The FHWA Travel Model Improvement Program Workshop over the Web

The Travel Model Development Series: Part I – Travel Model Estimation

presented by
Thomas Rossi
Cambridge Systematics, Inc.

October 16, 2008

1

Acknowledgments

- Webinar prepared by Cambridge Systematics, Inc.
  - Thomas Rossi
  - Yasasvi Popuri
- Prepared for Federal Highway Administration
  - H. Sarah Sun
- Arranged by Texas Transportation Institute
  - Gary Thomas

2

Webinar Objectives
Learning

- Intended for those who have a low level of familiarity with the estimation and validation of travel models
- Introduces
  - Development of model estimation data sets
  - Structures of various model components
  - Procedures for estimating models
- Does not train participants in model estimation procedures

3

Webinar Outline

- Session 1: Introduction – October 16, 2008
- Session 2: Data Set Preparation – November 6, 2008
- Session 3: Estimation of Non-Logit Models – December 11, 2008
- Session 4: Estimation of Logit Models – February 10, 2009

4

Webinar Outline (continued)

- Session 5: Application and Validation of Logit Models – March 12, 2009
- Session 6: Advanced Topics in Discrete Choice Models – April 14, 2009
- Session 7: Trip Assignment – May 7, 2009
- Session 8: Evaluation of Validation Results – June 9, 2009

5

Homework

(You thought you finished this forever decades ago!)

- Exercise will be assigned at the end of each session, based on some of the information discussed
- May require generally available software (e.g., Excel, text editor, etc.), but not specialized software such as model estimation or application programs
- You may e-mail questions on the homework to us up to one week before the next session
  - trossi@camsys.com, ypopuri@camsys.com
- Homework will be reviewed during the first part of the next session

6

Types of Urban Models

- Four-step models
- Three-step models (no mode choice)
- Tour and activity based models

7

Model Components
Four-Step Models

- Trip production
- Trip attraction
- Trip distribution
- Mode choice
- Assignment
- Time of day
- Auto ownership
- Other
  - Trucks/freight
  - External trips
  - Other?

8

The Model …

- … takes a set of *input data* …
- … and converts it to a set of *output data* …
- … using a set of *mathematical models* …
- … which use *parameters* to perform the conversions
- The input and output data for individual model components may be temporary, or interim

9

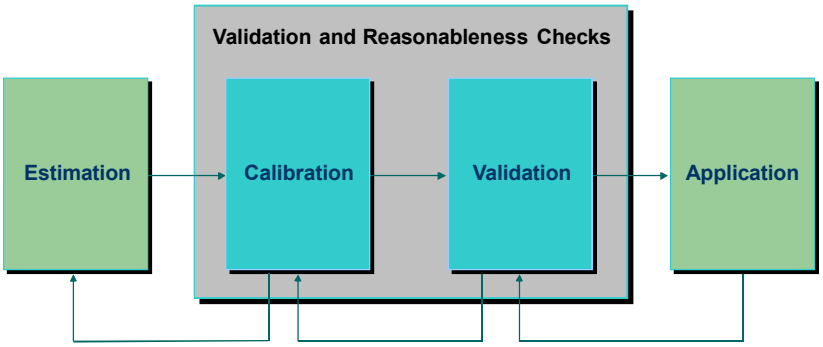
Model Parameters …

… may be:

- Estimated (usually from local data),
- "Borrowed" from another model or other data source, or
- Asserted, based on knowledge from other sources

10

The Model Development Process



Model Development Process

- Have a model development plan
- Assemble necessary data for estimation, validation, and application
- For each component
  - Define the mathematical model and the model structure
  - Estimate (or otherwise obtain) the parameters
  - Validate the model (a whole ‘nother topic!)
  - Recalibrate the parameters as necessary

12

Model Development Process (continued)

- Validate overall model system, including forecast years
  - Recalibrate as necessary
- Don't forget documentation!

13

Model Application Process

- Overall inputs
  - Socioeconomic data
  - Networks (highway, transit)
  - Other (parking costs, auto operating costs, etc.)
- Overall outputs
  - Link volumes → VMT
  - Link speeds → VHT
  - Transit boardings, line volumes
  - Trips by O-D, mode, purpose, time of day, etc.

14

Some Model Component Examples

- Trip generation
  - Input – Socioeconomic data
  - Output – Trip ends
- Trip distribution
  - Input – Trip ends/network skims
  - Output – Trip tables

15

Some Model Component Examples (continued)

- Mode choice
  - Input – Trip tables/network skims
  - Output – Trip tables by mode
- Highway assignment
  - Input – Trip tables for auto mode/highway network
  - Outputs – Volumes/speeds

16

Common Model Component Formulations

- Simple factoring
- Cross-classification
- Regression
- Logit (multinomial, nested)
- Assignment

17

Model Estimation Data Sources

- Household activity/travel survey (household, trip level)
- Transit on-board survey
- Critical nonsurvey data
  - Socioeconomic data
  - Networks
  - Other (area types, parking costs, auto operating costs, etc.)

18

Model Types

Model Component	Typical Model Type	Usual Estimation Data Source	Other Basic Data
Auto ownership	Multinomial/ ordered response logit	Household survey (household file)	Socioeconomic, network (skims)
Trip production	Cross-classification	Household survey (household file)	Socioeconomic
Trip attraction	Linear regression	Household survey (trip file)	Socioeconomic
Trip distribution	Gravity/ multinomial logit	Household survey (trip file)	Socioeconomic, network (skims)

19

Model Types (continued)

Model Component	Typical Model Type	Usual Estimation Data Source	Other Basic Data
Time of day	Simple factoring	Household survey (trip file)	
Mode choice	Multinomial/ nested logit	Household survey (trip file)	Socioeconomic , network (skims)
Highway assignment	Static user equilibrium		Highway network
Transit assignment	All-or-nothing, multipath		Transit network

20

Regression Model

$Y = B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_n X_n$

where:

Y = Dependent variable

B_i = Estimated coefficients

X_i = Independent variables

21

Regression Model Example
Trip Attraction Model

HBO Attractions = 1.32 (service employment)

+ 1.46 (retail employment)

+ 0.76 (households)

22

Cross-Classification Model

Independent Variable #2	Independent Variable #1			
	Value 1	Value 2	…	Value n
Value 1	Dep var value	Dep var value		Dep var value
Value 2	Dep var value	Dep var value		Dep var value
…				
Value n	Dep var value	Dep var value		Dep var value

23

Cross-Classification Model Example
Trip Production Model – Home Based Shop Trips

Persons/ Household	Vehicles/Household			
	0	1	2	3+
1	0.294	0.333	0.333	0.333
2	0.377	0.465	0.619	0.619
3	0.425	0.515	0.619	0.619
4+	0.433	0.601	0.740	0.767

24

Gravity Model
Trip Distribution

$$T_{ij} = \frac{P_i A_j F(t)_{ij} K_{ij}}{\sum_j P_i A_j F(t)_{ij} K_{ij}}$$

where:

- T_{ij} = number of trips produced in zone i and attracted to zone j
- P_i = trips produced in zone i
- A_j = trips attracted to zone j
- $F(t)_{ij}$ = friction factor from i to j (based on impedance t)
- K_{ij} = K factor from i to j
- i = origin zone
- j = destination zone

25

Multinomial Logit Model

$$P(\text{alt } 1) = \frac{\exp(V_1)}{\sum_j \exp(V_j)}$$

where:

- V_j = Deterministic component of utility of alternative j
- exp = exponential function (e^x)

26

Multinomial Logit Model
Utility Function

$$V_{ij} = B_{0j} + B_{1j} X_{1ij} + B_{2j} X_{2ij} + \ldots + B_{nj} X_{nij}$$

where:

V_{ij} = Utility (deterministic component) of alternative j for individual i

X_{kij} =Attributes (k = 1,n) for individual i for alternative j

B_{kj} =Estimated coefficients for attribute k for alternative j

B_{0j} = Alternative-specific constant for alternative j

27

Multinomial Logit Model Example
Vehicle Availability Model

Variable	Vehicle Availability Level			
	0	1	2	3+
Constant	Base alt (U = 0)	0.64	-0.45	-2.29
1 worker in household (0,1)		0.83	1.10	1.66
2+ workers in household (0,1)		0.54	2.47	3.32
Low-med income (0,1)		1.16	2.18	2.26
High-med income (0,1)		0.87	3.04	3.64
High income (0,1)		1.78	4.31	5.28
% employment within 15 min		-0.03	-0.08	-0.12

28

Multinomial Logit Model Example
Vehicle Availability Model (continued)

Utility functions:

$U_0 =$ 0

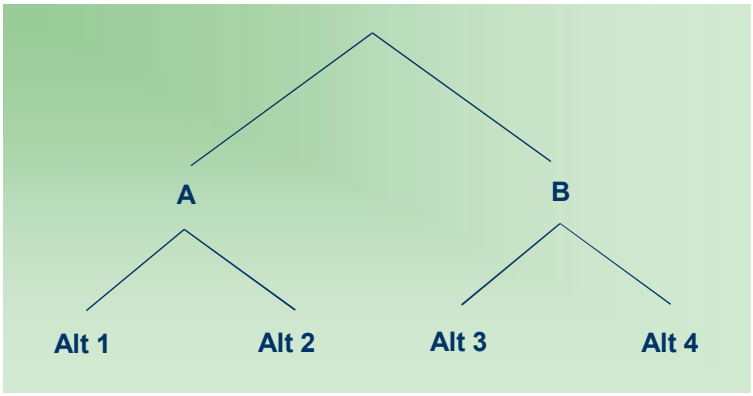
$U_1 =$ 0.64 + 0.83 (1 worker) + 0.54 (2+ worker)
+ 1.16 (Low-medium income) + 0.87 (High-medium income)
+ 1.78 (High income) -0.03 (% employer w/in 15 min)

$U_2 =$ -0.45 + 1.10 (1 worker) + 2.47 (2+ worker)
+ 2.18 (Low-medium income) + 3.04 (High-medium income)
+ 4.31 (High income) -0.08 (% employer w/in 15 min)

$U_3 =$ -2.29 + 1.66 (1 worker) + 3.32 (2+ worker)
+ 2.26 (Low-medium income) + 3.64 (High-medium income)
+ 5.28 (High income) -0.12 (% employer w/in 15 min)

29

Nested Logit Model



30

Nested Logit Model

$$P(\text{alt 1} \mid A) = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)}$$

$$P(A) = \frac{\exp[f(V_1, V_2)]}{\exp[f(V_1, V_2)] + \exp[f(V_3, V_4)]}$$

$$P(\text{alt 1}) = P(\text{alt 1} \mid A)\ P(A)$$

31

Highway Assignment

- Static equilibrium assignment
  - Assumes O-D travel times along all used paths are equal
- Link travel time is a function of travel time, e.g.,

  $T = T_0 + [\ 1 + a\ (v/c)^b\ ]$

  where:

  T = link travel time

  T_0 = free flow travel time

  v = link volume

  c = link capacity

  a, b model parameters

32

Model Parameter Development
Example

- Estimated (from household survey)
  - Trip generation rates
  - Friction factors for gravity model
  - Time of day percentages by purpose
- Borrowed
  - Mode choice utility coefficients (from other model)
- Asserted
  - BPR function parameters a, b (from literature)
  - K-factors (all set to 1.0)

33