I T L S

WORKING PAPER

ITLS-WP-10-08

The accuracy of proxy responses in a stated choice setting

By

Matthew J. Beck, John M. Rose and David A. Hensher

April 2010

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-10-08

TITLE: **The accuracy of proxy responses in a stated choice setting**

ABSTRACT: Data is typically gathered from an individual respondent who represents the group or the household. This individual is often identified as the "primary decision maker" and is often asked to provide responses as a proxy for the group as the cost of interviewing each member individually is impractical and/or expensive. The collection of joint preferences is rarely done (Arora 2006) and indeed in terms of travel behaviour research the use of proxy responses is not uncommon (Wargelin and Kostyniuk 2004, Richardson 2006). Under such a framework, there exists an assumption that the primary decision maker has perfect knowledge of other group member preferences and bargaining behaviour, and is able to synthesize this information when providing a response on their behalf. The validity of such an assumption however remains an open question, with recent research calling the reliability of proxy responses into account (Bateman and Munro 2009). In this paper, using three models estimated in willingness to pay space, we examine the accuracy of proxy responses in a stated choice experiment. We find that there is overlap between a proxy response and the own preferences of the individual providing the proxy choice, that the proxy responses fail to represent the full preference heterogeneity that exists in the actual choices made by individuals, and that when the preferences of another differ substantial from an individuals on preferences the proxy choice provided by that individual is a poor estimate. Overall we find that the ability of individuals to correctly predict the choice of other individuals in their household is poor and as a result proxy responses are unreliable estimators of preference.

KEY WORDS: ***Proxy response, stated choice experiment, willingness to pay, accuracy, preference estimation, sampling method, bias***

AUTHORS: Matthew J Beck, John M Rose and David A Hensher

CONTACT: Institute of Transport and Logistics Studies **(C37)**
The Australian Key Centre in Transport Management
The University of Sydney NSW 2006 Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itls@sydney.edu.au
Internet: http://www.itls.sydney.edu.au

DATE: April 2010

Acknowledgment:
This research was funded under an Australian Research Grant to David Hensher ARC DP0770618.

1. Introduction

It is well known that a significant amount of human activity takes place within a group context in which the group not only becomes the primary agent for socialization and learning but also in affecting decision making and preference formation. Indeed, it is the household that represent the basic consumption unit for the majority of consumer goods purchases, both in terms of consumer durables and non-durables. How the interaction of individual group members influence the group's decision making and preference formation processes therefore represents an important dimension of our understanding of economic behaviour.

The study of group decision making has as its early roots, research undertaken in the field of social psychology (see e.g., Thorndike 1938). Since its inception, the study of group decision making has included research into such facets as individual behaviour in social contexts, the impact of within and between group interactions on group performance, and the identification and categorization of means of aggregating individual beliefs and preferences into collective group consensus (Baron et al. 1992, Arrow 1963).

Accepting the possibility of the existence of significant interaction effects between agents requires an acknowledgement of the fact that preference formation may be conditional upon the preferences of other agents present within an agent's cohort. A growing body of literature (mainly in marketing) has recognized this fact. Dellaert et al. (1998), Arora and Allenby (1999), Aribarg et al. (2002) and Beharry-Borg et al. (2009) are recent examples of attempts to incorporate the effects of both individual preferences and influences into group decision making.

Aribarg et al (2002) and Arora and Allenby (1999) make use of a hierarchical Bayes model to yield estimates of influence and preference amongst household members whilst Dellaert et al. (1998) and Hensher and Puckett (2008) utilize a two stage conjoint approach to elicit similar information. Both approaches provide the analyst with information on the degree of influence asserted on the preference formation of one agent by another agent. Whilst informative, the above approaches require that the input from an agent's cohort be treated as an exogenous variable.

Decision contexts involving interaction between multiple agents involve elements of both cooperation and non-cooperation. Both elements will be observed whether individual agents attempt to act as a single agent entity such as family members acting as a single household in the context of an automobile purchase; or as separate agent entities in competition with one another, such as a car salesperson attempting to sell a motor vehicle to a family. In both cases the preferences of individual agents may be in opposition, however a convergence of preferences (which may or may not be the goal of all agents present) may be achieved through a process of preference revision and concession (Aribarg et al. 2002, Hensher et al. 2008). The end stage of this process of revision and concession is that of an equilibrium state represented by either agreement (preference convergence) or disagreement (where preferences fail to converge).

Given a situation in which two or more agents interact to some degree in the determination of choice of alternative the possible outcomes are choice and non-choice agreement, and choice or non choice disagreement. Choice agreement, the result of cooperation amongst all parties, arises when all agents select the same alternative. Non choice agreement arising from limited cooperation amongst parties results in the simultaneous rejection of an alternative concurrent with non-cooperation as to the choice agreement of a single alternative. Non-choice agreement thus represents the removal of an alternative from the group's consideration set. Choice and non-choice disagreement represent the inverse positions. Earlier research by Hensher introduced the idea of Interactive Agency Choice Experiments (IACE) (Hensher and Chow 1999, Brewer and Hensher 2000, Hensher 2002,) in which a network of agents assess a common set of alternatives either sequentially or simultaneously. More recent research has provided a number of opportunities to explore empirically the bargaining power of group members in the joint choice setting, developing a range of methods (Dosman and Adamowicz 2006, Hensher and Puckett 2008, Beharry et al. 2009), as well as the interactive agency choice experiment model of feedback and revision of group preferences (Rose and Hensher 2004, Hensher et al. 2008).

Sitting alongside studies designed to understand group decision making behaviour are studies that use proxy responses provided by a single sampled respondent to reflect the behaviour and preferences of the group. Indeed, the vast majority of studies dealing with decisions that are rightly the domain of multiple agents typically sample the primary decision maker and assume that their preferences reflect those of others in the group. Early research examining the ability of single group members to provide information on the characteristics and observed behaviour of other group members has demonstrated that errors in reported responses are systematically related to the precise relationship between group members, the information being sought, the characteristics of the proxy, whether the proxy participated in events being surveyed as well as the survey administration method (see e.g., Cash and Moss 1974, Mathiowetz and Groves 1985, Rodgers and Herzog 1987, Rodgers et al. 1988, Moore 1988, Groves 1989, Bickart et al. 1990, Kojetin and Miller 1993, Bliemer and Lyberg 2003).

In terms of using proxy responses to understand group member preferences, research has found significant differences between individual and group preferences (Krishnamurthi 1988, Corfman and Lehmann 1993, Arora and Allenby 1999). More specifically, Dellaert et al. (1998) found that family members are relatively poor predictors of preference, and seem to project other family members' preferences along lines of their own. Arora and Allenby (1999) found that individual-specific attribute sensitivities do not capture group preferences adequately. Arora (2006) conducted a study in which respondents were asked to complete a series of choice tasks in order to determine their own individual preferences, the anticipated preference of others, the anticipated preferences of the group as a whole, and finally the true group preferences. The study found that individuals do poorly in assessing the preference of others, and perceptions of joint preference are different to actual joint preference. More recently, Bateman and Munro (2009) examined the differences in individual and group willingness to pay for reductions in dietary health risks and found significant differences between household and individual values.

These findings have led to a growing concern amongst researchers of the ability to use and rely on the individual responses as a proxy for capturing data on household decision making. Unfortunately, whilst the validity of using individual responses as a proxy for group choice, or for the choice of other individuals, has been questioned, and the transition of individual to group preferences, the ability of individuals to specifically assess the preferences of others is yet to be tested in a meaningful way. In response to this, this paper investigates choice of automobile, where group members are required to provide their own choice as well as a prediction of they think the other group member will choose. The accuracy of these predictions is examined and the salient attributes in errors of prediction are identified. The paper is structured as follows: Section 2 outlines the modelling methodology; Section 3 briefly comment on the sample, and present the empirical evidence for three models that assess differences in preferences; and Section 4 provides discussion and concluding remarks.

2. Methodology

For the current study, we collect stated choice (SC) data, not only on the choices made by individual respondents, but also on the choices those same respondents believe other household members will make given the same choice task. To analyse this data, discrete choice models are used. To understand the estimated models, let U_{ntj} denote the utility of alternative j perceived by respondent n in choice situation t. U_{ntj} consists of two components, a modelled component V_{ntj} and an unobserved component ε_{ntj}, such that

$$U_{ntj} = V_{ntj} + \varepsilon_{ntj}. \tag{1}$$

As is common practice, we assume the modelled component of utility to be represented as a linear relationship of k attributes, x, related to each of the j alternatives and corresponding parameters weights such that

$$U_{ntj} = \sum_{k=1}^{K} \beta_{nk} x_{ntjk} + \varepsilon_{ntj}, \tag{2}$$

where β_{nk} represents the marginal utility or parameter weight associated with attribute k for respondent n and the unobserved component, ε_{nsj}, is assumed to be independently and identically (IID) extreme value type 1 (EV1) distributed. As well as containing information on the levels of the attributes, x may also contain up to J-1 alternative specific constants (ASCs) capturing the residual mean influences of the unobserved effects on choice associated with their respective alternatives; where x takes the value 1 for the alternative under consideration or zero otherwise.

Given that we are interested in establishing estimates of WTP, we further assume that Equation (2) is separable in price, c_{ntj} and other non price attributes x_{ntjk}, such that Equation (2) may be rewritten as

$$U_{ntj} = \beta_{nc} c_{ntj} + \sum_{k=1}^{K} \beta_{nk} x_{ntjk} + \varepsilon_{ntj}, \tag{3}$$

The marginal willingness to pay for attribute k may then be calculated as

$$WTP = \frac{\dfrac{d}{dx_{ntjk}} \beta_{nk} x_{ntjk}}{\dfrac{d}{dc_{ntj}} \beta_{nc} c_{ntj}} = \frac{\beta_{nk}}{\beta_{nc}}. \tag{4}$$

In writing out the utility function as we have in Equations (2) and (3), the subscript n associated with the parameter weights implies a particular econometric model form will be estimated. In this case, and under the IID EV1 error term assumption, the utility function shown in Equation (2) implies the use of the mixed multinomial logit (MMNL) model specification framework. The MMNL model allows for the analyst to specify that some or all of the parameter weights estimated be allowed to vary over the sampled population with density $f(\beta_{nk} \mid \Omega)$.Note that if a parameter is to be treated as non-random, the subscript n will simply cease to be associated with that parameter, as the parameter will be fixed or constant across individuals.

The utility specification in Equation (3) is flexible in that it allows for the possibility that different respondents may have different marginal utilities for each attribute being modelled. Unfortunately, in practice it is not generally feasible to estimate individual specific parameter weights. As such, it is typical to estimate parameter weights for the population moments of the sample, such that

$$\beta_{nc} = \left(\bar{\beta}_c + \sigma_c \omega_n\right) \tag{5}$$

where $\bar{\beta}_c$ represents the mean of the parameter distribution, σ_c represents the standard deviation of preferences (or deviation from the mean) over the sampled population, and w_n random draws from a standard normal distribution. Likewise, non-price parameters that are treated as random parameters are estimated as $\beta_{nk} = \left(\bar{\beta}_k + \Gamma v_n\right)$ where $\bar{\beta}_k$ represents the mean of the parameter distribution, Γ a lower triangular Cholesky Matrix and v_n random draws over the sampled population with covariance matrix I, so that $\mathrm{var}(\beta) = (\Gamma\Gamma')$. The vast majority of studies assume Γ to be a diagonal matrix and hence treat the estimated parameter distributions as being independent. Such assumptions are not strictly necessary with models allowing for off-diagonal values in Γ to have complex correlation structures across the parameter distributions. Further, the random parameters need not be multivariate normally distributed, however the non diagonal Cholesky Matrix assumes such a distribution. As such, studies that assume non-normal random parameters typically also assume that Γ is a diagonal matrix, although approximations of the Cholesky decomposition process have been applied to non Normally distributed random parameters in the past (see e.g., Greene and Hensher 2009).

Equation (3) is defined in 'preference space' (see Train and Weeks 2005, Sonnier et al. 2007 or Scarpa et al. 2008). It is possible to re-specify the utility function so as to estimate the WTP estimates directly. To do this, we rewrite Equation (3) as follows.

$$\begin{aligned} U_{ntj} &= \beta_{nc}\left[c_{ntj} + \frac{1}{\beta_{nc}}\sum_{k=1}^{K}\beta_{nk}x_{ntjk}\right] + \varepsilon_{ntj} \\ &= \beta_{nc}\left[c_{ntj} + \sum_{k=1}^{K}\theta_{nk}x_{ntjk}\right] + \varepsilon_{ntj}. \end{aligned} \tag{6}$$

In this case, the cost parameter, β_{nc}, simply becomes a normalising constant in the WTP representation or viewed another way, a scaling factor for the WTP estimates θ_{nk}.

In estimating the model, β_{nc} takes the form $\beta_{nc} = \bar{\beta}_c e^{\left(\frac{-\tau^2}{2}+\tau\omega_n\right)}$ where τ represents a variance parameter influencing scale of β_{nc}. The transformation insures that the sign of the parameter will be constrained, with no *a priori* expectation, as taking an exponential of any value will always produce a positive value, whilst $\bar{\beta}_c$ may take any sign. In estimation, depending on the estimate of τ, extremely large values of β_{nc} can occur depending on the values drawn from w_n. When such large values are observed, software overflows may occur and the estimator becomes unstable. As such, rather than use a standard Normal distribution for w_n, Fiebig et al. (2009) employ truncated standard Normal distribution with truncation at ±2. In taking this approach, any draw from outside this range is rejected and a new draw taken in its place. Rather than use a truncated standard Normal distribution with an acceptance/rejection of the random parameter draw approach, Greene and Hensher (2009) propose a method to directly restrict the values of w_n to be between ±1.96. This is achieved by setting $w_{nr} = \Phi^{-1}\left[0.025 + 0.95U_{nr}\right]$ where the value of w_n for the r^{th} draw is calculated from the inverse of the standard normal cumulative distribution function, $\Phi^{-1}[.]$ given a random draw from a standard uniform distribution bounded by 0 and 1. Assuming that draw *r* from U_{nr} is 0, then the probability is transformed to 0.025. At the other bound, a draw of 1 for U_{nr} corresponds to a probability of 0.975. As such, $\Phi^{-1}[.]$ will be naturally bounded at ±1.96. For the current paper, we utilize the approach suggested by Greene and Hensher (2009).

3. Empirical data

Data for the current study were collected as part of a project involving first and second year marketing undergraduate students at the University of Sydney in 2007. As part of the project, each student was required to recruit two participants from the same household (not necessarily from their own household) and administer an online survey. The central component of the online survey was an unlabelled SC experiment examining the use of motor vehicle choice. Vehicle choice was viewed as a purchase decision that would address many of the concerns in previous studies. Within the sample, all respondents currently use a motor vehicle and own or have access to a car. Additionally, a car is used by multiple stakeholders with differing purposes and, typically, a vehicle purchasing decision is a considered one where the preferences of other users are known and need to be accommodated due to the communal nature of the vehicles use. The participants in the experiment are also peers, meaning communication between them would be higher and preferences more likely to be closer together than individuals outside a participant's reference group.

It is reasonable to assume that each agent has a varying amount of knowledge about the preferences of others, and this knowledge can be placed on a continuum. At one extreme there are individuals who have no knowledge of member preferences, and at the other end those with full (perceived) knowledge of the preferences of others. To examine where respondents lie on this continuum, and thus test the ability of individuals to act as proxy respondents in a stated choice experiment, individuals were presented with four choice sets consisting of three unlabelled vehicle alternatives and a no choice alternative. To negate response bias engendered by the presence of other agents, separation between respondents was maintained. The first respondent's initial preferences were collected along with what they thought the other agent in the dyad would select. Once the choice task was completed, this person was asked to leave the room and was replaced by the second respondent who completed the same choice task undertaken by the first individual.

Prior to commencing, respondents were asked a series of questions relating to their currently owned vehicle, and this information was then used to assign agent pairs to different survey segments related specifically to the size of their current vehicle (small, medium, or large). This thus provided the respondents with a frame of reference when undertaking each of the choice tasks. The attributes and attribute levels of the stated choice experiment are shown in Table 1.

Table 1: Experimental design attributes

Attributes	Attribute Levels	
Engine Size	Small	(1.2; 1.3; 1.4; 1.5)
	Medium	(1.6; 1.8; 2.0; 2.2)
	Large	(2.3; 2.9; 3.4; 4.0)
Price	Small	($12,000; $13,500; $15,000; $16,500)
	Medium	($19,990; $21,990; $23,990; $25,990)
	Large	($28,000; $30,000; $32,000; $34,000)
Fuel Efficiency *(litres per 100km)*	Small	(6.2; 6.7; 7.4; 7.7)
	Medium	(7.6; 8.1; 8.5; 9.0)
	Large	(8.8; 9.8; 10.7; 11.7)
ABS	Yes	(1)
	No	(0)
Air-conditioning	Yes	(1)
	No	(0)
Transmission	Manual	(1)
	Automatic	(0)

Within the data, it was found that in 59 percent of the choice sets in the sample, the choice predicted by a respondent coincided with their own choice. Such high correlation is important as it is reasonable to assume that respondents who believe that their own preferences and the preferences of others align, will exhibit significant bias in terms of their ability to predict the choice of others. When comparing respondents' own choices to the actual choice of the other person, the choice intersects 52 percent of the time. These results indicate that while respondents acknowledge different preferences, they slightly under estimate the extent of those differences.

4. Model results

4.1 Own versus proxy choice

A number of models are estimated as part of the current study. Table 2 presents the first model which parameterises directly into WTP space, the sampled respondents own WTP for the various SC attributes alongside the proxy WTPs obtained from the same set of respondents. These proxy WTP values reflect the interviewed respondent's belief as to what the other member of their dyad would be willing to pay for each of the attributes modelled. In estimating this model, own and proxy choices were modelled simultaneously.

In estimation, all random WTP parameter estimates were specified using Normal distributions employing 500 Halton draws. Whilst models allowing for correlations amongst the random WTP parameters were tested, the final model reported in Table 2, which represents the best fitting model, does not directly account for correlation amongst the set of random parameters. Also, in estimating the model, the best fitting model was found to account for the pseudo panel nature of the SC data in the modelling process (see Train 2003).

Table 2: Model results for own versus proxy WTP

	Own Choice				Proxy Choice				
	Par.	(*t*-ratio)	Lower 95%	Upper 95%	Par.	(*t*-ratio)	Lower 95%	Upper 95%	(*t*-ratio of diff.)
Random Parameters									
Transmission (mean)	-24.373	(-4.06)	-36.150	-12.595	-18.652	(-3.67)	-28.625	-8.679	-0.73
Transmission (std dev.)	25.133	(3.69)	11.765	38.501	28.152	(3.45)	12.137	44.167	-0.28
Fuel Efficiency (mean)	3.225	(2.77)	0.942	5.507	4.069	(2.76)	1.180	6.958	-0.45
Fuel Efficiency (std dev.)	0.322	(0.41)	-1.206	1.850	1.072	(2.12)	0.081	2.062	-0.81
ABS brakes (mean)	8.918	(3.05)	3.194	14.642	10.993	(3.42)	4.687	17.300	0.48
ABS brakes (std dev.)	12.305	(3.15)	4.638	19.972	14.693	(2.58)	3.528	25.859	-0.35
Engine Capacity (mean)	-5.677	(-2.58)	-9.984	-1.371	-3.684	(-1.87)	-7.553	0.185	-0.67
Engine Capacity (std dev.)	6.653	(2.63)	1.700	11.606	0.194	(0.09)	-4.196	4.584	1.92
Air Conditioning (mean)	18.822	(-4.53)	10.674	26.969	24.169	(-3.87)	11.939	36.399	0.71
Air Conditioning (std dev.)	4.730	(1.29)	-2.466	11.927	9.427	(2.21)	1.056	17.797	-0.83
Non Random Parameters									
Constant Alt. A	2.567	(3.90)	1.278	3.855	4.468	(4.71)	2.608	6.328	-1.65
Constant Alt. B	2.931	(4.50)	1.654	4.208	4.889	(5.19)	3.042	6.735	-1.71
Constant Alt. C	2.599	(3.89)	1.291	3.907	4.693	(4.91)	2.819	6.568	-1.80

	Par.	(*t*-ratio)	Lower 95%	Upper 95%
Parameter for Cost (WTP space)				
Cost (mean)	-0.227	(-2.99)	-0.376	-0.078
Cost (std dev.)	0.105	(2.85)	0.033	0.177
Scale Parameter				
Variance Parameter in Scale (τ)	1.203	(4.76)	0.707	1.698
Sigma Parameter				
Sample Mean	0.788	-	-	-
Sample Std Dev.	0.837	-	-	-
Model Fits				
LL(0)	-2362.246			
LL(β)	-1231.010			
$\rho(0)$	0.479			
Adj. $\rho(0)$	0.465			
Number of Respondents	370			
Number of Observations	1136			

The overall fit of the final model is good with an adjusted ρ^2 of 0.465 relative to a base model estimated under the null parameter hypothesis. Examination of the WTP estimates reveals that the means of the WTP distributions are of the expected signs for both the own and proxy choices. Comparing the 95 percent confidence intervals for own versus proxy WTP, suggests that the two models closely align.

Examining the individual attribute related parameters, as expected, the model indicates that individuals are willing to pay for more fuel efficient cars, and that this preference is largely consistent with what the individual would expect the other party to be willing to pay. Similar effects are found for the ABS, Air Conditioning and Engine Size attributes which all have a significant role in determining an individual's own WTP value as well as what they predict will be the WTP for others; however the significant spreads around these parameters indicate that the range of WTP values is large. When looking at the values for the standard deviation parameters, we note greater heterogeneity exists for the proxy choices for all parameters except for engine capacity. It is interesting to note that in determining an individual's own choice engine size is significant, with significant heterogeneity in preferences. However, with respect to proxy responses, it is found to be insignificant in its mean influence with preferences being homogeneous. In conjunction with this finding, the proxy model predicts significant variance in preferences for fuel efficiency and air conditioning, which is absent in the own choice model. Table 2 also provides the *t*-ratios for the parameter differences between the proxy and own choices. Based on these *t*-ratios, it is clear that in the aggregate, the proxy WTP values are very similar to the respondent's actual own WTP values.

Examining the scale parameter (i.e., τ) of the model reveals that it is highly significant. This suggests that scale heterogeneity exists within the combined data sets. To breakdown this observed scale heterogeneity, we further tested models that allowed for correlation between the random scale term and the random parameters as specifications assuming uncorrelated WTP imply a pattern of correlation in utility coefficients that is difficult to implement in preference space (Scarpa et al. 2008); however the correlations were found not to be statistically significant and were removed from the model. Although an imperfect test, the mean values of the conditional WTP distributions were calculated and tests of correlation between the own and proxy choices were conducted. Given the lack of correlation structure allowed for in the final estimated model, it is not surprising to find very low correlations between the mean values of the own versus proxy conditional parameter distributions. The correlations are shown in Table 3. In the sample, own choice coincided with proxy choice in 59 percent of the choice sets in the sample, however the correlation structure presents contrary evidence to any assumed relationship between the parameters.

Table 3: Correlations of own versus proxy mean WTP conditional parameter distributions

		Own Choice					Proxy Choice				
		Trans.	Fuel Eff.	ABS	Eng. Cap.	Air Cond.	Trans.	Fuel Eff.	ABS	Eng. Cap.	Air Cond.
Own Choice	Trans.	1.00									
	Fuel Eff.	-0.06	1.00								
	ABS	-0.18	0.07	1.00							
	Eng. Cap.	0.05	0.27	0.10	1.00						
	Air Cond.	-0.04	0.39	0.00	0.46	1.00					
Proxy Choice	Trans.	**-0.02**	-0.09	0.01	-0.04	0.04	1.00				
	Fuel Eff.	-0.05	**0.18**	-0.07	0.08	0.18	0.00	1.00			
	ABS	0.02	0.04	**0.06**	-0.03	0.16	0.03	0.13	1.00		
	Eng. Cap.	0.02	0.08	0.12	**0.08**	-0.11	-0.29	0.02	-0.21	1.00	
	Air Cond.	0.02	0.15	-0.02	0.01	**-0.03**	-0.18	0.46	-0.07	0.06	1.00

4.2 Proxy versus real choice

Table 4 presents the results for a second model comparing the proxy WTP to the WTP of the group of respondents whom the proxy choices are supposed to represent. As with the original model, a single model was estimated on both data simultaneously. As with the first model, the WTPs for each design attribute were treated as random parameters drawn from a multivariate Normal distribution. Five hundred Halton draws were used in the estimation procedure, and the model was estimated to account for the pseudo panel nature of the SC data.

Table 4: Model results for proxy versus real WTP

	Proxy Choice				Real Choice				
	Par.	(*t*-ratio)	Lower 95%	Upper 95%	Par.	(*t*-ratio)	Lower 95%	Upper 95%	(*t*-ratio of diff.)
Random Parameters									
Transmission (mean)	-16.943	(-3.14)	-27.504	-6.381	-22.512	(-3.48)	-35.198	-9.825	(0.66)
Transmission (std dev.)	26.462	(3.50)	11.649	41.275	31.949	(3.53)	14.203	49.695	(-0.47)
Fuel Efficiency (mean)	3.920	(2.65)	1.021	6.818	4.617	(2.75)	1.320	7.913	(-0.31)
Fuel Efficiency (std dev.)	1.737	(1.99)	0.024	3.451	0.451	(0.46)	-1.457	2.358	(0.98)
ABS brakes (mean)	11.954	(3.33)	18.994	4.914	9.569	(2.85)	16.139	2.998	(-0.49)
ABS brakes (std dev.)	14.393	(3.06)	5.179	23.606	10.577	(2.32)	1.623	19.531	(0.58)
Engine Capacity (mean)	-4.124	(-1.98)	-8.217	-0.031	-5.230	(-1.86)	-10.729	0.270	(0.32)
Engine Capacity (std dev.)	2.701	(0.92)	-3.080	8.482	8.518	(2.92)	2.796	14.239	(-1.40)
Air Conditioning (mean)	21.985	(3.67)	33.714	10.255	19.210	(3.65)	29.536	8.885	(-0.35)
Air Conditioning (std dev.)	11.120	(2.40)	2.034	20.206	14.242	(2.79)	4.224	24.259	(-0.45)
Non Random Parameters									
Constant Alt. A	4.207	(3.49)	1.84	6.57	4.607	(4.00)	2.35	6.86	(-0.24)
Constant Alt. B	4.627	(3.92)	2.32	6.94	4.874	(4.18)	2.59	7.16	(-0.15)
Constant Alt. C	4.437	(3.67)	2.07	6.81	4.638	(3.99)	2.36	6.92	(-0.12)

	Par.	(*t*-ratio)	Lower 95%	Upper 95%
Parameter for Cost (WTP space)				
Cost (mean)	-0.136	(-3.96)	-0.376	-0.078
Cost (std dev.)	0.098	(3.85)	0.033	0.177
Scale Parameter				
Variance Parameter in Scale (τ)	0.289	(1.15)	-0.206	0.784
Sigma Parameter				
Sample Mean	0.984	-	-	-
Sample Std Dev.	0.247	-	-	-
Model Fits				
LL(0)	-2362.246			
LL(β)	-1220.939			
$\rho(0)$	0.483			
Adj. $\rho(0)$	0.47			
Number of Respondents	370			
Number of Observations	1136			

The overall fit of the model is similar to that of the first model, with an adjusted ρ^2 of 0.470. Similar to the first model, correlations between the random parameters as well as between the random parameters and the scale parameter were tested, however these models were rejected as they did not lead to statistically better model fits. Unlike the first model however, the scale parameter of the model is not statistically significant, whereas the majority of the random parameters are. The signs of the means of the WTP parameter distributions for both the proxy and real choices are of the expected signs and are all statistically significant at the 95 percent level.

Examining the confidence intervals around parameter estimates suggests that in the aggregate, the proxy choices are able to predict the real WTPs of the respondents the proxies are supposed to represent. Nevertheless, the mean WTP parameter estimates appear to be quite different for the engine capacity and transmission attributes. The *t*-ratios of the parameter differences however confirm that these differences are not statistically significant. Although not statistically significant, it is interesting to note that on average, there appear to be differences in the level of heterogeneity between the proxy and real WTP estimated values. With respect to fuel efficiency, the proxy responses indicate a significant spread of preferences for this variable where in relative the standard deviation for this estimate is not significant. Discrepancies are also apparent with the engine capacity attribute, where the proxy responses anticipate that this attribute will be a significant influence on taste with no significant preference heterogeneity where in reality the overall impact of engine size was insignificant, with significantly large variation of preferences around this estimate. Interestingly, the standard deviations for the transmission and air conditioning WTP parameter distributions for the real choice are, in the aggregate, larger than those for the proxy choices, indicating that respondents may have underestimated the differing preferences that exist for these attributes.

Again, we provide the correlation matrix for the mean values taken from the conditional distributions. Once more, the correlations of the mean values of the conditional parameter distributions reveal very limited correlations between the proxy and real WTP values. As with the first model, it must be acknowledged that part of this is possibly due to the fact that correlations between the random parameters have not been accounted for in the model; however it is worth noting here that, unlike the first model, the proxy choices were not made by the same individuals who made the observed choices (i.e., those whom the proxy choices were meant to represent). As such, low correlation coefficients are far more likely to be expected from this model.

Table 5: Correlations of proxy versus real mean WTP conditional parameter distributions

		Proxy Choice					Real Choice				
		Trans.	Fuel Eff.	ABS	Eng. Cap.	Air Cond.	Trans.	Fuel Eff.	ABS	Eng. Cap.	Air Cond.
Proxy Choice	Trans.	1.00									
	Fuel Eff.	0.00	1.00								
	ABS	-0.04	0.17	1.00							
	Eng. Cap.	-0.07	0.71	0.15	1.00						
	Air Cond.	-0.18	0.47	-0.02	0.33	1.00					
Real Choice	Trans.	**0.03**	0.04	-0.05	-0.01	-0.09	1.00				
	Fuel Eff.	-0.05	**-0.13**	0.09	-0.13	-0.15	0.14	1.00			
	ABS	0.04	0.00	**0.14**	0.04	-0.10	-0.03	0.24	1.00		
	Eng. Cap.	-0.03	0.03	-0.03	**0.01**	-0.08	0.13	0.58	0.13	1.00	
	Air Cond.	-0.02	-0.05	0.01	-0.04	**-0.06**	-0.09	0.24	-0.11	0.36	1.00

4.3 Proxy versus incorrectly predicted real choices

The models reported in Sections 4.2 and 4.3 examined the proxy choices against the choices made by the person providing the proxy response as well as against the actual choices made by the person whom the proxy was meant to represent. In estimating these models, all available data were used including choice observations where the proxy choices incorrectly predicted the actual outcome. Table 6 presents a model where separate parameters are estimated for those choice tasks in which the proxy choice correctly predicted the real choice (this occurred in 66 percent of choice tasks) as well as for choice tasks in which the proxy choice incorrectly predicted the real choice. As with the first two models, this model was jointly estimated using 500 Halton draws.

Table 6: Model results for correctly predicted versus incorrectly predicted choice tasks

	Correctly Predicted Choice				Incorrectly Predicted Choice				
	Par.	(t-ratio)	Lower 95%	Upper 95%	Par.	(t-ratio)	Lower 95%	Upper 95%	(t-ratio of diff.)
Random Parameters									
Transmission (mean)	-31.229	(-2.97)	-51.833	-10.625	-3.983	(-1.28)	-10.065	2.100	(-2.49)
Transmission (std dev.)	39.547	(2.88)	12.666	66.428	12.056	(2.14)	0.999	23.113	(2.43)
Fuel Efficiency (mean)	4.789	(2.23)	0.584	8.993	-0.985	(-0.96)	-2.997	1.027	(-1.42)
Fuel Efficiency (std dev.)	1.389	(0.96)	-1.463	4.242	0.809	(1.17)	-0.552	2.171	(-1.97)
ABS brakes (mean)	-13.089	(-2.77)	-22.342	-3.836	-5.304	(-1.91)	-10.748	0.141	(-2.05)
ABS brakes (std dev.)	12.298	(2.43)	2.363	22.234	12.016	(2.23)	1.458	22.573	(1.85)
Engine Capacity (mean)	-7.777	(-2.45)	-14.008	-1.546	0.376	(0.14)	-4.802	5.555	(0.36)
Engine Capacity (std dev.)	7.768	(2.01)	0.183	15.354	5.268	(0.87)	-6.573	17.109	(0.04)
Air Conditioning (mean)	-29.320	(-3.15)	-47.569	-11.072	-9.101	(-2.78)	-15.528	-2.673	(0.35)
Air Conditioning (std dev.)	15.383	(2.26)	2.012	28.753	4.153	(0.69)	-7.639	15.946	(1.23)
Non Random Parameters									
Constant Alt. A	3.682	(2.28)	0.523	6.841	4.521	(1.77)	-0.474	9.517	(-0.28)
Constant Alt. B	4.301	(2.67)	1.138	7.463	4.884	(1.92)	-0.095	9.863	(-0.19)
Constant Alt. C	3.979	(2.37)	0.694	7.265	4.878	(1.90)	-0.160	9.915	(-0.29)

	Par.	(t-ratio)	Lower 95%	Upper 95%
Parameter for Cost (WTP space)				
Cost (mean)	-0.279	(-2.61)	-0.376	-0.078
Cost (std dev.)	0.198	(2.53)	0.033	0.177
Scale Parameter				
Variance Parameter in Scale (τ)	0.719	(2.11)	0.052	1.386
Sigma Parameter				
Sample Mean	0.918	-	-	-
Sample Std Dev.	0.575	-	-	-
Model Fits				
LL(0)	-1181.123			
LL(β)	-591.199			
$\rho(0)$	0.499			
Adj. $\rho(0)$	0.473			
Number of Respondents	290			
Number of Observations	568			

Examining the differences in the parameter estimates across the correctly predicted and incorrectly predicted choice tasks, it is clear that statistically significant differences exist for the transmission and ABS parameters, thus suggesting that these two parameters were largely the source of proxy error. Interestingly, there also exists a statistically significant difference in the WTP heterogeneity for the transmission attribute, with less heterogeneity in the incorrectly predicted choice observations than in the correctly predicted choice tasks. This suggests that over the sampled population, there exists significant WTP variation for the transmission attribute, however this variation was not correctly accounted for in providing the proxy choice prediction. Moreover, with the exception of air conditioning, none of the mean estimates are significant in the incorrectly predicted outcome when in the correct responses all attributes have a significant influence.

5. Discussion and conclusions

In this paper, we investigate the accuracy of proxy respondents in a stated choice experiment, where the nature of the task was to predict vehicle choice. Given that all individuals sampled have experience with driving motor vehicles, that all respondent dyads were identified as being joint decision makers and part of a household group, that choosing a motor vehicle carries a large financial cost and the typical lifecycle of a vehicle post purchase is protracted relative to other consumer items, it is not an unrealistic assumption to expect that peoples preferences for types of vehicles and their inherent features would be known. Despite this, we found that the overall ability of individuals to predict the choice of others is poor, with incorrect predictions being made 40 percent of the time.

To better understand the source of errors in the prediction of choice, three separate models were estimated. Initially we examined how closely a respondent's own preferences coincided with what they predicted the other person would select. Whilst the proxy response elicited from an individual coincided with that individual's own choice in 59 percent of choice sets, the real choice intersected 52 percent of the time, suggesting a tendency for people to overestimate the degree of similarity in preferences. Willingness to pay space was used to examine the different role each attribute plays and examining 95 percent confidence intervals around the estimates reveals an alignment between proxy responses and the own choice made by the individual providing the proxy. While engine size was found to be a significant determinant in a respondent's own choice, the proxy model revealed an anticipation that it would not be a significant determinant of choice and that preferences around this would not vary. Generally wider standard deviation parameters in the proxy model show a more varied set of preferences than were revealed in the own choice model, perhaps indicating a degree of uncertainty as to the role these attributes would play in attempting to predict the choice of others relative to a respondents own choice.

After examining the alignment between individual's proxy response and their own choice, the accuracy of the proxy responses was investigated. Whilst in aggregate the proxy responses seem to predict the real WTPs, some discrepancies exist. The proxy model predicts that engine capacity would be significant with no significant variation in preferences; however in reality it was insignificant though preferences were significantly heterogeneous. The proxy model also suggests significant variation in the preference for fuel efficiency; however this was not supported by the real choice model. Generally, the real choices made by respondents exhibit a greater degree of preference heterogeneity than the proxy response indicate, suggesting that errors in prediction are most likely linked to an underestimation of the degree of preference variation exhibited by others, or an inability to predict the position of others on such wide spectrum of preferences, arguably a difficult task, which is contributing to the lack of precision in the proxy responses.

To further examine the sources of error in prediction, the choice sets where the proxy responses aligned with the real choice were compared to the choice sets in which the proxy responses were incorrect. This model revealed the transmission and ABS attributes as the main sources of confusion, with significant differences in these parameters across the two data sources. Moreover, the general lack of significant parameters when looking at instances where the proxy responses was incorrect suggests a great degree of randomness in these choice sets, perhaps indicating that when confronted with a choice that an individual could not predict a simple guess was provided rather than a systematic evaluation of the attributes presented.

Taken as a whole, the similarities between the choice made by an individual and the proxy response that person provides suggests that when providing the proxy response an individual will use their own preferences as a starting point and attempt to adjust their choices from this anchor. This "anchoring and adjustment" is a decision making heuristic first theorised by Tversky and Kahneman (1974) with an inherent bias in that the "anchor" is relied upon too heavily in the decision process and adjustment from this value is insufficient (Epley and Gilovich 2006). In this instance, by believing that a person's choice will be, to a large extent, similar to your own and referencing your own choice as a starting point for predicting the choice of others introduces a bias into the proxy response, particularly if the sensitivities to the attributes that are thought to be similar are misunderstood.

In comparing the proxy choice to the real choice, the generally smaller standard deviation parameters for the model estimated on the proxy choices compared to model using the real choice reveal a potential second decision making bias. “Overconfidence” is a bias in which individuals overestimate the accuracy of their estimates (Gilovich et al. 2002). In this instance it appears as if the proxy responses were too restrictive in the determination of the degree of heterogeneity present in the sample.

Lastly, in comparing the instances where the proxy is correct to where it is incorrect reveals the underlying uncertainty in these responses. When the proxy response was correct, the real choice coincided with a respondents own choice 66 percent of the time. However, when the proxy response was incorrect the level of coincidence dropped to 47 percent. Combined with the general lack of significant parameters for the model on the incorrect proxy responses, these results indicate that when confront with situations in which another individuals preferences differ markedly from their own, respondents are unable to provide informed responses.

In summary, by establishing a task where respondents are required to specifically attempt to predict the choice of others, using motor vehicle choice as a case study where preferences and partners are well known to respondents engaged in the task, we redress weakness inherent in earlier studies. Unlike Dellaert et al. (1998), however, we find that respondents are able to discern that other individuals have differing preferences and that they attempt to provide estimates of these preferences rather than projecting their own choice. Despite this finding our conclusions are similar to those found elsewhere (Menon et al. 1995, Arora and Allenby 1999, Arora 2006): the use of proxy respondents can lead to biased results as individuals are unable to fully synthesise the preferences of others, even when explicitly asked to. This finding brings into question the reliability of proxy responses as a sampling method and consequently the use of such data in estimating preferences.

References

Aribarg, A., Arora, N. and Bodur, H. O. (2002). "Understanding the Role of Preference Revision and Concession in Group Decisions." *Journal of Marketing Research* 39(3): 336-349.

Arora, N. (2006). "Estimating Joint Preference: A Sub-Sampling Approach." *International Journal of Research in Marketing* 23: 409-418.

Arora, N. and Allenby, G. M. (1999). "Measuring the Influence of Individual Preference Structures in Group Decision Making." *Journal of Marketing Research* 36(4): 476-487.

Arrow, K. J. (1963). *Social Choice and Individual Values*. New Haven, Yale University Press.

Baron, R. S., Kerr, N. L. and Miller, N. (1992). *Group Processes, Group Decision, Group Action*. Pacific Grove, CA, Brooks Cole.

Bateman, I. J. and Munro, A. (2009). "Household Versus Individual Valuation: What's the Difference." *Environmental and Resource Economics* 43(1): 119-135.

Beharry-Borg, N., Hensher, D. A. and Scarpa, R. (2009). "An Analytical Framework for Joint vs Separate Decisions by Couples in Choice Experiments: The case of coastal water quality in Tobago." *Environmental and Resource Economics* 43(1): 95-117.

Brewer, A. and Hensher, D. (2000). "Distributed Work and Travel Behaviour: The Dynamics of Interactive Agency Choices Between Employers and Employees." *Transportation* 27(1): 117-148.

Corfman, K. and Lehmann, D. (1993). "The Importance of Others' Welfare in Evaluating Bargaining Outcomes." *Journal of Consumer Research* 20(1): 124-137.

Dellaert, B. G. C., Prodigalidad, M. and Louviere, J. J. (1998). "Family Members' Projections of Each Other's Preference and Influence: A Two-Stage Conjoint Approach." *Marketing Letters* 9(2): 135-145.

Dosman, D. and Adamowicz, W. (2006). "Combining Stated and Revealed Preference Data to Construct an Empirical Examination of Intrahousehold Bargaining." *Review of Economics of the Household* 4(1): 15-34.

Epley, N. and Gilovich, T. (2006). "The Anchoring-and-Adjustment Heuristic: Why the adjustments are insufficient." *Pyschological Science* 17(4): 311-318.

Fiebig, D., Keane, M., Louviere, J., and Wasi, N. (2009) The generalized multinomial logit: accounting for scale and coefficient heterogeneity, *Marketing Science*, published online before print July 23, DOI:10.1287/mksc.1090.0508

Gilovich, T., Griffin, D. and Kahneman, D. (2002). *Heuristics and Biases: The psychology of intuitive judgment*. Cambridge, Cambridge University Press.

Greene, W. and Hensher, D. A. (2009). "Does Scale Heterogeneity Across Individuals Matter? A Comparative Assessment of Logit Models." submitted to *Transportation* August 29, 2009.

Hensher, D. and Chow, G. (1999). Interacting Agents and Discrete Choices in Logistics Outsourcing: A Conceptual Framework. *World Transport Research: Transport Modelling/Assessment*. H. Meersman, E. van de Voorde and W. Winkelman. Oxford, Perfamon Press. 3.

Hensher, D. A. (2002). "A Systematic Assessment of the Environmental Impacts of Transport Policy: An end use perspective." *Environmental and Resource Economics* 22(1-2): 185-217.

Hensher, D. A. and Puckett, S. M. (2008). "Power, Concession and Agreement in Freight Distribution Chains subject to Distance-Based User Charges (UGM Paper #11)." *International Journal of Logistics: Research and Applications* 11(2): 81-100.

Hensher, D. A., Rose, J. M. and Black, I. (2008). "Interactive Agency Choice in Automobile Purchase Decisions: The role of negotiation in determining equilibrium choice outcomes." *Journal of Transport Economics and Policy* 42(2): 269-296.

Krishnamurthi, L. (1988). "Conjoint Models of Family Decision Making." *International Journal of Research in Marketing* 5(3): 185-198.

Menon, G., Bickart, B., Sudman, S. and Blair, J. (1995). "How Well Do You Know Your Partner? Strategies for formulating proxy-reports and their effects on the convergence to self-reports." *Journal of Marketing Research* 32: 75-84.

Richardson, A. J. (2006). "Proxy Responses in Self-Completion Travel Diary Surveys." *Transportation Research Record*(1972): 1-8.

Rose, J. M. and Hensher, D. A. (2004). "Modelling Agent Interdependency in Group Decision Making: Methodological Approaches to Interactive Agent Choice Experiments." *Transportation Research Part E* 40(1): 63-79.

Scarpa, R., Thiene, M. and Train, K. (2008). "Utility in WTP Space: A tool to address confounding random scale effects in destination choice to the Alps." *American Journal of Agricultural Economics* 90(40): 994-1010.

Sonnier, G., Ainslie, A. and Otte, T. (2007). "Heterogeneity Distributions of Willingness-To-Pay in Choice Models." *Quantitative Marketing and Economics* 5(3): 313-331.

Thorndike, R. L. (1938). "On What Type of Task Will a Group do Well?" *Journal of Abnormal and Social Psychology* 33: 409-413.

Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge, Cambridge University Press.

Train, K. and Weeks, M. (2005). Discrete Choice Models in Preference Space and Willing-To-Pay Space. *Applications of Simulation Methods in Environmental and Resource Economics*. R. Scarpa and A. Alberini. Dordrecht, Springer: 1-16.

Tversky, A. and Kahneman, D. (1974). "Judgment Under Uncertainty: Heuristics and biases." *Science* 185: 1124-1130.

Wargelin, L. and Kostyniuk, L. (2004). Proxy Respondents in Household Travel Surveys. *Travel Survey Methods: Quality and Future Directions*. P. Stopher and C. Stecher, Emerald Group Publishing: 201-212.