \mid c_i)$ = the probability that the person 'i' chooses '*j*' from the choice set c $= f($*attributes of alternatives,*

    attribute of the decision maker

    attribute of the choice context

Discrete Choice Model

- Concept of random utility: Decision maker makes choices to maximize utility.
  1. Observed inconsistency in choice behaviour is mainly due to our observational deficiency.
  2. Utilities of alternative are not constant or not known to us with certainty.
  3. The choice probability of an alternative *j* to a person t is equal to the probability that the utility of the alternatives, U_j is greater than or equal to the utilities of all other alternatives in the choice set.

$$U_j = V_j + \varepsilon_j \qquad j \in C_i$$

Total Utility

Systematic utility

Random utility

Choice set

$$P_i(j) = \Pr\left[(U_j) \geq \max_{\substack{j \in C_t \\ i \neq j}} \left(U_j\right) \right]$$

$$P(j) = \Pr\left[(V_j + \varepsilon_j) \geq \max_{\substack{k \in C_i \\ k \neq j}} \left(V_k + \varepsilon_k\right) \right]$$

➢ Choice model (Probability function of choosing an alternative) is defined by the assumption of ε

Discrete Choice Model

- Distributional assumption of random utility component, ε
  - Normal Distribution: Probit model
  - Type I Extreme Value Distribution: Logit model
  - Generalized Extreme Value Distribution: GEV model
- Difference between utilities of two alternatives are relevant: Absolute value of utility function has little to no meaning
- Marginal rate of substitution of two variables influencing the choice of discrete choice alternative = ratio of coefficients of the variables in the utility function
- Consumer surplus = social welfare = expected maximum utility of all alternatives in the choice set
- Elasticity = rate of change of probability with respect to rate of change in variable values

Discrete Choice for 2 alternative: Binary

- Binary Logit model: $P_i(j) = \dfrac{1}{1+e^{\mu(V_k - V_j)}} = \dfrac{1}{1+e^{-\mu(V_j - V_k)}} = \dfrac{e^{\mu V_j}}{e^{\mu V_j}{}' + e^{\mu V_k}}$
- Binary Probit model: $P_i(j) = \Phi\left(\left(V_k - V_j\right)/\sigma\right)$

μ: Scale parameter

σ is the variance

- In either model, the systematic utility function: Linear-in-parameter function

$$V_j = \beta_{0j} + \beta_1 x_{1j} + \beta_2 \log(x_{2j}) + \beta_3 \exp(x_{4j}) + \ldots.$$

- Direct elasticity of alternative *j* with respect to x_j (Logit model) $= \beta x_j \left(1 - P_i(j)\right)$
- Cross elasticity of alternative *j* with respect to x_k (Logit model) $= -\beta x_j P_i(k)$
- Model parameters can be estimated by maximum likelihood estimation

Discrete Choice for multiple alternatives

- Multinomial Logit model (*μ is the scale parameter)*

$$P_i(j) = \frac{e^{\mu V_j}}{e^{\mu V_j} + e^{\mu V_k} + \ldots\ldots} = \frac{e^{\mu V_j}}{\sum_k e^{\mu V_k}}$$

- The systematic utility function: Linear-in-parameter function

$$V_j = \beta_j + \beta_1 x_{1j} + \beta_2 \log(x_{2j}) + \beta_3 \exp(x_{4j}) + \ldots.$$

- Direct elasticity of alternative *j* with respect to x_j (of alternative *k*) $= \beta x_j \left(1 - P_i(j)\right)$

- Cross elasticity of alternative *j* with respect to x_k (of alternative *k*) $= -\beta x P_i(k)$

- Model parameters can be estimated by maximum likelihood estimation

Issues with MNL: Independent and Irrelevant Alternatives- IIA

- Cross elasticities of any variable/attribute of alternative j (x_j) are equal

$$\text{Cross elasticity of } x_k \text{ on the choice of alternative } j = -\beta\, x_k P_i(k)$$

- Example: three alternatives for inter-city mode choice: Car, Transit, Air
  - Increase in parking cost for car will increase choice prob of Air and Transit equally
  - Increase in air fare will increase the choice prob of Car and Transit equally
  - Increase of transit fare will increase the choice prob of Car and Air equally
- IIA can also lead to serious over-prediction of choice alternatives in the choice set if they have overlapping properties:
  - Subset of alternatives has common attribute

Miss-Prediction if IIA are not Valid

- Consider 2 alternatives and assume they have exactly same properties

$$V_{car} = V_{bus} = v \qquad P(car) = P(bus) = \frac{1}{1+\exp(v-v)} = \frac{1}{1+\exp(0)} = \frac{1}{2}$$

- Paint half of the busses Red and half of the busses Blue: No change in services

$$V_{car} = V_{red\ bus} = V_{blue\ bus} = v$$

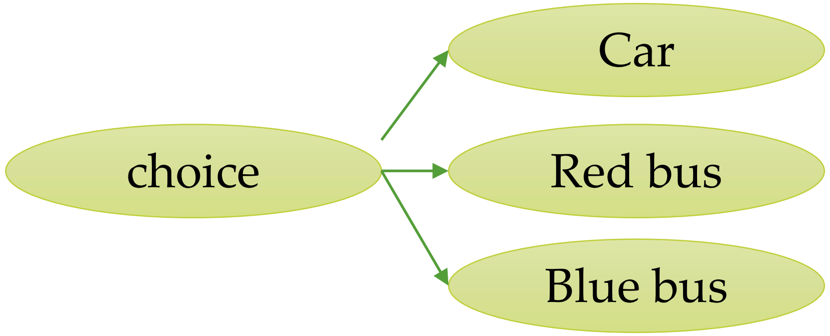
- Consider red and blue busses are separate alternative (even though they are not different in services) and use MNL

$$P(car) = P(blue\ bus) = P(red\ bus) = \frac{\exp(v)}{\exp(v)+\exp(v)+\exp(v)} = \frac{1}{3}$$

  - Car choice probability is under-predicted (it should be ½ as no new services are added)
  - Blue bus and red bus should have equal probability summed upto ½ , but the MNL over predicts to (2/3)

Overcoming Effects of IIA

- Overlooking overlapping properties of alternatives:

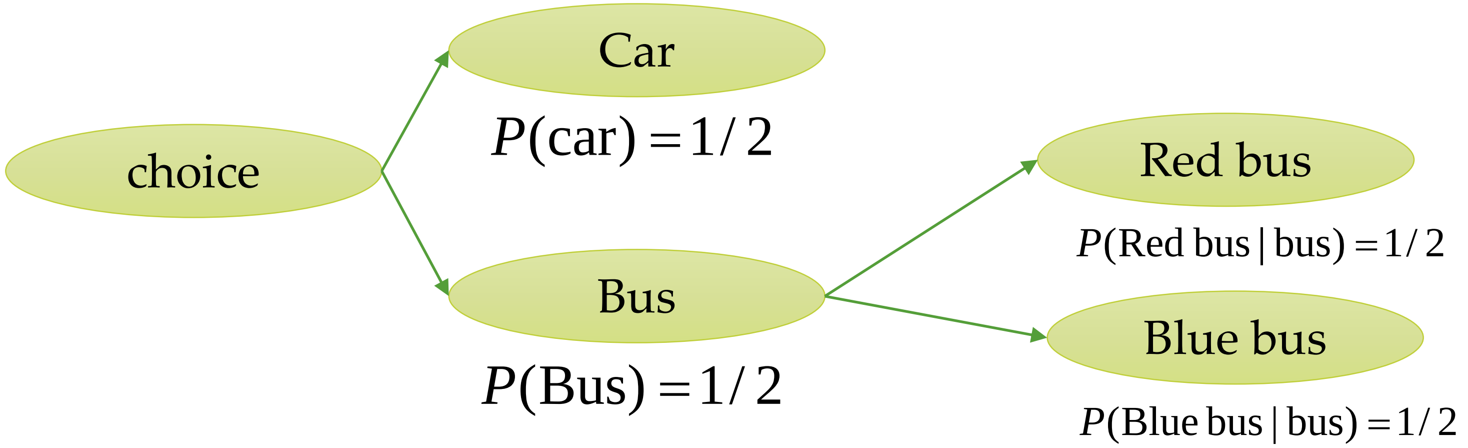


$$P(\text{car}) = 1/3$$

$$P(\text{Red bus}) = 1/3$$

$$P(\text{Blue bus}) = 1/3$$

- Considering the overlapping of properties of alternatives by nesting



$$\left.\begin{array}{l} P(\text{car}) = 1/2 \\ P(\text{Red bus}) = 1/2 \times 1/2 = 1/4 \\ P(\text{Blue bus}) = 1/2 \times 1/2 = 1/4 \end{array}\right.$$

Overcoming IIA: Nested Logit (NL) Model

Binary case between Car & Bus

$$P(\text{Car}) = \frac{1}{1+\exp(-\mu_1(v_{Bus} - v_{Car}))}$$

$$P(Bus) = 1 - P(Car)$$

Binary case between Red bus & Blue bus for the Bus nest

$$P(\text{Red bus} \mid \text{Bus}) = \frac{1}{1+\exp(-\mu_2(v_{\text{Blue bus}} - v_{\text{Red bus}}))}$$

$$P(\text{Blue bus} \mid \text{Bus}) = 1 - P(\text{Red bus})$$

Here, the utility of Bus nest, v_{Bus} will be a composite function of $v_{Red\ bus}$ & $v_{Blue\ bus}$

$$P(\text{Car}) = \frac{1}{1+\exp(-\mu_1(v_{Bus} - v_{Car}))}$$

$$P(Bus) = 1 - P(Car)$$

$$v_{car} = v$$

$$v_{car} = -\frac{1}{\mu_2}\log(e^{-\mu_2 v_{\text{Red bus}}} + e^{-\mu_2 v_{\text{Blue bus}}})$$

khandker.nurulhabib@utoronto.ca

Overcoming IIA: Nested Logit (NL) Model

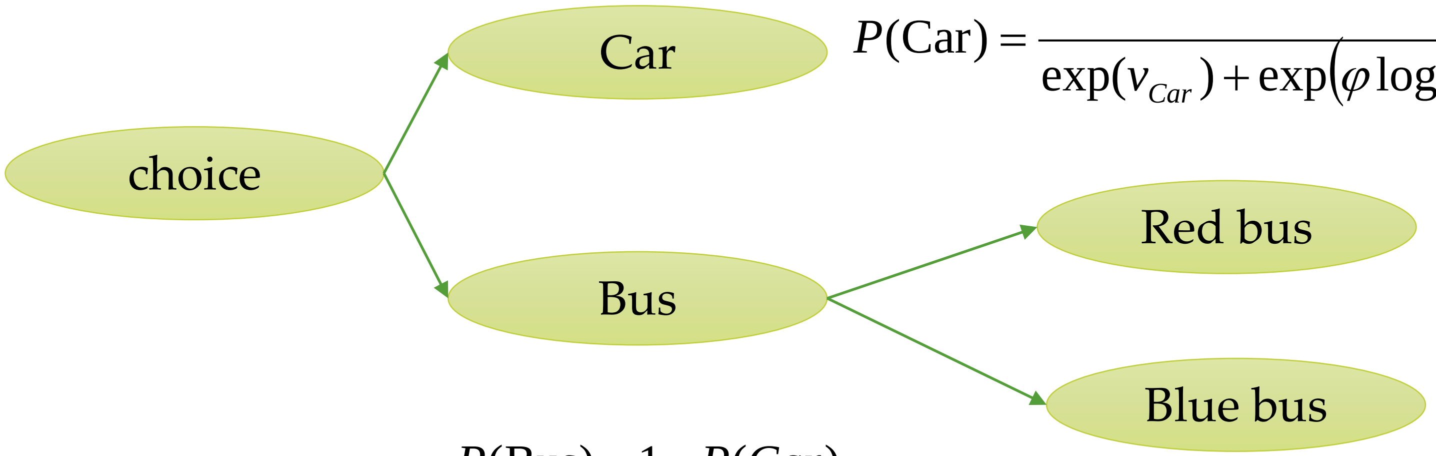
$$P(\text{Car}) = \frac{1}{1+\exp(-\mu_1(v_{Bus} - v_{Car}))} = \frac{\exp(\mu_1 v_{Car})}{\exp(\mu_1 v_{Car}) + \exp\left(\frac{\mu_1}{\mu_2}\log\left(\exp(\mu_2 v_{\text{Red Bus}}) + \exp(\mu_2 v_{\text{Red Bus}})\right)\right)}$$

Normalizing μ_1=1 and assuming $\varphi = \mu_1/\mu_2$

$$P(\text{Car}) = \frac{\exp(v_{Car})}{\exp(v_{Car}) + \exp\left(\varphi\log\left(\exp(v_{\text{Red Bus}}/\varphi) + \exp(v_{\text{Red Bus}}//\varphi)\right)\right)} \qquad P(\text{Bus}) = 1 - P(Car)$$

$$P(\text{Red bus} \mid \text{Bus}) = \frac{\exp(v_{\text{Red Bus}}/\varphi)}{\exp(v_{\text{Red Bus}}/\varphi) + \exp(v_{\text{Blue Bus}}/\varphi)} \qquad P(\text{Blue bus} \mid \text{Bus}) = 1 - P(\text{Red bus} \mid \text{Bus})$$

Overcoming IIA: Nested Logit (NL) Model



$$P(\text{Car}) = \frac{\exp(v_{Car})}{\exp(v_{Car}) + \exp\left(\varphi \log\left(\exp(v_{\text{Red Bus}} / \varphi) + \exp(v_{\text{Red Bus}} // \varphi)\right)\right)}$$

$$P(\text{Red bus} \mid \text{Bus}) = \frac{\exp(v_{\text{Red Bus}} / \varphi)}{\exp(v_{\text{Red Bus}} / \varphi) + \exp(v_{\text{Blue Bus}} / \varphi)}$$

$$P(\text{Blue bus} \mid \text{Bus}) = 1 - P(\text{Red bus} \mid \text{Bus})$$

$$P(\text{Bus}) = 1 - P(Car)$$

For $v_{car} = v_{Bus} = v_{Blue\ bus} = v_{Red\ bus} = v$, P(Car) should be ½

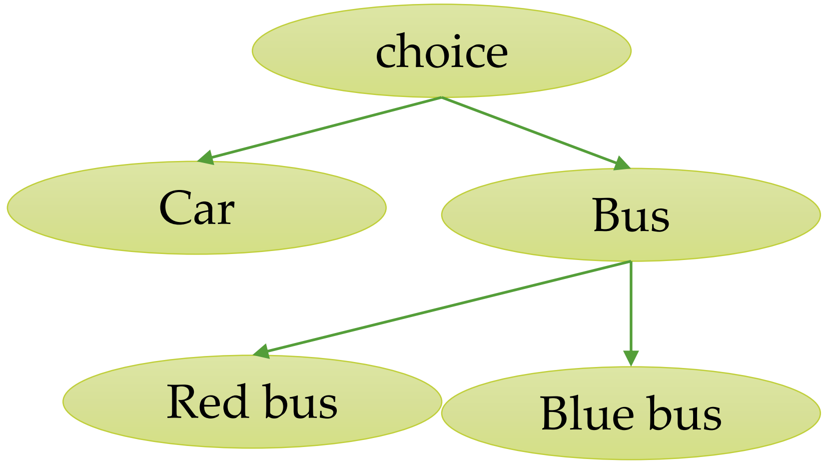
$$P(\text{Car}) = \frac{\exp(v_{Car})}{\exp(v_{Car}) + \exp\left(\varphi \log\left(\exp(v_{\text{Red Bus}} / \varphi) + \exp(v_{\text{Red Bus}} // \varphi)\right)\right)} = \frac{\exp(v)}{\exp(v) + \exp\left(\varphi \log\left(2 \exp(v / \varphi)\right)\right)}$$

Overcoming IIA: Nested Logit (NL) Model

$$P(\text{Car}) = \frac{\exp(v)}{\exp(v) + \exp\left(\varphi \log\left(2\exp(v/\varphi)\right)\right)}$$

$$= \frac{\exp(v)}{\exp(v) + \exp\left(\log(2^{\varphi}) + v\right)}$$

$$= \frac{\exp(v)}{\exp(v) + \left(2^{\varphi} \times \exp(v)\right)} = \frac{1}{1+2^{\varphi}}$$



For $\varphi = 0$,

$$P(\text{Car}) = \frac{1}{1+2^{0}} = \frac{1}{1+1} = \frac{1}{2}$$

For $\varphi = 1$,

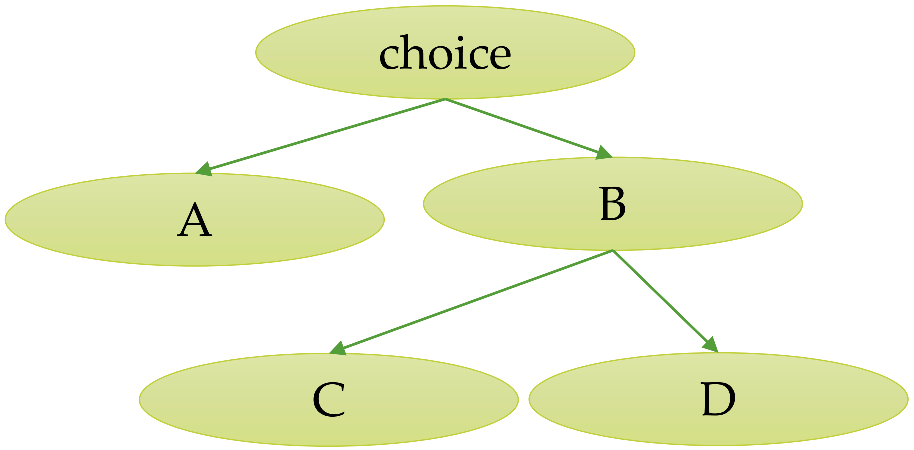
$$P(\text{Car}) = \frac{1}{1+2^{1}} = \frac{1}{1+2} = \frac{1}{3}$$

For $0 < \varphi < 1$,

$$\frac{1}{2} > P(\text{Car}) > \frac{1}{3}$$

Here $\varphi = \mu_1 / \mu_2$ is the Inclusive Value parameter capturing the degree of correlation among the nested alternatives

Nested Logit (NL) Model



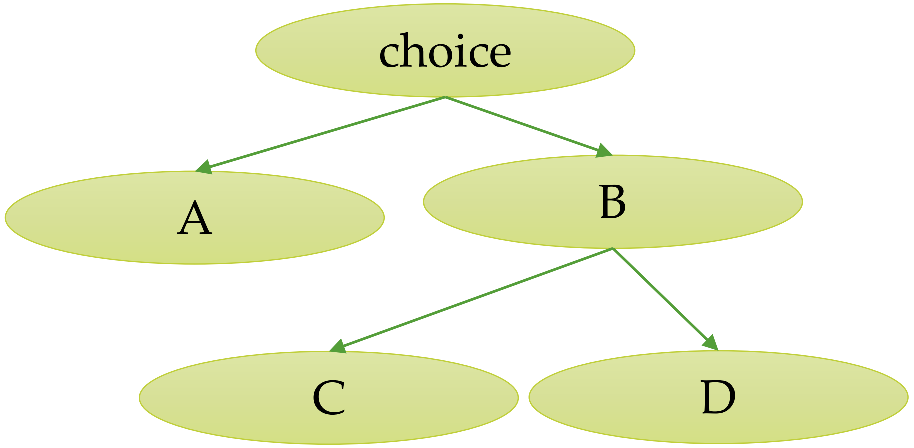
Alternative B nests 2 alternatives C & D

$$P(\mathrm{A}) = \frac{\exp(v_A)}{\exp(v_A) + \exp\left(\varphi \log\left(\exp(v_C/\varphi) + \exp(v_D/\varphi)\right)\right)}$$

$$P(\mathrm{C}) = P(C \mid B)P(B) = \frac{\exp(\exp(v_C/\varphi))}{e\exp(v_C/\varphi) + \exp(v_D/\varphi)} \times \left(1 - P(A)\right)$$

$$P(\mathrm{D}) = P(D \mid B)P(B) = \frac{\exp(\exp(v_C/\varphi))}{e\exp(v_C/\varphi) + \exp(v_D/\varphi)} \times \left(1 - P(A)\right)$$

Nested Logit (NL) Model



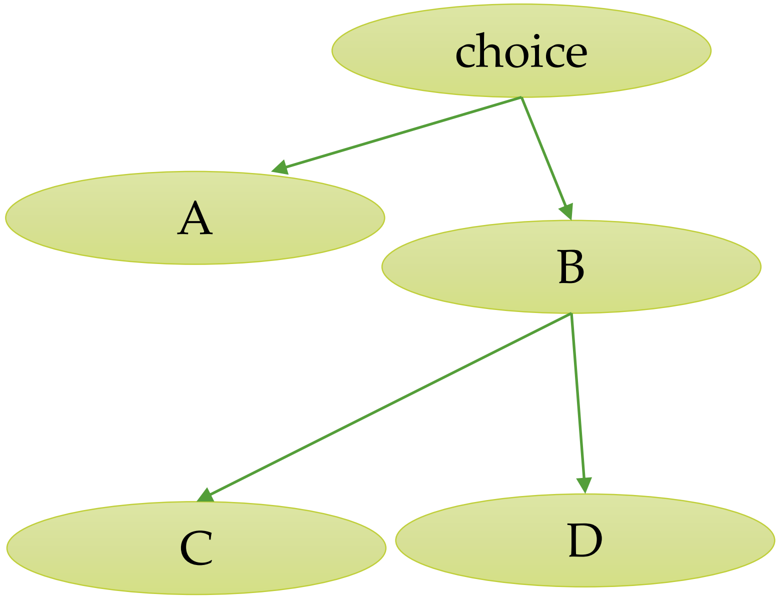
Alternative B nests 2 alternatives C & D

$$P(\mathrm{A}) = \frac{\exp(\mu_1 v_A)}{\exp(\mu_1 v_A) + \exp\left(\dfrac{\mu_1}{\mu_2}\log\left(\exp(\mu_2 v_C) + \exp(\mu_2 v_D)\right)\right)}$$

$$P(\mathrm{C}) = P(C \mid B)P(B) = \frac{\exp(\exp(\mu_2 v_C))}{\exp(\mu_2 v_C) + \exp(\mu_2 v_D)} \times \left(1 - P(A)\right)$$

$$P(\mathrm{D}) = P(D \mid B)P(B) = \frac{\exp(\exp(\mu_2 v_C)}{\exp(\mu_2 v_C) + \exp(\mu_2 v_D)} \times \left(1 - P(A)\right)$$

Nested Logit (NL) Model



Alternative B nests 2 alternatives C & D

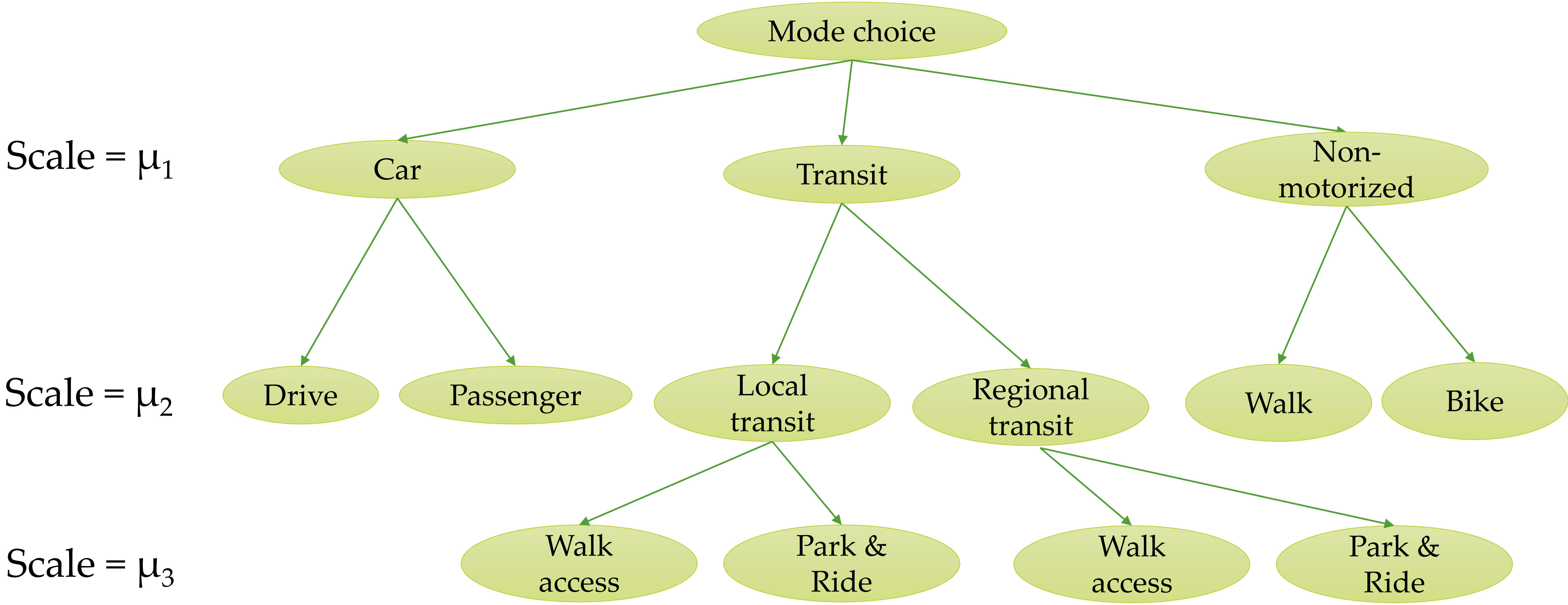
$$P(\mathrm{A}) = \frac{\exp(\mu_1 v_A)}{\exp(\mu_1 v_A) + \exp\left(\frac{\mu_1}{\mu_2}\log\left(\exp(\mu_2 v_C) + \exp(\mu_2 v_D)\right)\right)}$$

$$P(\mathrm{C}) = P(C \mid B)P(B) = \frac{\exp(\exp(\mu_2 v_C))}{\exp(\mu_2 v_C) + \exp(\mu_2 v_D)} \times P(B)$$

$$P(\mathrm{D}) = P(D \mid B)P(B) = \frac{\exp(\exp(\mu_2 v_C)}{\exp(\mu_2 v_C) + \exp(\mu_2 v_D)} \times P(B)$$

khandker.nurulhabib@utoronto.ca

Nested Logit (NL) Model

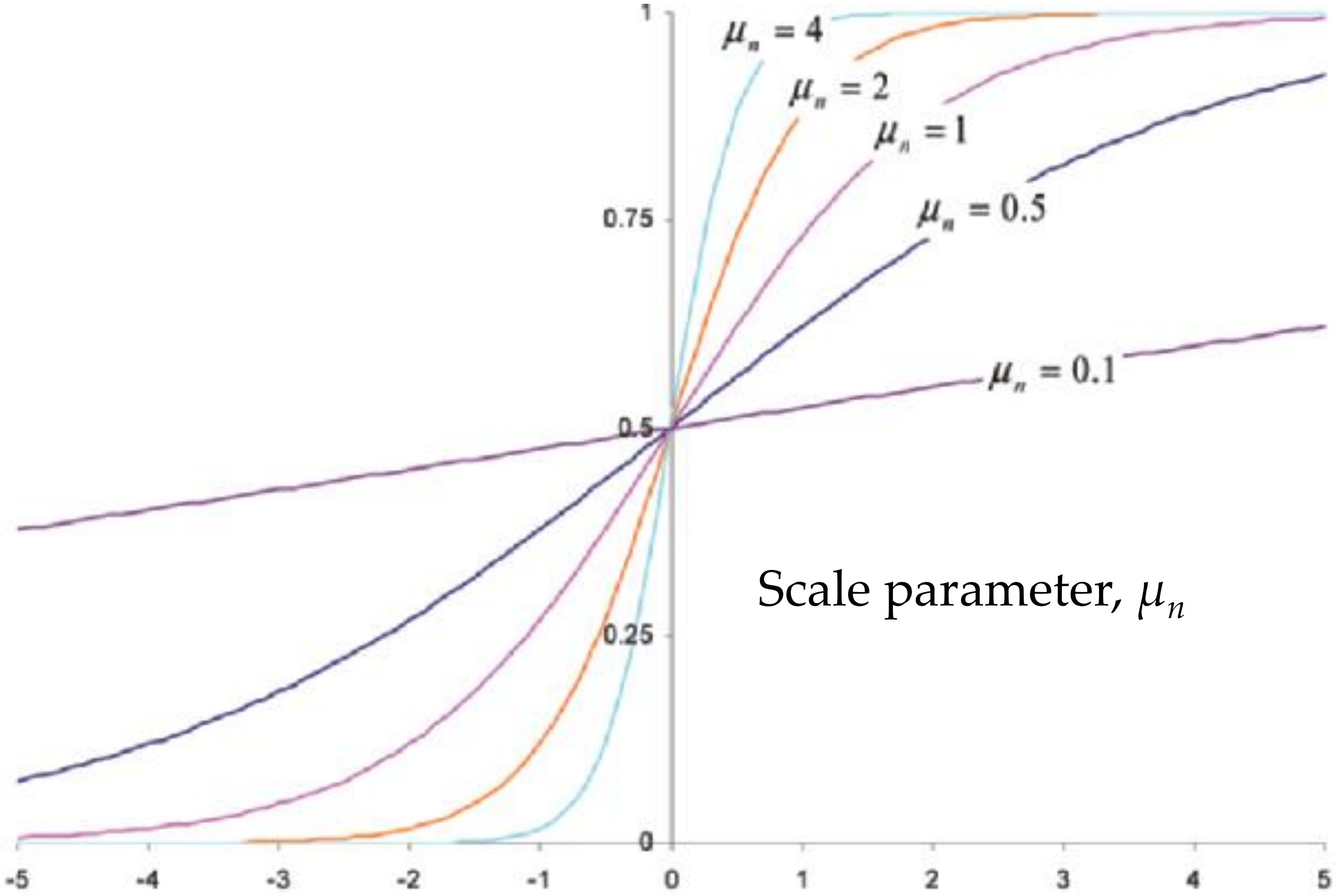


Key Elements of Utility Function

- Alternative Specific Constant (ASC): Captures the unexplained systematic preferences:
  - It is important to have full set of ASCs in the choice model: Necessary to capture market share
  - ASC values need to be updated to any context to make sure that the based model correctly captures the market shares
- Scale parameter (μ):
  - Explains the randomness of mode prediction
  - Can be updated to new context before applying the model
- Attributes of choice alternatives: Captures substitution
- Socio-economics variables: captures moderating effects of attributes of person/hh

Key Elements of Utility Function

Probability of choosing j



Difference in Systematic Utility of j and reference k

Advanced Choice Models: Closed Form

- Nested logit model with parameterized scale function:
  - Allows capturing difference in sources of choice randomness
- GEV model with complex substitution patterns:
  - Ordered Generalized Extreme Value (OGEV) model: Modeling choice of alternatives that are discrete and ordered in nature
  - Cross-Nested logit model: Alternatives sharing multiple nests
  - Generalized Nested Logit (GenL) model: Alternatives sharing multiple nests with varying weights per nests
- GEV model with choice set formation
- Elasticity (direct and cross) function changes with different formulations

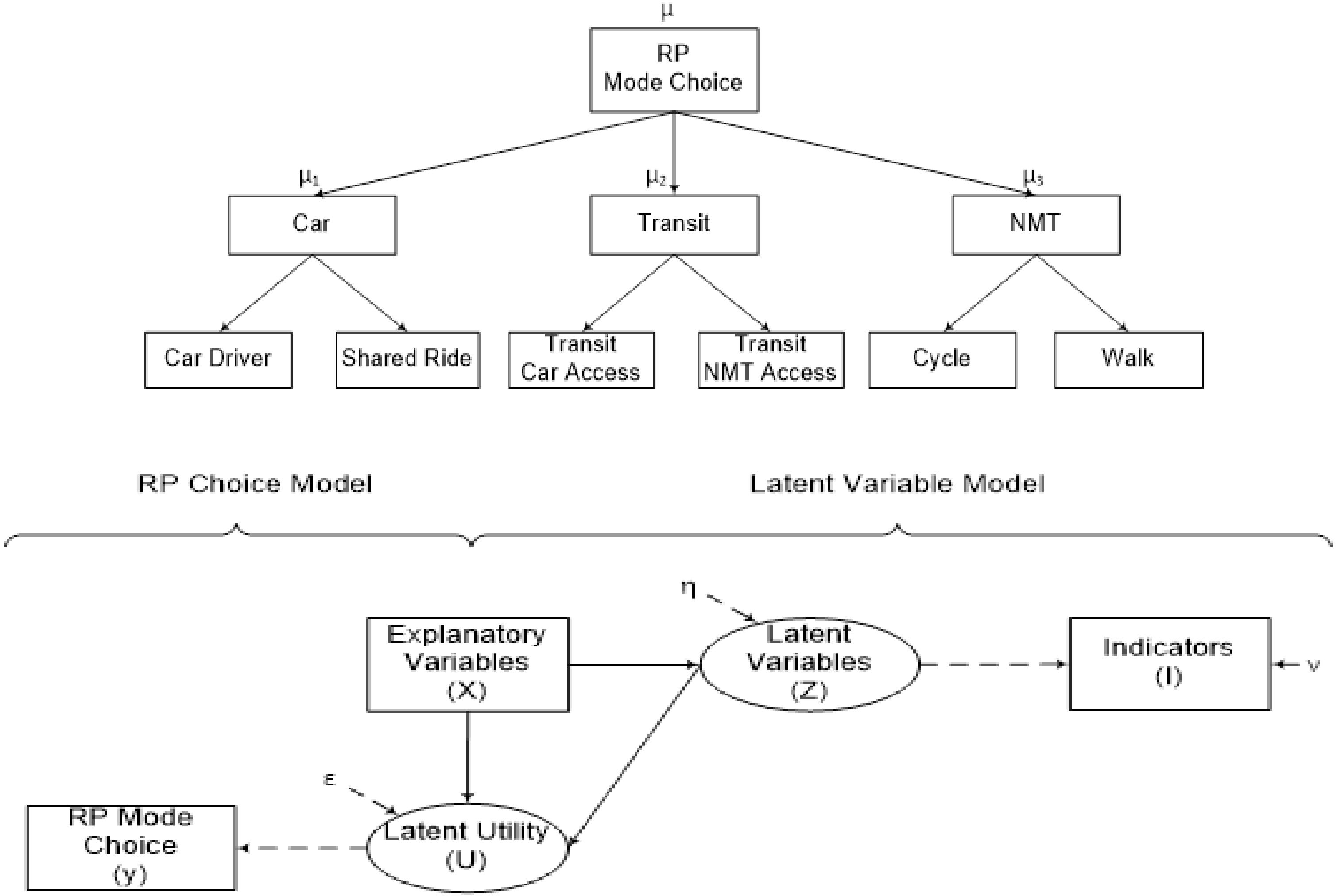
Advanced Choice Models: Non-Closed Form

- Mixed Logit model:
  - Completely flexible substitution patterns
- Mixed logit model with latent variables
- Mixed logit with latent segmentation

- Elasticity (direct and cross) function changes with the specification of mixed random errors
- Model estimation and prediction requires simulation techniques

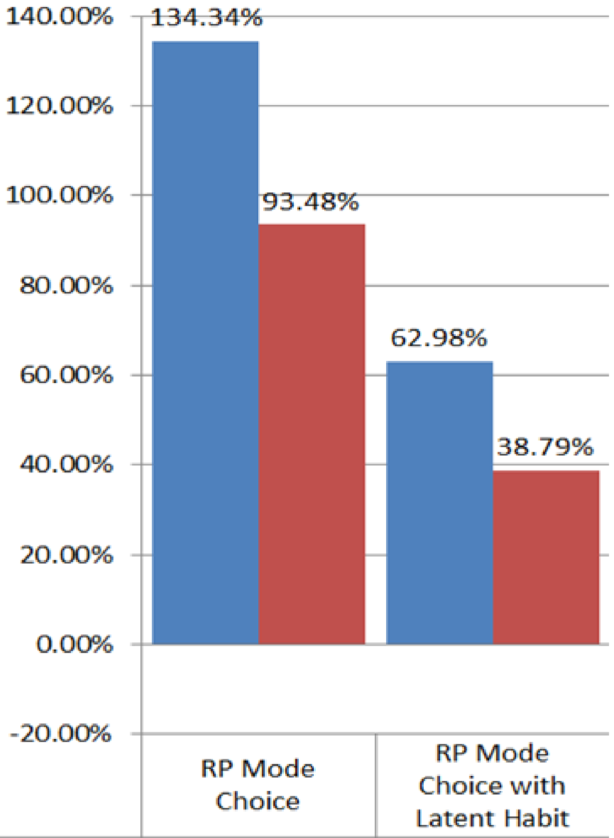
Example: Choice model with Latent variables

- Modelling the choice of switching to Transit:
  - Considering the effects of habit and inertia by accommodating latent variables
  - Without considering the effects of habit and Inertia
- Data: from a sample of commuter in the City of Toronto



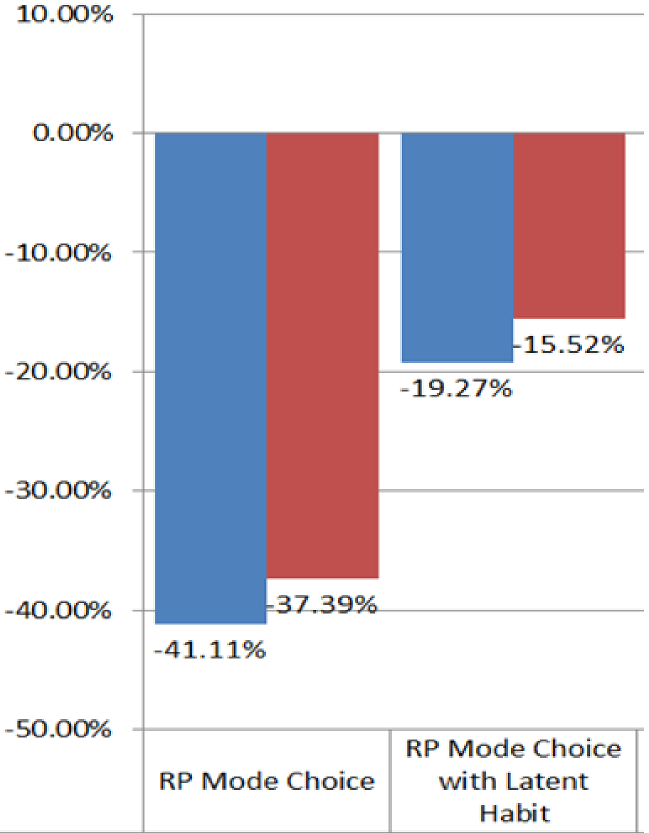
Example: Choice model with Latent variables

Over-prediction of Transit Share



	RP Mode Choice	RP Mode Choice with Latent Habit
239 Observations	134.34%	62.98%
1407 Observations	93.48%	38.79%

Under-prediction of Car Share



	RP Mode Choice	RP Mode Choice with Latent Habit
239 Observations	-41.11%	-19.27%
1407 Observations	-37.39%	-15.52%

Models of Multiple Choice Contexts

- Panel Data Model:
  - Multiple choice information per respondents:
    - ✓ Multiple choice observation of same choice or multiple choice information of different choices but of same respondents
    - ✓ Multiple choice records per respondents in the Stated Preference Experiment
- Dynamic Choice Model: Modelling multiple choices that are made in sequences:
  - Mode choices of sequences of trips in a tour: home-based tour, work-based tour etc.
  - Destination location choice of a sequences of trips in your or in day
  - Trip-activity scheduling choice model

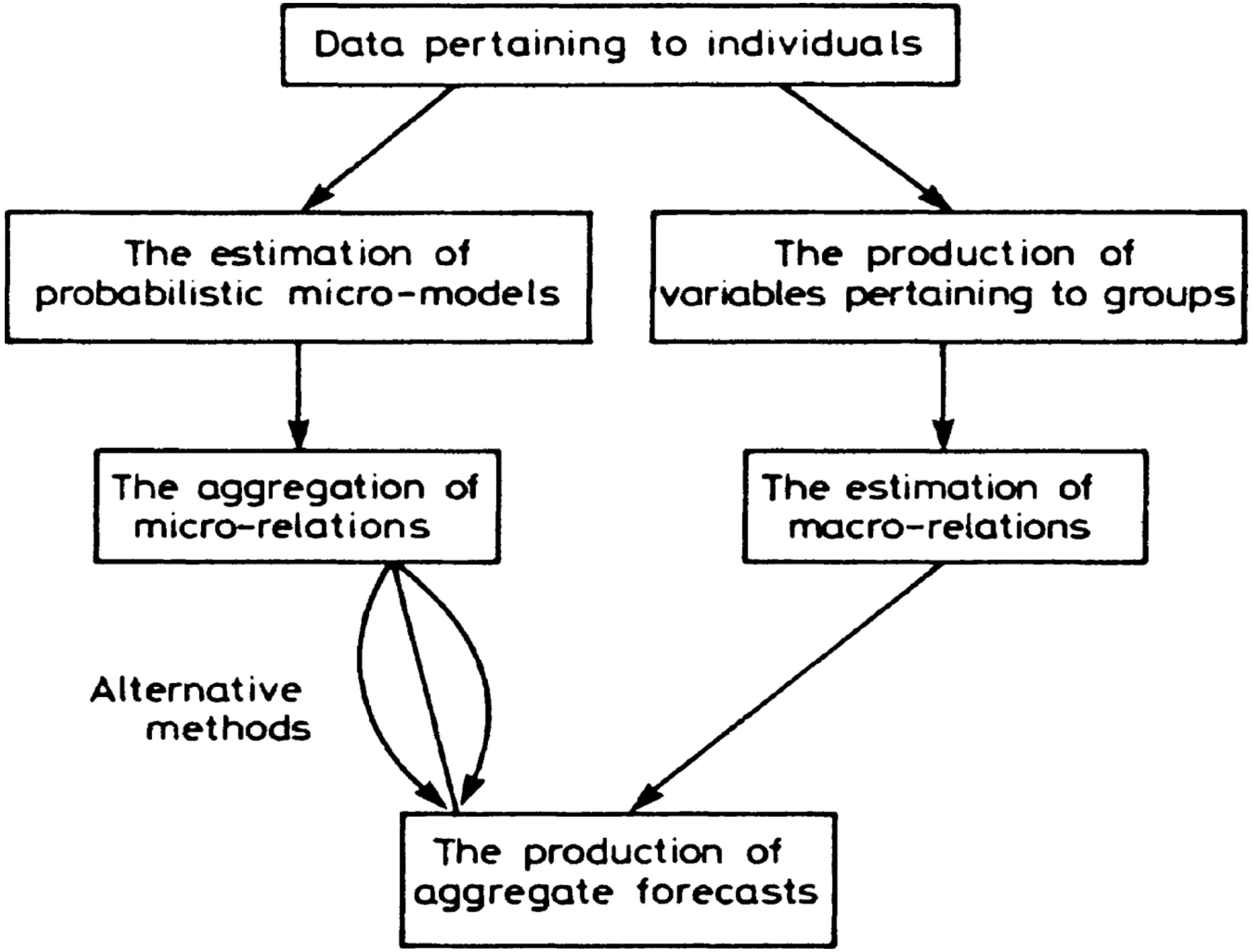
Discrete Choice Model Estimation

- Application of Disaggregate Models:
  - Sample observation: data of observed choices
  - Attributes of choice alternatives:
    - ✓ Feasible alternatives
    - ✓ Attributes of chosen alternatives
    - ✓ Attributes of non-chosen alternatives
  - Attributes of choice contexts:
    - ✓ Attributes of choice makers
    - ✓ Attributes of situations
- Appropriate specification of explanatory variables
- Estimating model parameters: R, LIMDEP, BIOGEME

Model Specifications

- If choices are distinct (labelled alternatives, e.g. mode choice)
  - Discrete choice model needs to have alternative-specific constant (ASC)
  - ASC captured systematic utility that are not explained by available variables
- In case of alternatives are very large in numbers: (un-labelled alternatives e.g. choice of destination location)
  - Discrete choice model may not have ASC
- For large number of alternatives:
  - It may be necessary to sample from the alternatives for model estimation
- Choice model needs to be calibrated before future prediction/scenario analysis:
  - Making sure that the model accurate predicts current market shares

Disaggregate Model for Forecasting



Disaggregate Model for Forecasting

$$\sum_{\substack{all \\ individuals}} Individual\ \ Behaviour = Aggregate\ (Collective)\ outcome$$

Alternative Methods of Aggregation

Microsimulation	Simulate 100% population and predict for all
Sample Enumeration	Generate a representative sample to predict for the sample and then expand to get the population
Naïve aggregation	Assumed population average value of all variables to predict the average shares
Classification with Naïve aggregation	Classify/segment population into finite number of categories and use Naïve aggregate for each category separately

Model Transferability

- Transferability issues:
  - Systematic bias
  - Location/context specific local conditions
- Updating model:
  1. Systematic bias of the model is captured by ASCs. Constants are error basket. At a minimum level, the constant terms changes place to place.
  2. Scaling issue: Scale parameter of random error term changes place to place.
  3. Model specification: model specification may also changes place to place

Thank You