The University of Sydney

FACULTY OF ECONOMICS AND BUSINESS

Putting even more Behavior in Behavioral Models

David A. Hensher
Institute of Transport and Logistics Studies
Faculty of Economics and Business
The University of Sydney

June 28 – June 30 2009

Ryuichi Kitamura Memorial Symposium
University of California Davis

ITLS

The questions Ryuichi would ask…

- How reliable is evidence
  - that is inferred from stated or hypothetical contexts? (incl. CV and SC experiments)
  - from surveys where we ask individuals to report on attributes associated with non-chosen alternatives?
- This is now part of a growing literature on hypothetical bias (HB)

What is Hypothetical Bias?

- Hypothetical Bias is
  - The extent to which individuals might behave inconsistently, when they do not have to back up their choices with real commitments
    - […and when the survey imposes conditions that are not consistent with real world choice settings e.g., forced non-chosen alternatives]

- Of particular interest is the influence of hypothetical bias on marginal WTP (MWTP) and total WTP (TWTP) in a choice experiment context
  - Establishing the Gap, and 'Closing' or 'Reducing' the Gap

ITLS

Hypothetical Bias 'Gaps'

	Marginal Willingness to Pay (MWTP)	Hypothetical Bias
TR	"Truth"	Gap?
BD	Behaviour observed ' at a distance'	Gap?
RP	Revealed Preference (RP) Model	Gap?
CE	Choice Experiment (CE) Model	Gap?
CV	Contingent Valuation (CV) Model	
	vs	
TR	"Truth"	Gap?
BD	Behaviour observed ' at a distance'	Gap?
R_CE	Referenced Choice Experiment Model	Gap?
RP	Revealed Preference (RP) Model	Gap?
CV	Contingent Valuation (CV) Model	

Summing up re Clues from Contingent Valuation

- The evidence from many quality studies is well summarised in Murphy et al. (2005)

- "…it is likely that a number of factors affect hypothetical bias and therefore no single technique will be the magic bullet that eliminates this bias" page 317.
  - Cheap talk
  - Context
  - Experience

Choice Experiments (CE) 'to the rescue'!

- CE's are typically framed in a manner that adds realism, in that they more closely resemble individual purchasing or usage *decisions.*
  - But not always (see below).
- There are surprisingly few published studies that test for HB in CE
- Exceptions being Alfnes and Steine 2005, Lusk and Schroeder 2004, Cameron et al. 2002, Carlsson and Martinsson 2001, Carlsson et al.(2005), Johansson-Stenman and Svedsäter (2003), Brownstone and Small 2005, and Isaccson 2007.
  - Carlsson and Martinsson (2001) and Cameron et al. (2002) **fail to reject** a hypothesis of equal MWTP in both a real and a hypothetical setting,
  - Johansson-Stenman and Svedsäter (2003) **reject** the equality of MWTPs, and
  - Lusk and Schroeder (2003) find that hypothetical TWTP for the good exceeds real TWTP, but **fail to reject** the equality of MWTPs for changes in the single attributes.

ITLS

Transport Studies and HB

- There are two key transport studies that investigate HB:
  - Brownstone and Small (B and S)(2005)
  - Isacsson (2007)
- Using a simple dichotomous choice experiment with two attributes, Isacsson (2007) suggests that there is a bias in estimates of the VOTS associated with public transport, based on hypothetical choices.
  - Real values tend to be *higher* than values derived from hypothetical choices.
  - This replicates the findings of B and S (2005).
  - 'Real' choices in Isaccson produced an estimated mean value of time savings which was **twice as large** as the corresponding hypothetical value.

Transport Studies and HB

- This ‘limited’ evidence in CE’s in transportation applications, is the *inverse* to the general findings in CV studies, where CV studies conclude that
  - hypothetical WTP estimated in stated preference surveys is most often found to be an *overstatement* of ‘true’ WTP
- WARNING – when looking at studies, ensure you know if CV or CE

Mark Wardman's Commentary

- Wardman (2001) and B and S (2005) suggest a range of reasons as to why CE and RP MWTP may differ.
- Wardman (2001, 120) suggests that a ***lower CE MWTP*** can be explained (in part) by
  - (i) ***strategic response bias***, especially on the parameter of cost, associated with greater sensitivity to cost variation that a CE generates (i.e., **higher marginal disutility**);
  - (ii) the ability in a CE to "… adopt simplified decision rules such as **ignoring attributes** of lesser importance or which vary less" (confirmed recently by Hensher and Layton 2009); although this is also an issue with RP data, and
  - (iii) a variation on (ii), to **ignore attribute variations** which are not realistic, thereby reducing mean parameter estimates. Wardman suggests this is more likely to be an issue for a parameter estimate which is the numerator of the WTP calculation (such as travel time).

ITLS

Brownstone and Small's Contribution

- B and S (2005) also offer some explanations for the differences, also cited in Isacsson (2007).
- The most appealing is that individuals display (time) inconsistency in their actual behaviour, or more generally
  - ***time budget or scheduling*** constraints associated with real actions that are not accounted for in SP experiments.
  - **A point made well by Ryuichi on many occasions (linked to activity research)**
- It is suggested that these constraints tend to result in higher-cost choices more frequently in real life than in hypothetical surveys.
  - **Hence lower marginal disutility of cost in reality**
- They also consider the misperception of travel time:
  - They ask individuals to report the time savings they think could be realised by using express lanes. This belief elicitation was non-incentivised.
  - Individuals typically report an estimate (based on the mean), twice the actual time-savings.

ITLS

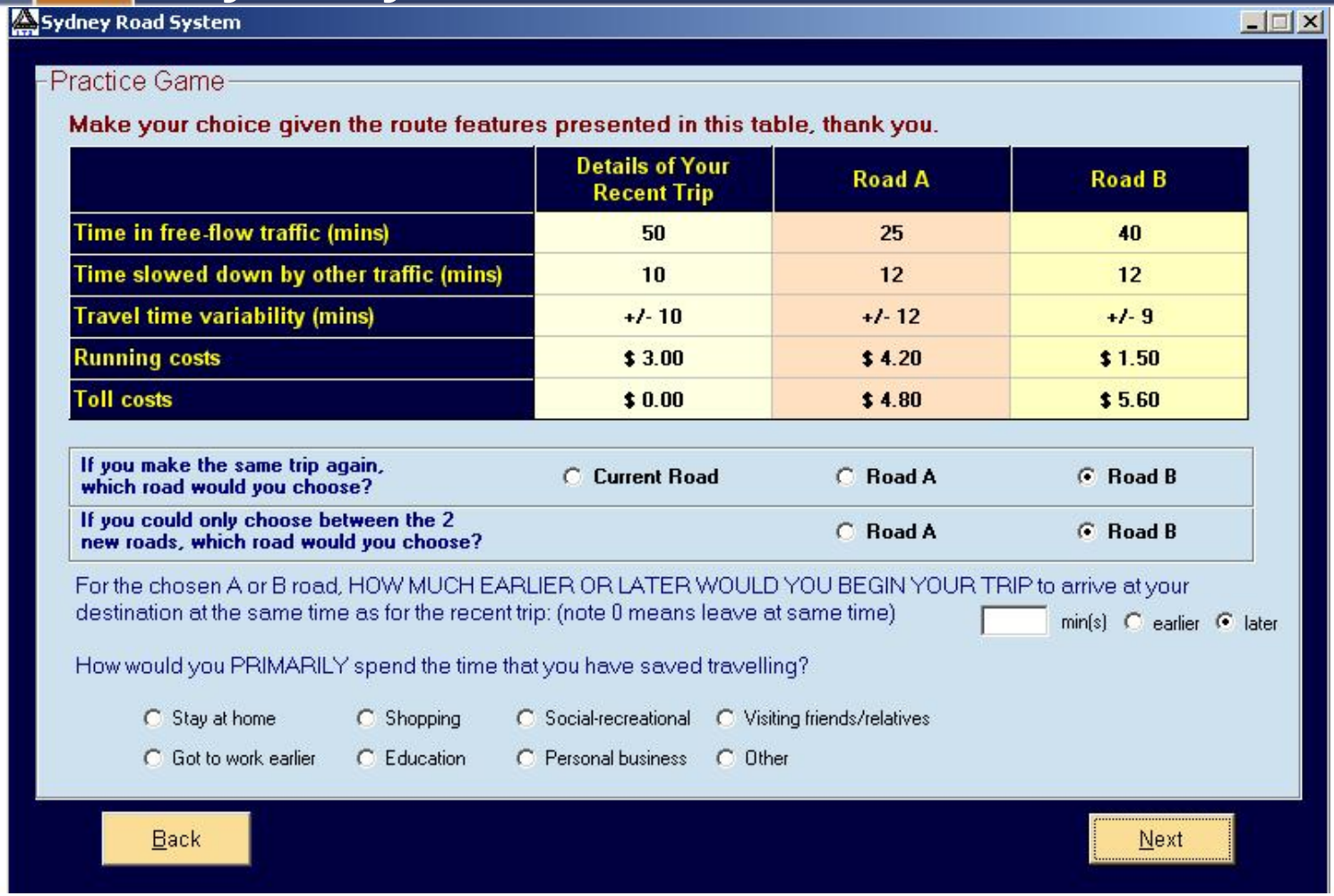
Brownstone and Small's Contribution

- B and S suggest two possible explanations of misperception (2005, 288):
  - individuals focus on total delays on *part of the trip* instead of the full origin-destination trip,
  - impatience with heavy traffic leads to exaggeration of actual delay time
    - Hensher and Layton (2009) refer to this as common-metric parameter transfer.
- These reasons are then used to suggest that the same level of an attribute in a CE will lead to a ***different reaction*** and a lowering of the parameter estimate for time.

Pivot designs – elements of RP and CE

- A Proposition:
  - Given the different context of an SP experiment in general:
  - CE studies with Referencing might be CONSIDERED an alternative to traditional RP studies, where ‘traditional’ RP data is deficient.
  - **Offering a way to reduce the HB gap?**
- Pivot (or Referencing) designs may well be a way forward
  - Since they offer relevant RP data
  - ***Especially in habit contexts***
- ***Point often made by Ryuichi about inertia in travel behaviour***

Sydney CAPI Choice Scenario Screen



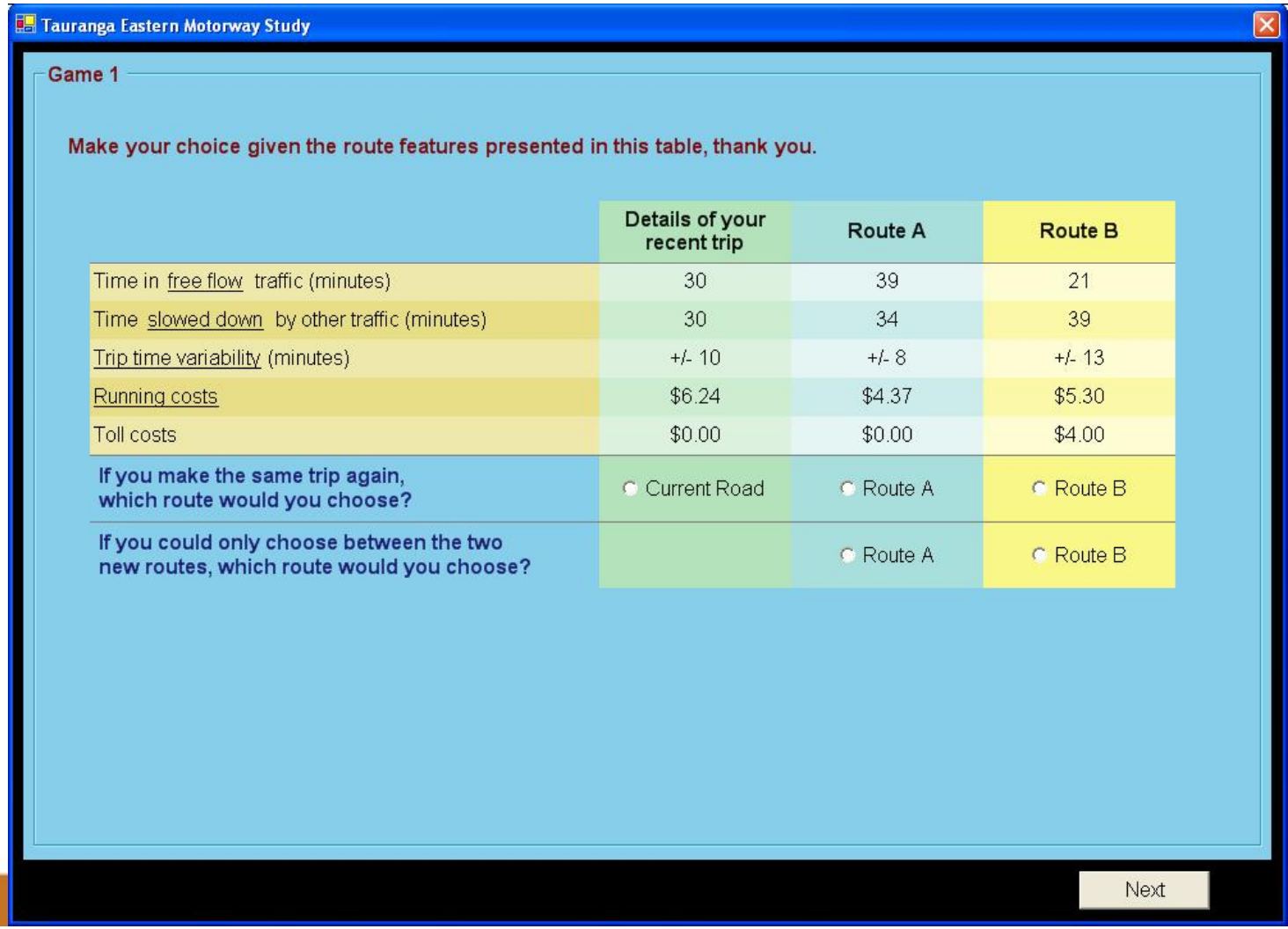
Some Initial Evidence – Sydney
(Mixed Logit, Error Component, Panel, 500 Halton Draws)

- For the Sydney study

$$\text{Prob}(y_{it} = j) = \frac{\exp\left[\alpha_{ji} + \beta_i' \mathbf{x}_{jit} + \Sigma_{m=1}^{M} d_{jm} \theta_m E_{im}\right]}{\sum_{q=1}^{J_i} \exp\left[\alpha_{qi} + \beta_i' \mathbf{x}_{qit} + \Sigma_{m=1}^{M} d_{qm} \theta_m E_{im}\right]}$$

  - The mean VTTS for the **reference alternative is $26.99** per person hour with a standard deviation of $7.94;
  - The mean for the **CE alternatives is $17.92** (standard deviation of $7.82), derived from the model that includes the reference alternative.
  - The ratio of the Reference to CE alternatives VTTS is **1.51**.
  - (note – fully generic model mean VTTS =$18.80, close to CE VTTS).

NZ CAPI Choice Scenario Screen



Some Initial Evidence – New Zealand
(Mixed Logit, Error Component, Panel, 500 Halton Draws)

- For the New Zealand study
  - The mean VTTS for the **reference alternative is $27.34** per person hour, with a standard deviation of $7.46;
  - The mean for the **CE alternatives is $13.65** (standard deviation of $4.31), derived from the model that includes the reference alternative.
  - The ratio of the Reference to CE alternatives VTTS is **2.00**.

ITLS

An Empirical Revelation?

- We find that the marginal disutility associated with travel time in the reference alternative is substantially higher than that associated with the CE design alternatives, and likewise for cost
  - **BUT** the marginal disutility for travel time outweighs the marginal disutility of cost, resulting in the higher VTTS for the reference (or real market experienced) alternative.
- The evidence from other studies co-authored by Hensher and Louviere (over a number of years) that **the attribute range has the greatest influence on MWTP than any other dimension of choice experiments**, with MWTP being higher with a reduced attribute range, supports the findings herein
  - the CE design alternatives have a wider attribute range relative to the range of attributes of other alternatives that people face in real choices, and
  - hence a lower mean VTTS than the mean VTTS from the real market alternative.

ITLS

***An Empirical Revelation! ***

- In CE studies, it is common to have a wider range of an attribute to assess; that is essentially what choice experiments are all about – creating behaviourally richer variance.
  - However this may come at a price, in that real markets are not so rich in variability, and hence when actual market data is used, we observe after estimation, higher MWTP compared to a choice experiment.

- This naturally begs the question – *does the ratio of the range of each attribute in the numerator and denominator of the calculation of MWTP for the Reference and CE observations account for part or all of the difference in the mean MWTP?*

  - Sydney example: RP VTTS =1.51of SC VTTS
  - Ratio of SC/RP: time range = 1.42; cost=1.48
  - **Coincidence or of Empirical interest?**

What have we (maybe) failed to see!

- The empirical evidence herein may suggest that, for all the years of interest in choice experiments, and the debate about the role of RP and CE data, *we may have missed or masked an important message:*
  - namely that choice experiments with referencing back to a real market activity may provide a suitable specification
    - especially where repeated behaviour or habit is dominant
  - short of capturing data ‘at a distance’, where the latter has evaded almost every single travel study to date.

Modelling Behavioural Variance is what it is all about

- If we recognise that the requirement to seek data on at least one non-chosen alternative in RP modelling is linked to the creation of variance necessary to estimate a model,
  - then this imposition in the context of habitual behaviour may be better (or also) accommodated by variance revelation through an CE pivot design,
  - where the only information required from real markets relates to the habitually selected alternative.
  - Something that Ryuichi would find intriguing and not inconsistent with his views on habitual behavior.

ITLS

Final Message

- Pivot based CE data has the power of richness to enable respondents to express preferences involving not only the actual memory, but also ***related*** *memories constructed from it* (Hensher 2006).
- We do, however, emphasise that the evidence herein in support of the directional and magnitude differences between WTP associated with RP and CE alternatives should not be seen as anything more than **encouraging consistency**.
- Natural field experiments are required to test this preliminary finding.

We still have useful information, but….

- **The point here is to distinguish, as a matter of logic:**
  - **the fact that we can predict real valuations given the hypothetical valuations *and* some assumption about the HB, and**
  - **the fact that the hypothetical valuations in and of themselves were wrong.**

Hypothetical Bias 'Gaps'

	Marginal Willingness to Pay (MWTP)	Hypothetical Bias
TR	"Truth"	Gap?
BD	Behaviour observed ' at a distance'	Gap?
RP	Revealed Preference (RP) Model	Gap?
CE	Choice Experiment (CE) Model	Gap?
CV	Contingent Valuation (CV) Model	
	vs	
TR	"Truth"	Gap?
BD	Behaviour observed ' at a distance'	Gap?
R_CE	Referenced Choice Experiment Model	Gap?
RP	Revealed Preference (RP) Model	Gap?
CV	Contingent Valuation (CV) Model	

Putting more Behaviour In…..

- Attribute Processing
  - Relevancy vs. Complexity
  - Heuristics in Choice Analysis
    - Through Model Specification
    - Through Supplementary questions
- Group Choosing

A Reminder

- Statistical non-significance of a parameter of an attribute should **not** be seen as necessarily meaning ‘not relevant’
  - unless the analyst has been able to quarantine the information that is ‘not relevant’ (at least up to a probability)

Ryuichi as recent as 2008

Paper Title: "SHALL WE MIXED LOGIT? ESTIMATION STABILITY AND PREDICTION RELIABILITY OF ERROR COMPONENT MIXED LOGIT MODELS"

- "It follows from (i) that it is in general impossible, through model estimation, to discern whether heteroskedasticity exists, or error terms are correlated across alternatives, or error terms are heteroskedastic *and* correlated with each other"
- This is a recognition of the research that accounts for scale and taste heterogeneity

Generalised MNL: Scale and Preference Heterogeneity

$$U_{qjt} = (\sigma_q \beta + \gamma \eta_q + (1-\gamma)\sigma_q \eta_q) x_{qjt} + \varepsilon_{qjt}$$

$$\sigma_q = \exp(\bar{\sigma} + \delta z_q + \tau \varepsilon_{0q})$$

MNL: $\gamma_q = 0$, $\text{var}(\eta_q)=0$, $\sigma_q=\sigma=1$

Scale heterogeneity = $\sigma_q\beta$

Taste heterogeneity = $\beta + \eta_q$

Links between Scale and Preference Heterogeneity

- If the standard deviation of residual taste heterogeneity is independent of the scaling of β, then:
- $\gamma\eta_q$ indicates that taste heterogeneity is independent of scale heterogeneity
- If the standard deviation of η_q is proportional to σ_q, then
- $(1-\gamma)\sigma_q\eta_q$ indicates proportionality between taste and scale heterogeneity
- The parameter γ governs how the variance of residual taste heterogeneity varies with scale, in a model that includes both. γ *lies between 0 and 1*

Generalised MNL

- τ is the key parameter that indicates if scale heterogeneity is present in the data.
- As τ 0, G-MNL approaches the standard mixed logit (or random parameters) model in which all unobserved heterogeneity is associated with taste

The University of Sydney

FACULTY OF ECONOMICS AND BUSINESS

There is Much more to Do

We thank Ryuichi for Giving us Countless Clues

Thank You

ITLS