



Munich Personal RePEc Archive

Manipulating a stated choice experiment

Fosgerau, Mogens and Börjesson, Maria

Technical University of Denmark, Denmark, Royal Institute of
Technology, Sweden

2015

Online at <https://mpra.ub.uni-muenchen.de/67053/>
MPRA Paper No. 67053, posted 04 Oct 2015 06:34 UTC

Manipulating a stated choice experiment

Mogens Fosgerau

Technical University of Denmark, Denmark

Royal Institute of Technology, Sweden

mf@transport.dtu.dk

Maria Börjesson

Royal Institute of Technology, Sweden

maria.borjesson@abe.kth.se

September 24, 2015

Abstract

This paper considers the design of a stated choice experiment intended to measure the marginal rate of substitution (MRS) between cost and an attribute such as time using a conventional logit model. Focusing the experimental design on some target MRS will bias estimates towards that value. The paper shows why this happens. The resulting estimated MRS can then be manipulated by adapting the target MRS in the experimental design.

JEL: C8, C9, C25

Keywords: stated choice; willingness to pay; misspecification; experimental design

1 Introduction

This paper considers a stated choice experiment designed to measure the marginal rate of substitution (MRS) between cost and another attribute of the choice alternatives. We present a case where the design is centered on some target MRS and where the MRS is estimated by applying a conventional logit model. We demonstrate that if the random individual response heterogeneity is mainly driven by heterogeneity of the MRS rather than by response errors, then the estimated MRS will tend to be biased towards the target. This bias hinges on misspecification of the logit model, which has constant marginal utilities and additive random residuals. Thus, this bias will tend to confirm the target MRS and may thus not be informative about consumer preferences.

Seasoned choice experiment designers probably know this already. The objective of this paper is to make clear how the mechanism works and to make this insight more widely accessible.

Stated choice experiments are used within a range of applied fields of economics, including energy, transportation, health, tourism, agricultural and environmental economics. They are also used by consultants in a variety of contexts and in marketing. The mechanism explored in this paper may not only be present in stated choice data, but also when the MRS is estimated on revealed preference data.

The mechanism described in the present paper should not be confused with other sources of bias on the MRS inferred from stated choice data. The choice experiment can influence the output for a number of reasons. Many cognitive processes are at play when people make decisions, typically simplified by various decision rules, heuristics, implying that the experimental design can influence the result (Hensher, 2014; Leong and Hensher, 2012). There are empirical (Bliemer and Rose, 2011) and theoretical indications (Burgess and Street, 2005; Sandor and Wedel, 2005) that attribute levels and the order of them within the experiment influence output (see Rose and Bliemer, 2014, for a review). De Borger and Fosgerau (2008) and Hess et al. (2008) show that preference asymmetries such as different valuation of gains or losses, and reference effects influence respondent's stated choices, with De Borger and Fosgerau (2008) linking these phenomena to prospect theory (Tversky and Kahneman, 1991). In particular, the value of time from stated choice experiments has been found to depend on the size and sign of the attribute differences between the alternatives. A commonly found sign effect is that gains are valued less than losses; this is called loss aversion. Size effects refer to cases where the estimated MRS is found to depend on the size of the attribute difference between alternatives (e.g., Mackie et al., 2001; Hultkrantz and Mortazavi, 2001; Bates and Whelan, 2001; Fosgerau, 2006). A commonly found size effect is that small time savings are valued less per minute than large time

savings.

Adaptive design techniques can introduce bias in the estimated MRS through yet another mechanism, namely that attribute levels and unobserved heterogeneity in the respondents preferences become correlated (Bradley and Daly, 1993). Adaptive designs then lead to a self-imposed endogeneity problem, violating the statistical assumptions underlying standard models, and making subsequent statistical inference invalid.

There is evidence that stated choices are sometimes subject to hypothetical bias (Harrison, 2014). In the context of risky choices, nonlinearities regarding attitudes and perceptions of risks imply that the design of the experiment in terms of risk levels and presentation of those influence the choices (Bates et al., 2001; Liu and Polak, 2007; Borjesson and Eliasson, 2011; Loomes and Blackburn, 2014). These other potential sources of bias are not in focus in this paper.

The mechanism we discuss in the present paper relates to the boundary value design approach (Fowkes and Wardman, 1988). The idea of this design approach is to choose boundary or trade-off values within a range where the analyst think the MRS distribution is located. Fowkes and Wardman point out that the boundary values should cover a reasonable range of potential variation in taste and uncertainty, but that it is often desirable to have them closer together in the range where the actual values are expected to be located. An implication of our results is that the boundary value approach should be avoided, since it runs the risk of biasing the estimated MRS toward the chosen boundary values. Moreover, it is not suited to models that account for preference heterogeneity.

A remedy to avoid the bias arising from the misspecification of the logit model is to explore the error structure of the data, using non-parametric techniques, before defining a parametric model. If such analysis indicates heterogeneity in the MRS, the parametric model should allow for heterogeneity in the MRS. The choice of parametric distribution for the MRS is crucial and should be tested (Fosgerau and Bierlaire, 2007; Fosgerau and Mabit, 2013).

To estimate the distribution of the MRS, a key condition is that the full distribution of the MRS is uncovered by the data. If this condition is not met, then the estimates of the mean and other moments of the MRS distribution have to rely on assumptions about the shape of the distribution in the range where it cannot be identified by the data. Such assumptions are hard to verify. Fosgerau (2006) shows that when the tail of the MRS distribution is not revealed, then the choice of parametric distributions can result in arbitrarily high estimates of the mean MRS. The resulting estimated of the mean MRS will depend on the parts of the distribution that are extrapolated outside the range of data.

The paper is organized as follows. Section 2 describes the experimental context and choice generating process assumed when applying the standard logit model. Section 3 describes a different choice generating process, governed by

the random MRS model, and describes the mechanism that tends to bias the estimated MRS. Section 4 describes the experimental design used in the simulation exercise in section 5. Section 6 uses part of the Danish value of time data to empirically validate the theoretical predictions of section 3. Section 7 concludes.

2 Choice setup and the model to estimate

The context of a trip is used for concreteness. The alternatives can be different routes by some mode of transportation, for example; section 6 uses data concerning rail trips. Each subject is asked to choose between two alternatives for a trip and each route is characterized by an associated monetary cost and a travel time.

This setting can be used to reveal the subjects' rate of substitution between travel time and cost. By design, one route will be faster but also more expensive than the other. With everything else being equal, subjects reveal through their choices whether their willingness to pay for the time saving associated with the faster route is greater or smaller than the cost difference between the two routes.

The canonical model for this situation is the logit model with the foundation in the theory of consumer demand developed by [McFadden \(1974\)](#). The conventional and widely used specification of the binary logit model in this setting uses indirect utilities for the two alternatives that are linear indices $v_{ij} = \alpha c_{ij} + \beta t_{ij} + \gamma 1_{\{j=2\}} + \varepsilon_{ij}$, where subscript i indexes individuals, $j = 1, 2$ indexes choice alternatives, c_{ij} is the travel cost for individual i in alternative j with corresponding marginal utility α , t_{ij} is the travel time with corresponding marginal utility β , γ is a constant specific to alternative 2 and ε_{ij} are i.i.d. extreme value type 1 random residuals. The marginal utilities α, β are expected to be negative. Each individual chooses the alternative yielding the highest indirect utility. Then the probability that alternative 1 is chosen is

$$\begin{aligned} p_{i1} &= \frac{e^{\alpha c_{i1} + \beta t_{i1}}}{e^{\alpha c_{i1} + \beta t_{i1}} + e^{\alpha c_{i2} + \beta t_{i2} + \gamma}} \\ &= \frac{1}{1 + e^{\alpha \Delta c_i + \beta \Delta t_i + \gamma}}, \end{aligned} \tag{1}$$

where $\Delta c_i = c_{i2} - c_{i1}$ and $\Delta t_i = t_{i2} - t_{i1}$. The marginal rate of substitution between time and cost in this model is β/α .

This model can be extended in various ways. The most common are to interact parameters with individual characteristics, to apply nonlinear transformations to the cost and time variables or to use random parameter specifications ([McFadden and Train, 2000](#)). Here we shall stick with the most basic specification in order to gain maximal intuition.

Figure 1 shows the shape of the logit probability p_1 as a function of $(\Delta c, \Delta t)$. In the quadrant shown, alternative 1 has a lower travel cost but a longer travel time than alternative 2 (Δc_i is positive and Δt_i is negative). In the light gray area in the upper left corner the probability of choosing the cheaper alternative 1 is high (ranging from 0.8 to 1). In the darker area in lower right corner the probability is low (ranging from 0 to 0.2). The point of the figure is to show that the probability is constant on parallel straight lines in the $(\Delta c, \Delta t)$ -plane. From the logit probability (1) we see that constant p_1 implies that $\alpha\Delta c + \beta\Delta t + \gamma = C$, where C is a constant given by probability p_1 . This implies $\Delta c = C - \gamma - \frac{\beta}{\alpha}\Delta t$, which shows that the equiprobability lines have the constant slope $-\beta/\alpha$.

Keeping the ratio $-\beta/\alpha$ fixed, the distance between equiprobability lines decreases as α and β increase in absolute value. This is easily seen on the vertical Δc -axis in the figure, along which $\Delta t = 0$. The distance between the equiprobability lines equals $\frac{\Delta C}{\alpha}$, where ΔC is a function of the difference in probability between the lines (Δp_1) which is 0.2 in the figure. Hence, a larger α , implies a larger distance between the lines. On the Δt -axis the distance between the equiprobability lines equals $\frac{\Delta C}{\beta}$.

The equiprobability lines can be shifted perpendicularly by changing γ . The probability traces out a logistic distribution in the direction perpendicular to the equiprobability lines.

Fitting the logit model to responses to $(\Delta c, \Delta t)$ by estimating (α, β, γ) amounts to determining the best slope $-\beta/\alpha$ of equiprobability lines, determining the scale of (α, β) to match the distance between equiprobability lines and determining γ to match the location of the equiprobability lines.

3 A more realistic data generating process

When fitting the standard logit model described in the previous section to data, the modeler makes the assumption that all the response heterogeneity in the data arises from response error and none from variation in the MRS. To illustrate how the logit model may bias the result if this assumption does not hold, we continue by making the assumption in the other extreme: that all the response heterogeneity in the data arises from variation in MRS and none of it arises from response error. We assume that this data is the result of the choice generation process ruled by the random MRS model proposed by [Cameron and James \(1987\)](#).

In the random MRS model, individuals each have a MRS w_i , which has some absolutely continuous population distribution Φ . Individual i evaluates alternatives with the generalized cost $c_{ij} + w_i t_{ij}$ and chooses the alternative with the least generalized cost. Assuming $\Delta t_i \neq 0$, the probability that alternative 1 is

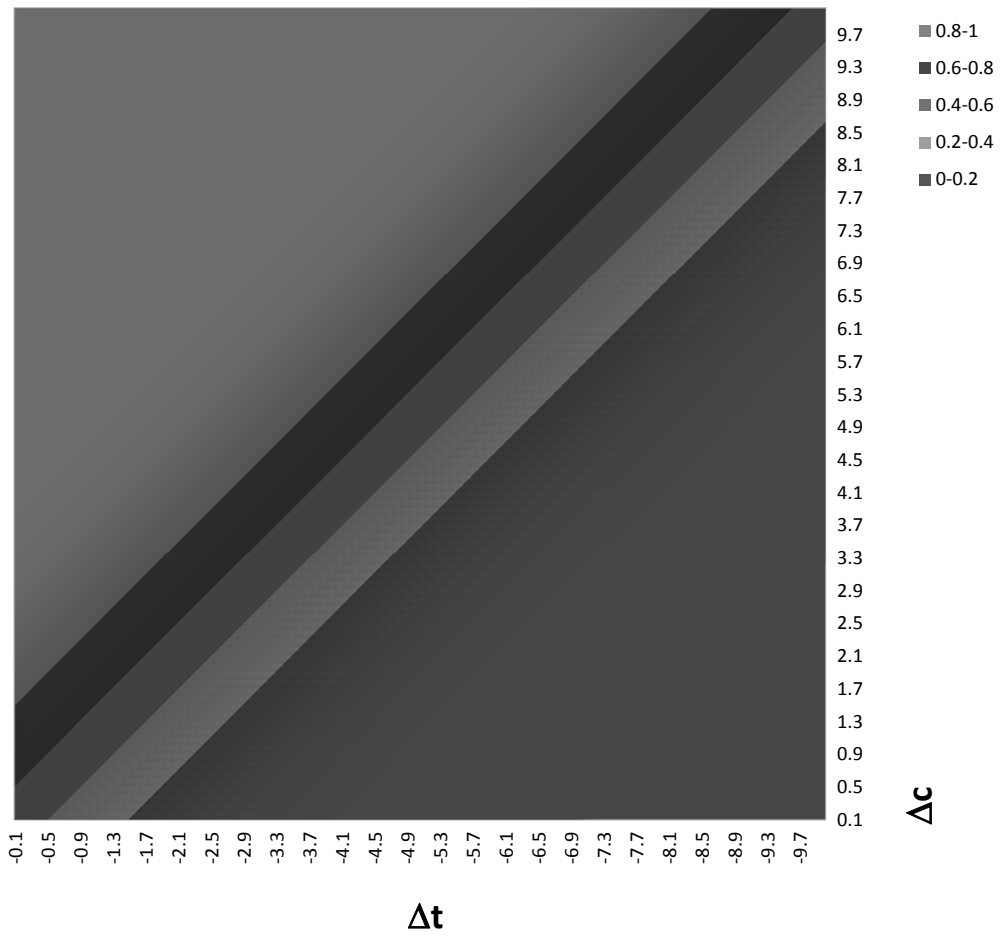


Figure 1: Probabilities in the logit model

chosen by individual i is

$$\begin{aligned}
\hat{p}_{i1} &= P(c_{i1} + w_i t_{i1} < c_{i2} + w_i t_{i2}) \\
&= P(-w_i \Delta t_i < \Delta c_i) \\
&= \begin{cases} \Phi\left(-\frac{\Delta c_i}{\Delta t_i}\right) & \Delta t_i < 0 \\ 1 - \Phi\left(-\frac{\Delta c_i}{\Delta t_i}\right) & \Delta t_i > 0 \end{cases} \quad (2)
\end{aligned}$$

The difference between this model and (1) is that all response heterogeneity here is captured by w_i , while there is no ε added to capture random error.

This model is similar to the transfer price model of Hensher [Hensher \(1976\)](#), which assumes that $c_u + TP - c_a = a_0 + a_1(t_u - t_a)$, where c is travel, t is travel time, u indicates the usual mode, a indicates alternative mode, and the transfer price TP is defined as 'the amount of cost change that would have to occur in the usual mode journey for the individual to consider an alternative mode of transport'.

Figure 2 shows the probability that alternative 1 is chosen against $(\Delta c, \Delta t)$, using as example a lognormal distribution for Φ . From (2), we have that $\Delta c_i = -C\Delta t_i$, where C is constant for a given probability. Hence, the probability is constant on rays from the origin in $(\Delta c, \Delta t)$ -space as illustrated in the figure. From (2) we also see that the slope of the equiprobability lines increases in p_1 in the quadrant shown in the figure (where $\Delta t_i < 0$). The equiprobability lines have a large slope in the light area in the upper left corner where the probability of choosing the alternative 1 is high (ranging from 0.8 to 1). In the dark gray area in the lower right corner, where the probability is low (ranging from 0 to 0.2), the equiprobability lines have smaller slope.

Previous studies suggest that the random MRS model is a better description of the true data generating process than the logit model (1): [Fosgerau \(2007\)](#) and [Borjesson et al. \(2012\)](#) use nonparametric techniques to distinguish between the logit and the random MRS choice generating processes, with data from the Danish and the Swedish national value of time studies, respectively. Both studies find clear evidence in favor of the random MRS model. A part of the Danish dataset is used in section 6.

4 A bad experimental design

The objective of the present paper is not to explore the pros and cons of the many different SP design techniques that exist. It is merely to show what happens if the trade-offs in the data are centered around a target MRS and the standard logit model is estimated even if the random MRS model describes the data better. So we

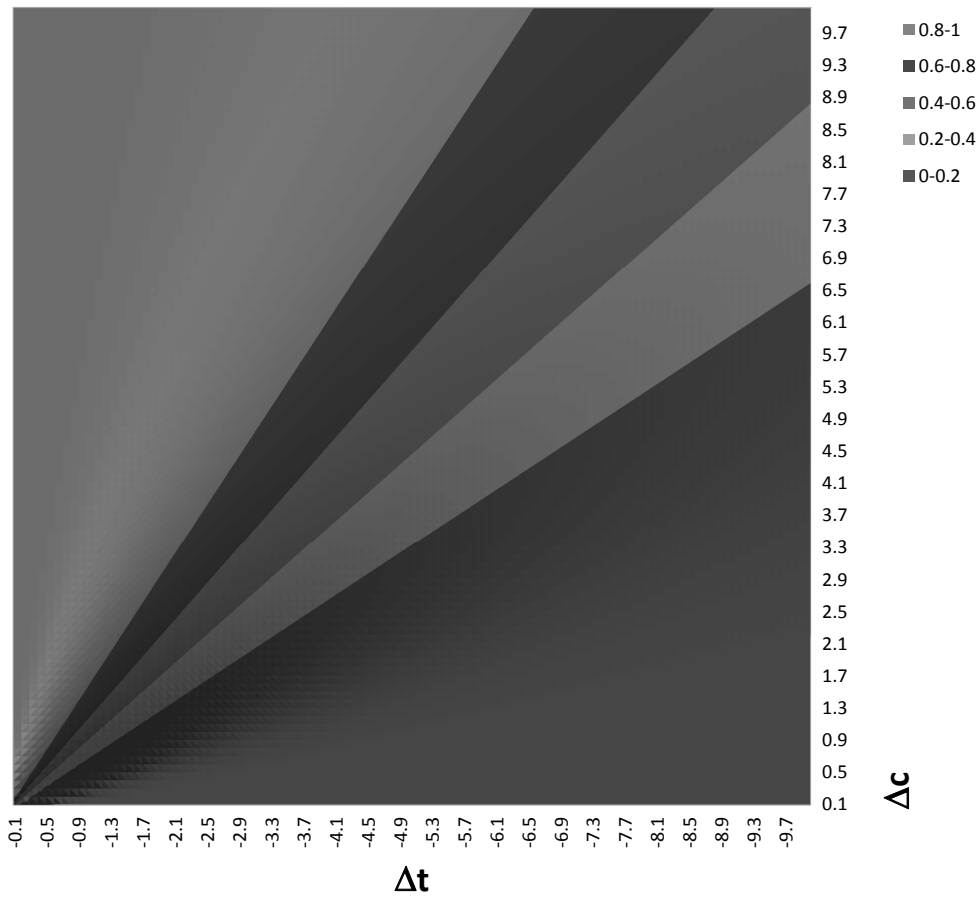


Figure 2: Probabilities in the random MRS model

shall see what happens when a data generating process that really is the random MRS model of Section 3 meets a modeler who has the logit model of Section 2 in mind. This is the core insight of this paper.

Say that the experimental designer assumes a constant MRS across the population and has a target MRS in mind, denoted \hat{w} . He¹ may first select time differences Δt_i from some range. The range would be chosen to yield time differences that are noticeable and that subjects can accept as being realistic. Next, he may select corresponding values of cost differences Δc_i that are near $-\hat{w}\Delta t_i$, but with an equal mix of values that are larger and smaller such that the design has points $(\Delta c_i, \Delta t_i)$ that are centered on the ray with slope $-\hat{w}$. A common rationale for such a design strategy is that it should ensure that the MRS $-\beta/\alpha$ (assumed constant by the designer) is estimated with high precision provided it is close to \hat{w} . Having collected responses to this design, the final point in this methodology is to fit a logit model as in (1) to these data.

Under the maintained and plausible assumption that the main source of heterogeneity is actually random MRS with some distribution, the expected response will be similar to Figure 2. By design, there will be only data near the target ray and the expected response will be roughly constant.

If the sample is sufficiently large, then the estimated logit model (or a similar model) will approximate this shape. Therefore the estimated logit model will have constant probability on lines that have slope close to the target MRS and thus the estimated MRS will be close to the target. This is, of course, extremely problematic, since the point of the exercise is to estimate the MRS of experimental subjects and not the target MRS of the experimental designer. The scale of the coefficients α, β is determined by the slope of the response surface in the perpendicular direction while the constant γ adjusts to match the estimated probability to the observed.

5 Simulation exercise

The theoretical predictions made in section 2 and section 3 are now validated in a small simulation exercise. In the simulation exercise we adopt the choice generation process described in section 3, the random MRS model. We generate five datasets by applying an experimental design as discussed in section 4. In the simulation of each of the five data sets we assume a different target MRS \hat{w} when generating the experimental design. The standard logit model (1) is then estimated on the five simulated datasets, such that one MRS estimate is achieved for each dataset and target MRS.

¹We determined the gender of this representative person by flipping a Swedish krona.

Table 1: Estimation results, simulation data, standard errors in parentheses

Target MRS	α	β	γ	Estimated MRS
0.5	-0.246 (0.062)	-0.162 (0.037)	-0.613 (0.28)	0.662
1	-0.149 (0.026)	-0.159 (0.030)	0.174 (0.24)	1.064
2	-0.080 (0.015)	-0.175 (0.035)	1.426 (0.28)	2.205
3	-0.068 (0.013)	-0.224 (0.043)	2.248 (0.36)	3.292
4	-0.056 (0.012)	-0.259 (0.054)	3.022 (0.45)	4.598

More specifically, a choice experiment was designed by first fixing a target MRS. Time differences Δt_i were sampled from a uniform distribution on $[10, 20]$ and the cost differences were set to the target MRS times a random noise term sampled from a uniform distribution on $[0.8, 1.2]$.

A dataset was then generated where 2000 individuals make one choice each with an MRS drawn from a standard lognormal distribution. Applying the binary logit model (1) the MRS is then estimated from the generated dataset. The whole procedure is then repeated four times assuming different target MRS in the design of the choice experiment.

Table 1 reports the logit estimation results, including the estimated MRS α/β corresponding to each of the five target MRS. Evidently, it is possible to generate estimated MRS that are quite close to the target over a wide range. The parameters of interest α and β are estimated with good precision, which would not lead one to suspect problems. This happens even though the model, by construction, is misspecified.

6 Empirical test

The empirical test uses data from a Danish value of time study (Fosgerau et al., 2007). The dataset concerns binary choices between train trips differentiated by travel time and cost. The design was created pivoting around the travel time and cost of actual train trips undertaken by subjects. Eight choice situations were generated for each subject by first drawing two absolute travel time differences from a set. Second, eight MRS were drawn from the interval $[3:200]$ Danish Kroner (DKK) per hour², using stratification to ensure that all subjects were presented with both low and high values. The absolute cost difference was then found for

²1 EUR \simeq 7.5 DKK. The maximum value corresponds to about twice the average after tax hourly wage in Denmark.

Table 2: Estimation results, rail data, standard errors in parentheses

MRS	Est MRS	#obs	α	β	γ	LL
[0, 19.5]	24.212	1728	-0.652 (0.044)	-15.8 (0.65)	-0.0881 (0.029)	-935.575
[19.5, 36]	41.726	1689	-0.147 (0.020)	-6.14 (0.56)	-0.0463 (0.025)	-1112.29
[36, 60]	43.293	2122	-0.138 (0.017)	-5.98 (0.80)	-0.0752 (0.025)	-1456.54
[60, ∞]	49.536	1868	-0.105 (0.006)	-5.18 (0.52)	-0.0516 (0.027)	-1092.77

each choice situation by multiplying the absolute time difference by the MRS. Third, the sign of the cost and time differences relative to the reference were determined such that they were opposite and such that each of the four possible permutations was used twice. The differences were added to the reference to get the numbers that were presented to subjects on screen. The sequence of choice situations was randomly scrambled. Travel costs were rounded to the nearest 0.5 DKK. The data have been trimmed prior to analysis by removing the 5% of observations having the largest journey times and time and cost differences.³

The choice situations in these data have a wide range of implicit MRS. They can therefore be used to simulate the effect of using a narrower range of MRS centered on different target MRS. The data are thus split according to quartiles of the MRS implicit in choice situations and a model is estimated for each quartile. This simulates designs that are based on different intervals of MRS. The following table shows the results of estimating binary logit models on these data. The data in each split are pooled and standard errors are not corrected for repeated observations. Our interest is just to see how the estimated MRS can be manipulated.

The results in Table 2 show that the estimated MRS does change in the appropriate direction as the MRS design interval is changed and that there is a factor two difference between the highest and the lowest estimated MRS. This shows that the approach outlined above leads to estimation results that are strongly influenced by the target MRS. As in the simulation exercise, the precision of the estimates is good, which does not reveal that the model is misspecified.

The relationship between estimated and target MRS is not as tight as in the simulation. This is unsurprising since the simulation model was created to precisely deliver the desired effect. The simulation model, however, is not a perfect representation of the choice behavior of real subjects, and we do expect that some response heterogeneity arises from other sources than variation in MRS in real data.

³This is done to avoid outliers and obtain more robust estimates.

7 Concluding remarks

There is empirical evidence indicating that a substantial part of the response heterogeneity in stated choice experiments arises from variation of the MRS in the population. That, in combination with an SP design based on a target MRS and in combination with a misspecified model, will lead to bias. This can be a severe problem.

We have illustrated the problem in a very simple setting with just two alternatives and two attributes and a naive SP design. While this is very simple, it is close to what is sometimes done, in research and in consultancy projects. Of course current state-of-the-art design techniques are much more sophisticated, see [Rose and Bliemer \(2009\)](#) for an overview. We do, however, hope that the intuition one may gain from the examples we have presented in this paper will be useful when working with more advanced design methods. It seems to be an open question to which extent various types of designs and consequent model estimation are robust against misspecification.

Extrapolating from our results, it is possible to offer some guidance. One general point is that it is a very good idea to perform specification testing of estimated models. We have seen that it is not sufficient to examine the precision of parameter estimates. It is possible to have quite precise estimates from a model that is completely misspecified.

In order to be able to do specification testing, it is important that there is sufficient variation in the data that allows potential misspecification to be detected. This provides an argument for SP designs with wide attribute ranges. Wide attribute ranges are also useful for identification of parameter distributions, where the distribution of willingness-to-pay parameters is of particular interest. Where the distribution of a random MRS is of interest, then experimental design should allow nonparametric identification of that distribution. Our results illustrate the importance of using SP designs that are robust in the sense that they will allow the identification of a range of models.

Analysis of SP data often uses the plain logit model similar to (1) as a starting point and elaborates from that. In principle, this model can be generalised into any other model, for example adding (many) random parameters ([McFadden and Train, 2000](#)). In practice, however, this is never possible since datasets are always finite and mostly not that large. It is then a better idea to begin from a parsimonious model that describes the main features of the data. Depending on the circumstances, this could be a model where the basic source of randomness is random MRS rather than random noise. Non-parametric tools exist that can assist in choosing a base model.

Finally, we note that the misspecification issue that we have discussed is not exclusive to SP but may arise also with revealed preference (RP) data. It is the

structure of the data that matters, not how they were generated. We note also that it is not essential that the sign of the random MRS is known, the results in section 3 can be extended to cases where the sign of the random MRS is negative or where it may attain both positive and negative values.

We hope that these results will inspire further research on SP design and analysis.

8 Acknowledgements

We thank Glenn Harrison and Stefan Mabit for comments. Mogens Fosgerau has received financial support from the Danish Strategic Research Council.

References

- Bates, J., Polak, J., Jones, P. and Cook, A. (2001) The valuation of reliability for personal travel *Transportation Research Part E* **37**(2-3), 191–229.
- Bates, J. and Whelan, G. (2001) Size and sign of Time Savings *ITS Working Paper* (561).
- Bliemer, M. C. J. and Rose, J. M. (2011) Experimental design influences on stated choice outputs: An empirical study in air travel choice *Transportation Research Part A: Policy and Practice* **45**(1), 63–79.
- Borjesson, M. and Eliasson, J. (2011) On the use of “average delay” as a measure of train reliability *Transportation Research Part A: Policy and Practice* **45**(3), 171–184.
- Borjesson, M., Fosgerau, M. and Algers, S. (2012) Catching the tail: Empirical identification of the distribution of the value of travel time *Transportation Research Part A: Policy and Practice* **46**(2), 378–391.
- Bradley, M. and Daly, A. (1993) New issues in stated preferences research *Technical report* 21st PTRC Summer Annual Meeting Manchester, U.K.
- Burgess, L. and Street, D. J. (2005) Optimal designs for choice experiments with asymmetric attributes *Journal of Statistical Planning and Inference* **134**(1), 288–301.
- Cameron, T. A. and James, M. D. (1987) Efficient Estimation Methods for "Closed-Ended" Contingent Valuation Surveys *Review of Economics and Statistics* **69**(2), 269–276.

- De Borger, B. and Fosgerau, M. (2008) The trade-off between money and time: a test of the theory of reference-dependent preferences *Journal of Urban Economics* **64**(1), 101–115.
- Fosgerau, M. (2006) Investigating the distribution of the value of travel time savings *Transportation Research Part B: Methodological* **40**(8), 688–707.
- Fosgerau, M. (2007) Using nonparametrics to specify a model to measure the value of travel time *Transportation Research Part A: Policy and Practice* **41**(9), 842–856.
- Fosgerau, M. and Bierlaire, M. (2007) A practical test for the choice of mixing distribution in discrete choice models *Transportation Research Part B: Methodological* **41**(7), 784–794.
- Fosgerau, M., Hjorth, K. and Vincent Lyk-Jensen, S. (2007) The Danish Value of Time Study - Final Report *Technical report* www.dtf.dk.
- Fosgerau, M. and Mabit, S. L. (2013) Easy and flexible mixture distributions *Economics Letters* **120**(2), 206–210.
- Fowkes, T. and Wardman, M. (1988) The Design of Stated Preference Travel Choice Experiments: With Special Reference to Interpersonal Taste Variations *Journal of Transport Economics and Policy* **22**(1), 27–44.
- Harrison, G. W. (2014) Real choices and hypothetical choices *Handbook of Choice Modelling* Edward Elgar Publishing pp. 236–254.
- Hensher, D. (2014) Attribute processing as a behavioural strategy in choice making Edward Elgar pp. 268–289.
- Hensher, D. A. (1976) The Value of Commuter Travel Time Savings: Empirical Estimation Using an Alternative Valuation Model *Journal of Transport Economics and Policy* **10**(2), 167–176.
- Hess, S., Rose, J. M. and Hensher, D. A. (2008) Asymmetric preference formation in willingness to pay estimates in discrete choice models *Transportation Research Part E: Logistics and Transportation Review* **44**(5), 847–863.
- Hultkrantz, L. and Mortazavi, R. (2001) Anomalies in the Value of Travel-Time Changes *Journal of Transport Economics and Policy* **35**(2), 285–300.
- Leong, W. and Hensher, D. A. (2012) Embedding Decision Heuristics in Discrete Choice Models: A Review *Transport Reviews* **32**(3), 313–331.

- Liu, X. and Polak, J. (2007) Nonlinearity and Specification of Attitudes Toward Risk in Discrete Choice Models *Transportation Research Record: Journal of the Transportation Research Board* **2014**, 27–31.
- Loomes, G. and Blackburn, S. (2014) Towards a more complex model of risky choice Edward Elgar pp. 73–98.
- Mackie, P., Jara-Diaz, S. and Fowkes, A. (2001) The value of travel time savings in evaluation *Transportation Research Part E: Logistics and Transportation Review* **37**(2-3), 91–106.
- McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behaviour *Frontiers in Econometrics* Academic Press New York pp. 105–142.
- McFadden, D. and Train, K. (2000) Mixed MNL Models for discrete response *Journal of Applied Econometrics* **15**, 447–470.
- Rose, J. and Bliemer, M. (2014) Stated choice experimental design theory: the who, the what and the why Edward Elgar pp. 152–177.
- Rose, J. M. and Bliemer, M. C. J. (2009) Constructing Efficient Stated Choice Experimental Designs *Transport Reviews* **29**(5), 587–617.
- Sandor, Z. and Wedel, M. (2005) Heterogeneous Conjoint Choice Designs *Journal of Marketing Research* **42**(2), 210–218.
- Tversky, A. and Kahneman, D. (1991) Loss Aversion in Riskless Choice: A Reference-Dependent Model *The Quarterly Journal of Economics* **106**(4), 1039–1061.