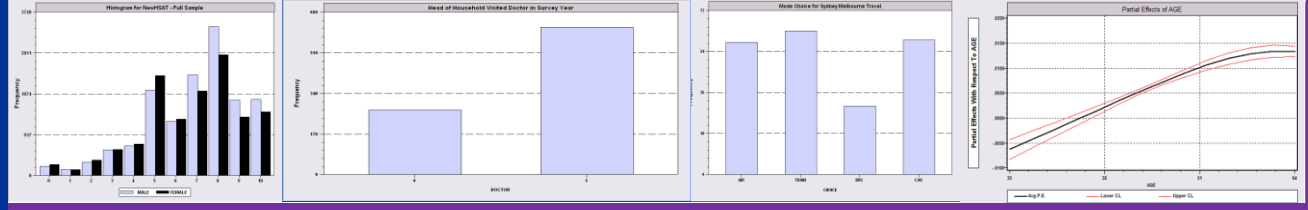
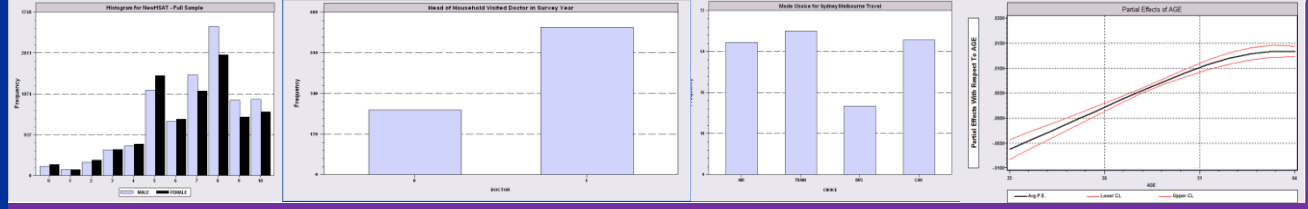


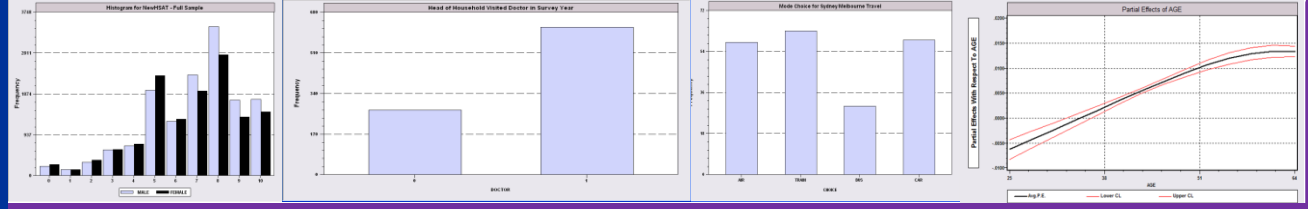
Discrete Choice Modeling

William Greene
Stern School of Business
New York University



Revealed and Stated Preference Data

- Pure RP Data
  - Market (ex-post, e.g., supermarket scanner data)
  - Individual observations
- Pure SP Data
  - Contingent valuation
  - (?) Validity
- Combined (Enriched) RP/SP
  - Mixed data
  - Expanded choice sets

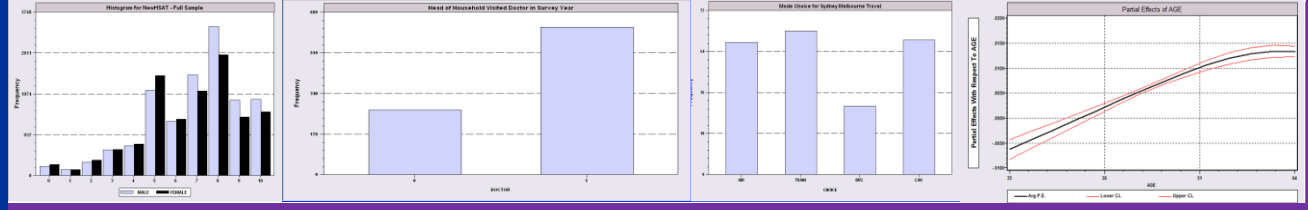


Application

Survey sample of 2,688 trips, 2 or 4 choices per situation
Sample consists of 672 individuals
Choice based sample

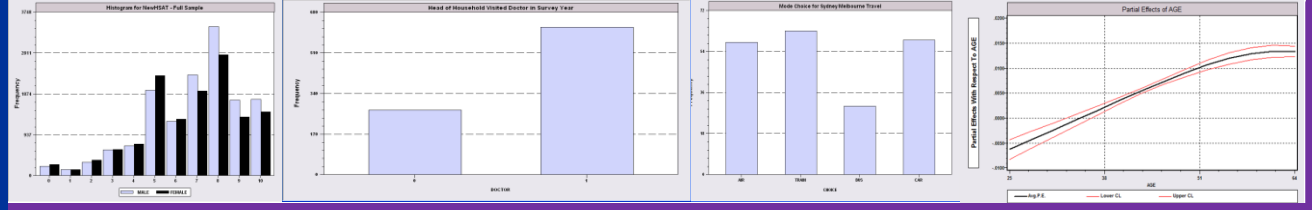
Revealed/Stated choice experiment:
- Revealed: Drive,ShortRail,Bus,Train
- Hypothetical: Drive,ShortRail,Bus,Train,LightRail,ExpressBus

Attributes:
- Cost –Fuel or fare
- Transit time
- Parking cost
- Access and Egress time



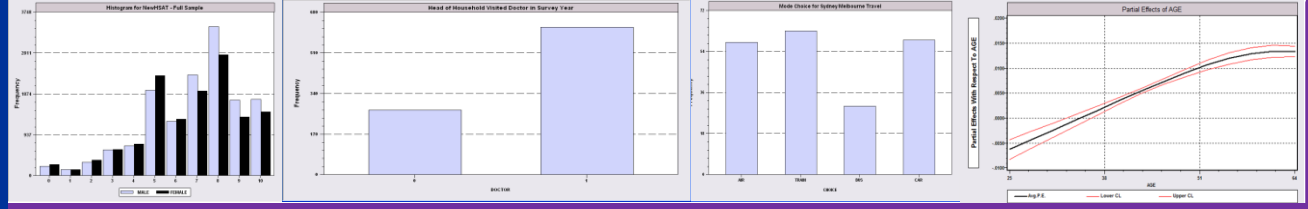
Application: Shoe Brand Choice

- Simulated Data: Stated Choice,
  - 400 respondents,
  - 8 choice situations, 3,200 observations
- 3 choice/attributes + NONE
  - Fashion = High / Low
  - Quality = High / Low
  - Price = 25/50/75,100 coded 1,2,3,4
- Heterogeneity: Sex (Male=1), Age (<25, 25-39, 40+)
- Underlying data generated by a 3 class latent class process (100, 200, 100 in classes)



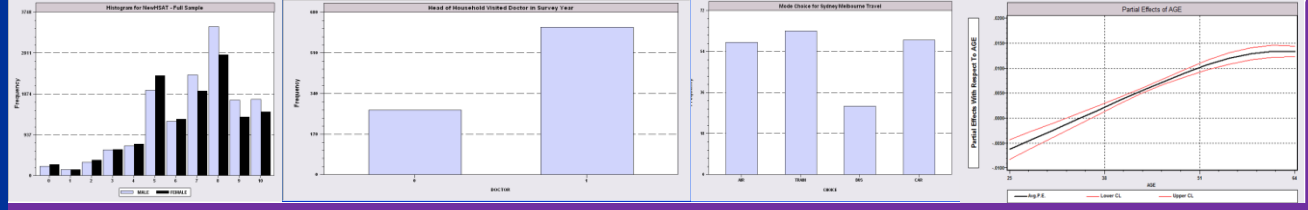
Stated Choice Experiment: Unlabeled Alternatives, One Observation

		ID	BRAND	CHOICE	FASH	QUAL	PRICE	PRICESQ	ASC4	
Brand 1 Brand 2 Brand 3 None	1 »	1	1	0	0	0	0.12	0.0144	0	t=1
	2 »	1	2	1	1	0	0.12	0.0144	0	
	3 »	1	3	0	0	1	0.08	0.0064	0	
	4 »	1	4	0	0	0	0	0	1	
	5 »	1	1	1	1	1	0.12	0.0144	0	t=2
	6 »	1	2	0	0	1	0.12	0.0144	0	
	7 »	1	3	0	1	0	0.12	0.0144	0	
	8 »	1	4	0	0	0	0	0	1	
Brand 1 Brand 2 Brand 3 None	9 »	1	1	0	0	1	0.08	0.0064	0	t=3
	10 »	1	2	0	1	1	0.2	0.04	0	
	11 »	1	3	1	1	0	0.08	0.0064	0	
	12 »	1	4	0	0	0	0	0	1	
	13 »	1	1	0	0	0	0.08	0.0064	0	t=4
	14 »	1	2	1	0	1	0.16	0.0256	0	
	15 »	1	3	0	1	1	0.2	0.04	0	
	16 »	1	4	0	0	0	0	0	1	
Brand 1 Brand 2 Brand 3 None	17 »	1	1	1	0	0	0.04	0.0016	0	t=5
	18 »	1	2	0	1	0	0.12	0.0144	0	
	19 »	1	3	0	1	0	0.08	0.0064	0	
	20 »	1	4	0	0	0	0	0	1	
	21 »	1	1	0	0	0	0.08	0.0064	0	t=6
	22 »	1	2	0	0	1	0.12	0.0144	0	
	23 »	1	3	1	1	0	0.08	0.0064	0	
	24 »	1	4	0	0	0	0	0	1	
Brand 1 Brand 2 Brand 3 None	25 »	1	1	0	1	1	0.2	0.04	0	t=7
	26 »	1	2	1	0	0	0.08	0.0064	0	
	27 »	1	3	0	0	1	0.08	0.0064	0	
	28 »	1	4	0	0	0	0	0	1	
	29 »	1	1	0	0	1	0.08	0.0064	0	t=8
	30 »	1	2	1	1	0	0.12	0.0144	0	
	31 »	1	3	0	0	0	0.04	0.0016	0	
	32 »	1	4	0	0	0	0	0	1	



Customers' Choice of Energy Supplier

- California, Stated Preference Survey
- 361 customers presented with 8-12 choice situations each
- Supplier attributes:
  - Fixed price: cents per kWh
  - Length of contract
  - Local utility
  - Well-known company
  - Time-of-day rates (11¢ in day, 5¢ at night)
  - Seasonal rates (10¢ in summer, 8¢ in winter, 6¢ in spring/fall)



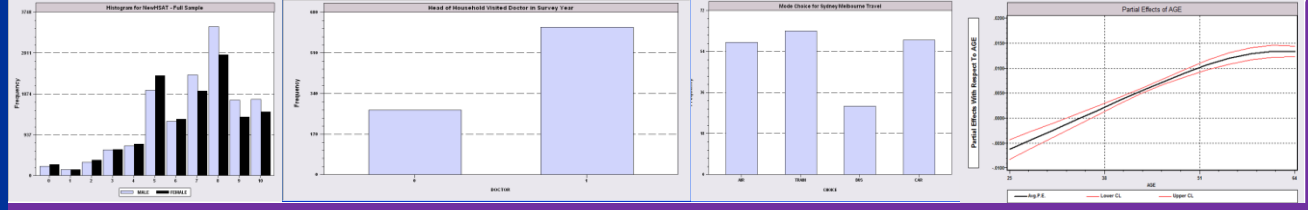
HEALTH ECONOMICS
Health Econ. **22**: 554–567 (2013)
Published online 20 April 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/hec.2824

INVESTIGATING ATTRIBUTE NON-ATTENDANCE AND ITS CONSEQUENCES IN CHOICE EXPERIMENTS WITH LATENT CLASS MODELS

MYLENE LAGARDE*

Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London, UK

Based on the vector of coefficient estimates β_i representing taste intensities, the probability that respondents would prefer a new set of guidelines to manage malaria in pregnancy over the current ones can be simulated by computing the probability associated with the utility derived from the new guidelines.

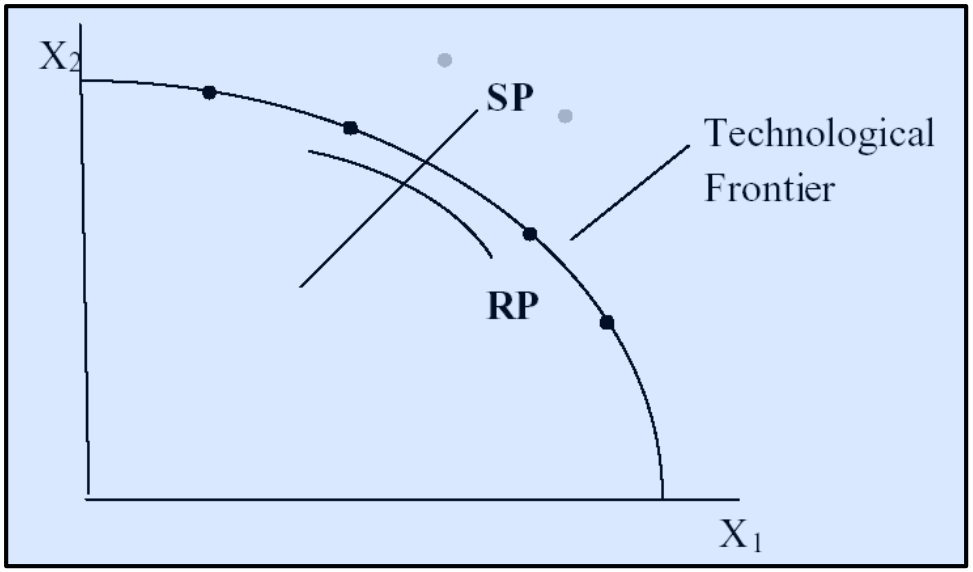


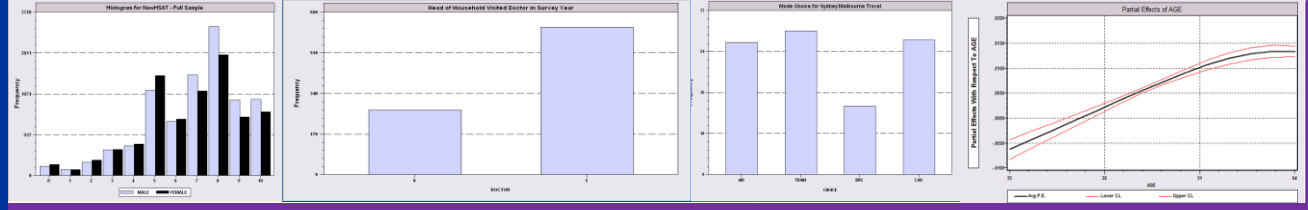
Panel Data

- Repeated Choice Situations
- Typically RP/SP constructions (experimental)
- Accommodating “panel data”
  - Multinomial Probit [marginal, impractical]
  - Latent Class
  - Mixed Logit

Revealed Preference Data

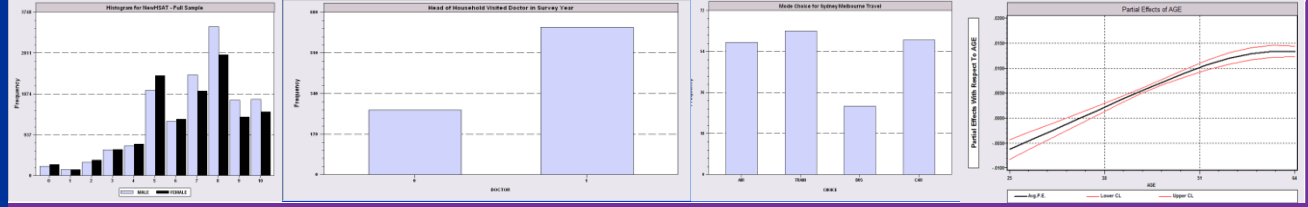
- Advantage: Actual observations on actual behavior
- Disadvantage: Limited range of choice sets and attributes – does not allow analysis of switching behavior.





Pooling RP and SP Data Sets - 1

- Enrich the attribute set by replicating choices
- E.g.:
  - RP: Bus,Car,Train (actual)
  - SP: Bus(1),Car(1),Train(1)
    Bus(2),Car(2),Train(2),…
- How to combine?



Discrete Choice Modeling

Stated Preference

101/900 Vars; 11111 Rows: 9408 Ot Cell: 0

	ID	CITY	SPRP	SPEXP	ALTIJ
1 »	1000	1	1	0	1
2 »	1000	1	1	0	4
3 »	1000	1	2	1	5
4 »	1000	1	2	1	6
5 »	1000	1	2	1	8
6 »	1000	1	2	1	10
7 »	1000	1	2	2	5
8 »	1000	1	2	2	6
9 »	1000	1	2	2	9
10 »	1000	1	2	2	10
11 »	1000	1	2	3	5
12 »	1000	1	2	3	6
13 »	1000	1	2	3	7
14 »	1000	1	2	3	8
15 »	1001	1	1	0	1

Each person makes four choices from a choice set that includes either two or four alternatives.

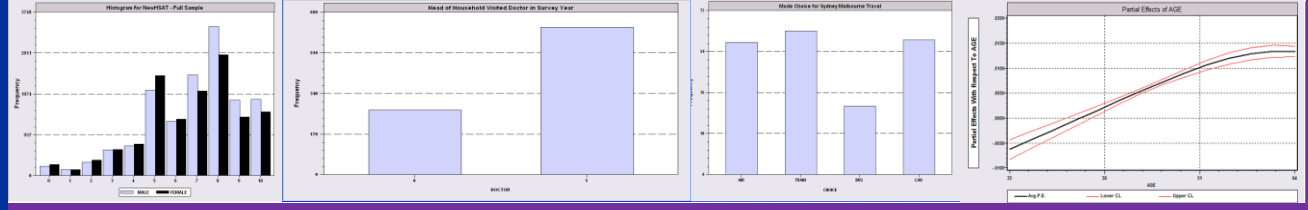
The first choice is the RP between two of the RP alternatives

The second-fourth are the SP among four of the six SP alternatives.

There are ten alternatives in total.

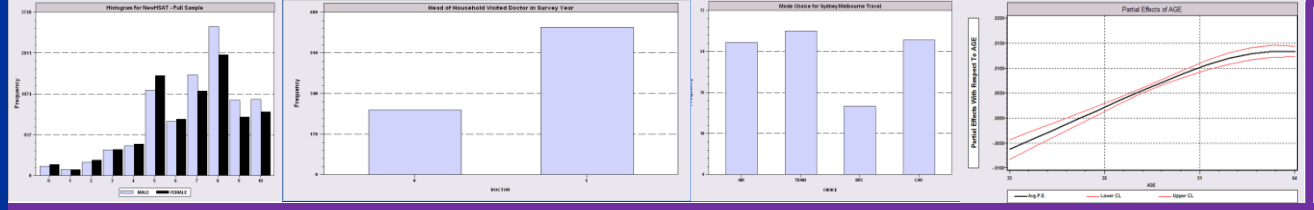
101/900 Vars; 11111 Rows: 9408 Ot Cell: 0

	ID	CHSNMODE	ALTMODE	SPCHOIC	CHOSEN	CSET
1 »	1000	11	2	0	1	2
2 »	1000	11	2	0	0	2
3 »	1000	0	0	5	1	4
4 »	1000	0	0	5	0	4
5 »	1000	0	0	5	0	4
6 »	1000	0	0	5	0	4
7 »	1000	0	0	10	0	4
8 »	1000	0	0	10	0	4
9 »	1000	0	0	10	0	4
10 »	1000	0	0	10	1	4
11 »	1000	0	0	8	0	4
12 »	1000	0	0	8	0	4
13 »	1000	0	0	8	0	4
14 »	1000	0	0	8	1	4
15 »	1001	11	12	0	1	2



Stated Preference Data

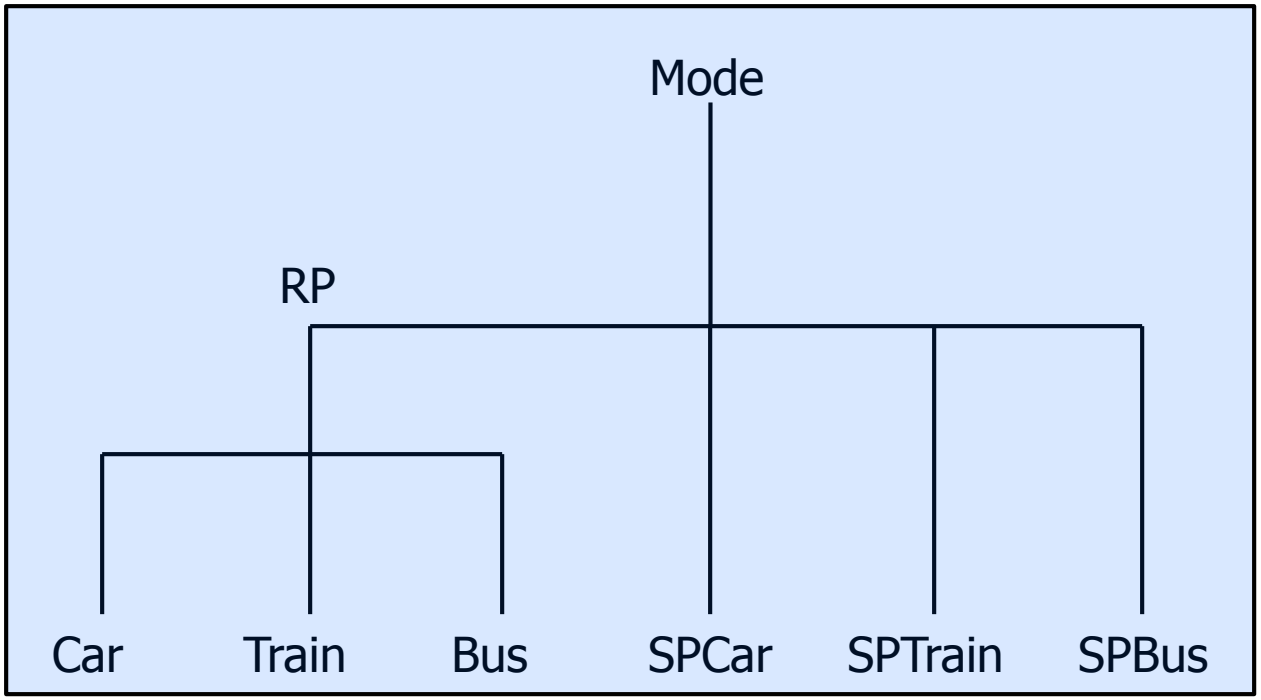
- Pure hypothetical – does the subject take it seriously?
- No necessary anchor to real market situations
- Vast heterogeneity across individuals



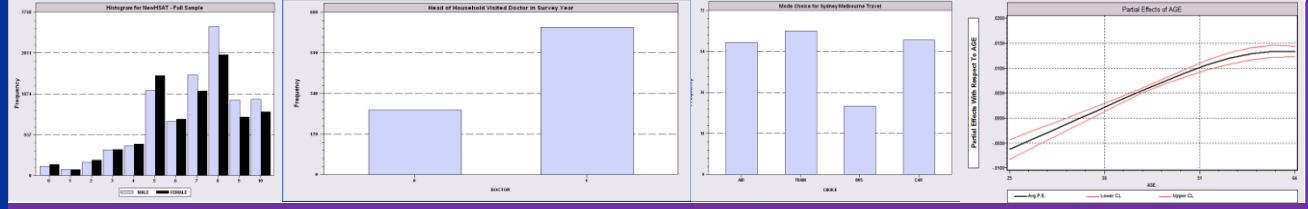
An Underlying Random Utility Model

$$\begin{aligned}
U(\textit{choice}_{rp}) &= \alpha + \boldsymbol{\beta}'\mathbf{x}_{rp} + \boldsymbol{\gamma}'\mathbf{y} + \varepsilon_{rp} \\
U(\textit{choice}_{sp}) &= \delta + \boldsymbol{\beta}'\mathbf{x}_{sp} + \boldsymbol{\theta}'\mathbf{z} + \varepsilon_{sp} \\
\sigma^2 &= \mathrm{Var}[\varepsilon_{rp}]/\mathrm{Var}[\varepsilon_{sp}] \\
&= \text{a scaling parameter such that } \mathrm{Var}[U_{rp}] = \sigma^2\mathrm{Var}[U_{sp}] \\
&\quad \text{so that pooling of the two data sets is valid,} \\
\mathbf{x}_{rp}, \mathbf{x}_{sp} &= \text{attributes common to the RP and SP data sets,} \\
\mathbf{y}, \mathbf{z} &= \text{observed attributes specific to the RP or SP data sets,} \\
[\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\delta,\boldsymbol{\theta}] &= \text{the unknown parameters to be estimated,} \\
\varepsilon_j,\ j = \mathrm{RP},\mathrm{SP} &= \text{unobserved individual effects.}
\end{aligned}$$

Nested Logit Approach



Use a two level nested model, and constrain three SP IV parameters to be equal.



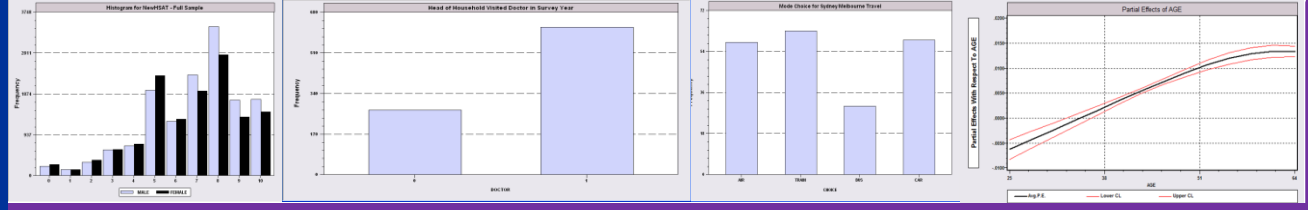
Enriched Data Set – Vehicle Choice

Choosing between Conventional, Electric and LPG/CNG Vehicles in Single-Vehicle Households

David A. Hensher
Institute of Transport Studies
School of Business
The University of Sydney
NSW 2006 Australia

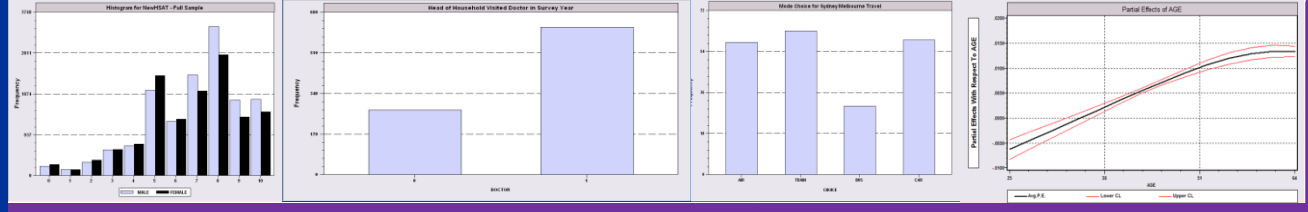
William H. Greene
Department of Economics
Stern School of Business
New York University
New York USA

September 2000



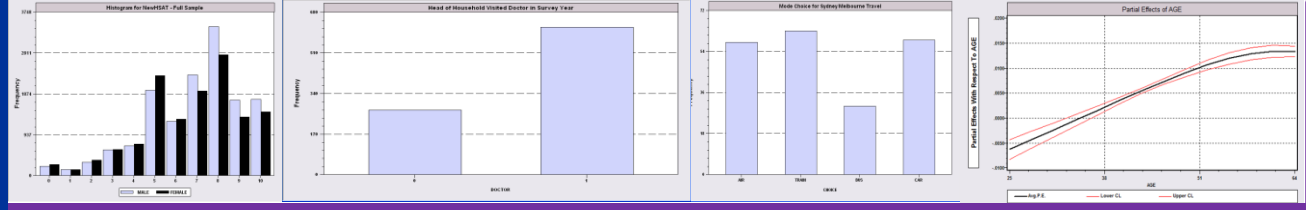
Fuel Types Study

- Conventional, Electric, Alternative
- 1,400 Sydney Households
- Automobile choice survey
- RP + 3 SP fuel classes
- Nested logit – 2 level approach – to handle the scaling issue



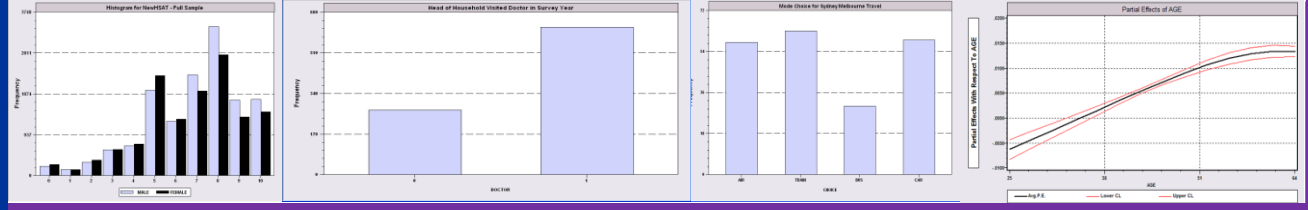
Attribute Space: Conventional

Conventional Vehicle		**OPTION 1**	**OPTION 2**	**OPTION 3**
Attributes		Small Car	Medium Car	Large Car
Price:	New vehicle ($'000)	15,25,35	30,40,50	30,55,80
	2 yr old vehicle ($'000)	11,18,26	22,30,37	22,42,62
	5 yr old vehicle ($'000)	7,12,17	15,20,25	15,30,45
	10 yr old vehicle ($'000)	5,7,10	9,12,16	10,20,30
Other costs:	Regn. fee (excl. any ins.) $	150,250,350	200,350,500	250,400,550
	Fuel cost to travel 500kms $	20,40,60	30,60,90	40,80,120



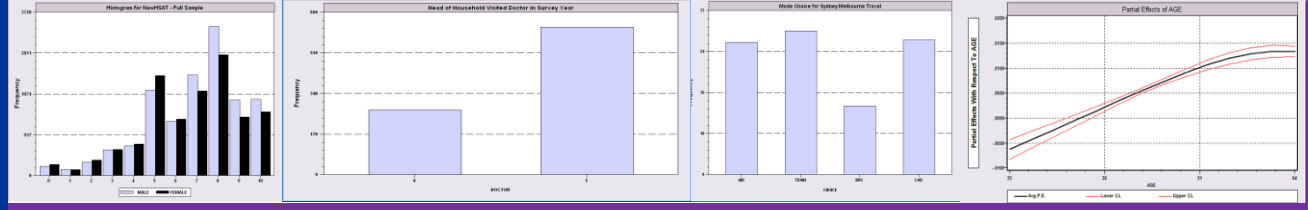
Attribute Space: Electric

Electric Vehicle		**OPTION 4**	**OPTION 5**	**OPTION 6**
Attributes		Small Car	Medium Car	Large Car
Price:	New vehicle ($'000)	15,28,40	25,40,55	30,70,110
	2 yr old vehicle ($'000)	12,22,32	20,32,44	24,55,90
	5 yr old vehicle ($'000)	9,17,24	15,24,33	18,42,65
	10 yr old vehicle ($'000)	6,12,17	11,17,25	13,30,50
Other costs:	Regn. fee (excl. any ins.) $	50,100,150	75,125,175	100,150,200
	Fuel cost to travel 500kms $	10,20,30	15,30,45	20,40,60
Features: (compared to	Fully fuelled range (% of)	90,70,50	90,70,50	90,70,50
conventional vehicles)	Acceleration	s,ss,cs	s,ss,cs	s,ss,cs
	Boot size (% of)	90,60	90,60	90,60



Attribute Space: Alternative

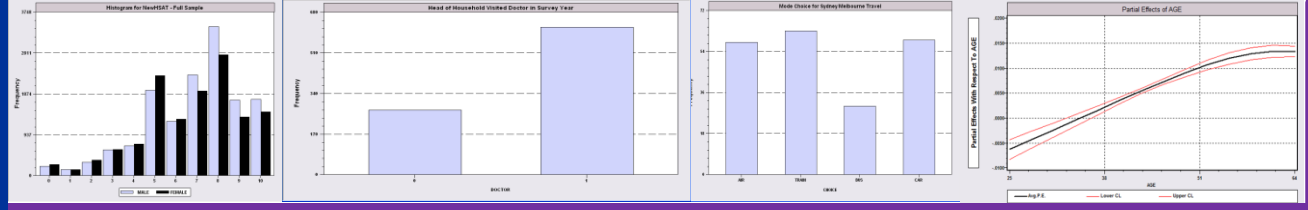
Alternative Fuel Vehicle		**OPTION 7**	**OPTION 8**	**OPTION 9**
Attributes		Small Car	Medium Car	Large Car
Price:	New vehicle ($'000)	15,25,35	30,40,50	30,55,80
	2 yr old vehicle ($'000)	11,18,26	22,30,37	22,42,62
	5 yr old vehicle ($'000)	7,12,17	15,20,25	15,30,45
	10 yr old vehicle ($'000)	5,7,10	9,12,16	10,20,30
Other costs:	Regn. fee (excl. any ins.) $	50,100,150	75,125,175	100,150,200
	Fuel cost to travel 500kms $	15, 30, 45	20,40,60	25,55, 85
Features: (compared to	Fully fuelled range (% of)	100,85,70	100,85,70	100,85,70
conventional vehicles)	Acceleration	s,ss,cs	s,ss,cs	s,ss,cs
	Boot size (% of)	90,70	90,70	90,70



Experimental Design

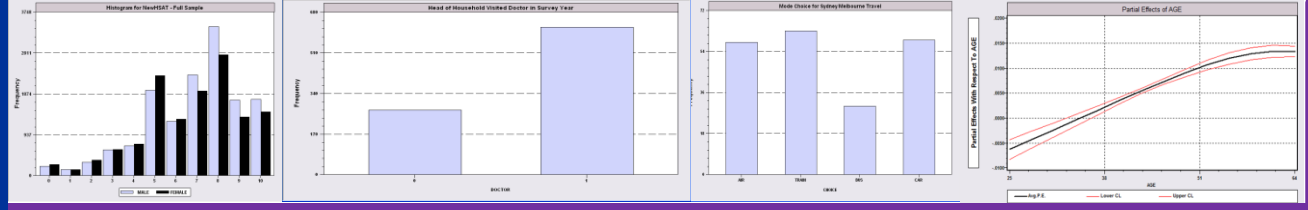
Table 1 Trip Attributes in Stated Choice Design
(Times are in minutes and costs are in $s)

Public Transport Attributes	Light rail	New heavy rail	Bus	Existing Busway	Existing train	Car attributes
Fare (one-way)	√	√	√	√	√	Running Cost
In-vehicle travel time	√	√	√	√	√	In-vehicle Travel time
Service frequency (per hour)	√	√	√	√	√	Toll Cost (One way)
Access Mode (AM) Walk time	√	√	√	√	√	Daily Parking Cost
AM Car time	√	√	√	√	√	Egress time
AM Bus time	√	√		√	√	
AM Bus fare	√	√		√	√	
Egress time	√	√	√	√	√	
Transfer time	√	√				
# attributes	9	9	6	8	8	5



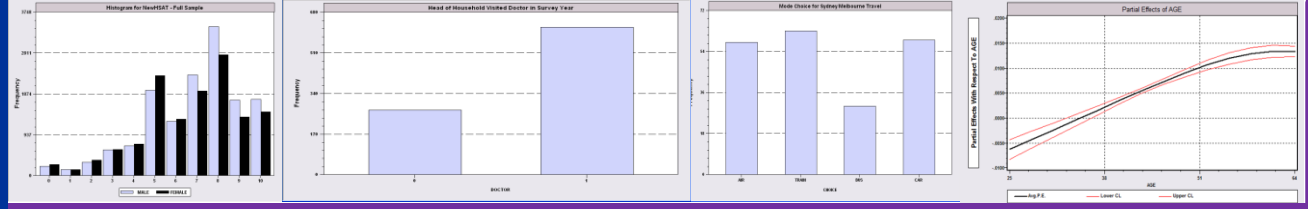
Valuation of Travel Time Savings in WTP and Preference Space in the Presence of Taste and Scale Heterogeneity

David A. Hensher
Institute of Transport and Logistics Studies
Faculty of Economics and Business
University of Sydney
NSW 2006 Australia
D.Hensher@itls.usyd.edu.au
William H. Greene
Department of Economics
Stern School of Business
New York University, New York 10012
wgreene@stern.nyu.edu



Mixed Logit Approaches

- Pivot SP choices around an RP outcome.
- Scaling is handled directly in the model
- Continuity across choice situations is handled by random elements of the choice structure that are constant through time
  - Preference weights – coefficients
  - Scaling parameters
    - Variances of random parameters
    - Overall scaling of utility functions

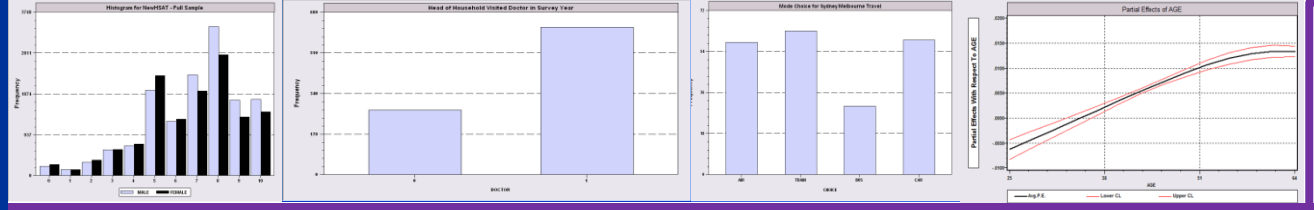


Application

Survey sample of 2,688 trips, 2 or 4 choices per situation
Sample consists of 672 individuals
Choice based sample

Revealed/Stated choice experiment:
- Revealed: Drive,ShortRail,Bus,Train
- Hypothetical: Drive,ShortRail,Bus,Train,LightRail,ExpressBus

Attributes:
- Cost –Fuel or fare
- Transit time
- Parking cost
- Access and Egress time



Each person makes four choices from a choice set that includes either 2 or 4 alternatives.

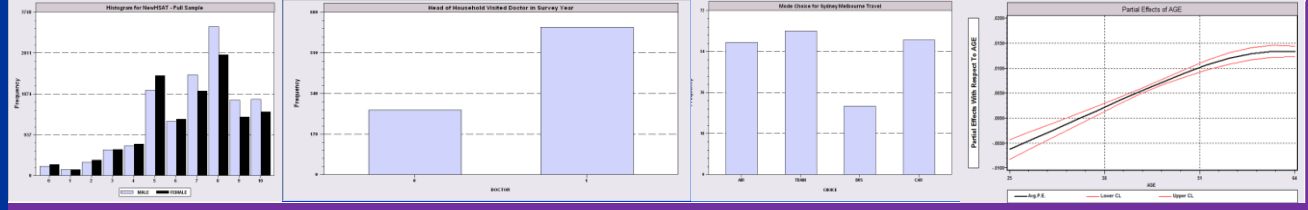
The first choice is the RP between two of the 4 RP alternatives

The second-fourth are the SP among four of the 6 SP alternatives.

There are 10 alternatives in total.

	ID	CITY	SPRP	SPEXP	ALTIJ	CHSNMODE	ALTMODE	SPCHOIC	CHOSEN	CSET
1 »	1000	1	1	0	1	11	2	0	1	2
2 »	1000	1	1	0	4	11	2	0	0	2
3 »	1000	1	2	1	5	0	0	5	1	4
4 »	1000	1	2	1	6	0	0	5	0	4
5 »	1000	1	2	1	8	0	0	5	0	4
6 »	1000	1	2	1	10	0	0	5	0	4
7 »	1000	1	2	2	5	0	0	10	0	4
8 »	1000	1	2	2	6	0	0	10	0	4
9 »	1000	1	2	2	9	0	0	10	0	4
10 »	1000	1	2	2	10	0	0	10	1	4
11 »	1000	1	2	3	5	0	0	8	0	4
12 »	1000	1	2	3	6	0	0	8	0	4
13 »	1000	1	2	3	7	0	0	8	0	4
14 »	1000	1	2	3	8	0	0	8	1	4

A Stated Choice Experiment with Variable Choice Sets



Experimental Design

Table 1 Trip Attributes in Stated Choice Design
(Times are in minutes and costs are in $s)

Public Transport Attributes	Light rail	New heavy rail	Bus	Existing Busway	Existing train	Car attributes
Fare (one-way)	√	√	√	√	√	Running Cost
In-vehicle travel time	√	√	√	√	√	In-vehicle Travel time
Service frequency (per hour)	√	√	√	√	√	Toll Cost (One way)
Access Mode (AM) Walk time	√	√	√	√	√	Daily Parking Cost
AM Car time	√	√	√	√	√	Egress time
AM Bus time	√	√		√	√	
AM Bus fare	√	√		√	√	
Egress time	√	√	√	√	√	
Transfer time	√	√				
# attributes	9	9	6	8	8	5