Biased estimates in discrete choice models: the appropriate inclusion of psychometric data into the valuation of recycled wastewater

Fiona L. Gibson

And

Michael Burton

A Contributed Paper to the Australian Agricultural & Resource Economics Society's Annual Conference, Cairns, February 11-13, 2009.

Biased estimates in discrete choice models: the appropriate inclusion of psychometric data into the valuation of recycled wastewater

Fiona L. Gibson and Michael Burton

School of Agricultural and Resource Economics

University of Western Australia

Abstract

The introduction of measurement bias in parameter estimates into non-linear discrete choice models, as a result of using factor analysis, was identified by Train *et al.* (1987). They found that the inclusion of factor scores, used to represent relationships amongst like variables, into a subsequent discrete choice models introduced measurement bias as the measurement error associated with each factor score is excluded. This is an issue for non-market valuation given the increase in popularity of including psychometric data, such as primitive beliefs, attitudes and motivations, in willingness to pay estimates. This study explores the relationship between willingness to pay and primitive beliefs through a case study eliciting Perth community values for drinking recycled wastewater. The standard discrete decision model, with sequential inclusion of factor scores, is compared to an equivalent discrete decision model, which corrects for the measurement bias by simultaneously estimating the underlying latent variables using a measurement model. Previous research has focused on the issue of biased parameters. Here we also consider the implications for willingness to pay estimates.

Keywords discrete choice models, attitudes, factor analysis, measurement models, recycled wastewater

1 Introduction and Objectives

Measurement error is a prominent issue in almost every statistical field, notably biometrics, econometrics and psychometrics (Durbin, 1954; Cochran, 1968; Skrondal and Rabe-Hesketh, 2007; Goldberger, 1972; Wansbeek and Meijer, 2000; Carroll *et al*., 2006; Rabe-Hesketh *et al.,* 2003). Unfortunately, the issue has been relatively ignored in the applied literature, particularly in econometrics (Train *et al.* 1987; Morikawa *et al.* 2002; Burton, 2008). Wang (2002) cites two reasons for this: the first being that, whilst practitioners are aware that measurement error is likely to be present in their data, "they believe that their effects in the model are likely to be insignificant". To the best of our knowledge, few studies have shown the sensitivity of the latent variables to varying levels of measurement error (Carroll *et al.,* 1984) and no studies have presented evidence as to the effect of measurement error on partworth estimates which economists are most interested in. Secondly, few statistical packages used by economists have procedures to correct for measurement error easily available.

The term measurement error refers to one of two types of error: error in the raw data or error in capturing a latent variable (Wansbeek and Meijer, 2000). The former refers to, for example respondents overstating their income or errors in the data collection technique used. The latter, which this paper investigates, is where observed variables used as a proxy for the unobservable latent may not entirely capture its true form (Wansbeek and Meijer, 2000; Greene, 1997). For example, intelligence could be measured singly by the level of ones education. Education level may be a poor proxy for intelligence, in which case intelligence is included into a regression model with a significant amount of unexplainable uncertainty, or error. The theory of how to correct for measurement error applies to both causes.

Latent variables, such as attitudes, social rules, motivations and primitive beliefs (often termed as psychometric data), are used in various disciplines, particularly social psychology, to better understand the behavioural decision making process (Morikawa et al., 2002; Ajzen *et al.,* 2004;

Nancarrow *et al*., 2008; Cook *et al.,* 2002; Spence and Townsend, 2006; Lobb *et al.,* 2007). A latent variable is a construct which captures these unobservable theoretical counterparts. Latent variables, for instance risk or environmental concern, have been included in stated preference surveys to explain heterogeneity in an economic value estimate or willingness to pay (Bateman *et al.* 2006; Burton *et al.* 2001; Smith, 1996).

The standard practice in environmental valuation, and other disciplines, for incorporating latent variables in a discrete choice model (DCM) is by a factor score, produced using factor analysis, even though there are superior techniques available. An increasingly common inclusion in WTP studies is the New Ecological Paradigm (NEP) theory, which states that respondents have primitive beliefs towards the environment, and these beliefs drive their preferences (Dunlap *et al.*, 2000; Kotchen and Reiling, 2000; Milon and Scrogin, 2006). These latent beliefs are revealed through answers to a bank of 15 attitudinal questions, and the use of factor analysis as a means to recreate these beliefs is explicitly advocated by Dunlap *et al,* (2000) in analysing the NEP data:

"Items can be treated as an internally consistent summated rating scale" (p425) *although*

"We encourage researchers to at least factor analyse the entire set at the outset" (p430)

Train *et al.* (1987) finds that by using factor scores, whereby the uncertainty in these variables is ignored, there is an issue with biased parameters in the subsequent non-linear model. Ben-Akiva *et al.* (1999), present a framework which seemingly overcomes this issue. Morikawa *et al.* (2002) developed the theory and outline a sequential estimation, inputting factor scores and variance into the DCM, and a simultaneously estimation of the DCM and measurement models. The simultaneous estimation yielded efficient estimators and more significant parameters. Simultaneous estimation can be done using a Generalised Linear Latent And Mixture Model (GLLAMM) in Stata 10.1 (StataCorp, 2007).

Burton (2008) investigated Train *et al.* (1987) findings by undertaking a Monte Carlo simulation. Bias in the parameter estimates resulted from using factor scores in a logit regression. However the impact of the bias on combinations of parameters (in particular those involving ratios) was not investigated. These are of particular interest in the context of valuation studies.

The objectives of this study are;

1. to use measurement models as a way of including latent variables into discrete choice models,
2. to determine the extent to which ignoring measurement error in factor scores gives biased parameter estimates in a subsequent discrete choice model, and
3. to determine whether this bias affects a median and marginal WTP estimate.

1.1Measurement bias

The use of primitive beliefs, attitudes and social rules to explain an economic value estimate (willingness to pay) has been increasing in stated preference surveys (Kotchen and Reiling, 2000; Milon and Scrogin, 2006; Spash, *et al*. 2006). These are often considered to be strictly unobservable, or latent, and can only be inferred from a set of observed indictor variables. The common method for moving from multiply indictors to a single measure has been thus far been through factor scores, which can be produced from a factor analysis. Factor analysis is a multivariate technique used to identify underlying structure among a set of variables, and then interrelated variables are grouped into factors. The factor score is a composite measure created for each observation on each latent extracted from the factor analysis and used in conjunction with the original variables value to calculate an observations score (Hair *et al.,* 2006).

A significant issue with using factor scores in non-linear models is that the parameters will be biased. This was identified by Train *et al.* (1987), in an investigation of attitudes towards energy consumption. Train *et al.* (1987) point out that the factor score computed for each individual is not the true value of the latent, rather a distribution of each individual's latent given their response to the indicator variables. Green (1997) also refers to this issue, noting that measurement error is inevitably present in most regression models where proxy variables are used as they are rarely true measurements of their theoretical counterparts. In other words, the factor scores are assumed as certain measures of the variable of interest when included in the response model, when in fact they are not. Simply omitting the variable induces a worse bias (McCallum, 1972; Wickens, 1972; Wansbeek and Meijer, 2000).

This section draws heavily upon Train *et al.* (1987) to indicate the nature of the problem. Assume a standard logit model of a discrete decision (e.g. to accept a new source of drinking water):

$$\begin{aligned} &y^* = \beta x + \varepsilon \\ &y = 1 \;\; if \;\; y^* > 0 \\ &y = 0 \;\; if \;\; y^* <= 0 \\ &P(y=1) = \frac{\exp(\beta x)}{1+\exp(\beta x)} \end{aligned} \tag{1}$$

Where y^* is an underlying perception of the utility of acceptance, dependent on observable variable x, and observed acceptance is given by $y = 1$. For a given value of x, the expected probability of acceptance is:

$$E[P(y=1|x)] = \frac{\exp(\hat{\beta} x)}{1+\exp(\hat{\beta} x)} \tag{2}$$

The relationship between the acceptance probability and x is given in Figure 1.

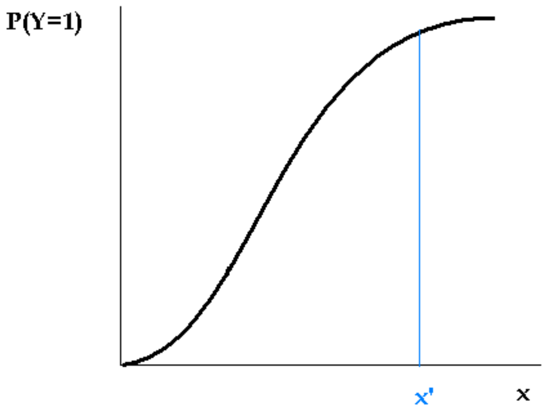

Figure 1: the logit response function.

However, consider the case where *x* is not observable, but has to be inferred (for example, it is a “primitive” attitude towards risk).

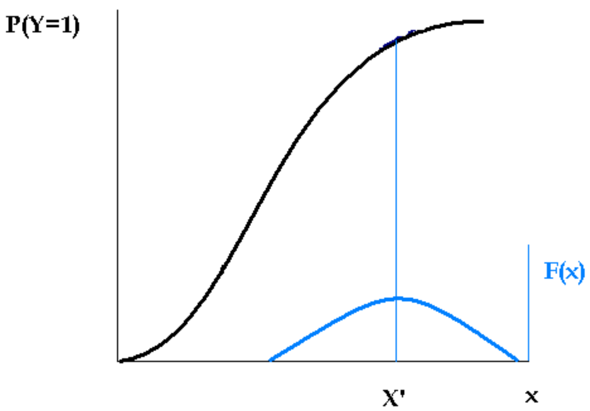

Figure 2: the logit response function with uncertainty about x.

Illustrated in Figure 2, a particular value of *x* now has a distribution F(*x*) (measured on the right hand scale, which is in different units to the left). Now,

$$E[P(y=1\mid x)]=\int_{-\infty}^{+\infty}\left[\frac{\exp(\hat{\beta}x)}{1+\exp(\hat{\beta}x)}\right].f(x)dx \qquad (3)$$

and

$$E[P(y=1\mid x)]\neq[P(y=1|E(x))] \qquad (4)$$

The implication is that estimating non-linear models using the expected value of uncertain exogenous variables will lead to biased and inconsistent results. Factor analysis and factor scores (based on a number of intermediate questions) give estimates of the mean of unobserved (and uncertain) latent exogenous variables (Train *et al.,* 1987). The factor scores are treated as a variable and as such any measurement properties that go along with forming the latent construct (through multiple items) are disregarded in the regression model (Hair *et al.,* 2006). To overcome the issue, a well accepted solution is to use measurement models with the discrete choice model (Green, 1997; Wansbeek and Meijer, 2000; Ben-Akiva *et al*., 1999; Morikawa *et al,*. 2002; Rabe-Hesketh and Skrondal, 2003; Eymann *et al,*. 2007; Burton, 2008) [1].

2 Methodology

[1] There are other ways to overcome measurement error. Wang (2002) suggests imposing a fixed variance, however you must first know what the value of the variance is. A fixed variance can also be imposed in the Stata program `gllamm` (Rabe-Hesketh *et al.*, 2003).

The DCM has already been explained in equations (1) and (2). The mode by which latent variables are to be incorporated into this model can be by a factor score or measurement models. Let us outline the two procedures.

2.1 Measurement models

In its simplest form the measure of the latent variable (*L*) is simply a number (ζ), normally distributed,

$$L = \zeta \ , \ \zeta \sim N(\theta, 1) \tag{5}$$

As previously mentioned, the latent cannot be directly measure, but must be inferred by a set of observed, indicator, variables.

The Ben- Akiva *et al.* (1999) method allows for flexible disturbances and explicit modelling of latent variables, heterogeneity and latent segmentation in DCM's. The measurement model, of which there are a set of measurement equations, relates latent variables and their indicators (Morikawa *et al.*, 2002). A measurement equation is defined as:

$$I = \lambda + \delta L + \varepsilon \tag{6}$$

where *I* is the score for the indicator variable. The location, λ, is the amount of offset between the latent and the score for the indicator variable. The loading, δ, is the correlation between the indicator score and the latent variable and ε is a random error component, normally distributed.

2.2 Factor scores

Essentially, a factor analysis uses a measurement model to produce a factor score. A standard factor analysis in Stata 10.1 (StatCorp, 2007) also normalises the latent variable. Additionally, the score for the indicator variable is normalised, having a mean of 0 and standard deviation of 1. The factor loadings are the correlation between the normalised indicator variable and underlying latent variable (factor), reflecting the strength of the relationship between the indicator and latent variable. Loadings cannot take a value of greater than or equal to 1. As the indicator variable is normalised, the location parameter in the measurement equation is 0, so that:

$$I = \delta' L + \varepsilon \qquad (7)$$

The factor score for each respondent is produced, and this variable is then used in the DCM.

2.3 GLLAMM

Rabe-Hesketh *et al.* (2004) provide a program in Stata (StatCorp, 2007), `gllamm`, which can fit a wide range of Generalised Linear Latent and Mixed Models (GLLAMM). GLLAMM's are a class of multilevel latent variable models for multivariate responses (for example dichotomous, ordered categorical and ranking). Latent class models also fall into this category of models.

The program `gllamm` can fit a dichotomous response model with latent and observed variables simultaneously. The latent variable score and variance is estimated using responses to observed, indicator, variables in conjunction with a response variable. Unlike other SEM software, the `gllamm` program allows for flexibility in the measurement and decision models: relationships can be linear or non-linear. Given that psychometric data is often measure on a Likert scale, there is an advantage to be able to manipulate the measurement model to reflect ordinal rather than continuous data, and even to model determinants of the latent directly.

Another advantage of `gllamm` is its handling of missing data. In other statistical packages, the factor analysis component of the SEM would only use a complete set of observations. Whereas `gllamm` uses all observations regardless of whether some are missing.

2.4 Monte Carlo simulation

This section outlines the Monte Carlo simulations used to identify each technique's sensitivity to measurement error. Varying levels of error in the measurement of the latent and its affect on subsequent parameter estimates and ratios tested. The Cronbach alpha is selected as the variance measure, due to its wide spread use in measuring a latent's internal consistency i.e. how well the indicators reflect the underlying latent. For a set of items to be considered a scale, the widely-accepted cut-off point for the alpha score of a latent variable is 0.7 or higher, however it can be as lenient as 0.6 (Garson, 2008). We shall investigate bias at varying alpha scores.

A draw of data consists of a sample of 500 observations. The response model is defined as

$$Y^* = \beta_0 - \beta_1 z - \beta_2 p - \varepsilon \tag{8}$$

The expected probability of the response is a standard probit model where Y^* is conditional on z.

$$P(Y^* = 1|x) = \Phi(x', \beta) \tag{9}$$

where $x = (z, p)$ and z is assumed to be the true value of a primitive belief held by respondent i, and p is the bid amount they are responding to within a choice experiment.

$$z = \zeta\,, \qquad \zeta \sim N(1, \theta) \tag{10}$$

For simplicity, it is assumed the *z* is not to be influenced by observed variables[2].

As *z* is not observed directly, there are assumed to be 4 repeated measures for each individual, and the relationship between observed measure and latent (suppressing the individual subscript *i*) is given by:

$$Z_j = \lambda_j + \delta_j z + \varepsilon_j \text{ and } j = (1,4) \tag{11}$$

Where λ_j is a location parameter for measure j, and δ_j the loading.

For identification, one parameter in the measurement model for each latent must be fixed hence $\delta_{z1} = 1$. This makes no difference to the results, as the value of the latent is scaled appropriately.

To determine the sensitivity of the ratio of the parameters (WTP) to bias, the median WTP is determined by:

$$median\, WTP = -\left(\frac{\beta_0+\beta_1 z}{\beta_2}\right) \tag{12}$$

And given a marginal unit change in the latent

$$marginal\, WTP = \left(\frac{\beta_1}{\beta_2}\right) \tag{13}$$

With each simulation parameters from three models are captured using;

1. a probit model, estimated using the actual value of *z*,
2. A probit model, estimated using an estimate of *z* (factor score) derived from a factor analysis on the Z_{ij} , and

[2] Accounting for indirect effects of an observed variable (eg. age, gender) are common in epidemiology research. Indirect effects are the impact on the latent and are an addition to equation (3) would be termed an exposure model (Rabe-Hesketh *et al*., 2003).

3. a joint estimation of a probit model, using an endogenous estimate of z, estimated simultaneously with the measurement model using `gllamm`.

The three techniques used are hereafter termed the actual, ignore and `gllamm`. Estimates are stored and the average estimate of the parameter (over the 100 simulations) are reported in section 3.1.

2.5 An application to recycled wastewater demand

The case study compares the Perth community's economic and social values of two future drinking water sources for Perth: a second desalination plant, and injecting recycled wastewater into an aquifer of stored future drinking water, which is known as Managed Aquifer Recharge (MAR)[3]. The most recent indication of the Perth communities level of acceptance of MAR is from Po *et al.* (2005), which found that less than one third (31.3%) of respondents would unconditionally accept this option for future supply. For further reading on Australian studies into preferences for future water supplies, from a social psychology perspective see Porter *et al.* (2005); Nancarrow *et al.* (2008); Hurlimann *et al*. (2008); Dolnicar and Schafer (2008).

2.5.1 Survey design

A double bounded discrete choice stated preference technique is employed. Respondents are asked to accept or decline a proposal for the introduction of the MAR scheme as opposed to the status quo, which is a second desalination plant. Their choice is constrained by the two available options, the price of the proposed recycled water scheme and their budget. The social psychology component to the survey was designed and tested by Porter *et al*. (2005). It contains a series of attitudinal questions,

[3] MAR or groundwater replenishment, is the injection of water into aquifers for later use as a drinking water, while improving groundwater quality and environmental values. MAR could also be used to mitigate or control saltwater intrusion into coastal aquifers (Water Corporation, 2006).

measured on a Likert scale, towards desalination and the MAR scheme. Generic attitudes and attitudes specific to the MAR scheme are (currently) used in the statistical analysis.

The elicitation of WTP accounted for positive and negative preferences in a block design. Respondents were assigned at random to either a version of the survey which has positive bid amounts associated with MAR (testing WTP for MAR) or negative amounts (testing willingness to accept MAR). The bid amounts offered range, by intervals of $30, from -$130 to $150.

Because there is no prefilter question, there is no reason to assume that the respondent will be assigned to the survey which reflects their true values, and hence one would expect 50% of the sample to respond no-no to the positive bid amounts, and yes-yes to negative bid amounts (the double bound bids never straddle zero). However, there is a possibility that some proportion of the sample is strictly indifferent between the two. This implies a discontinuity in the response function at a zero bid. In order to assist in identifying this, those who respond no-no or yes-yes are offered a third $1 bid (Mitchell and Carson, 1989).

The payment vehicle is the annual water service fee, which is a fixed additional fee to ratepayer's. The water service fee is independent of water usage, rather it covers the cost of sourcing and supplying the water to the household.

The survey was administered on the internet via an online survey company in August 2007. A total of 475 completed responses were collected. The response rate is approximately 25%.

2.5.2 Analysis

The latent variables that may influence WTP were identified through Porter *et al.* (2005). Community trust in authorities and the information they provide, the fairness of the scheme to various users, the perception of risk to various users, the perceived outcomes of the scheme (i.e. longevity, logic, sustainability etc.) and the subjective assessment of the scheme (i.e. whether the risks outweighed the benefits) were found to be significant drivers of acceptance. Po et al (2005) and subsequent work also identified emotion or the "yuck" factor as being an important driver of behaviour.

Confirmatory factor analysis was conducted on the specified indicators, and 11 latent constructs scored a Cronbach alpha greater than 0.7. A standard probit regression was done using the first bid amount offered, the response to this bid amount and the 11 latent's. Three latent's had a significant influence on MAR acceptance at the 1% significance level. They are;

1. the perceived fairness of the MAR scheme to Perth households, the Perth environment and future Perth generations (higher scores mean greater perceived fairness of MAR)
2. the respondents emotive feelings towards drinking the water or "yuck" factor (higher scores mean a lesser "yuck" factor), and
3. the respondents trust in authorities to manage and provide information on water systems (higher scores mean greater trust).

The three latent's are hereafter named *fairness*, *emotion* and *trust* respectively. A summary of each variable is given in Table 1.

Table 1. Description of latent variables.

Latent variables (L)	*Fairness*	*Emotion*	*Trust*
Number of indicators (I)	3	4	4
Cronbach alpha	0.95	0.95	0.86

In `gllamm`, a structural equation model with a linear relationship between latent variable and indicators and non-linear probit relationship between latent variables and the decision is implemented.

3 Results

3.1 Monte Carlo simulation

In the following figure's, each data point represents an average of 100 replications.

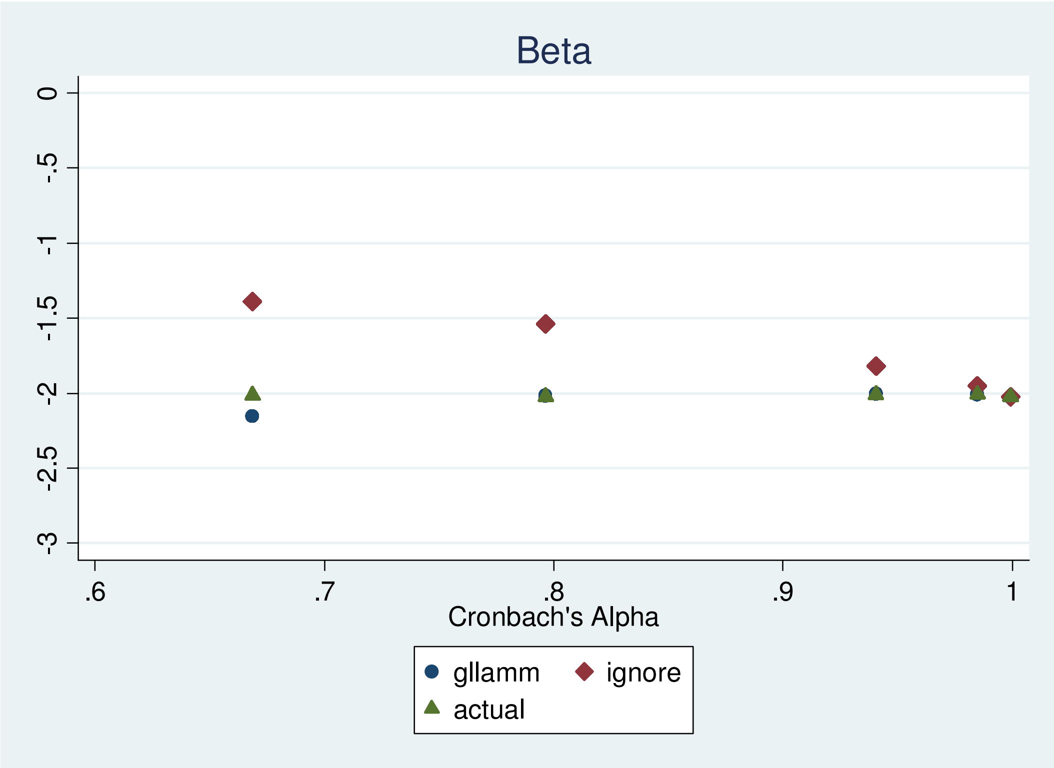

Figure 3. Averaged coefficients (β) of z at varying Cronbach's alpha using the actual value of z, an estimate of z from factor analysis (ignore) and an estimate of z from `gllamm`.

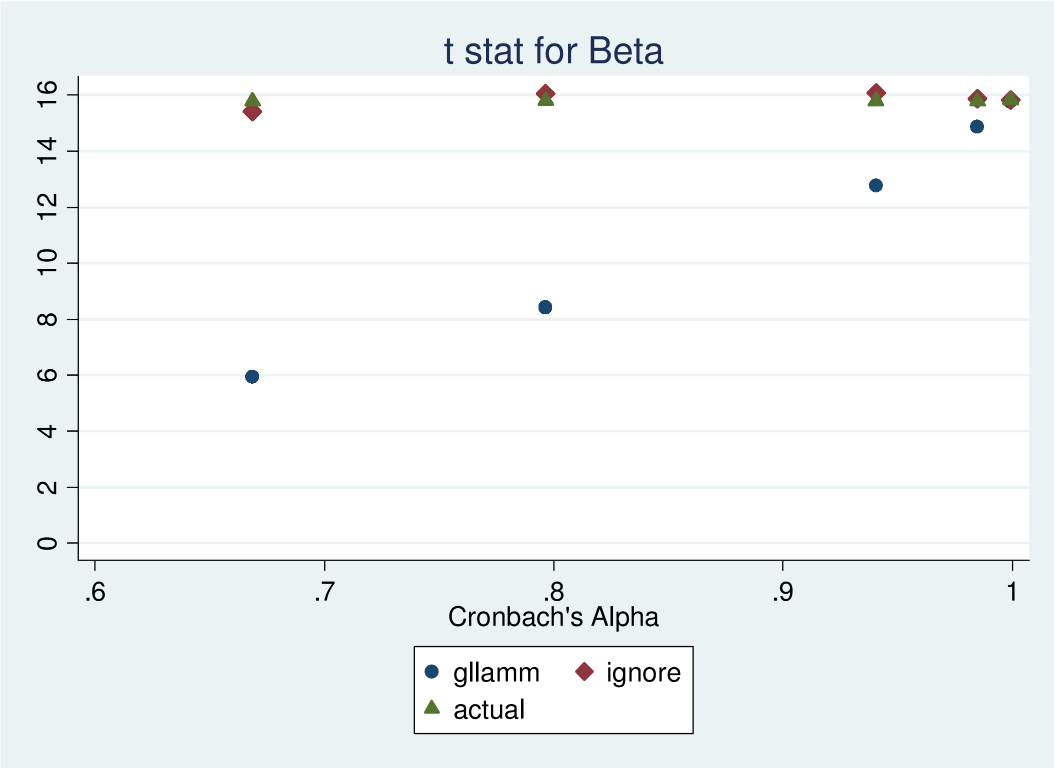

Figure 4. Averaged standard error of coefficients (β) of z at varying Cronbach's alpha using the actual value of z, an estimate of z from factor analysis (ignore) and an estimate of z from `gllamm`.

In Figure 3, the average coefficient on the latent (β) is becoming increasing biased towards zero when uncertainy in the measurement of the latent is ignored. The average coefficients on the latent using the actual and `gllamm` methods remain constant at -2. In Figure 4 gllamm is recognising the uncertainty in the latent score and hence its significance. As the factor analysis estimate of *z* hasn't allowed for any uncertainty in the latent, one wouldn't expect a decline in the t-value with increasing error.

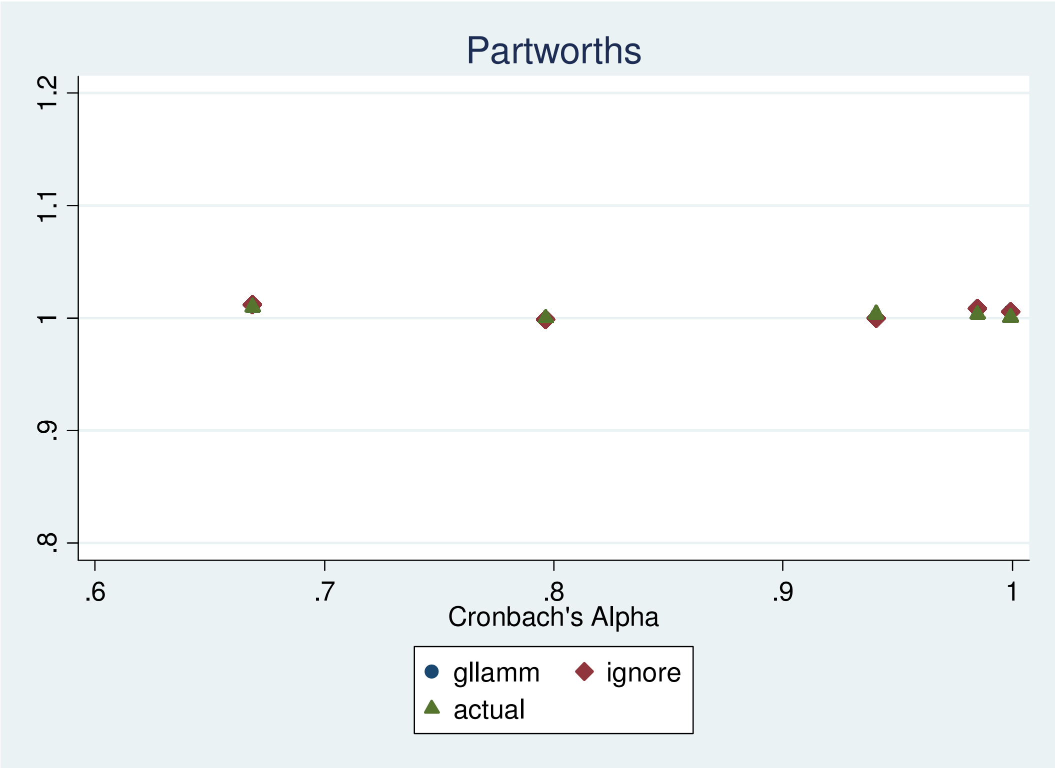

Figure 5. The average median WTP, at varying Cronbach alpha's, using the actual value of z, an estimate of z from factor analysis (ignore) and an estimate of z from `gllamm`.

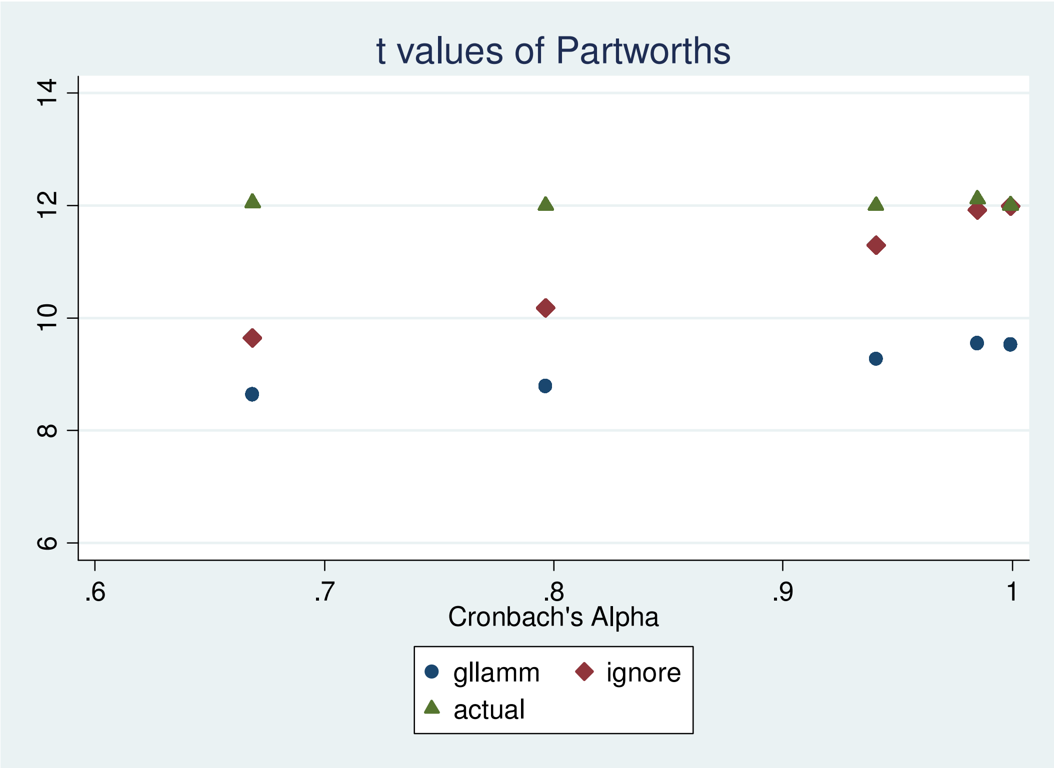

Figure 6. Average t-values for the median WTP, at varying Cronbach alpha's, using the actual value of z, an estimate of z from factor analysis (ignore) and an estimate of z from `gllamm`.

In Figure 5 the average median WTP (given by equation12) remains constant for all three methods over the range of alphas. The t-values however, plotted in Figure 6, remain constant when using the actual value of z and start declining when using the estimated z (ignore and `gllamm`). The t-values given from `gllamm` don't converge back to the t-value given by the actual. At higher levels of uncertainty in the measurement model it is to be expected that the confidence interval for the partworth should be higher with `gllamm`, as it is reflecting that uncertainty. It is more surprising that using the measurement model when there is relatively little measurement error comes with a penalty. In particular, it appears as if the `gllamm` estimates have an increased standard error for the *constant* in the response model, even when there is effectively no uncertainty in the measurement model. This is an unexpected result which requires further investigation.

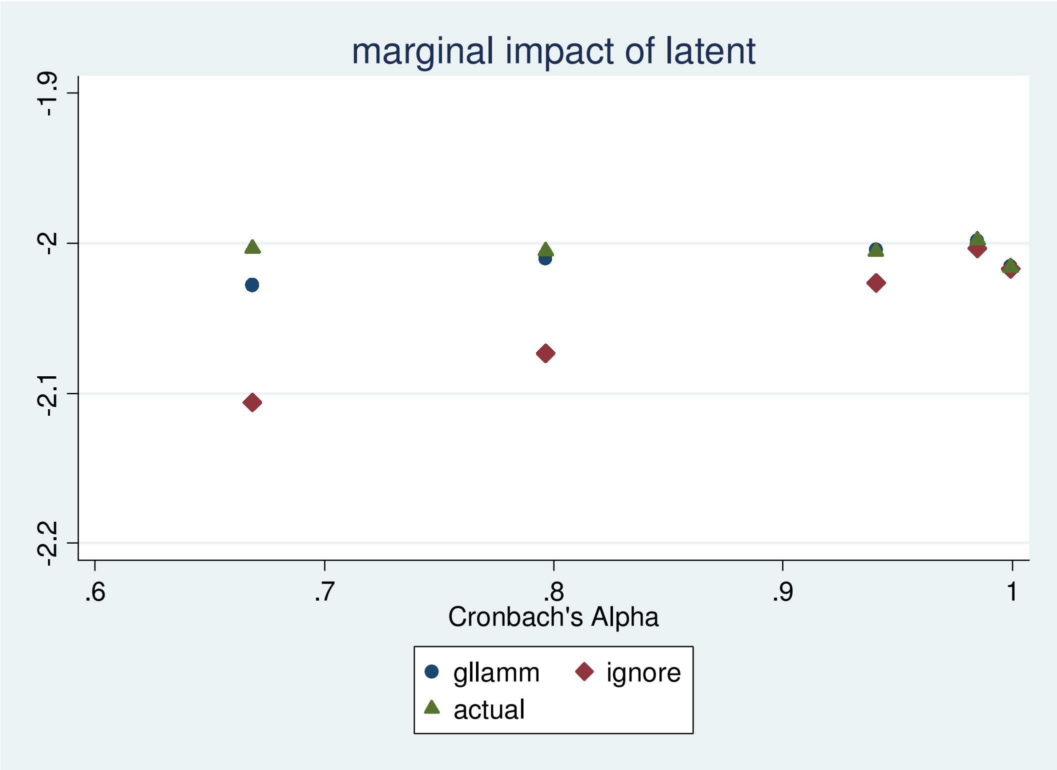

Figure 7. The average marginal impact of the latent on the WTP, at varying Cronbach alpha's, using the true value of z, an estimate of z from factor analysis (ignore) and an estimate of z from `gllamm`.

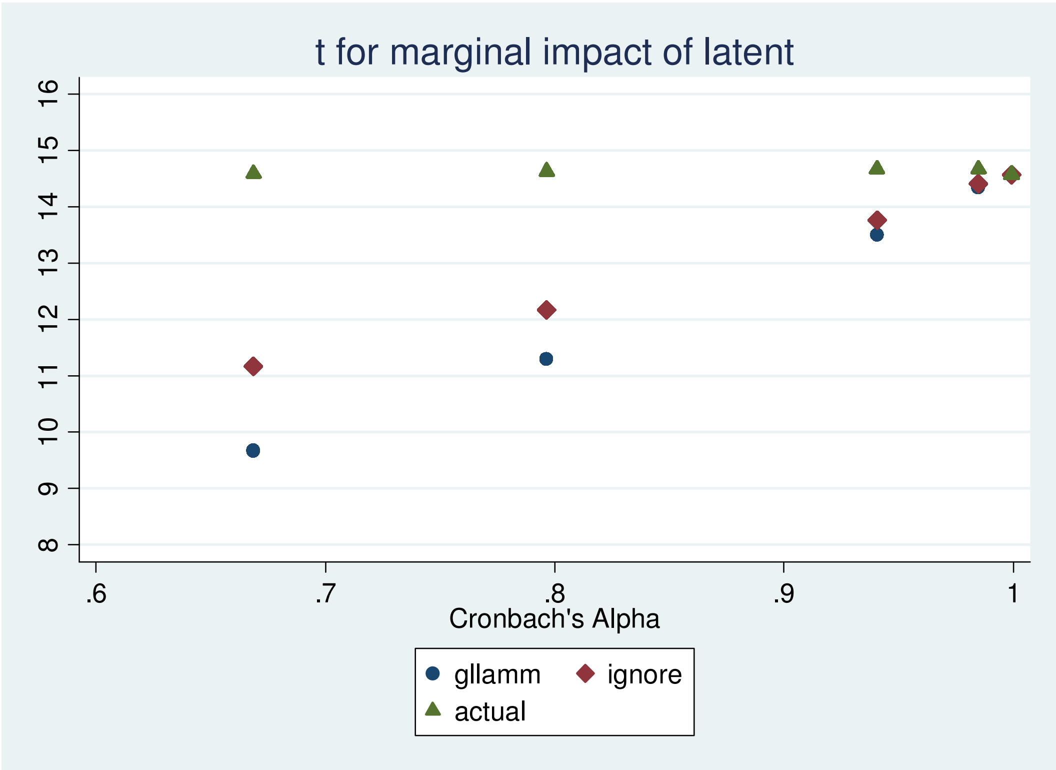

Figure 8. The average marginal impact of the latent on the WTP, at varying Cronbach alpha's, using the true value of z, an estimate of z from factor analysis (ignore) and an estimate of z from `gllamm`.

Figure 7 shows that by ignoring the uncertainty in z a greater effect on the average WTP (given by equation 13) given a marginal change in the latent can be expected. The t-values in Figure 8 converge back to the actual t-values as the Cronbach's alpha approaches 1, however they are still understated when using `gllamm`.

3.2 Willingness to pay for a MAR scheme

The parameters from the decision models, derived using factor scores and measurement models, are presented in Table 2. All parameters are significant at the 1% level. The factor analysis and measurement model parameters presented are the loadings (δ) and locations (λ) used in the estimation of each measurement equation (equation 6). As the indicator scores are normalised by the factor analysis, the locations are 0.

The coefficients differ marginally between the factor analysis and `gllamm` models, however significance levels are constant. As *Bid* decreases, more respondents are likely to accept the MAR scheme. Consistent with Porter et al (2005), high perceptions of fairness and low emotive feelings towards drinking the water improves acceptability of the scheme. Perhaps surprisingly, low levels of trust improves acceptability, meaning those with a higher level of trust in the ability of authorities generally would prefer desalination (the status quo).

Table 2. Parameter estimates for MAR from a standard probit with factor analysis and `gllamm`.

	Factor analysis	gllamm
Constant	-0.2***	-0.204***
Bid	-0.00511***	-0.00518 ***
Fairness	0.351***	0.334***
Emotion	0.442***	0.4***
Trust	-0.286***	-0.364***
	Factor analysis	**Measurement model**
Fairness	(δ'_1) 0.9 (δ'_2) 0.94 (δ'_3) 0.94	(δ_{11}) 1(fixed) (λ_{11}) 3.53 (δ_{12}) 1.04 (λ_{12}) 3.58 (δ_{12}) 1.11 (λ_{12}) 3.59
Emotion	(δ'_1) 0.9 (δ'_2) 0.85 (δ'_3) 0.95 (δ'_4) 0.92	(δ_{21}) 1(fixed) (λ_{21}) 3.28 (δ_{22}) 0.94 (λ_{22}) 3.19 (δ_{23}) 0.99 (λ_{23}) 3.26 (δ_{24}) 0.86 (λ_{24}) 3.10
Trust	(δ'_1) 0.8 (δ'_2) 0.86 (δ'_3) 0.67 (δ'_4) -0.76	(δ_{31}) 1(fixed) (λ_{31}) 3.02 (δ_{32}) 1.09 (λ_{32}) 2.9 (δ_{33}) 0.96 (λ_{33}) 3.66 (δ_{34}) -0.98 (λ_{34}) 2.75

**** 1% significant level, ** 5% significant level, *10% significance level*

In Table 3, the median WTP for a MAR scheme in Perth is -\$39, regardless of which method is used. On average respondents must have their yearly water service fee reduced by \$39 to accept the MAR scheme. The significance of this estimate is less when estimated with `gllamm` (5% significance compared with 1% with factor analysis). A unit increase (equivalent to 1 s.d) in *emotion* (respondents are less emotive towards the scheme) significantly increases WTP from the median of -\$39 to \$47, which is significant at the 5% level. `gllamm` produced an estimate of \$37, which was not significant.

A marginal decrease in the latent shifts WTP equally downwards from the median. In reducing the perception of fairness of the scheme and increasing emotive feelings toward drinking the water respondents must be significantly compensated. Decreasing respondents trust level further does shift WTP but it does not become significantly different from \$0. Increasing respondents trust levels significantly decreases WTP[4].

[4] Further investigation is required here as this is an unexpected result, conflicting with other literature on recycled water acceptance (Porter *et al.,* 2005; Nancarrow *et al.*, 2008)

Table 3. The median WTP at alternatives values, by estimation method.

	Factor analysis	`gllamm`
All latents = 0 (sample average)	-$39***	-$39**
Fairness +1 -1	$29 -$108***	$25 -$103***
Emotion +1 -1	$47** -$125***	$37 -$116***
Trust +1 -1	-$95*** $17	-$109*** $31

**** 1% significant level, ** 5% significant level, *10% significance level*

4 Discussion

4.1 Measurement bias

Figure 3 adds to the evidence (Carroll *et al.,* 1984; Burton, 2008) that model parameters are biased when factors scores are included into a non-linear DCM without accounting for error. However Figure5 shows that the WTP evaluated at the mean (zero) of the latent's, is not affected by the bias due to it being a ratio of parameters. The estimate of the marginal impact of the latent on the WTP does suffers some bias when error is ignored (Figure 7). The level of bias introduced seems to be trivial above Cronbach alpha >0.9.

By correcting for the bias in `gllamm`, we do lose significance in our partworths (Figure 6 and 8). This may partly be due to `gllamm` introducing uncertainty on the parameters where there is very little (Cronbach alpha > 0.9). Hence there seems to be a trade off between precise parameters and diminished significance of the parameters.

4.2 Case study

The bias has not affected the median WTP estimate (both are stable at -$39), however the significance of this estimate is slightly reduced when estimated with `gllamm`.

The affect of bias on marginal changes to the latent seems to following the same pattern. By improving respondents emotive feelings towards drinking the water they are WTP $47 on top of their current annual water service fee, however the `gllamm` estimate produces a WTP that (although positive) is not significantly different from $0.

Emotion and fairness are strongly influencing WTP for MAR, consistent with Porter *et al*. (2005). Those who found the scheme (any one of) highly "disgusting" or "revolting" or highly unfair to others, future generations and the environment (i.e. those in the top 15% of the distribution) needed to be compensated significantly in order to accept it (Table 3). Importantly, respondents require a reduction of at least $100 in their annual water service fee to accept MAR.

In contrast to other studies (Porter *et al.,* 2005; Nancarrow *et al.,* 2008) trust in authorities and information does not seem to be linked to perceptions of risk and hence those with lower levels of trust were more likely to accept the scheme, whilst those with high levels of trust preferred a desalination scheme (Table 2). Further investigation is required to untangle the relationship between trust and its high significance in WTP for MAR.

5 Conclusion

We have shown that including latents with error in DCM's will lead to biased parameters. However, the ratio of the parameters, WTP, may not be highly affected. There does seem to be a trade off between bias in the parameter and efficiency of the estimate, such that one may reduce the estimates efficiency if corrections are used on a latent with a small level of error (Carroll *et al.,* 1984)**.** Currently there is no guidance on when one should switch to correcting for error.

The application of measurement error correction to water demand in Perth produced robust and similar effects in both ignoring error (through factor analysis) and accounting for error (through `gllamm`). The median WTP is negative, implying, on average within the sample, an aversion to the introduction of the new MAR scheme. The individuals WTP seems to be highly sensitive to attitude variables in both forms of the model.

Further work aims to exploit `gllamm`'s capacity to specify non-linear measurement models (i.e. ordered probit) to account for the Likert scales used in the indicator questions. Secondly, it is possible to apply an 'exposure' model approach to explaining determinants of primitive beliefs, based on socio-economic data.

5 Acknowledgments

F.L Gibson is supported by a University of Western Australia postgraduate scholarship. The data collection was funded by a University of Western Australia Small Grant, and the Commonwealth Scientific and Industrial Research Organisation (CSIRO). The helpful comments from Blair Nancarrow, Zoe Leviston and Dr Sorada Tapsuwan of CSIRO: Sustainable Ecosystems are greatly acknowledged.

6 References

Ajzen, I., Brown, T.C. and Carvajal, F. (2004). Explaining the discrepancy between intentions and actions: the case of hypothetical bias in contingent valuation. *Personality and Social Psychology Bulletin* 30(9), 1108-1121.

Bateman, I.J., Cole, M.A. Georgiou, S. and Hadley, D.J. (2006). Comparing contingent valuation and contingent ranking: A case study considering the benefits of urban river water quality improvements. *Journal of Environmental Management* 79, 221-231.

Burton, M. (2008). Latent exogenous variables in nonlinear models: Prospects for biased and inconsistent estimates in adoption and valuation studies. In proceedings from the 52nd Australian Agricultural and Resource economics Conference, Canberra, February 2008.

Cochran, W.G. (1968). Errors of measurement in statistics. *Technometrics* 4, 637-666.

Durbin, J. (1954). Errors in variables. *International Statistics Review* 22, 2332.

Dunlap, R.E., VanLiere, K.D., Mertig, A.G. and Jones, R.E. (2000). Measuring endorsement of the New Ecological Paradigm: A revised NEP scale. *Journal of Social Issues* 56, 425-442.

Carroll, R.J., Spiegelman, C.H., Gordon Lan, K.K., Bailey, K.T. and Abbott, R.D. (1984). On errors-in-variables for binary regression models. *Biometrika* 71(1), 19-25.

Carroll, R.J., Ruppert, D. and Stefanski, L.A. (2006). *Measurement error in nonlinear models: a modern perspective,*2nd. Ed. Chapman and Hall: Florida.

Cook, A.J., Kerr, G.N. and Moore, K. (2002). Attitudes and intentions towards purchasing GM food. *Journal of Applied Social Psychology* 23(5), 557-572.

Dolnica, S. and Schafer, A.I. (2008). Desalinated versus recycled water: public perceptions and profiles of the accepters. *Journal of Environmental Management* article in press, 1-13.

Garson, D. (2008). Structural equation modelling. Retrieved http://faculty.chass.ncsu.edu/garson/PA765/structur.htm#biblio

Goldberger, A.S. (1972). Econometrics and psychometrics. *Psychometrika* 36(2), 83-107.

Greene, W.H. (1997). *Econometric Analysis, 3rd Ed.* Prentice Hall: New Jersey.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (2006). *Multivariate data analysis, 6th ed.* Pearson International Edition.

Hurlimann, A., Hemphill, E., McKay, J. and Geursen, G. (2008). Establishing components of community satisfaction with recycled water use through a structural equation model. *Journal of Environmental Management* 88, 1221-1232.

Kotchen, M.J. and Reiling, S.D. (2000). Environmental attitudes, motivations and contingent valuation of non-use values: a case study involving endangered species. *Ecological Economics* 32, 93-107.

Lobb, A.E., Mazzocchi, M. and Traill, W.B. (2007). Modelling risk perception and trust in food safety information within the theory of planned behaviour. *Food Quality and Preference* 18, 384–395.

McCallum, B. (1972) Relative asymptotic bias from errors of omission and measurement. *Econometrica* 40, 757-758.

Milon, J.W. and Scrogin, D. (2006). Latent preference and valuation of wetland ecosystem restoration. *Ecological Economics* 56(2), 162-175.

Mitchell, R.C. and Carson, R.T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington, D.C: Resources for the Future.

Morikawa, T., Ben-Akiva, M. and McFadden, D. (2002). Discrete choice models incorporating revealed preferences and psychometric data. *Econometric Models in Marketing* 16, 29-55.

Nancarrow, B.E., Leviston, Z., Po, M., Porter, N.B. and Tucker, D.I. (2008). What drives communities' decisions and behaviours in the reuse of wastewater. *Water Science and Technology* 57(4), 485-491.

Po, M., Nancarrow, B.E., Leviston, Z., Porter, N.B., Syme, G.J. and Kaercher, J.D. (2005). *Predicting Community Behaviour in Relation to Wastewater Reuse: What drives decisions to accept or reject?* CSIRO: Water for a Healthy Country National Research Flagship, Land and Water : Perth.

Porter, N.B., Leviston, Z., Nancarrow, B.E., Po, M. and Syme, G.J. (2005). *Interpreting Householder Preferences to Evaluate Water Supply Systems: An Attitudinal Model*, CSIRO: Water for a Healthy Country National Research Flagship, Land and Water : Perth.

Rabe-Hesketh, S., Skrondal, A. and Pickles, A., (2003). Maximum likelihood estimation of generalised linear models with covariate measurement error, *The Stata Journal* 3(4), 386-411.

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2004). Generalised multilevel structural equation modelling. *Psychometrika* 69, 167-190.

Skrondal and Rabe-Hesketh, (2007) *Multilevel and longitudinal modelling using Stata*, 2nd edition. College Station, TX: StataCorp LP.

Skrondal, A. and Rabe-Hesketh. S.(2004). *Generalised latent variable modelling: multilevel, longitudinal, and structural equation models.* Chapman and Hall/CRC, Boca Raton, Florida.

Smith, K.V. (1996). Can contingent valuation distinguish economic values for different public goods? *Land Economics* 72(2), 139.

Spash, C.L., Urama, K., Burton, R., Kenyon, W., Shannon, P. and Hill, G. (2006). Motives Behind Willingness to pay for improving biodiversity in a water ecosystem: economics, ethics and social psychology. *Ecological Economics* article in press, 1-10.

Spence, A. and Townsend, E. (2006). Examining consumer behavior toward genetically modified (GM) food in Britain. *Risk Analysis* 26(3), 657-670.

StataCorp. (2007). *Stata Statistical Software: Release 10.* College Station, TX: StataCorp LP.

Train, K.E., McFadden, D.L. and Goett, A.A. (1987). Consumer attitudes and voluntary rate schedules for public utilities. *The Review of Economics and Statistics* 69(3), 383-391.

Wang, L. (2002) A simple adjustment for measurement errors in some limited dependent variable models. *Statistics and Probability Letters* 58, 427 433.

Wansbeek, T. and Meijer, E. (2000). *Measurement Error and Latent Variables in Econometrics.* Elsevier Science B.V: The Netherlands.

Water Corporation. (2006). Managed Aquifer Recharge. Retrieved http://www.watercorporation.com.au/M/mar.cfm?uid=2402-1988-6089-1527.

Wickens, M. (1972) A note on the use of proxy variables. *Econometrica* 40, 759-760.