JENA ECONOMIC RESEARCH PAPERS

2010 – 003

Incomplete Preferences in Choice Experiments: A note on avoidable noise and bias in welfare estimates

by

Mitesh Kataria
Jason F. Shogren

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Incomplete Preferences in Choice Experiments:
A note on avoidable noise and bias in welfare estimates

Mitesh Kataria[A]
Jason F. Shogren[B]

Abstract

How does a choice experiment model derived under standard preference axioms perform for respondents with incomplete preferences? Using simulated data, we illustrate how this preference-model mismatch generates noise and bias in welfare estimates, and we show how it can be avoided.

Keywords: Choice experiment, Ordered Logit, Bias, Preference Axioms.

JEL Classification: D61, Q51.

Acknowledgments: We have received valuable comments from Astrid Matthey, Elina Lampi and Karin Larsén and Riccardo Scarpa. The usual disclaimer applies.

[A] Max Planck Institute of Economics, Strategic Interaction Group, Kahlaische Strase 10, D-07745 Jena, Germany. Tel: +49 3641 686 632. E-mail: kataria@econ.mpg.de
Kahlaische Straße 10, D-07745 Jena, Germany. Tel: : +49 3641 686 632. E-mail: kataria@econ.mpg.de
[B] University of Wyoming, Department of Economics and Finance, Laramie, WY 82071-3985

1. Introduction

Choice experiments (CE) are a popular method to elicit preferences (see Louviere, 2000). CEs are designed to reveal preferences by asking a person to choose between alternative consumption bundles—assuming people have a complete preference ordering. But in complex or hypothetical decisions such as those for public goods, a person can be indecisive or indifferent to a choice between alternatives (see e.g., Hey and Orne 1994; Wang 1997; Ariely et al. 2003; Cantillo et al. 2009; Hanley et al., 2009).[1] Eliaz and Ok (2006) argue that substantial evidence exists which suggests the explanatory power of the standard theory of individual choice is unsatisfactory; in response, they developed a choice-theoretic foundation for incomplete preferences. People with incomplete preferences might prefer being given a "no-opinion" option in a CE survey. Excluding this "no-opinion" option could yield biased estimates of preferences if the indecisive/indifferent person was treated in the econometric analysis as if he had complete preferences.[2] But adding the no-opinion option also raises estimation issues in CE. As noted by Fenichel et al. (2009): "Including no-opinion response options means that respondents will select them, which reduces the sample size of yes and no responses. However, if there is a way to recover information from some no-opinion responses, then adding no-opinion response options may be beneficial."[3]

[1] These people are violating the fundamental completeness axiom underpinning demand theory. Recall the completeness axiom assumes a person choosing between two bundles x_1 and x_2, can rank the alternatives as either: (i) x_1 is preferred to x_2, (ii) x_1 is indifferent to x_2, or (iii) x_2 is preferred to x_1.

[2] Respondents´ might still answer to please the experimenter. Another reason might be that they gain compensation if all questions in the survey are answered. In web-surveys it is not unusual that respondents´ cannot proceed to next question if they have not answered previous questions.

[3] Fenichel et al. (2009) used a split-sample design to explore the implications of including no-opinion responses in CE application to estimate preferences for inland, freshwater wetland mitigation. They found 25 percent of the responses to be no-opinion responses. For more complicated surveys one could of course expect higher amount of no-opinion responses.

Within the contingent valuation (CV) literature, Groothuis and Whitehead (2002) examined the effects of including a “Don´t Know” (DK) alternative in referendum contingent valuation (CV) study. They suggest a DK option avoids a large amount of protest responses.[4] Wang proposed a random valuation model that explicitly treats indecisive responses assuming that uncertainty arises because the alternatives have a similar level of utility and respondents have complete preferences if the thresholds of the utilities are exceeded. Balcombe and Fraser (2009) propose a model that simultaneously deals with misreporting and DK responses; here a reported DK might be a YES, NO or DK response, and YES could be a NO and vice versa.

In the CE literature researchers are interested in how to deal with choice task complexity. For example, respondents do not always consider all attributes in choosing the utility maximizing alternative (e.g. Hensher et al 2005; Campbell et al. 2008), people have lexicographical preferences (e.g. Burton and Rigby 2009), and people adopt a simplified strategy in making decision with high level of task complexity (e.g. Swait and Adamowicz 2001). Randomly choosing the alternatives could be another example of such simplified behavior but that easily can be avoided by including a no-opinion option.

In this note, we illustrate how an *ordered logit* model could be used to recover the information from no-opinion responses in CE that allow for incomplete preference orderings when similarities between alternatives lead people to be indecisive/indifferent between alternatives.[5] Extending the work of Krishnan (1977) and Wang (1997) we find

[4] Strazzera et al. (2003) proposed a mixture model with sample selection to account for both the true zero values (i.e. respondents who are indifferent to whether the public good is provided) and the protest responses.

[5] Contillo et al. (2009) develops a similar model using random thresholds and evaluates error in part-whole values (WTP) using synthetic and real data. We compare a model with and without fixed thresholds using Monte Carlo simulations and extend the scope of welfare measures that are evaluated. One of the strengths of the Monte Carlo Simulation method is that it uses repeated sampling that generates a large number of

three key results based on simulated data. First, the traditional binary logit model becomes noisier and nosier as the fraction of indecisive/indifferent respondent grows. Second, the ordered logit approach estimates values without much more noise regardless of the indecisive fraction of the population. Third, while absolute welfare estimates are not significantly biased in the binary logit model, relative values are biased in proportion to the fraction of incomplete preferences.

2. Econometric Model and Monte Carlo Simulation

First, we define our benchmark model. The traditional model used in choice experiments is the Logit model, which assumes a complete preference ordering and the absence of indecision/indifference, also called the "no-opinion" response. In the binary choice model, a respondents´ choice between two alternatives, 1 and 2, is modeled as an index function:

$$y = 1 \text{ if } \Delta U^* (= U_2 - U_1) = \alpha + \beta' x + \varepsilon > 0, \tag{1}$$

$$y = 0 \text{ if } \Delta U^* (= U_2 - U_1) = \alpha + \beta' x + \varepsilon < 0 \tag{2}$$

The latent function U* can be interpreted either as general index function or as a net-utility function. Assuming each error terms is independently and identically Logistic distributed we have the Logit Model. This is the standard model in which the no-opinion response is not elicited.

Second, we now develop our *ordered logit* model which can recover information from no-opinion responses that allow for indecision/indifference in the preference ordering. We incorporate the notion of indecision/indifference by adding thresholds into

synthetic data sets that can be used to evaluate various statistics without being contingent on any single number of utilized samples of synthetic data.

the standard binary model. Assume the respondent chooses alternative 2 if $\alpha + \beta' x + \varepsilon > k_1$; alternative 1 if $\alpha + \beta' x + \varepsilon < k_2$ .The net-utility of the alternatives must exceed a threshold values for the respondent to choose one of the alternatives. The respondent is defined as *indecisive/indifferent* when $k_2 < \alpha + \beta' x + \varepsilon < k_1$.

If each error term is independently and identically normal distributed, we have the order probit model (McKelvey and Zavoina, 1975). Assume the error term is logistic distributed. For the standard model, the probability of choosing alternative 1 or 2 is:

$$P_{n,1} = P(\varepsilon < k_2 - \alpha - \beta' x) = \frac{\exp(k_2 - \alpha - \beta' x)}{1 + \exp(k_2 - \alpha - \beta' x)}, \tag{3}$$

$$P_{n,2} = 1 - P(\varepsilon < k_1 - \alpha - \beta' x) = 1 - \frac{\exp(k_1 - \alpha - \beta' x)}{1 + \exp(k_1 - \alpha - \beta' x)} \tag{4}$$

Write the joint probability a person will choose the indecisive/indifferent alternative as:

$$\begin{aligned} P_{n,Ind} &= P(k_2 - \alpha - \beta' x < \varepsilon < k_1 - \alpha - \beta' x) = P(\varepsilon < k_1 - \alpha - \beta' x) \\ &- P(\varepsilon < k_2 - \alpha - \beta' x) = \frac{\exp(k_1 - \alpha - \beta' x)}{1 + \exp(k_1 - \alpha - \beta' x)} - \frac{\exp(k_2 - \alpha - \beta' x)}{1 + \exp(k_2 - \alpha - \beta' x)} \end{aligned} \tag{5}$$

Let $y1 = 1$ if alternative 2 is chosen, 0 otherwise; and $y2 = 1$ if alternative 1 is chosen, 0 otherwise. Consequently, $1 - y1 - y2 = 1$ if the no-opinion alternative is chosen; $1 - y1 - y2 = 0$ otherwise.

The log likelihood function for the three response categories is:

$$\begin{aligned} LL(\beta) &= \sum y1 * \log\left[\frac{\exp(-k + \alpha + \beta' x)}{1 + \exp(-k + \alpha + \beta' x)}\right] + \sum y2 * \log\left[\frac{\exp(k - \alpha - \beta' x)}{1 + \exp(k - \alpha - \beta' x)}\right] \\ &+ \sum (1 - y1 - y2) * \log\left[\left(\frac{\exp(k - \alpha - \beta' x}{1 + \exp(k - \alpha - \beta' x)}\right) - \left(\frac{\exp(-k - \alpha - \beta' x}{1 + \exp(-k - \alpha - \beta' x)}\right)\right] \end{aligned} \tag{6}$$

We impose an identification restriction $k_2 = -k_1$ in equation (6), which implies the two thresholds are symmetrically placed around zero in the net-utility space. Assuming

$k = |k_1| = |k_2|$ we map the preferences of the respondent accordingly to: $U_1 \succ U_2$ (U_1 is preferred to U_2) if $U_1 > U_2 + k$; $U_2 \succ U_1$ if $U_2 > U_1 + k$ and $U_1 \sim U_2$ (perceived as equal) if $|U_1 - U_2| \leq k$. The threshold k maps what is "too similar" in the utility space and identifies the indecision-indifference responses. The standard binary logit model is a special case with $k = 0$. If the threshold is incorrectly neglected the variance will increase as $|\Delta U^*|$ decreases.[4]

Third, we use Monte Carlos simulations to explore the relative validity of the benchmark binary logit and ordered logit choice models. The benefit of the simulations is that the true parameters and thresholds of the utility function are known. The choices are simulated based on the difference in utility from the deterministic part and a randomly drawn error term from a standardized logistic distribution. Except for when the absolute difference in utility was smaller than the threshold value, a value of 1 was assigned to the choice alternative that produced the greatest utility and 0 to the other choice alternative(s). When the absolute difference in utility was smaller than the threshold, one of the alternatives in the binary choice models were randomly assigned the value of 1. In the ordered logit model, a value of 1 was assigned to the indifferent/indecisive alternative and 0 to the other choice alternative(s). These steps were repeated 2000 times using two sample sizes: 1152 and 2304 observations.

We ran the Monte Carlo study assuming a linear and additive utility function. Equation (7) reflects the *true* difference in utility between the two alternatives:

$$\Delta U^* (= U_2 - U_1) = 2{,}0 + 2{,}0 x_1 + 1{,}0 x_2 - 0{,}01 Cost \tag{7}$$

[4] It is well-known that heteroscedasiticty in non-linear models is problematic and results in inconsistent parameters (Yatchew and Griliches, 1984).

The choice sets were created from the collective factorial (Louviere 1988). The first two attribute (x_1, x_2) are dummy variables; followed by the cost attribute taking the levels 100, 200, 300, 400, 600, and 800. Based on the utility function, economic measures of value is retrieved using Hanneman's (1984) classic formula to calculate (a) *total willingness to pay* (TWTP) for both attributes; the (b) *willingness to pay* for each attribute separately (WTP1 and WTP2); and (c) the *relative willingness to pay* (RWTP = WTP2/WTP1) to illustrate relative values, which can be useful for public policy decisions.[5] The validity of the choice models are evaluated by comparing the true welfare values with the estimated.

We have two indicators of success— *bias* and *precision*. We calculate *bias* by taking the difference of the average welfare estimates and the true value, in which we calculate the average welfare estimates from the 2000 simulated welfare observations. As a measure of *precision* the distribution of the 2000 simulated welfare observations is used, where a wider distribution indicates less precision.

3. Result

Tables 1 and 2 summarize the simulation results based on sample size, n = 2304 or n = 1152. Three key findings emerge. First, as the share of indifferent/indecisive choices is increased, the standard binary model produces welfare estimates with more noise, i.e., more variance and wider distributions (see Table 1). This is intuitive and expected—as more respondents cannot decide between alternatives, methods that assume they can decide become more imprecise.

[5] TWTP = -(2+2+1)/-0.01=500, WTP1=-(2/-0.01)=200, WTP2=-(1/-0.01)=100 and RWTP=2/1=2

Second, in contrast, the precision of the ordered logit model remains good. The reason is that the ordered logit model includes an additional alternative that captures the indecisive responses. Those that worry that including a no-opinion response will reduce the amount of yes and no responses will be pleased to see the model is fairly accurate even when the amount of no-opinion responses increases.

Third, the additional noise has more impact on RWTP relative to WTP and TWTP. Also note that the bias in TWTP never exceeds 5 percent. The bias in WTP is slightly greater, with an upper-limit of 13 percent. In either case, it could be concluded that the bias in TWTP and WTP is relatively small. If share of randomized choices is more than around 40 percent (see column 11 in Table 2) the bias in RWTP is considerable. [6] This suggests resource allocation advice (i.e. share of a budget to spend on different attributes) could be misleading. In practice, however, it seems plausible to avoid such high share of randomized choices by designing a good survey through focus groups and pilot studies.

4. Conclusions

People might have a sense of what they are willing to pay for a quart of milk, but for more complex goods such as environmental services it seems plausible that they only know their WTP within an order of magnitude (also see Hanley et al. 2009 on how some respondents prefer to give a range of values). In this note we illustrate how to estimate such preferences by including a no-opinion alternative and compare it with a traditional

[6] The welfare estimates are unbounded, meaning that when the parameter in the denominator goes to zero, the welfare measure goes to infinity. Increased error variance because of neglected threshold implies higher risk for this to occur, which explains the occurrence of extreme RWTP estimates.

CE model that assumes complete preference orderings. A useful feature of the suggested CE model is that it is straightforward to apply.

Our results show how a CE model that does not address the no-opinion alternative could suffer from unnecessarily noisy welfare measures. This noise can produce misleading conclusions on the significance of WTP, and on significant difference of WTP across attributes and differences across treatments. On the positive side, we show the problem is less serious when the amount of no-opinion responses is low and the sample size is high.

References

Ariley, D., G. Lowenstein, and D. Prelec. 2003. “Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences.” *Quarterly Journal of Economics* 118(1), 73-15.

Balcombe, K., and I. Fraser. 2009. “Dichotomous Choice Contingent Valuation with Don´t Know Responses and Misreporting.” Journal of Applied Econometrics 24:7, 1137–1152

Burton and Rigby. 2009. “Hurdle and Latent Class Approaches to Serial Non-Participation in Choice Models.” *Environmental and Resource Economics* 42:211–226.

Cantillo, V., J. Amaya, and J. Ortúzar. 2009. “Thresholds and indifference in stated choice surveys.” *Transportation Research Part B* 44, 753-763.

Campbell, D., G. Hutchinson and R. Scarpa. 2008. “Incorporating discontinuous preferences into the analysis of discrete choice experiments.” *Environmental and Resource Economics* 41 (3), 101-117.

Eliaz, K. and Ok, E. (2006). "Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences." *Games and Economic Behavior*, 56(1), 61-86.

Fenichel, E., F. Lupi., J. P., Hoehn and M. D. Kaplowitz. 2009. “Split-Sample Tests of “No Opinion” Responses in an Attribute-Based Choice Model.” *Land Economics* 85, *348-362*.

Groothuis, P.A., and J.C. Whitehead. 2002. “Does ‘Don’t Know’ Mean No? Analysis of Don’t Know Responses in Dichotomous Choice Contingent Valuation Questions.” *Applied Economics* 34(15), 1935–1940.

Hanley, N., B. Kriström, and J. F. Shogren. 2009. “Coherent Arbitrariness: On Value Uncertainty for Environmental Goods.” *Land Economics* 85, 41-50.

Hanemann, M. 1984. “Welfare evaluations in contingent valuation experiments with discrete responses.” *American Journal of Agricultural Economics* 66, 332–341.

Hensher, D., J. Rose and W. Greene. 2005. “The implications on willingness to pay of respondents ignoring specific attributes”. *Transportation* 32, 203-222.

Hey, J., and C. Orme. 1994. “Investigating Generalizations of Expected Utility Theory Using Experimental Data.” *Econometrica* 62, 1291-1326.

Krishnan, K.S. 1977. “Incorporating thresholds of indifference in probabilistic choice models.” *Management Science* 23, 1224-1233.

Louviere, J.J. 1988. *Analyzing Decision Making: Metric Conjoint Analysis*. Newbury Park: Sage.

Louviere, J.J, D.A. Hensher and J.U.D. Swait. 2000. *Stated Choice Methods: Analysis and Applications*. Cambridge: Cambridge University Press.

McKelvey, R. D., and Zavoina, W. 1975. “A statistical model for the analysis of ordinal level dependent variables”. *Journal of Mathematical Sociology*, 4, 103–120.

Strazzera, E., Scarpa, R., Calia, P., Garrod, G. D. and Willis, K. G. 2003 "Modelling zero values and protest responses in contingent valuation surveys" Applied Economics, 35, 133–8.

Yatchew, A. and Z. Griliches. 1985. “Specification Error in Probit Models.” *Review of Economics and Statistics* 67, 134-139.

Wang, H. 1997. “Treatment of ‘don’t know’ responses in contingent valuation survey: a random valuation model.” *J. Environ. Econom. Manage*. 32, 219–232.

Table 1: Large Sample: Logit and Ordered Logit model (2304 observations).

Threshold	**Nr. Obs.**	**Random Choices (%)**	**TWTP** (100 & 90 % CI)	**Bias (%)**	**WTP1** (100 & 90 % CI)	**Bias (%)**	**WTP2** (100 & 90 % CI)	**Bias (%)**	**Relative part-whole value (RWTP)** (100 & 90 % CI)	**Bias (%)**
Binary Logit,										
0,00	2304	0	500	0	200	0	100	0	2,03	1,5
			(464 - 542)		*(160 - 238)*		*(54 - 140)*		*(1,29 - 3,56)*	
			(482 - 519)		*(179 - 220)*		*(80 - 119)*		*(1,64 - 2,55)*	
1,50	2304	36	498	-0,4	200	0	101	1	2,04	2,0
			(457 - 545)		*(152 - 250)*		*(52 - 154)*		*(1,25 - 4,14)*	
			(476 - 520)		*(175 - 225)*		*(76 - 125)*		*(1,55 - 2,70)*	
3,00	2304	65	489	-2,2	203	1,5	103	3	2,12	6,0
			(410 - 558)		*(121- 307)*		*(3 - 194)*		*(1,00 - 54,33)*	
			(454 - 523)		*(163- 246)*		*(65 - 142)*		*(1,36 - 3,22)*	
4,50	2304	84	480	-4,0	219	9,5	111	11	0,44*10^11	2,2*10^12
			(339 - 632)		*(54 - 440)*		*(-48 - 328)*		*(-0,73*10^15 - 0,81*10^15)*	
			(410 - 551)		*(137 - 311)*		*(35- 194)*		*(1,00 - 5,91)*	
Ordered Logit										
1,50	2304	37	500	0	200	0	100	0	2.03	1,5
			(472 - 529)		*(166 - 232)*		*(70 - 132)*		*(1,43 - 2.87)*	
			(486 - 514)		*(185 - 216)*		*(83 - 116)*		*(1,70 - 2.43)*	
3,00	2304	65	500	0	200	0	100	0	2,02	1,0
			(472 - 530)		*(164 - 240)*		*(53 - 137)*		*(1,47 - 3,70)*	
			(486 - 515)		*(181 - 218)*		*(83 - 218)*		*(1,67 - 2,46)*	
4,50	2304	84	500	0	200	0	100	0	2,05	2,5
			(460 - 544)		*(148 - 261)*		*(56 - 152)*		*(1,14 - 3,84)*	
			(481 - 519)		*(173 - 229)*		*(77 - 123)*		*(1,59 - 2,65)*	

True TWTP = 500, True WTP1=200, True WTP2=100, True RWTP=2.

Table 2: Small Sample: Logit and Ordered Logit model (1152 observations).

Threshold	Nr. Obs.	Random Choices (%)	TWTP (100 & 90 % CI)	Bias (%)	WTP1 (100 & 90 % CI)	Bias (%)	WTP2 (100 & 90 % CI)	Bias (%)	Relative part-whole value (RWTP) (100 & 90 % CI)	Bias (%)
Binary Logit										
0,00	1152	0	499	-0,2	200	0	99	-1	2,08	4
			(451 - 550)		*(147 - 259)*		*(45 - 164)*		*(1,19 - 5,37)*	
			(474 - 526)		*(171 - 228)*		*(71 - 127)*		*(1,52 - 2,86)*	
1,50	1152	38	498	-0,4	200	0	101	1	2,10	5
			(428 - 563)		*(130 - 275)*		*(29 - 186)*		*(1,05 - 7,08)*	
			(467 - 530)		*(165 - 238)*		*(66 - 136)*		*(1,40 - 3.18)*	
3,00	1152	64	491	-1,8	206	3	104	4	0,19*10^13	9,5*10^13
			(398 - 606)		*(85 - 332)*		*(-2 - 221)*		*(-93,48 - 0,39*10^16)*	
			(443 - 543)		*(145 - 269)*		*(48 – 163)*		*(1,16 - 4,42)*	
4,50	1152	83	483	-3,4	224	12	113	13	0,15*10^12	7,5*10^12
			(254 - 720)		*(12 - 534)*		*(-143 - 441)*		*(-0,74*10^16 - 0,16*10^17)*	
			(386 - 585)		*(109 - 353)*		*(4 - 231)*		*(0,46 - 8,93)*	
Ordered Logit										
1,50	1152	35	500	0	200	0	99	1	2,05	2,5
			(458 - 534)		*(138 - 248)*		*(54 - 144)*		*(1,09 - 4,17)*	
			(480 - 519)		*(179 - 224)*		*(77 - 121)*		*(1,61 - 2,65)*	
3,00	1152	64	500	0	200	0	100	0	2,06	3
			(463 - 542)		*(150 - 258)*		*(43 - 151)*		*(1,30 - 4,37)*	
			(479 - 521)		*(174 - 227)*		*(75 - 151)*		*(1,57 - 2,75)*	
4,50	1152	83	500	0	201	0	100	0	2,10	5
			(451 - 562)		*(131 - 291)*		*(29 - 181)*		*(0,94 - 5,28)*	
			(474 - 527)		*(164 - 242)*		*(68 - 133)*		*(1,45 - 3,02)*	

True TWTP = 500, True WTP1=200, True WTP2=100, True RWTP=2.