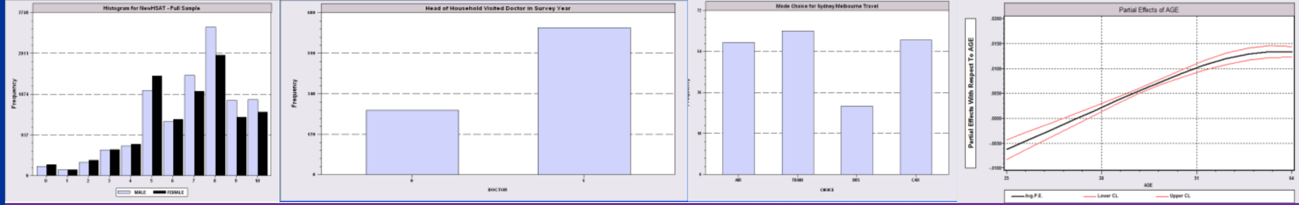
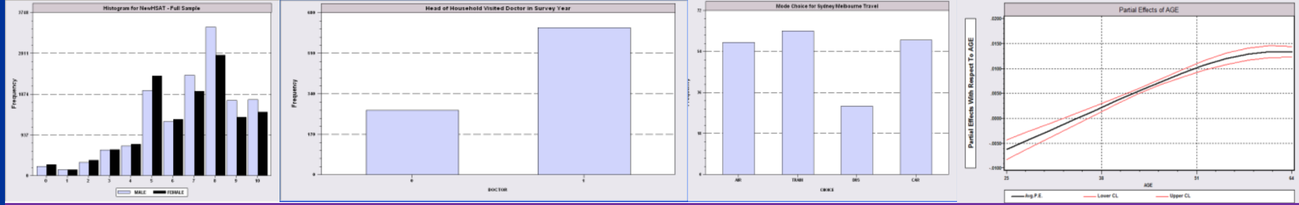


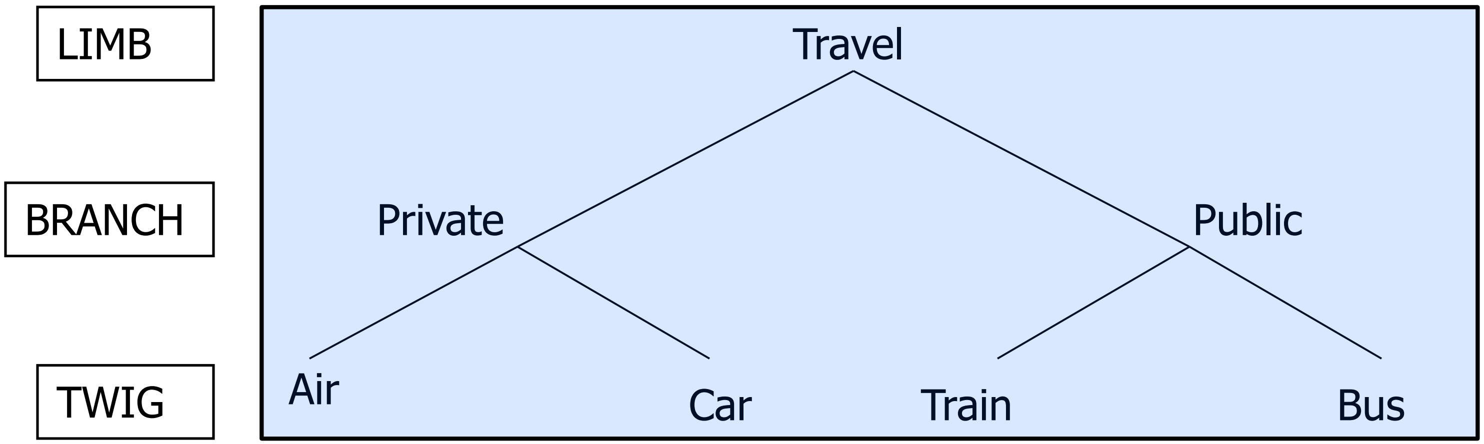
Discrete Choice Modeling

William Greene
Stern School of Business
New York University



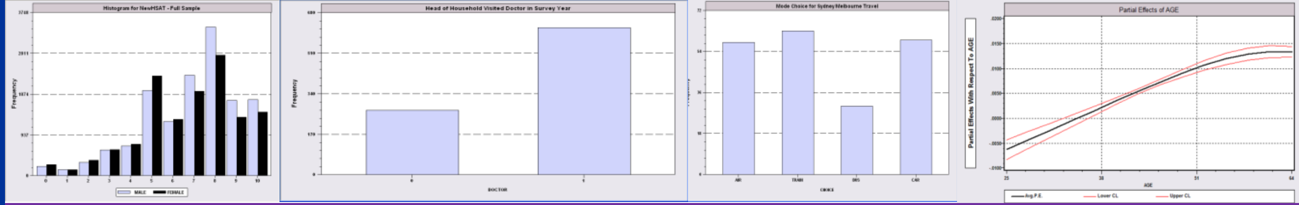
Extended Formulation of the MNL

Sets of similar alternatives



Compound Utility: U(Alt)=U(Alt|Branch)+U(branch)

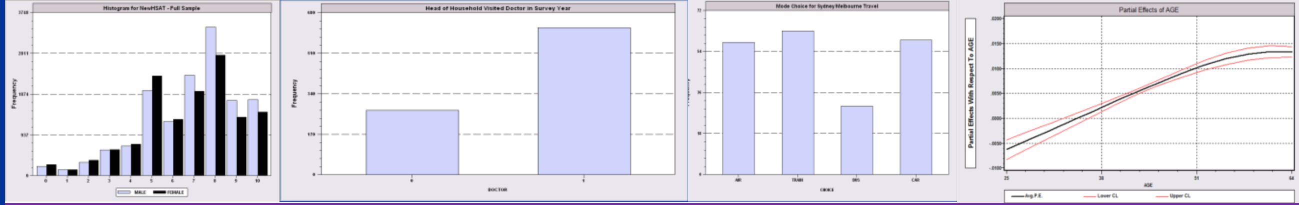
Behavioral implications – Correlations within branches



Correlation Structure for a Two Level Model

- Within a branch
  - Identical variances (IIA (MNL) applies)
  - Covariance (all same) = variance at higher level
- Branches have different variances (scale factors)
- Nested logit probabilities: Generalized Extreme Value

Prob[Alt,Branch] = Prob(branch) * Prob(Alt|Branch)



Probabilities for a Nested Logit Model

Utility functions; (Drop observation indicator, i.)

Twig level: $k \mid j$ denotes alternative k in branch j

$$U(k \mid j) = \alpha_{k|j} + \boldsymbol{\beta}'\mathbf{x}_{k|j}$$

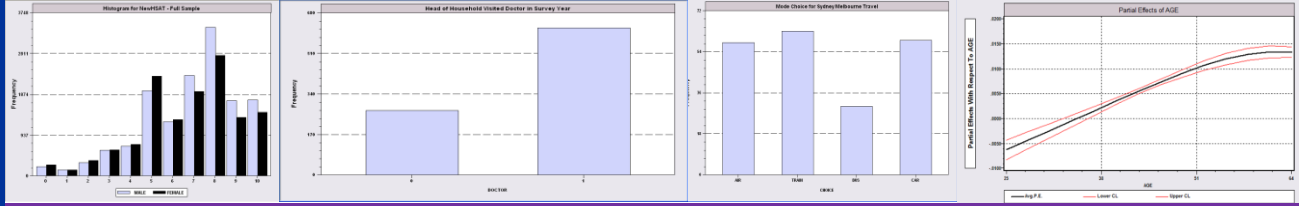
Branch level $U(j) = \boldsymbol{\gamma}'\mathbf{y}_j$

Twig level probability: $P(k \mid j) = P_{k|j} = \boxed{\dfrac{\exp(\alpha_{k|j} + \boldsymbol{\beta}'\mathbf{x}_{k|j})}{\sum_{m=1}^{K|j} \exp(\alpha_{m|j} + \boldsymbol{\beta}'\mathbf{x}_{m|j})}}$

Inclusive value for branch $j = IV(j) = \log\left[\Sigma_{m=1}^{K|j} \exp(\alpha_{m|j} + \boldsymbol{\beta}'\mathbf{x}_{m|j})\right]$

Branch level probability: $\boxed{P(j) = \dfrac{\exp\left[\lambda_j\left(\boldsymbol{\gamma}'\mathbf{y}_j + IV(j)\right)\right]}{\sum_{b=1}^{B} \exp\left[\lambda_b\left(\boldsymbol{\gamma}'\mathbf{y}_b + IV(b)\right)\right]}}$

$\lambda_j = 1$ for all branches returns the original MNL model



Model Form RU1

Twig Level Probability

$$\text{Prob}(\text{Choice}=k\,|\,j)=\frac{\exp(\boldsymbol{\beta}'\mathbf{x}_{k|j})}{\sum_{m=1}^{K|j}\exp(\boldsymbol{\beta}'\mathbf{x}_{m|j})}$$

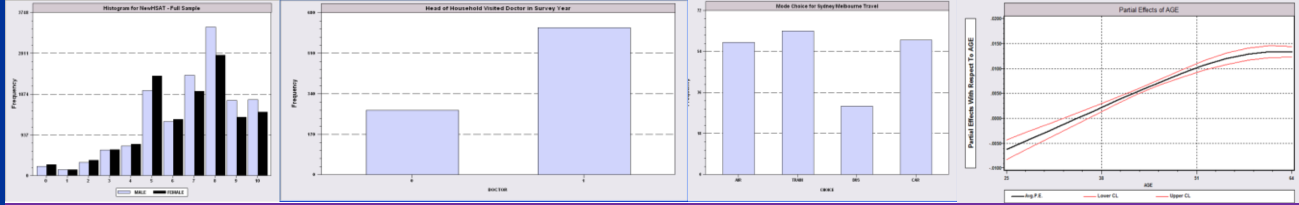
Inclusive Value for the Branch

$$\text{IV}(j)=\log\left[\sum_{m=1}^{K|j}\exp(\boldsymbol{\beta}'\mathbf{x}_{m|j})\right]$$

Branch Probability

$$\text{Prob}(\text{Branch}=j)=\frac{\exp\left[\lambda_j\left(\boldsymbol{\gamma}'\mathbf{y}_j+\text{IV}(j)\right)\right]}{\sum_{b=1}^{B}\exp\left[\lambda_b\left(\boldsymbol{\gamma}'\mathbf{y}_b+\text{IV}(b)\right)\right]}$$

$\lambda_j = 1$ Returns the Multinomial Logit Model



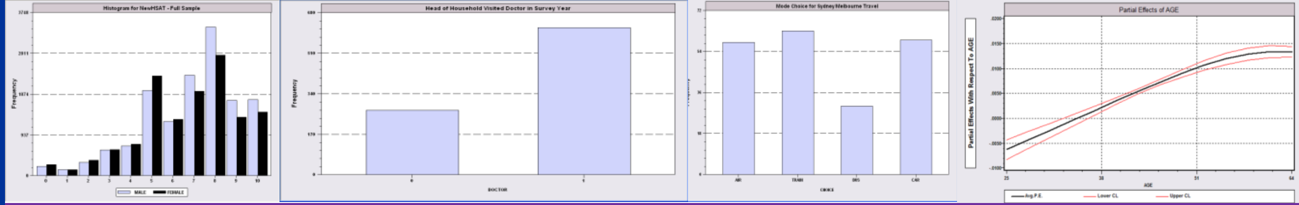
Moving Scaling Down to the Twig Level

RU2 Normalization

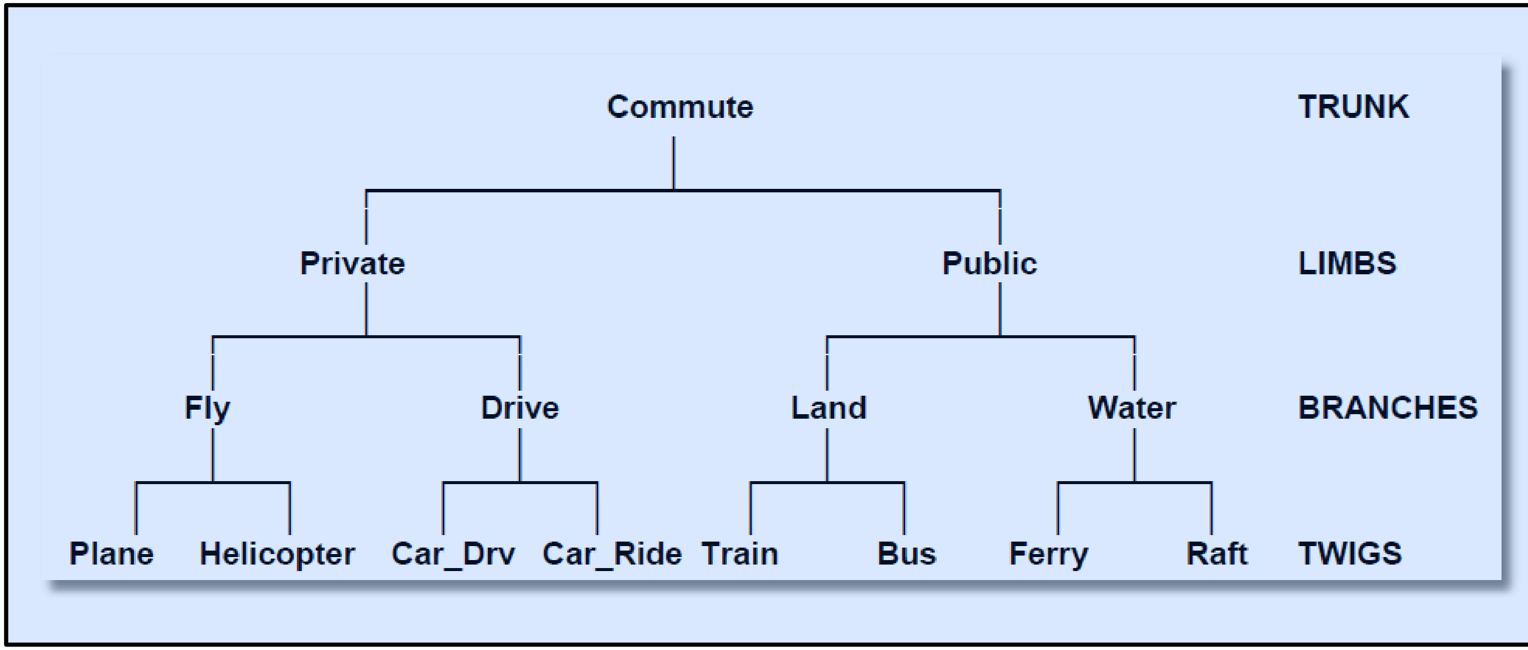
Twig Level Probability : $$P_{k|j} = \frac{\exp\left[\dfrac{\boldsymbol{\beta}'\mathbf{x}_{k|j}}{\mu_j}\right]}{\sum_{m=1}^{k|j}\exp\left[\dfrac{\boldsymbol{\beta}'\mathbf{x}_{m|j}}{\mu_j}\right]}$$

Inclusive Value for the Branch : $$IV(j) = \log\left(\sum_{m=1}^{k|j}\exp\left[\dfrac{\boldsymbol{\beta}'\mathbf{x}_{m|j}}{\mu_j}\right]\right)$$

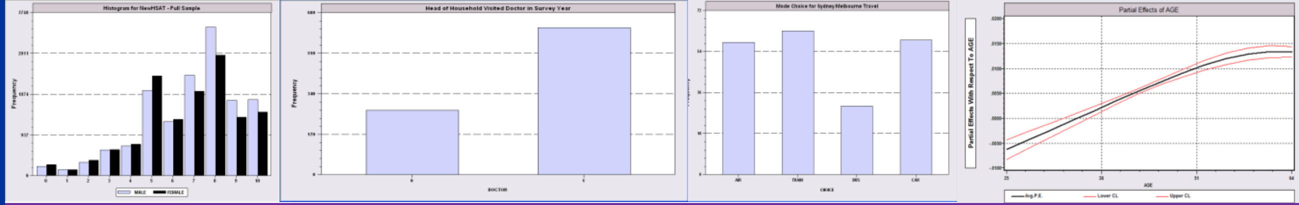
Branch Probability : $$P_j = \frac{\exp\left[\boldsymbol{\gamma}'\mathbf{y}_j + \mu_j IV(j)\right]}{\sum_{b=1}^{B}\exp\left[\gamma' y_b + \mu_b IV(b)\right]}$$



Higher Level Trees



E.g., Location (Neighborhood)
Housing Type (Rent, Buy, House, Apt)
Housing (# Bedrooms)

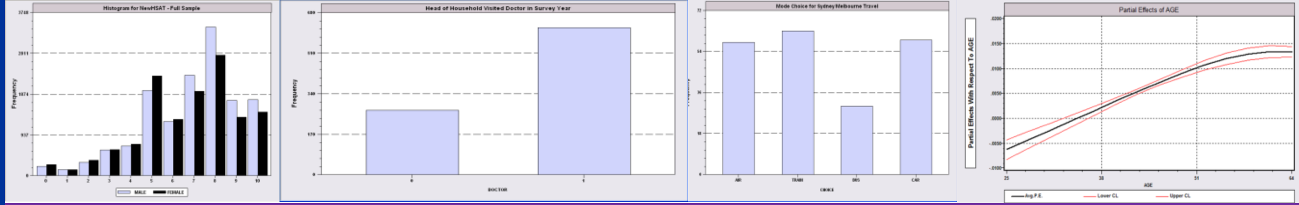


Estimation Strategy for Nested Logit Models

- **Two step estimation (ca. 1980s)**
  - For each branch, just fit MNL
    - Loses efficiency – replicates coefficients
  - For branch level, fit separate model, just including **y** and the inclusive values in the branch level utility function
    - Again loses efficiency

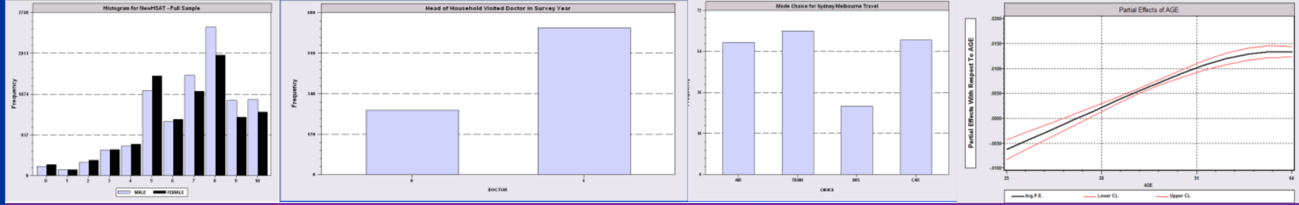
- **Full information ML (current)**

  Fit the entire model at once, imposing all restrictions



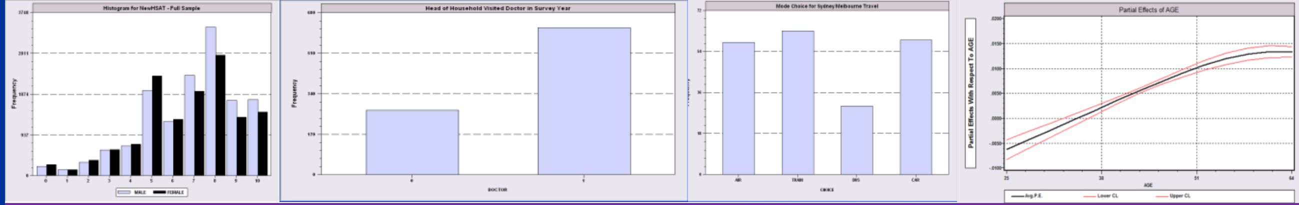
MNL Baseline

```
-----------------------------------------------------------
Discrete choice (multinomial logit) model
Dependent variable                Choice
Log likelihood function       -172.94366
Estimation based on N =    210, K =  10
R2=1-LogL/LogL* Log-L fncn R-sqrd R2Adj
Constants only   -283.7588  .3905 .3787
Chi-squared[ 7]          =     221.63022
Prob [ chi squared > value ] =    .00000
Response data are given as ind. choices
Number of obs.=   210, skipped    0 obs
--------+--------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]
--------+--------------------------------------------------
      GC|    .07578***        .01833        4.134   .0000
    TTME|   -.10289***        .01109       -9.280   .0000
    INVT|   -.01399***        .00267       -5.240   .0000
    INVC|   -.08044***        .01995       -4.032   .0001
   A_AIR|   4.37035***       1.05734        4.133   .0000
AIR_HIN1|    .00428           .01306         .327   .7434
 A_TRAIN|   5.91407***        .68993        8.572   .0000
TRA_HIN3|   -.05907***        .01471       -4.016   .0001
   A_BUS|   4.46269***        .72333        6.170   .0000
BUS_HIN4|   -.02295           .01592       -1.442   .1493
--------+--------------------------------------------------
```



FIML Parameter Estimates

```
-----------------------------------------------------------
FIML Nested Multinomial Logit Model
Dependent variable                        MODE
Log likelihood function             -166.64835
The model has 2 levels.
Random Utility Form 1:IVparms = LMDAb|l
Number of obs.=    210, skipped     0 obs
--------+--------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]
--------+--------------------------------------------------
        |Attributes in the Utility Functions (beta)
      GC|      .06579***       .01878        3.504   .0005
    TTME|    -.07738***        .01217       -6.358   .0000
    INVT|    -.01335***        .00270       -4.948   .0000
    INVC|    -.07046***        .02052       -3.433   .0006
   A_AIR|    2.49364**        1.01084        2.467   .0136
AIR_HIN1|      .00357          .01057         .337   .7358
 A_TRAIN|    3.49867***        .80634        4.339   .0000
TRA_HIN3|    -.03581***        .01379       -2.597   .0094
   A_BUS|    2.30142***        .81284        2.831   .0046
BUS_HIN4|    -.01128           .01459        -.773   .4395
        |IV parameters, lambda(b|l),gamma(l)
 PRIVATE|    2.16095***        .47193        4.579   .0000
  PUBLIC|    1.56295***        .34500        4.530   .0000
--------+--------------------------------------------------
```



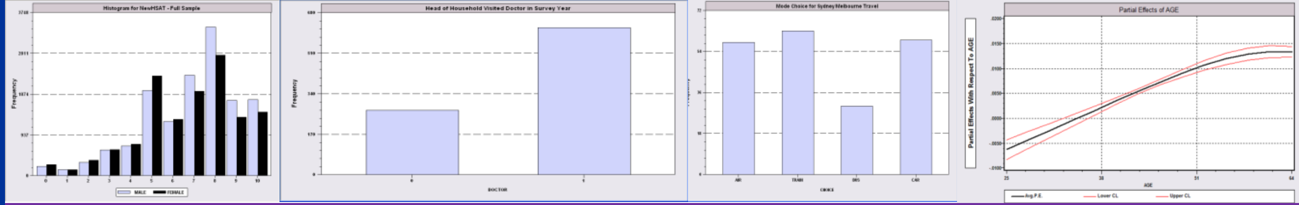
Elasticities Decompose Additively

$$\frac{\partial \log P(alt=j, limb=l, branch=b, trunk=r)}{\partial x(k) \mid alt=J, limb=L, branch=B, trunk=R} = D(k \mid J,B,L,R) = \Delta(k) \times F \;,$$

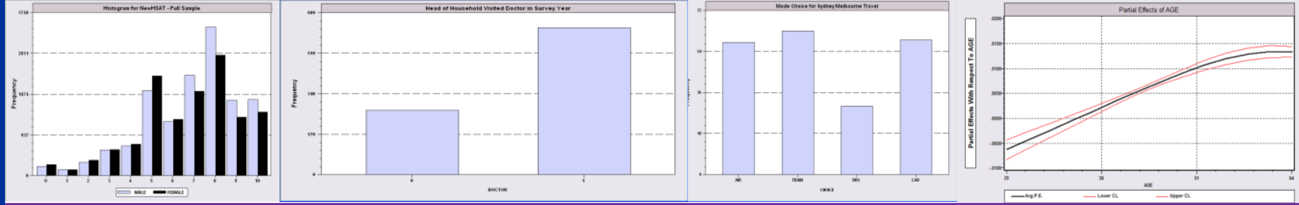
where $\quad \Delta(k)$ = coefficient on $x(k)$ in $U(J|B,L,R)$

and

$$
\begin{aligned}
F = \;& \mathbf{1}(r{=}R) \times \mathbf{1}(l{=}L) \times \mathbf{1}(b{=}B) \times [\mathbf{1}(j{=}J) - P(J|BLR)] && \text{(trunk effect)},\\
& \mathbf{1}(r{=}R) \times \mathbf{1}(l{=}L) \times [\mathbf{1}(b{=}B) - P(B|LR)] \times P(J|BLR) \times \tau_{B|LR} && \text{(limb effect)},\\
& \mathbf{1}(r{=}R) \times [\mathbf{1}(l{=}L) - P(L|R)] \times P(B|LR) \times P(J|BLR) \times \tau_{B|LR} \times \sigma_{L|R} && \text{(branch effect)},\\
& [\mathbf{1}(r{=}R) - P(R)] \times P(L|R) \times P(B|LR) \times P(J|BLR) \times \tau_{B|LR} \times \sigma_{L|R} \times \phi_R && \text{(twig effect)}.
\end{aligned}
$$

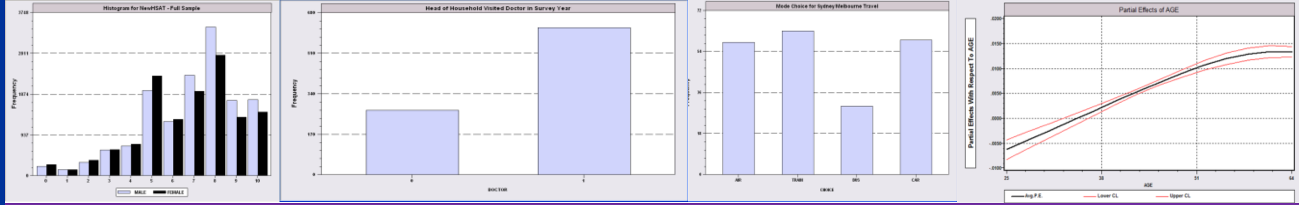



```
+--------------------------------------------------------------------------+
| Elasticity             averaged over observations.                       |
| Attribute is INVC     in choice AIR                                      |
|                        Decomposition of Effect if Nest    Total Effect|
|                        Trunk    Limb    Branch    Choice      Mean   St.Dev|
|    Branch=PRIVATE                                                        |
| *       Choice=AIR       .000     .000   -2.456   -3.091     -5.547   3.525 |
|         Choice=CAR       .000     .000   -2.456    2.916       .460   3.178 |
|    Branch=PUBLIC                                                         |
|         Choice=TRAIN     .000     .000    3.846     .000      3.846   4.865 |
|         Choice=BUS       .000     .000    3.846     .000      3.846   4.865 |
+--------------------------------------------------------------------------+
| Attribute is INVC     in choice CAR                                      |
|    Branch=PRIVATE                                                        |
|         Choice=AIR       .000     .000    -.757     .650      -.107    .589 |
| *       Choice=CAR       .000     .000    -.757    -.830     -1.587   1.292 |
|    Branch=PUBLIC                                                         |
|         Choice=TRAIN     .000     .000     .647     .000       .647    .605 |
|         Choice=BUS       .000     .000     .647     .000       .647    .605 |
+--------------------------------------------------------------------------+
| Attribute is INVC     in choice TRAIN                                    |
|    Branch=PRIVATE                                                        |
|         Choice=AIR       .000     .000    1.340     .000      1.340   1.475 |
|         Choice=CAR       .000     .000    1.340     .000      1.340   1.475 |
|    Branch=PUBLIC                                                         |
| *       Choice=TRAIN     .000     .000   -1.986   -1.490     -3.475   2.539 |
|         Choice=BUS       .000     .000   -1.986    2.128       .142   1.321 |
+--------------------------------------------------------------------------+
| * indicates direct Elasticity effect of the attribute.                   |
+--------------------------------------------------------------------------+
```

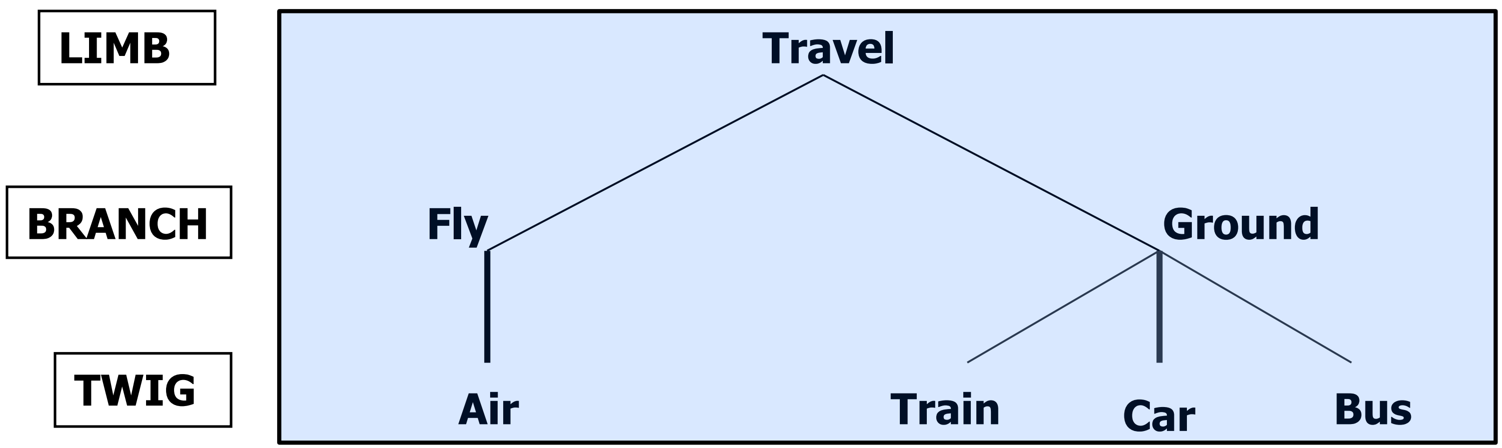


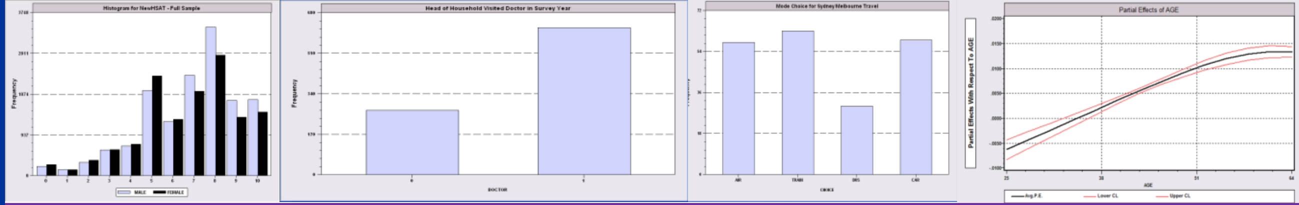
Testing vs. the MNL

- Log likelihood for the NL model
- Constrain IV parameters to equal 1 with
  ; IVSET(list of branches)=[1]
- Use likelihood ratio test
- For the example:
  - LogL (NL) = -166.68435
  - LogL (MNL) = -172.94366
  - Chi-squared with 2 d.f. = 2(-166.68435-(-172.94366))
    = 12.51862
  - The critical value is 5.99 (95%)
  - The MNL (and a fortiori, IIA) is rejected



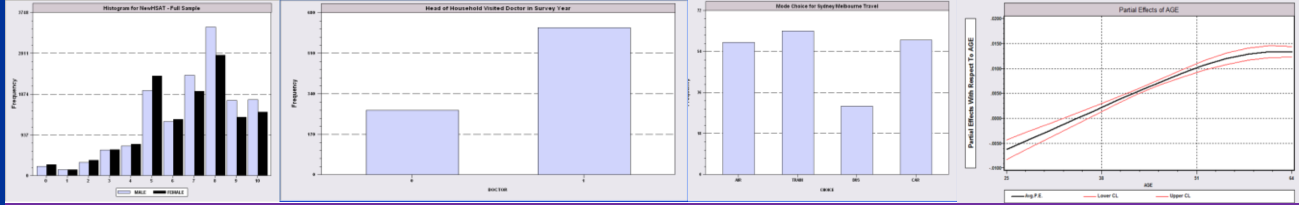
Degenerate Branches



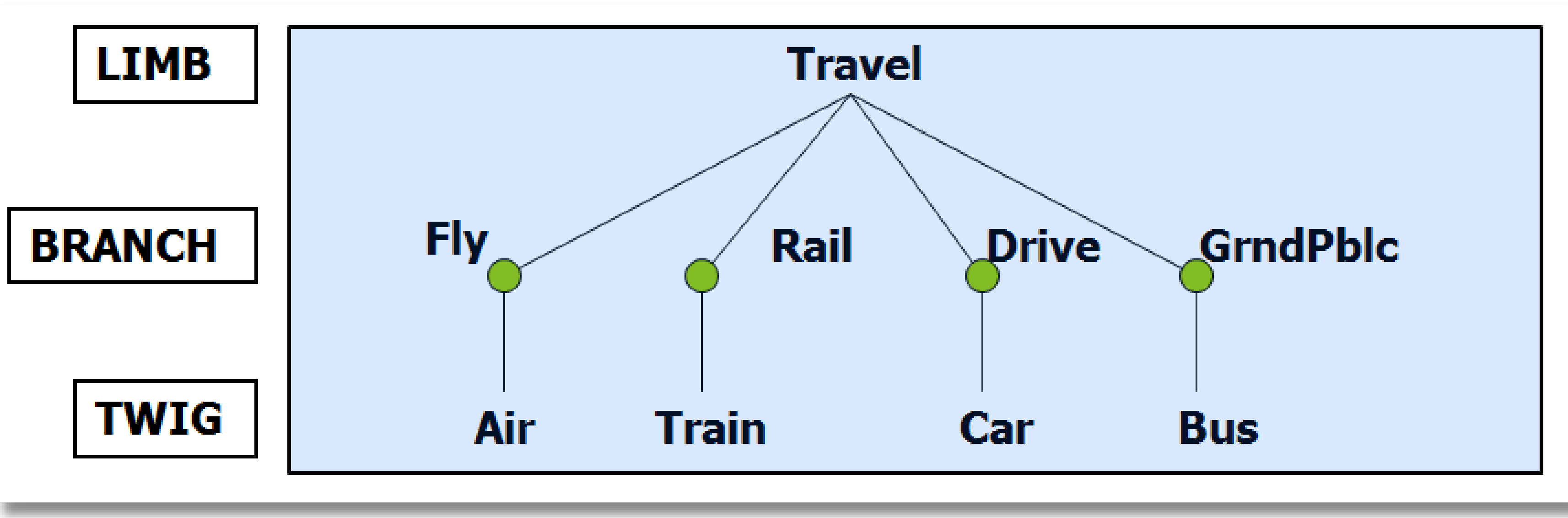


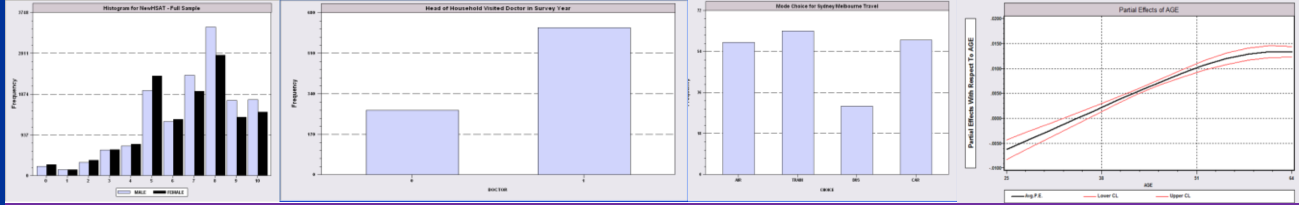
NL Model with a Degenerate Branch

```
-----------------------------------------------------------
FIML Nested Multinomial Logit Model
Dependent variable                  MODE
Log likelihood function      -148.63860
--------+--------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]
--------+--------------------------------------------------
        |Attributes in the Utility Functions (beta)
      GC|      .44230***       .11318       3.908   .0001
    TTME|     -.10199***       .01598      -6.382   .0000
    INVT|     -.07469***       .01666      -4.483   .0000
    INVC|     -.44283***       .11437      -3.872   .0001
   A_AIR|     3.97654***      1.13637       3.499   .0005
AIR_HIN1|      .02163          .01326       1.631   .1028
 A_TRAIN|     6.50129***      1.01147       6.428   .0000
TRA_HIN2|     -.06427***       .01768      -3.635   .0003
   A_BUS|     4.52963***       .99877       4.535   .0000
BUS_HIN3|     -.01596          .02000       -.798   .4248
        |IV parameters, lambda(b|l),gamma(l)
     FLY|      .86489***       .18345       4.715   .0000
  GROUND|      .24364***       .05338       4.564   .0000
--------+--------------------------------------------------
```



Using Degenerate Branches to Reveal Scaling

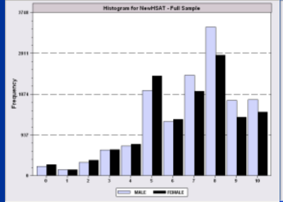
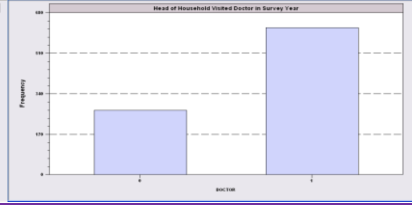
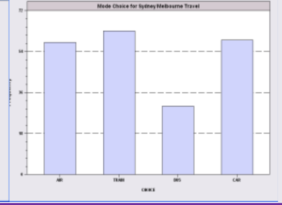
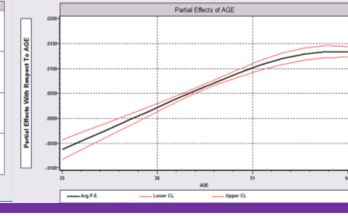




Scaling in Transport Modes

```
NLOGIT ; Lhs=mode
; Rhs=gc,ttme,invt,invc,one
; Choices=air,train,bus,car
; Tree=Fly(Air),
       Rail(train),
       LoclMass(bus),
       Drive(Car)
; ivset:(drive)=[1]$
```

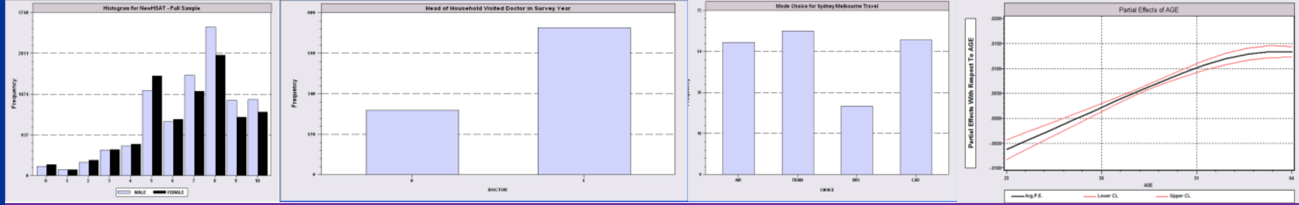
```
-----------------------------------------------------------
FIML Nested Multinomial Logit Model
Dependent variable                     MODE
Log likelihood function          -182.42834
The model has 2 levels.
Nested Logit form:IVparms=Taub|l,r,Sl|r
& Fr.No normalizations imposed a priori
Number of obs.=    210, skipped    0 obs
--------+--------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]
--------+--------------------------------------------------
        |Attributes in the Utility Functions (beta)
      GC|     .09622**         .03875        2.483    .0130
    TTME|    -.08331***        .02697       -3.089    .0020
    INVT|    -.01888***        .00684       -2.760    .0058
    INVC|    -.10904***        .03677       -2.966    .0030
   A_AIR|    4.50827***       1.33062        3.388    .0007
 A_TRAIN|    3.35580***        .90490        3.708    .0002
   A_BUS|    3.11885**        1.33138        2.343    .0192
        |IV parameters, tau(b|l,r),sigma(l|r),phi(r)
     FLY|    1.65512**         .79212        2.089    .0367
    RAIL|     .92758***        .11822        7.846    .0000
LOCLMASS|    1.00787***        .15131        6.661    .0000
   DRIVE|    1.00000        ......(Fixed Parameter)......
```

Simulation

NLOGIT ; lhs=mode;rhs=gc,ttme,invt,invc ; rh2=one,hinc; choices=air,train,bus,car
; tree=Travel[Private(Air,Car),Public(Train,Bus)] ; ru1
; simulation = * ; scenario:gc(car)=[*]1.5

```
|Simulations of Probability Model                              |
|Model: FIML: Nested Multinomial Logit Model                   |
|Number of individuals is the probability times the            |
|number of observations in the simulated sample.               |
|Column totals may be affected by rounding error.              |
|The model used was simulated with     210 observations.|
Specification of scenario 1 is:
Attribute  Alternatives affected            Change type            Value
---------  -------------------------------  ------------------- ---------
GC         CAR                              Scale base by value     1.500
Simulated Probabilities (shares) for this scenario:
+----------+--------------+--------------+------------------+
|Choice    |     Base     |   Scenario   | Scenario - Base  |
|          |%Share Number |%Share Number |ChgShare ChgNumber|
+----------+--------------+--------------+------------------+
|AIR       | 26.515    56 |  8.854    19 |-17.661%      -37 |
|CAR       | 29.200    61 |  6.836    14 |-22.364%      -47 |
|TRAIN     | 29.782    63 | 12.487    26 |-17.296%      -37 |
|BUS       | 14.504    30 | 71.824   151 | 57.320%      121 |
|Total     |100.000   210 |100.000   210 |   .000%        0 |
+----------+--------------+--------------+------------------+
```

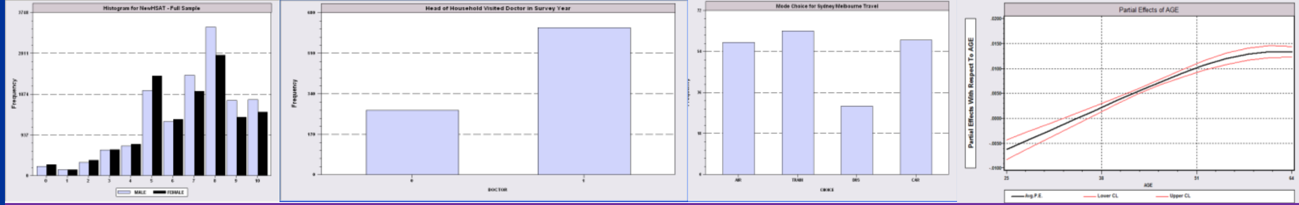


Nested Logit Approach for Best/Worst

```
nlogit
;lhs=choice,cset,altij
;choices=A,B,C,D
;model:
U(A) = Seat + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(B) = Scrn + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(C) = Alco + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(D) =        in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 $
```

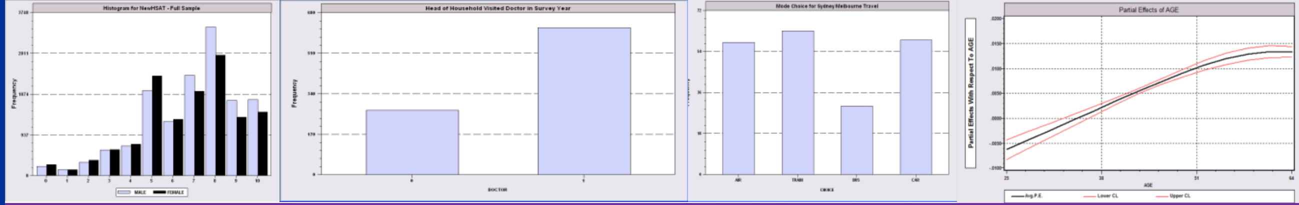
Note that you could potentially run any model for this data. For example, if one wanted to test for differences in scale between the best and worst alternatives, one could use the NL model (note however that you need the altij variable to take different values for best and worst now – see altn in the two examples above).

Uses the result that if U(i,j) is the lowest utility, -U(i,j) is the highest.



Nested Logit Approach.

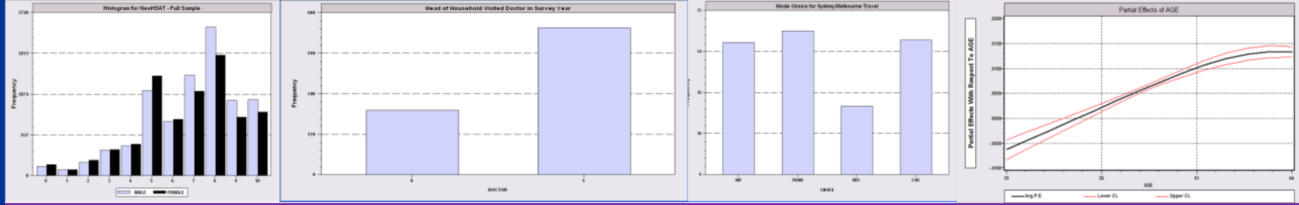
Resp	Set	Altij	Altn	Cset	Bestworst	Inch28	Inch30	CabScr	LimMov	Pay	Hour1	Hour3	Choice
1	1	1	1	4	1	0	1	0	0	0	0	0	0
1	1	2	2	4	1	0	0	0	1	0	0	0	0
1	1	3	3	4	1	0	0	0	0	0	0	0	1
1	1	4	4	4	1	0	0	0	0	0	0	0	0
1	1	1	5	4	-1	0	-1	0	0	0	0	0	0
1	1	2	6	4	-1	0	0	0	-1	0	0	0	1
1	1	3	7	4	-1	0	0	0	0	0	0	0	0
1	1	4	8	4	-1	0	0	0	0	0	0	0	0
1	2	1	1	4	1	1	0	0	0	0	0	0	0
1	2	2	2	4	1	0	0	1	0	0	0	0	0
1	2	3	3	4	1	0	0	0	0	1	0	0	1
1	2	4	4	4	1	0	0	0	0	0	0	1	0
1	2	1	5	4	-1	-1	0	0	0	0	0	0	1
1	2	2	6	4	-1	0	0	-1	0	0	0	0	0
1	2	3	7	4	-1	0	0	0	0	-1	0	0	0
1	2	4	8	4	-1	0	0	0	0	0	0	-1	0



Nested Logit Approach – Different Scaling for Worst

```
nlogit
;lhs=choice,cset,altn
;choices=Ab,Bb,Cb,Db, Aw,Bw,Cw,Dw
;tree =Bst(Ab,Bb,Cb,Db),Wst(Aw,Bw,Cw,Dw)
;ru1
;ivset:(Bst)=[1.0]
;model:
U(Ab) = Seat + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Bb) = Scrn + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Cb) = Alco + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Db) =        in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Aw) = Seat + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Bw) = Scrn + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Cw) = Alco + in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 /
U(Dw) =        in28*inch28 + in30*inch30 + scr*cabscr + limmov*limmov + pay*pay + h1*hour1 + h3*hour3 $
```

8 choices are two blocks of 4. Best in one brance, worst in the second branch



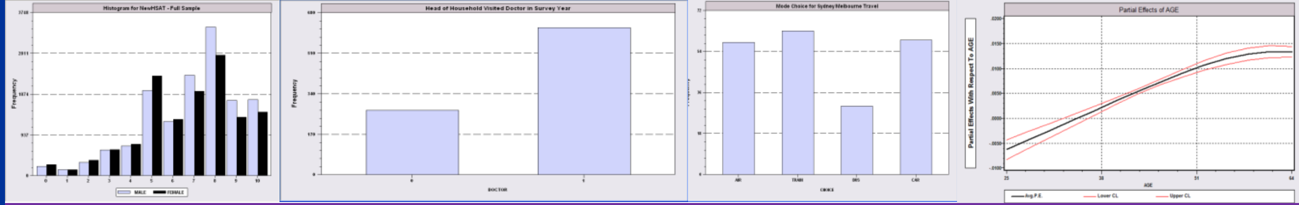
An Error Components Model

Random terms in utility functions share random components

$$\begin{array}{llllll}
U(Air,i) & = \alpha_{AIR} & + \beta_1 INVC_{i,AIR} & + \ldots + \varepsilon_{i,AIR} & + & w_{i,1} \\
U(Car,i) & = & \beta_1 INVC_{i,CAR} & + \ldots + \varepsilon_{i,CAR} & + & w_{i,1} \\
U(Train,i) & = \alpha_{TRAIN} & + \beta_1 INVC_{i,TRAIN} & + \ldots + \varepsilon_{i,TRAIN} & + & w_{i,2} \\
U(Bus,i) & = \alpha_{BUS} & + \beta_1 INVC_{i,BUS} & + \ldots + \varepsilon_{i,BUS} & + & w_{i,2}
\end{array}$$

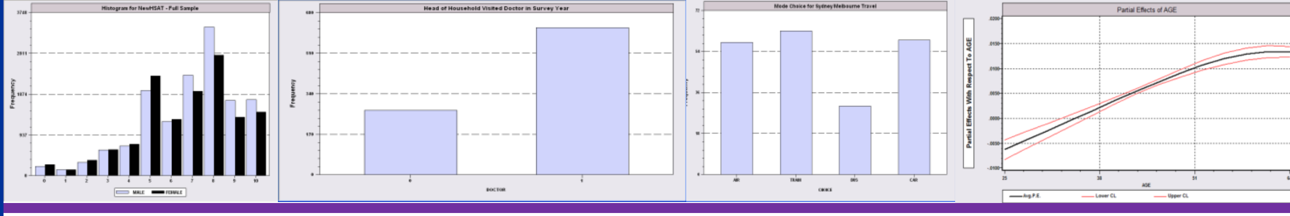
$$\text{Cov}\begin{pmatrix} \text{Air} \\ \text{Car} \\ \text{Train} \\ \text{Bus} \end{pmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 + \theta_1^2 & \theta_1^2 & 0 & 0 \\ \theta_1^2 & \sigma_\varepsilon^2 + \theta_1^2 & 0 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 + \theta_2^2 & \theta_2^2 \\ 0 & 0 & \theta_2^2 & \sigma_\varepsilon^2 + \theta_2^2 \end{bmatrix}$$

This model is estimated by maximum simulated likelihood.



Error Components Logit Model

```
-----------------------------------------------------------
Error Components (Random Effects) model
Dependent variable                      MODE
Log likelihood function           -182.27368
Response data are given as ind. choices
Replications for simulated probs. =   25
Halton sequences used for simulations
ECM model with panel has        70 groups
Fixed number of obsrvs./group=          3
Hessian is not PD. Using BHHH estimator
Number of obs.=   210, skipped    0 obs
--------+--------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]
--------+--------------------------------------------------
        |Nonrandom parameters in utility functions
      GC|      .07293***       .01978        3.687   .0002
    TTME|     -.10597***       .01116       -9.499   .0000
    INVT|     -.01402***       .00293       -4.787   .0000
    INVC|     -.08825***       .02206       -4.000   .0001
   A_AIR|     5.31987***       .90145        5.901   .0000
 A_TRAIN|     4.46048***       .59820        7.457   .0000
   A_BUS|     3.86918***       .67674        5.717   .0000
        |Standard deviations of latent random effects
SigmaE01|      .27336         3.25167         .084   .9330
SigmaE02|     1.21988          .94292        1.294   .1958
--------+--------------------------------------------------
```



The Multinomial Probit Model

$$U(i,t,j) = \alpha_j + \boldsymbol{\beta}'\mathbf{x}_{itj} + \boldsymbol{\gamma}_j'\mathbf{z}_{it} + \varepsilon_{i,t,j}$$

$$[\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_J] \sim \text{Multivariate Normal}[\mathbf{0}, \boldsymbol{\Sigma}]$$

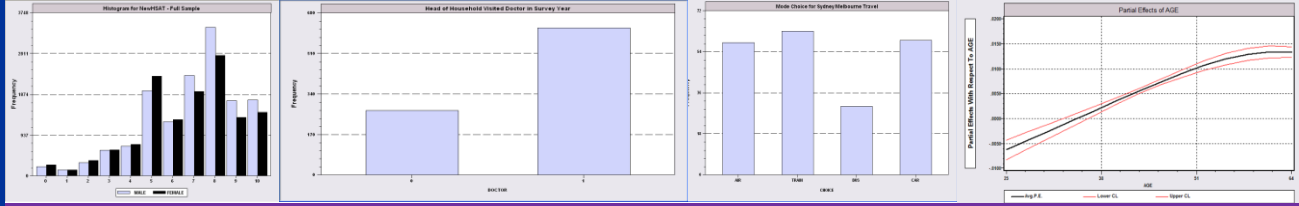
Correlation across choices

Heteroscedasticity

Some restrictions needed for identification

Sufficient: Last row of $\boldsymbol{\Sigma}$ = last row of I

One additional diagonal element = 1.



Multinomial Probit Probabilities

$$U(i,j) > U(i,1)$$
$$U(i,j) > U(i,2)$$
$$\ldots$$
$$U(i,j) > U(i,J)$$

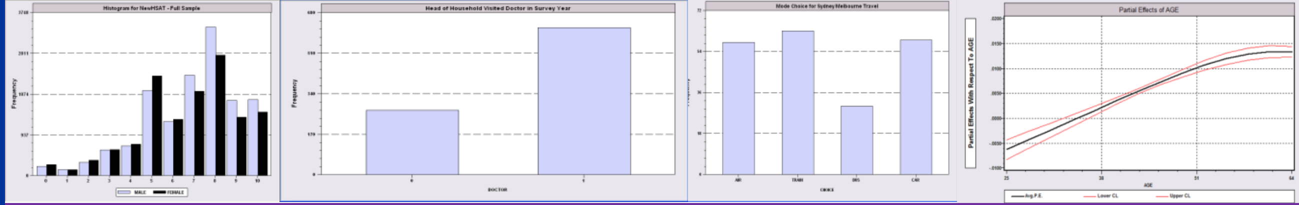
$$\text{Prob}(event) = \int_{-\infty}^{U(i,1)-U(i,J)} \int_{-\infty}^{U(i,1)-U(i,J-1)} \ldots \int_{-\infty}^{U(i,1)-U(i,1)} \phi_{J-1}(\varepsilon_J, \varepsilon_{J-1} \ldots \varepsilon_1 \,|\, \boldsymbol{\Sigma}_{(j)}) d\boldsymbol{\varepsilon}$$

Requires (J-1)-variate multivariate normal integration with a full covariance matrix.

The GHK simulator uses simulation to compute these probabilities accurately

using a simulation of the form $\text{Prob}(event) \approx \frac{1}{R}\sum_{r=1}^{R} \prod_{j=1}^{J-1} \Phi[h(w_{kr})]$.

w_{kr} = sequences of random N[0,1] draws using a simulator.



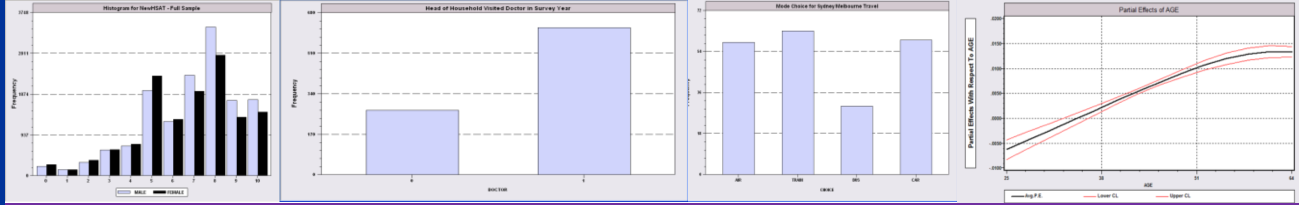
The problem of just reporting coefficients

```
Alternative-specific multinomial probit        Number of obs      =       840
Case variable: id                              Number of cases    =       210

Alternative variable: mode                     Alts per case: min =         4
                                                              avg =       4.0
                                                              max =         4

Integration sequence:      Hammersley
Integration points:               200              Wald chi2(5)   =     32.05
Log simulated-likelihood = -190.09418              Prob > chi2    =    0.0000

------------------------------------------------------------------------------
      choice |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
mode         |
  travelcost |    -.00977   .0027834    -3.51   0.000    -.0152253   -.0043146
    termtime |  -.0377095   .0094088    -4.01   0.000    -.0561504   -.0192686
-------------+----------------------------------------------------------------
air          |  (base alternative)
-------------+----------------------------------------------------------------
train        |
      income |  -.0291971   .0089246    -3.27   0.001     -.046689   -.0117052
       _cons |   .5616376   .3946551     1.42   0.155    -.2118721    1.335147
-------------+----------------------------------------------------------------
bus          |
      income |  -.0127503   .0079267    -1.61   0.108    -.0282863    .0027857
       _cons |  -.0571364   .4791861    -0.12   0.905    -.9963239     .882051
-------------+----------------------------------------------------------------
car          |
      income |  -.0049086   .0077486    -0.63   0.526    -.0200957    .0102784
       _cons |  -1.833393   .8186156    -2.24   0.025     -3.43785   -.2289357
-------------+----------------------------------------------------------------
      /ln12_2|  -.5502039   .3905204    -1.41   0.159     -1.31561    .2152021
      /ln13_3|  -.6005552   .3353292    -1.79   0.073    -1.257788    .0566779
-------------+----------------------------------------------------------------
        /l2_1|   1.131518   .2124817     5.33   0.000     .7150612    1.547974
        /l3_1|   .9720669   .2352116     4.13   0.000     .5110606    1.433073
        /l3_2|   .5197214   .2861552     1.82   0.069    -.0411325    1.080575
------------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------
Multinomial Probit Model
Dependent variable                 MODE
Log likelihood function      -189.43560
Restricted log likelihood    -291.12182
Chi squared [ 13](P= .000)    203.37243
Significance level               .00000
McFadden Pseudo R-squared      .3492909
Estimation based on N =    210, K =  13
Inf.Cr.AIC  =    404.9 AIC/N =    1.928
R2=1-LogL/LogL* Log-L fncn R-sqrd R2Adj
No coefficients  -291.1218  .3493  .3356
Constants only   -283.7588  .3324  .3183
At start values  -215.7845  .1221  .1036
Response data are given as ind. choices
Replications for simulated probs. = 100
Pseudo random draws (Mersenne twister).
Number of obs.=    210, skipped    0 obs
--------+--------------------------------------------------------------------
        |                Standard            Prob.       95% Confidence
    MODE|  Coefficient     Error       z    |z|>Z*          Interval
--------+--------------------------------------------------------------------
        |Attributes in the Utility Functions (beta)
      GC|   -.02377***     .00905    -2.63  .0086     -.04150    -.00604
    TTME|   -.09303***     .03596    -2.59  .0097     -.16351    -.02255
   A_AIR|   4.53688**     2.04884     2.21  .0268      .52123    8.55254
AIR_HIN1|    .01181        .02848      .41  .6785     -.04401     .06762
 A_TRAIN|   5.89123***    1.93928     3.04  .0024     2.09031    9.69215
TRA_HIN2|   -.06043***     .02181    -2.77  .0056     -.10317    -.01768
   A_BUS|   4.36644***    1.26372     3.46  .0005     1.88960    6.84328
BUS_HIN3|   -.01934        .01811    -1.07  .2854     -.05484     .01615
        |Std. Devs. of the Normal Distribution.
  s[AIR]|   2.81207**     1.27772     2.20  .0277      .30777    5.31636
s[TRAIN]|   1.91542**      .96835     1.98  .0479      .01749    3.81334
  s[BUS]|    1.0      .....(Fixed Parameter).....
  s[CAR]|    1.0      .....(Fixed Parameter).....
        |Correlations in the Normal Distribution
rAIR,TRA|   -.04362       1.18925     -.04  .9707    -2.37451    2.28726
rAIR,BUS|   -.12346       1.24418     -.10  .9210    -2.56200    2.31508
rTRA,BUS|    .56226        .47185     1.19  .2334     -.36254    1.48707
rAIR,CAR|    0.0      .....(Fixed Parameter).....
rTRA,CAR|    0.0      .....(Fixed Parameter).....
rBUS,CAR|    0.0      .....(Fixed Parameter).....
--------+--------------------------------------------------------------------
```

Stata: AIR = “base alternative”

Normalizes on CAR



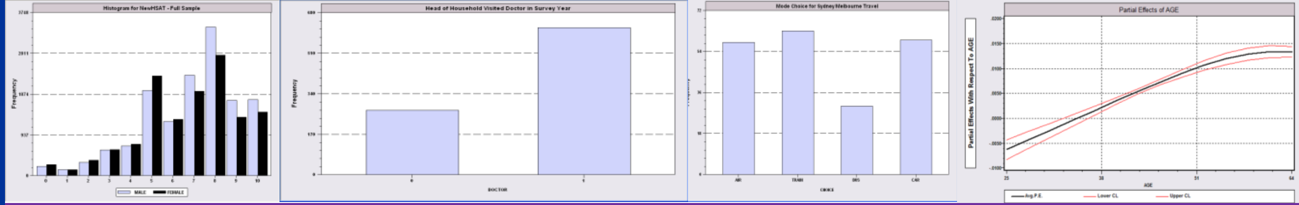
Multinomial Probit Model

```
+---------------------------------------------+
| Multinomial Probit Model                    |
| Dependent variable                MODE      |
| Number of observations             210      ||
| Log likelihood function      -184.7619      | Not comparable to MNL
| Response data are given as ind. choice.     |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
---------+Attributes in the Utility Functions (beta)
 GC      |    .10822534       .04339733     2.494   .0126
 TTME    |   -.08973122       .03381432    -2.654   .0080
 INVC    |   -.13787970       .05010551    -2.752   .0059
 INVT    |   -.02113622       .00727190    -2.907   .0037
 AASC    |   3.24244623      1.57715164     2.056   .0398
 TASC    |   4.55063845      1.46158257     3.114   .0018
 BASC    |   4.02415398      1.28282031     3.137   .0017
---------+Std. Devs. of the Normal Distribution.
 s[AIR]  |   3.60695794      1.42963795     2.523   .0116
 s[TRAIN]|   1.59318892       .81711159     1.950   .0512
 s[BUS]  |   1.00000000    ......(Fixed Parameter).......
 s[CAR]  |   1.00000000    ......(Fixed Parameter).......
---------+Correlations in the Normal Distribution
 rAIR,TRA|    .30491746       .49357120      .618   .5367
 rAIR,BUS|    .40383018       .63548534      .635   .5251
 rTRA,BUS|    .36973127       .42310789      .874   .3822
 rAIR,CAR|      .000000    ......(Fixed Parameter).......
 rTRA,CAR|      .000000    ......(Fixed Parameter).......
 rBUS,CAR|      .000000    ......(Fixed Parameter).......
```

Correlation Matrix for

Air, Train, Bus, Car

$$\begin{bmatrix} 1 & .305 & .404 & 0 \\ .305 & 1 & .370 & 0 \\ .404 & .370 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



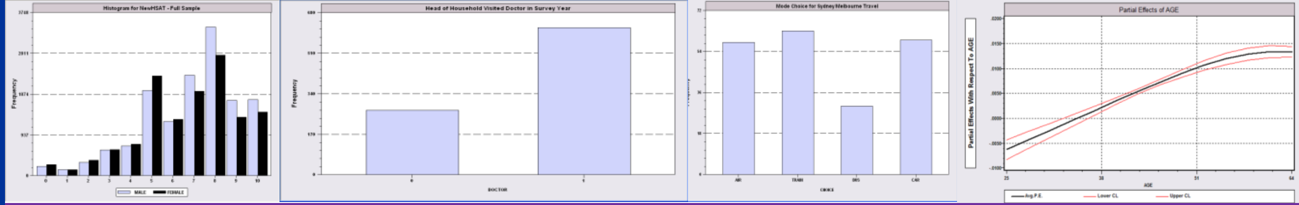
Multinomial Probit Elasticities

Elasticity averaged over observations.

Attribute is INVC in choice AIR		Mean	St.Dev
Effects on probabilities of all choices in model:			
* = Direct Elasticity effect of the attribute.			
*	Choice=AIR	-4.2785	1.7182
	Choice=TRAIN	1.9910	1.6765
	Choice=BUS	2.6722	1.8376
	Choice=CAR	1.4169	1.3250
Attribute is INVC in choice TRAIN			
	Choice=AIR	.8827	.8711
*	Choice=TRAIN	-6.3979	5.8973
	Choice=BUS	3.6442	2.6279
	Choice=CAR	1.9185	1.5209
Attribute is INVC in choice BUS			
	Choice=AIR	.3879	.6303
	Choice=TRAIN	1.2804	2.1632
*	Choice=BUS	-7.4014	4.5056
	Choice=CAR	1.5053	2.5220
Attribute is INVC in choice CAR			
	Choice=AIR	.2593	.2529
	Choice=TRAIN	.8457	.8093
	Choice=BUS	1.7532	1.3878
*	Choice=CAR	-2.6657	3.0418

Multinomial Logit

INVC in AIR	Mean	St.Dev
*	-5.0216	2.3881
	2.2191	2.6025
	2.2191	2.6025
	2.2191	2.6025
INVC in TRAIN		
	1.0066	.8801
*	-3.3536	2.4168
	1.0066	.8801
	1.0066	.8801
INVC in BUS		
	.4057	.6339
	.4057	.6339
*	-2.4359	1.1237
	.4057	.6339
INVC in CAR		
	.3944	.3589
	.3944	.3589
	.3944	.3589
*	-1.3888	1.2161



Not the Multinomial Probit Model MPROBIT

The MNP model is frequently motivated using a latent-variable framework. The latent variable for the jth alternative, $j = 1, \ldots, J$, is

$$\eta_{ij} = \mathbf{z}_i \boldsymbol{\alpha}_j + \xi_{ij}$$

where the $1 \times q$ row vector $\mathbf{z}_i$ contains the observed independent variables for the ith decision maker. Associated with $\mathbf{z}_i$ are the J vectors of regression coefficients $\boldsymbol{\alpha}_j$. The $\xi_{i,1}, \ldots, \xi_{i,J}$ are distributed independently and identically standard normal. The decision maker chooses the alternative k such that $\eta_{ik} \geq \eta_{im}$ for $m \neq k$.

This is identical to the multinomial logit – a trivial difference of scaling that disappears from the partial effects.
(Use ASMProbit for a true multinomial probit model.)