

Demonstration of NLOGIT Software

- Preparation of Data set
- Reading Data
- Various data analysis commands
- Input for MNL model
- Interpretation of MNL output
- Input file for Nested logit model
- Interpretation of NL output

NLOGIT

- Extension of **LIM**ited **DEP**endent Variable Models (LIMDEP)
- **N**ested **LOGIT** Models (NLOGIT)
 - Descriptive statistics
 - Linear regressions
 - Developing Models like
 - Multinomial Logit models
 - Nested Logit models
 - Random parameter logit models
 - Probit models etc.

Choice	twTT	tTT	bTT	cTT	twTC	tTC	bTC	cTC	Hhinc	Hhsiz
2	35	24	38	30	15	7	10	23	50	4
3	30	20	32	25	13	6	8	22	40	3
1	50	35	59	45	23	13	15	40	70	2

ALOGIT

Choices: 1-TW, 2-Train, 3-Bus, 4-Car

NLOGIT

Choice	TT	TC	Hhinc	Hhsiz
0	35	15	50	4
1	24	7	50	4
0	38	10	50	4
0	30	23	50	4
0	30	13	40	3
0	20	6	40	3
1	32	8	40	3
0	25	22	40	3
1	50	23	70	2
0	35	13	70	2
0	59	15	70	2
0	45	40	70	2

Data set

	Cset	AltID	Choice	TT	TC
TW	4	1	0	20	27
Train	4	2	0	15	22
Bus	4	3	0	45	52
Car	4	4	1	60	70
TW	4	1	0	15	22
Train	4	2	1	12	19
Bus	4	3	0	42	49
Car	4	4	0	70	100
TW	4	1	0	34	41
Train	4	2	0	30	37
Bus	4	3	0	80	87
Car	4	4	1	100	140
TW	4	1	0	55	62
Train	4	2	0	42	49
Bus	4	3	1	100	107
Car	4	4	0	150	200

Reading Data

- Manual entry
- Spread sheet file program (Microsoft Excel)
- Data file (.dat)
- By using “Read” command
- By using project file (.lpj)

Reading data

- Text;_ _ _ _ _ \$
- Read
;File=E:\Nlogit.prn
;Nobs=1000
;Nvar=7
;Names=Choice,TT,TC,GC,WT,HHInc,Hhsize \$

Instructions to NLOGIT

- Two methods of giving instructions to NLOGIT
 - Dialog boxes
 - Self documenting Command lines

Data analysis commands

- REGRESS; Lhs = dependent variable; Rhs = independent variable \$
- HISTOGRAM; Rhs = a variable \$
- DSTATS; RHS = the list of variables \$? For descriptive statistics.
- CREATE; name = expression; name = expression . . . \$
- CROSSTAB; Lhs = variable; RHS = Variable \$
- SHOW
- Simulation
 - ;Scenario: TC(Rail)=[*]2
 - ;Scenario: TT(Bus)=[+]4

Types of data on the choice variable

- Individual data

The Lhs variable consists of zeros and a single which indicates the choice that the individual made

- Proportions data

Lhs variable consists of a set of sample proportions ranges from zero to one

- Frequency data

Lhs variable consists of a set of frequency counts for the outcomes (non negative integers)

- Ranks data

Lhs variable consists of a complete set of ranks of the alternatives in the individual's choice set

- $[0, 1, 0, 0, 0]$ ----- unranked
- $[4, 1, 3, 2, 5]$ ----- ranked



Multinomial Logit model

- The choice variable, **Choice**=1,2,3,4 for tw, train, bus and car
- For each mode, **TT, TC, GC, WT** --- Differ by choices
- For the individual, **Hhinc, Hhsize** --- Does not differ by choice
- **NLOGIT**
 - ; Lhs= Choice
 - ; Choices = tw, train, bus, car
 - ; Rhs = TT, TC, GC, WT
 - ;Rh2 = One, hhinc, hhsiz \$

Multinomial choice model (MNL)

NLOGIT

;lhs = choice

;choices =<names of alternatives>

;Model:

U(alternative 1 name) = <utility function 1> /

U(alternative 2 name) = <utility function 2> /

U(alternative i name) = <utility function i> \$

Multinomial choice model (MNL)

NLOGIT

;lhs = choice

;choices =Metro,Rail,Bus,Car

;Model:

U(Metro) =

a_metro+tt*TT+tc*TC+wt*WT+gc*GC+a_hhinc*HHInc /

U(Rail) =b_rail+ tt*TT+tc*TC+wt*WT+gc*GC+b_hhinc*HHInc /

U(Bus) = c_bus+tt*TT+tc*TC+wt*WT+gc*GC+c_hhinc*HHInc/

U(Car)=d_car+tt*TT+tc*TC+wt*WT+gc*GC+d_hhinc*HHInc \$

Maximum Likelihood Estimation

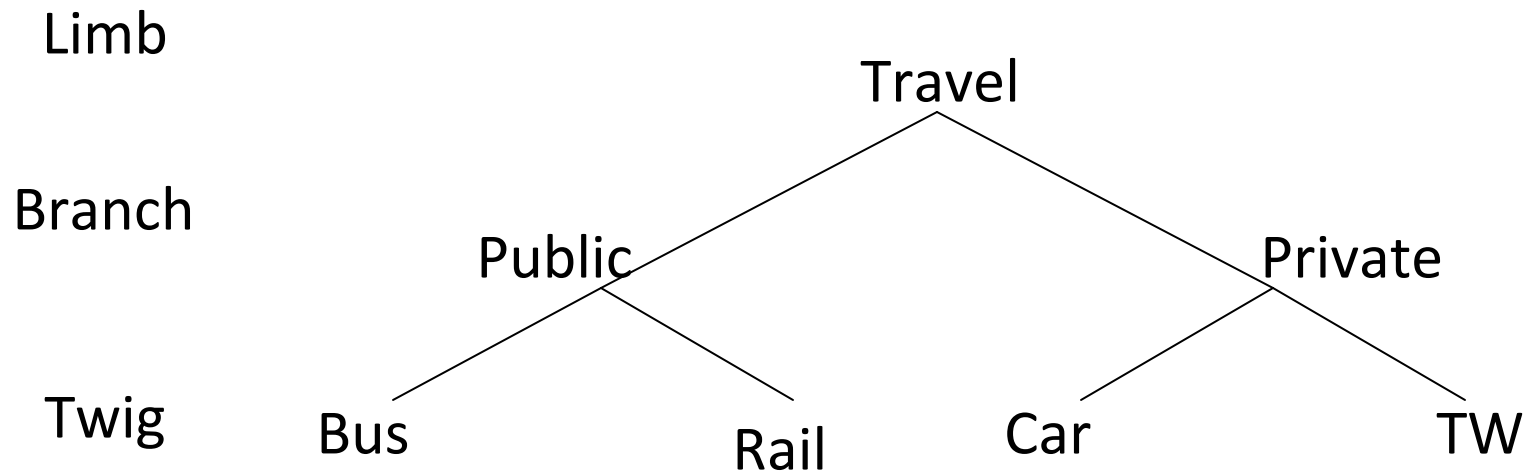
- $LL = (n! / h!(n - h)!) p^h (1 - p)^{(n-h)}$
- Searching for the value of “p” which will maximize the likelihood
- $-2(LL_{\text{base model}} - LL_{\text{estimated model}}) \sim \text{Chi squared}(\text{number of new parameters estimated in the estimated model})$
- $\text{Prob} [\text{chi squared} > \text{value}] = .000000$

Interpretation

To Determine whether an explanatory variable is statistically significant or not

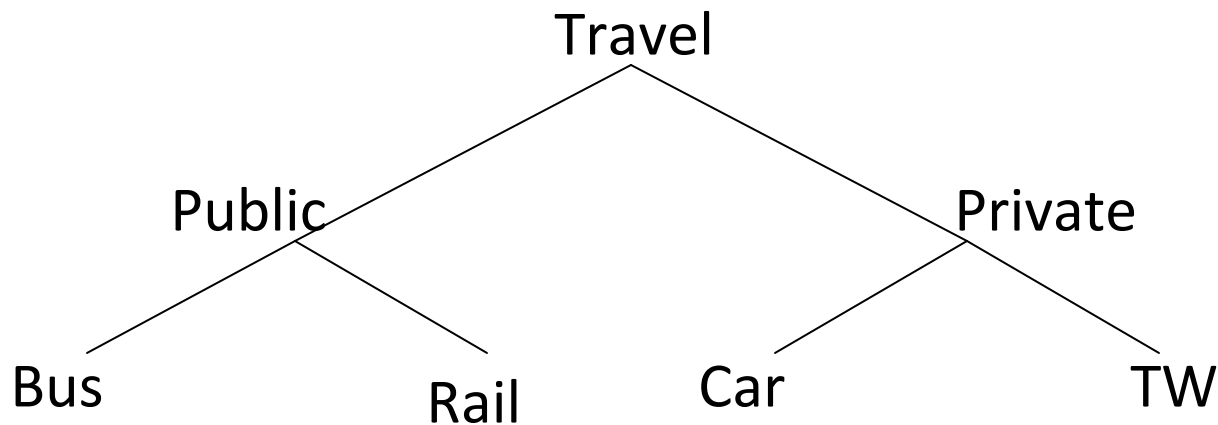
- Logical sign of the parameter
 - Wald statistic (model) > critical wald value (1.96) [at 95 percent confidence level (i.e. $\alpha=0.05$)]
 - P-value < 0.05 (alpha)
- [Null hypothesis: Estimated model is no better than base model
Alternate hypothesis: Estimated model is better than base model]

Nested Logit Model



- $U(\text{Bus}) = \beta_0 + \beta_1 \cdot \text{Var1} + \beta_2 \cdot \text{Var2}$
- $U(\text{Public}) = \sum \alpha \cdot Z + \Phi \cdot \text{EMU}$
- $\text{EMU}(\text{Bus}, \text{Rail}) = \ln \{ \exp V(\text{bus}) + \exp V(\text{Rail}) \}$

- In defining the tree structure, the following NLOGIT conventions apply:
- {} specifies a trunk (level 4)
- [] specifies a limb within a trunk (level 3)
- () specifies a branch within a limb within a trunk (level 2).



- NLOGIT may have up to a maximum of five trunks, 10 limbs, 25 branches, and 100 alternatives
- NLOGIT can estimate upto 4 level nested structure

Scale Parameter

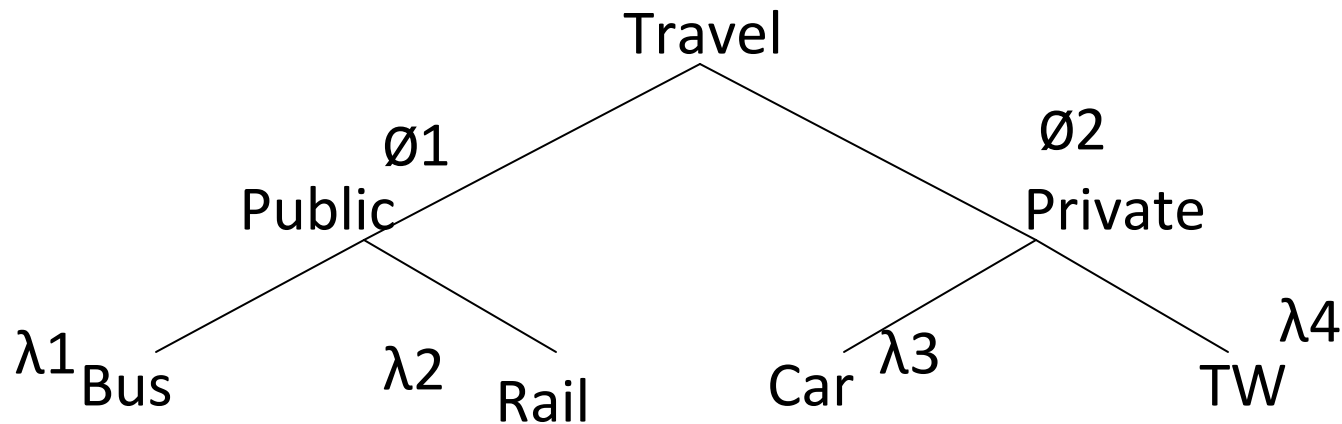
- The mean and variance of the Gumbel distribution are:

$$\text{Mean} = \eta + (0.577 / \lambda)$$

$$\text{Variance} = (\pi^2 / 6 \lambda^2)$$

Where, λ is the scale parameter (inverse function of variance) and η is the location or mode parameter

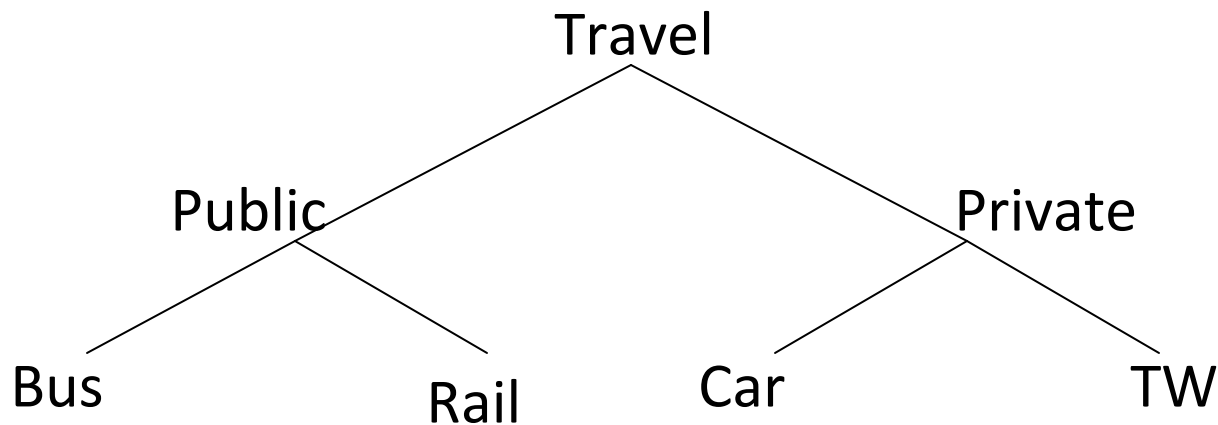
- $U(\text{Bus}) = \lambda \beta_0 + \lambda \beta_1 \text{Var}_1 + \lambda \beta_2 \text{Var}_2$



- Parameter of Inclusive value at level 2 = Ratio of the scale parameter at level 2 to scale parameter at level 1 = (\emptyset/λ)
- Normalization at lower level = $RU1 = \emptyset$
- Normalization at higher level = $RU2$

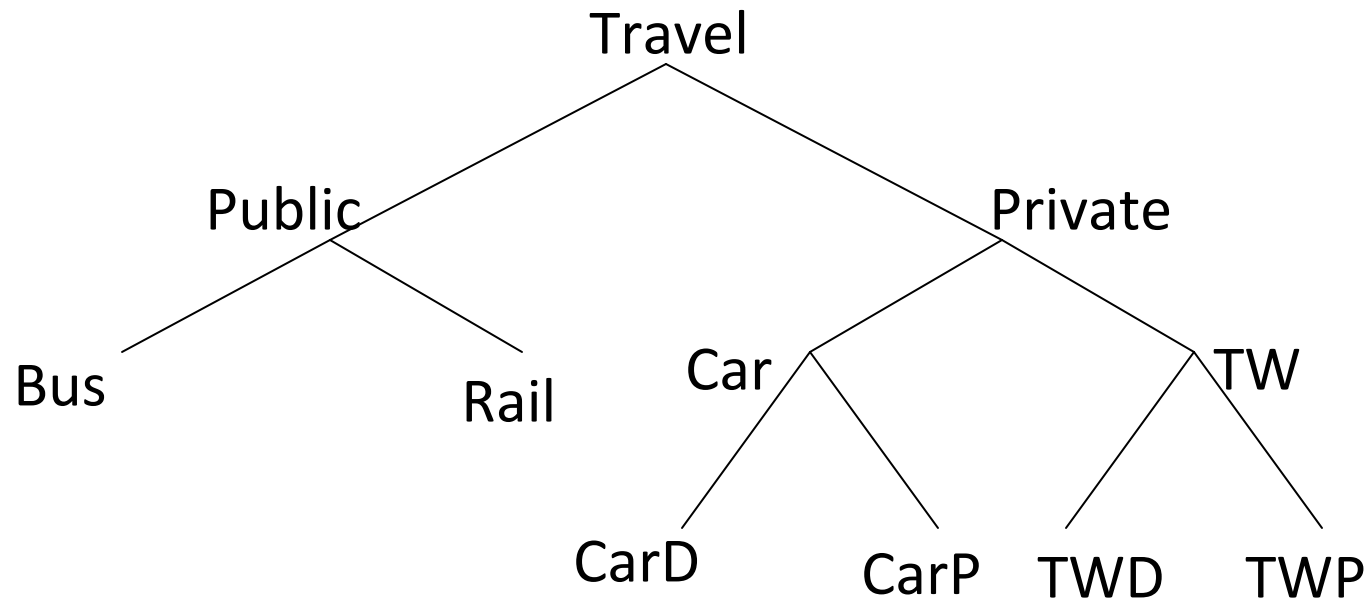
Range of IV parameter

- Must lie within 0-1 range
- Scale parameter at upper level must be lower than scale parameter of lower level



Consistency of IV parameter

- The structural parameter should satisfy the condition: $0 < \emptyset \leq 1$
- If $\emptyset < 0$, an increase in the utility of an alternative in the nest, which should increase the value of EMU, would actually diminish the probability of selecting the nest
- If $\emptyset = 0$, such an increase would not effect the nest's probability of being selected, as EMU would not effect the choice between car and PT.
- If $\emptyset > 1$, an increase in the utility of an alternative in the nest would tend to increase not only its selection probability but also those of the rest of the options in the nest.
- If $\emptyset = 1$, the model becomes equivalent to MNL.



Tree =
Travel{Public[Bus,Rail],Private[Car(CarD,CarP),TW
(TWD,TWP)]}

Command Line for NL

- Nlogit

;Lhs=CHOICE

;Choices=Bus,Rail,BR,TW,IPT,CarD,CarP,WNMT

;Tree=Private(CarD,CarP,TW),Public(Bus,Rail,BR),Others(IPT,WNMT)

;ias=Carp

;Show Tree

;start=logit

;Maxit=100

;Model: Utility functions

Interpretation

- FIML: Nested Multinomial Logit Model
- Nested logit models can be estimated: sequentially or simultaneously
- Sequential estimation (known as limited information maximum likelihood estimators or LIML) involves the estimation of separate levels of the NL tree in sequential order from the lowest level of the tree to the highest level.
- The simultaneous estimation of the branches, limbs, and trunks of an NL model is achieved using FIML

- $-2(\text{LLbase model} - \text{LLestimated model}) \sim \text{Chi squared}(\text{difference in number of parameters estimated between two models})$
- $\text{Prob} [\text{chi squared} > \text{value}] = .00000$

Normalizing IV parameters

- Constraining IV parameters

`;ivset:(private,public)`

- Normalizing to fixed value

`;ivset:(private)=[0.8]/(public)=[0.9]`

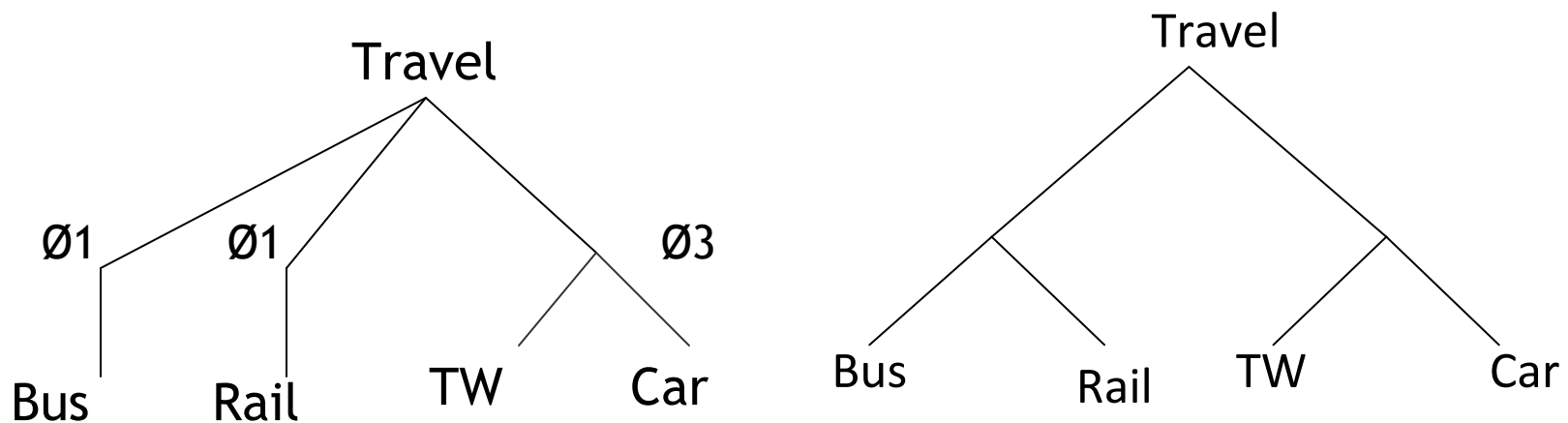
- Trail and Error process

- IV parameters closer to 1.0 not only indicate a smaller difference in the variance between adjoining levels, but also smaller correlation structures between the utility functions of alternatives present within the lower level of the nest

$$\text{corr}(\text{bus}, \text{train}) = 1 - (\text{IV})^2$$

- Nlogit
- ;Lhs=CHOICE
- ;Choices=Bus,Rail,BR,TW,IPT,CarD,CarP,WNMT
- ;Tree=Private(CarD,CarP,TW),Public(Bus,Rail,BR),Others(IPT, WNMT)
- ;Start=logit
- ;ivset: (Car)=[1.0]
- ;Maxit=100
- ;Model:
- $U(\text{private}) = \text{numbvehs} * \text{numbvehs} /$
- $U(\text{public}) = \text{asc1} /$
- $U(\text{Others}) = \text{asc2} /$
- $U(\text{carD}) = /; U(\text{carP}) = /; U(\text{TW}) = /; U(\text{train}) = /; U(\text{bus}) = /;$
 $U(\text{BR}) = /; U(\text{IPT}) = ; U(\text{WNMT}) = \quad \$$

- If there is two degenerate alternatives, both the scale parameters are normalized to 1 , which is equivalent to treating both the alternatives are in the same nest



Converting Data set

Choice	twTT	tTT	bTT	cTT	twTC	tTC	bTC	cTC	Hhinc	Hhsiz
2	35	24	38	30	15	7	10	23	50	4
3	30	20	32	25	13	6	8	22	40	3
1	50	35	59	45	23	13	15	40	70	2

ALOGIT

NLOGIT

Choice	TT	TC	Hhinc	Hhsiz
0	35	15	50	4
1	24	7	50	4
0	38	10	50	4
0	30	23	50	4
0	30	13	40	3
0	20	6	40	3
1	32	8	40	3
0	25	22	40	3
1	50	23	70	2
0	35	13	70	2
0	59	15	70	2
0	45	40	70	2

Converting the data set to multiple line format

- The choice variable, **Choice**=1,2,3,4 for tw, train, bus and car
- For each mode, **TT, TC** --- Differ by choices
- For the individual, **Hhinc, Hhsiz** --- Does not differ by choice

- **NLCONVERT**

; Lhs= Choice

; Choices = tw, train, bus, car

; Rhs = twTT, tTT, bTT, cTT, twTC, tTC, bTC, cTC

;Rh2 = hhinc, hhsiz

;Names = choice, TT, TC, Hhinc, HHsiz

Data set

	Cset	AltID	Choice	TT	TC
Bus	4	1	0	20	27
Train	4	2	0	15	22
Car	4	3	0	45	52
Walk	4	4	1	60	0
Bus	4	1	0	15	22
Train	4	2	1	12	19
Car	4	3	0	42	49
Walk	4	4	0	70	0
Bus	4	1	0	34	41
Train	4	2	0	30	37
Car	4	3	0	80	87
Walk	4	4	1	100	0
Bus	4	1	0	55	62
Train	4	2	0	42	49
Car	4	3	1	100	107
Walk	4	4	0	150	0

Variable choice set

- Number of choices is not constant from one observation to next
 - Universal choice set (Tw, bus, car, train, metro)
 - Non availability of alternatives
- NLOGIT
 - ;Lhs = choice, cset, altij
 - ;Choices = tw, train, bus, car, metro
 - ;Rhs = TT, TC \$

Choice	Cset	AltID	TT	TC	Hhinc	Hhsiz
0	3	1 (TW)	35	15	50	4
1	3	2 (Train)	24	7	50	4
0	3	4 (Car)	38	10	50	4
0	4	1 (TW)	30	13	40	3
0	4	2 (Train)	20	6	40	3
1	4	3 (Bus)	32	8	40	3
0	4	4 (Car)	25	22	40	3
1	2	3 (Bus)	50	23	70	2
0	2	4 (Car)	35	13	70	2

Restricting the choice set

- IIA test is performed by fitting the model to a restricted choice set
- Comparing the two sets of parameter estimates

- NLOGIT
 - ;Lhs = choice, cset, altij
 - ;Choices = tw, (bus), train, car
 - ;Rhs = TT, TC \$

- In out put, restricted choices are marked with “*”

Generalized Nested Logit Model

- Extension of Nested logit model
- Alternatives may appear in more than one branch

- **GNLOGIT**

;Lhs = Mode

;Choices=tw, train, bus, car

;Rhs=one,tt,tc,gc

;Tree=private(car,plane),ground(car,train,bus) \$