Phd Program in Transportation

Transport Demand Modeling

Filipe Moura
(fmoura@tecnico.ulisboa.pt)

Session 2

Discrete Choice Models

An Application of Multinomial Logit and Nested Logits

(Aknowledgements are given to Gonçalo Correia who prepared the previous version of this slides)

Application of a Multinomial Logit

Revealed Preference Experiments

- When calibrating discrete choice models we may recur to two different types of data: **Revealed Preference (RP)** or **Stated Preference (SP)**.

- **Revealed Preference:** We survey the population to what they are doing now, with the choice set they perceive to have available now. In mode choice (our most important choice in Transportation) this means having the attributes of the respondents, the attributes of the modes they have available (this is not straightforward) and the choices they make everyday.

A Modal Choice Revealed Preference Survey in Australia – Lab (I)

- This survey is part of the Book “Applied Choice Analysis: A Primer” by David Hensher, et al (2005). They have a Revealed and a Stated Preference Survey.

- We start just with the Revealed Preference (RP) survey for understanding how to estimate MNL models in NLOGIT (Econometric Software Inc.).

- Each respondent in the sample answered a questionnaire about the trip they had the day before of the survey.

- For this Lab we have filtered the data, selecting just some of the available explanatory variables.

A Modal Choice Revealed Preference Survey in Australia – Lab (II)

Variable	Meaning
ALTIJ	Alternative Number: **1 = DRIVE ALONE, 2 = RIDE SHARE, 3 = BUS, 4 = TRAIN, 5 = WALK, 6 = BICYCLE;**
CHOICE	1, chosen, 0 not chosen
CSET	Number of alternatives in each comparison. In this case there are always 2.
MPTRFARE	Cost of public transport ($AUS)
MPTRTIME	Time on public transport (min)
WAITTIME	Time waiting for public transport (min)
AUTOTIME	Time inside the automobile (min)
VEHPRKCT	Cost of parking in destination ($AUS)
VEHTOLCT	Paid toll ($AUS)
NUMBVEHS (SDC)	Number of vehicles in household
WKROCCUP (SDC)	Occupation category:1 = Managers and Admin, 2 = Professionals, 3 = Para-professional, 4 = Tradespersons, 5 = Clerks, 6 = Sales, 7 = Plant operators, 8 = Laborers , 9 = Other
PERAGE (SDC)	Age

INSTITUTO SUPERIOR TÉCNICO

FEUP

A Modal Choice Revealed Preference Survey in Australia – Lab (III)

Variable (continue)	Meaning
DISDWCBD (like an SDC)	Distance to the Central Business District (CBD) (km)
TRIPTIME	Trip time in Bicycle or walking (min)

- Variables ALTIJ, CHOICE and CSET, allow building the dependent variable of the DCM.

- ALTIJ will identify the mode of transportation that each line of data represents, CHOICE will tell which of the lines (modes) has been chosen from the Choice set, and finally CSET will tell how many lines (alternatives) are in each choice which respondents have answered.

- In this RP survey each choice is made between the mode of transportation that the user has selected the day before and in the questionnaire they were also asked to give attribute levels of a single alternative means of traveling to work as perceived by that respondent. This second mode was deemed the alternative mode.

Data Sample

ALTIJ	CHOICE	CSET	MPTRFARE	MPTRTIME	WAITTIME	AUTOTIME	VEHPRKCT	VEHTOLCT	NUMBVEHS	WKROCCUP	PERAGE	DISDWCBD	TRIPTIME
1	1	2	0	0	0	20	0	0	2	3	39	25	0
4	0	2	4,98	20	20	0	0	0	2	3	39	25	0
1	1	2	0	0	0	15	0	0	2	1	36	20	0
2	0	2	0	0	0	15	0	0	2	1	36	20	0
1	1	2	0	0	0	30	0	0	2	1	41	23	0
2	0	2	0	0	0	30	0	0	2	1	41	23	0
1	1	2	0	0	0	20	0	0	1	1	44	16	0
4	0	2	2,5	30	10	0	0	0	1	1	44	16	0
1	0	2	0	0	0	20	0	0	4	5	45	16	0
2	1	2	0	0	0	25	0	0	4	5	45	16	0
2	1	2	0	0	0	25	0	0	1	2	30	16	0
4	0	2	2,6	30	5	0	0	0	1	2	30	16	0
2	1	2	0	0	0	15	0	0	4	5	22	13	0
4	0	2	2	40	10	0	0	0	4	5	22	13	0
2	1	2	0	0	0	20	0,5	0	1	2	32	12	0
6	0	2	0	0	0	0	0	0	1	2	32	12	50
2	0	2	0	0	0	20	0	0	1	1	39	13	0
4	1	2	2,1	35	10	0	0	0	1	1	39	13	0
1	1	2	0	0	0	15	0	0	2	5	26	15	0
4	0	2	6	40	35	0	0	0	2	5	26	15	0
1	1	2	0	0	0	20	0	0	2	5	28	14	0
4	0	2	2	30	5	0	0	0	2	5	28	14	0
2	1	2	0	0	0	15	0	0	2	5	45	15	0
4	0	2	4	15	5	0	0	0	2	5	45	15	0
1	1	2	0	0	0	15	0	0	1	2	32	14	0
4	0	2	4,98	20	10	0	0	0	1	2	32	14	0
1	1	2	0	0	0	10	0	0	4	0	35	22	0
2	0	2	0	0	0	10	0	0	4	0	35	22	0
2	0	2	0	0	0	20	0	0	0	5	39	18	0
4	1	2	2,5	40	10	0	0	0	0	5	39	18	0

(…)

Building the Utility Functions for an MNL

The first step on running an MNL is thinking of an initial structure for the Utility functions. My proposal is the following:

Coefficient Variable

Drive alone Utility:
U(DA) =ASDR+TDRDA*AUTOTIME+**COST***VEHPRKCT+**COST***VEHTOLCT+VEHD*NUMBVEHS+AGE *PERAGE+MANAGE***MANAGERS**/

New Variable!

Ride Share Utility:
U(RS) = ASRS +**COST***VEHPRKCT+**COST***VEHTOLCT/

Bus Utility:
U(BUS) = ASBU +**COST***MPTRFARE + TPTBUS*MPTRTIME+TW*WAITTIME/

Train Utility:
U(TRAIN) = ASTR + **COST***MPTRFARE/ +TPTTRA*MPTRTIME+TW*WAITTIME+DISTAN*DISDWCBD

Walk Utility:
U(WALK) = ASWA+TRPEDBI*TRIPTIME/

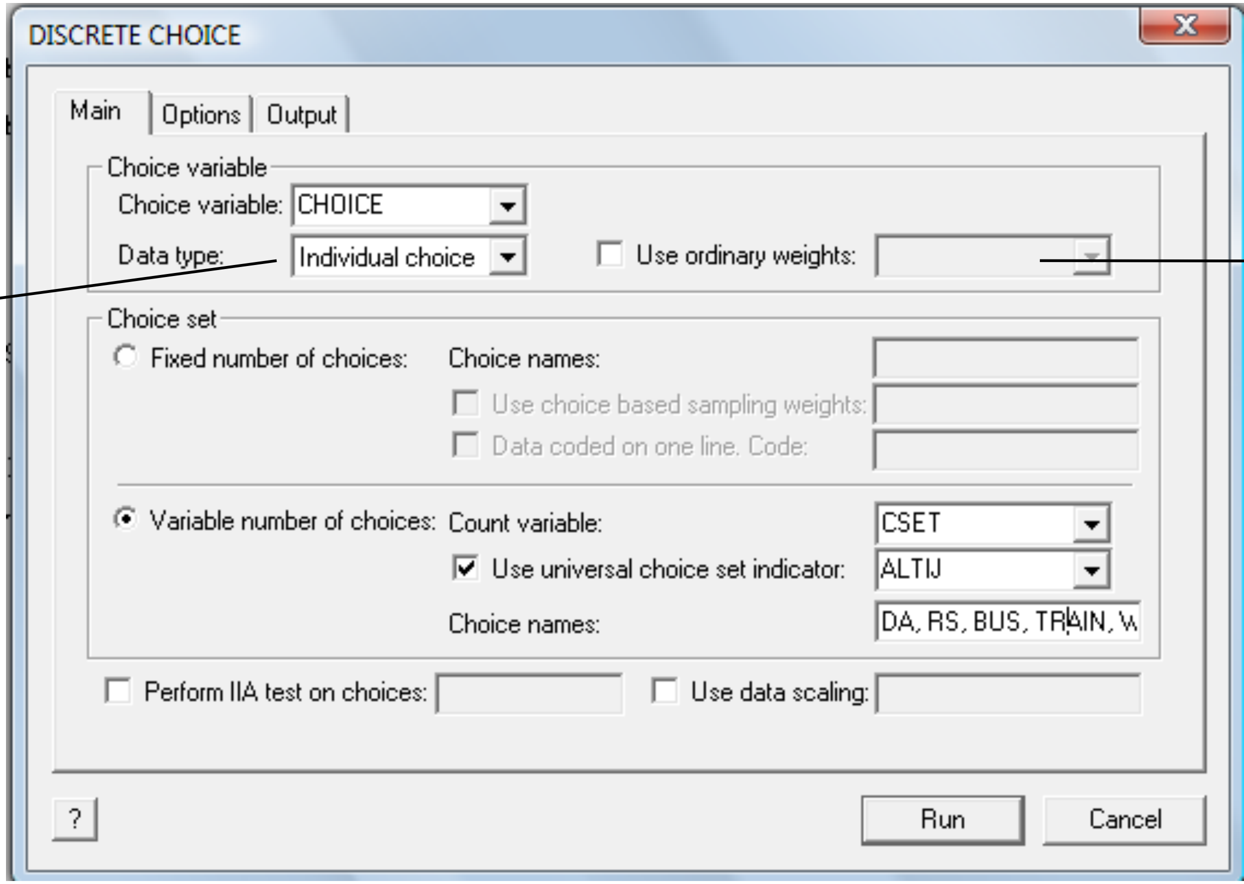
Cycle Utility:
U(CYCLE) = TRPEDBI*TRIPTIME

Reference Alternative

MNL's in Nlogit

- The command in Nlogit to create MNL models is the DISCRETECHOICE (or NLOGIT) command. Go to Model->Discrete Choice->Discrete Choice.

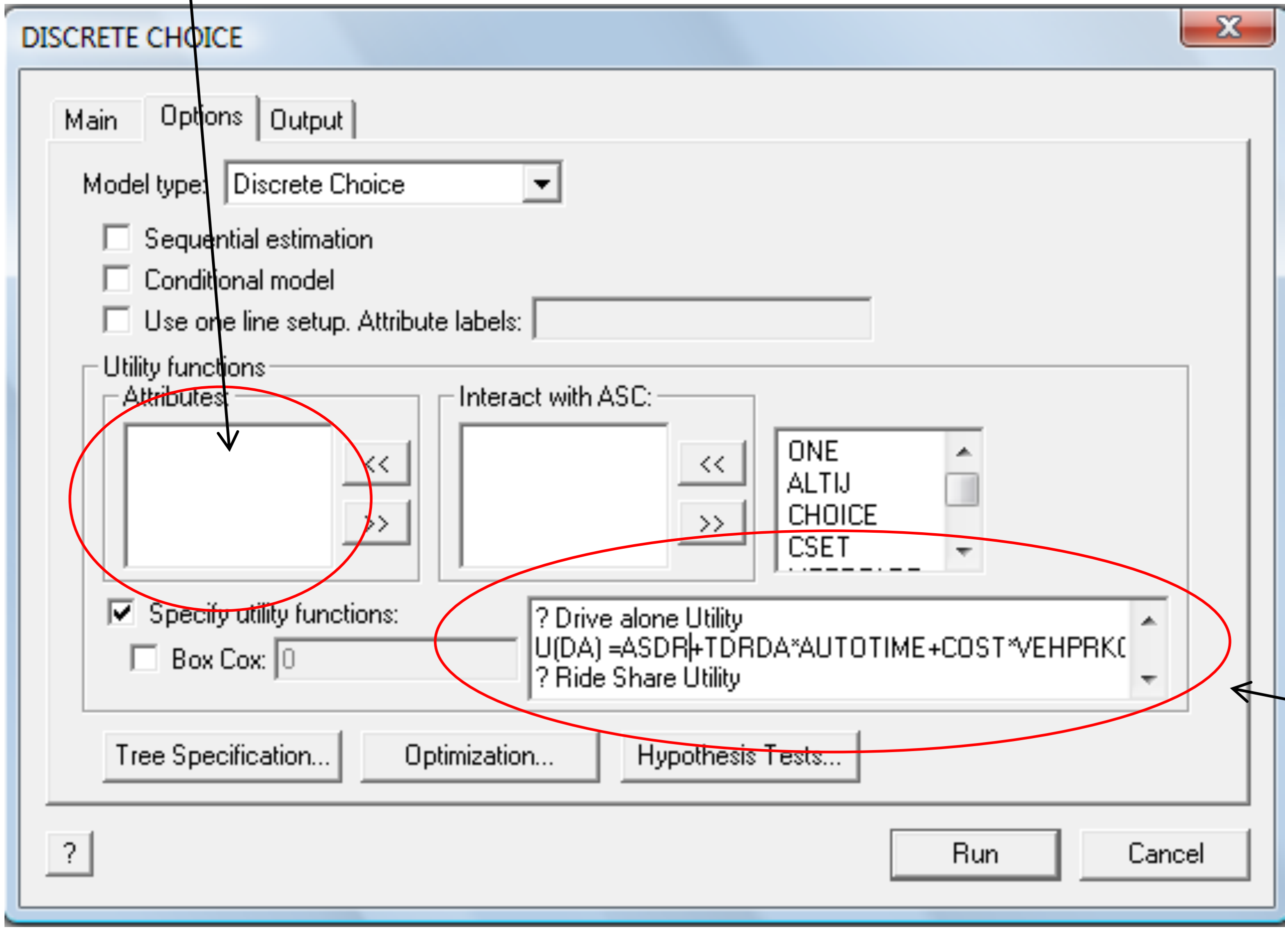
Instead of having the choice we could have a frequency of choices, which Nlogit transforms in probabilities.



If we want to give a higher weigh to some choices, for instance, due to sample stratification

MNL's in Nlogit

It does not allow to specify different weighs for the same alternative attribute like travel time of BUS and travel time of Car.



Best Option! We have all the freedom we want

The code in Nlogit

But the best to do is really to write the code it self:

```
SAMPLE ; all$
? Create new variable
CREATE
;if(WKROCCUP<2)MANAGERS=1
;if(WKROCCUP>1)MANAGERS=0$
? Multinomial Logit
DISCRETECHOICE
;lhs=CHOICE, CSET, ALTIJ
;choices=DA,RS,BUS,TRAIN,WALK,CYCLE
;MODEL:
? Drive alone Utility
U(DA) =ASDR+TDRDA*AUTOTIME+COST*VEHPRKCT+COST*VEHTOLCT+VEHD*NUMBVEHS+AGE*PERAGE+MANAGE*MANAGERS/
? Ride Share Utility
U(RS) = ASRS +COST*VEHPRKCT+COST*VEHTOLCT/
? Bus Utility
U(BUS) = ASBU +COST*MPTRFARE + TPTBUS*MPTRTIME+TW*WAITTIME/
? Train Utility
U(TRAIN) = ASTR + COST*MPTRFARE + TPTTRA*MPTRTIME+TW*WAITTIME+DISTAN*DISDWCBD/
? Walk Utility
U(WALK) = ASWA+TRPEDBI*TRIPTIME/
? Cycle Utility
U(CYCLE) = TRPEDBI*TRIPTIME
;Crosstab
;Describe
;Prob=Prob
;CHECKDATA$
```

- SAMPLE ; all$ → Considers all the data for estimating the model
- CREATE → Creates the new variable according to an existing one. We should not introduce a categorical variable directly in the model because we are implicitly considering a linear effect of the categories on the utility which is most of the times false.
- ;Crosstab → Cross tabulation of true versus predicted choices
- ;Describe → Describes all utility functions and their coefficients estimates
- ;Prob=Prob → Will produce a new variable called “Prob” with the probability of the alternative being chosen in its choice set
- ;CHECKDATA$ → Verifies the data before estimation

❑Run the model by selecting all text with your cursor and pressing go!

The output

```
+--------------------------------------------+
| Discrete choice and multinomial logit models|
+--------------------------------------------+
+-------------------------------------------------------+
| Inspecting the data set before estimation.            |
| These errors mark observations which will be skipped.  |
| Row Individual = 1st row then group number of data block |
+-------------------------------------------------------+
No bad observations were found in the sample

Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Discrete choice (multinomial logit) model    |
| Maximum Likelihood Estimates                 |
| Model estimated: Oct 30, 2010 at 04:55:50PM.|
| Dependent variable               Choice     |
| Weighting variable                 None     |
| Number of observations              854     |
| Iterations completed                  7     |
| Log likelihood function       -344.1835     |
| Number of parameters                 15     |
| Info. Criterion: AIC =           .84118     |
|   Finite Sample: AIC =           .84185     |
| Info. Criterion: BIC =           .92461     |
| Info. Criterion:HQIC =           .87313     |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| Constants only.  Must be computed directly. |
|                  Use NLOGIT ;...; RHS=ONE $ |
| Chi-squared[10]         =     1473.22277    |
| Prob [ chi squared > value ] =    .00000    |
| Response data are given as ind. choice.     |
| Number of obs.=   854, skipped   0 bad obs. |
+---------------------------------------------+
```

It is the Log Likelihood

Choice Makers, not data lines

Be careful this Chi-squared is not computed correctly. It is supposed to be the test comparing LL(c) and LL(*) that is why it has 10 degrees of freedom: 15-5. However what is doing is wrong! He is applying the following: 2*(LL(c)+LL(*))=2*(392,51+344,18)=1473

Comparing to a model with equal shares

- Due to the error in Nlogit we must compute ourselves, the Log Likelihood ratio and the pseudo-R2 for a base model with equal shares.

- Each respondent chooses one of two alternatives, thus equal shares means 0.5 probability of choice in each choice set for each of the alternatives, thus we have:

$LL(0) = (1 * \ln(0.5) + 0 * \ln(0.5)) * 854$=-591,94

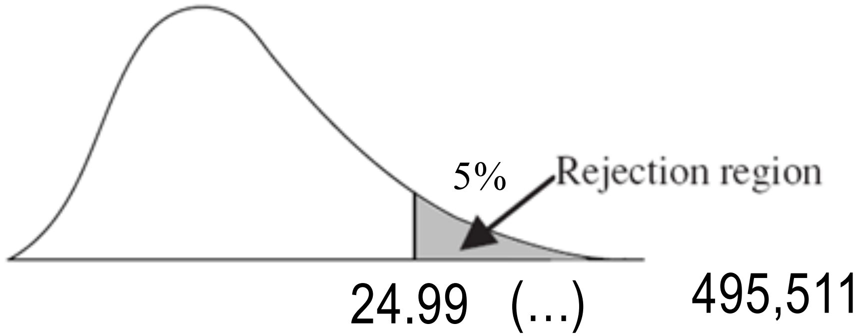
- The Log Likelihood ratio will then be:

$-2(L(0) - L(*)) = -2(-591.94 - (-344.1845)) = 495{,}511$

With degrees of freedom=15-0=15

In Excel:
=INV.CHI(0.05,15)=24,99

We reject the hypothesis that our model is the same as a base model with equal shares.



Pseudo R^2 (McFadden)= 1-(-344.1845/-591,94)=0.419

Comparing to a model with ASCs

- NLOGIT;Lhs=CHOICE,CSET,ALTIJ;Choices=DA,RS,BUS,TRAIN,WALK ,CYCLE;Rh2=ONE$

```
+---------------------------------------------+
| Discrete choice (multinomial logit) model   |
| Maximum Likelihood Estimates                 |
| Model estimated: Oct 30, 2010 at 05:53:26PM.|
| Dependent variable               Choice     |
| Weighting variable                 None     |
| Number of observations              854     |
| Iterations completed                  5     |
| Log likelihood function       -392.5100     |
| Number of parameters                  5     |
| Info. Criterion: AIC =           .93094     |
|   Finite Sample: AIC =           .93102     |
| Info. Criterion: BIC =           .95875     |
| Info. Criterion:HQIC =           .94159     |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| Constants only.  Must be computed directly. |
|                  Use NLOGIT ;...; RHS=ONE $ |
| Response data are given as ind. choice.     |
| Number of obs.=   854, skipped   0 bad obs. |
+---------------------------------------------+
```

$-2(L(c) - L(*)) =$

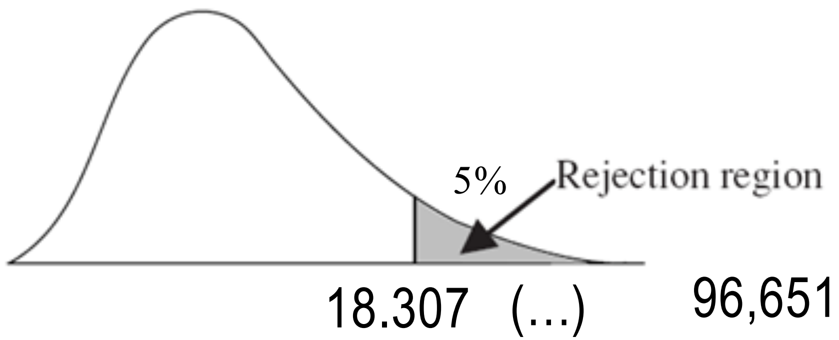
$-2(\mathbf{-392.51} - (-344.1845)) =$

$= 96.651$

degrees of freedom=15-5=10

In Excel:
=INV.CHI(0.05,10)=18.307

We reject the hypothesis that our model is the same as a base model with just the alternative specific constants.



Crosstab of Actual vs Predicted

- The predicted choices are obtained by computing the probabilities and if $P(i) > P(j)\ \forall\ j \epsilon J$ we say that alternative i is chosen.

```
+--------------------------------------------------------+
| Cross tabulation of actual vs. predicted choices.      |
| Row indicator is actual, column is predicted.          |
| Predicted total is F(k,j,i)=Sum(i=1,...,N) P(k,j,i).   |
| Column totals may be subject to rounding error.        |
+--------------------------------------------------------+
```

Predicted

Actual		DA	RS	BUS	TRAIN	WALK	CYCLE	Total	
	DA	446	19	17	11	3	3	499	58,4%
	RS	11	79	16	21	3	4	135	15,8%
	BUS	20	15	53	7	0	0	95	11,12%
	TRAIN	15	18	9	33	2	0	78	9,13%
	WALK	3	2	0	2	9	4	20	2,34%
	CYCLE	4	2	0	4	3	15	27	3,16%
	Total	499	135	95	78	20	27	854	
		58,4%	15,8%	11,12%	9,13%	2,34%	3,16%		

$$\%\ Correct\ Predictions = \frac{635}{854} = 74\%$$

Good prediction ability.

MNL Coefficients results

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
ASDR	-1.42080681	.80538373	-1.764	.0777
TDRDA	-.02112582	.00658611	-3.208	.0013
COST	-.15101493	.05036466	-2.998	.0027
VEHD	.77097604	.18661110	4.131	.0000
AGE	.01658750	.01351151	1.228	.2196
MANAGE	-.27113166	.34478143	-.786	.4316
ASRS	-1.73200621	.56851134	-3.047	.0023
ASBU	-1.62077187	.64099298	-2.529	.0115
TPTBUS	.01524565	.00751420	2.029	.0425
TW	-.04005296	.01846677	-2.169	.0301
ASTR	-1.30254759	.63557526	-2.049	.0404
TPTTRA	-.01916878	.00863789	-2.219	.0265
DISTAN	.00582623	.00495277	1.176	.2395
ASWA	-.19848961	.36459281	-.544	.5862
TRPEDBI	-.09000302	.02225632	-4.044	.0001

Number of vehicles in the household is very positive for choosing to drive alone

Not being able to explain part of the disutility of these modes against the reference alternative, the Bicycle

Time walking and bicycling is very negative for those alternatives

Experiment at home: Try considering one alternative specific coefficient for the travel time in bicycle and another for the time walking … see what happens. Determine the value of driving time alone.

Wald Statistic

Test if we can reject the hypothesis of the cost parameter being zero:

WALD;FN1:COST-0$

```
+---------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of    |
| nonlinear restrictions.                      |
| Wald Statistic              =      8.99058   |
| Prob. from Chi-squared[ 1] =       .00271   |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |    -.15101493       .05036466    -2.998   .0027
```

This results in exactly the same value as in the previous table. Because we are testing the same hypothesis.

Test if we can reject the hypothesis of the travel time inside the bus and the time waiting for the bus have the same weight in the Utility function:

WALD;FN1:TPTBUS-TW$

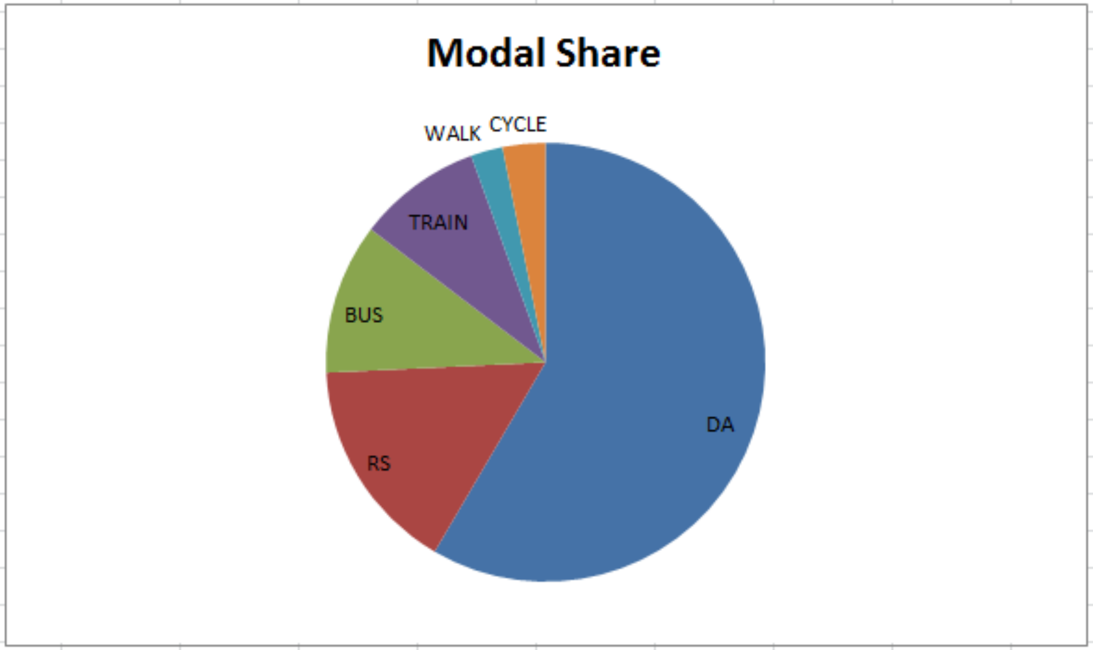
```
+---------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of    |
| nonlinear restrictions.                      |
| Wald Statistic              =      6.72661   |
| Prob. from Chi-squared[ 1] =       .00950   |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |     .05529862       .02132141     2.594   .0095
```

We reject the hypothesis that both are the same

Aggregating across Individuals

- With the Prob variable we may aggregate across individuals and obtain the modal shares using Excel:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1				DA	RS	BUS	TRAIN	WALK	CYCLE				
2	ALTIJ	PROB		1	2	3	4	5	6				
3	1	0,96889910000	1	58%	16%	11%	9%	2%	3%				
4	4	0,03110090000							100%				
5	1	0,86556280000											
6	2	0,13443720000	854										
7	1	0,83594110000											
8	2	0,16405890000											
9	1	0,87531890000											
10	4	0,12468110000											
11	1	0,97631690000											
12	2	0,02368310000											
13	2	0,65594800000											
14	4	0,34405200000											
15	2	0,72394410000											
16	4	0,27605590000											
17	2	0,93659080000											
18	6	0,06340920000											
19	2	0,70751980000											
20	4	0,29248020000											
21	1	0,98928550000											
22	4	0,01071450000											
23	1	0,92139240000											



- See that the model is reproducing the shares in the sample. This must always happen because it is a direct result of the estimation process.

Application of a Nested Logit

A mode choice SP experiment Australia (I)

SA101	*1. car, toll road*		*2. car, non-toll road*	
Travel time to work	10 min.		15 min.	
Time variability	None		None	
Toll (one way)	$1.00		Free	
Pay toll if you leave at this time (otherwise free)	6–10 am		–	
Fuel cost (per day)	$1.00		$3.00	
Parking cost (per day)	Free		Free	
	3. bus		*4. train*	
Total time in the vehicle (one way)	10 min.		10 min.	
Time from home to your closest stop	Walk 5 min.	Car/Bus 4 min.	Walk 5 min.	Car/Bus 4 min.
Time to your destination from the closest stop	Walk 5 min.	Bus 4 min.	Walk 5 min.	Bus 4 min.
Frequency of service	Every 5 min.		Every 5 min.	
Return fare	$1.00		$1.00	

(Hensher et al, 2005)

- The experiment showed 4 transportation alternatives to the respondents in several cities in Australia.
- However the choice set was between: Car with toll; Car with no toll; bus; train; busway and light rail (these last two, were non-existent at that time).

INSTITUTO SUPERIOR TÉCNICO

FEUP

A mode choice SP experiment Australia (II)

- To initiate this experiment the trip length in terms of travel time relevant for each respondent’s current commuting trip was first established so that the travel choices could be given in a context that had some reality for the respondent. The travel choice sets were divided into trip lengths of:
  - Less than 30 minutes: Short trip
  - 30–45 minutes: Medium trip
  - Over 45 minutes: Long trip
- In participating in the choice experiments, each respondent was asked to consider a context in which the offered set of attributes and levels represented the only available means of undertaking a commuter trip from the current residential location to the current workplace location. It was made clear that the purpose was to establish each respondent’s coping strategies under these circumstances.

The set of attributes and their levels

short (<30 mins.)	*car no toll*	*car toll rd*	*public transport*	*bus*	*train*	*busway*	*light rail*
Travel time to work	15, 20, 25	10, 12, 15	Total time in the vehicle (one-way)	10, 15, 20	10, 15, 20	10, 15, 20	10, 15, 20
Pay toll if you leave at this time (otherwise free)	None	6–10, 6:30–8:30, 6:30–9	Frequency of service	Every 5, 15, 25	Every 5, 15, 25	Every 5, 15, 25	Every 5, 15, 25
Toll (one-way)	None	1, 1.5, 2	Time from home to closest stop	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25
Fuel cost (per day)	3, 4, 5	1, 2, 3	Time to destination from closest stop	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25
Parking cost (per day)	Free, $10, 20	Free, $10, 20	Return fare	1, 3, 5	1, 3, 5	1, 3, 5	1, 3, 5
Time variability	0, ±4, ±6	0,±1,±2					
medium (30–45 mins.)	*car no toll*	*car toll rd*	*public transport*	*bus*	*train*	*busway*	*light rail*
Travel time to work	30, 37, 45	20, 25, 30	Total time in the vehicle (one-way)	20, 25, 30	20, 25, 30	20, 25, 30	20, 25, 30
Pay toll if you leave at this time (otherwise free)	None	6–10, 6:30–8:30, 6:30–9	Frequency of service	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25
Toll (one-way)	None	2, 3, 4	Time from home to closest stop	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25
Fuel cost (per day)	6, 8, 10	2, 4, 6	Time to destination from closest stop	Walk 5, 15, 25 Bus 4, 6, 8	Walk 5, 15, 25 Bus 4, 6, 8	Walk 5, 15, 25 Bus 4, 6, 8	Walk 5, 15, 25 Bus 4, 6, 8
Parking cost (per day)	Free, $10, 20	Free, $10, 20	Return fare	2, 4, 6	2, 4, 6	2, 4, 6	2, 4, 6
Time variability	0, ±7, ±11	0, ±2, ±4					
long (>45 mins.)	*car no toll*	*car toll rd*	*public transport*	*bus*	*train*	*busway*	*light rail*
Travel time to work	45, 55, 70	30, 37, 45	Total time in the vehicle (one-way)	30, 35, 40	30, 35, 40	30, 35, 40	30, 35, 40
Pay toll if you leave at this time (otherwise free)	None	6–10, 6:30–8:30, 6:30–9	Frequency of service	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25
Toll (one-way)	None	3, 4.5, 6	Time from home to closest stop	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25	Walk 5, 15, 25
Fuel cost (per day)	9, 12, 15	3, 6, 9	Time to destination from closest stop	Walk 5, 15, 25 Bus 4, 6, 8	Walk 5, 15, 25 Bus 4, 6, 8	Walk 5, 15, 25 Bus 4, 6, 8	Walk 5, 15, 25 Bus 4, 6, 8
Parking cost (per day)	Free, $10, 20	Free, $10, 20	Return fare	3, 5, 7	3, 5, 7	3, 5, 7	3, 5, 7
Time variability	0, ±11, ±17	0, ±7, ±11					

(Hensher et al, 2005)

INSTITUTO SUPERIOR TÉCNICO

FEUP

Stated preference database

- We will use this data in the next session to use Nested Logit models.

ID	CITY	SPRP	SPEXP	ALTISPRP	ALTIJ	SPCHOICE	CHOICE	CSET	SPCART	SPCARNT	SPBUS	SPTN	SPBW
1000	1	2	1	7	1	2	0	4	0	1	0	0	0
1000	1	2	1	8	2	2	1	4	0	1	0	0	0
1000	1	2	1	10	4	2	0	4	0	1	0	0	0
1000	1	2	1	12	6	2	0	4	0	1	0	0	0
1000	1	2	2	7	1	6	0	4	0	0	0	0	0
1000	1	2	2	8	2	6	0	4	0	0	0	0	0
1000	1	2	2	11	5	6	0	4	0	0	0	0	0
1000	1	2	2	12	6	6	1	4	0	0	0	0	0
1000	1	2	3	7	1	4	0	4	0	0	0	1	0
1000	1	2	3	8	2	4	0	4	0	0	0	1	0
1000	1	2	3	9	3	4	0	4	0	0	0	1	0
1000	1	2	3	10	4	4	1	4	0	0	0	1	0
1001	1	2	1	7	1	2	0	4	0	1	0	0	0
1001	1	2	1	8	2	2	1	4	0	1	0	0	0
1001	1	2	1	10	4	2	0	4	0	1	0	0	0
1001	1	2	1	12	6	2	0	4	0	1	0	0	0
1001	1	2	2	7	1	1	1	4	1	0	0	0	0
1001	1	2	2	8	2	1	0	4	1	0	0	0	0
1001	1	2	2	9	3	1	0	4	1	0	0	0	0
1001	1	2	2	10	4	1	0	4	1	0	0	0	0
1001	1	2	3	7	1	6	0	4	0	0	0	0	0
1001	1	2	3	8	2	6	0	4	0	0	0	0	0
1001	1	2	3	11	5	6	0	4	0	0	0	0	0
1001	1	2	3	12	6	6	1	4	0	0	0	0	0

(…)

(…)

INSTITUTO SUPERIOR TÉCNICO

FEUP

The Hausman IIA Test of the IIA hypothesis (I)

- As we have seen Logit is built under the hypothesis that the error terms on the utility functions are IID which has the important behavioral imposition of the IIA hypothesis for multinomial Logit models.
- The main test for verifying if this is acceptable or not, is the **Hausman-test**. This encompasses two main steps:
  - First step – Run an unrestricted model with all the alternatives included
  - Second step – Run a model with only a restricted number of alternatives and compute the statistic *q* (shown in next slide). Compare the statistic with a Chi-Squared distribution.
- What the test does is to see if extracting one alternative from the choice set changes the relation between the remaining alternatives. If it changes we reject the IIA.

The Hausman IIA Test of the IIA hypothesis (II)

- The test statistic:

transposed   inverse

$$q = [b_u - b_r]' \, [V_r - V_u]^{-1} [b_u - b_r]$$

Column vector of the coefficients estimates for the unrestricted model

Column vector of the coefficients estimates for the restricted model

Variance-Covariance matrix for the restricted and unrestricted model

(not including alternative specific coefficients)

- The test statistic *q*, is given as a Chi-Squared statistic with the number of degrees of freedom equal to the number of estimated parameters.

The Hausman IIA Test of the IIA hypothesis (III)

- Nlogit has an option which allows to use this test in the MNL menu, but is very limited and fails a lot of times because it is not prepared for all ways of representing survey information in Nlogit. The best option is to specify the test itself.
- Let's consider the following model applied to the SP Lab data (Australia's SP experiment):

```
NLOGIT
;lhs= choice, cset, altij
;Choices = cart, carnt, bus, train, busway, LR
;Model:
U(cart) = asccart + fuel*fuel /
U(carnt) = asccarnt + fuel*fuel /
U(bus) = ascbus + fare*fare /
U(train) = asctn + fare*fare /
U(busway) = ascbusw + fare*fare /
U(LR) = fare*fare $
```

ASCCART
FUEL
ASCCARNT
ASCBUS
FARE
ASCTN
ASCBUSW

- This is the unconstrained model because it has all the alternatives

The Hausman IIA Test of the IIA hypothesis (IV)

- For each estimated model, NLOGIT saves the parameters in a column vector named *B* (this matrix is overwritten each time a new model is estimated).

- The *B* column vector may be accessed via the project dialog box under the Matrices folder.

- Remember that the Hausman-test of the IIA assumption is performed using only those parameters which are not constant terms. As such, not all of the parameters of the *B* column vector are required (i.e. b_u and b_r do not include any parameters that are constant terms).

- For the above example, b_u will have only two parameters, those being for the fuel and fare attributes.

The Hausman IIA Test of the IIA hypothesis (V)

- To construct b_u, we must first create a permutation matrix, *J1,* which is used to extract the relevant parameters from the *B* column vector. The number of rows of the *J1* matrix will be equal to the number of parameters required for the b_u column vector, and the number of columns will equal the number of parameters present within the *B* column vector. This is because each column of the *J1* matrix is associated with each row of the *B* matrix (e.g. column one of the *J1* matrix is related to the *asccart* parameter, column two to the fuel parameter etc.).
- In the construction of the *J1* matrix, a zero is placed in each column associated with a row in the *B* column vector for which the parameter in *B* is to be discarded. A "1" means that the related parameter located in the *B* column vector is to be retained. For the above example, the *J1* matrix is shown below. The command **MATRIX; is used to either create** or manipulate matrices in NLOGIT.

MATRIX; J1 = [0,1,0,0,0,0,0 /

0,0,0,0,1,0,0] $

The Hausman IIA Test of the IIA hypothesis (VI)

- Once *J1* is created, b_u and V_u (as with b_u , the elements of the variance–covariance matrix for the test, V_u , are not inclusive of elements related to constant terms) are created using the **MATRIX** command:

MATRIX; Bu = J1*B $

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -0.226588 \\ -0.202418 \\ 0.379942 \\ 0.0314588 \\ -0.193473 \\ -4.09564e-006 \\ 0.111259 \end{bmatrix} = \begin{bmatrix} -0.202418 \\ -0.193473 \end{bmatrix}$$

The Hausman IIA Test of the IIA hypothesis (VII)

- The variance-covariance matrix is stored in VARB, under Matrices, thus:

Vu = J1*VARB*J1' $

$$
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}
\times
\begin{bmatrix}
0.00686456 & -0.00037865 & 0.00646231 & 0.00265986 & 0.000692354 & 0.00227902 & 0.00221727 \\
-0.00037865 & 0.000204405 & -0.000887322 & 1.24995e-005 & 6.04582e-005 & -5.63867e-006 & 6.61332e-006 \\
0.00646231 & -0.000887322 & 0.00982311 & 0.0026238 & 0.000517349 & 0.00229555 & 0.00219447 \\
0.00265986 & 1.24995e-005 & 0.0026238 & 0.00597961 & 1.21631e-005 & 0.00278552 & 0.00254353 \\
0.000692354 & 6.04582e-005 & 0.000517349 & 1.21631e-005 & 0.000264105 & -1.61575e-005 & 8.09373e-006 \\
0.00227902 & -5.63867e-006 & 0.00229555 & 0.00278552 & -1.61575e-005 & 0.00523008 & 0.0018625 \\
0.00221727 & 6.61332e-006 & 0.00219447 & 0.00254353 & 8.09373e-006 & 0.0018625 & 0.0043179
\end{bmatrix}
$$

$$
\times
\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}
=
\begin{bmatrix} 0.000204405 & 6.04582e-005 \\ 6.04582e-005 & 0.000264105 \end{bmatrix}
$$

The Hausman IIA Test of the IIA hypothesis (VIII)

- Now we have to do the same for a **restricted model**. To estimate the restricted model we must take out all observations (i.e. choices) from the sample for which the alternative(s) which are to be removed were chosen.
- Fortunately the data base has a series of dummy variables which are equal to one for choice sets in which specific alternatives where chosen (or zero otherwise) we have decided to take out the **Bus** alternative (arbitrary). For the bus alternative, this variable is called **spbus**. <u>But be careful you can´t take out an alternative which leaves the remaining model undetermined!</u>
- As well as removing choice sets in which the alternative(s) to be removed (in this case is the Bus) were chosen, we are also required to remove any rows of data related to the alternative(s) to be removed from choice sets in which those alternatives were not chosen.
- Assuming that we are to remove the bus alternative (altij equal to three), the following **REJECT** commands will remove all reference to the bus alternative from the sample to be used:

The Hausman IIA Test of the IIA hypothesis (IX)

REJECT; spbus=1 $
REJECT; altij=3 $

Rejects the lines of the survey where the respondent answered bus. In this case all 4 lines of the experiment are erased

Rejects the lines of the survey where the bus option appeared but was not chosen

- The rejection of an alternative such as the bus alternative means that for observations where that alternative was once present, the choice set size will be smaller by the number of alternatives removed;
- (e.g. if bus was one of four alternatives, removing this alternative will leave three remaining alternatives and hence the choice set size decreases from four to three). This has implications for both altij and cset in the model commands, neither of which can be used without modification.

The Hausman IIA Test of the IIA hypothesis (X)

```
CREATE
;if(altij<3)altijn=altij
;if(altij>3)altijn=altij-1
;if(cn<3)csetn=3
;if(cn>2)csetn=cset $
```

Alternatives are renumbered

If the experiment had Bus, cn<3 then the choice set should be reduced to three.

The restricted model should now be:

```
NLOGIT
;lhs= choice, csetn, altijn
;Choices = cart, carnt, train, busway, LR
;Model:
U(cart) = asccart + fuel*fuel /
U(carnt) = asccarnt + fuel*fuel /
?U(bus) = ascbus + fare*fare /
U(train) = asctn + fare*fare /
U(busway) = ascbusw + fare*fare /
U(LR) = fare*fare $
```

Bus has disappeared

This utility disapears

The Hausman IIA Test of the IIA hypothesis (XI)

The resulting coeficients are:

ASCCART
FUEL
ASCCARNT
ASCTN
FARE
ASCBUSW

- To obtain the required b_r and V_r matrices we may use the same method we employed to obtain b_u and V_u. The number of parameters for the restricted model is smaller by the number of alternative specific parameters related to alternatives included in the unrestricted model. As such, the number of columns in the permutation matrix will also be fewer by this number.

The Hausman IIA Test of the IIA hypothesis (XII)

MATRIX; J2 = [0,1,0,0,0,0 /
0,0,0,0,1,0] $

MATRIX; Br = J2*B $

MATRIX; Vr = J2*VARB*J2’ $

$$q = [b_u - b_r]'[V_r - V_u]^{-1}[b_u - b_r]$$

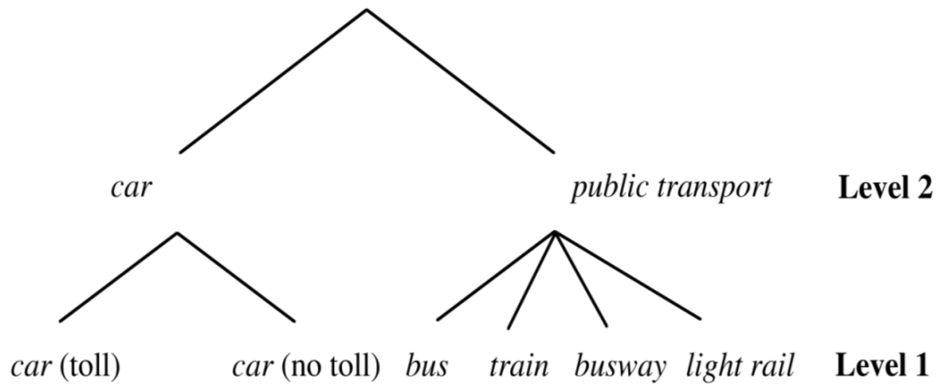
- **MATRIX; bd = Bu-Br ;Vd = Vr-Vu $**
- **MATRIX; vdinv=[vd] $** → Matrix inverse
- **MATRIX; list; q = bd’*vninv*bd $**
- **CALC; list; p = 1-chi(q,2) $** → Computes the p value (all probability to the right) for the test statistic q for a chi-squared with two degrees of freedom (two parameters), which are the estimated parameters for building p

```
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
P        =         .026792
```

Reject the IIA hypothesis!

Applying the Nested Logit Structure to the SP data experiment

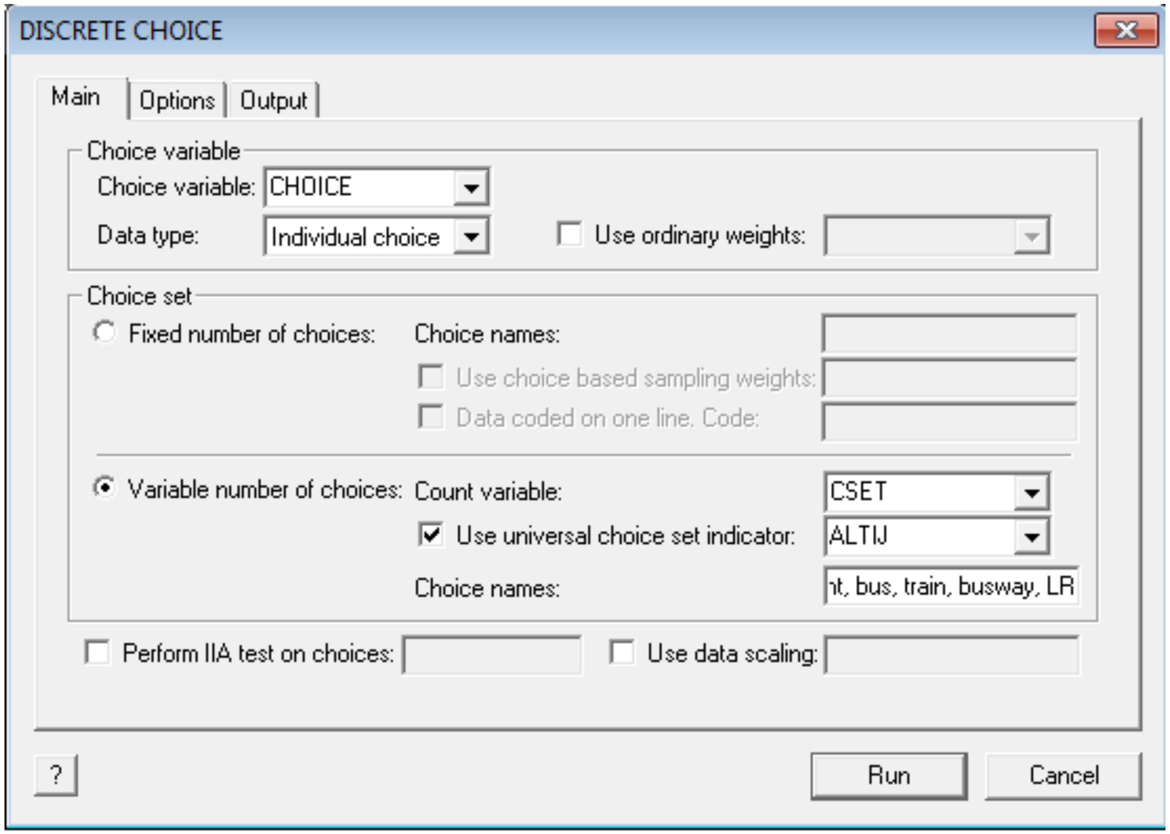
- The decision regarding the shape of the nested structure is the analyst choice. We first start with a structure which has proved many times to be significant:



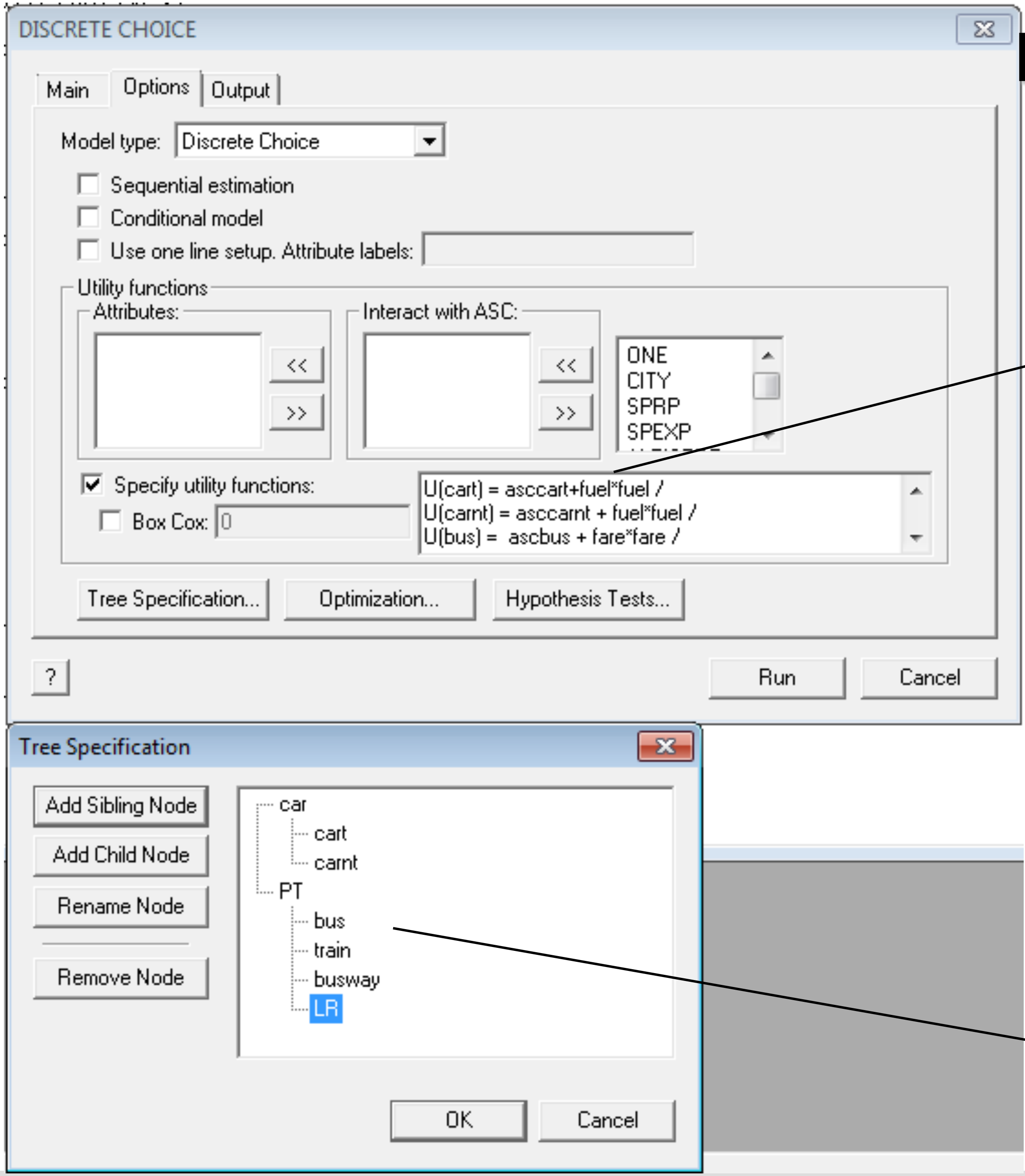
- In this perspective we are saying that both car options are correlated in their error components, that is, all the attributes that we are forgetting in the systematic part of utility may be correlated in that branch.
- We are proposing the same for the public transport branch.

First NL model (I)

- The menu to run a Nested Model in Nlogit: Model -> Discrete Choice -> Nested Logit.
- In the first screen you do exactly the same as you did for the MNL's.



First NL model (II)

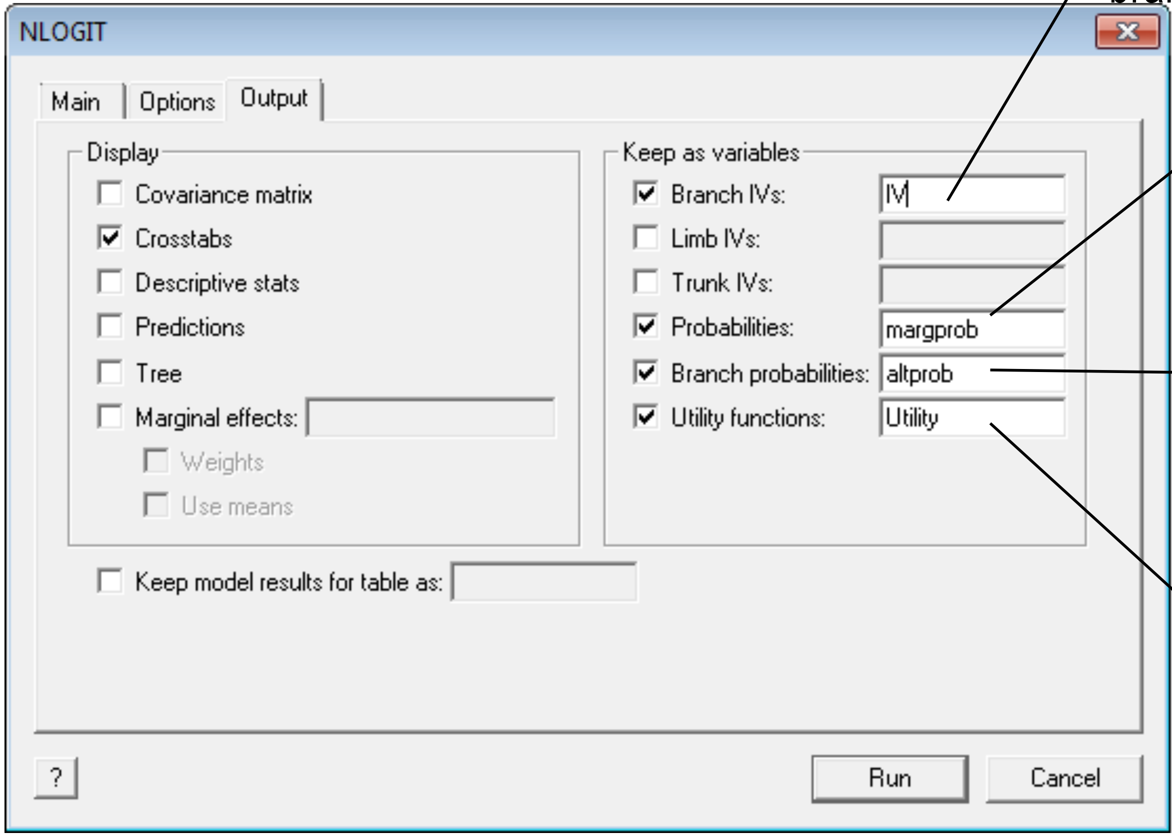


Utility Functions

It does not allow to specify RU1 or RU2. By default it will use RU1: normalization in the Level 1 alternatives!

The tree

First NL model (III)



The IV values of each branch

Marginal probabilities: The probabilities of each of the four alternatives in each comparison (sum=1)

Conditional probability: The probabilities for each alternative given that they are inside each branch

The Utility of each alternative

First NL model (IV)

```
NLOGIT
;lhs = choice, cset, altij
;choices = cart, carnt, bus, train, busway, LR
;tree = car(cart, carnt),PT(bus,train, busway,LR)
;RU1
;start = logit
;ivset: (car)=[1.0]
;maxit = 100
;Prob=MARGPROB
;cprob=ALTPROB
;ivb=IVBRANCH
;Utility=U1
;model:
U(cart) = asccart + fuel*fuel /
U(carnt) = asccarnt + fuel*fuel /
U(bus) =  ascbus + fare*fare /
U(train) = asctn + fare*fare /
U(busway) = ascbusw + fare*fare /
U(LR) = fare*fare
;Crosstab$
```

Branch1(alt1, alt2),Branch2(alt3, (…))

Begins by running an MNL for having initial values for estimating the parameters.

IV parameter normalization: the IV parameter will be 1 which normalizes the scale of the car branch to 1.

We choose to normalize the IV parameter of the car branch because it has the highest scale.

First NL model (V)

```
+----------------------------------------------------+
| Discrete choice (multinomial logit) model          |
| Maximum Likelihood Estimates                       |
| Model estimated: Nov 03, 2010 at 11:51:30AM.       |
| Dependent variable                  Choice         |
| Weighting variable                    None         |
| Number of observations                3587         |
| Iterations completed                     5         |
| Log likelihood function          -4804.548         |
| Number of parameters                     7         |
| Info. Criterion: AIC =             2.68277         |
|   Finite Sample: AIC =             2.68278         |
| Info. Criterion: BIC =             2.69484         |
| Info. Criterion:HQIC =             2.68707         |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj        |
| Constants only.  Must be computed directly.        |
|                  Use NLOGIT ;...; RHS=ONE $        |
| Chi-squared[ 2]          =    2981.12332           |
| Prob [ chi squared > value ] =    .00000           |
| Response data are given as ind. choice.            |
| Number of obs.=  3587, skipped   0 bad obs.        |
+----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
ASCCART	-.22658772	.08285265	-2.735	.0062
FUEL	-.20241829	.01429703	-14.158	.0000
ASCCARNT	.37994204	.09911161	3.833	.0001
ASCBUS	.03145882	.07732794	.407	.6841
FARE	-.19347334	.01625132	-11.905	.0000
ASCTN	-.409564D-05	.07231932	.000	1.0000
ASCBUSW	.11125885	.06571070	1.693	.0904

- ❑ Nlogit will start by calibrating an MNL to generate initial coefficients for the iterative calibration.
- ❑ It gives exactly the same results as if you run an MNL with those 6 alternatives.

First NL model (VI)

```
+-------------------------------------------------------------+
| FIML Nested Multinomial Logit Model                         |
| Maximum Likelihood Estimates                                |
| Model estimated: Nov 03, 2010 at 11:53:18AM.                |
| Dependent variable                 CHOICE                   |
| Weighting variable                   None                   |
| Number of observations               3587                   |
| Iterations completed                   15                   |
| Log likelihood function         -4768.225                   |
| Number of parameters                    8                   |
| Info. Criterion: AIC =            2.66307                   |
|   Finite Sample: AIC =            2.66309                   |
| Info. Criterion: BIC =            2.67687                   |
| Info. Criterion:HQIC =            2.66799                   |
| Restricted log likelihood       -6324.968                   |
| McFadden Pseudo R-squared        .2461267                   |
| Chi squared                      3113.487                   |
| Degrees of freedom                      8                   |
| Prob[ChiSqd > value] =           .0000000                   |
| Constants only.  Must be computed directly.                 |
|                  Use NLOGIT ;...; RHS=ONE $                 |
| At start values  -4804.5480  .00756 *******                 |
| Response data are given as ind. choice.                     |
+-------------------------------------------------------------+
```

$$Pseudo\ R^2 = 1 - \frac{L(*)}{L(0)} =$$

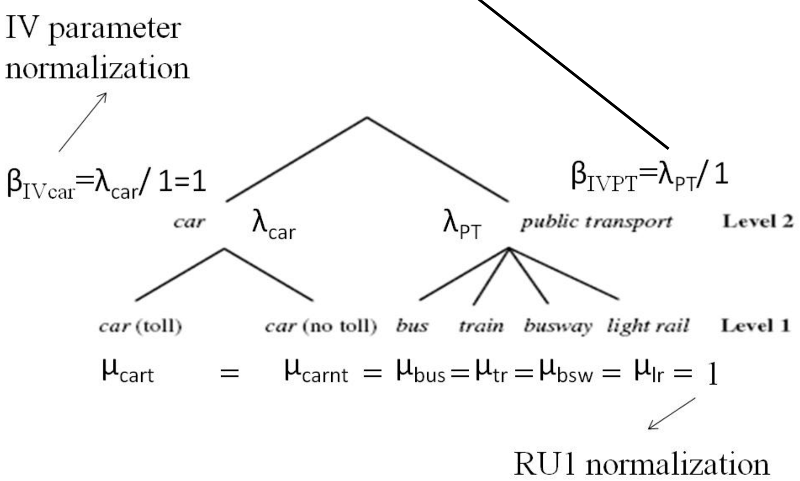
$$1 - \left(\frac{-4768.225}{-6324.968}\right) = 0.2437$$

- Then the output shows the results for the Nested structure that we want to fit.
- The pseudo R-Squared with no information base model is presented immediately.
- In Nested Logits we will use only this, because a model with just the alternative specific constants in nests is not the same as in an MNL.

First NL model (VII)

IV(Car) normalization
IV(PT) = 0.01986

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
---------+Attributes in the Utility Functions (beta)				
ASCCART	-.42688862	.05746897	-7.428	.0000
FUEL	-.16106782	.01480054	-10.883	.0000
ASCCARNT	.06085689	.08508272	.715	.4744
ASCBUS	.18651365	.09875140	1.889	.0589
FARE	-.30785954	.02220022	-13.867	.0000
ASCTN	.10895001	.08674916	1.256	.2091
ASCBUSW	.13549235	.07489564	1.809	.0704
---------+IV parameters, lambda(b|l),gamma(l)				
CAR	1.00000000	......(Fixed Parameter)......		
PT	.01986084	.08874732	.224	.8229
---------+Underlying standard deviation = pi/(IVparm*sqr(6))				
CAR	1.28254980	......(Fixed Parameter)......		
PT	64.5768238	288.558847	.224	.8229



- We can't reject the hypothesis that the IV_{PT} is zero.
- There is too much correlation between the alternatives in the PT branch. Huge common variance.

First NL model (IX)

```
--> WALD;FN1:PT-0$

+------------------------------------------------+
| WALD procedure. Estimates and standard errors  |
| for nonlinear functions and joint test of      |
| nonlinear restrictions.                        |
| Wald Statistic               =       .05008    |
| Prob. from Chi-squared[ 1]  =       .82292    |
+------------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |    .01986084      .08874732       .224    .8229
--> WALD;FN1:PT-1$

+------------------------------------------------+
| WALD procedure. Estimates and standard errors  |
| for nonlinear functions and joint test of      |
| nonlinear restrictions.                        |
| Wald Statistic               =    121.97335    |
| Prob. from Chi-squared[ 1]  =       .00000    |
+------------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |   -.98013916      .08874732    -11.044    .0000
```

We reject with great certainty the hypothesis that the coefficient may be 1, an MNL is definitely not advisable! Hence we should search for a new structure of the NL.

- You may proceed with several Wald tests to the data.
- The first test is exactly the same as the output test from the previous slide, we are testing the hypothesis of the coefficient being zero.
- The second one is about testing the hypothesis that the IV_{PT} is 1, which would mean equal scales, (equal variances) between both levels thus pointing to an MNL.

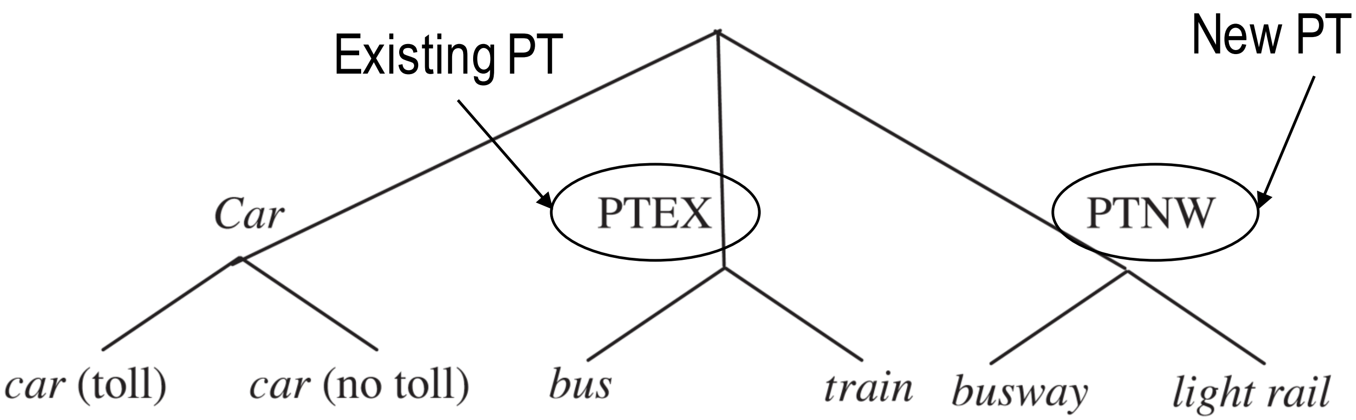
INSTITUTO SUPERIOR TÉCNICO

FEUP

Second Nested Model (I)

- We are not yet satisfied with the Nested Model we have just tested.

- A second alternative model can be having three branches. We may study the following structure:



Second Nested Model (II)

```
NLOGIT
;lhs = choice, cset, altij
;choices = cart,  carnt, bus, train, busway, LR
;tree = car(cart, carnt),PTEX(bus,train),PTNW (busway, LR)
;RU1
;start  = logit
;ivset: (car)=[1.0]
;maxit = 100
;Prob=MARGPROB
;cprob=ALTPROB
;ivb=IVBRANCH
;Utility=U1
;model:
U(cart) = asccart + fuel*fuel /
U(carnt) = asccarnt + fuel*fuel /
U(bus) =  ascbus + fare*fare /
U(train) = asctn + fare*fare /
U(busway) = ascbusw + fare*fare /
U(LR) = fare*fare
;Crosstab$
```

Second Nested Model (III)

```
+----------------------------------------------+
| Discrete choice (multinomial logit) model    |
| Maximum Likelihood Estimates                 |
| Model estimated: Nov 03, 2010 at 01:41:02PM. |
| Dependent variable               Choice      |
| Weighting variable                 None      |
| Number of observations             3587      |
| Iterations completed                  5      |
| Log likelihood function       -4804.548      |
| Number of parameters                  7      |
| Info. Criterion: AIC =          2.68277      |
|   Finite Sample: AIC =          2.68278      |
| Info. Criterion: BIC =          2.69484      |
| Info. Criterion:HQIC =          2.68707      |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj  |
| Constants only.  Must be computed directly.  |
|                  Use NLOGIT ;...; RHS=ONE $  |
| Chi-squared[ 2]          =   2981.12332      |
| Prob [ chi squared > value ] =   .00000      |
| Response data are given as ind. choice.      |
| Number of obs.=  3587, skipped   0 bad obs.  |
+----------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
ASCCART	-.22658772	.08285265	-2.735	.0062
FUEL	-.20241829	.01429703	-14.158	.0000
ASCCARNT	.37994204	.09911161	3.833	.0001
ASCBUS	.03145882	.07732794	.407	.6841
FARE	-.19347334	.01625132	-11.905	.0000
ASCTN	-.409564D-05	.07231932	.000	1.0000
ASCBUSW	.11125885	.06571070	1.693	.0904

- The initial MNL model is the same
- Notice that we haven't been worrying much about the significance of the variables, looking mainly at the model structure.
- Coefficient of variable ASCTN is irrelevant for now.

INSTITUTO SUPERIOR TÉCNICO

FEUP

Second Nested Model (IV)

```
+----------------------------------------------+
| FIML Nested Multinomial Logit Model          |
| Maximum Likelihood Estimates                 |
| Model estimated: Nov 03, 2010 at 01:46:05PM. |
| Dependent variable              CHOICE       |
| Weighting variable                None       |
| Number of observations            3587       |
| Iterations completed                17       |
| Log likelihood function      -4798.884       |
| Number of parameters                 9       |
| Info. Criterion: AIC =         2.68073       |
|   Finite Sample: AIC =         2.68074       |
| Info. Criterion: BIC =         2.69625       |
| Info. Criterion:HQIC =         2.68626       |
| Restricted log likelihood    -6427.041       |
| McFadden Pseudo R-squared     .2533292       |
| Chi squared                   3256.314       |
| Degrees of freedom                   9       |
| Prob[ChiSqd > value] =        .0000000       |
| Constants only.  Must be computed directly.  |
|                  Use NLOGIT ;...; RHS=ONE $  |
| At start values  -4804.5480  .00118 *******  |
| Response data are given as ind. choice.      |
+----------------------------------------------+
```

- Model has improved against a model with equal shares (no information)

```
+---------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+---------+--------------+----------------+--------+--------+
---------+Attributes in the Utility Functions (beta)
 ASCCART |  -.34016067       .08514441     -3.995   .0001
 FUEL    |  -.19924460       .01429222    -13.941   .0000
 ASCCARNT|   .25725625       .10265187      2.506   .0122
 ASCBUS  |  -.13734992       .16336122      -.841   .4005
 FARE    |  -.24786595       .02675499     -9.264   .0000
 ASCTN   |  -.18033197       .16301087     -1.106   .2686
 ASCBUSW |   .12774225       .07128270      1.792   .0731
---------+IV parameters, lambda(b|l),gamma(l)
 CAR     |  1.00000000      ......(Fixed Parameter).......
 PTEX    |   .62271038       .10163515      6.127   .0000
 PTNW    |   .79117650       .10190404      7.764   .0000
---------+Underlying standard deviation = pi/(IVparm*sqr(6))
 CAR     |  1.28254980      ......(Fixed Parameter).......
 PTEX    |  2.05962490       .33615993      6.127   .0000
 PTNW    |  1.62106659       .20879441      7.764   .0000
```


INSTITUTO SUPERIOR TÉCNICO

FEUP

Second Nested Model (V)

```
--> WALD;FN1:ptex-0$

+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic            =      37.53915     |
| Prob. from Chi-squared[ 1] =       .00000     |
+-----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |      .62271038       .10163515     6.127   .0000
--> WALD;FN1:ptex-1$

+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic            =      13.78040     |
| Prob. from Chi-squared[ 1] =       .00021     |
+-----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |     -.37728962       .10163515    -3.712   .0002
--> WALD;FN1:ptnw-0$

+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic            =      60.27871     |
| Prob. from Chi-squared[ 1] =       .00000     |
+-----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |      .79117650       .10190404     7.764   .0000
--> WALD;FN1:ptnw-1$

+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic            =       4.19929     |
| Prob. from Chi-squared[ 1] =       .04044     |
+-----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |     -.20882350       .10190404    -2.049   .0404
```

- The IV parameters are in the expected interval, neither 0 nor 1, meaning that the branches that we defined are significant for the data we are analyzing and for the Utility functions which we have proposed.

INSTITUTO SUPERIOR TÉCNICO

FEUP

Second Nested Model (VI)

- NL Three branches

	CART	CARNT	BUS	TRAIN	BUSWAY	LR	Total
CART	182	195	97	91	122	111	798
CARNT	187	214	96	105	119	118	838
BUS	94	97	126	49	62	0	428
TRAIN	95	97	51	126	0	58	428
BUSWAY	125	117	58	0	184	84	569
LR	115	118	0	58	82	154	526
Total	798	838	428	428	570	525	3587

$$\%\ Correct\ Predictions = \frac{986}{3587} = 27.48\%$$

- MNL model

	CART	CARNT	BUS	TRAIN	BUSWAY	LR	Total
CART	183	195	97	91	122	111	798
CARNT	186	213	97	105	119	118	838
BUS	92	96	125	54	61	0	428
TRAIN	95	97	55	125	0	57	428
BUSWAY	126	118	54	0	183	88	569
LR	116	119	0	54	85	153	526
Total	798	838	428	428	569	526	3587

$$\%\ Correct\ Predictions = \frac{982}{3587} = 27.37\%$$

No great difference!

INSTITUTO SUPERIOR TÉCNICO

FEUP

Specifying utility functions at higher levels of the NL tree (I)

- Up to now we have only specified the utility functions at level 1, the level of the alternatives. But what if there are variables which better explain the choice between the branches (level 2) and not the conditional probabilities (probability in each nest)?

- The nested Logit model allows to specify these utility functions. Let’s consider the same Nested Logit structure of the current example, but let’s now include as explanatory variables on the option to use Car the number of licensed drivers at the home of the respondent and the number of vehicles available. Intuitively these should motivate the choice for driving in either tolled or non tolled roads.

Specifying utility functions at higher levels of the NL tree (II)

```
NLOGIT
;lhs = choice, cset, altij
;choices = cart, carnt, bus, train, busway, LR
;tree = car(cart, carnt),PTEX(bus,train),PTNW (busway,LR)
;RU1
;start = logit
;ivset: (car)=[1.0]
;maxit = 100
;Prob=MARGPROB
;cprob=ALTPROB
;ivb=IVBRANCH
;Utility=U1
;model:
U(Car)=ndrivlic*ndrivlic+numbvehs*numbvehs/
U(ptex)=asptex/
U(cart) = asccart + fuel*fuel /
U(carnt) = asccarnt + fuel*fuel /
U(bus) =  ascbus + fare*fare /
U(train) = asctn + fare*fare /
U(busway) = ascbusw + fare*fare /
U(LR) = fare*fare
;Crosstab$
```

- This is off course something that only a Nested Logit can do.

Specifying utility functions at higher levels of the NL tree (III)

```
+---------------------------------------------+
| FIML Nested Multinomial Logit Model         |
| Maximum Likelihood Estimates                |
| Model estimated: Nov 03, 2010 at 11:02:29PM.|
| Dependent variable               CHOICE     |
| Weighting variable                 None     |
| Number of observations             3587     |
| Iterations completed                 24     |
| Log likelihood function       -4724.031     |
| Number of parameters                 12     |
| Info. Criterion: AIC =          2.64066     |
|   Finite Sample: AIC =          2.64069     |
| Info. Criterion: BIC =          2.66136     |
| Info. Criterion:HQIC =          2.64804     |
| Restricted log likelihood     -6427.041     |
| McFadden Pseudo R-squared      .2649757     |
| Chi squared                    3406.020     |
| Degrees of freedom                   12     |
| Prob[ChiSqd > value] =         .0000000     |
| Constants only.  Must be computed directly. |
|                  Use NLOGIT ;...; RHS=ONE $ |
| At start values  -4754.1916  .00634 ******* |
| Response data are given as ind. choice.     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+
|Variable| Coefficient   | Standard Error |b/St.Er.|P[|Z|>z]|
+---------+--------------+----------------+--------+---------+
---------+Attributes in the Utility Functions (beta)
 ASCCART |   -.87023131       .11332931     -7.679    .0000
 FUEL    |   -.18719390       .01446976    -12.937    .0000
 ASCCARNT|   -.30913690       .12785424     -2.418    .0156
 ASCBUS  |   -.46747995       .29695367     -1.574    .1154
 FARE    |   -.25578431       .02666065     -9.594    .0000
 ASCTN   |   -.43944077       .27711101     -1.586    .1128
 ASCBUSW |    .14775049       .07271463      2.032    .0422
---------+Attributes of Branch Choice Equations (alpha)
 NDRIVLIC|   -.24644270       .05470899     -4.505    .0000
 NUMBVEHS|    .55819644       .05285971     10.560    .0000
 ASPTEX  |    .35870254       .23034299      1.557    .1194
---------+IV parameters, lambda(b|l),gamma(l)
 CAR     |   1.00000000     ......(Fixed Parameter).......
 PTEX    |    .57213474       .10367102      5.519    .0000
 PTNW    |    .78683045       .10086590      7.801    .0000
---------+Underlying standard deviation = pi/(IVparm*sqr(6))
 CAR     |   1.28254980     ......(Fixed Parameter).......
 PTEX    |   2.24169189       .40619537      5.519    .0000
 PTNW    |   1.63002055       .20895669      7.801    .0000
```

Pseudo-R^2 has increased (using a null base model/equal shares model as reference)

As there are more members of the household with a driver's license there is more competition for using the vehicle

The presence of more vehicles increases the probability of using the automobile

Specifying utility functions at higher levels of the NL tree (IV)

```
--> WALD;FN1:ptex-1$

+----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic            =     17.03332      |
| Prob. from Chi-squared[ 1] =      .00004      |
+----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |    -.42786526       .10367102    -4.127   .0000

--> WALD;FN1:ptnw-1$

+----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic            =      4.46644      |
| Prob. from Chi-squared[ 1] =      .03457      |
+----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 Fncn(1) |    -.21316955       .10086590    -2.113   .0346
```

- The two IV parameters are statistically different from 1 and from 0 (previous table)

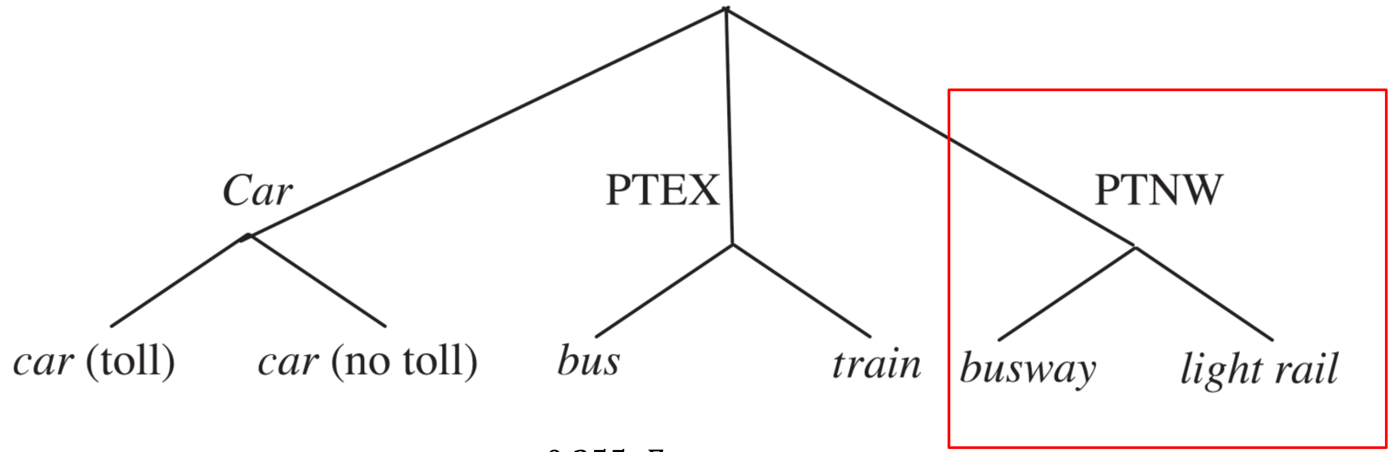
Specifying utility functions at higher levels of the NL tree (V)

	CART	CARNT	BUS	TRAIN	BUSWAY	LR	Total
CART	193	204	94	87	115	105	798
CARNT	196	222	91	98	117	114	838
BUS	87	91	132	55	64	0	428
TRAIN	91	94	55	129	0	60	428
BUSWAY	123	114	59	0	185	88	569
LR	109	113	0	57	88	159	526
Total	798	838	431	425	569	526	3587

$$\%\ Correct\ Predictions = \frac{1022}{3587} = 28.49\%$$

- This is not to say that the model will predict everything wrong: the shares as you remember are estimated through expectancy, aggregating probabilities across individuals so every probability will contribute.
- Modeling has as much of science as it has of art. It is difficult to say you have reached the best model. This model still does not have many explanatory variables.

Computing probabilities (I)



Conditional probabilities:

$$P(LR|PTNW) = \frac{e^{-0.255*Fare}}{e^{-0.255*Fare} + e^{0.14775-0.255*Fare}}$$

Branch Probabilities:

$$P(PTNW) = \frac{e^{0.786830*V_{PTNW}}}{e^{1*V_{Car}} + e^{0.572134*V_{PTEX}} + e^{0.786830*V_{PTNW}}}$$

$$= \frac{e^{0.786830*ln(e^{-0.255*Fare}+e^{0.14775-0.255*Fare})}}{e^{1*(-0.24644*ndrivlic+0.55819*numbvehs+IV_{Car})} + e^{0.572134*(0.35870+IV_{PTEX})} + e^{0.786830*(IV_{PTNW})}}$$

Same thing

Marginal probabilities:

$$P(LR) = P(LR/PTNW) * P(PTNW)$$

Computing probabilities (II)

- The marginal probability:

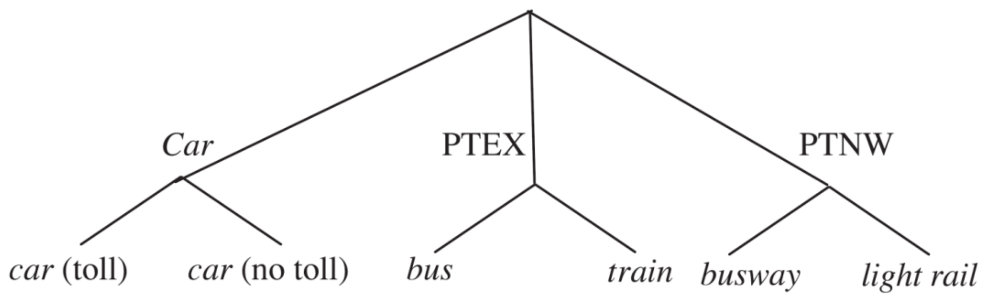
$$P(Bus) = P(Bus|PT) * P(PT)$$

marginal probability

Probability of the alternative inside each branch

Utility of each alternative

ALTIJ		IVBRANCH	MARGPROB	ALTPROB	U1
cart	1	-0.453722	0.26958	0.546783	-1.05743
carnt	2	-0.453722	0.223449	0.453217	-1.24511
train	4	-0.695225	0.279505	1	-0.695225
LR	6	-0.767353	0.227467	1	-0.767353
cart	1	-0.347147	0.224469	0.407599	-1.24462
carnt	2	-0.347147	0.326242	0.592401	-0.870719
busway	5	0.0925465	0.132152	0.294135	-1.13117
LR	6	0.0925465	0.317138	0.705865	-0.255784

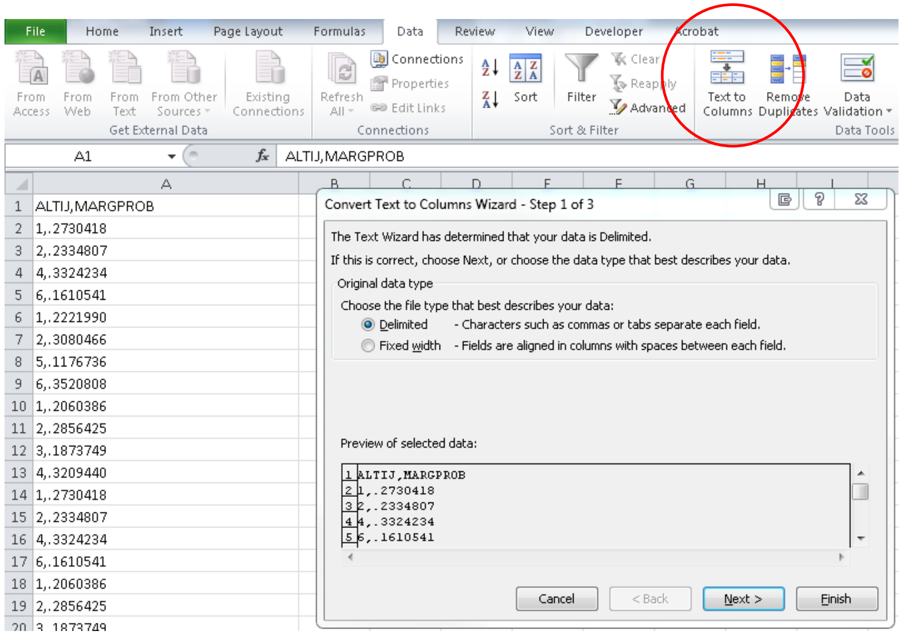


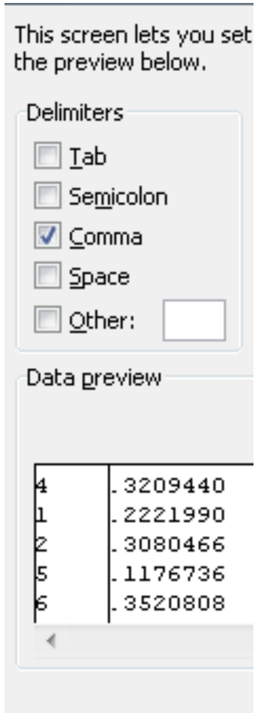
- Remember that in each choice the respondent had 4 alternatives, the first two were cart and carnt, the two other were Public Transport alternatives picked from 4 possible.

Aggregating across alternatives (I)

- Regarding aggregation be careful because you can´t just copy the Probability attribute to excel, it will only bring 1900 lines. You have to export the variables:

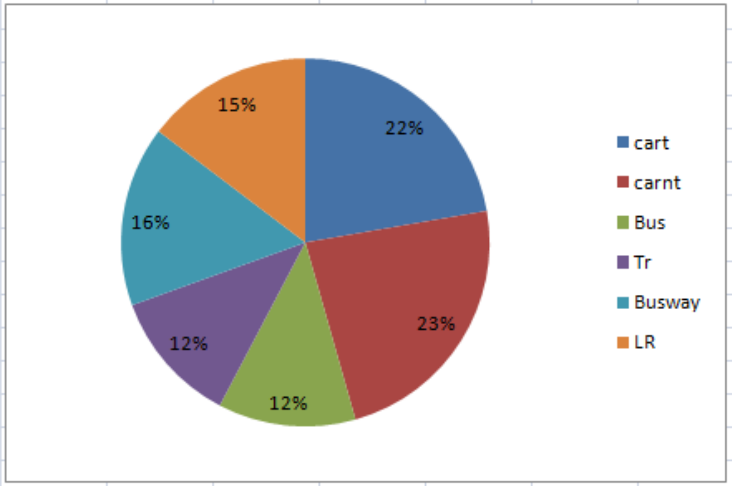
- Go to project -> Export -> Variables then Choose Excel Worksheet, give the name for your file and choose the variables you want to export: ALTIJ and MARGPROB.





Aggregating across alternatives (II)

			cart	carnt	Bus	Tr	Busway	LR		
ALTIJ	MARGPROB		1	2	3	4	5	6		
1	0.2695796		22%	23%	12%	12%	16%	15%	Sum	100%
2	0.2234488								Sum	3587
4	0.2795047									
6	0.2274669									
1	0.224469									
2	0.3262415									
5	0.1321516									
6	0.3171379									
1	0.2352972									
2	0.3419793									
3	0.1556801									
4	0.2670434									
1	0.2695796									
2	0.2234488									
4	0.2795047									
6	0.2274669									
1	0.2352972									
2	0.3419793									
3	0.1556801									
4	0.2670434									
1	0.224469									
2	0.3262415									
5	0.1321516									



- Be aware that we are aggregating across alternatives which have been produced synthetically, thus outputting indicative shares which you should be careful on using! An advanced topic on avoiding these issues is combining RP and SP data, but we will not see that on this course.

Bibliography

- Ben-Akiva, M. and Lerman, Steven R. (1985) "Discrete Choice Analysis: Theory and Applications to Travel Demand", MIT Press.
- Hensher, Rose and Greene (2005) "Applied Choice Analysis: A Primer" Cambridge.
- Ortúzar J. and Willumsen L. (2001) Modelling Transport. 3rd Edition. John Wiley and Sons. West Sussex, England.

Case study data is free to be accessed from: http://www.cambridge.org/0521605776