# Behavioural Realism: Heterogeneity in Decision Processes and Experience Conditioning


David A. Hensher, PhD, FASSA

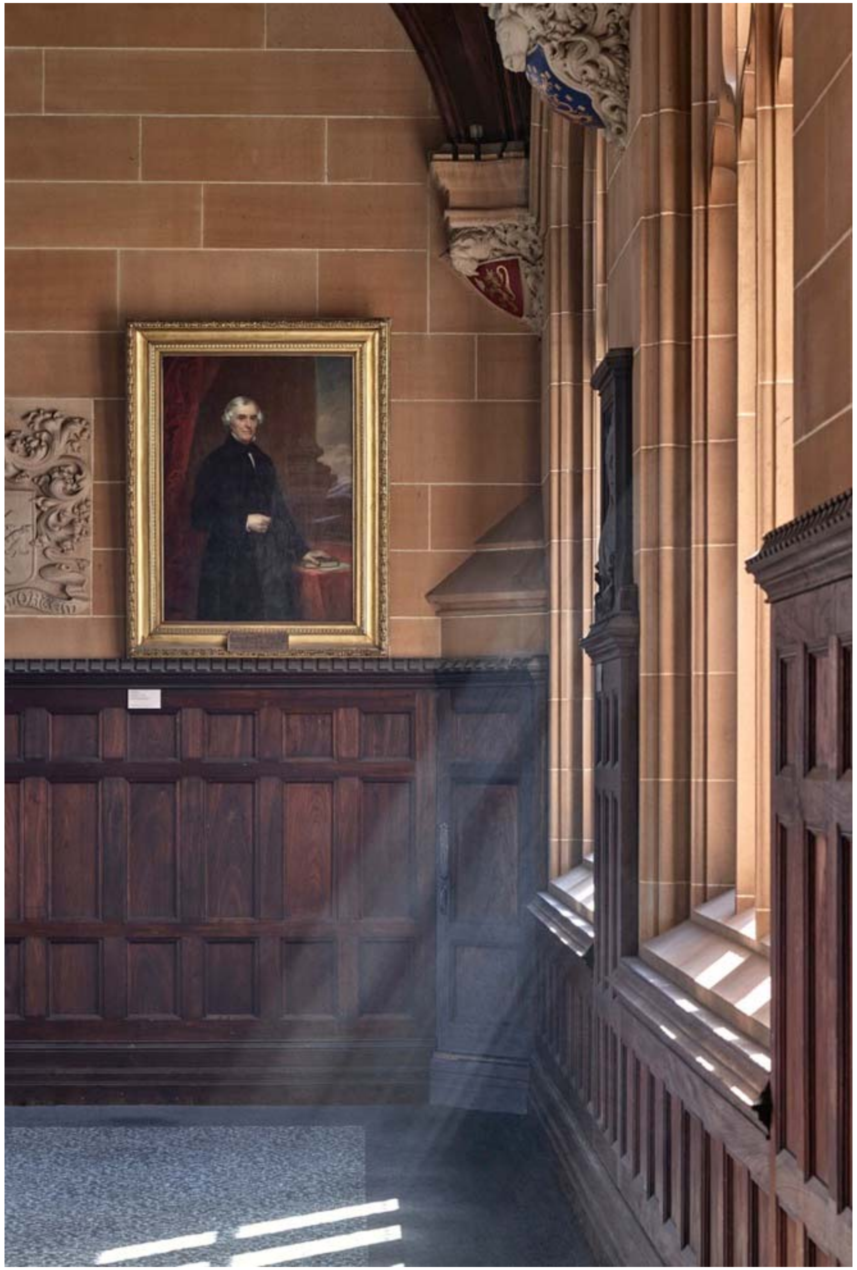
Founder and Director of the Institute of Transport and Logistics Studies (ITLS)

The University of Sydney Business School

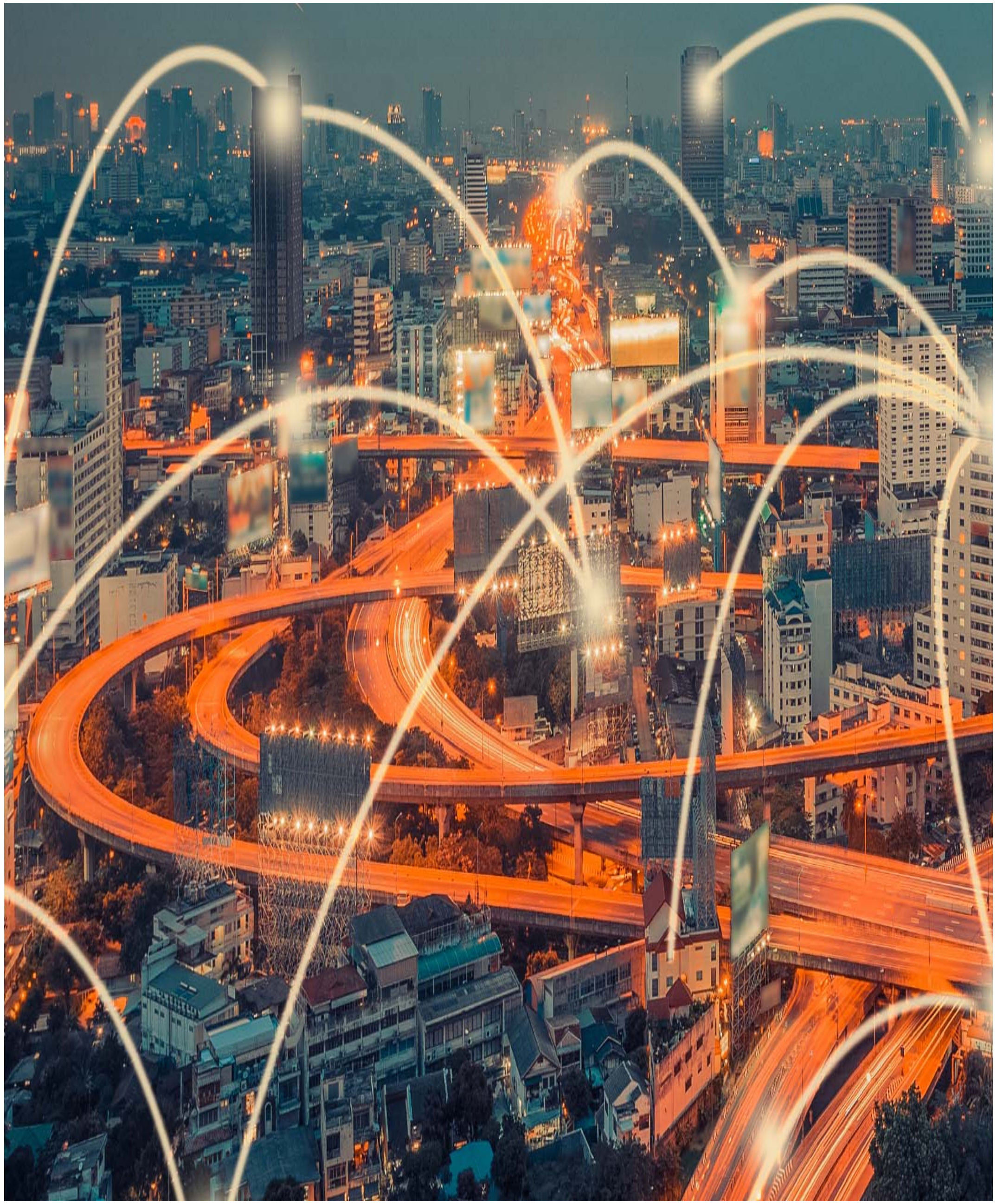
The University of Sydney NSW 2006 Australia

http://sydney.edu.au/business/itls


Keynote address at STRC May 15 – May 17, 2019, Monte Verita above Ascona, Switzerland.



# Heterogeneity in Decision Processes: Embedding Extremeness Aversion, Risk Attitude and Perceptual Conditioning in Multiple Heuristics Choice Making



# Focus of this Paper – Continuing exploration of role of process heuristics

- We have selected two behavioural heuristics to study jointly in explaining (through preference revelation) choice making:
  - Both of which reflect risk attitude in different ways, in a nonlinear additive form
  - where each heuristic contributes, up to a probability within a sampled population, the selection of a relevant multiple-heuristic utility expression, both within and between respondents.

- The jointly estimated model of interest (for a sample) herein is designed to account for:
  - Extremeness aversion as one form of risk-accommodating heuristic, and the
  - Fully compensatory attribute rule with risk attitude and perceptual conditioning as an extended expected utility attribute transformation (FC-EEUT) (as set out initially in Hensher *et al*. 2011).
- We compare the findings with separate stand alone models.
- We also make some comments on risk vs uncertainty and 'potential limitations' of stated choice settings in studying preferences under uncertainty – a challenge!

Hensher, D.A., Greene, W.H. and Li, Z. (2011) Embedding risk attitude and decisions weights in non-linear logit to accommodate time variability in the value of expected travel time savings, ***Transportation Research Part B*** 45, 954-972.

Hensher, D.A., Balbontin, C., and Collins, A. (2018) Heterogeneity in decision processes: embedding extremeness aversion, risk attitude and perceptual conditioning in multiple process rules choice making, (Paper presented at *The Fifth International Choice Modelling Conference* 3 – 5 April 2017, Cape Town, South Africa, April 2016, and the *Western Economic Association Annual Conference*, Newcastle Australia January 12 2018, special session chaired by Daniel L. McFadden), and the *International Transport Economics Association* (ITEA) Conference June 2018 Hong Kong, *Transportation Research Part A*, 111, 316-325.

# How to Capture Heterogeneity in Decision Processes

– A suggested alternative to the popular **latent class model (LCM)** form in studying process rules is to probability weight each heuristic or rule directly in a single utility function associated with each alternative in a choice set (e.g., Leong and Hensher 2012, 2012a)
  – Treating the model as a standard logit form such as MNL or MMNL (mixed logit).
  – Notes: LCMs are essentially a set of jointly estimated MNL (or MMNL) models.

– Within a single utility function, this approach allocates the proportional contribution of each process rule to overall utility (recognising we are dealing with a sample from a population)
  – with the possibility of linking the share outcome (HW) to the characteristics of respondents and other possible contextual influences (including beliefs, awareness etc.).
  – $U_i = HW_{H1} * U_{H1} + HW_{H2} * U_{H2} + \varepsilon_i$

– In a model with a total of $M$ heuristics or rules, the weights of each heuristic, denoted by $HW_m$, $m=1,2,\ldots,M$, can be given as a logistic function:

$$HW_m = \frac{\sum_i \exp(U_{m,i})}{\sum_{m=1}^{M} \sum_i \exp(U_{m,i})}$$

Leong, W., Hensher, D.A. (2012) Embedding multiple heuristics into choice models: An exploratory analysis, ***Journal of Choice Modelling***, 5, 131–144.

Leong, W. and Hensher, D.A. (2012a) Embedding decision heuristics in discrete choice models: a review, ***Transport Reviews***, 32 (3), 313-331. (ERA A)

Greene, W.H. and Hensher, D.A. (2003) A Latest Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit, ***Transportation Research Part B***, 37, 681-698.

# Some background: Contextual Concavity Model (CCM)

- One of the most robust empirical generalisations about human perception and decision making is **diminishing returns or sensitivity** (e.g., Meyer and Johnson 1995).
  - A specific feature is the mapping of objective attribute values onto psychological value expressed typically as a concave function (e.g., Thaler 1985; Tversky and Kahneman 1991).

- An extension is the compromise effect (relative to some setting) as a behavioural appealing idea that can be mathematically modelled by combining the notions of concavity and context dependence; in other words, via "**contextual concavity**."
  - Can include referencing, status quo etc.
  - Examples are studies by Hensher, Fosgerau, de Borger, De Palma…

- Within the DC setting, the deterministic component of utility of alternative *j* (for consumer *i*) equals the sum across attributes of concave functions of the MU gains between this alternative and the alternative with the minimum MU for each attribute *k* in the (local) choice set *S*.

- If concavity exists we have risk aversion behaviour.
- If convexity exists we have risk seeking behaviour.
- Otherwise risk neutral
- Note – this is separate from the issue of perceptual conditioning (see later)

# Compromise Models – A Class of Models

## A Nice Touch

– The CCM models are "general compromise" models in the sense that they can capture any form of compromise (or extremeness aversion);

– that is, compromise which is of either equal or different magnitude across attributes

– Very flexible – any number of attributes and alternatives

## Be aware

- We are taking only two interesting ‘process rules’, but any possibilities exist.
- In ongoing research, we are designing a study in which any number of relevant process rules are revealed that are used in making a choice
- We can then estimate a joint model of process and outcome  to represent heterogeneous choice making settings.
- Watch this space!

# Heuristic (Process Rule) 1:
## Extremeness Aversion

$$\text{U(current)} = \text{RefASC} \;+\left(\beta_t * (\text{time}_{\max} - \text{time})\right)^{\varphi_t} + \left(\beta_c * (\text{cost}_{\max} - \text{cost})\right)^{\varphi_c} + \varepsilon_0$$

$$\text{U(alt1)} = (1 + \beta_{tmsyr} * \text{TmsYr}) * \left(\beta_t * (\text{time}_{\max} - \text{time})\right)^{\varphi_t} + \left(\beta_c * (\text{cost}_{\max} - \text{cost})\right)^{\varphi_c} + \varepsilon_1$$

$$\text{U(alt2)} = (1 + \beta_{tmsyr} * \text{TmsYr}) * \left(\beta_t * (\text{time}_{\max} - \text{time})\right)^{\varphi_t} + \left(\beta_c * (\text{cost}_{\max} - \text{cost})\right)^{\varphi_c} + \varepsilon_2$$

In the model, we allow $\varphi_k$ (*phi*$_k$) to be freely estimated (the concavity parameter), and have chosen a form where the reference (i.e., "worst") attribute level, defined as the maximum of each of the time and cost components in the choice set, precedes the minus sign; hence the prior expectation is for $\hat{\beta}_k$ to be positive.

Sign of $\varphi_k$ : If the notion of *diminishing returns* in the contextual concavity model is accepted, the prior expectation is for $\varphi_k$ to satisfy the inequality $0 < \varphi_k < 1$

## ‘Process Rule’ 2: Recognising Risk Attitude and Perceptual Conditioning

Hensher, D.A., Greene, W.H. and Li, Z. (2011) Embedding Risk Attitude and Decisions Weights in Non-linear Logit to Accommodate Time Variability in the Value of the Expected Travel Time Savings, *Transportation Research Part B* 45, 954-972. (ERA A*)

Suggested replacing VoT and VoR with a single VETT

Li, Hao, Tu, H. and Hensher, D.A. (2016) Integrating the mean–variance and scheduling approaches to allow for schedule delay and trip time variability under uncertainty, *Transportation Research Part A*, 89, 151-163. (ERA A*)

Combing the two popular ways to value reliability

Li, B. and Hensher D.A. (2017) Risky weighting in choice analysis with risky prospects, *Transportation Research Part B*, 102, 1-21. (ERA A*)

A way to reveal functional form of DW without imposing *a priori*

THE UNIVERSITY OF SYDNEY

# Heuristic (Process Rule) 2:

## *Extended* Expected Utility (EEUT) Attribute form with Risk Attitude and Perceptual Conditioned Travel Time under the Fully Compensatory Paradigm

$$\text{U(currrent)} = \text{RefASC} + EEUT(U) + \sum_{z=1}^{Z} \beta_z S_z$$

$$\text{U(alt1,alt2)} = (1+\beta_{tmsyr} * TmsYr) * (EEUT(U) + \sum_{z=1}^{Z} \beta_z S_z)$$

$$EEUT(U) = \beta_x \{[W(P_1)\frac{x_1^{1-\alpha}}{1-\alpha} + W(P_2)\frac{x_2^{1-\alpha}}{1-\alpha} + \ldots + W(P_R)\frac{x_R^{1-\alpha}}{1-\alpha}]\}$$

$$w(P) = \exp(-(-\ln P)^{\gamma})$$

Prelec, D., 1998. The probability weighting function. Econometrica 66 (3), 497–527.

This form allow for over-weighting of low probabilities and under-weighting of high probabilities for $0 < \gamma < 1$. Allows for curvature but not elevation.

Assume CRRA form for risk attitude: Relative risk aversion (RRA) is a rate at which marginal utility decreases when wealth (herein x) is increased by one per cent. Arrow considered that RRA is likely to be constant or perhaps increasing. The parameter $\alpha$ measures the degree of RRA that is implicit in the utility function.

Concave utility function ($\alpha>0$): risk-averse attitude, i.e., a sure alternative is preferred to a risky alternative (i.e., with multiple possible outcomes) of equal expected value.
Convex utility function ( <0): risk taking attitude, i.e., a risky alternative is preferred to a sure alternative of equal expected value.

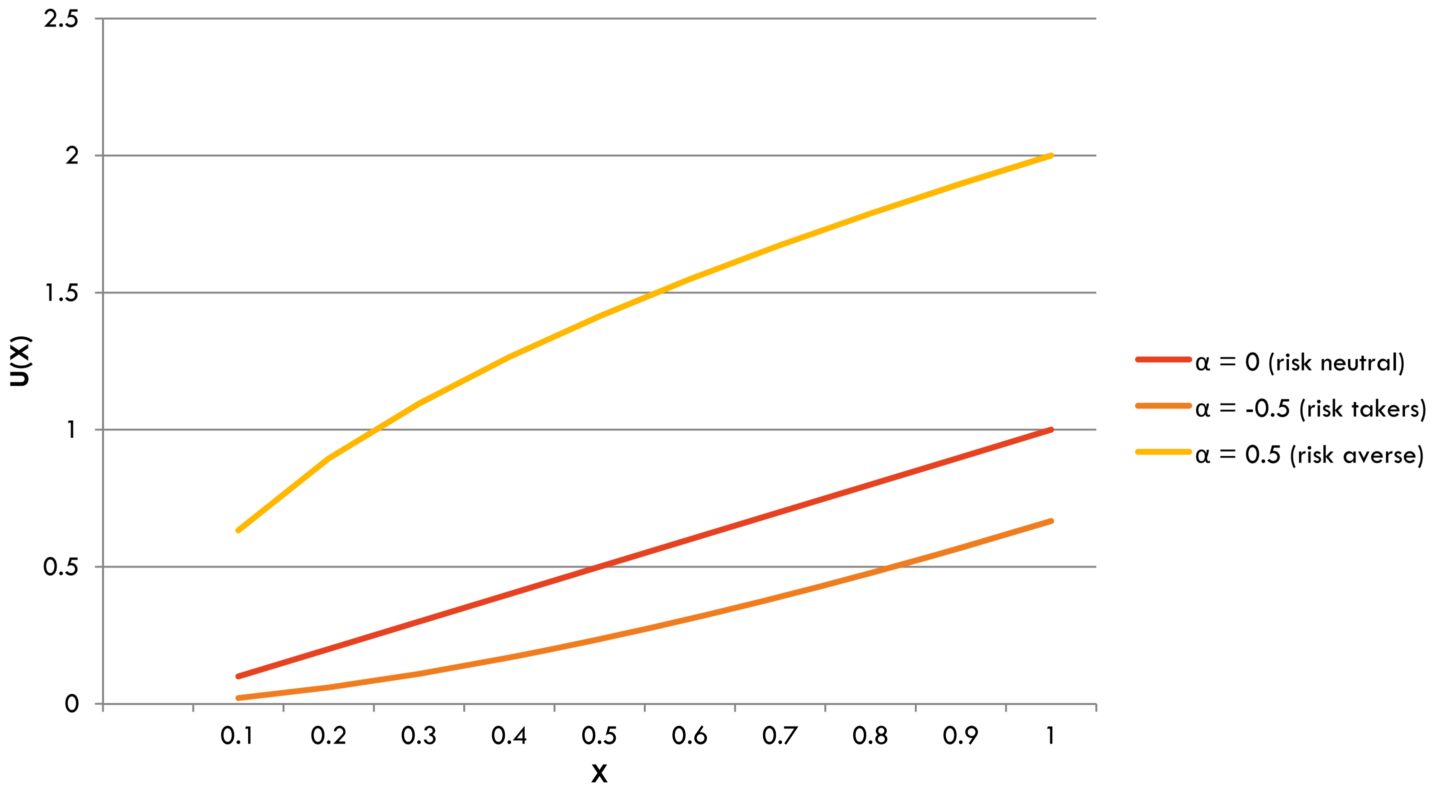
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–292.

# Recognising risk attitude

Implications of risk attitude on choice behaviour:

Risk Averse: U(sure)>U(risky)

Risk Taking: U(sure)<U(risky)

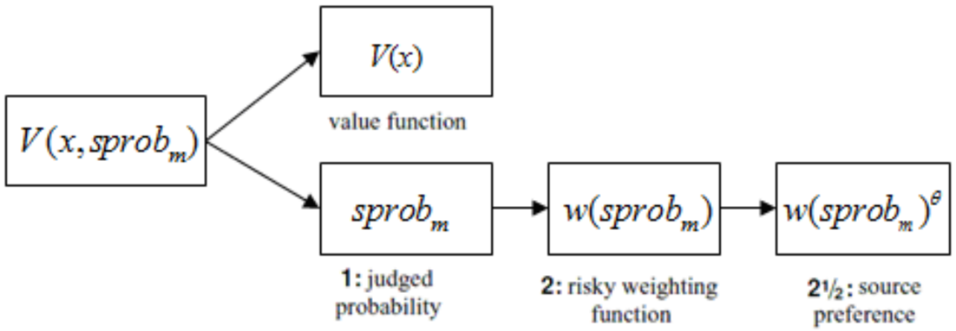


# Summary Table of Some Functional Forms of Perceptual Conditioning

| Name | Equation | References |
| --- | --- | --- |
| Linear form with discontinuous end points | Example: $w(p)=\begin{cases}(1-r)p & \text{if } p<1\\ 1 & \text{otherwise}\end{cases}$ | Loomes et al. (2002); (Abdellaoui et al., 2010) |
| Power | $w(p)=p^r$ | Simplification of the models: Goldstein and Einhorn (1987); Quiggin (1982); Tversky and Kahneman (1992) |
| Goldstein-Einhorn | $w(p)=\dfrac{sp^\gamma}{sp^\gamma+(1-p)^\gamma}$ | Goldstein and Einhorn (1987) |
| Tversky and Kahneman | $w(p)=\dfrac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{1/\gamma}}$ | Quiggin (1982); Tversky and Kahneman (1992); Camerer and Ho (1994); Tversky and Fox (1995) |
| Wu-Gonzalez | $w(p)=\dfrac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^s}$ | Wu and Gonzalez (1999) |
| Prelec I | $w(p)=\exp^{-(-\ln p)^\gamma}$ | Prelec (1998) |
| Prelec II | $w(p)=\exp^{-s(-\ln p)^\gamma}$ | Prelec (1998) |
| Exponential-power | $w(p)=\exp^{-\frac{r}{s}(1-p^s)}$ | Prelec (1998) |
| Hyperbolic-logarithm | $w(p)=(1-r\cdot\ln p)^{-s/r}$ | Prelec (1998) |

# A side issue: Risk and Uncertainty

- The focus of the model is on risk as well as ways to accommodate objective probabilities
  - In reality individuals form beliefs and opinions about likely travel time, referred to as subjective probability
  - Risk relates to a given or known probability of occurrence distribution.
- Potential issues with Stated Choice data compared to RP data
  - Assumes pre-defined attribute levels including occurrence probabilities (all studies using SC designs to value travel time variability or reliability with one exception we are aware of – ref below modified RP)
  - Focus mainly on statistical efficiency but less on behavioural relevance in settings of uncertainty
- Uncertainty ambiguity
  - Ellsberg paradox – when choice is made under uncertainty individuals have to assess the probs of potential outcomes with some degree of vagueness associated with their beliefs (i.e., subjective probabilities).
  - How well does perceptual conditioning on objective probs handle this?
  - We assume that DWs (i.e., the W(P) transformation) acts as a proxy to identify subjective probabilities
  - This gives us what psychologists refer to as a belief-based measure of outcomes.
  - In SC, I refer to it as 'equivalent subjective or belief adjusted attribute-specific outcome probabilities.
  - These parameterised transformations can be made a function of observed or reported contextual effects known as source preferences (Tversky and Fox) to explain and account for subjective probability (uncertainty) ambiguity.
  - **The process for modelling decision under uncertainty:**




Full details in: Hensher, D.A., Li., Z. and Ho, C. (2015) The role of source preference and subjective probability in valuing expected travel time savings, *Travel Behaviour and Society, 2, 42-54*.

# Four levels of subjectivity and objectivity in experiments

| | | |
|---|---|---|
| Level i | FO | = OPs+OAs |
| Level ii | PS(1) | = OPs+SAs |
| Level iii | PS(2) | = SPs+OAs |
| Level iv | FS | = SPs+SAs |

- FO: fully objective, PS(1): Partially subjective
- PS(2): Partially subjective, FS: fully subjective

- SPs: Subjective probabilities, SAs: Subjective attributes
- OPs : Objective probabilities, OAs: Objective attributes

# Joint Heuristics Model Summary

$$\mathrm{U}_{m1,current} = \left(\beta_t * (\mathrm{time}_{\max} - \mathrm{time})\right)^{\varphi_t} + \left(\beta_c * (\mathrm{cost}_{\max} - \mathrm{cost})\right)^{\varphi_c} + \varepsilon_0$$

$$\mathrm{U}_{m1,atl1} = \left(\beta_t * (\mathrm{time}_{\max} - \mathrm{time})\right)^{\varphi_t} + \left(\beta_c * (\mathrm{cost}_{\max} - \mathrm{cost})\right)^{\varphi_c} + \varepsilon_1$$

$$\mathrm{U}_{m1,alt2} = \left(\beta_t * (\mathrm{time}_{\max} - \mathrm{time})\right)^{\varphi_t} + \left(\beta_c * (\mathrm{cost}_{\max} - \mathrm{cost})\right)^{\varphi_c} + \varepsilon_2$$

$$U_{m2,i} = EEUT(U) + \sum_{z=1}^{Z} \beta_z S_z$$

The weights can be calculated as

$$HW_{m1} = \frac{(\exp^{U_{m1,1}} + \exp^{U_{m1,2}} + \exp^{U_{m1,3}})}{(\exp^{U_{m1,1}} + \exp^{U_{m1,2}} + \exp^{U_{m1,3}}) + (\exp^{U_{m2,1}} + \exp^{U_{m2,2}} + \exp^{U_{m2,3}})}$$

$$HW_{m2} = 1 - HW_{m1}$$

The total utility will be conditioned by the number of times the route has been used in the last year, as follows (the experience effect referred to by Dan) and treated as heteroscedastic conditioning

$$\mathrm{U(current)} = \mathrm{RefASC} + HW_{m1} * U_{m1,current} + HW_{m2} * U_{m2,current} + \varepsilon_0$$

$$\mathrm{U(alt1)} = (1 + \beta_{tmsyr} * TmsYr) * \left(HW_{m1} * U_{m1,alt1} + HW_{m2} * U_{m2,alt1} + \varepsilon_1\right)$$

$$\mathrm{U(alt2)} = (1 + \beta_{tmsyr} * TmsYr) * \left(HW_{m1} * U_{m1,alt2} + HW_{m2} * U_{m2,alt2} + \varepsilon_2\right)$$

# Recognising risk attitude and perceptual conditioning (*an aside given nature of data used*)

**Example of travel time variability valuation: State of practice**

- In reality, the travel time for the same repeated trip varies due to travel time variability
  - Which leads to a travel time distribution (time and probabilities of occurrence)

- In choice experiments (and RP data ideally), we can have multiple travel times per respondent alternative

- The mean-variance and the scheduling model are two dominant approaches

# Attribute Package in Road Context - Toll vs. Free Route Choice

Game 5

## Illustrative Choice Experiment Screen

Make your choice given the route features presented in this table, thank you.

| | Details of your recent trip | Route A | Route B |
|---|---|---|---|
| **Average travel time experienced** | | | |
| Time in free flow traffic (minutes) | 25 | 14 | 12 |
| Time slowed down by other traffic (minutes) | 20 | 18 | 20 |
| Time in stop/start/crawling traffic (minutes) | 35 | 26 | 20 |
| **Probability of travel time** | | | |
| 9 minutes quicker | 30% | 30% | 10% |
| As above | 30% | 50% | 50% |
| 6 minutes slower | 40% | 20% | 40% |
| **Trip costs** | | | |
| Running costs | $2.25 | $3.26 | $1.91 |
| Toll costs | $2.00 | $2.40 | $4.20 |
| If you make the same trip again, which route would you choose? | Current Road | Route A | Route B |
| If you could only choose between the two new routes, which route would you choose? | | Route A | Route B |

# Results: Accounted for multiple choice scenarios per respondent, $U_i = HW_{H1}*U_{H1}+HW_{H2}*U_{H2}+\varepsilon_i$

## Non-Linear in Parameters and Attributes

| Attribute | Joint Heuristics | Extremeness Aversion (EA) | FC-EEUT Risk Attitude and Perceptual Conditioning (RA_PC) |
|---|---|---|---|
| Reference alternative constant | 0.817 (26.61) | 0.91 (25.24) | 0.946 (28.81) |
| ***RA_PC*** | | | |
| Total cost ($) | -0.079 (4.69) | | -0.288 (26.8) |
| Gamma (PC) ($\gamma$)[1] | 1.39 (2.03) | | 1.08 (0.80) |
| Alpha (PC) ($\alpha$) | -0.183 (1.65) | | 0.167 (2.37) |
| Expected time (mins) | -0.017 (2.04) | | -0.117 (3.64) |
| ***EA*** | | | |
| Max time-time | 0.068 (16.35) | 0.056 (25.10) | |
| Max cost-cost | 1.32 (6.29) | 0.404 (13.53) | |
| Contextual concavity cost ($\varphi_c$) | 0.492 (14.56) | 0.551 (18.46) | |
| Contextual concavity time ($\varphi$t) | - | - | |
| ***Heteroscedastic conditioning*** | | | |
| Number of times recent trip is undertaken per annum | -0.002 (1.93) | 0.004 (2.67) | -0.001 (3.08) |
| **Log likelihood at zero** | | **-13,148.19** | |
| **Log-likelihood at convergence** | **-9,070.24** | **9,106.89** | **-9,189.231** |
| **AIC/sample size** | **1.509** | **1.515** | **1.528** |
| **McFadden pseudo $R^2$** | **0.310** | **0.307** | **0.301** |

# Results

– The results obtained from the joint model and the stand alone EA model show that all the $\hat{\varphi}$ (*phi*$_k$) for cost and $\beta$s are statistically significant at the 5% level (and many at the one percent level), with the $\hat{\beta}_s$ possessing the correct signs.

– Relative to the stand alone EA model, results from the Vuong (non-nested) test suggest that embedding a contextual heuristic into the joint model provides a better overall statistical fit at the 1% level of significance.

# Results – Extreme Aversion

– Concavity of the power parameter can be tested by comparing the null hypothesis (linear in the attributes) against the alternative hypothesis (concavity).

– We can reject the null hypothesis for **cost**, in favour of the alternative hypothesis (=0.492), at the five percent significance level.

  – This finding is consistent with a concavity power parameter. With all else equal, respondents are **extremeness averse** when evaluating the cost attribute. (Playing conservatively!)

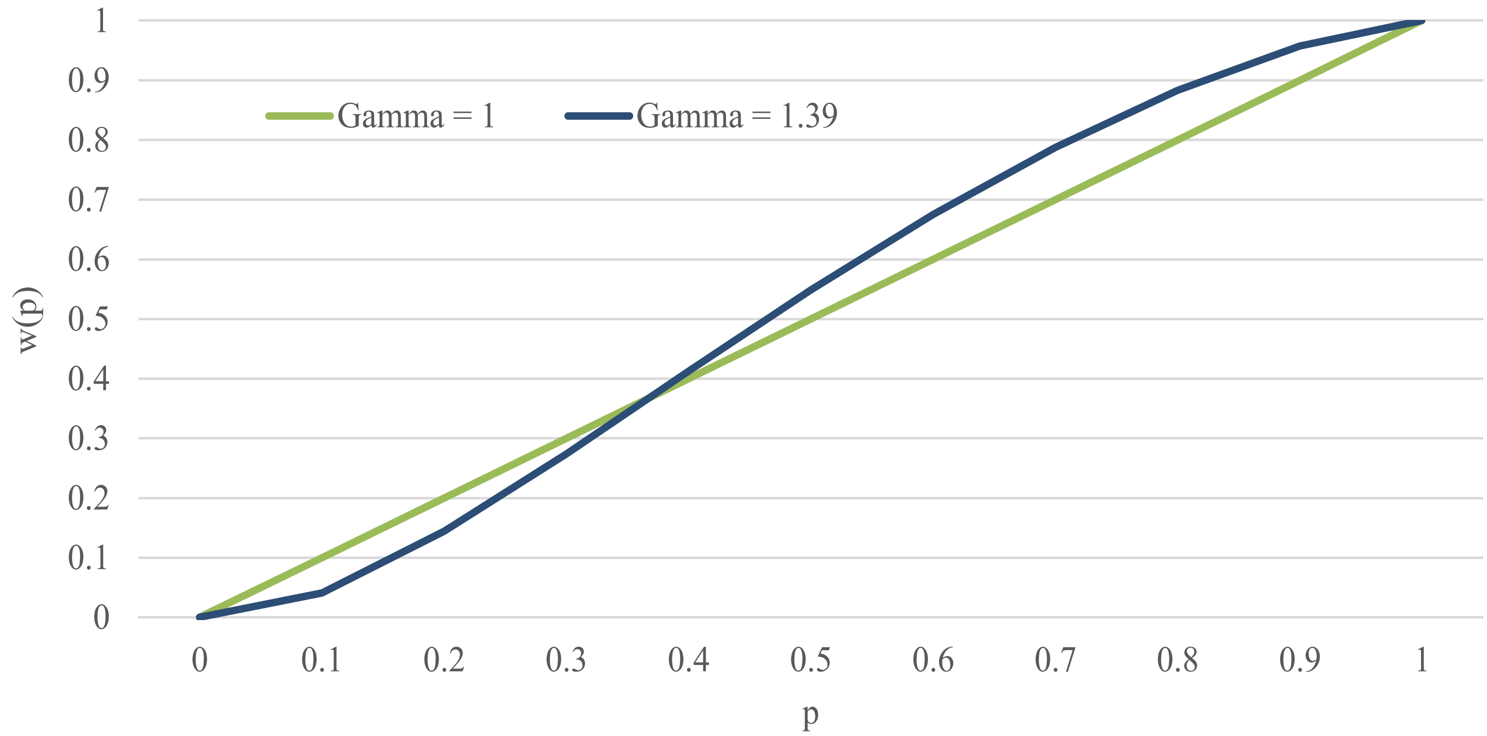
– For **travel time**, the parameter is not statistically significant; hence we have a linear effect (risk neutral).

## Results FC-EEUT

- Relative to the stand alone FC-EEUT model with risk attitude and decisions weights,

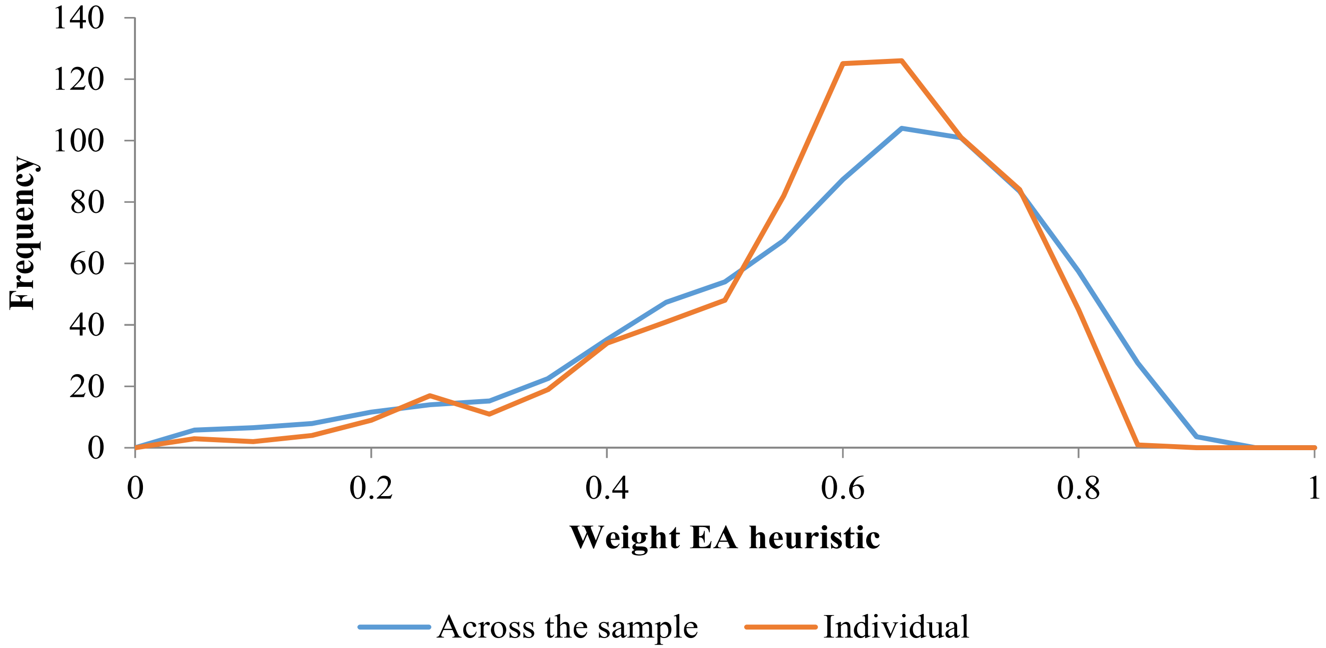
- Results from the Voung test suggest that embedding the FC-EEUT model heuristic into the joint model provides a better statistical fit at the 1% level.

# Non-linear probability weighting function in the joint model (overweighting at low probs….)



# Distribution of probability of process rule contributions to overall utility of an alternative (H1 is EA)

Distribution of probability of EA process rule contribution to overall utility of all alternatives at a choice set level (across the sample) and at the set of choice sets (at an individual respondent level) (Bin width is 0.05)




The extremeness aversion (seeking) heuristic (H1) has, on average, a 0.568 probability of relevance compared to a 0.432 probability of relevance for the FC-EEUT process rule (H2). At respondent level the probabilities sum to 1.0

# WTP: The Value of Travel Time Savings ($/person hour)

**Heuristic 1: Extremeness Aversion**

$$\frac{\partial V}{\partial T} = \beta_t \qquad \frac{\partial V}{\partial C} = -\beta_c^{\varphi_c} \varphi_c (\mathrm{C}_{\max} - C_j)^{\varphi_c - 1}$$

**Heuristic (Process Rule) 2: Extended Expected Utility Attribute form with Risk Attitude and Perceptual Conditioned Travel Time under the Fully Compensatory Paradigm**

$$\frac{\partial V}{\partial T_k} = \sum_{k=1}^{3} [\exp(-(\log(\mathrm{ProbT}_k))^{\gamma})] * [-\beta_t (\mathrm{T}_k)^{(-\alpha)}] \qquad \frac{\partial V}{\partial C} = -\beta_c$$

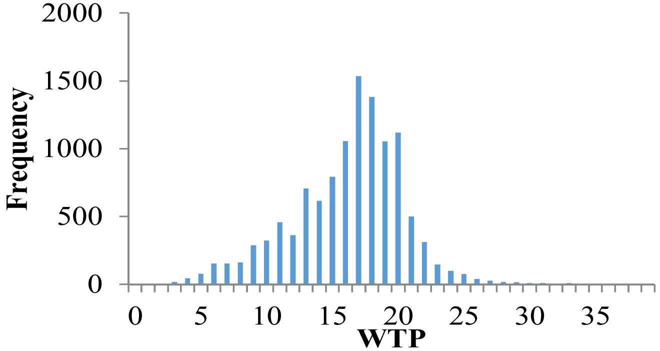
where $k$ = 1 (early), 2 (on time), and 3 (late)

| | Joint Heuristics | Extremeness Aversion (seeking) (EA(S)) | FC-EEUT Risk Attitude and Perceptual Conditioning (RA_PC) |
|---|---|---|---|
| | 17.73 (3.43) | 15.83 (4.58) | 13.38 (1.09) |
| **EA(S) (H1)** | 6.97 (3.10) | | |
| **RA_PC (H2)** | 10.93 (5.18) | | |

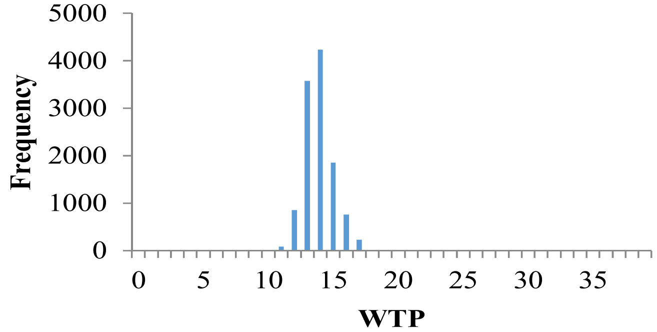
*Standard deviation estimates are in brackets

**These mean differences are HUGE**

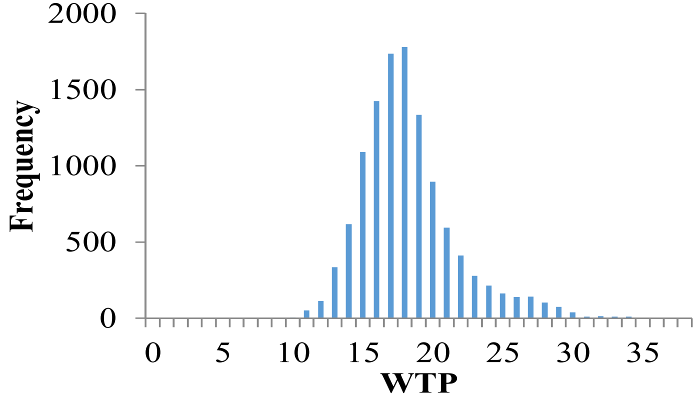
# Histograms of the distribution of VTTS across the sample




EA(S) model




FC-EEUT RA_PC model




Joint Model

# Concluding Comment – Food for Thought!

- **Is this a potentially more behaviourally appealing way of capturing notable preference heterogeneity than random parameters?**

- We explore this point in Balbontin et al. 2017 and Balbontin et al. 2018
- Taste heterogeneity in standard models may in fact be heterogeneity in decision rules?
  - One might improve on the behavioural interpretation of preference heterogeneity by conditioning identification of random marginal (dis)utility with a systematic conditioning through an attribute processing rule. (Balbontin et al used value learning VL)
  - The evidence shows that this approach is appropriate, and that there is a significant systematic relationship between the mean and Std Dev of random parameters and VL, which influences the WTP estimates and distribution.
  - The inclusion of process heuristic conditioning hence appears to add behaviourally relevant information that changes the distribution of preference heterogeneity.
  - There is, however, no clear pattern of a reduction or increase in the standard deviation associated with willingness to pay estimates when incorporating the VL process heuristic in the mean or standard deviations

Balbontin, C., Hensher, D.A., Collins, A.T., (2017) Is there a systematic relationship between random parameters and process heuristics? (Paper Presented in The Fifth International Choice Modelling Conference 3 – 5 April 2017, Cape Town, South Africa) ***Transportation Research Part E***, 106, 160-177. DOI: 10.1016/j.tre.2017.07.013.

Balbontin, C., Hensher, D.A. and Collins, A.T. (2018) Process homogeneity, process heterogeneity, and preference heterogeneity: How to better represent decision making and preferences, for IATBR 2018, Santa Barbara, California, July 2018; Interdisciplinary Choice Workshop (ICW), Santiago de Chile, 7-10 August 2018, ***Transportation Research Part B, 122,*** 2019.

# On Going Research: Conditioning of Random Process Heterogeneity (CRPH)

- The relationship between process and preference heterogeneity - Conditioning of Random Process Heterogeneity (CRPH).
- The approach recognises that the parameters defined under LPAA may be conditioned by a process strategy.

The random parameter specification, as was analysed previously, decomposes parameter $\theta_q$ in its mean, $\theta^m$, and standard deviation, $\sigma \cdot v$:

$U_{iqt} = \left(\theta + \sigma \cdot v\right) \cdot X_{iqt} + \varepsilon_{iqt}$

To incorporate process heuristics using the CRPH approach, the mean and standard deviation of attribute $x_{inqt}$ under an LPAA mixed logit model, are written as a function of each of the process heuristics. The utility expression can be written as follows:

$$U_i = \sum_n \left( \begin{bmatrix} \theta_{in} + \lambda^m_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^m_{RAM,in} \cdot RAM\left(x_{inqt}\right) \\ + \left[\sigma_{in} + \lambda^s_{VL,in} \cdot VL\left(x_{inqt}\right) + \lambda^s_{RAM,in} \cdot RAM\left(x_{inqt}\right)\right] \cdot v \end{bmatrix} \cdot x_{inqt} \right) + \varepsilon_{iqt}$$

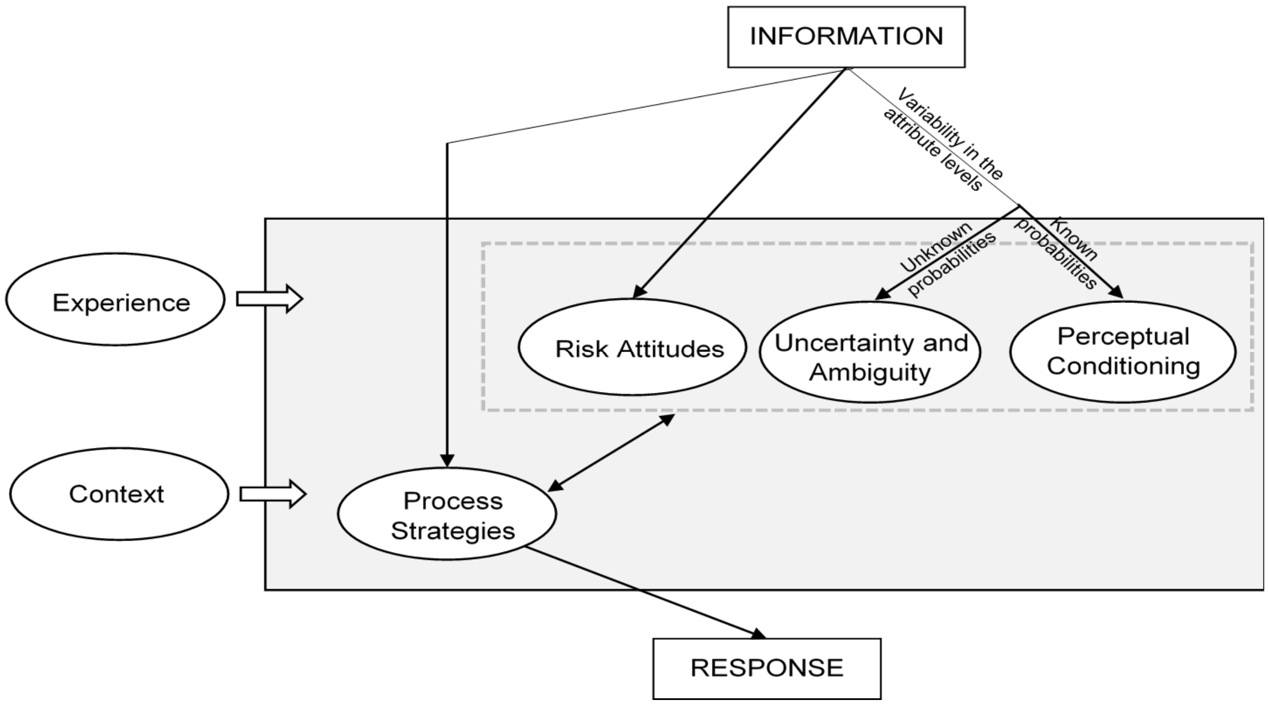
where $VL\left(x_{inqt}\right)$ represents the transformation of $x_{inqt}$ for the VL heuristic; $RAM\left(x_{inqt}\right)$ for the RAM heuristic; $\lambda^m_{VL,in}$ represents the relationship between the mean estimate and VL; $\lambda^m_{RAM,in}$ represents the relationship between the mean estimate and RAM; $\lambda^s_{VL,inq}$ is the relationship between the standard deviation estimate of the random parameter distribution and VL; and $\lambda^s_{RAM,inq}$ is the relationship between the standard deviation and RAM.

Balbontin, C., Hensher, D.A. and Collins, A.T. How to better represent preferences in choice models: the contributions to preference heterogeneity attributable to the presence of process heterogeneity, ***Transportation Research Part B***, 2019.

# On going Research

- The preferred preference revelation model form, *conditioning of random process heterogeneity* (CRPH), supports a behavioural paradigm in which individuals use more than one process heuristic in decision making, supporting heterogeneity in processing information related to alternatives on offer. The impact on important behavioural outputs such as willingness to pay is profound, and has important policy relevance in project appraisal.

- Proposed conceptual framework for decision making



# Some Comments on behaviourally more realistic Models

- Trade off between behavioural relevance and economic theory based on social welfare analysis
- ‘Irrational’ behaviour is otherwise captured by the existence of the error term in RUM models
- The Process heuristics are more aligned with behavioural decision theory than with strict axioms of economic rationality.
- Useful papers: Dekker, T and Chorus, C. (2018) Consumer surplus for random regret minimisation models, *Journal of Environmental Economics and Policy*, 7(3), 269-286, and McConnell, K. E. (1995) Consumer Surplus for Discrete Choice Models. Journal of Environmental Economics and Management 29 (3): 263–270.
- Choice probabilities interpreted ‘as if they were’ probabilistic demand functions so can derive Marshallian consumer surplus (linked to WTP).

- **Hensher, D.A. Context dependent process heuristics and choice analysis: a note on the behavioural setting guiding the research focus, January 2019**.

- Assume no income effect (reasonable in many transport applications) and so path dependency issue resolved.
- We focus on valuing changes in attributes of a single alternative but recognise that they are context dependent.
- In our work we have tested for starting values, covariance matrix effects (essentially correlations between process heuristics) and implications on WTP to conclude that model is identified and not confounded.
- The welfare effect is:

$$\mathrm{CV} = [\ln(1-\pi_1^0) - \ln(1-\pi_1^*)]/\beta$$

- This assumes that choice is independent of income. What we have here is a simple comparison of the choice probability before and after a change in the level of an attribute. McConnell shows that this is equivalent to the Marshallian change in consumer surplus as the area under a demand curve.

# Ken Train's Position

## 2.3 Derivation of Choice Probabilities

"Discrete choice models are usually derived under an assumption of utility-maximizing behavior by the decision maker. Thurstone (1927) originally developed the concepts in terms of psychological stimuli, leading to a binary probit model of whether respondents can differentiate the level of stimulus. Marschak (] 960) interpreted the stimuli as utility and provided a derivation from utility maximization. Following Marschak, models that can be derived in this way are called random utility models (RUMs).

It is important to note, however, that models derived from utility maximization can also be used to represent decision making that does not entail utility maximization. The derivation assures that the model is consistent with utility maximization; it does not preclude the model from being consistent with other forms of behavior. The models can also be seen as simply describing the relation of explanatory variables to the outcome of a choice, without reference to exactly how the choice is made."

# Experience Conditioning

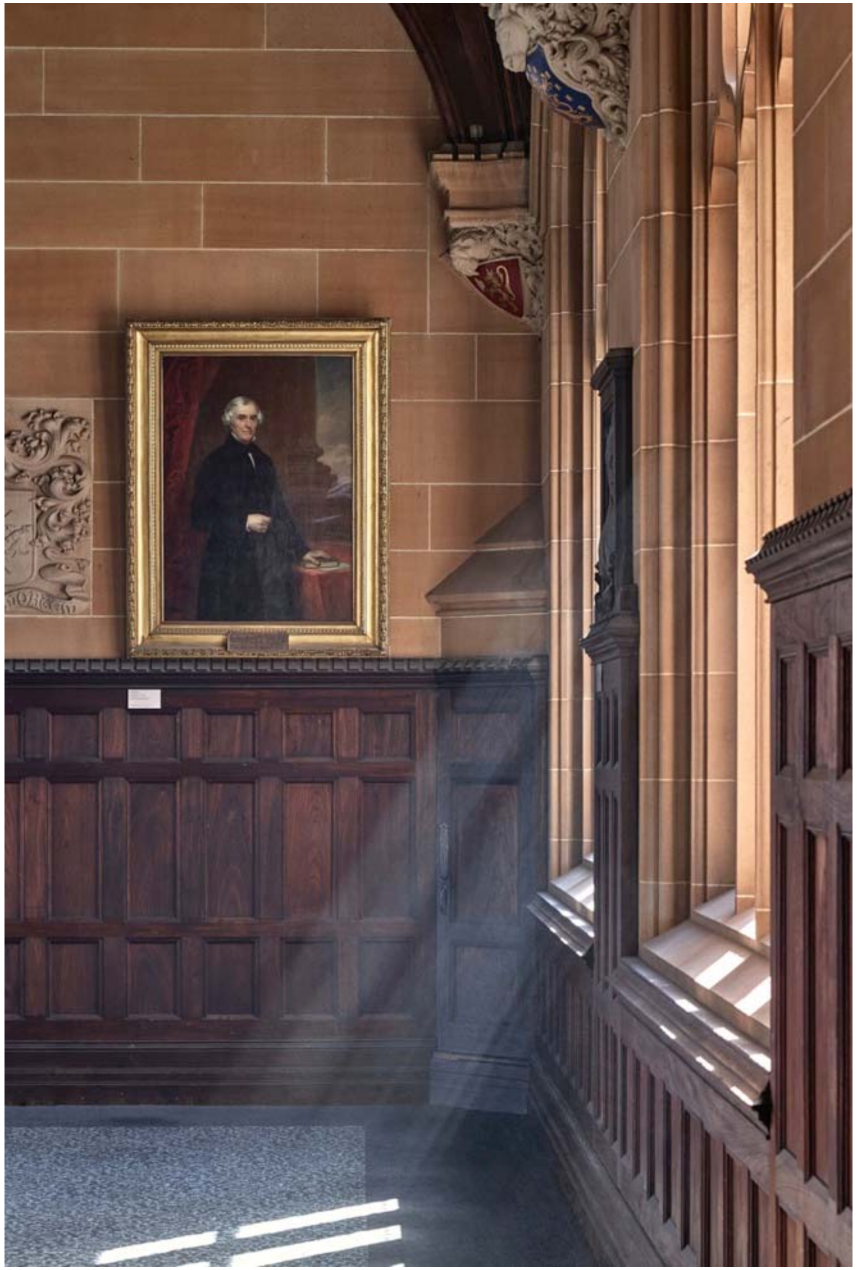

Hensher, D.A. and Ho., C. (2016) Experience conditioning in mode choice modelling – does it make a difference? *Transportation Research Part E*, 95, 164-176.

Balbontin, C., Hensher, D.A. and Collins, A.T. How to better represent preferences in choice models: the contributions to preference heterogeneity attributable to the presence of process heterogeneity, *Transportation Research Part B,* 24 August 2018, revised 4 February 2019.

Hensher, D.A., Greene, W.G. and Balbontin, C. Experience as a conditioning effect on choice – does it matter whether it is exogenous or endogenous?, for the *Sixth International Choice Modelling Conference* in Kobe, Japan, August 2019.

Buckell, J., Hensher, D.A. and Hess, S. Capturing the role of addiction in smokers' choices: an addiction-conditioned, hybrid choice model approach applied to smokers in the US, abstract prepared for the *2109 World Congress on Health Economics* and *Sixth International Choice Modelling Conference* in Kobe, Japan, August 2019.


THE UNIVERSITY OF SYDNEY



# Role of Experience

– Beginning with the standard utility expression associated with the $j^{th}$ alternative contained in a choice set of $j = 1,\ldots,J$ alternatives, we assume that an index defining overt experience with the $j^{th}$ alternative and $q^{th}$ individual, referred to as $E_{qi}$, conditions the utility expression. The functional form can be denoted by:

$$U_{qi}^{*} = \mu(E_q)U_{qi} = \mu(E_q)(V_{qi} + \varepsilon_{qi})$$

where $U_{qi}^{*}$ is the standard utility expression, $U_{qi}$, conditioned on the overt experience (and other possible influences) with an alternative. This conditioning is a form of heteroscedasticity. $E_q$ recognises that individual-specific experience, proxied by some metric such as frequency of use, conditions the marginal (dis)utility of each and every attribute, observed and unobserved, associated with the $j^{th}$ alternative in a pre-defined choice set. Therandom variables $\mu(E_q)\varepsilon_j$, for all $q$ and $j$ contained in an individual's choice set are IID Gumbel but with scale factors $\mu(E_q)$ that can vary as required across the sample. Dividing both the left and right hand by $\mu(E_q) > 0$ produces the standard basis of the RUM model. The probability behind random utility maximisation is unchanged by the positive scale factor, as shown in next slide.

# Heteroscedastic ConditioOning (linked to Scale)

$$\begin{aligned}\Pr[U^*_{qj} > U^*_{qj'}] &= \Pr[U_{qj} > U_{qj'}] \\ &= \Pr[V_{qj} - V_{qj'} > \varepsilon_{qj} - \varepsilon_{qj'}] \\ &= \Pr[\mu(E_q)(V_{qj} - V_{qj'}) \geq \mu(E_q)(\varepsilon_{qj} - \varepsilon_{qj'})]\end{aligned}$$

Given the IID property of the error difference, it follows that the probability of choosing an alternative is an MNL-like model with the observed sources of utility $\mu(E_q)V_{qj}$ as:

$$\Pr_{qj} = \frac{\exp\left[\mu(E_q \mid \gamma_j)\cdot V_{qj}(X_{qj} \mid \beta)\right]}{\sum_{j' \in J_q} \exp\left[\mu(E_q \mid \gamma_{j'})\cdot V_{qj'}(X_{qj'} \mid \beta)\right]}$$

where we have parameters $\gamma_j$ and $\beta$, and the observed variables $E$ and $X$ associated with each alternative and each individual. By making the parameters in the scale function vary across the alternatives (for identification), we have transformed the MNL model to one in which the utility functions are nonlinear in the parameters.

# Specific Functional Form for HC

– The specific functional form of heteroscedastic conditioning chosen is:

$$E_{q,car} = \exp\left(\gamma_{car} \ln(FR_{q,car} + 1)\right)$$

$$E_{q,PT} = \exp\left(\gamma_{PT} \ln(FR_{q,PT} + 1)\right)$$

where $FR_{q,j}$ is usage frequency, defined by the number of times the $q^{th}$ individual used mode $j$ ($j$=car or PT) over the last two months, and $\gamma_j$ ($j$ = car, PT) are parameters to be estimated. Socio-economic variables such as age and income can also be included to recognise the residual heterogeneity effect after individual experience has been accounted for.

When the experience effects are not significant (identified by $\gamma_j$ being not statistically different from zero), the experience functions receive the value of one, and hence, the choice model collapses to the standard utility expression.

## Experience and HC

– In summary, when we allow for this form of heteroscedasticity, the standard logit model takes the form shown below where $V_{qj}$ is linear-in-parameters.

$$\Pr_{jq} = \frac{\exp[E_{q,j} \times V_{q,j}]}{\sum_{j'=1}^{J} \exp[E_{q,j'} \times V_{qj'}]} = \frac{\exp\left[\exp\left(\gamma_j \ln(FR_{q,j}+1)\right) \times \beta' X_{q,j}\right]}{\sum_{j'=1}^{J} \exp\left[\exp\left(\gamma_{j'} \ln(FR_{q,j'}+1)\right) \times \beta' X_{qj'}\right]}$$

This model is non-linear-in-parameters since the parameter associated with the experience effect ($\gamma_j$) interacts with the parameters $\beta$ associated with attributes $X_{qj}$.

# Experience as a conditioning effect on choice – Does it matter whether it is exogenous or endogenous?

- There are a number of ways to set up a discrete choice model that embeds the presence of endogeneity associated with a specific inclusion in the representative component of a utility expression.
- Some key papers on this topic are applications by Train and Wilson (2009), Petrin and Train (2010), Guevara, and Ben Akiva (2010) and Guevara and Hess (2019) and a mainstream econometric review by Wooldridge (2015).
- Our research focuses on the potential endogeneity induced by including accumulated experience in using each of the modes in a mode choice application.
  - As set out above, experience, proxied by exogenous frequency of use has similar features to a reference alternative that is fixed and hence it might not induce endogeneity, but is worthy of consideration.

# Control Functions

- Control functions are statistical methods to correct for endogeneity problems by modelling the endogeneity in the relevant random components.
- The approach differs in important ways from other models that try to account for the same econometric problem. Instrumental variables, for example, attempt to model an endogenous variable $X$ as an often invertible model with respect to a relevant and exogenous instrument $Z$. Panel data use special data properties to difference out unobserved heterogeneity that is assumed to be fixed over time.
- Control functions were introduced by Heckman and Robb (1985a), although the principle can be traced back to earlier papers such as Heckman (1979). A particular reason why they are popular is because they work for non-invertible models (such as discrete choice models) and allow for heterogeneous effects, where effects at the individual level can differ from effects at the aggregate.
- Classic examples using the control function approach is the Heckit model and the Heckman (1979) correction.

# Control Functions

- It involves two stages. First, the endogenous variable is regressed on exogenous instruments; then, the residual (or a function of it) is incorporated into the utility function as an additional explanatory variable called the control function.
- A general advantage of the control function approach is that the test that the coefficient on the CF is zero is broadly equivalent to a test of exogeneity.

- Multiple Indicator Solution (MIS) method:
  - We extend Hensher and Ho (2017) by conditioning, at the first stage, the *entire* utility expression associated with all attributes in a utility expression, on the prior experience with an alternative.
  - This captures possible correlates associated with *each and every attribute* and not just one selected attribute (i.e., crowding in Guevara et al. 2019).
  - The second stage implemented is the control function method.
  - Standard errors for the two stage approach are computed using bootstrapping (Karaca-Mandic and Train, K. 2003 and Hensher et al. 2015).

# Application Example

## Revealed Preference Mode Choice

– The data collected to test the proposition on experience was obtained from an online survey undertaken in March 2014, using a sample of car and public transport commuters in the Sydney metropolitan area.

– Respondents were asked to report three perceived travel times and the likelihood of experiencing each travel time. The survey also included questions relating to travel cost, fuel consumption of the car used for travel, number of times a car and public transport were used for travel in the last two months, as well as socio-economic characteristics such as age, income, occupation and household car ownership.

# Main Questions

THE UNIVERSITY OF SYDNEY

## Travel Time Uncertainty Survey

### Car Commuter

You are qualified for this survey.

You said that you commute by **car as driver** and we know that your travel times will vary each time you travel to work for many reasons such as congestion, accidents, breakdowns, and road works.

Could you please share with us **3 possible travel outcomes** that you believe could occur on your **car trip to work**? *(these could be based on your recent experience on your regular commuter trip, or your perceptions of what it is likely to be.)*

The 3 possible travel outcomes must include: **one with the longest travel time, one with the shortest travel time and one with the most likely travel times.**

The likelihood of the most likely travel time must be highest amongst 3 possible outcomes and the likelihoods across 3 possible outcomes must add up to 100.

| | Outcome 1 | Outcome 2 | Outcome 3 |
|---|---|---|---|
| Door-to-door travel time (minutes) | | | |
| Travel distance by car (km) | | | |
| Toll paid ($) | | | |
| The likelihood of this outcome actually occurring (%) | | | |
| Rank possible travel outcomes from 1 (most preferred) to 3 (least preferred) | | | |

What is the average fuel consumption of the car that you would use for commuting? litres/100km

How many times in the last 2 months did you drive to work?

You said that you could use **public transport** to commute if you wanted to.

What would be the **3 possible outcomes** of your commuting trip by **public transport**? *(use the same principles provided above for car commuting trip to provide your answers)*

| | Outcome 1 | Outcome 2 | Outcome 3 |
|---|---|---|---|
| Door-to-door travel time (minutes) | | | |
| Fare ($) | | | |
| The likelihood of this outcome actually occurring (%) | | | |
| Rank possible travel outcomes from 1 (most preferred) to 3 (least preferred) | | | |

How many times in the last 2 months did you ride public transport to work?

Next

# Choice Models

– ***Experience Conditioning:***

ECondcr=1+tau1*log(ca2month)+agec*age

ECondpt=1+tau2*log(pt2month)

***M1:***

$U_{PT}$ = Econdpt*(ptfasc+betattp*avtim +betac*cost1h)

$U_{car}$ = Econdcr*(betattc*avtim+betac*cost1h)

***M2:***

$U_{PT}$ = Econdpt*(ptfasc+betattp*avtim +betac*cost1h) +**ptres*pt2res**

$U_{car}$ = Econdcr*(betattc*avtim+betac*cost1h) +**crres*ca2res**

# Control Functions (Ordinary Least Squares Regression)

| | Car experience (use per month) | Public transport experience (use per month) |
|---|---|---|
| Constant | 9.384 916.8) | 2.597 (7.22) |
| Personal income | -0.0047 (-2.28) | |
| Number of adults | 0.1370 (1.29) | 0.3685 (3.16) |
| Commuter trip (1,0) | -9.2806 (-47.9) | 11.279 (48.4) |
| Average travel time (mins) | -0.0039 (-3.18) | -0.0040 (-2.67) |
| Ages of respondent (years) | 0.0221 (3.39) | |
| Hours worked per week | 0.0612 (7.10) | |
| Male (1,0) | | -1.421 (-6.17) |
| Age less than 15 years (1,0) | | 0.1163 (0.30) |
| Age 25 to 34 years (1,0) | | -1.5569 (-4.93) |
| Age 35 to 44 years (1,0) | | 0.9759 (-3.34) |
| Age 55 to 64 years (1,0) | | -0.6026 (-2.26) |
| Dependent variable mean and Std Dev | 7.63 (5.86) | 6.75 (7.11) |
| Sample size | 1518 | |
| Adjusted R-squared | 0.674 | 0.617 |

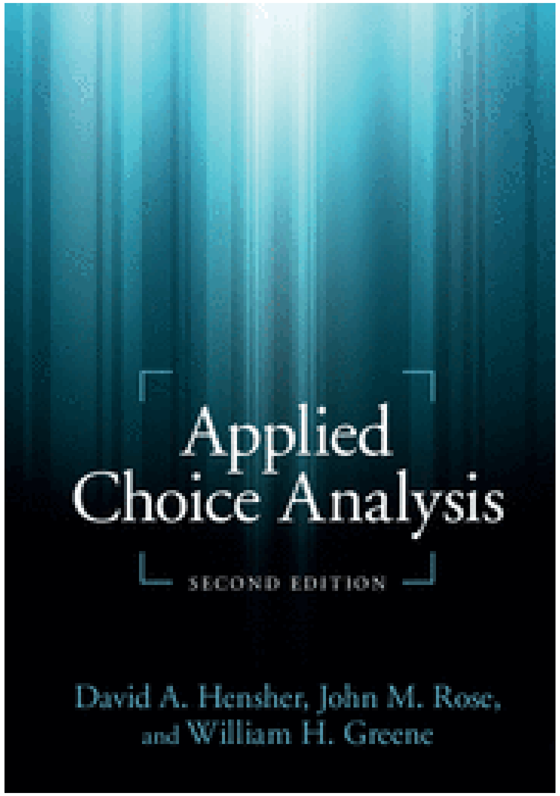
## Models with and without inclusion of explanatory variables to assess the presence or absence of endogeneity (z values in brackets)

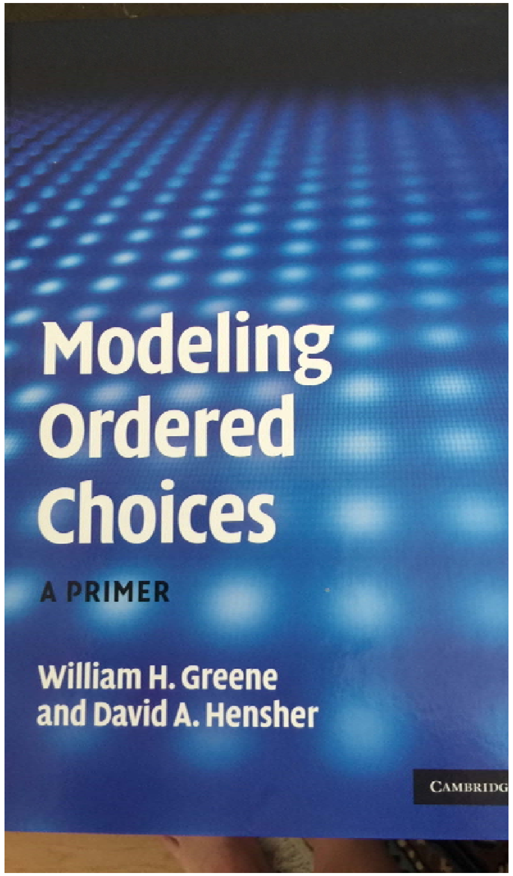
| | Model 1 (M1) | Model 2 (M2) | Bootstrapping M2 |
|---|---|---|---|
| **Mean of random parameters:** | | | |
| **Tau1** | -0.2486 (-4.98) | -0.2804 (-4.56) | -0.2804 (-5.54) |
| **Tau2** | -0.2172 (-7.86) | -0.3314 (-4.31) | -0.3314 (-14.07) |
| **Car residual** | | -4.9047 (-2.32) | -4.9047 (-6.64) |
| **Public Transport residual** | | -4.9912 (-2.45) | -4.9912 (-6.28) |
| **Non-random parameters:** | | | |
| **PT constant** | -2.3671 (-2.43) | -36.31 (-2.58) | -36.31 (-7.72) |
| **Average trip cost ($)** | -0.6265 (-4.40) | -3.2968 (-2.02) | -3.2968 (-2.10) |
| **Average travel time PT (mins)** | -0.01502 (-0.97) | -0.0254 (-0.29) | -0.0254 (-0.70) |
| **Average travel time Car (mins)** | -0.0782 (-7.22) | -0.7075 (-3.26) | -0.7075 (-4.11) |
| **Age of individual** | -0.0004 (-0.09) | -0.0041 (-1.07) | -0.0041 (-1.28) |
| **Std Dev of random parameters:** | | | |
| **Tau1 (constrained triangular)** | -0.2486 (-4.98) | 0.2804 (4.56) | 0.2804 (5.54) |
| **Tau2 (constrained triangular)** | -0.2172 (-7.86) | 0.3314 (4.31) | 0.3314 (14.07) |
| **Car residual** | | 12.598 (6.40) | 12.598 (2.61) |
| **Public Transport residual** | | 6.6403 (3.19) | 6.6403 (4.03) |
| **Goodness-of Fit:** | | | |
| **Log-likelihood at convergence** | -117.43 | -95.68 | |
| **McFadden pseudo $R^2$** | 0.777 | 0.818 | |
| **Sample size** | 759 | | |
| **AIC/N** | 0.328 | 0.281 | |

# Value of Travel ($/person hour) and Direct Choice Probability Elasticity Results (standard errors in brackets) * Bootstrapping took 6 hours to run.

| | Model 1 (M1) | Model 2 (M2) | Bootstrapping M2 (20 repetitions) t-values* |
|---|---|---|---|
| Car Value of travel time ($/person hr) | 7.49 (2.38) | 12.87 (1.74) | 3.03 |
| PT Value of travel time ($/person hr | 1.44 (1.71) | 0.66 (0.26) | 0.79 |
| Choice Probability Direct Elasticity Estimates: | | | |
| PT average travel time | -0.67 (0.039) | -0.421 (1.8) | 2.11 |
| Car average travel time | -2.32 (0.176) | -5.55 (2.5) | 2.20 |
| PT average travel cost | -1.58 (0.118) | -3.55 (2.80) | 1.92 |
| Car average travel cost | -1.084 (0.117) | -1.67 (1.41) | 1.47 |

# Commercial Break: Our Software and Books on Choice Experiments and Discrete Choice Modelling




http://www.limdep.com/

http://www.choice-metrics.com/download.html

http://sydney.edu.au/business/itls/courses/choice_analysis

# THANK YOU

**David Hensher PhD FASSA**
Institute of Transport and Logistics Studies (ITLS)
The University of Sydney Business School

**THE UNIVERSITY OF SYDNEY**
**E** david.hensher@sydney.edu.au | **W** http://sydney.edu.au/business/itls

# On Going Research: Conditioning of Random Process Heterogeneity (CRPH)

The parameters $\theta_{in}$ can be considered common between LPAA and RAM, but not for VL, which will have its own parameters $\theta_{in}^{VL}$. The transformations of $x_{inqt}$ associated with VL and RAM are as follows:

$$VL\left(x_{inqt}\right)=\theta_{in}^{VL}\cdot\left(x_{inqt}-ref_n\right)^{\varphi} \qquad RAM\left(x_{inqt}\right)=\theta_{in}\cdot x_{inqt}+\sum_{j\in S}R(i,j)$$

The expression for CRPH that includes VL and RAM results in the following form:

$$U_i=\sum_n\left(\left[\begin{array}{l}\theta_{in}+\lambda_{VL,in}^{'m}\cdot\left(x_{inqt}-ref_n\right)^{\varphi}+\lambda_{RAM,in}^{m}\cdot\left(\theta_{in}\cdot x_{inqt}+\sum_{j\in S}R(i,j)\right)\\+\left[\sigma_{in}+\lambda_{VL,in}^{'s}\cdot\left(x_{inqt}-ref_n\right)^{\varphi}+\lambda_{RAM,in}^{s}\cdot\left(\theta_{in}\cdot x_{inqt}+\sum_{j\in S}R(i,j)\right)\right]\cdot v\end{array}\right]\cdot x_{inqt}\right)+\varepsilon_{iqt}$$

$$\lambda_{VL,in}^{'}=\lambda_{VL,in}\cdot\theta_{in}^{VL}$$

This form also allows process strategies to have an influence over the mean but not the standard deviation of an attribute, with $\lambda_{in}^{m}=0$ and $\lambda_{in}^{s}\neq 0$ or, oppositely, over its standard deviation but not over its mean, with $\lambda_{in}^{m}\neq 0$ and $\lambda_{in}^{s}=0$.