

Environmental values and valuation over time

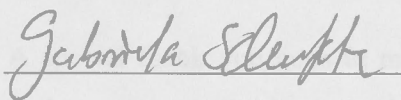
Gabriela Scheufele

February 2012

A thesis submitted for the degree of Doctor of Philosophy of
The Australian National University

DECLARATION

The contents of this thesis are original work by the PhD candidate. It contains no material that has been presented for a degree or diploma at any university or other institute of tertiary education. To my knowledge, it contains no material previously published or written by another person, except where duly acknowledged in the text.



Gabriela Scheufele

February 2012

ACKNOWLEDGMENTS

It is a pleasure to thank the many people who made this thesis possible.

First and foremost I want to thank my supervisor, Prof. Jeff Bennett. The enthusiasm he has for his research has been contagious and motivational for me. His inspiration, encouragement, moral support, sound advice, constructive feedback, and support in presenting my work at national and international conferences, have all made my PhD candidature a most positive experience, which I will keep in good memories.

A special thank you goes to my co-supervisor, Dr. David Godden for his continuous assistance, great guidance, and for all the invaluable comments he made on the final draft of my thesis. I especially appreciate the priceless insights into the world of policy-making I gained through him.

I am indebted to the Australian Government (Department of the Environment, Water, Heritage and the Arts) for funding the Environmental Economics Research Hub through the Commonwealth Environment Research Facilities. I thank them for providing me with the opportunity to write this PhD thesis.

I wish to thank Prof. Michael Burton from the University of Western Australia, Prof. Trevor Breusch, Dr. Michael Ward, Dr. Wendy Noble, Robyn Walter, Christine Adams, Tony Varty, and all other staff from the Crawford School of Economics and Governance at the Australian National University for interesting discussions, valuable advice and great support.

I want to thank Judy Bell for her excellent work in copy-editing my PhD thesis.

A big thank you goes to my PhD colleagues and friends Ben McNair, Marit Kragt, Kasia Mazur, and Sonia Akter for many inspiring, controversial, and sometimes emotional discussions about the discrete choice experiment method and many other things. Being able to write my PhD thesis in such an inspiring and exciting research environment was a privilege I am very grateful for.

I am very thankful to the ANU mountaineering club, and especially to its active members who made sure I wouldn't forget that there is a life besides writing a PhD thesis.

A thousand thanks go to my dear friends Ingrid Prem, Matthias Müller, Markus Schöfer, Fiona Jinman, Frank Summa, Sabrina Schröpfer, Martin Ribbeck, and Farah Beaini who, even though some of them are scattered around the world, always made me see the light at the end of the tunnel.

More than a thousand thanks go to my partner, David Brittain, for his endless patience and loving support.

Lastly, and most importantly, I wish to thank my family, and especially my parents Maria and Kurt who never lost faith in me or my endeavours, and supported me unconditionally along the way. To them I dedicate this PhD thesis.

*Gabriela Scheufele
Australian National University
February 2012*

ABSTRACT

The natural environment produces a range of goods and services that are valuable to society but not bought and sold in markets. Non-market valuation methods facilitate the estimation of the values society attaches to such non-marketed environmental goods and services. Considering such values in policy decisions has the potential to reduce inefficiencies introduced by incomplete property rights structures.

Policy decisions concerned with the natural environment often have long-term and irreversible consequences. Failing to account for changes in the values society attaches to non-marketed environmental goods and services over extended time periods may result in biased estimates of long-term benefits and costs. Grounding policy decisions on biased value estimates may reduce potential improvements in economic efficiency.

However, little is known about either the temporal value variability over extended time periods or the impact of the ongoing methodological evolution on value estimates used in cost-benefit analysis. The study presented in this thesis addresses this gap by making comparisons of marginal value estimates and the evolution and performance of stated preference valuation methods over a 30-year time period. Both the original stated preference valuation method (open-ended contingent valuation) and a state-of-the-art stated preference valuation method (discrete choice experiment) are used to make these comparisons with reference to the preservation value of Nadgee Nature Reserve, a protected natural area on the far south coast of New South Wales.

The results suggest that the marginal value respondents from Canberra attach to the preservation of Nadgee Nature Reserve in 1979 was not significantly different from comparable estimates made in 2009. One possible explanation for this apparent value stability across time is that proportional shifts in supply and/ or demand side factors might have cancelled each other out.

Alternative explanations for this observed temporal value stability are related to the elicitation format of the open-ended contingent valuation method used to estimate the preservation value of Nadgee Nature Reserve in 1979 and 2009. The findings indicate that this elicitation format might have stabilised value estimates across time through prominence in response. That is, the preservation value of Nadgee Nature Reserve

might have changed from 1979 to 2009, but the elicitation format did not reveal this change. These findings stress that the use of marginal value estimates obtained in the past by the open-ended contingent valuation method for current policy decisions might not be sound.

The findings further suggest that changes in the incentive compatibility properties of elicitation formats, and the associated potential changes in strategic response and strategic learning, which accompanied the evolution of stated preference methods, might confound an analysis of temporal value variability.

In general, this thesis argues that the possibility of value change across extended time periods questions the current practice of using snap-shot value estimates adjusted by a discount rate to predict future values. This research has provided first insights into the variability of marginal values related to non-marketed environmental goods and services over extended time periods. These insights warrant future research, which will be required to draw further policy relevant conclusions.

2.3	The concept of short economic value	22
2.4	Conclusion	23
3	ENVIRONMENTAL VALUE ACROSS TIME	25
3.1	Introduction	25
3.2	Temporal value variability – demand side	26
3.2.1	Proliferate	26
3.2.2	Project constraints	28
3.3	Temporal value variability – supply side	29
3.3.1	Scale of provision	29
3.3.2	Availability of close substitutes and associated relative scarcity	30
3.3.3	Reference point reliability	32
3.4	Conclusion	33

TABLE OF CONTENTS

1	INTRODUCTION	1
1.1	Research context	1
1.2	Research motivation	3
1.3	Research goals	5
1.4	Thesis outline	6
2	DEFINING VALUE	10
2.1	Introduction	10
2.2	Theoretical background	11
2.3	Economic value of non-marketed environmental goods and services	14
2.4	The notion of economic value from a broader perspective	17
2.5	The concept of total economic value	21
2.6	Conclusion	23
3	ENVIRONMENTAL VALUE ACROSS TIME	25
3.1	Introduction	25
3.2	Temporal value variability – demand side	26
3.2.1	Preferences	26
3.2.2	Budget constraint	28
3.3	Temporal value variability – supply side	29
3.3.1	Scale of provision	29
3.3.2	Availability of close substitutes and associated relative scarcity	30
3.3.3	Reference point relativity	32
3.4	Conclusion	32

4	ENVIRONMENTAL VALUATION ACROSS TIME	34
4.1	Introduction	34
4.2	Contingent valuation method	35
4.2.1	Background	35
4.2.2	Prominence in response in open-ended contingent valuation	38
4.2.3	Econometric modelling of open-ended contingent valuation data	41
4.3	Discrete choice experiment method	49
4.3.1	Background	49
4.3.2	Incentive compatibility properties of discrete choice experiments	51
4.3.3	Econometric modelling of discrete choice experiment data	58
4.4	Conclusion	64
5	RESEARCH QUESTIONS AND HYPOTHESES	66
5.1	Introduction	66
5.2	Value variability across time	68
5.3	Valuation variability across time	70
5.4	Conclusion	73
6	CASE STUDY	74
6.1	Introduction	74
6.2	Nadgee Nature Reserve	74
6.3	Reference study	76
6.4	Hypotheses revisited	81
6.5	Conclusion	82
7	RESEARCH DESIGN AND HYPOTHESES TESTING	84
7.1	Introduction	84
7.2	Value variability across time	85
7.2.1	Stated marginal willingness to pay	85
7.2.2	Explaining marginal willingness to pay – demand side	86
7.2.3	Explaining marginal willingness to pay – supply side	88
7.3	Valuation variability across time	89
7.3.1	Discrete choice experiments as an alternative to contingent valuation	89
7.3.2	Prominence in response in open-ended contingent valuation	91
7.3.3	Incentive compatibility properties of discrete choice experiments	93
7.4	Conclusion	100

8	DATA COLLECTION	102
8.1	Introduction	102
8.2	Contingent valuation method	103
8.2.1	Sampling	103
8.2.2	Questionnaire design	105
8.2.3	Data coding	107
8.2.4	Sample characteristics	109
8.3	Choice experiment method	112
8.3.1	Sampling	112
8.3.2	Questionnaire design and data coding	113
8.3.3	Experimental design	116
8.3.4	Focus groups	118
8.3.5	Pilot survey	120
8.3.6	Sample characteristics	121
8.4	Conclusion	125
9	ANALYSING VALUE VARIABILITY ACROSS TIME	127
9.1	Introduction	127
9.2	Stated marginal willingness to pay	128
9.3	Explaining marginal willingness to pay - demand side	133
9.4	Explaining marginal willingness to pay – supply side	157
9.5	Conclusion	159
10	ANALYSING IMPACTS OF METHODOLOGICAL EVOLUTION	163
10.1	Introduction	163
10.2	Discrete choice experiments as an alternative to contingent valuation	164
10.3	Prominence in response in open-ended contingent valuation	175
10.4	Incentive compatibility properties of discrete choice experiments	177
10.5	Conclusion	190
11	CONCLUSION	195
11.1	Discussion of research findings	195
11.2	Challenges and limitations	199
11.3	Recommendations for future research	204
11.4	Policy implications	205

APPENDICES	208
Appendix A: CV questionnaire of the reference study	208
Appendix B: CV questionnaire of the replicated study	208
Appendix C: Photo-test material	208
Appendix D: Model result comparison (linear OLS regression)	208
Appendix E: DCE questionnaire of RB split sample	209
Appendix F: DCE questionnaire of Pilot 1 split sample	209
Appendix G: Results of CL model	209
Appendix H: Data codebook	210
REFERENCES	211
Table 10: Sample characteristics in DCE survey (age)	131
Table 11: Sample characteristics in DCE survey (household income)	132
Table 12: Sample characteristics in DCE survey (gender)	133
Table 13: Sample characteristics in DCE survey (non-school education)	133
Table 14: Correspondence of DCE survey format	134
Table 15: Comparison of P/P	136
Table 16: Comparison of stated marginal WTP	136
Table 17: Comparison of stated marginal WTP per HH	137
Table 18: Comparison of stated zero marginal WTP rate	137
Table 19: Comparison of stated annual WTP	137
Table 20: Model results of log-linear regression model – marginal WTP per household	138
Table 21: Results of log-linear regression model – marginal WTP per HH	141
Table 22: Results of logit model	144
Table 23: Comparison of age – CV sample 1979/2009	146
Table 24: Comparison of non-school education – CV sample 1979/2009	146
Table 25: Comparison of household income – CV sample 1979/2009	147
Table 26: Comparison of gender and place of birth – CV sample 1979/2009	148

LIST OF TABLES

Table 1: Research design for hypotheses 2.3 – 2.6	100
Table 2: Coding of variables obtained in CV surveys	108
Table 3: Sample characteristics in CV surveys: age	110
Table 4: Sample characteristics in CV surveys: household income	111
Table 5: Sample characteristics in CV surveys: gender and place of birth.....	111
Table 6: Sample characteristics in CV surveys: non-school education	112
Table 7: Attributes and their levels in DCE survey	115
Table 8: Coding of variables obtained in DCE survey	116
Table 9: Pilot survey split sample treatments in DCE survey.....	120
Table 10: Sample characteristics in DCE survey: age	122
Table 11: Sample characteristics in DCE survey: household income	122
Table 12: Sample characteristics in DCE survey: gender.....	123
Table 13: Sample characteristics in DCE survey: non-school education	123
Table 14: Consequentiality of DCE survey format.....	124
Table 15: Comparison of PTP.....	130
Table 16: Comparison of stated marginal WTP.....	130
Table 17: Comparison of stated marginal WTP per HM.....	131
Table 18: Comparison of protest zero marginal WTP bids	132
Table 19: Comparison of protest motivations.....	132
Table 20: Model results of log-linear regression model – marginal WTP per household	138
Table 21: Results of log-linear regression model – marginal WTP per HM	141
Table 22: Results of logit model.....	144
Table 23: Comparison of age – CV sample 1979:2009	146
Table 24: Comparison of non-school education – CV sample 1979:2009	146
Table 25: Comparison of household income – CV sample 1979:2009	147
Table 26: Comparison of gender and place of birth – CV sample 1979:2009.....	148

Table 27: Comparison of behavioural and attitudinal variables – sample 1979:2009..	149
Table 28: Comparison of age – Census 1976:2006	151
Table 29: Comparison of household income – Census 1976:2006.....	152
Table 30: Comparison of gender and place of birth – Census 1976:2006.....	153
Table 31: Comparison of non-school education – Census 1976:2006.....	153
Table 32: Comparison of payment vehicle choice – CV sample 1979:2009.....	157
Table 33: Number of protected areas in New South Wales.....	158
Table 34: Extent of protected areas in New South Wales.....	158
Table 35: Equivalent surplus obtained through DCE-RT and open-ended CV	164
Table 36: Model results of pRPL(H) and pRPL(H-C) estimation in WTP space.....	167
Table 37: Model results of pRPL(P) and pRPL(P-C) estimation in WTP space	168
Table 38: Model results of pRPL(H-C-SD) and pRPL(H-C-HI) estimation in WTP space.....	170
Table 39: Comparison of age – CV(replicated sample):DCE-RT	172
Table 40: Comparison of non-school education – CV(replicated sample):DCE-RT ...	173
Table 41: Comparison of household income – CV(replicated sample):DCE-RT	174
Table 42: Comparison of gender– CV(replicated sample):DCE-RT	174
Table 43: Frequency distribution of prominent numbers.....	175
Table 44: Regrouped prominent number categories	176
Table 45: CL Model Results for RB1, SB and FirstRB.....	178
Table 46: Marginal WTP estimates for RB1, RB2, RB3, RB4, and SB.....	179
Table 47: Test results for Attribute and Scale Factor Equality across RB split samples and SB	180
Table 48: Test results Attribute and Scale Factor Equality for SB and FirstRB	181
Table 49: Model results of pRPL estimation – <i>strategy-type</i> and <i>positions</i> in the sequence separately.....	183
Table 50: Test results for equality for attribute and scale factors between P1, P2, P3 and P4	186
Table 51: Marginal WTP of P1, P2, P3, and P4	187

Table 52: Model results of pRPL estimation – <i>strategy-type</i> and <i>positions</i> in the sequence combined	189
Figure 1: Comparing water and electricity surplus	16
Figure 2: Effect of individual characteristics on marginal WTP	27
Figure 3: Effect of household size on marginal WTP	30
Figure 4: Effect of relative poverty on marginal WTP	31
Figure 5: Data characteristics of reference and the replicated sample	134
Figure 6: Results in log-linear regression model – marginal WTP per household ..	137
Figure 7: Results in the log-linear regression model – marginal WTP per RM	140
Figure 8: Comparison of changes in age between CV sample and Census data	154
Figure 9: Comparison of changes in non-school education between CV sample and Census data	154
Figure 10: Comparison of household income between CV sample and Census data ..	155
Figure 11: Distribution of promised numbers inflation-adjusted	176
Figure 12: Distribution of promised numbers inflation-adjusted	177
Figure 13: B/S acceptance curves	185

LIST OF FIGURES

Figure 1: Compensating surplus and equivalent surplus 16

Figure 2: Effect of indifference curve shape on marginal WTP 27

Figure 3: Effect of provision scale on marginal WTP 30

Figure 4: Effect of relative scarcity on marginal WTP 31

Figure 5: Data characteristics of reference and the replicated sample..... 134

Figure 6: Residuals in log-linear regression model – marginal WTP per household ... 137

Figure 7: Residuals in the log-linear regression model – marginal WTP per HM 140

Figure 8: Comparison of changes in age between CV sample and Census data 154

Figure 9: Comparison of changes in non-school education between CV sample and
Census data 154

Figure 10: Comparison of household income between CV sample and Census data... 155

Figure 11: Distribution of prominent numbers inflation-unadjusted 176

Figure 12: Distribution of prominent numbers inflation-adjusted 177

Figure 13: Bid acceptance curves 185

LAD	Least absolute deviations estimator
MA	Maintained assumption
ML	Mixed logit
MNL	Multinomial logit
OLS	Ordinary least squared
%	Percentage
PPI	Personal paper personal interview
PR	post random parameter logit
REP	Replicated reference
SS	Stochastic Surrogate Surrogate
SP	Stochastic Preference
SSD	Stochastic Surrogate Surrogate

LIST OF ACRONYMS

AA	Additional assumptions
AIC	Akaike information criterion
\$A	Australian dollar
BIC	Bayesian information criterion
BT	Benefit transfer
CAPI	Computer assisted personal interview
CPI	Consumer price index
CBA	Cost-benefit analysis
CL	Conditional logit
CV	Contingent valuation
DCE	Discrete choice experiment
EV1	Extreme value type 1
H	Hectares
HM	Household member
HQIC	Hannan-Quinn Information Criterion
IIA	Independence-from-Irrelevant Alternatives axiom
IID	Independently and identically distributed
LAD	Least absolute deviations estimator
MA	Maintained assumption
ML	Mixed logit
MNL	Multinomial logit
OLS	Ordinary least squared
P	Percentage
PPPI	Pen and paper personal interviews
pRPL	panel random parameter logit
PTP	Preparedness to pay
RB	Repeated-binary split sample
RPL	Random parameter logit
RT	Repeated-triple split sample
SB	Single-binary split sample
SD	Statistical division
SSD	Statistical subdivision

SLA	Statistical local areas
VIF	Variance inflation factor
WTP	Willingness to pay

The concepts of “cost” and “benefit” are closely related as both influence human behaviour. However, not environmental economists studies have attempted to link these two concepts to design a better understanding of the temporal variability of values related to environmental environmental goods and services with public good character. The main goal of this thesis is to address this gap.

1.4 Research context

The natural environment provides a flow of environmental goods and services that are valuable to humans and thus contribute to the economic welfare of society. However, since many of these environmental goods and services are not traded in markets, they lack a market price and have often been treated as having a default value of zero. Hence, considering the value of benefits derived from these environmental goods and services in decision making has the potential to improve or worsen efficiency in resource allocation.

Cost-benefit analysis (CBA) is a widely used tool to analyse the relative efficiency of alternative resource allocations (for example, Hanley and Barbier, 2009). It facilitates exploring potential welfare changes in society by comparing benefits and costs associated with a reallocation of resources from a status quo. A resource allocation is said to be Pareto optimal if no other feasible change makes anyone better off without making at least one person worse off. Since Pareto optimal resource allocations are only a hypothetical concept, economists use the term “Pareto improvement” to denote an increase in efficiency. Clearly, the Pareto criterion is restrictive since almost any reallocation of resources will lead to a net loss for some people. While the Pareto criterion still holds if the gains actually compensate the losses, the practical application of such a restriction is challenging. A solution to overcome the restriction of the Pareto criterion is to employ the Kaldor-Hicks criterion (for example, Hanley and Barbier, 2009). The Kaldor-Hicks criterion states that a change in resource allocation results in a net benefit to society if the gainers could, in theory, compensate the losers even though no compensation payments are made.

1 INTRODUCTION

The concepts of ‘time’ and ‘value’ are clearly related as both influence human behaviour. However, few environmental economics studies have attempted to link these two concepts to obtain a better understanding of the temporal variability of values attached to non-marketed environmental goods and services with public good character. The overall goal of this thesis is to address this gap.

1.1 Research context

The natural environment provides a flow of environmental goods and services that are valuable to humans and thus contribute to the economic welfare of society. However, since many of these environmental goods and services are not traded in markets they lack a market price and have often been treated as having a default value of zero. Hence, considering the value of benefits derived from these environmental goods and services in decision making has the potential to improve economic efficiency in resource allocation.

Cost-benefit analysis (CBA) is a widely used tool to analyse the relative efficiency of alternative resource allocations (for example, Hanley and Barbier, 2009). It facilitates exploring potential welfare changes to society by comparing benefits and costs associated with a reallocation of resources from a status quo. A resource allocation is said to be Pareto optimal if no other feasible change makes anyone better off without making at least one person worse off. Since Pareto optimal resource allocations are only a hypothetical scenario economists use the term ‘Pareto improvement’ to denote an increase in efficiency. Clearly, the Pareto criterion is restrictive since almost any reallocation of resources will lead to a net loss for some people. While the Pareto criterion still holds if the gainers actually compensate the losers, the practical application of such a transaction is challenging. A solution to overcome the restriction of the Pareto criterion is to employ the Kaldor-Hicks criterion (for example, Hanley and Barbier, 2009). The Kaldor-Hicks criterion states that a change in resource allocation results in a net benefit to society if the gainers could, in theory, compensate the losers even though no compensation payments are made.

Including costs and benefits of non-marketed environmental goods and services in a CBA provides the opportunity to reduce market failure that is introduced by incomplete property rights structures. Sources of market failure due to incomplete property rights associated with many non-marketed environmental goods and services include their public good character (for example, Tietenberg and Lewis, 2009). Public goods are characterised by non-excludability and indivisibility in consumption. Indivisibility in consumption implies that the consumption of the public good by one person does not diminish its availability (in quantity and quality) for others. Non-excludability is the consequence of incomplete property rights and implies that it is either infeasible or too costly to exclude anyone who does not pay for consuming the public good. Both non-excludability and indivisibility in consumption provide incentives that lead to an economically inefficient undersupply of public goods by the private sector.

However, incorporating welfare changes related to non-marketed environmental goods and services with public good character in CBA may pose challenges. A range of technical and conceptual challenges as well as moral reservations associated with the inclusion of non-marketed benefits in CBA has been debated extensively in the literature¹. Arguments proposed include concerns about distributional effects; double counting of benefits and costs; and critique addressed towards the anthropocentric worldview underpinning CBA. A particular challenge lies in obtaining value estimates for non-marketed environmental goods and services with public good character. Since these types of goods and services are not sold and bought in markets, market valuation methods are not applicable. Methods that allow the valuation of non-marketed benefits and costs can be divided into (i) revealed preference methods, such as hedonic pricing, and the travel cost method, and (ii) stated preference methods, including the contingent valuation (CV) method, and the discrete choice experiment (DCE) method. The former estimates values through data on market behaviour, whereas the latter infers values using data obtained by consumer surveys (for example, Hanley and Barbier, 2009). While revealed preference methods are restricted to an elicitation of use values, stated preference methods, and in particular CV and DCE, allow the elicitation of both use and non-use values associated with environmental goods and services (see Section 2.5).

¹ For a detailed discussion see, for instance, Graves (2007), Hanley (2001), Miller and Patassini (2005), and Pearce et al. (2006).

The impact of potential temporal value variability and effects of the ongoing evolution of stated preference methods on CBA outcomes have not received much attention. Policy decisions concerned with the allocation of natural resources often have long-term and irreversible impacts on the environment. Failing to account for changes in values and valuation across time may result in an over- or underestimation of long-term benefits and costs. A policy decision based on a CBA that ignores these dimensions may not be able to exploit potential improvements in economic efficiency.

1.2 Research motivation

For over forty years, economists have been engaged in developing, employing and refining non-market valuation methods to estimate non-marketed environmental benefits and costs, in monetary terms. This history of non-market valuation provides an opportunity to make comparisons of value estimates and the evolution and performance of valuation methods over time.

Only a limited number of studies has been conducted in this context. Most studies have been concerned with comparing value estimates obtained at different points in time over relatively short time spans as well as with analysing the temporal reliability of non-market valuation methods. A temporally reliable valuation method produces test-retest value estimates which are not significantly different from each other, assuming that the underlying variables impacting on these estimates (for example, income, prices of related goods, and preferences) have not changed significantly, and carry-over (also called recall) effects are absent (Whitehead and Hoban, 1999). Some studies suggest that, over relatively short time periods, the CV method produces reliable results (for example, McConnell et al., 1998; Reiling et al., 1990). The assumption of stable preferences may be justifiable over short time periods but is questionable when exploring extended time spans. Hence, separating the effects of a temporarily unreliable non-market valuation method from those of preference change poses a considerable challenge.

Little is known about the impact on value estimates of the evolution on non-market valuation methods and value changes across extended time periods. A better understanding of preference formation and evolution provides the opportunity to gain insights into the direction and extent of changes in environmental values over time.

Such knowledge has the potential to improve the explanatory power of economic models concerned with consumer behaviour. Questions about the 'shelf life' of, for instance, existence value estimates, are of key policy importance (Stevens et al., 1994, p.355). Decisions about policy changes affecting the natural environment often have long-term and irreversible impacts and thus require a forecast of future preferences and an extrapolation of value estimates over extended time periods. Frequently, benefit estimates employed in a CBA are obtained some significant time prior and are transferred across space and time using the benefit transfer method (BT) (for example, Garrod and Willis, 1999). Even though the valuation method itself may exhibit temporal stability, factors that determine environmental value such as the underlying preferences or constraints may change over time.

If, for example, a value were to grow faster than the discount rate used in a CBA, benefit estimates would have to be adapted before it would be meaningful to transfer them across time and space. Krutilla and Cicchetti (1972) discussed the relationship between increasing marginal benefit and the discount rate used in the Hells Canyon context. If the increase in marginal benefit was greater than the discount rate, the time horizon of provision would be infinite, implying an infinite present value. Hence, ignoring such value changes across time (represented by a value growth rate that is larger than the discount rate) would result in an underestimation of marginal environmental value due to an incorrectly calculated time horizon. On the other hand, as marginal environmental values change over time, it is likely that people will adjust by changing both supply and demand. Even though an infinite present value is unlikely, this thought experiment demonstrates that economic choices concerned with the temporal allocation of natural resources that ignore dynamic values may reduce intertemporal efficiency. A better understanding of the relationship between values and underlying variables impacting on them through time would thus support more informed decision making.

Moreover, extreme events such as large-scale environmental pollution (for example, the Exxon Valdez oil spill), unusual weather events (for example, Hurricane Katrina) or long-term threats (for example, climate change) may change environmental values and underlying preferences. Using non-market valuation methods such as CV that involve 'snap-shot' scenarios to estimate compensation for environmental damage raises the question of the 'right' point in time to be used as a reference. According to Carson et al.

(1997), immediate short-term responses to an extreme event may change over time with increased availability and distribution of more information regarding the causes and consequences of that event. Additional information gives the individual the opportunity to reassess their value associated with the incident.

Apart from value variability caused by changes in supply and demand, the evolution of non-market valuation methods may have affected value estimates, even though the underlying true value has remained stable. The evolution of stated preference methods over more than forty years may have influenced value estimates through procedural changes in sampling methods, survey designs, elicitation formats, underlying economic models, econometric analysis and ways of interpreting the results. A comparison of value estimates obtained by superseded preference valuation tools with those elicited using state-of-the-art tools may provide information to evaluate the usability of the former in a current or future context. Analysing incentive compatibility properties of elicitation formats used in stated preference methods has been one focus of the history of developing, employing and refining non-market valuation methods.

A possible explanation for the emphasis on this particular issue is that the efficiency level of decisions concerning the allocation of non-market environmental goods and services with public good character depends on individuals truthfully disclosing their preferences, which are only known by each individual. According to mechanism design theory, however, revealing true preferences might not be an individual's optimal strategy for a given social choice function (for example, Mas-Colell et al., 1995). Hence, analysing incentive compatibility properties of alternative elicitation formats supports a better understanding of impacts related to the evolution of the stated preference methods on value estimates. A better understanding of incentive compatibility properties of elicitation formats provides the opportunity to decrease strategic response in surveys used to elicit non-market values, and thus increase the efficiency of natural resource allocations.

1.3 Research goals

The overall goal of this thesis is to enhance the understanding of temporal impacts on values attached to non-marketed environmental goods and services with public good character. This goal will be achieved by making comparisons of value estimates and the

evolution and performance of stated preference valuation methods over time. In 1978, the first CV application undertaken in Australia was focused on Nadgee Nature Reserve on the far south coast of New South Wales (Bennett, 1981). The intention is to replicate that study using both the original and state-of-the-art stated preferences valuation methods. The comparison over time afforded by this replication will provide insights into the extent and direction of changes in non-marketed environmental benefits in Australia and the impact of methodological evolution on value estimates. Such insights will be supportive of improved efficiency in resource allocation.

Hence, the overall research goal can be broken down into a conceptual goal, a methodological goal, and a policy goal. The conceptual goal is to gain a better understanding of the direction and underlying drivers of a value change or stability across extended time periods. The methodological goal is to expand the knowledge about potential impacts of the evolution of stated preference valuation methods on value estimates. A focus will be on the impact of alternative elicitation formats on value estimates associated with the evolution of stated preference methods. Drivers of such impacts investigated in this thesis are potentially stabilising effects of CV elicitation formats and potential impacts of varying incentive compatibility properties associated with alternative DCE elicitation formats. The results generated through the conceptual and the methodological goal feed into the policy goal. Information about the direction and motivation of temporal value change as well as insights about potential impacts of the evolution of stated preference methods will assist policy makers to improve efficiency in resource allocation.

1.4 Thesis outline

Chapter 1 (Introduction) has provided an overview of the research context, discussed the research motivation, and stated the overall research goal, which was disaggregated into a conceptual, a methodological, and a policy goal. The remainder of this thesis is structured as follows.

Chapter 2 (Defining value) outlines the economic theory required as a base for an analysis of value variability across time. This chapter reviews the principles of neo-classical microeconomic theory of consumer choice as a foundation for defining economic value. Hence, consumer preferences, the role of constraints, and the concept

of utility maximisation are introduced. This is followed by a specification of the definition of economic value related to non-marketed environmental goods and services with public good character. The strictly economic perspective is then widened by deliberating the notion of value from a broader perspective using social-psychological models as an example. Also presented in this chapter are definitions of different types of economic value which people attach to the environment. The concept of total economic value is introduced and its limitations and relevance to this thesis are discussed.

Chapter 3 (Environmental values across time) uses the definition of economic value related to non-marketed environmental goods and services with public good character presented in Chapter 2 to introduce the economic theory behind potential temporal value variability. The introduction leads into a discussion of possible drivers of value change across time, which can be grouped into demand side and supply side drivers. This is followed by a discussion of the potential impact of changing preferences and budget constraints, changing scales of provision, margining availability of close substitutes and associated relative scarcity, and reference point relativity on the temporal value variability.

In Chapter 4 (Environmental valuation across time) the methodological evolution of stated preference methods is introduced. Following a general introduction about the history of stated preference methods, this chapter leads into a discourse about DCE as an alternative to the CV method. This is succeeded by an argument about potential impacts of varying elicitation formats associated with the methodological evolution on value estimates. In detail, this chapter debates the possibility of prominence in response in open-ended CV and its consequences on eliciting value estimates across time. This is followed by a deliberation of incentive compatibility properties and strategic response in DCE .

Chapter 5 (Research questions and hypotheses) develops the research questions and testable hypotheses to achieve the research goals stated in Chapter 1. The research questions and the testable hypotheses are derived from the literature review provided in Chapters 1 to 4. After introducing the two research questions this chapter states the testable hypotheses related to temporal value change and valuation variability across time, respectively.

The case study used to investigate the two research questions stated in Chapter 5 is introduced in Chapter 6 (Case study). This chapter describes Nadgee Nature Reserve, defines its economic values, and gives an overview of policy change, which have affected its environmental characteristics. This is followed by a presentation of the reference study used as a baseline for the comparison across time. Based on this information, the need to adjust the hypotheses to the case study is emphasised.

The adjusted research hypotheses and the research design and methods for testing are presented in Chapter 7 (Research design and hypothesis testing) presents. This chapter develops the research design and methods related to the analysis of value change across time, which is concentrated on analysing changes in stated marginal willingness to pay (WTP) and explaining such changes from the demand and the supply side. This is followed by a section that establishes the research design and methods associated with the investigation of impacts related to the methodological evolution of stated preference methods on value estimates. A focus rests on examining effects of alternative stated preference methods (CV versus DCE), prominence in response in open-ended CV, and incentive compatibility properties and strategic response in DCE.

This is followed by Chapter 8 (Data collection), which describes the data collection process and associated challenges and limitations. The introduction leads into the data collection process, which is separated through the use of two alternative stated preference methods. This is succeeded by a report on the sampling, questionnaire design, data coding, and sample characteristics related to the CV survey. This chapter then communicates the sampling, the questionnaire design, the data coding, the experimental design, the conduction and results of the focus groups and the pilot survey, and the sample characteristics of the DCE survey.

Chapter 9 (Analysing values across time) opens with an introduction that sets forth an analysis of the temporal variability of the values people attach to non-marketed environmental goods and services with public good character using the preservation value of Nadgee Nature Reserve as an example. In detail, this chapter explores temporal changes in stated marginal WTP obtained in an open-ended CV, and attempts to explain these changes from the demand side and supply side. This section additionally summarises the maintained assumptions formulated throughout the thesis.

Chapter 10 (Analysing impacts of methodological evolution) concentrates on an investigation of the impact of methodological evolution of stated preference methods on value estimates. After summarising the relevant maintained assumptions stated throughout this thesis, this chapter focuses on three elements that may influence the variability of value estimates across time. Hence, this chapter explores the influences of DCE as an alternative to CV, examines a potential stabilising effect of prominence in response in open-ended-CV, and analyses incentive compatibility properties and strategic response in DCE.

Finally, Chapter 11 (Conclusion) opens with a discussion of the main research findings presented in this thesis. This is followed by an overview and assessment of the challenges and limitations this study presented. Together the discussion of the main research findings and the assessment of challenges and limitations provide the basis for a debate about their policy implications and recommendations for future research.

2 DEFINING VALUE

2.1 Introduction

Analysing temporal variability of non-market values attached to environmental goods and services with public good character requires a definition of the notion of value. The first goal of this chapter is to derive such a definition.

‘Value’ is a widely used term in the academic literature. However, its interpretation differs significantly across disciplines and depends on the respective underlying paradigms. Brown (1984) suggested a general differentiation between held value and assigned value. Held values are defined by psychologists as higher-order guiding principles for individuals’ daily behaviour (for example, Rokeach, 1973; Schwartz, 1992). For example, psychologists analyse people’s value orientations and ethical beliefs with regard to natural resources. In contrast to this, economists estimate assigned value underpinned by consumer preferences and budget constraints (for example, Varian, 2006). It has to be noted though, that philosophers, environmentalists, and economists who reject a pure anthropocentric worldview discuss the existence and relevance of intrinsic environmental value irrespective of human preferences and needs (for example, Turner, 2003). Apart from that, assigned value has been misinterpreted as being equivalent to revenue, cost or price. Brown (1984), for example, points out that prices of environmental goods and services reflect and signal resource scarcity, but must not be confused with the notion of value.

In this thesis, value is specified in economic terms as a change in individual welfare using the neo-classical microeconomic theory of consumer choice as a foundation. Some of the welfare measures suggested by this theory are estimated with respect to a change in price of the good under consideration. Many environmental goods and services, however, have a public good character and are not sold and bought in markets (see Section 1.1). This implies that welfare measures that are either based on changes in price or continuous changes in quantity are not applicable. Hence, the second goal of this chapter is to specify the definition of economic value for this type of environmental goods and services.

The remainder of this chapter is structured as follows. Section 2.2 reviews individual welfare measures based on the principles of neo-classical microeconomic theory of consumer choice as the foundations of economic valuation. Based on this general theoretical background of defining economic welfare, Section 2.3 then focuses on welfare measures used to express the economic value of non-marketed environmental goods and services with public good character. Special attention is given to a discussion of the underlying concepts of preferences and constraints. This is followed by Section 2.4, which aims to broaden the understanding of the notion of value. Models suggested by social psychological research will be discussed and their relevance to understanding the value people attach to non-marketed environmental goods and services will be explored. Section 2.5 sets out the different types of economic value attached to environmental goods and services. The concept of total economic value is introduced and its relevance and its limitations in this thesis are discussed. Section 2.6 provides a conclusion.

2.2 Theoretical background

The neo-classical microeconomic theory of consumer choice provides the foundation for the definition of economic value (for example, Varian, 2006). It is founded on the assumption that consumers make economic choices based on individual preferences. This preference based approach to consumer choice describes consumers' preference relations over a set of feasible, mutually exclusive goods and services bundles in a given consumption set (for example, Mas-Colell et al., 1995)². Following Frank (2003), a preference relation can be represented by means of a utility function, which reflects the amount of satisfaction (measured in *utils*) that is provided by consuming a bundle of goods and services, and, likewise to an indifference map, provides information about an individual's preference ordering:

$$U(X, Y) = XY, \quad [1]$$

with utility U , and quantity of good or service X and Y (for example, Frank, 2003). Bundles that are located on a higher indifference curve yield a higher level of utility than those on a lower indifference curve. The assumed desirability condition underlying

² Much of neo-classical microeconomic theory of consumer choice is based on a range of assumptions about these preference relations, including (i) rationality, (ii) desirability, (iii) continuity, and (iv) convexity (for example, Mas-Colell et al., 1995).

the concept of preference relations ('more is better') implies that an individual tries to maximise their utility by choosing the bundle on the highest feasible indifference curve (for example, Mas-Colell et al., 1995). Since all possible bundles of X and Y that yield the same amount of utility are represented by an indifference curve, trade-offs along an indifference curve between any two goods can be expressed either (1) as the marginal rate of substitution defined as the absolute value of the first derivative of the indifference curve, or (2) as a change in marginal utility MU while satisfying the following equation (for example, Frank, 2003):

$$MU_X \Delta X = MU_Y \Delta Y. \quad [2]$$

That is, the change in utility due to a change in consumption of good X is fully offset by the change in utility due to a change in consumption of good Y .

However, utility maximisation behaviour is not solely a result of individual preferences but bounded by an individual's consumer budget. The so-called competitive budget set is defined as the set of all goods and services bundles feasible to an individual with given wealth facing a given price vector (for example, Mas-Colell et al., 1995). Following the neo-classical microeconomic theory of consumer choice, an individual maximises their utility by choosing their most preferred consumption bundle in the competitive budget set, assuming rational, continuous and locally non-satiated preference relations; this consumer choice problem is denoted as the utility maximisation problem (for example, Mas-Colell et al., 1995). The function that assigns the set of optimal consumption bundles in the utility maximisation problem for each price-wealth situation is defined as the uncompensated or Marshallian demand function (for example, Mas-Colell et al., 1995). The Marshallian demand function can be used to measure individual welfare effects. An individual welfare effect due to a change in price at a fixed wealth level can be measured in terms of a change in Marshallian consumer surplus (for example, Varian, 2006).

An alternative welfare measure is Hicksian consumer surplus (Hicks, 1945). The Hicksian consumer surplus is based on a Hicksian demand function, defined as the function that assigns the set of optimal consumption bundles in the expenditure minimisation problem³ for each price-utility situation. The Hicksian or compensated demand function adjusts the level of wealth by means of the Hicksian wealth

³ The expenditure minimisation problem is the dual of the utility maximisation problem.

compensation as prices vary while holding the level of utility constant. Thus, in comparison to the Marshallian demand function, the Hicksian demand function eliminates the income effect, and hence, reflects solely the substitution effect (for instance, Mitchell and Carson, 1989). Hicks (1939) used the money metric indirect utility function to develop two measures of individual welfare change, the compensating variation, denoted by CV , and the equivalent variation denoted by EV ⁴. These two measures of individual welfare are defined as the wealth change needed to compensate for a price change as to keep utility constant at some fixed level, and with

$$e(p^0, u^0) = e(p^1, u^1) = w, \quad [3]$$

the compensating (equivalent) variation is the difference between the wealth level w at the utility level u^1 and the wealth level w at the utility level u^0 after a price change from p^0 to p^1 with respect to prices p^1 (p^0) (for example, Mas-Colell et al., 1995):

$$CV(p^0, p^1, w) = w - e(p^1, u^0) = e(p^1, u^1) - e(p^1, u^0), \quad [4]$$

$$EV(p^0, p^1, w) = e(p^0, u^1) - w = e(p^0, u^1) - e(p^0, u^0). \quad [5]$$

The Marshallian consumer surplus, as well as the Hicksian compensating variation and equivalent variation, express welfare changes triggered by a change in price. In an environmental economics context, however, individual welfare change is frequently triggered by a change in quantitative or qualitative characteristics of an environmental good or service rather than a change in price. Hence, the individual welfare effect due to a change from a current level of provision to a suggested alternative one has to be analysed. Such a change can influence directly an individual's utility function (for instance, a change in an individual's utility due to a change in the pristine nature of a national park), or impact indirectly through entering into a household or firm production function (for example, Bockstael and Freeman, 2005; Freeman, 1993). The remainder of this thesis will focus on the former, with the following section defining economic value of non-marketed environmental goods and services with public good character.

⁴ Both the compensating variation and the equivalent variation are measures to express individual welfare effects due to a price change. In general, however, the monetary value for an individual welfare change obtained by these two measures differs. However, if the income effect is zero the compensating variation and the equivalent variation are equally represented by the Marshallian consumer surplus (Marshall, 1920).

2.3 Economic value of non-marketed environmental goods and services

Analogous to the Hicksian measures of individual welfare with respect to a change in price, Hicks (1945) also suggested the compensating surplus, denoted by CS , and equivalent surplus denoted by ES , with respect to a change in quantity as:

$$CS(q^0, q^1, p, w) = w - e(p^0, q^1, u^0) = e(p^0, q^0, u^0) - e(p^0, q^1, u^0), \quad [6]$$

$$ES(q^0, q^1, p, w) = e(p^0, q^0, u^1) - w = e(p^0, q^0, u^1) - e(p^0, q^1, u^1) \quad [7]$$

Mäler (1971; 1994) connected the Hicksian individual welfare measures to environmental goods and services by incorporating a vector of non-marketed environmental goods and services with public good character – henceforth abbreviated as *environmental goods and services* (see Section 1.1). He emphasised that some of these environmental goods and services are available only in fixed quantities, which behave as constraints and thus limit individuals' choices. Hence, an individual holds preferences towards a private good vector x (variable consumption level) with an associated price vector for all market goods p and a public good vector q (given consumption level), which can be represented by a utility function⁵.

Integrating the inverse Hicksian demand function with respect to q , the compensating surplus and the equivalent surplus can be expressed as (Mitchell and Carson, 1989):

$$CS = \int_{q^0}^{q^1} h(p, q, u^0) dq, \quad [8]$$

$$ES = \int_{q^0}^{q^1} h(p, q, u^1) dq. \quad [9]$$

For a discrete change in q , both the compensating surplus and the equivalent surplus represent the maximum a consumer is willing to pay (willing to accept) to enjoy (forgo) an increase or to prevent (accept) a decrease in the level of provision of some environmental asset (Mitchell and Carson, 1989).

⁵ The utility function is assumed to be continuous and non-decreasing in x and q , and quasi-concave in q . Hanemann (1999) points out that the utility function representing this preference relation cannot necessarily be assumed to display quasi-concavity with respect to q , as usually assumed with respect to x since q is given and not freely selectable by the consumer.

The choice of whether to use the compensating measure or the equivalent measure depends on the property rights allocation. The compensating measure reflects a situation with the consumer having the property rights to keep the current level of utility, whereas the equivalent measure represents a situation in which the consumer is entitled to some alternative level of utility (Mitchell and Carson, 1989). For instance, for an increase in the level of biodiversity the compensating surplus expresses the maximum a consumer is willing to pay to benefit from the additional biodiversity to keep their utility level constant. In contrast, for an announced decrease in biodiversity, equivalent surplus can be interpreted as the maximum a consumer is willing to pay in order to prevent this biodiversity loss to maintain his utility constant, only having a right to the diminished supply.

The concept of compensating surplus and equivalent surplus measured in terms of marginal WTP are demonstrated in Figure 1. It displays a situation where an individual is indifferent between consuming bundle A (q_c^1, q_e^0) and bundle B (q_c^0, q_e^1), with q_c and q_e representing the quantities of the composite good W (expressed as wealth/ income) and the environmental good or service E , respectively. An increase in supply of E from q_e^0 to q_e^1 would increase the individual's utility, represented by moving to a higher indifference curve with associated higher utility from bundle A to bundle C (q_c^1, q_e^1). The marginal rate of substitution between W and E quantifies the increase in utility an individual would enjoy from the increase in the supply of E . This increase in utility is expressed as moving from bundle A to bundle B (compensating surplus), or as preventing a move from bundle B to bundle A (equivalent surplus). An individual would be willing to pay $q_c^1 - q_c^0$ of W in order to enjoy an increase (prevent a decrease) in supply in E from q_e^0 to q_e^1 (q_e^1 to q_e^0), while keeping their utility level constant at utility level u_0 (u_1), with $u_1 > u_0$.

As shown in Figure 1, the magnitudes of equivalent surplus and compensating surplus are partially driven by the shape of the indifference curves representing individuals' preferences. The absolute value of the first derivative of the indifference curves determines the marginal rate of substitution at any consumption bundle (for instance, Frank, 2003). The steeper the indifference curve, the more of W has to be given up to enjoy q_e^1 of E . That is, individuals with steeper indifference curves are willing to pay

more for an increase in E from q_e^0 to q_e^1 (see Figure 1a) than those with flatter indifference curves (see Figure 1b). However, as illustrated in equations [4] and [5], marginal WTP expressed as compensating surplus and equivalent surplus are not only a function of preferences but are bounded by wealth/ income w . That is, marginal WTP is subject to a budget constraint.

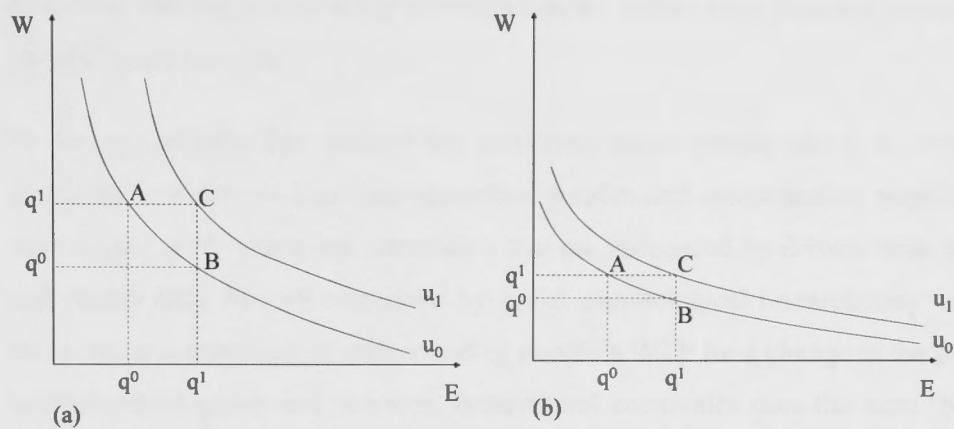


Figure 1: Compensating surplus and equivalent surplus

In general, microeconomic theory of consumer choice suggests that Hicksian compensating surplus and equivalent surplus are influenced by factors of demand and supply. Demand side factors include socio-economic characteristics of individuals such as age, income/ wealth, gender, and education level. These characteristics may influence individuals’ preferences for environmental goods and services. This influence would be represented by the shape of indifference curves and thus the marginal rate of substitution. Additionally, demand side factors may act directly and indirectly as constraints for marginal WTP. Income/ wealth acts directly as a budget constraint or indirectly through the impact of other demand side factors such as the education level or age on income/ wealth.

On the supply side, Hicksian compensating surplus and equivalent surplus may be influenced by the availability of close substitutes for the non-marketed environmental goods and services under consideration. In turn, close substitutes that are perceived as being (nearly) perfect substitutes impact on the relative scarcity of these goods and services. The effect of relative scarcity is represented by the change in the marginal rate of substitution along the indifference curve for different consumption bundles. That is, the more abundant the environmental goods and services the smaller the marginal rate

of substitution. That is, individuals' WTP for a proposed change in the provision of environmental goods and services follows the law of diminishing marginal utility and would be lower the less scarce it is currently.

This discussion shows that drivers on the demand and supply side may influence the value individuals attach to non-marketed environmental goods and services with public good character through their impact on individuals' preferences and constraints. Hence, analysing whether and how these factors change across time supports explaining value variability across time.

So far, this chapter has defined the economic values people attach to environmental goods and services as Hicksian equivalent surplus and compensating surplus based on individuals' preferences and constraints that are influenced by drivers from the demand and supply side. Models suggested by social psychological research may complement this economic approach of understanding people's WTP for a change in the provision of environmental goods and services. Behavioural economics uses the term 'behavioural failures' to describe individuals' behaviour, which challenges neo-classical microeconomic theory of consumer choice (for example, Shogren and Taylor, 2008). Shogren and Taylor (2008) distinguish between bounded rationality⁶, bounded self-interest⁷ and bounded willpower⁸ (for example, Mullainathan and Thaler, 2000; Shogren and Taylor, 2008). The goal of the following section is to widen the perspective on value beyond a strictly neo-classical economic understanding.

2.4 The notion of economic value from a broader perspective

This section explores exemplarily the potential of a range of models that are based on social-psychological research in their capacity to improve our comprehension of people's WTP for a change in the provision of environmental good and services. The purpose of this section is to widen the perspective beyond a strictly economic understanding of economic value and to discuss if social-psychological research

⁶ Bounded rationality reflects individuals' limited abilities to process all information relevant for rational choices (for example, Mazzotta and Opaluch, 1995, Mullainathan and Thaler, 2000, Shogren and Taylor, 2008, Simon, 1957).

⁷ Bounded self-interest implies that individuals can exhibit altruistic behaviour (for example, Shogren and Taylor, 2008; Smith, 1998). It can be argued, however, that altruistic behaviour may increase utility motivated by self-interest. This would imply that bounded self-interest is impossible.

⁸ Bounded willpower describes individuals' lack of self-control, which sometimes results in choices that contradict their long-term interests (for example, Mullainathan and Thaler, 2000; Shogren and Taylor, 2008).

findings may be supportive in understanding the role of social-psychological factors in explaining marginal WTP. The social-psychological models discussed in this section are the theory of planned behaviour (Ajzen, 1985), the moral norm-activation theory of altruism (Schwartz, 1977), and the value-belief-norm theory (Stern, 2000).

According to the theory of planned behaviour (Ajzen, 1985), human action is based on attitudes toward behaviour (a function of the strength of a belief that the behaviour will result in a certain outcome and the evaluation of that outcome); subjective norms (personal perceptions about a particular behaviour affected by the judgment of relevant others); and perceived behavioural control (personal beliefs about the existence of variables facilitating or impeding the performance of the behaviour)^{9,10}. Ajzen and Peterson (1988) suggest that individuals may exhibit different attitudes toward different aspects of a given scenario and thus attach different values to different value dimensions. Hence, they stress the importance of differentiating between attitudes toward the good itself (for example, maintaining a national park), attitudes toward a certain behaviour as to the good (for example, paying an entrance fee), and attitudes toward a policy inducing a change as to the good (for example, access regulations). Together, attitudes toward behaviour, subjective norms, and perceived behavioural control determine behavioural intentions, defined as an indication of an individual's willingness to perform a particular behaviour (intention) (Ajzen, 1985; 1991; 2001; Fishbein and Ajzen, 1975). A limited range of studies have incorporated the theory of planned behaviour to test its applicability and relevance in the context of assessing non-marketed values (Ajzen and Driver, 1992; Bernath and Roschewitz, 2008; Larson and Lach, 2008; Pouta and Recola, 2001; Werner et al., 2002). Some of these studies suggest that applying an attitude-behaviour framework significantly improves the explanatory power of economic models explaining marginal WTP.

A range of studies concerned with non-marketed values have included the new ecological paradigm (NEP)¹¹ as a variable representing general environmental attitudes.

⁹ In general, a belief is defined as '... the subjective probability of a relation between the object of the belief and some other object, value, concept, or attribute. Thus a person may believe that he possesses certain attributes (for example, that he is intelligent, honest, punctual, etc.), that a given behaviour will lead to certain consequences, that certain events occur, etc.' (Fishbein and Ajzen, 1975 p.131).

¹⁰ Besides behavioural, normative and control beliefs, additional kinds of beliefs are assumed to explain behaviour such as personal moral obligation or feeling of responsibility (Beck and Ajzen, 1991; Gorusch and Ortberg, 1983; Pomazal and Jaccard, 1976; Schwartz and Tessler, 1972)

¹¹ The NEP, originally proposed by Dunlap and Van Liere (1978) and revised by Dunlap et al. (2000), is a widely used measure of public concern for the environment reflecting five hypothesised aspects of an ecological worldview:

These studies provide evidence for significant relationships between environmental attitudes, marginal WTP and underlying motivation such as egoistic motives and rights-based beliefs (for example, Cooper et al., 2004; Kotchen and Reiling, 2000; Spash, 2006). The literature suggests that the NEP is related to a broad range of ecological attitudes and behaviours and correlated with socio-economic variables such as income, age, and education (Casey and Scott, 2006; Dunlap et al., 2000).

The moral norm-activation theory of altruism (Schwartz, 1977) states that altruistic behaviour results from personal moral obligations activated in individuals who are aware of possible threats posed to others by particular conditions such as environmental problems (awareness of consequences), and the perceived ability to reduce those threats by making choices (ascription of responsibility to self). Several studies suggest the theory's applicability in a range of environmental contexts (e.g., Black et al., 1985; Guagnano et al., 1995; Schultz and Zelenzky, 1999; Wiidegren, 1998). Following the rationale of the norm-activation theory of altruism, awareness of environmental problems as well as awareness that certain choices could attenuate those problems may influence preferences and thus marginal WTP.

The value-belief-norm theory of environmentalism developed by Stern (2000) links value theory, the norm-activation-theory of altruism (Schwartz, 1977) and the NEP (Dunlap and Van Liere, 1978; Dunlap et al., 2000) that has been discussed earlier in this chapter. In comparison to the norm-activation-theory of altruism (Schwartz, 1977) suggesting awareness of consequences, Stern's theory is based on beliefs about adverse consequences (perceived consequences). The value-belief-norm theory (for example, Snelgar, 2006; Stern, 2000; Stern et al., 1999; Stern et al., 1993; Stern et al., 1995) hypothesises that individuals' environmental behaviour is motivated by the extent to which choices have adverse consequences for valued objects, which in turn depends on individuals' value orientations, that is, to what extent these consequences impact on others (altruistic orientation), on the self (egoistic orientation), and on the biosphere (biospheric orientation).

Consider, for example, a proposed development of a national park. For supporters of the project, beliefs about adverse consequences if the development does not proceed may be, for instance, the loss of employment opportunities for self (egoistic orientation) or

(1) reality of growth limits, (2) fragility of the balance of nature, (3) anti-anthropocentrism, (4) rejection of exemptionalism, and (5) possibility of an eco-crisis.

others (altruistic orientation). On the other hand, opponents may hold beliefs about adverse consequences in the case of a development such as the impact of biodiversity loss on society (altruistic orientation) or on other species (biospheric orientation). In this way, the value-belief-norm theory interprets not just exclusively pro-environmental but environmental attitude-behaviour relations in general (Snelgar, 2006). Based on these value orientations reflected in individuals' ecological worldviews, the value-belief-norm theory suggests that beliefs about adverse consequences for valued objects and the perceived ability to attenuate those consequences activate personal norms for pro-environmental behaviour (sense of obligation to take corrective actions). Consequently, information impacting on the beliefs about adverse consequences as well as about the ability to take corrective actions may influence personal norms, and thus preferences and marginal WTP. A range of studies have empirically tested models based on the value-belief-norm theory (for example, Gärling et al., 2003; Hansla et al., 2008; Joireman et al., 2001; Nordlund and Garvill, 2003; Slimak and Dietz, 2006; Snelgar, 2006; Stern et al., 1999). Few studies have applied this theory in the context of assessing non-marketed values (for example, Spash, 2006).

The discussion of these theories demonstrates that social-psychological research may be supportive in understanding the role of social-psychological factors in explaining marginal WTP. In particular, the roles of information and experience may be relevant in forming, shaping and expanding preference sets. However, the available data is often limited – especially if the analysis depends on a comparison with an existing study as is the case in this thesis. Furthermore, the validity of models that include such social-psychological factors strongly depends on how the information about these factors was obtained. The validity of using information on these factors depends on whether the respondent accounts for opportunity costs or is just stating their attitude as a matter of principle without considering any trade-offs. The latter may lead to biased results. For example, someone may have a strong positive attitude towards the provision of a higher level of some environmental goods and services but is not willing to give up something in return. Using this attitudinal variable to explain marginal WTP would not be justified by the microeconomic theory of consumer choice (see Section 2.2). According to this theory, preferences are assumed to reflect choices between alternatives subject to constraints considering opportunity costs, whereas, following psychological theory, attitudes are assumed to express the desirability of a single good or behaviour without the requirement for either prioritisation or consideration of trade-offs (Green and

Tunstall, 1999). For these reasons, the examination of value variability across time presented in this thesis focuses on an investigation of traditional variables of demand and supply as well as behavioural variables representing exposure to and experience with the environmental goods and services. Available data on attitudes, however, will not be ignored but treated with caution.

2.5 The concept of total economic value

Section 2.4 defined the notion of value and suggested compensating surplus and equivalent surplus as measures for value people attach to environmental goods and services. This section will provide a definition of the range of value types people derive from these environmental goods and services. This is considered relevant since it is assumed that temporal variability of value people attach to environmental goods and services varies across the different value types they provide. These different value types have been organised within the concept of total economic value (for example, Turner, 1999).

Total economic value comprises use and non-use values. These two value types are divided further into a range of sub-types such as direct use value, indirect use value, option value, quasi-option value, existence value, vicarious value and bequest value. The classification of these sub-types has been extensively and controversially discussed in the economic literature (for example, Arrow and Fisher, 1974; Freeman, 1993; Johansson, 1993; Mitchell and Carson, 1989; Pearce, 1993; Randall and Stoll, 1982; Turner, 1999; 2003; Turner et al., 2000; Young, 1991).

A frequently used classification of total economic value is the one established by Turner et al. (2000) connecting functions, uses and values of wetlands. Even though this classification has been developed using the example of wetlands, it can be adopted to classify values of any ecosystem. According to the classification of Turner et al. (2000), use value originates from direct or indirect utilisation of ecosystem goods (for example, timber or recreational opportunities) and services (for example, groundwater recharge or toxics retention), respectively. Value sub-types frequently associated with use value are option value and quasi-option value. Option value is described as the benefit obtained from preserving a resource for potential use in the future (Turner, 2003; Turner et al., 2000), whereas quasi-option value is defined as the benefit derived from waiting for

revised information before forgoing the option of conserving a resource for potential utilisation in the future (Arrow and Fisher, 1974).

Non-use value (also called ‘passive-use value’) describes benefits which do not depend on the current or future use of a resource. Various sub-types of non-use value have been debated in the literature, including bequest, existence and vicarious value. The notion of non-use value goes back to Krutilla’s observation that individuals may value natural amenities simply for their existence (Krutilla, 1967)¹². Existence value has been described as benefits derived from the knowledge that an environmental resource such as an individual species or a whole ecosystem continues to exist, independent of any direct or indirect use (Turner, 2003; Turner et al., 2000). Bequest value is associated with benefits derived from conserving a resource to ensure its availability for future generations (Turner, 2003). However, Pearce (1993) notes that the preferences of future generations may be uncertain. Thus, bequest value is not necessarily associated with an expectation that the resource will be used in the future.

The concept of total economic value is a useful framework for valuing the natural environment since both market and non-market values are explicitly considered. However, it is not without limitations. Environmental goods and services may influence two or more value components simultaneously, whereas competing uses may reduce each others benefits (Turner et al., 2000). Therefore, ecosystem complexity frequently prevents a precise categorisation of the economic value of environmental resources. Ignoring ecosystem complexity and neglecting trade-offs between competing uses may result in confounding effects when analysing value variability across time. The concept of total economic value will be used to define the value types associated with the case study (Chapter 6). Hence, Section 2.5 has provided the theoretical foundation for that task.

¹² The hypothesis that altruism may be the underlying motive of non-use value triggered an ample and controversial debate about the validity of non-use value, and especially existence value, as a measure of economic benefits of public goods (for example, Ajzen and Driver, 1992; Andreoni, 1990; Becker, 1993; Chilton and Hutchinson, 2000; Kahneman and Knetsch, 1992; Rolfe and Bennett, 1996; Stevens et al., 1991; Stevens et al., 1994; Turner, 1999). This thesis argues that even though non-use value may be motivated by altruism it is still justified on the grounds of the microeconomic theory of consumer choice since altruism may increase an individual’s utility.

2.6 Conclusion

Chapter 2 has discussed and defined the notion of value. This was necessary since its definition differs significantly across disciplines and depends on the respective underlying paradigms. This thesis follows the microeconomic theory of consumer choice, which is based on the assumption that consumers have an anthropocentric worldview and make rational economic choices based on individual preferences and constraints. It was argued that due to the public good character of many non-marketed environmental goods and services implying fixed quantities and the absence of prices, the appropriate welfare measures to quantify the value of marginal changes in supply are Hicksian compensating surplus and equivalent surplus. In this regard, the compensating surplus and the equivalent surplus are defined as the maximum a consumer is willing to pay to enjoy an increase or to prevent a decrease in the level of provision of some environmental goods and services. This implies that the definition of economic value in terms of Hicksian surplus depends on the distribution of property rights. The compensating measure accounts for a situation with the consumer having the property rights to keep the current level of utility, whereas the equivalent measure is appropriate for a situation in which the consumer is entitled to some alternative level of utility. This implies that choosing the appropriate welfare measure requires information about the property rights situation. This will need to be considered in the case study (Chapter 6).

Additionally, Chapter 2 investigated in more detail the influence of preferences and constraints on marginal WTP. An emphasis was given to the effect of the shape of indifference curves on the marginal rate of substitution. This was followed by a discussion of factors on the demand and supply side that are assumed to influence marginal WTP through their effect on preferences and constraints. This discussion was complemented by widening the perspective on value beyond a strictly economic understanding. A range of models that are based on social-psychological research were examined in their capacity to improve our comprehension of WTP for a change in the provision of environmental goods and services. These discussions provided the foundation for exploring temporal value variability (Chapter 3) and valuation across time (Chapter 4). Based on the general definition of value, this chapter additionally introduced the concept of total economic value. This concept differentiates between different values people derive from environmental goods and services. Differentiating

between different value types and possible overlaps was considered important to avoid confounding effects in the analysis of temporal value variability. Hence, this chapter additionally provided the theoretical foundation for defining the value types that will be analysed in this thesis (Chapter 6).

3 ENVIRONMENTAL VALUE ACROSS TIME

3.1 Introduction

Chapter 2 defined the notion of economic value as Hicksian compensating and equivalent surplus measured in WTP for a change in the level of provision of non-marketed environmental goods and services with public good character. Based on this definition, the aim of Chapter 3 is to discuss the theory behind value variability across time associated with these types of goods and services. This discussion will provide the foundation for one of the objectives of this thesis: analysing temporal value variability over extended time periods.

Neo-classical microeconomic theory of consumer choice suggests that values for non-marketed environmental goods and services with public good character are a function of drivers on both the demand side and the supply side (see Section 2.3). This implies that values may change along with temporal changes in these drivers. So far, only a limited number of studies have examined temporal value variability. Some of these studies have reported relatively stable values over relatively short time spans, whereas others looked at longer time periods and found indications for value variability (Brouwer, 2006; Brouwer and Bateman, 2005; Carson et al., 1997; Loomis, 1989; McConnell et al., 1998; Reiling et al., 1990; Stevens et al., 1994; Teisl et al., 1995; Whitehead and Aiken, 2007; Whitehead and Hoban, 1999; Zandersen et al., 2005a; b). Nevertheless, little is known about the extent and direction of possible changes in inherent environmental values over extended time periods of 30 years or more.

The remainder of this chapter is structured as follows. Section 3.2 discusses possible demand side drivers of value change across time through their influence on preferences and budget constraints. Such drivers include socio-economic variables and changes in the level of information about and familiarity/ experience with non-marketed environmental goods and services. Section 3.3 debates factors on the supply side that may support temporal value variability directly. Such factors include changes in the scale of provision; changes in the availability of close substitutes and associated

changes in relative scarcity; and changing reference points. Finally, Section 3.4 provides a conclusion.

3.2 Temporal value variability – demand side

3.2.1 Preferences

As outlined in Section 2.2, the definition of economic value used in this thesis is based on individuals maximising their utility subject to a budget constraint. This implies that preferences are key determinants of economic value. Consequently, changes in preference relations over time are expected to drive temporal value variability. However, Norton et al. (1998) argue that preferences often are estimated from ‘snapshots’ rather than as a dynamic process, assuming that they are at least stable for the duration of the analysis¹³. They further argue that since preferences are assumed to change relatively slowly the hypothetical construct of temporal stable preferences makes sense in the short term. Many economic choices about policy changes regarding the natural environment, however, have long-term consequences. Assuming time consistency and thus ignoring the dynamics of preferences may lead to reduced efficiency in resource allocation.

As outlined in Section 2.3, a preference change is reflected in a change of indifference curve shapes. As a result, non-marketed environmental goods and services may become more or less attractive relative to the composite good, which would be reflected in the marginal rate of substitution (and thus marginal WTP) for all consumption bundles along these re-shaped indifference curves. As illustrated in Figure 2, marginal WTP associated with a steeper indifference curve is higher than marginal WTP related to a flatter indifference curve. A preference change can be explained in two ways that are based on and complement each other. One way assumes that preference relations are a function of a range of factors including socio-demographic characteristic of individuals such as age, gender, income, cultural background, and education. This suggests that preference relations change if these functional relationships fluctuate. Norton et al. (1998) suggests that individuals’ preferences may fluctuate over time, for instance,

¹³ This view is consistent with the theory of time consistency with respect to intertemporal utility (for example, Mas-Colell et al., 1995).

under the influence of educational campaigns, advertising or changing cultural assumptions. However, these functional relationships are still not well understood. Other still unknown factors may likewise influence these functional relationships. A second way of explaining preference change suggests that preferences may change as a result of a shift in socio-economic characteristics that underpin these functional relationships. For example, an aging population may have different preferences than a population with a high proportion of young people because of a socio-economic shift in age rather than a changed functional relationship between age and preferences.

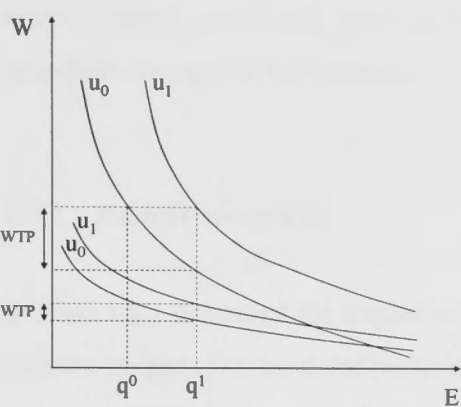


Figure 2: Effect of indifference curve shape on marginal WTP

Temporal fluctuations in preference relations may not be restricted to exiting preferences but may represent an expansion of a preference set. A study undertaken by Tisdell et al. (Tisdell et al., 2008) suggest that people cannot be expected, necessarily, to be familiar with the environmental good. And individuals that are not familiar with certain environmental goods and services do not know their respective preferences. Plott (1996) proposed that respondents may discover their underlying preferences through a learning process rather than possessing stable unknown preferences (value learning). Such learning processes are expected to change preferences and/ or expand the preference set. Such an expansion may be triggered by a shift in socio-economic characteristics. For example, a higher wealth/ income level may provide opportunities for experiences that were unavailable beforehand resulting in the discovery of new preferences. This line of reasoning contradicts the assumption of completeness in preference relations, which proposes that individuals have well-defined preferences, which enables them to compare any two goods and services bundles (for example, Mas-Colell et al., 1995). Consequently, individuals’ preferences for unfamiliar non-marketed environmental goods and services such a good may be controlled over time by an

increasing level of information, knowledge, and experience. This process may reshape existing preferences or may enable consumers to detect unknown preferences and thus expand their preference set (Loomis, 1989; Plott, 1996). For instance, Carson et al. (1997) suggest that immediate short-term responses to an extreme event may change over time with increased availability and distribution of more information regarding the causes and consequences of that event, which gives the individual the opportunity to reassess the incident. Apart from short-term response to extreme events, increasing availability of general information about the environment and environmental education may influence preferences over longer time periods. Hence, values for non-marketed environmental goods and services may be affected by both short-term responses and long-term changes in preferences.

3.2.2 Budget constraint

Section 3.2.1 described the impact of preference change on economic value assuming an infinite budget. For most individuals this assumption does not hold. This implies that marginal WTP is bounded by a budget constraint. That is, a change in disposable wealth/ income is expected to change an individual's competitive budget set affecting marginal WTP. Hence, a shift in disposable income/ wealth across time is predicted to cause a change in marginal WTP. Horowitz (2002) provides an overview of the relevance of income as a variable impacting on marginal WTP by reviewing numerical results of a range of studies that report income elasticities of demand. According to these studies income elasticities of demand range from 0.12 to 0.55 (Krisström and Riera, 1996; McFadden, 1994). It needs to be noted that the effects of changes in individual preferences and competitive budget sets may head in different directions. This may lead to a cancellation of both effects. That is, a stable value does not necessarily imply stable preferences and an unchanged competitive budget set. Of course, socio-economic characteristics such as age, gender, and education may impact indirectly on marginal WTP through their influence on the budget constraint, suggesting, for example, a relationship between the education level and the wage rate. Thus, a change towards better educated people may result in a shift of an individual's budget constraint to the right, reflecting a higher disposable wealth/ income. In summary, the impact of the budget constraint itself is expected to be stable across time.

A change in the competitive budget set due to a change in the budget constraint, however, is supposed to impact on marginal WTP.

3.3 Temporal value variability – supply side

3.3.1 Scale of provision

According to neo-classical microeconomic theory of consumer choice an individual's marginal WTP should be responsive to changes in the scale of provision of non-marketed environmental goods and services – known as scale effect (for example, Carson, 2000; Carson and Mitchell, 1995; Kahneman and Knetsch, 1992; Rolfe et al., 2011; Smith and Osborne, 1996). Of course, according to the microeconomic principle of diminishing marginal utility marginal WTP is not expected to increase proportional to the magnitude of a change. Rather, marginal WTP is expected to increase at a decreasing rate.

This expectation is grounded in the assumption of desirability in preference relations¹⁴. Desirability refers to the assumption that a larger quantity of goods and services is always preferred to a smaller one, assuming that larger quantities are feasible, *ceteris paribus* (for example, Mas-Colell et al., 1995). The desirability assumption is reflected in downward sloping indifference curves, implying that marginal WTP increases with increasing change in the scale of provision (for example, Frank, 2003). As illustrated in Figure 3, a change from provision level q^0 to q^1 yields a lower marginal WTP (Figure 3a) than a change from q^0 to q^2 (Figure 3b), with $q^0 < q^1 < q^2$. For instance, an individual's marginal WTP for the maintenance of a nature reserve covering 50,000 hectares should be higher than for one covering only 10,000 hectares (both compared to the same baseline).

¹⁴ Desirability can be captured by assuming a monotone preference relation, which is satisfied if the commodities are *goods* (instead of *bads*) (for example, Mas-Colell et al., 1995).

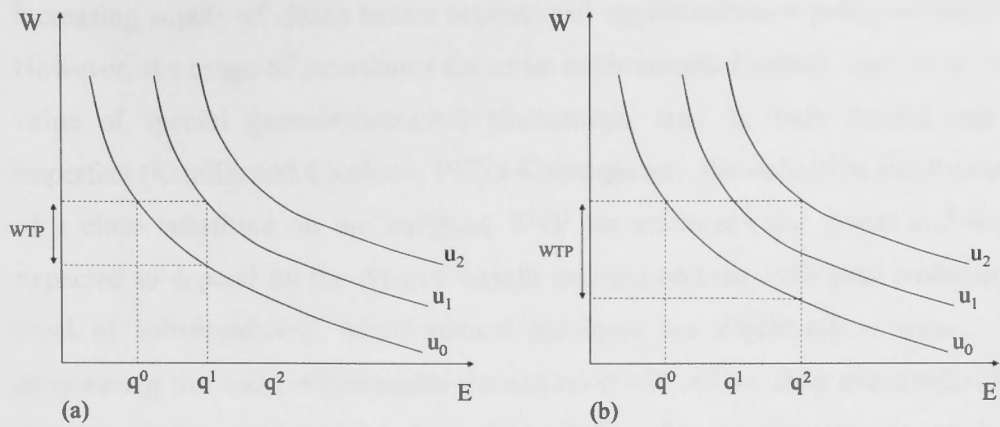


Figure 3: Effect of provision scale on marginal WTP

This discussion demonstrates that value variability across time will be confounded with scale effects if the change in the level of provision used to measure marginal WTP varies across time. Consequently, in order to analyse value variability across time the change in the level of provision needs to be kept constant.

3.3.2 Availability of close substitutes and associated relative scarcity

Other important determinants of marginal WTP suggested by neo-classical microeconomic theory of consumer choice are the availability of close natural or human-made substitutes and the associated relative scarcity of non-marketed environmental goods and services under consideration. Their influence is grounded in the convexity assumption of preference relations. Convexity refers to trade-offs between the different consumption bundles individuals face in choice situations implying that consumption bundles that consist of a mixture of commodities are preferred to extreme ones (for example, Frank, 2003). Convexity thus reflects diminishing marginal rates of substitution along a downward sloping indifference curve¹⁵. A change in supply due to a change in the availability of close natural or human-made substitutes is expected to influence marginal WTP through a change in relative scarcity of the non-marketed environmental goods and services. As illustrated in Figure 4, a change from provision level q^0 to q^1 yields a higher marginal WTP (Figure 4a) than a change from q^1 to q^2 (Figure 4b), with $q^0 < q^1 < q^2$ and changes of equal scale.

¹⁵ As pointed out previously, the downwards sloping property of the indifference curves is also a consequence of the desirability assumption in preference relations (for example, Frank, 2003).

For example, the recreational value of a national park may decrease through time with increasing supply of cheap indoor recreational opportunities or protected natural areas. However, the range of substitutes for some environmental values, such as the aesthetic value of special geomorphological phenomena, may be both limited and mainly imperfect (Krutilla and Cicchetti, 1972). Consequently, the definition and thus influence of a close substitute on the marginal WTP for environmental goods and services is expected to depend on the type of benefit enjoyed and the individual perception of its level of substitutability. Many natural resources are depletable in nature. That is, progressing use and environmental deterioration will reduce their availability. Both the absence of close substitutes and the depletable of many natural resources inhibit an increase in supply (Krutilla and Cicchetti, 1972). Consequently, over time, decreasing availability (for instance, the extent of protected areas of land) will increase their scarcity value, and thus, in turn, may increase their marginal value to consumers (Krutilla and Cicchetti, 1972). For example, the less available irreplaceable pristine natural areas are the more valuable the remaining areas are to consumers.

However, Söderquist et al. (2000) argue in the context of wetlands that this holds only to a certain degree of degradation before the ecosystems begin to be bio-physically dysfunctional and the marginal value drops to zero (marginal value paradox). Following this reasoning, the marginal value of pristine areas increases with increasing scarcity across time up to a point at which the size is too small to have any value for consumers at all. The findings of Krutilla and Cicchetti (1972) are also supposed to be relevant if considering an increasing supply. Rollins and Lyke (1998), for example, found that marginal WTP significantly decreased with an increased number of (hypothetically) existing national parks. That is, policy directed to an ongoing conversion of areas into national parks may decrease their marginal value across time.

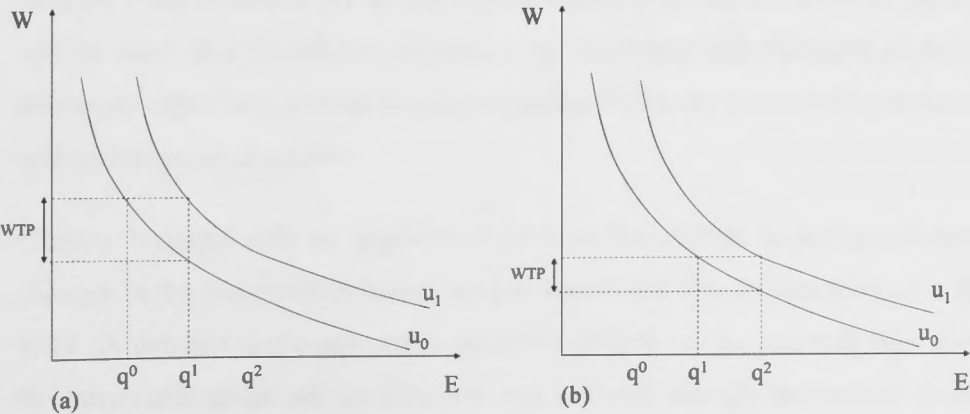


Figure 4: Effect of relative scarcity on marginal WTP

3.3.3 Reference point relativity

Reference dependency is another factor that has been suggested to affect value variability across time (for example, Horowitz, 2002; Tversky and Kahneman, 1991). Since the definition of value as Hicksian surplus is based on a change in utility relative to a baseline (reference point), marginal WTP depends on how this reference point is perceived relative to the alternative level of provision. Following Horowitz (2002), reference dependencies may result in different acceptance levels of environmental degradation at different points in time. Following this rationale, ‘... future decision-makers will “accept” low environmental quality that present decision-makers would want to avoid’ (Horowitz, 2002, p.252). This raises the question about the comparability of reference points used to analyse value variability across time. That is, it raises the question whether the same baseline used at two different points in time provides the same reference point while holding the change in level of provision constant. For example, the perspective of losing 20% of 10,000 hectares of pristine natural area of land might be less (or more) concerning to individuals today than it was 30 years ago. That is, a level of provision that was perceived highly scarce 30 years ago might be perceived as being only slightly scarce today, and vice versa. Since relative scarcity is expected to influence marginal WTP, a change in the perception of relative scarcity might influence marginal WTP across time. However, including these factors in an analysis of temporal value variability requires a rich data set, which may not be collectable or available to the analyst.

3.4 Conclusion

Chapter 3 has discussed the theory behind value variability across time. This discussion will be used as a foundation for analysing the extent and direction of value changes across extended time periods related to non-marketed environmental goods and services with public good character.

Chapter 3 started with an explanation of how fluctuations in preference relations and changes in the competitive budget set influence Hicksian surplus measured as marginal WTP. A detailed discussion about potential drivers on the demand side indicated that the interrelationships are too complex and not well enough understood to differentiate

between the effects of preference fluctuations and changes in the budget constraint. An analysis will be limited to differentiating effects on Hicksian surplus triggered by changes in the underlying structure between marginal WTP and demand side factors; and those driven by a shift in demand side factors. This discussion also emphasised the role of changing levels of information and experience in shaping preferences and preference sets. Hence, including behavioural and attitudinal variables in an analysis of value variability across time is considered important. However, it was concluded that including attitudinal variables must be handled with caution.

Furthermore, Chapter 3 revealed the importance of constant marginal provision scales while analysing temporal value variability to avoid confounding effects. Different marginal provision scales result in different marginal utilities and thus different Hicksian surplus. Also explained was the relationship between changes in the availability of close substitutes and changes in the relative scarcity of environmental goods and services as well as their combined impact on value variability across time. However, it was noted that the magnitude of this effect will depend on the individuals' definition of close substitutes. The chapter concluded with a discussion about the effect of reference point relativity on value change across time and the difficulties in considering these effects when analysing temporal value variability.

In summary, Chapter 3 looked at the theory underpinning temporal value variability assuming that the evolution of stated preference methods used to obtain these values has no impact. This thesis hypothesises that this assumption does not hold. Potential effects of methodological evolution on value estimates are discussed in the next chapter (Chapter 4).

4 ENVIRONMENTAL VALUATION ACROSS TIME

4.1 Introduction

Chapter 3 discussed the theory of potential value change related to environmental goods and services across time, ignoring potential effects of the evolution of stated preference methods used to elicit these values. This chapter aims to discuss potential impacts of temporal variability of stated preference methods on value estimates related to non-marketed environmental goods and services with public good character.

Considering non-marketed values generated by environmental goods and services in CBA has the potential to improve economic efficiency in resource allocation (Chapter 1). The development of stated preference methods has rendered it possible to express non-marketed use and non-use benefits and costs in monetary terms, and thus enabled their inclusion into CBA. Since the CV method was first suggested in theory by Ciriacy-Wantrup (1947) reams of applications have caused an ongoing methodological evolution in stated preference valuation. This methodological evolution expressed itself both in form of refining the CV method and through the development of alternatives including the DCE method. This methodological evolution has been reflected in an ongoing adjustment in elicitation formats. Open-ended elicitation formats used in early CV applications have been largely replaced by dichotomous choice CV elicitation formats and variations thereof. These formats were developed further into repeated multiple choice formats through the emergence of the DCE method.

The stream of experience in applying stated preference methods provides an opportunity to analyse impacts of the evolution of stated preference methods on value estimates. Some effects of, for example, alternative elicitation formats, payment vehicles and survey instruments used in CV and DCE that have evolved over time have already been investigated widely (for example, Alberini and Kahn, 2006; Birol and Koundouri, 2008; Mitchell and Carson, 1989). Within the scope of these efforts, though, two issues remain largely unexplored: potentially stabilising effects of the open-ended CV elicitation format on the temporal variability of value estimates and the impact of strategic response on value estimates due to the incentive compatibility properties of

elicitation formats used in DCE. This chapter concentrates on a discussion of the theoretical foundations underpinning these two issues.

The remainder of Chapter 4 is structured as follows. Sections 4.2 and 4.3 introduce the CV method and DCE method respectively. Methodological evolutions are outlined and econometric models are reviewed. While both sections touch on a wide range of problems associated with these two valuation methods, Section 4.2 focuses on the potentially stabilising effects of the open-ended CV elicitation format on the temporal variability of value estimates and the evolution of econometric modelling of open-ended CV data, whereas Section 4.3 places emphasis on strategic response and incentive compatibility properties of elicitation formats used in DCE.

Both CV and DCE, necessitate the collection of primary choice data to obtain marginal WTP estimates. An alternative to conducting a specific stated preference study is the BT method. BT is used to transfer existing data on non-marketed value related to environmental goods and services across space and time (for example, Bateman et al., 2011; Johnston and Rosenberger, 2010; Morrison and Bennett, 2004). BT limits requirements regarding data collection, and thus lowers costs. However, misleadingly assuming that the transferred value is temporarily stable may lead to a reduction in economic efficiency in resource allocations (see Section 1.2). Even though BT is not applied in the context of this thesis, the implication of potential value change across time on the efficiency of resource allocations based on cost-benefit analysis involving BT data is considered important and discussed in Chapter 11. Finally, Section 4.4 provides a summary and a conclusion.

4.2 Contingent valuation method

4.2.1 Background

CV is a method used to elicit compensating surplus (or equivalent surplus) by asking people how much they are willing to pay to enjoy an increase (or to avoid a decrease) in the provision of environmental goods and services due to a marginal change in a

particular resource allocation¹⁶. The CV method is based on the description of a hypothetical market presented to survey participants in the form of a questionnaire (survey instrument); the survey instrument provides information about the current situation (status quo) and the proposed marginal change in resource allocation, the associated required payment to facilitate the marginal change, and the way the payment would be collected (payment vehicle) (Bateman et al., 2002; Hanley and Barbier, 2009). A more detailed description of a CV questionnaire design is given in Chapters 6 and 8.

The CV method was first suggested in theory by Ciriacy-Wantrup (1947). The first practical application pioneered by Davis (1963) initiated a range of further early CV studies (for example, Brown and Hammack, 1973; Hammack and Brown, 1974; Randall, 1974; Randall et al., 1974). The first CV study applied in Australia was carried out in 1979 estimating the value of a natural area of land at the south coast of New South Wales (Bennett, 1981; 1984). Since then, CV has been widely applied worldwide in various contexts including wildlife conservation (for example, Bandara and Tisdell, 2004), the preservation of tropical rainforests (for example, Carson, 1998), weed control (for example, Champ et al., 2005), recreational fishing (for example, Rolfe and Prayaga, 2007), and climate protection (for example, Solomon and Johnson, 2009).

The accumulating experience since the first practical application in 1963 has resulted in a continuous methodological evolution. Consequently, the published literature has advanced significantly. Hence, from a current perspective, some of the first CV applications are potentially problematic with respect to the validity of their estimates. Part of the methodological evolution of the CV method has been the development of alternative ways to elicit marginal WTP (elicitation formats). Some early CV studies used pen-ended valuation questions asking survey participants directly how much they were willing to pay to enjoy an increase (or to avoid decrease) in the provision of certain environmental goods and services. That is, survey participants were asked to provide a single monetary value to express their maximum marginal WTP without being provided with any cues. Later applications of open-ended CV include Tisdell et al. (2008) and Carlsson and Martinsson (2006a).

¹⁶ Even though the CV method can be used to elicit compensating surplus (or equivalent surplus) by asking people how much they are willing to accept to forgo an increase (or to suffer a decrease) in the provision of environmental goods and services due to a policy change, this thesis focuses on WTP estimates.

However, the open-ended CV elicitation format is theoretically incentive incompatible (Arrow et al., 1993), providing respondents with strategic opportunities for free-riding or overstating their marginal WTP (see Section 4.3.2). That is, the open-ended elicitation format encourages respondents to mis-state their preferences. A theoretically incentive compatible¹⁷ and widely used elicitation format is the single-bounded dichotomous choice. In a single-bounded CV survey respondents are asked whether or not they are willing to pay a certain amount for a proposed marginal change in a particular resource allocation. Examples include Baral et al. (2008), Blumenschein et al. (2008), Brouwer and Bateman (2005), and Farmer and Lipscomb (2008). Further developments in elicitation formats include multiple-bounded dichotomous choice (for example, Aprahamian et al., 2008; Svedsater, 2007; Vossler et al., 2004), payment cards (for example, Brouwer et al., 2008; Frör, 2008; MacMillan et al., 2006), random payment card sorting approaches (for example, Covey et al., 2007), and a combination of open-ended and single-bounded dichotomous choice (for example, Ojeda et al., 2008)¹⁸.

However, only the single-bounded dichotomous choice format is theoretically incentive compatible. Strategic bias introduced by incentive incompatible CV elicitation formats has been explored intensively (for example, Arrow et al., 1993; Carson and Groves, 2007; 2011; Carson et al., 1995; Schläpfer and Braeuer, 2007). Many other effects of alternative CV elicitation formats have been extensively investigated, such as anchoring bias (for example, Blumenschein et al., 2008; Shaikh et al., 2007), hypothetical bias (for example, Carson and Groves, 2011; Cummings et al., 1997; Cummings et al., 1995; Mozumder and Berrens, 2007), ‘yeah saying’ bias (for example, Frör, 2008), information effects (for example, Bergstrom et al., 1989; Hoehn and Randall, 2002; MacMillan et al., 2006; Munro and Hanley, 1999), or uncertainty effects (for example, Aadland et al., 2007; Chang et al., 2007; Shaikh et al., 2007; Sun and van Kooten, 2009). Hanley and Barbier (2009), Mitchell and Carson (1989), and Alberini and Kahn (2006) provide a comprehensive overview.

Still unexplored, however, remain potential impacts of alternative CV elicitation formats on the temporal variability of value estimates, including the open-ended CV

¹⁷ An incentive compatible elicitation format additionally requires the inclusion of a provision rule into the questionnaire, a consequential survey and a take-it-or-leave-it offer with no linkage to other issues (Carson and Groves, 2007).

¹⁸ For a detailed discussion about alternative elicitation formats see, for instance, Alberini and Kahn (2006) and Hanley and Barbier (2009).

elicitation format. Investigating values across extended time periods (see Chapter 3) requires the use of early open-ended CV applications as a baseline. For this reason it is crucial to explore whether potential temporal value variability is masked due to a stabilising effect of the open-ended elicitation method itself. Values may have changed over long time periods. Nonetheless, the open-ended CV method may not disclose this change. In consequence, investigating such potentially stabilising effects is of particular importance when analysing temporal value variability.

The evolution of elicitation formats has been accompanied by a refinement of econometric models suitable for analysing open-ended CV data. Econometrics has progressed significantly since the time when the use of linear and log-linear regression models was state-of-the-art practice. Ergo, from a contemporary perspective, econometric models used in the early days of the CV method may be problematic with respect to their econometric properties. Exploring alternative state-of-the-art econometric models for open-ended CV data may enable a more rigorous comparison of values across extended time periods.

4.2.2 Prominence in response in open-ended contingent valuation

Social psychology suggests that heuristics may influence value estimates obtained by stated preference valuation methods. The theory of prominence originated by Albers (1997) and the model of consumer decisions championed by Weinberg (1981) suggest alternative social psychological processes to explain the role and effect of heuristics in stated preference valuation.

The term ‘prominence’ was first introduced by Schelling (1960) who observed that people’s choices are guided by the prominence of given choice alternatives. This observation led to the theory of prominence (Albers, 1997), which explores how people manage the problem of numerical response. The theory of prominence argues that people use a particular subset of numbers to deal with the difficulty of choices that involve quantitative comparisons. These numbers are called ‘prominent numbers’ and are defined as a sequence determined by the algorithm $a \times 10^i : a \in \{1, 2, 5\}$, with i being an integer number (Albers, 1997; 2001). This algorithm generates the sequence 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, et cetera (Covey and Smith, 2006). ‘These are the integer powers of ten, their doubles and their halves. They have the property that every

number is either approximately or exactly the double of the preceding one. Halving and doubling are the basic mental processes for the comparison of numerical data' (Albers, 2001, p.298). Of course, numerical response restricted to this sequence provides interval values with upper and lower bounds rather than point values. These interval bounds, however, are only known to the individual respondent and are not revealed to the analyst. Nevertheless, the sequence suggested by the theory of prominence can be used to construct these intervals bounds (IB) between two subsequent prominent numbers a_1, a_2 of the sequence $a \times 10^i : a \in \{1, 2, 5\}$, with $IB = (a_1 + a_2) / 2$. However, the distribution of the numbers within these intervals remains unidentified. Put simply, the use of prominent numbers implies a reduction in response exactness with the extent of this inexactness remaining unknown.

Few studies have explored potential impacts of prominent numbers on stated preference valuation. Covey and Smith (2006) and Whynes et al. (2007; 2005) found evidence of prominence in response data obtained by the CV method using both open-ended and payment card elicitation formats. They argue that respondents' use of heuristics (prominent numbers) to express marginal WTP is a way to reduce transaction costs if numerical choices are intrinsically difficult or if the consequences of inexact responses are rendered irrelevant. However, potential effects of such heuristics on temporal variability of open-ended CV value estimates remain unknown. Prominent numbers that are stable across time may have a stabilising effect on variable values. A value change that is too small to provoke a change from one prominent number to another may not be picked up if marginal WTP is expressed as an interval rather than a point value. While values may have changed over long time periods, if respondents use prominent numbers that are stable across time, the open-ended CV method may not disclose this value change. In other words, the results of an open-ended CV method may suggest stable preferences whereas, in fact, they reflect a stabilising effect of prominent numbers.

One might argue that, from a statistical perspective, prominent numbers are exogenous. If that is the case and a continuous shift in the underlying WTP distribution is assumed then the theory of prominent numbers should have no more impact on measuring a random variable at a finite number of locations and power to discern a shift than any other common measurement situations. However, even though the prominent numbers might be stable over time, it is likely that their underlying value might have changed. If this is the case, prominent numbers may cause confounding effects as described above.

The effect of prominent numbers on value estimates assumes cognitive controlled behaviour of survey participants. However, value estimates obtained by stated preference methods, including the open-ended CV method, have frequently been criticised as being based, at least partially, upon non-rational motivations such as habits and impulsive emotions, raising doubt about their validity as measures of economic benefits of public goods (Kahneman et al., 1993). Nevertheless, habits and impulsive emotions may impact on marginal WTP, irrespective of their recognition by neo-classical microeconomic theory of consumer choice. Consequently, habits and impulsive emotions may have an influence on value variability through time¹⁹. Fischer and Hanley (2007) apply the model of consumer decisions (Kroeber-Riel and Weinberg, 1996; Weinberg, 1981) to a CV study to differentiate between economically meaningful marginal WTP estimates and those which do not reveal information about respondents' preferences. They argue that the model of consumer decisions hypothesises four types of decision making (extensive, limited, impulsive, and habitualised), motivated by three underlying psychological processes (cognitive control, emotional involvement, and reactive-spontaneous behavioural responses). Limited and extensive decisions are assumed mainly to be driven by cognitive control, that is, respondents express their preferences toward the good under consideration and gather information as far as necessary. They differ by the degree of emotional involvement and the need for new information. Fisher and Hanley (2007) conclude that both types are assumed to be consistent with the assumptions of neo-classical consumer theory; in contrast, motivations such as habits and impulsive emotions do not provide information about preferences underlying a choice and are thus assumed to be inconsistent with neo-classical consumer theory. According to the model of consumer decisions, impulsive and habitualised decisions are not driven by cognitive control (Kroeber-Riel and Weinberg, 1996; Weinberg, 1981). Kroeber-Riel and Weinberg (1996) and Weinberg (1981) postulate that impulsive decisions result from emotional involvement and reactive-spontaneous behavioural responses. In other words, they represent emotional spontaneous reactions to a stimulus such as a cue (for example, a certain policy or payment vehicle) (Fischer and Hanley, 2007). Habitualised decisions are, however, hypothesised to be motivated by reactive-spontaneous behavioural responses without emotions involved (Kroeber-Riel and Weinberg, 1996; Weinberg, 1981). In this way, responses reflect a routine choice (Fischer and Hanley, 2007).

¹⁹ One way of testing this hypothesis is the performance of a scope test.

Consequently, temporal value variability may either be driven by cognitive controlled behaviour or governed by impulsive-emotional decisions in reaction to a cue. The latter would not provide information about how and why values change over time. On the other hand, relatively invariant consumer behaviour over time may reflect either stable preferences (assuming cognitive controlled behaviour) or the presence of habits (assuming reactive-spontaneous behaviour) (Hodgson, 1988). This line of reasoning is vividly summarised by Stevens et al. (1994, p.359) who question if ‘the value of eagles or the habit of contribution’ is stable over time. Thus, unchanging value estimates over time may be mistaken as an indication of stable preferences while instead expressing habits independent of the good under consideration.

4.2.3 Econometric modelling of open-ended contingent valuation data

The evolution of stated preference methods triggered a continuous development and refinement of econometric models that are, in theory, suitable to analyse open-ended CV data²⁰. In this way, econometrics has made significant progress in modelling data obtained by open-ended CV. Of course, the open-ended elicitation format facilitates the calculation of marginal WTP means, medians, and confidence intervals of stated marginal WTP without the need for any model estimations. However, econometric models are being used to analyse explanatory variables of, and their impact on, stated marginal WTP. Early open-ended CV studies used linear and log-linear regression models applying ordinary least square (OLS) estimators to investigate drivers of marginal WTP. However, these types of econometric models are considered inadequate for the data characteristics of open-ended CV studies.

Commonly, open-ended CV data are characterised by a distribution over strictly positive values and a cluster of zero values; that is, a significant number of respondents maximise their utility by choosing the corner solution (Wooldridge, 2006). Put simply, a corner solution response describes an outcome where respondents state that they are not willing to pay any amount for a proposed change in a particular resource allocation²¹.

²⁰ Even though this thesis focuses on the evolution of econometric models with respect to open-ended CV, progress in econometric modelling is not restricted to this particular elicitation format. A detailed discussion of state-of-the-art econometric models suitable for dichotomous elicitation formats, payment cards and variations thereof is given in, for example, Wooldridge (2010).

²¹ The fraction of zero values may include censored WTP bids: The open-ended CV elicitation format does not allow for negative WTP bids. Respondents who hold a negative value may choose to provide a zero WTP bid as a surrogate for their true negative WTP.

The econometric literature suggests that using OLS linear regression models to analyse this type of data might be inappropriate for a range of reasons.

First of all, a limited dependent variable that is represented by a significant proportion of zero values cannot be normally distributed (Wooldridge, 2006). Hence, hypothesis testing and confidence intervals based on open-ended CV data using OLS linear regression models is only asymptotically justified by the Central Limit Theorem (Hill et al., 2008). Additionally, applying an OLS linear regression model to data that contains a substantial fraction of corner solution response may lead to negative predictions of the dependent variable (Wooldridge, 2010). Another concern strongly supported by econometric theory is the supposition that applying an OLS linear regression model to data that contains a significant number of corner solution responses will lead to inconsistent estimators of the regression parameters; limiting the analysis to positive observations only will not avoid this problem (Greene, 2008; Wooldridge, 2010). This issue is best explained by reviewing corner solution models²² (for example, Greene, 2008; Long, 1997; Wooldridge, 2010).

Corner solution models assume that the dependent variable y is approximately continuously distributed over strictly positive values but takes on the value of zero for a significant proportion of the population²³. For a randomly drawn observation i from the population the model can be written as

$$y_i^* = x_i' \beta + \varepsilon_i \text{ with } \varepsilon_i | x_i \sim N(0, \sigma^2), \quad [10]$$

$$y_i = \max(0, y_i^*), \quad [11]$$

where σ is the standard deviation, y_i is a dependent variable, x_i is a vector of explanatory variables, β is a vector of parameter estimates, and ε_i is a random error term.

Based on these assumptions the expected value can be derived as

²² The corner solution model, also called Tobit model, was first suggested by Tobin (1958) who analysed households purchases for durable goods. The corner solution model is frequently discussed as a special case of a censored regression model assuming a zero limit. That might be the reason why it is also called a censored regression model. However, as outlined by Wooldridge (2010) this is misleading. A corner solution model represents utility maximisation. That is, the non-trivial fraction of zero values is not a result of censoring but of perfectly observing corner solutions. A censored regression model, on the other hand, is characterised by a censored variable. That is, the exact value of data that falls below or above a certain limit is not observed and assumed to be the limit value.

²³ The discussion of corner solution model follows, if not stated otherwise Wooldridge (2010).

$$\begin{aligned}
E(y_i | x_i) &= \Pr(y_i = 0 | x_i) \cdot E(y_i | x_i, y_i = 0) + \Pr(y_i > 0 | x_i) \cdot E(y_i | x_i, y_i > 0) \\
&= \Pr(y > 0 | x) \cdot E(y | x, y > 0).
\end{aligned} \tag{12}$$

Using a probit model and defining a binary variable $z = 1$ if $y > 0$, $z = 0$ if $y = 0$, $\Pr(y_i > 0 | x_i)$ can be derived as follows:

$$\begin{aligned}
\Pr(z = 1 | x_i) &= \Pr(y_i^* > 0 | x_i) = \Pr(\varepsilon_i > -x_i' \beta | x_i) \\
&= \Pr(\varepsilon_i / \sigma > -x_i' \beta / \sigma) = \Phi(x_i' \beta / \sigma),
\end{aligned} \tag{13}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

The expected value for $y_i > 0$ given x_i can be found by

$$\begin{aligned}
E(y_i | x_i, y_i > 0) &= x_i' \beta + E(\varepsilon_i | \varepsilon_i > -x_i' \beta) = x_i' \beta + \sigma \left[\frac{\phi(x_i' \beta / \sigma)}{\Phi(x_i' \beta / \sigma)} \right] \\
&= x_i' \beta + \sigma \lambda(x_i' \beta / \sigma),
\end{aligned} \tag{14}$$

where $\phi(\cdot)$ is the standard normal probability distribution function.

$$\text{and } \lambda(x_i' \beta / \sigma) = \left[\frac{\phi(x_i' \beta / \sigma)}{\Phi(x_i' \beta / \sigma)} \right].$$

Hence, the expected value of y_i given x_i can be written as

$$E(y_i | x_i) = \Phi(x_i' \beta / \sigma) [x_i' \beta + \sigma \lambda(x_i' \beta / \sigma)]. \tag{15}$$

Equations [14] and [15] demonstrate why applying an OLS linear regression model leads to inconsistent estimators irrespective of whether the zero observations are excluded or not. Equation [14] shows that excluding zero observations results in the omission of a variable, $\sigma \lambda$, and equation [15] clearly illustrates that $E(y_i | x_i)$ is non-linear in x_i , β , and σ .

Instead of applying an OLS estimator, corner solution models use maximum likelihood estimation. Given that the probability distribution function for $y_i | x_i$ can be written as

$$f(y_i | x_i) = \mathbb{I}[y_i = 0][1 - \Phi(x_i' \beta / \sigma)] + \mathbb{I}[y_i > 0]\{\phi[(y_i - x_i' \beta) / \sigma] / \sigma\}, \tag{16}$$

with the probability distribution function of $y_i < 0$ being zero and $\theta \equiv (\beta', \sigma^2)'$ denoting the parameter vector $(K + 1) \times 1$,

the conditional log likelihood can be written as

$$l_i(\theta) = 1[y_i = 0] \log[1 - \Phi(x_i\beta / \sigma)] + 1[y_i > 0] \{ \log \phi[(y_i - x_i\beta) / \sigma] - \log(\sigma^2) / 2 \}. \quad [17]$$

Corner solution models are used to accommodate the specific characteristics of open-ended CV data (for example, Carlsson and Martinsson, 2007). However, it is important to note that corner solution models are based on the assumption of normally distributed error terms [10]. Violating this assumption may result in inconsistent estimators (Wooldridge, 2010). A range of tests are available to check for non-normality, including a conditional moment test suggested by Pagan and Vella (1989) and Skeels and Vella (1999). Corner solution models that allow the assumption of alternative distributions such as the exponential, Weibull, and lognormal have been developed (Kalbfleisch and Prentice, 2002). However, as pointed out by Greene (2008), assuming alternative distributions does not necessarily guarantee a solution to the problem. Hence, more research is required to explore possible effects of these model specifications, which is beyond the scope of this thesis. An alternative approach to deal with corner solution data that exhibits non-normally distributed error terms is Powell's (1984) least absolute deviations (LAD) estimator:

$$y_i^* = x_i' \beta + \varepsilon_i, \quad \text{with } Med(\varepsilon_i | x_i = 0), \quad [18]$$

$$y_i = \text{Max}(0, y_i^*). \quad [19]$$

The LAD model is a quantile regression model, which estimates the median of the dependent variable conditional on a set of explanatory variables by minimizing the sum of the absolute residuals rather than the sum of the squares of the residuals as in OLS regressions (StataCorp, 2007). Since this estimator does not produce estimates for σ , a bootstrapping method is available to estimate the asymptotic covariance matrix (Greene, 2002).

Apart from the potential violation of the normality assumption, Greene (2008) argues that the use of corner solution models might be inappropriate for many applications since this model not only predicts a pile of zeros but values near the corner solution outcome. Hence, fitting this model to data that is characterised by a discontinuous distribution, in other words by a gap between zero and positive observations, might lead to a significant underestimation of parameter estimates. Other econometricians, however, have not voiced this concern (for example, Long, 1997; Wooldridge, 2010).

Fin and Schmidt (1984) note that, in a corner solution model, the impact of an explanatory variable on the probability of an observation being a corner solution response has the same sign as its impacts on the mean of the positive observations. This is supported by Greene (2008) who argues that corner solution response and positive marginal WTP bids might be underpinned by two different decision processes: the decision to participate at all, and, if that is the case, the strength of this participation.

Ignoring information on positive marginal WTP bid levels, binary choice models can be used to analyse these two decision processes in open-ended CV data. A binary choice model (for example, Greene, 2008) can be written as the index function model²⁴

$$y_i = x_i' \beta + \varepsilon_i, \quad [20]$$

$$y_i = 1 \text{ if } y_i^* > 0,$$

$$y_i = 0 \text{ if } y_i^* \leq 0$$

with ε_i having a zero mean and being either normally distributed (probit model),

$$\Pr(y_i = 1 | x_i) = \Phi(x_i' \beta), \quad [21]$$

or logistically distributed (logit model),

$$\Pr(y_i = 1 | x_i) = \Lambda(x_i' \beta). \quad [22]$$

The probability distribution function of $y_i | x_i$ can be written as

$$f(y_i | x_i) = [F(x_i' \beta)]^{y_i} [1 - F(x_i' \beta)]^{1-y_i}, \quad y_i = 0, 1. \quad [23]$$

Based on the probability function, the log likelihood function can be written as

$$l_i(\beta) = y_i \log[F(x_i' \beta)] + (1 - y_i) \log[1 - F(x_i' \beta)]. \quad [24]$$

Alternatives that might be used to address the concerns of Greene (2008) and Fin and Schmidt (1984) without ignoring information on the positive bids are double hurdle or two-part models (for example, Wooldridge, 2010). A double hurdle model distinguishes between the decision of whether $y = 0$ or $y > 0$ and the decision about the value of y if $y > 0$ ²⁵. Cragg (1971) proposed a model that combines a probit model with a truncated regression model. As with corner solution models, truncated regression

²⁴ The discussion of the binary choice model follows Greene (2008).

²⁵ The discussion of the hurdle model follows Wooldridge (2010).

requires normally distributed error terms. The probability distribution function of $f(y_i | x_i, y_i > 0)$ and $f(y_i | x_i)$ can be expressed as

$$f(y_i | x_i, y_i > 0) = [\Phi(x_i' \beta / \sigma)]^{-1} \{\phi[(y_i - x_i' \beta) / \sigma] / \sigma\}, \quad y_i > 0, \quad [25]$$

$$f(y_i | x_i) = 1[y_i = 0][1 - \Phi(x_i' \omega)] + 1[y_i > 0]\Phi(x_i' \beta / \sigma)^{-1} \{\phi[(y_i - x_i' \beta) / \sigma] / \sigma\}^{24}. \quad [26]$$

If $\omega = \beta / \sigma$, Cragg's double hurdle model becomes the corner solution model (see above). Fin and Schmidt (1984) suggested a test based on the Lagrange multiplier statistic to check whether a corner solution model or Cragg's double hurdle model are more appropriate for a particular data set. As with the corner solution model the conditional mean is non-linear in x_i , β , and σ . Thus, the violation of the normality assumption will lead to inconsistent estimators (Wooldridge, 2010).

As previously outlined, corner solution models and truncated regression models assume that positive values are approximately continuously distributed. However, the open-ended CV method frequently produces discrete values. Therefore, neither corner solution models nor hurdle models may be completely appropriate for this type of data. Models well suited for discrete data are count models and negative binominal regression models (for example, Long, 1997; Wooldridge, 2010). These models are not only designed for handling discrete data. Being based on a Poisson or a negative binominal distribution, respectively, they also fit skewed data distributions. The drawback of these models is that data characterised by a non-monotonic distribution produce an unreasonable model fit. Therefore, count models and negative binominal regression models are inadequate for open-ended CV data that are characterised by a non-monotonic marginal WTP distribution.

Following the theory of prominent numbers (see Section 4.2.2) discrete values produced by the open-ended CV method may represent intervals rather than point values. Censored interval models, which are a more general form of censored regression models (for example, Long and Freese, 2006), would be a promising alternative if the boundaries of the intervals were known to the analyst. However, they are only known to the respondent. Of course, assumptions about those boundaries could be based on the theory of prominent numbers. However, if this theory does not apply to the data the results will be biased. Whether it does apply or not, however, is unknown to the analyst.

Protest votes are another challenge frequently associated with stated preference methods, and thus with the open-ended CV method. Protest can manifest itself by non-response, extremely high marginal WTP bids, and zero marginal WTP bids (for example, Haddad and Howarth, 2006). With respect to the latter, the key issue is whether the respondent would vote 'no'/ contribute zero given the conditions of the contingent market. In consequence, the fraction of zero marginal WTP bids may be a result of two alternative underlying processes. One part of this fraction consists of valid zero marginal WTP bids. Respondents do have a zero value for a marginal change in a particular resource allocation or are restricted by a budget constraint (see Section 2.2.3). A second part of this fraction may be composed of protest zero marginal WTP bids: respondents may hold a positive or negative value but refuse to disclose it. In other words, respondents choose not to reveal their true underlying preferences. Protest marginal WTP bids frequently account for a significant proportion of total marginal WTP bids. As reviewed by Haddad and Howarth (2006), the proportion of protest votes found in CV studies ranges from five per cent up to 50 per cent of the total votes. The motivations for protest votes are manifold. Jorgensen and Syme (2000) provide a list of possible reasons for protest marginal WTP bids in CV data involving public goods. The reasons include concerns about fairness, aversion against certain payment vehicles such as taxes, rejection of individual responsibility and mistrust of the payment collecting agency.

The management of protest votes (and in particular protest zero marginal WTP bids) is still subject to debate. One approach is to declare them outliers and to exclude them from the analysis. However, as pointed out by Boyle and Bergstrom (1999) and supported by findings of Meyerhoff and Liebe (2006), such an exclusion may not be justified. From an econometrics perspective, discarding zero marginal WTP bids might be problematic. Following the econometric literature (for example, Wooldridge, 2010) protest zero marginal WTP bids can be regarded as a result of sample selection due to non-response, with non-response being defined as those marginal WTP bids, which do not reveal true underlying preferences. In general, Wooldridge (2010) defines a selected sample as a non-random sample with one possible selection mechanism being non-response. He points out that, even though respondents are randomly drawn from the population, data are missing for one or more variables for some respondents; excluding those respondents from the sample and thus restricting the analysis on the remaining sub-sample may lead to a sample selection problem.

Wooldridge (2010) argues that if the reasons for non-response are endogenous such an approach may have significant effects on estimating population parameters. To illustrate his point be provided the following example. If voters are those respondents whose true values are high compared to the average value, excluding them would underestimate the true marginal WTP. If, on the other hand, the protest sub-sample is a random sample of the non-protest sample (that is, the distribution of true values is the same for the protesters' sub-sample and non-protesters sub-sample), excluding protest bids would not lead to inconsistent estimators. Whether this is the case or not, however, is only known to the respondent and not observable by the analyst.

A range of sample selection models have been developed including the Heckit model (Heckman, 1976) and the type II Tobit model (Amemiya, 1985). As emphasised by Wooldridge (2010), the type II Tobit model is a sample selection model and must not be confused with a model capable of addressing corner solution response. In general, sample selection models cannot accommodate corner solution response since all zero observations, true zero marginal WTP bids and protest zero marginal WTP bids would be treated as missing data. Yet, treating true zero marginal WTP bids as protest zero marginal WTP bids would be inappropriate.

The discussion about the evolution of econometric models for open-ended CV data can be summarised as follows. A binary choice model such as the probit or the logit model gives the opportunity to explore variables explaining respondents' decisions of whether to pay for a marginal change in a particular resource allocation. Furthermore, corner solution models and hurdle models offer the possibility for analysing explanatory variables of marginal WTP while accounting for a non-trivial fraction of zero observations. However, corner solution models rely heavily on the normality assumption, which might be violated by open-ended CV data. Additionally, corner solution models assume approximately continuously distributed positive values. The open-ended CV method, though, frequently produces discrete data. Count models and negative binominal models suited to discrete data assume, however, a monotonic value distribution. Applying such models to non-monotonic value distributions would result in poor model fits. Moreover, as suggested by the theory of prominent numbers, the open-ended CV method may produce not only discrete but at least partial interval data. Since the upper and lower interval bounds of each respondent are unknown to the analyst censored interval models might not be a suitable alternative. An alternative approach to

deal with corner solution data that exhibits non-normally distributed error terms are quantile regression models such as Powell's LAD estimator. However, neither quantile regression models nor corner solution models are yet capable of differentiating between corner solutions and protest zero marginal WTP bids.

Hence, even though econometric modelling has made substantial progress, open-ended CV data may still pose a notable challenge to the analyst.

4.3 Discrete choice experiment method

4.3.1 Background

The DCE method can be seen as setting up a contingent market in which a good is set up, part of this as the mode of survey administration, part as the actual design of the DCE, and part as a data collection effort designed to help explain the DCE responses. Its potential to elicit implicit prices for individual attributes that jointly describe a particular good or service is one advantage of the DCE method over other methods such as CV (Bennett and Blamey, 2001). The DCE method, which originated in marketing and transport economics (for example, Louviere and Hensher, 1982; Louviere and Woodworth, 1983), was first applied in environmental economics by Carson et al. (Carson et al., 1990). It is now a widely used alternative to the CV method to estimate non-marketed values as inputs in CBA to improve efficiency in resource allocation. Birol and Koundouri (2008) detail some recent European examples while Bennett and Birol (2010) provide contemporary developing country case studies.

The DCE method is based on Lancaster's 'characteristics value theory', which he summarised in the following statements:

The good, per se, does not give utility to the consumer; it possesses characteristics, and these characteristics give rise to utility. ... In general, a good will possess more than one characteristic, and many characteristics will be shared by more than one good. ... Goods in combination may possess characteristics different from those pertaining to the goods separately (Lancaster, 1966, p.134).

Following Lancaster, the DCE method (for example, Bennett and Blamey, 2001) asks respondents to make trade-offs between characteristics that describe non-marketed goods and services. One of these characteristics, or attributes, is commonly represented in monetary terms as a cost or benefit attribute. All attributes take on different levels and are bundled in choice options, which are offered to respondents in choice sets. One of the choice options usually represents the ‘status quo’, involving no change in a particular resource allocation. Hence, this option is characterised by constant attribute levels. The alternative choice options that do involve marginal changes in a particular resource allocation are described by differing attribute levels. Choice sets are thus distinguished by a constant ‘status quo’ and differing alternative choice options.

The process of designing a DCE includes a sequence of steps: the choice of a sampling procedure; the preparation of a sampling protocol; the decision about the sampling mode (such as internet based survey, telephone survey, and mail survey); the definition of the valuation question (establishing the status quo/ zero cost choice option and alternative choice options); the definition of the attributes and their levels; the choice of an experimental design (such as orthogonal design or efficient design); the choice of a payment vehicle; the design of the questionnaire (description of the scenario, choice sets; formulating questions about respondents’ socio-demographics, attitudes and behaviour; the formulation of follow-up questions) (for example, Bennett and Blamey, 2001; Birol and Koundouri, 2008). A detailed discussion of the components of the DCE design process and the challenges they pose to the analyst are discussed in detail in Chapter 8.

The DEC method provides information about whether the attributes used to describe a good or service are significant determinants of respondents’ preferences (Hanley and Barbier, 2009). It also facilitates the estimation of monetary values of changes in the provision of a particular attribute (implicit prices), and thus allows the estimation of WTP²⁶ (compensating surplus or equivalent surplus) for a marginal change in a particular resource allocation due to a marginal change in the level of more than one attribute at once (for example, Bennett and Blamey, 2001; Hanley and Barbier, 2009). A detailed explanation of how implicit prices, compensating surplus and equivalent surplus can be derived is given in Section 4.3.3.

²⁶ Even though the DCE method can be used to elicit compensating surplus (or equivalent surplus) by asking people how much they are willing to accept to forgo an increase (or to suffer a decrease) in the provision of environmental goods and services due to a policy change, this thesis focuses on WTP estimates.

Typically, environmental policy involves choices about marginal changes in quality and quantity of attributes describing environmental goods and services rather than a decision to gain or lose these goods and services completely (Hanley and Barbier, 2009). Birol and Koundouri (2008) suggest that another advantage of the DCE method over the CV method is its ability to avoid the ‘yeah saying’ bias found in dichotomous CV as well as strategic bias in open-ended CV due to the opportunity to over- or understate marginal WTP. However, these challenges may only have taken on a different form. ‘Yeah saying’ might have been replaced by respondents always choosing the status quo choice option. Strategic response might have shifted from over- or understating of marginal WTP towards taking advantage of strategic opportunities provided by a sequential elicitation format (see Section 4.3.2). Many challenges associated with the DCE method have been explored intensively, including hypothetical bias (Bennett and Blamey, 2001; Carlsson and Martinsson, 2001; List et al., 2006); learning effects (Savage and Waldman, 2008); design effects (Carlsson and Martinsson, 2006b; Caussade et al., 2005; DeShazo and Fermo, 2002); and scale and scope effects (Rolfe et al., 2011). Yet, incentive compatibility properties and strategic response in DCE have been given less attention.

4.3.2 Incentive compatibility properties of discrete choice experiments²⁷

As outlined previously, DCE involves respondents making trade-offs between attributes that describe non-market goods and services. A variety of attribute levels are bundled in choice options and offered to respondents in choice sets. Choice sets are thus distinguished by differing choice options. The number of choice options and choice sets varies widely across studies. Each respondent is commonly presented with a sequence of choice tasks containing more than two choice options rather than restricting the elicitation format to a single referendum – a single-binary choice set²⁸. In contrast to a single referendum, a single-multiple elicitation format requires respondents to make one

²⁷ Parts of the literature review presented in Section 4.3.2, parts of the research design presented in Section 7.3.3, parts of the data collection presented in Section 8.3, and parts of the findings presented in Section 10.4 were reported in one article that has been published in *Environmental and Resource Economics* (Scheufele and Bennett, 2012b) and another article that has been revised and re-submitted to *The Australian Journal of Agricultural Economics* (Scheufele and Bennett, 2012a). The candidate conducted all analysis and took the primary role in writing the article, with Bennett providing input at the research design stage and contributions during the writing of the article.

²⁸ It is common practice to present respondents with a choice between a zero cost choice option (often the status quo) and one or more choice options with positive cost where the goods and services are assumed to be positively valued. This thesis does not explore cases where choice options are associated with disutility or where none is the status quo.

choice between more than two choice options presented in one single choice set, whereas a repeated-binary elicitation format asks respondents to make repeated trade-offs between two choice options. A repeated-multiple elicitation format, finally, combines the two offering respondents repeated choices between more than two choice options presented in a sequence of choice sets. Of course, using an elicitation format other than a single referendum increases the statistical efficiency of the data for a given number of respondents. However, both mechanism design theory (Hurwicz, 1960) and empirical evidence (Bateman et al., 2008; McNair et al., 2011) suggest that repeated choice tasks per respondent may induce strategic behaviour. Respondents may mis-state their true preferences to gain a strategic advantage.

The efficiency of decisions concerning resource allocation depends on individuals truthfully disclosing their privately known preferences. However, revealing true preferences in a DCE might not be an individual's optimal strategy for a given social choice function (for example, Mas-Colell et al., 1995). Samuelson (1954) concluded that there exists no mechanism that can guarantee an efficient level of public goods since individuals have a strong incentive to conceal their true preferences. Despite Samuelson's findings, economists have continued to seek incentive compatible demand revealing mechanisms. The analysis of demand revealing mechanisms is the province of mechanism design theory, originally introduced by Hurwicz (1960). Mechanism design theory compares equilibrium outcomes of alternative mechanisms in non-cooperative games of incomplete information with self-interested participants. Hurwicz (1972) defined a mechanism as a communication system used by utility-maximizing participants to reveal private information, such as true or simulated preferences, where the aggregated private information assigns the outcome. He further argues that the social choice function defines the aggregation process. Accordingly, the social choice function, called the provision rule in DCE, is the link between respondents' choices and the corresponding policy outcome. Whether revealing true preferences is a dominant strategy thus depends on both, the mechanism and the expectations of respondents about the provision rule used to aggregate their choices (Gibbard, 1973; Moulin, 1994; Satterthwaite, 1975).

A widely cited example of an incentive compatible mechanism is a binding referendum between two candidates in an election. Carson and Groves (2007) provided evidence to suggest that replacing the binding character of the referendum by an advisory

referendum does not change the incentive compatibility properties of the mechanism. Green and Laffont (1978) showed that this also holds for a sample rather than population based referenda. This is important since the majority of choice experiments use statistical samples and, when dealing with public goods, frequently simulates an advisory referendum.

The theorems of Gibbard (1973) and Satterthwaite (1975) provide a theoretical foundation to analyse the incentive compatibility properties of mechanisms used in stated preference methods such as DCE. The theorems show that all non-dictatorial mechanisms other than the single-binary choice format are generically incentive incompatible. Consequently, any departure from the single-binary format thus changes the incentive properties of DCE (Carson and Groves, 2007). The theorems of Gibbard and Satterthwaite also hold for Nash implementations if provision rules are required to be singleton-valued (Maskin, 1977; Muller and Satterthwaite, 1985). In other words, a non-singleton provision rule may result in potentially incentive compatibility. Many policy decisions that are concerned with the provision of environmental goods and services, however, are confronted with mutually exclusive policy scenarios, that is, the choice of a single scenario is required. Therefore, using a mechanism with a Nash implementation is not a feasible alternative. Carson and Groves (2007) pointed out that in the case of private and quasi-public goods the provision of more than one good may be possible, that is, the provision rule is not singleton-valued. Hence, for private and quasi-public goods, this provides the possibility of an incentive compatible Nash implementation, that is, respondents' incentives to untruthfully reveal their preferences may be reduced. For pure public goods, however, this possibility does not exist.

By introducing a provision rule that is based on a randomly drawn choice question to be binding, a repeated-multiple choice format can be disaggregated into a sequence of independent single-binary choices. Such a provision rule increases the probability that respondents disclose their true preferences. Laboratory choice experiments have employed a provision rule that is based on a randomly drawn choice question to be binding (for example, Collins and Vossler, 2009). Policy decisions about public goods valued in field studies that are based on random draws, however, raise credibility concerns (Carson and Groves, 2007; 2011). Hence, the results of laboratory experiments that apply a random provision rule are inapplicable to explain strategic behaviour in the context of public goods. Another issue with randomly drawn choice questions to be

binding is that expected utility is generally required for consistent preference revelation (Holt, 1986).

Incentive properties of DCE are mainly changed in two ways. One effect of repeated choice tasks is the introduction of conditional preferences. The literature on incentive compatibility proposes that respondents may adjust their preferences on expectations about the choices of other survey participants (Carson and Groves, 2007; 2011). Carson and Groves (2007; 2011) argue, the dominant strategy for some respondents may be to choose a less preferred option across choice questions if they expect that their most preferred option has no chance of being selected for policy implementation. This is also true for a single-multiple choice format. In that case, a single-multiple choice format collapses to a single-binary choice between the two choice options that the respondent perceives to be other respondents' most preferred choice option if a plurality vote provision rule is applied. However, a single-multiple elicitation format may be potentially incentive compatible in this sense if respondents have uniform priors about other respondents' preferred choices (Moulin, 1994).

Another consequence of the sequential choice format is dependency across multiple choice questions given to each respondent. Respondents may exploit strategic opportunities by involving information about previous choice sets and choice decisions (Carson and Groves, 2007; 2011). As a result, Carson and Groves (2007; 2011) argue that a dominant strategy for some respondents may be to choose a less preferred option in one or more binary choice questions in an attempt to ensure a policy outcome that is more favourable to them.

Despite these problematic incentive properties, DCE using repeated-multiple choice formats have been frequently conducted to inform policy decisions. Results that are thus potentially biased are accepted for the increase in statistical efficiency of the data for a given number of survey participants²⁹. The extent of this bias remains largely unknown. Few empirical studies have investigated dependencies across multiple choice questions, whereas the effect of dependencies across respondents is still unexplored.

²⁹ Rose et al. (2009) used simulated data to investigate the statistical impact of panel data in discrete choice experiments. Their study provided evidence that for a given sample size increasing the number of observations per respondents yields less biased estimates and larger t-ratios. However, this advantage diminished with increasing sample size.

Racevskis and Lupi (2008) used a split sample design to explore the effect of a single-binary versus a repeated-binary elicitation format. They found a significant difference between fits across the two models based on pooling the data of the two response formats in two different ways: the first model included generic attributes whereas the second model included split sample specific attributes. This study fell short of accounting for effects of differing sample sizes implied by each choice format and was focused on a comparison of the model fit between the two split samples. It missed the opportunity to explore impacts on further DCE dimensions. Holmes and Boyle (2005) considered a sequence of four binary choice questions, one of which was the status quo. They found a structural influence of previous and successive choice questions on current choices. Their findings suggest that respondents may compare the levels of cost and non-cost attribute across choice questions and thus learn about their preferences as they progress through the choice task.

Bateman et al. (2008) added a further dimension to the discussion about incentive properties of elicitation formats. They found evidence that repeated choice dynamically increases awareness of strategic opportunities as respondents progress through the choice tasks. This increased awareness might occur through either or both the information provided to the respondent before the choice task (choice set awareness effect)³⁰ and a dynamic increase in awareness of strategic opportunities as progress is made through the sequence of choice questions (ordering effect). Bateman et al. (2008) argue that strategic opportunities provide incentives to mis-state true preferences. Previous and successive choice sets may contain alternative costs for the same or a similar level of provision of a particular good or, vice versa, the same or similar cost for alternative levels of provision of a particular good. According to Bateman et al. (2008), this may cause respondents either to question the credibility of the survey or to learn to take advantage of such perceived inconsistencies by rejecting a preferred choice option when the same or a similar level of provision was offered in a previous or successive choice question at a lower cost. Bateman et al. (2008) conclude that repeated choice may cause learning about strategic opportunities and how to exploit them.

Bateman et al. (2008) thus expand the well-established notions of learning, which have been described as institutional learning and value learning. Institutional learning

³⁰ Mail surveys disclose all choice questions before a choice has to be made whereas an internet based survey can be programmed to reveal only one choice question at a time.

describes a process where respondents become increasingly familiar with the choice context, the offered good, and the choice task (Braga and Starmer, 2005). This process of learning results in more accurate choices reflected in the scale factor³¹ rather than in a preference change. Braga and Starmer (2005) argue that as respondents progress through the choice questions, their responses become more accurate (increase in scale factor) until fatigue sets in (decrease in scale factor). In this context, Swait and Adamowicz (2001) discuss a ‘smaller noise to signal ratio’ and ‘larger noise to signal ratio’, respectively. Value learning suggests that respondents ‘discover’ their true underlying preferences through a learning process rather than possessing stable preferences (Plott, 1996). Such learning processes are expected to change preferences, and thus parameter estimates in DCE. In contrast, strategic learning as proposed by Bateman et al. (2008) hypothesises a process where respondents become increasingly familiar with the strategic opportunities provided by the choice tasks they face and adjust their choices accordingly without having changed their preferences. In other words, strategic opportunities provide respondents with the incentive to mis-state rather than to disclose their true preferences.

The findings of Bateman et al. (2008) are supported by McNair et al. (2011) who provide evidence that increasing the number of choice sets per respondent decreases estimates of marginal WTP. McNair et al. (2011) argue that this effect may be explained by the ordering of alternative cost levels offered across a sequence of choice questions. The influence of ordering effects on DCE is further analysed by Day and Prades (2010). Using non-parametric analysis they observed ordering effects in both the cost and non-cost dimensions of the choice task. Their findings also suggest that these effects compound across choice sets.

Bateman et al. (2008) discuss alternative choice response strategies including strong and weak versions of cost-minimisation and good deal/ bad deal heuristics. They suggest that respondents who follow a strong cost-minimisation strategy assume that the good can be provided at the lowest cost offered. Hence, according to them, they are expected to never choose an option if a similar level of provision was offered in a previous choice set at lower cost. In contrast, respondents who employ weak cost-minimisation or good deal/ bad deal heuristics are assumed to trade-off between minimizing costs and reducing risks that the provision at a low cost level might not eventuate. Bateman et al.

³¹ The scale factor is inversely related to the variance of the error distribution (Swait and Louviere, 1993).

(2008) explain that the difference between a weak cost minimizing strategy and good deal/ bad deal heuristics lies in the assumption that respondents who follow the former have stable preferences at the beginning of the choice task whereas those who employ the latter do not.

Carson and Groves (2007; 2011) discuss an additional dimension of incentive compatibility. They argue that, for respondents to disclose private information truthfully, a consequential survey format is required. They define consequentiality as the environmental goods and services being of relevance to the respondent and respondents believing that their choices have an impact on the outcome. They point out that in inconsequential surveys respondents perceive choice options as indistinguishable and equally non-beneficial. They suggest that under such circumstances, it remains unknown whether or not respondents reveal their true preferences. Associated drivers postulated to additionally influence choice behaviour include the properties of the payment vehicle, plausibility of the choice questions, credibility of the policy scenario, and comprehensibility of the choice task (Carson and Groves, 2007). Surveys lacking a payment vehicle that respondents perceive as coercive may induce free-riding behaviour (Arrow et al., 1993). Implausible choice task may result in choices that are based on a different choice set than the one presented by the researcher (Carson and Groves, 2007). Carson and Groves (2007) further argue that if respondents are presented with a non-credible policy scenario they may be unsure whether presented options will be deliverable. They suggest that, if that is the case, respondents may include their perceived probability of provision into their choice rule. Carson and Groves (2007) propose that respondents who misunderstand the choice task may answer the question they think has been asked instead of the one the researcher intended to have answered. They conclude that, in an extreme case, respondents may choose without providing any information about their preferences if answering the choice questions lies beyond their capabilities.

Commonly, analysts using DCE assume consequentiality of the survey, plausibility of the choice questions, credibility of the policy scenario, and comprehensibility of the choice task. These assumptions, however, may be violated. Including follow-up questions concerned with exploring these issues is a possible means of investigating the extent to which true preferences are disclosed. However, the incentive properties of

such follow-up questions are unknown. Hence, the actual opinions of respondents may not be reflected in their answers.

4.3.3 Econometric modelling of discrete choice experiment data

Econometric models used to analyse DEC data have evolved in the course of the evolution of stated preference methods. As for dichotomous CV choice data, econometric models used to analyse discrete choice data are motivated by random utility theory (Manski, 1977; McFadden, 1974; Thurstone, 1927)³². In comparison to dichotomous choice CV, which asks respondents to answer a single referendum, respondents involved in the DEC format we use are asked to answer a sequence of choice questions with two or more choice options. This discrepancy in elicitation formats requires different econometric models to analyse the data. This section presents econometric models available to analyse DCE data.

Random utility theory interprets utility U that a sampled individual $i = 1, \dots, N$ obtains from choosing option $j = 1, \dots, J$ in choice situation $s = 1, \dots, S$ as a latent variable that can be elicited indirectly by observing individuals' choices. Utility U_{ijs} is described by an observable systematic component V_{ijs} and an unobservable random error term ε_{ijs} :

$$U_{ijs} = V_{ijs} + \varepsilon_{ijs}. \quad [27]$$

The systematic component of utility V_{jis} is defined as a linear, additive function in x_{jis} :

$$V_{jis} = \alpha_{jis} + \beta_i' x_{jis} + \delta_i' z_i, \quad [28]$$

where α_{jis} , β_i and δ_i are parameter vectors, x_{jis} is a vector of explanatory variables (attributes) and z_i is a vector of choice invariant characteristics of individuals.

Random utility theory assumes that a sampled individual $i = 1, \dots, N$ presented with alternative options J in choice situation $s = 1, \dots, S$ chooses the option that provides the highest level of utility to them. This implies that the probability that an individual chooses option m is equal to the probability that the utility of option m is greater than

³² The derivation of the basic MML/CL model follows, if not stated otherwise, Louviere et al. (2000).

(or equal to) all other options $j = 1, \dots, m, \dots, J$ in choice situation $s = 1, \dots, S$, which can be written as

$$\Pr(y_{is} = m) = \Pr(U_{mi} \geq U_{ji}) \forall j \in j = 1, \dots, J; m \neq j, \quad [29]$$

which is equivalent to

$$\begin{aligned} \Pr(y_{is} = m) &= \Pr\{(V_{mi} + \varepsilon_{mi}) \geq (V_{ji} + \varepsilon_{ji}) \forall j \in j = 1, \dots, J; m \neq j\} \\ &= \Pr\{(\varepsilon_{ji} - \varepsilon_{mi}) < (V_{mi} - V_{ji}) \forall j \in j = 1, \dots, J; m \neq j\}, \text{ assuming } U_j \neq U_m. \end{aligned} \quad [30]$$

Since this probability is a cumulative density function it can be formulated as a multidimensional integral over the probability density function $f(\varepsilon_i)$ (Train, 2009):

$$\Pr(y_{is} = m) = \int_{\varepsilon} I\{(\varepsilon_{ji} - \varepsilon_{mi}) < V_{mi} - V_{ji} \forall j \neq m\} f(\varepsilon_i) d(\varepsilon_i), \quad [31]$$

where $I = 1$, if $\{\}$ is true; $I = 0$ otherwise. For identification, one dimension of the probability density function needs to be normalised, which implies a scaling of the systematic component of utility V_{jis} .

The specification of this density function determines the type of discrete choice model. A range of distributions are applicable yet the extreme value type 1 (EV1) distribution (also called Gumbel distribution) is the one typically used. The EV1 distribution is characterised by a location parameter τ and a scale parameter λ . The cumulative distribution function has the form

$$F(\varepsilon, \tau, \lambda) = \exp(-\exp-(\varepsilon - \tau)\lambda). \quad [32]$$

The standard deviation of this distribution is

$$\sigma = 1 / \lambda(\pi / \sqrt{6}). \quad [33]$$

The derivation of an operational choice model is based on the Independence-from-Irrelevant Alternatives (IIA) axiom. This axiom states that ‘... the ratio of the probabilities of choosing one alternative over another (given that both alternatives have a non-zero probability of choice) is unaffected by the presence or absence of any additional alternatives in the choice set’ (Louviere et al., 2000, p. 44). The IIA axiom implies that the random error terms are independently (that is, uncorrelated) and identically distributed (that is, homoscedastic) across choice options (IID).

If the error terms ε_{jis} are assumed to be IID-EV1 distributed, the cumulative density function becomes (Ben-Akiva and Lerman, 1985):

$$F(\varepsilon_{jis}, \lambda_{jis}) = \exp(-\exp(-(\varepsilon_{jis})\lambda_{jis})). \quad [34]$$

The standard deviation of this distribution is

$$\sigma_{jis} = 1/\lambda_{jis}(\pi/\sqrt{6}). \quad [35]$$

Hence, the scale parameter is inversely related to the standard deviation. Assuming IID-EV1 distributed error terms implies a homoscedastic variance, $\sigma_{jis}^2 = (\pi^2/6)$, which is obtained by arbitrarily fixing the scale parameter to unity.

Additionally, assuming α_i and β_i to be invariant across individuals and repeated choices to be uncorrelated yields the conditional logit (CL) model³³ for individuals $i = 1, \dots, N$:

$$\Pr(y_{is} = j) = \exp(\lambda V_{jis}) / \sum_{m=1}^J \exp(\lambda V_{mis}), \quad [36]$$

$$\text{where } \lambda = 1/\sigma(\pi/\sqrt{6}). \quad [37]$$

In a CL model all scale parameters are arbitrarily fixed to unity. Hence it does not account for scale heterogeneity. If the scale parameter varies randomly the random utility parameters are correlated, which in turn means that assuming uncorrelated random parameters implies assuming a constant scale parameter (Scarpa et al., 2008).

The CL model is quite restrictive (for instance, Hensher et al., 2005). Firstly, a CL model assumes that preferences are homogenous; that is, the parameter vector β is the same across all individuals. Secondly, the IID assumption implies that choices individuals make in a sequence of choice situations are uncorrelated. Thirdly, the ratio of choice probabilities of any given pair of choice options is independent of any other choice options offered in each choice situation (IIA axiom). Mixed logit (ML) models^{34,35} overcome these limitations by disaggregating the error component in a

³³ The CL model is used for explanatory variables that vary across choice options (for example, attributes), yet invariant variables can be included in this specification by interacting the invariant variables with the constant; in contrast, the multinomial logit (MNL) model is specified for explanatory variables that are invariant across choice options (for example, socio-demographics) (Cameron and Triverdi, 2005).

³⁴ ML models are frequently called random parameter logit (RPL) models. In contrast to RPL models, ML model are able to include both random parameters and latent error components.

stochastic IID-EV1 error term and error terms that are based on underlying parameter vectors and observed data associated with choice options and respondents (for example, Brownstone and Train, 1999; Greene and Hensher, 2007; Greene, 2007; Hensher and Greene, 2003; McFadden and Train, 2000).

This relaxation provides the opportunity to model preference heterogeneity associated with individual specific preference parameters that are assumed to be distributed continuously over respondents around a fixed mean. The elements of the random parameter vector β_i and α_{ji} are defined as

$$\beta_{ki} = \beta_k + \sigma_k \nu_{ik}, \forall k \in k = 1, \dots, K, \quad [38]$$

and

$$\alpha_{ji} = \alpha_j + \sigma_j \nu_{ji}, \quad [39]$$

where β_k is the population preference mean for the k^{th} attribute, ν_{ik} is the individual specific preference heterogeneity (with zero mean and standard deviation one), σ_k is the standard deviation of the distribution of β_{ik} around β_k . This model can be further modified by assuming the means of the randomly distributed individual specific preference parameters to be heterogeneous. This can be specified as

$$\beta_{ik} = \beta_k + \delta_i' z_i + \sigma_k \nu_{ik}, \quad [40]$$

where z_i is a vector of choice invariant characteristics of individuals and δ_i a vector of parameters³⁶. The full vector of K random parameters can be written as

$$\rho_i = \rho + \Delta z_i + \Gamma \nu_i, \quad [41]$$

where $\Gamma = \text{diag}(\sigma_1, \dots, \sigma_K)$ is a triangular matrix with zero off-diagonal elements, that is, the random parameters are assumed to be uncorrelated. This assumption can be relaxed by assuming the lower triangular matrix to contain non-zero elements. In a random parameter specification, preference parameters can be assumed to be random across both respondents and choice tasks (cross-sectional) or across respondents but not choice tasks (panel). Cross-sectional data assume a single choice task per respondent whereas panel data assumes repeated choices per respondent.

³⁵ The discussion of the ML model follows, if not stated otherwise, Greene (2007).

³⁶ This model assumes that the variance is homoscedastic. The variance can be modelled as being heteroscedastic with observed individual characteristics (for example, Greene, 2007).

An ML model can be further extended to allow the variances to differ across choice options. The error component model is defined as (for example, Greene, 2007)

$$U_{jis} = \alpha_{jis} + \beta_i' x_{jis} + \delta_i' z_i + \eta_{ji} + \varepsilon_{jis}, \quad [42]$$

where the error component η_j is a random parameter with zero mean and estimated variance. The error component model can be integrated into an ML model specification.

The unconditional probability of an ML model is obtained by integrating v_i out of the conditional probability. The integral is approximated by averaging repeated random draws since it does not exist in closed form (for example, Revelt and Train, 1998).

Most discrete choice models, including the CL and ML model, apply maximum likelihood estimation to estimate parameter vectors. The log-likelihood function of a panel ML model is defined as (for example, Greene, 2007)

$$l(\theta) = \sum_{i=1}^N \log \frac{1}{R} \sum_{r=1}^R \prod_{s=1}^S \exp[\lambda(\alpha_{ji} + \beta_{ir}' x_{jis})] / \sum_{m=1}^J \exp[\lambda(\alpha_{mi} + \beta_{ir}' x_{mis})], \quad [43]$$

$$\beta_{ir} = \beta + \Delta z_i + \Gamma v_{ir}, \quad [44]$$

where R is the number of draws, β_{ir} is the r^{th} draw of β_i , v_{ir} is the r^{th} multivariate draw of individual i , and $\theta = \beta, \Delta, \Gamma$.

Implicit prices (IP) for attributes are derived by taking the ratio of estimated distributions of cost and non-cost parameters obtained from a choice model defined in utility space (for example, Hanley and Barbier, 2009):

$$IP_k = -\frac{\beta_k}{\gamma}, \quad [45]$$

where $\beta_k \forall k \in 1, \dots, K$ represents a vector of non-cost parameter estimates and γ is a cost parameter estimate. Commonly, a bootstrapping procedure is used to account for sampling errors (Krinsky and Robb, 1986). Compensating surplus, denoted by CS , or equivalent surplus, denoted by ES , can be calculated as follows (Hanley and Barbier, 2009):

$$CS = -\frac{1}{\gamma}(V_1 - V_0), \quad [46]$$

$$ES = -\frac{1}{\gamma}(V_0 - V_1), \quad [47]$$

where V_0 and V_1 represent the utility function individuals derive before and after the proposed marginal change in resource allocation, respectively.

Calculating implicit prices for any attribute k the cost parameter is often treated as a fixed parameter to ensure a negative sign, avoid potential identification problems (Ruud, 1996), and to facilitate the calculation of the marginal WTP distribution from the distribution of the non-cost parameter (Scarpa and Alberini, 2005). Treating the cost parameter as random may lead to unreasonably high or low mean estimates for implicit prices (Hensher and Greene, 2011). A fixed cost parameter, on the other hand, implies that the marginal utility of money is homogeneous across respondents, which is an unreasonable assumption in most cases. It also implies that the scale parameter is constant across all observations (Scarpa et al., 2008). If the assumption of homogeneous marginal utility of money is violated, scale heterogeneity will be ascribed incorrectly to variation in marginal WTP (Scarpa and Alberini, 2005). An alternative approach is to constrain the distribution of the cost parameter by assuming, for example, a constrained triangular distribution (Hensher et al., 2005). This may help to overcome the limitations mentioned above but essentially truncates preference heterogeneity (Hensher and Greene, 2011).

These limitations can be overcome by estimating implicit prices directly in WTP space (Scarpa and Alberini, 2005; Scarpa et al., 2008; Sonnier et al., 2007; Train and Weeks, 2005)^{37,38,39}. Separating the attribute vector into the cost attribute c_{ijs} and the vector of non-cost attributes x_{ijs} allows a reformulation of the utility function in preference space such that:

$$U_{jis} = -(\gamma_i c_{jis} + \beta_i' x_{jis} + \varepsilon_{jis}), \quad [48]$$

Following equation [48] the utility function in preference space can be redefined in WTP space by letting $\omega_i = \beta_i / \gamma_i$:

³⁷ As pointed out by Scarpa et al. (2008), the idea of willingness-to-pay space models was originated by Cameron and James (1987) and Cameron (1988) and extended to multinomial choice models with random preferences by Sonnier et al. (2007) and Train and Weeks (2005).

³⁸ To the author's knowledge, Scarpa et al. (2008) were the first to apply maximum simulated likelihood to random parameter models in willingness to pay space.

³⁹ The discussion of the WTP space model follows Scarpa et al. (Scarpa et al., 2008).

$$U_{ijs} = -(\gamma_i c_{jis} + (\gamma_i \omega_i)' x_{jis} + \varepsilon_{jis}). \quad [49]$$

In this model, the variation in IP_k is independent of scale λ_i , which cancels out, and separated from the variation in the cost parameter γ_i .

4.4 Conclusion

Chapter 4 has discussed both the opportunities for and the importance of analysing impacts of the evolution of stated preference methods on value estimates. Therefore, this chapter has established the foundation and justification for analysing the impact of the evolution of stated preference methods on value estimates related to environmental goods and services.

Chapter 4 opened with a brief review of the history of stated preference methods. This review demonstrated that even though most of the problems associated with the open-ended CV elicitation format have been investigated intensively, their potential to mask temporal value variability remains unexplored. The theory of prominent numbers suggests that even though values may have changed over long time periods, the open-ended CV method may not disclose this change. However, open-ended CV applications are needed as a baseline to explore values over extended time periods (see Chapter 3). Consequently, such a comparison of value estimates across time needs to be complemented by an analysis of potentially stabilising effects of open-ended CV elicitation formats.

Moreover, the argument in this chapter indicated that information processing and decision-making types may influence human behaviour since the availability of information and experience alone does not necessarily result in cognitive controlled behaviour (Fischer and Hanley, 2007). Therefore, unchanging value estimates over time may be mistakenly interpreted as indicative of stable preferences while instead expressing habits independent of the good under consideration, whereas value variability may reflect a change in the importance of certain cues rather than changes in rational consumer decisions. Considering social psychological models as conceptual frameworks to improve the interpretation of temporal value variability seems reasonable and may reveal a more detailed insight into choice behaviour across time.

As outlined in this chapter, the evolution of elicitation formats used in the CV method was accompanied by an evolution of econometric models available to analyse the obtained data. Binary choice models, hurdle models, corner solution models, and semi-parametric quantile regression models have the potential to facilitate a more rigorous analysis of values across time and may produce outcomes different from those obtained by linear and log-linear OLS regression models. This chapter showed that even though significant progress is being made in handling open-ended CV data, their specific characteristics still pose a significant challenge to the analyst.

It was additionally pointed out in this chapter that the incentive compatibility properties vary across alternative elicitation formats used in stated preference methods. While the incentive compatibility properties of most CV elicitation formats are widely explored, their impact on value estimates in DCE has not received much attention. Hence, Chapter 4 discussed compatibility properties and strategic response in DCE and their potential implications on value estimates.

In summary, the potential effects of methodological evolution on an analysis of variability across time have been explored in Chapter 4, providing the base to explore the effect of unexplored elements associated with methodological evolution on the temporal variability of value estimates related to environmental goods and services. Together, Chapter 3 and Chapter 4 have built a foundation to formulate the research questions and state the associated hypotheses to be tested in this thesis.

5 RESEARCH QUESTIONS AND HYPOTHESES

5.1 Introduction

Considering environmental goods and services in CBA provides the potential for improvements in economic efficiency through a reduction in market failure introduced by incomplete property rights structures (see Section 1.1). Non-market valuation methods based on the neo-classical microeconomic theory of consumer choice (see Section 2.2) are used to obtain marginal WTP estimates of non-marketed values attached to environmental goods and services (see Section 2.3). Only stated preference methods, such as CV and DCE, are capable of eliciting both use and non-use values (see Section 2.1).

Policy decisions concerned with the allocation of natural resources often have long-term and irreversible impacts on the environment. Hence, policy decisions based on CBA have to consider potential instabilities of non-market values attached to environmental goods and services across time. Ignoring these instabilities will reduce economic efficiency in resource allocation across time (see Section 1.1).

Few studies have examined the potential variability of non-market values attached to environmental goods and services across time. Most studies have been concerned with comparing marginal WTP estimates obtained at different points in time over relatively short time spans or have analysed the temporal reliability of stated preference methods. Consequently, little is known about either the potential change of non-market values attached to environmental goods and services across extended time periods (see Chapter 3) or the impact of the evolution of stated preference methods on marginal WTP estimates (see Chapter 4).

The economic literature suggests that potential temporal variability of non-market values attached to environmental goods and services is motivated by changes in factors on both the supply and the demand side. Hence value variability across time is expected to be driven by changes in preferences (see Section 3.2.1) and budget constraints (see Section 3.2.2); the scale of provision (see Section 3.3.1); the availability of close

substitutes and associated relative scarcity (see Section 3.3.2); and reference point relativity (see Section 3.3.3).

Apart from value changes across time, advances in stated preference methods may affect marginal WTP estimates of non-market values attached to environmental goods and services (Chapter 4). These impacts can be disaggregated into effects of improvements of existing stated preference methods, development of new stated preference methods, and progress in econometric modelling. Both improvements in existing stated preference methods and the development of new stated preference methods have been accompanied by an evolution of elicitation formats and changes in their respective impact on value estimates. These improvements include potentially stabilising effects of the open-ended CV elicitation format (see Section 4.2.2) and impacts of incentive compatibility properties (see Section 4.3.2). The effects of the incentive compatibility properties of CV elicitation formats on marginal WTP estimates are relatively well understood, whereas those of DCE elicitation formats still remain less clear.

Against this background, the objective of this research is to analyse values and valuation across time by exploring the following two research questions:

- (1) Do non-market values attached to environmental goods and services change across extended time periods (value variability across time)?
- (2) Has the evolution of stated preference valuation methods affected marginal WTP estimates for non-market values attached to environmental goods and services (valuation variability across time)?

The two research questions are analysed by testing a number of hypotheses, expressed as the null.

The remainder of this chapter is organised as follows. Sections 5.2 and 5.3 develop the hypotheses tested to explore the two research questions stated above. Section 5.4 provides a summary and a conclusion.

5.2 Value variability across time

The first research question investigates whether non-market values attached to environmental goods and services change across extended time periods. For this purpose, the first hypothesis is concerned with investigating whether the marginal WTP estimates of the non-use values attached to environmental goods and services have changed across extended time periods. It is assumed that the marginal WTP estimates represent the true, but for the researcher, unknown, value. That is, potential impacts of the stated preference method used to obtain the marginal WTP estimates are ignored. The marginal WTP estimates are obtained by an open-ended CV since it is the stated preference method that has been in use the longest (see Section 4.2), which allows, theoretically, a comparison of marginal WTP estimates over 40 years. This leads to the following hypothesis:

Hypothesis 1.1:

H_0 : The marginal WTP estimates of non-use values attached to environmental goods and services obtained by an open-ended CV are stable across extended time periods.⁴⁰

Also of interest is whether the influence of drivers of demand on marginal WTP estimates of non-use values attached to environmental goods and services are stable across extended time periods. This question gives rise to the following two hypotheses:

Hypothesis 1.2:

H_0 : The underlying relationships between marginal WTP estimates of non-use values attached to environmental goods and services and explanatory variables of demand obtained by an open-ended CV remain stable over extended time periods.

⁴⁰ The choice of whether to define the null hypothesis as WTP stability or WTP variability remains disputable in the case of an analysis concerned with the temporal variability of marginal WTP across extended time periods since prior expectations might be agnostic.

Hypothesis 1.3:

H_0 : The magnitudes of explanatory variables of demand that exhibit a significant underlying relationship with marginal WTP estimates of non-use values attached to environmental goods and services obtained by an open-ended CV do not change across extended time periods.

The quantitative and qualitative characteristics of the environmental goods and services under consideration are kept constant to avoid confounding effects (see Section 3.3.1). Additionally, following Hanemann's separability condition (1988), the relative prices of composite goods are assumed to be negligible. Hence marginal WTP related to environmental goods and services under consideration is assumed to be uninfluenced by price changes of private goods.

As outlined in Section 3.1, analysing value variability across time requires the assumption of stable preferences. This assumption may be justifiable over short time periods. However, separating the effects of an unreliable non-market valuation method from those of changed impacts of explanatory variables of demand over extended time periods poses a considerable challenge. Hence, Hypothesis 1.2 is based on the maintained assumption that the stated preference method used to obtain the marginal WTP estimates is reliable across extended time periods.

The examination of demand side drivers suggested by social psychological research are limited and restricted to the influence of changes in behaviour, environmental attitudes, and attitudes towards the payment vehicle. The reasons for this restriction include limited availability of reference data as well as concerns about the compatibility of social psychological concepts in an economic context (see Section 2.4).

Since the introduction of stated preference methods, alternative payment vehicles and their impact on marginal WTP estimates have been subject to scrutiny. This scrutiny has led to their continuous improvement. However, the payment vehicle applied in the reference study is used for the comparison across time despite its shortcomings from a contemporary perspective. Analysing changes in attitudes towards the payment vehicle can be regarded as being subject to both the first research question (values across time) and the second (valuation across time). In this study, it is analysed under the former. This leads to the following hypothesis:

Hypothesis 1.4:

H_0 : The attitudes of respondents towards the payment vehicle presented in an open-ended CV survey associated with non-use values attached to environmental goods and services remain stable across extended time periods.

Possible impacts of changes in supply variables on marginal WTP estimates related to non-marketed environmental goods and services are investigated by exploring changes in the availability of close and cheap substitutes and associated changes in the relative scarcity of the environmental goods and services (see Section 3.3.2). Assuming that changes in the availability of close and cheap substitutes directly affect the relative scarcity of the environmental goods and services may imply that an increase in supply may decrease its marginal WTP across time. The qualitative and quantitative characteristics of the environmental asset are held constant across time to avoid potentially confounding effects (see Section 3.3.1). This leads to the following hypothesis:

Hypothesis 1.5:

H_0 : The availability of close and cheap substitutes for environmental goods and services that provide non-use benefits and their associated relative scarcity remains stable over extended time periods.

The available data for a rigorous analysis of potential impacts of varying perceived reference points is insufficient and assumed to be negligible (see Section 3.3.3).

5.3 Valuation variability across time

The progress in stated preference methods has been substantial over the last 40 years. Exploring exhaustively its impact on marginal WTP estimates of non-market values attached to environmental goods and services is beyond the scope of this study. Hence, to explore the second research question, the research hypotheses are focused on the effect related to the introduction of DCE and the evolution of elicitation formats used in CV and DCE (see Chapter 4).

To examine the impact of the introduction of the DCE method as an alternative to the CV method (see Section 4.2 and Section 4.3) the following hypothesis is tested:

Hypothesis 2.1:

H_0 : Hicksian surplus of non-use values attached to environmental goods and services obtained by an open-ended contingent valuation is not different from that elicited by a state-of-the-art DCE.

Assessing the drivers for differing marginal WTP estimates obtained by these two methods, however, is beyond the capacity of this comparison. The aim of such a comparison is rather to raise awareness that using outdated stated preference surveys for current policy decisions may not be advisable.

The use of the open-ended elicitation format in CV has been criticised, among other things, for being incentive incompatible (see Section 4.2); hence, most current CV studies use alternative elicitation formats (see Section 4.1). Apart from its incompatible incentive properties, open-ended CV may also suffer from a ‘prominence in response’ bias (see Section 4.2.2). This bias may result in stable marginal WTP estimates across extended time periods, even though the values are, in fact, variable. Hence, results obtained by open-ended CV may suggest stable values across extended time periods, whereas state-of-the-art CV outcomes may suggest the opposite. The analysis of the impact of the open-ended elicitation format used in CV applications on the variability of marginal WTP estimates of non-use values attached to environmental goods and services yields the following hypothesis:

Hypothesis 2.2:

H_0 : Prominence in response does not influence possible temporal variability of marginal WTP estimates of non-use values attached to environmental goods and services elicited through open-ended CV across extended time periods.

Another aspect of changes in the elicitation format across time is targeted by analysing the incentive compatibility properties of and potential strategic behaviour in DCE (see Section 4.3.2). Contrary to the preceding hypotheses, however, the conducted DCE consider both use and non-use values attached to environmental goods and services. This aspect gives rise to the following hypothesis:

Hypothesis 2.3:

H_0 : DCE outcomes associated with use and non-use values attached to environmental goods and services obtained by a single-binary elicitation format are the same as those obtained by a repeated-binary elicitation format that contains four choice sets.

Furthermore, it has been proposed that the respondents' sole awareness of having multiple choices may induce strategic behaviour leading to the following hypothesis:

Hypothesis 2.4:

H_0 : DCE outcomes associated with use and non-use values attached to environmental goods and services obtained by a single-binary elicitation format are the same as those obtained by the first choice question of repeated-binary elicitation formats that contain four choice sets.

The literature suggests that the order in which choice sets are presented to respondents provides strategic opportunities that may affect choice decisions. This is investigated by testing the following hypothesis:

Hypothesis 2.5:

H_0 : Strategic opportunities provided by the order in which choice sets concerned with use and non-use values attached to environmental goods and services are presented to respondents in DCE do not affect choice decisions.

To examine whether strategic response coexists with other effects discussed in the literature such as strategic learning, value learning, and institutional learning, it is additionally investigated whether choice decisions are stable across a sequence of choice tasks:

Hypothesis 2.6:

H_0 : Choice decisions concerned with use and non-use values attached to environmental goods and services are stable across a sequence of DCE choice tasks.

5.4 Conclusion

In Chapter 5, the research objective - analysing value and valuation variability across time - was systematically deduced from the literature review (see Chapter 1 through Chapter 4). This research objective has provided the framework for the formulation of a range of hypotheses that were grouped into two subsets by the following two core research questions:

- (1) Do non-market values attached to environmental goods and services change across extended time periods (value variability across time)?
- (2) Has the evolution of stated preference methods affected marginal WTP estimates of non-market values attached to environmental goods and services (valuation variability across time)?

Apart from the research objective and the hypotheses, this chapter has stated three maintained assumptions (MA) that are required for the analysis:

MA 5.1: The applied stated preference method is temporally reliable.

MA 5.2: Changes in the relative prices of composite goods do not influence marginal WTP.

MA 5.3: The effect of varying perceptions of reference points is negligible.

Clearly, the effects of potential violations of these assumptions have to be acknowledged in the interpretation of the results.

The deduction of the research objective and the development of the hypotheses defined the general scope of the research. However, the stated hypotheses have to be embedded in a specific case study, which will be introduced in Chapter 6. This embedding will necessitate adjustments of the stated hypotheses that will be discussed in Chapter 7.

6 CASE STUDY

6.1 Introduction

As outlined in Chapter 5, the objective of this thesis is to analyse value and valuation across time, whereby the respective hypotheses are tested with data obtained by stated preference surveys concerned with estimating non-market values related to environmental goods and services. The purpose of Chapter 6 is to present the reference study for the comparison across time, an open-ended CV study carried out in 1979 (Bennett, 1981; 1984). This study elicited the preservation value of Nadgee Nature Reserve. Embedding the analysis in this specific case study necessitates an adjustment of the stated hypotheses.

The remainder of this chapter is structured as follows. Section 6.2 introduces Nadgee Nature Reserve, provides an overview of its evolutionary history, and discusses the public good character of the preservation of Nadgee Nature Reserve. The reference study is described in Section 6.3. This section gives information about the survey logistics, the questionnaire, and the econometric framework used to analyse the data. It also debates, from today's perspective, the shortcomings of the reference study and its implications for this thesis. This is followed by Section 6.4, which revisits and refines the hypotheses. Finally, a summary and a conclusion are provided in Section 6.5.

6.2 Nadgee Nature Reserve⁴¹

Nadgee Nature Reserve is situated on the south coast of New South Wales between Wonboyn Lake and the Victorian border. The reserve consists of a large variety of landscape features including low tablelands, plateau remnants, coastal ranges, coastal plains, estuaries, coastal lagoons, beaches, cliffs, dunes, sea caves and other coastal landforms. These landscape features support a diversity of habitats with a highly pristine character; that is, habitat fragmentation, habitat degradation and habitat loss are

⁴¹ The information provided in Section 6.2 was obtained, if not stated otherwise, from the New South Wales National Parks and Wildlife Service (2003).

minimal. These pristine habitats contain a large number of rare, endangered and biogeographically significant native plant and animal species. Examples include the Green Scentbark (*Eucalyptus ignorabilis*), the Burnettia (*Burnettia cuneata*), the Aromatic Peppercress (*Lepidium hyssopifolium*), the Green and Golden Bell Frog (*Litoria aurea*), the Eastern Bristlebird (*Dasyornis brachypterus*), and the Southern Brown Bandicoot (*Isoodon obesulus*). Apart from this rich natural endowment, the reserve contains valuable Aboriginal heritage and several other significant historic sites.

The declaration of a fauna reserve (11,430 hectares) in 1957 paved the way for Nadgee Nature Reserve, which was officially established in 1967. Successive expansions have increased the protected area to its present 20,671 hectares. The reserve embodies the only coastal wilderness area in New South Wales. Nadgee Wilderness was declared in 1994. Its current extension incorporates approximately 18,800 hectares or about 91% of Nadgee Nature Reserve. Nadgee Wilderness joins Cape Howe Wilderness in Victoria. Together they represent the largest undisturbed coastal area in south-eastern Australia.

Wilderness areas in New South Wales are declared under the Wilderness Act 1987 and

... shall be managed as so (a) to restore (if applicable) and to protect the unmodified state of the area and its plant and animal communities, (b) to preserve the capacity of the area to evolve in the absence of significant human interference, and (c) to permit opportunities for solitude and appropriate self-reliant recreation (New South Wales Government, 1987, Section 2).

The current management plan for the reserve was adopted by the Minister of the Environment in 2003. The main objective of this management plan is to conserve its natural and cultural values. The management plan thus implies restrictions on the use of the reserve, and in particular, of Nadgee Wilderness. The use of Nadgee Wilderness is limited to management activities, aboriginal cultural activities, scientific research, environmental education, and low-impact recreational activities. Low-impact recreational activities are managed by means of a permit system for overnight bush walkers, minimising vehicle tracks and facilities, camping restrictions, and information campaigns. Horse riding, cycling and orienteering are explicitly prohibited.

The natural endowment of Nadgee Nature Reserve supplies a range of economic benefits (see Section 2.5). Provided use benefits include direct use values, indirect use

values, option values, and quasi-option values. Non-use benefits supplied by Nadgee Nature Reserve include bequest values and existence values.

Non-use values and indirect use values are non-diminishable and non-excludable in nature, whereas many of the provided direct use values are not. However, as outlined previously, direct use is currently restricted to scientific research, education, aboriginal cultural activities, and low-impact recreation. Hence, the currently consumed direct use benefits are limited to those that are indivisible in consumption. Direct use benefits derived from environmental education and aboriginal cultural activities are additionally considered non-excludable. However, whether direct use benefits associated with scientific research and recreation are non-excludable in nature is subject to debate. Individuals can be excluded from recreational activities through the established permit system. Nonetheless, the permit system excludes individuals on the basis of visitor numbers. That means anyone willing and able to pay the associated permit fee is able to eventually receive a permit. The permit fee is low (\$A5 per person per day) and thus only a fraction of the benefits provided to visitors are likely to be reflected in the revenue received by the provider, the New South Wales Government. Benefits associated with low-impact recreation can therefore be classified as non-excludable. Similarly, scientific research is encouraged by the provision of access and accommodation. On the other hand, some scientific results may generate exclusively private benefits. Assuming the extent of private benefits to be insignificant, benefits flowing from scientific research can be classified as non-excludable. In conclusion, the preservation of Nadgee Nature Reserve is predominantly characterised by indivisibility in consumption and non-excludability and can therefore be classified as a public good.

6.3 Reference study

The reference for the comparison across time is an open-ended CV study carried out in 1979 (Bennett, 1981; 1984) (henceforth referred to as ‘reference study’). The reference study elicited the preservation value of Nadgee Nature Reserve. Value was measured as Hicksian equivalent surplus since consumers were assumed not to have the property rights to keep the current level of utility.

The survey material consisted of a questionnaire, show cards and photographs. The questionnaire was part of an omnibus survey carried out by the Survey Research Centre,

Australian National University, in September and October 1979. The questionnaire about the preservation value of Nadgee Nature Reserve was bundled with several other surveys related to the quality of ACT schools, drug use, family planning, and public opinions on ACT politics. Fortunately, the questionnaire concerned with the preservation value of Nadgee Nature Reserve (henceforth referred to as ‘Nadgee Nature Reserve Questionnaire’) was positioned directly after the first part of the omnibus survey that asked respondents about their socio-demographic characteristics. Consequently, the impact of the subsequent questionnaires on the survey of interest is most likely negligible⁴².

In the first part of the Nadgee Nature Reserve Questionnaire, respondents were asked about their general experience of visiting protected areas in Australia or worldwide. This was followed by a sequence of questions about the importance respondents attached to preserved areas of land in general.

In the second part of the Nadgee Nature Reserve Questionnaire respondents were provided with background information about a specific preserved area of land. The described features included beaches, sand dunes, a salt water lake, river estuaries, areas of coastal heathland, dry and wet eucalypt forest, and upland heath. This description was supported by photographs showing different landscape types, native plants and native animals. The preserved area of land was described in term of the features of Nadgee Nature Reserve without revealing its identity. This concealment was required since the natural area of land was already protected. Asking for a payment to have something protected that is already protected would have reduced the credibility of the survey. This concealment was also an attempt to isolate existence benefits from all other benefits provided by Nadgee Nature Reserve. This attempt was further supported by describing the area as isolated and extremely difficult to access. Additionally, respondents were told that in order to preserve this area of land in its natural state visits by the public, including access for educational purposes, would have to be very limited; that is, access would mainly be restricted to scientists and rangers. Finally, respondents were advised that even if the scientific research alone produced no benefits, existence benefits would still be generated.

⁴² Note, however, that respondents were aware of the range of topics included in the omnibus survey.

Despite these efforts, it is debatable whether the isolation of existence values was successful. It is likely that respondents not only considered existence values but also bequest values, indirect use values, direct use values associated with scientific research, and option values. Hence, the remainder of this thesis uses the umbrella term ‘preservation value’ for the value elicited by the reference study. The second part of the survey concluded with questions about the importance of the preservation of this particular natural area of land.

In the third part of the Nadgee Nature Reserve Questionnaire, respondents were told that funds had to be raised to enable the government to purchase the land and conserve the area. Respondents were asked whether they were willing to pay an extra tax or give a donation to ensure the preservation of the natural area of land. The next question asked whether respondents preferred the government (extra tax) or a conservation organisation (voluntary donation) to collect the funds. This was followed by asking respondents how much they were willing to pay. The final part of the survey concluded with questions about environmental attitudes and behaviour.

A pilot survey was carried out to test and refine the Nadgee Nature Reserve Questionnaire and other survey material. This pilot survey was based on a randomly selected sample of 30 members of the population of Canberra, drawn from the electoral rolls of the Federal divisions of Fraser and Canberra. The complete and final Nadgee Nature Reserve Questionnaire of the reference study is attached in Appendix A.

Sampling and interviewing for the main survey were carried out by the Survey Research Centre, Australian National University. The sample was based on a 1-in-100 systematic random sample of block and section references from a Department of the Australian Capital Territory Rates File. This procedure yielded, after the omission of temporary residents, diplomats and vacant lots, a sample of 590 households. After having collected information on the socio-demographics of each member of these households (1,883 respondents)⁴³, 593 respondents were randomly chosen for personal interviews using an identification code and their position in the household. Respondents were informed of their selection by letter prior to the interview. Refusals, contact failure and language difficulties of some respondents reduced the final sample to 544 completed and usable

⁴³ The sampling procedure described by Bennett (1981) did not provide detailed information about the collection method.

interviews⁴⁴. A response rate was not provided by Bennett (1984). The sample was randomly divided into four groups of 380, 54, 51, and 59 respondents respectively, with each group receiving a different version of the questionnaire. This division was required to conduct tests designed to analyse the impact of differing levels of information. To avoid confounding effects of differing questionnaire versions, the sample used as reference for the research presented in this thesis is restricted to the group of 380 respondents. Furthermore, 51 questionnaires with contradictory answers regarding preparedness to pay (PTP) and marginal WTP were listwise deleted⁴⁵, leaving 329 valid questionnaires. The personal interviews were conducted by trained interviewers who were familiar with the questionnaire and with potential problems associated with questions that were expected to arise.

The data were preserved as Excel files by the Australian Social Science Data Archives. The original completed questionnaires, however, were not traceable. Therefore, potential data entry errors cannot be detected. Although a codebook was available, the coding remains partially unclear.

The data were analysed using descriptive statistics, discriminant analysis, and OLS estimations based on linear and log-linear regression models (see Section 4.2.3). Zero values of marginal WTP were excluded from the econometric analysis in order to avoid negative marginal WTP predictions and allow the estimation of log-linear models.

The reference study was the first CV study applied in an Australian context (Bennett, 1981; 1984). Since then the published literature has advanced significantly. Hence, compared to a state-of-the-art CV questionnaire design (for example, Alberini and Kahn, 2006; Carson, 2000) the reference study is problematic in many respects. This becomes clear when looking at the study's key characteristics.

First of all, the reference study used an open-ended elicitation format, which is known to be incentive incompatible; that is, revealing private information truthfully was probably not a dominant strategy for all participants (see Section 4.3.2). Furthermore, the reference study did not include a provision rule. The provision rule is the link between respondents' choices and the corresponding policy outcome. Whether revealing

⁴⁴ The sampling procedure described by Bennett (1981) did not provide detailed information about contact establishment and management.

⁴⁵ Responses are defined as contradictory if the PTP question was answered with 'yes' and the WTP questions was answered with zero.

true preferences is a dominant strategy thus depends on both, the elicitation format and the expectations of respondents about the provision rule used to aggregate their choices.

An additional problematic characteristic of the reference study is its payment vehicle (see Section 4.1). The survey asked respondents to choose between a voluntary donation, a compulsory tax, or a mixture of both. This payment vehicle mixture influenced the incentive structure of the survey. The option of a voluntary donation most likely caused free-riding behaviour even though the income tax was a coercive payment vehicle for the majority of the respondents. A tax may be a biased payment vehicle since only respondents subject to this tax perceive it as coercive. However, since all respondents had the opportunity to choose the voluntary donation, this potential impact is considered irrelevant.

Moreover, for respondents to disclose private information truthfully, a consequential survey format is required. Consequentiality, in terms of Carson and Groves (2007), means that the goods and services have to be of relevance to the respondent and respondents have to believe that their choices have an impact on the outcome (see Section 4.3.2). In inconsequential surveys, respondents may perceive choice options as indistinguishable and equally non-beneficial. Associated drivers postulated to additionally influence choice behaviour include plausibility of the choice questions, credibility of the policy scenario, and comprehensibility of the choice task. However, the reference study did not include follow-up questions concerned with exploring these issues. Consequently, it remained unclear to what extent true preferences were disclosed.

Economic choices are not solely a result of individual preferences but are constrained by an individual consumer budget (see Section 2.2). In the reference study, however, respondents were not reminded of their budget constraint. Respondents who were unaware of their individual budget constraint may have overstated their stated marginal WTP.

Furthermore, the study design process relied on comments and suggestions of colleagues rather than on focus groups input. Refining survey material by means of collegial feedback is problematic since a group of selected colleagues is most likely an unrepresentative sample of the population.

Finally, the data of the reference study were analysed using OLS estimation based on linear and log-linear models. Zero values of marginal WTP were excluded from the econometric analysis in order to avoid negative marginal WTP predictions and allow the estimation of log-linear models. However, econometrics has made significant progress in modelling data obtained by open-ended CV (see Section 4.2.3). The development of new models such as (double) hurdle models, corner solution models, sample selection models, or count models has provided alternative treatments for samples with a non-trivial fraction of zero marginal WTP values. Hence, from today's perspective, using OLS linear regression models for analysing open-ended CV data may be regarded as inappropriate.

Irrespective of its shortcomings, the estimates of the reference study are still listed in the Environmental Valuation Reference Inventory (2011), a benefit transfer data base for policy applications. This has motivated the idea to re-estimate the preservation value of the Nadgee Nature Reserve not only by replicating the reference study using the original survey material and method (henceforth referred to as 'replicated study') but also by using a state-of-the-art stated preference method. Testing for causes explaining potential differences across methods, however, is beyond the capacity of this comparison. The aim of such a comparison is rather to raise awareness that using outdated stated preference surveys for current policy decisions is not advisable. In general, both state-of-the-art CV and DCE could be used for this re-estimation. However, this study uses DCE to re-estimate the preservation value of Nadgee Nature Reserve. The reasons for this decision are the following. Firstly, an important aim of this thesis is to analyse the incentive compatibility properties of DCE. Secondly, the number of split samples required to rigorously test for impacts of changes in CV would be unmanageably large. Thirdly, the current trend, especially in Australia, is towards using DCE. Fourthly, future replications of the study should be based on what is currently considered state-of-the-art.

6.4 Hypotheses revisited

The chosen reference study necessitates adjusting the hypotheses stated in Sections 5.2 and 5.3. As outlined in Section 6.2 and 6.3, this thesis examines non-marketed use and non-use values with public good character. Firstly, this study examines changes in the

preservation value of Nadgee Nature Reserve, over a 30-year time period (1979 – 2009). Over this time period, the area of Nadgee Nature Reserve has been expanded. Thus, the area size presented to respondents had to be held constant in order to minimise confounding effects of changing supply conditions. The stated hypotheses are therefore based on the maintained assumption that respondents were not aware that the size of the area of land presented in the questionnaire in 2009 differed from its actual size in 2009. Since the identity of the area of land described in the survey had not been revealed this maintained assumption is likely to hold.

Additionally, the original CV study aimed to restrict the estimation to non-use values (preservation value). The DCE study includes values associated with low-impact recreation. As outlined in Section 4.3.1, the DCE method allows the elicitation of implicit prices for each of the attributes used to describe the environmental goods and services under consideration. Therefore, in theory, it is possible to separate the equivalent surpluses of use and non-use values.

The adjusted hypotheses are presented in Chapter 7.

6.5 Conclusion

Chapter 6 has introduced the public good to be used as the test-bed for the hypotheses, the preservation of Nadgee Nature Reserve, as well as the reference study used to achieve the research objective. In Section 6.2, Nadgee Nature Reserve was characterised as a relatively undisturbed area of land on the south coast of New South Wales with a rich natural and cultural endowment. The reserve, founded as a fauna reserve in 1957, has been successively expanded to its present 20,671 hectares, of which about 91 % is declared as a wilderness area.

Section 6.2 emphasised that the management plan for Nadgee Nature Reserve, adopted in 2003 by the Minister of the Environment, implies restrictions on the use of the reserve, and in particular, Nadgee Wilderness. Hence, it was argued that, since the consumption of benefits flowing from Nadgee Nature Reserve is currently restricted to those exhibiting a non-exclusive and non-diminishable character, Nadgee Nature Reserve can be classified as a public good.

The reference for the comparison across time was introduced in Section 6.3. This reference, an open-ended CV study that was carried out in 1979, elicited the preservation value of Nadgee Nature Reserve (Bennett, 1981; 1984). Detailed information about the survey logistics, the questionnaire, and the econometric framework used to analyse the data were given. This was followed by a discussion about the shortcomings of the reference study from today's perspective. The potential impacts of the methodological evolution of stated preference methods over the last 40 years as outlined in Chapter 4 were emphasised. Section 6.3 also justified the decision to restrict the re-estimation of the preservation value of the Nadgee Nature Reserve on a state-of-the-art DCE and a replication of the reference study using the original survey material and method.

Furthermore, the successive expansion of Nadgee Nature Reserve outlined in Section 6.2 required the formulation of a maintained assumption to minimise confounding effects of changing supply conditions:

MA 6.1: Respondents were unaware that the size of the area of land presented in the questionnaire in 2009 was held constant at 1979 levels and thus differed from its current size.

Embedding the analysis in this specific case study necessitated an adjustment of the stated null hypotheses (Section 6.4). Firstly, adjustments were conducted with respect to the value type. The adjusted null hypotheses focus on preservation and recreational values. Secondly, the reference study defined the time horizon to be 30 years.

The adjusted hypotheses (henceforth again simply referred to as 'Hypotheses') are a concise base to analyse the research questions associated with the research objective. They will be presented together with the research design and methods in Chapter 7.

7 RESEARCH DESIGN AND HYPOTHESES TESTING

7.1 Introduction

Chapter 5 stated the hypotheses to be tested to explore the two main research questions of this study:

- (1) Do non-market values attached to environmental goods and services change across extended time periods (value variability across time)?
- (1) Has the evolution of stated preference methods affected marginal WTP estimates of non-market values attached to environmental goods and services (valuation variability across time)?

This was followed by Chapter 6, which introduced the case study and provided a rationale for adjusting the hypotheses stated in Sections 5.2 and 5.3. This chapter presents the adjusted hypotheses as well as the research methods used to test them.

Following the grouping of the hypotheses under the two research questions, the research design is divided into two parts. This is reflected in the organisation of the remainder of this chapter. Section 7.2 outlines the research design and methods used to explore potential value variability across time analysing stated marginal WTP and explanatory variables of demand and supply. Section 7.3 describes the research design and methods applied to investigate potential impacts of the methodological evolution of stated preference methods on value estimates attached to environmental goods and services. There is a focus on exploring potential impacts of the CV open-ended elicitation format on the stability of value estimates through prominence in response, exploring the performance of DCE as an alternative to open-ended CV, and analysing the influence of the evolution of elicitation formats on the incentive compatibility properties of stated preference methods.

In this study, stated preference methods are both the research object and the research method. Hence, DCE and CV in their capacity as research objects were discussed in

detail in Chapter 4. To avoid unnecessary repetition, references to the relevant sections of Chapter 4 will be provided as necessitated.

7.2 Value variability across time

Microeconomic theory of consumer choice suggests that environmental values may be variable across time (see Chapter 3). A re-estimation of the preservation value of Nadgee Nature Reserve using the original survey material and method (replicated study) is used to afford a comparison across a 30-year time period. The reference study described in Section 6.3 serves as the baseline. This comparison focuses on stated marginal WTP, as well as on the impact of explanatory variables of demand and supply.

7.2.1 Stated marginal willingness to pay

The first hypothesis is concerned with investigating whether stated marginal WTP for the preservation of Nadgee Nature Reserve obtained by an open-ended CV has changed across a 30-year time period. Exploring the temporal variability of stated marginal WTP obtained by an open-ended CV gives rise to the following adjusted hypothesis:

Hypothesis 1.1:

H_0 : Stated marginal WTP for the preservation of Nadgee Nature Reserve obtained by an open-ended CV in 1979 is the same as stated marginal WTP obtained in 2009.

To test this hypothesis, marginal WTP stated in the reference study is compared to inflation adjusted⁴⁶ marginal WTP stated in the replicated study. In particular, this hypothesis is tested by exploring differences in the proportion of positive marginal WTP bids, the mean and median of stated marginal WTP, and the distribution of stated marginal WTP across the two samples. This comparison is expanded by dividing stated marginal WTP bids by the number of household members (HM). A range of statistical tests is used to examine statistically significant differences, including the Pearson's chi-squared test (independence of observations), t-test on the equality of means, non-

⁴⁶ Consumer Price Index Canberra (1979-2009: Factor 3.82191780821918)

parametric K-sample test on the equality of medians, Mann-Whitney test on the equality of distributions (Mann and Whitney, 1947), and a test on equality of standard deviations robust to non-normality (Brown and Forsythe, 1974; Levene, 1960).

Protest zero marginal WTP bids are examined separately. A test on equal proportions (pr-test) is used to check for statistically significant differences in the proportion of protest zero marginal WTP bids across the two samples. This is complemented by an analysis of differences in protest motivation using a Pearson's chi-squared test (independence of observations).

7.2.2 Explaining marginal willingness to pay – demand side

The influence of explanatory variables of demand on the preservation value of Nadgee Nature Reserve across time is explored on the basis of the following hypothesis:

Hypothesis 1.2:

H_0 : The underlying relationships between marginal WTP estimates for the preservation of Nadgee Nature Reserve and explanatory variables of demand obtained by an open-ended CV are the same in 1979 and 2009.

Hypothesis 1.3:

H_0 : The magnitudes of explanatory variables of demand that exhibit a significant underlying relationship with marginal WTP estimates for the preservation of Nadgee Nature Reserve obtained by an open-ended CV are the same in 1979 and 2009.

Based on economic theory discussed in Chapter 3, the following variables are hypothesised to explain marginal WTP for the preservation of Nadgee Nature Reserve:

- socio-demographics (age, household income⁴⁷, gender, non-school education, place of birth)
- attitudinal variables (interest in nature study, importance of Nadgee Nature Reserve to the respondent)

⁴⁷ In this thesis, differences in income across time are treated as real rather than as relative differences. However, this might not be true for extended time periods. Political scientists and sociologists have long argued that the relative position and generational cohort effects may affect results.

- behavioural variables (visit to preserved areas in Australia; subscription to any conservation organisation; reading books or watching TV programmes featuring the study of nature, formal training in biology or related studies)

Hypothesis 1.2 is tested by comparing these underlying relationships represented by parameters estimates obtained by a state-of-the-art econometric model at two different points in time using the reference data (1979) and the replicated data (2009). This requires choosing a state-of-the-art econometric model that is most appropriate to deal with the characteristics of the collected data. Alternatives available today include probit and logit models, corner solution models, hurdle models, interval models, and quantile regression models. The choice of a state-of-the-art econometric model that is most appropriate to deal with the characteristics of the collected data is facilitated by investigating the data characteristics of the samples obtained in the reference and the replicated study. Based on that information, several models are estimated and compared.

However, neither of the state-of-the-art econometric models are yet capable of dealing with challenges posed by open-ended CV data and, at the same time, differentiating between corner solutions and protest zero marginal WTP bids. Therefore, protest zero WTP bids are excluded from the analysis. This requires stating the maintained assumption that the bias introduced by excluding the protest votes is the same across the two samples and thus does not confound the comparison. The validity of this maintained assumption is partially tested by comparing protest votes across the two samples (see Section 7.2.1). The most appropriate econometric models are estimated using both the stated marginal WTP bids per household and stated marginal WTP bids divided by the number of HM to control for potentially confounding effects. Confidence intervals are calculated to test for statistically significant differences in parameter estimates across the two samples.

Testing Hypothesis 1.3 involves a comparison of the magnitudes of those explanatory variables that exhibit a significant underlying relationship with marginal WTP estimates for the preservation of Nadgee Nature Reserve. This comparison is conducted by using the reference and the replicated sample as well as the census data from 1976 and 2006. Mann-Whitney tests of equality of distributions (Mann and Whitney, 1947), t-test for equality of means, pr-tests of equal proportions, and Pearson's chi-squared tests (independence of observations, goodness of fit) are used to test for statistically significant differences. These tests disclose both whether the magnitudes of those

explanatory variables are different across the two samples and whether they are representative of the populations.

As discussed in Section 2.4, attitudes towards the payment vehicle may influence marginal WTP. As outlined in Sections 6.3 and 8.2.2, the surveys asked respondents to choose between a voluntary donation, a compulsory tax, or a mixture of both. Hence, exploring if the attitudes towards the payment vehicle have changed over a 30-year time period gives rise to the following hypothesis:

Hypothesis 1.4:

H_0 : The attitudes of respondents towards the payment vehicle presented in an open-ended CV survey associated with the preservation value of Nadgee Nature Reserve are the same in 1979 and 2009.

This hypothesis is tested by examining statistically significant differences in the distribution of the preferred payment vehicle using Pearson's chi-squared test (independence of observations). A statistically significant difference may indicate that a change in marginal WTP for the preservation of Nadgee Nature Reserve across time may be partially caused by a change in the attitudes towards the payment vehicle. However, whether a change in attitudes towards the payment vehicle affected the marginal WTP is a non-testable hypothesis, since respondents who were not willing to pay were not asked to choose a payment vehicle.

7.2.3 Explaining marginal willingness to pay – supply side

Possible impacts of changes in variables of supply on stated marginal WTP for the preservation of Nadgee Nature Reserve are investigated by exploring changes in the availability of close substitutes and associated changes in the relative scarcity of preserved natural areas of land (see Section 3.3.2). Assuming that changes in the availability of close substitutes directly affects the relative scarcity of preserved natural areas of land implies that an increase in supply decreases the preservation value of Nadgee Nature Reserve across time, and vice versa. Furthermore, to avoid confounding effects, the same changes in qualitative and quantitative characteristics of Nadgee Nature Reserve were proposed in the two studies (see 3.3.1). The available reference

data for rigorously analysing the potential impact of potentially varying reference points (see Section 3.3.3) is insufficient and assumed to be negligible.

This leads to the following hypothesis:

Hypothesis 1.5:

H_0 : The availability of close substitutes for and associated relative scarcity of the preservation of Nadgee Nature Reserve are the same in 1979 and 2009.

Testing this hypothesis requires an assumption of how respondents define close substitutes for the preservation of Nadgee Nature Reserve. A wider definition may incorporate protected and unprotected natural areas of land in a similar condition, whereas a more stringent definition may restrict such substitutes to protected natural areas of land only. Of course, how respondents defined such substitutes is unknown to the analyst. In consequence, rigorously testing this hypothesis is impossible. Nevertheless, the hypothesis is examined by comparing the number and extent of protected areas in New South Wales in 1979 and 2009 assuming that for the respondents they are (nearly) perfect substitutes for Nadgee Nature Reserve. However, the interpretation of the results obtained by this analysis will be discussed from a broader perspective acknowledging alternative definitions.

7.3 Valuation variability across time

7.3.1 Discrete choice experiments as an alternative to contingent valuation

As outlined in Section 4.3, the DCE method is being increasingly used as an alternative to the CV method. This provides the opportunity to explore the impact of methodological evolution across a 30-year time period using the development of the DCE method and associated econometric models as an example. The impact of methodological evolution over 30 years is investigated by testing the following adjusted hypothesis:

Hypothesis 2.1:

H_0 : Equivalent surplus of the preservation of Nadgee Nature Reserve in 2009 obtained by an open-ended CV is not different from that elicited by a state-of-the-art DCE.

This hypothesis is tested by comparing equivalent surplus for the preservation of Nadgee Nature Reserve obtained by the replicated open-ended CV study with equivalent surplus estimated through a state-of-the-art DCE using a repeated-triple elicitation format (RT). In both samples, protest votes are excluded from the analysis. The open-ended CV equivalent surplus is calculated as the mean value of the stated WTP bids obtained in the replicated study. As outlined in Section 6.3, respondents were told that if development were to be allowed the pristine character of this natural area of land would be lost. Hence, to ensure the preservation of this area, respondents were told that funds would have to be raised to buy the land. This implies that the equivalent surplus is represented by a 100% change from a pristine to a developed area of land.

The DCE-RT equivalent surplus is estimated by applying a pRPL model in WTP space that allows free correlations across random parameters (see Section 4.3.3). The underlying experimental design used to collect the DCE-RT data (see Section 8.3) contained a total of 16 choice sets. Each choice set consisted of three choice options: two non-zero choice options that varied across choice sets and one zero cost choice option that was available in each choice set. The zero cost option (status quo) was held constant across all choice sets and respondents. The 16 choice sets were divided into four blocks of four choice sets per respondent. The internet survey mechanism (see Section 8.3.1) was used to randomise the four choice set blocks across respondents. The order in which choice sets in each block were presented to respondents was not randomised for the DCE-RT sample⁴⁸. This may have caused order effects potentially influencing the magnitude of the calculated equivalent surplus of this split sample (see Section 10.2). The equivalent surplus is calculated based on the difference in utility functions assuming two alternative scenarios with respect to percentage area of land in a pristine condition: 100% and 0%. For both utility functions the access attribute is set to ‘no access allowed’.

⁴⁸ This was not intended but resulted from a miscommunication with the survey research company.

Confidence intervals are calculated using the delta method to test if the equivalent surpluses obtained by the two different methods are statistically significantly different from each other. Additionally, the characteristics of the open-ended CV sample are compared with those of the DCE-RT sample to analyse potentially confounding sampling effects. A Pearson's chi-squared test (independence of observations) is used to test for statistically significant differences between the two samples.

Testing for causes explaining a potential difference in equivalent surplus across the two methods, however, is beyond the capacity of this comparison. Potential drivers are too manifold as to be controllable in this research design. Potential drivers include differences in sampling, questionnaire design, survey mode, definition of the environmental goods and services, econometric analysis, and elicitation format (see Sections 8.2 and 8.3). The aim of this comparison is therefore rather to raise awareness that using outdated stated preference methods for current policy decisions may not be advisable. Nevertheless, potential impacts introduced due to the elicitation format will be further analysed in Section 7.3.2 (effects of prominence in response in open-ended CV) and Section 7.3.3 (effects of incentive compatibility properties in DCE).

7.3.2 Prominence in response in open-ended contingent valuation

The theory of prominence suggests that respondents may use prominent numbers reflecting interval rather than point values when stating their marginal WTP in an open-ended CV (see Section 4.2.2). The use of such heuristics may affect the temporal variability of open-ended CV value estimates. This implies that a change in marginal WTP that remains within an interval of a prominent number may not be detected by the open-ended CV elicitation format. In consequence, respondents' use of prominent numbers that are temporally stable may result in marginal WTP estimates that are seemingly stable across extended time periods, even though they are, in fact, variable. The impact of the open-ended elicitation format used in CV applications on the variability of marginal WTP estimates of non-use values attached to environmental goods and services is explored by testing the following adjusted hypothesis:

Hypothesis 2.2:

H_0 : Prominence in response does not influence possible temporal variability of marginal WTP estimates for the preservation of Nadgee Nature Reserve elicited through an open-ended CV in 1979 and 2009.

As a first step, this hypothesis is tested by exploring if the marginal WTP distributions of both the reference and the replicated study contain prominent numbers and if these prominent numbers are the same. It is assumed that a stabilising effect of prominent numbers increases with an increasing proportion of prominent numbers in the marginal WTP distribution. This is followed by a second step, which investigates if the proportion of prominent numbers is stable across a 30-year time period. For this purpose the proportions of prominent numbers found in each sample are compared. A test on equal proportions (pr-test) is used to test for statistically significant differences between the two samples. A statistically significant difference between the two samples would indicate that the prominence in response in 1979 is different from that in 2009. A third step analyses if the distribution of prominent numbers is stable across the 30-year time period. For this purpose, the distribution of prominent numbers found in each of the two samples is compared. This comparison is extended by adjusting the prominent numbers found in the reference study by the inflation over the last 30 years⁴⁹, regrouping these adjusted numbers into prominent number categories by using the boundaries suggested by the theory of prominence, and comparing their distribution to that of the prominent numbers found in the replicated study. The Mann-Whitney test of equality of distributions (Mann and Whitney, 1947) is used to test for statistically significant differences. A statistically insignificant difference would indicate that the distribution of prominent numbers found in the replicated study might have shifted by the inflation relative to the reference study.

However, even though theory suggests that temporally stable prominence in response may stabilise marginal WTP estimates, the extent of such a stabilising effect remains unknown. This extent may be zero in the case where these marginal WTP estimates actually are stable across long time periods. Hence, it is not possible to differentiate between actual stable marginal WTP estimates and those that are variable but are stabilised by stable prominence in response.

⁴⁹ Consumer Price Index Canberra (1979-2009)

As outlined in Section 4.2.2, a potential stabilising effect introduced by prominent numbers requires cognitive controlled behaviour of survey participants. However, value estimates obtained by the open-ended CV method may be based partially on non-rational behaviour triggered by habits and impulsive emotions. To date, distinguishing effects of prominence in response from those of non-rational behaviour presents an insuperable challenge, limiting the explanatory power of the analysis described in this section. Further research is required which is beyond the scope of this thesis.

7.3.3 Incentive compatibility properties of discrete choice experiments⁵⁰

In order to increase the statistical efficiency of a DCE for a given number of respondents, it is common practice to present each respondent with a sequence of choice questions rather than restricting the choice format to a single referendum. However repeated choice tasks may influence choice outcomes through strategic response, strategic learning, institutional learning and value learning (see Section 4.3.2). The review of the literature suggests that few empirical studies have investigated effects of repeated choice, and in particular, strategic behaviour caused by incentive incompatible DCE elicitation formats. There does not appear to be any work apart from McNair et al. (2011), Bateman et al. (2008), and Day and Prades (2010) that has explored ordering effects and strategic response in sequential DCE empirically. This thesis expands the research on effects of repeated binary choice on choice behaviour in DCE. The following adjusted hypothesis investigates the potential effect of alternative elicitation formats on choice outcomes:

Hypothesis 2.3:

H_0 : DCE outcomes associated with preservation and recreational values attached to Nadgee Nature Reserve obtained by a single-binary elicitation format are the same as those obtained by a repeated-binary elicitation format that contains four choice sets.

⁵⁰ Parts of the literature review presented in Section 4.3.2, parts of the research design presented in Section 7.3.3, parts of the data collection presented in Section 8.3, and parts of the findings presented in Section 10.4 were reported in one article that has been published in *Environmental and Resource Economics* (Scheufele and Bennett, 2012b) and another article that has been revised and re-submitted to *The Australian Journal of Agricultural Economics* (Scheufele and Bennett, 2012a). The candidate conducted all analysis and took the primary role in writing the article, with Bennett providing input at the research design stage and contributions during the writing of the article.

It is hypothesised that the choice between a repeated-binary and a single-binary choice format implies a trade-off between decreased choice accuracy under the single-binary format and increased strategic behaviour under the repeated-binary choice format because of its incentive incompatibility.

A split sample approach based on DCE field surveys⁵¹ (see Section 8.3) using a single-binary elicitation format (with a majority vote) as the baseline is employed to test this hypothesis. In this way, choice outcomes of a single-binary DCE format are compared with those of a repeated-binary DCE format. Choice outcomes are defined as the attribute parameter estimates, scale parameter estimates, and welfare measures (implicit prices). Choice outcomes are estimated through CL models in preference space (see Section 4.3.3). A bootstrapping procedure applying 1000 draws (Krinsky and Robb, 1986) is used to estimate implicit prices. Obtaining only a single choice observation per respondent does not allow the discovery of random parameters that are statistically significantly different from zero (for example, Rose et al., 2009). A possible explanation is that in the absence of a very large sample, it is impossible to disentangle the assumed distribution of random terms associated with preference parameters or alternatives from the assumed EV1 distribution of the remaining random term that is assumed to be IID across alternatives and individuals. This implies that the analysis of SB data, and thus the comparison of the single-binary (SB) and repeated-binary (RB) split samples, is restricted to the CL model specification.

This thesis further explores whether, as suggested by Bateman et al. (2008), awareness of having repeated choices influences choice outcomes. This possibility gives rise to the following adjusted hypothesis:

Hypothesis 2.4:

H_0 : Choice outcomes of a DCE associated with preservation and recreational values attached to Nadgee Nature Reserve obtained by a single-binary elicitation format are the same as those obtained by the first choice question of repeated-binary elicitation formats that contain four choice sets.

This hypothesis is tested by comparing SB choice outcomes with the choice outcomes derived from the first choice questions of an RB elicitation format. Choice outcomes are

⁵¹ The reserve was described in terms of the features of Nadgee Nature Reserve, even though its identity was not revealed. This may have increased the hypothetical nature of the questionnaire and thus may have affected the results.

defined as the attribute parameter estimates, scale parameter estimates and welfare measures (implicit prices). Choice outcomes are estimated through CL models in preference space. A bootstrapping procedure applying 1000 draws (Krinsky and Robb, 1986) is used to estimate implicit prices. As explained above, the single-binary DEC format precludes the use of more advanced models such as panel random parameter logit (pRPL) models and hence the estimation of implicit prices directly in WTP space (see Section 4.3.3).

In contrast to Bateman et al. (2008), who defined the first choice question of a repeated choice task as the incentive compatible baseline, this thesis employs a split sample approach based on field surveys using a single-binary elicitation format with a majority vote. The author of this thesis is unaware of any work other than the research of Racevskis and Lupi (2008) and McNair et al. (2011) that has tested sequence effects focused on the incentive compatibility properties of elicitation formats in DCE using field data and a split sample approach with a single-binary elicitation format as a baseline. The approach of Racevskis and Lupi (2008) is expanded by exploring additional outcome dimensions, testing for choice set awareness, and adjusting the number of observations in the DCE with a single-binary elicitation format to reduce confounding influences. While McNair et al. (2011) focus on a public good with private elements, a pure public good context is investigated. Finally, this study uniquely includes follow-up questions to examine the properties of the potentially incentive compatible baseline.

The effects of repeated choice on choice behaviour are investigated further by exploring whether strategic opportunities provided by the order in which choice sets are presented to respondents affect choice decisions. Using the study of Bateman et al. (2008) as a starting point, choice sets can be categorised by strategy types⁵², which provide incentives to mis-state true preferences - opportunities for strategic response.

- (1) Choice sets presented first in the sequence of choices (*first*) - The first choice set in a sequence necessarily provides both the maximum and the minimum cost level to the respondents at that position in the sequence. Since the first choice set does not provide any strategic opportunities, it is assumed to be incentive compatible.

⁵² The categorisation is based on McNair et al. (2011), and was further developed in discussion with these authors.

- (2) Choice sets presented at any position in the sequence after the first choice set – Choice sets with a cost level that is the minimum presented to the respondent relative to previous offers (*min*).
- (3) Choice sets presented at any position in the sequence after the first choice set – Choice sets with a cost level that is the maximum presented to the respondent relative to previous offers (*max*).
- (4) Choice sets presented at any position in the sequence after the first choice set – Choice sets with a cost level that is neither the minimum nor the maximum presented to the respondent relative to previous offers in the sequence (*none*).

This choice set categorisation facilitates the formulation of the following adjusted hypothesis:

Hypothesis 2.5:

H_0 : Strategic opportunities provided by the order in which choice sets concerned with preservation and recreational values attached to Nadgee Nature Reserve are presented to respondents in a DCE do not affect choice decisions.

To investigate this hypothesis, preferences for preservation and recreational values of a RB data set are estimated using panel random parameter multinomial logit (RPL) models in preference space (see Section 4.3.3). Variables representing the strategy type of the choice set are interacted with the cost variable (*first*cost*, *min*cost*, *max*cost* and *none*cost*) and included into the pRPL model. These interaction variables allow the investigation of potential heterogeneity in the mean of the cost parameter estimate. Positive (negative) parameter estimates indicate that respondents who are presented with a choice set in a particular strategy-type category are less (more) cost-sensitive than on average. For example, a *min*cost* (*max*cost*) parameter estimate that is statistically significantly different from zero and positive (negative) would indicate that respondents who are presented with a cost level that is the minimum (maximum) presented to the respondent relative to previous offers are less (more) cost-sensitive than on average. This implies that positive (negative) parameter estimates indicate that respondents who are presented with a choice set in the *min* (*max*) category have a higher (lower) marginal WTP than on average, such that

$$\text{Marginal } WTP = \frac{\beta^k}{-[\beta^m + \beta^{\text{strategy-type*cost}}]}, \quad [50]$$

where

β^k = *non-cost* attribute parameter estimates

β^m = *cost* attribute parameter estimate

$\beta^{\text{strategy-type*cost}}$ = *strategy-type*cost* parameter estimate.

To allow a comparison across the *strategy-type*cost* parameter estimates, the *first*cost* variable is used as the incentive compatible baseline since the first choice set is assumed to lack strategic opportunities. This approach extends the study of Bateman et al. (2008) by analysing the effects of choice set ordering in sequential DCE relating choices to observed strategy-types. Contrary to McNair et al. (2011), these effects are explored using a pure public good that provides use and non-use values rather than a public good with private elements. Finally, this study adds to the non-parametric approach of Day et al. (2010) by employing a parametric econometric analysis.

To examine whether strategic response coexists with other effects discussed in the literature such as strategic learning (Bateman et al., 2008), value learning (Plott, 1996), and institutional learning (Braga and Starmer, 2005), it is additionally investigated whether choice decisions are stable across a sequence of choice tasks through testing the following adjusted hypothesis:

Hypothesis 2.6:

H_0 : Choice decisions concerned with preservation and recreational values attached to Nadgee Nature Reserve are stable across a sequence of DCE choice tasks.

To test this hypothesis, variables indicating the position of the choice question in the sequence interacted with the cost variable (*position-1*cost*; *position-2*cost*; *position-3*cost*; *position-4*cost*) are included into the pRPL model. These interaction variables allow the investigation of potential heterogeneity in the mean of the cost parameter estimate. The study of Bateman et al. (2008)⁵³ is extended by investigating the relative

⁵³ Bateman et al. (2008) used the logarithm of the position to account for the assumption that having 17 choice sets the effect will decline more rapidly within the first few choice sets. In this study, however, respondents are only given four choice sets.

influence of choice set positions in a sequence by treating them as nominal rather than ordinal variables. A decreasing value of the *position*cost* parameter estimates indicates that respondents' cost sensitivities decline along the sequence of choice sets. This implies a declining marginal WTP while progressing through the choice task, such that

$$wtp = \frac{\beta^k}{-[\beta^m + \beta^{position*cost}]}, \quad [51]$$

where

β^k = *non-cost* attribute parameter estimates

β^m = *cost* attribute parameter estimate

$\beta^{position*cost}$ = *position*cost* parameter estimate.

To investigate this hypothesis further, the RB data set is divided according to choice set position (P1, P2, P3, P4). Acceptance rates of non-zero cost options, bid acceptance curves, magnitudes of the parameter vector, scale factors, and marginal WTP are compared according to the choice set positions.

Hypotheses 2.5 and 2.6 explore strategic response and choice stability separately. Combined models are estimated to test whether these two effects coexist in the data. Since the *first*cost* variable and the *position-1*cost* variable represent the same data points, perfect multicollinearity prevents the inclusion of both variables into one model. An alternative approach is used to avoid this problem: the strategy-types are re-categorised by an effects coded nominal *minimum*cost* variable and an effects coded nominal *maximum*cost* variable. It cannot be ruled out, however, that the result of the combined model may be affected by a potential correlation between the nominal *position-1*cost* variable and the nominal *minimum*cost/ maximum*cost* variables.

All hypothesis testing discussed in Section 7.3.3 is based on a data set that is divided into six split samples. The split samples differed in the number of choice sets per respondent, the choice set order, and the wording of some explanations and instructions so necessitated. It is assumed that the marginal differences in wording of the choice questions across the split samples do not make any statistically significant difference to the choice incentives. The split samples used the same experimental design with a total of 16 choice sets each containing two choice options: one invariant zero cost choice

option that was available in each choice set and one non-zero cost choice option that varied across choice sets (see Sections 8.3.2 and 8.3.3)⁵⁴.

For the SB elicitation format split sample the 16 choice sets were divided into 16 blocks each containing one choice set. That is, each respondent was asked to answer one choice question only. Respondents in the SB split sample were told that they would receive a single choice question. For the RB split sample the 16 choice sets were divided into four blocks of four choice questions. That is, each respondent received four choice questions. The internet survey mechanism (see Section 8.3.1) was used to randomly assign the four blocks to the respondents. As in the SB split sample, respondents were told how many choice questions they would be asked. The order of the choice sets of each block was rotated four times and randomly assigned to respondents to reduce potentially confounding influences of ordering effects. This rotation resulted in four RB split samples each containing all 16 choice sets but in a different order (RB1, RB2, RB3, RB4). This randomisation ensures that the mix of choice tasks across the sample is the same at all positions in the choice sequence. A sixth split sample (FirstRB) was created by pooling the first choice questions of RB1, RB2, RB3, and RB4. The choice set rotation described above guaranteed that each choice set was presented in the first position approximately the same number of times across RB1, RB2, RB3, and RB4. Consequently, the choice questions in FirstRB were the same as those in SB. To avoid the confounding effects of having different numbers of observations across the SB and FirstRB split samples as well as across the SB and the RB1, RB2, RB3, and RB4 split samples, the number of responses collected for the SB split sample (1444) was about four times that for the split samples RB1 (367), RB2 (371), RB3 (369) and RB4 (376)⁵⁵. A Bayesian D-efficient experimental design (Bliemer et al., 2008) was used to generate the choice sets (see Section 8.3.3).

The complete research design is summarised in Table 1.

⁵⁴ A DCE with a binary elicitation format differs from a CV study in that in the former all attribute levels vary across choice sets whereas in the latter variation is limited to the cost variable.

⁵⁵ Rose et al. (2009) used simulated data to investigate the statistical impact of panel data in discrete choice experiments. They showed that increasing the number of choice observations per respondent, while holding sample size constant, results in less biased estimates and larger t-ratios. However, this advantage diminishes with increasing sample size.

Table 1: Research design for hypotheses 2.3 – 2.6

Hypothesis	Test method
2.3	Comparison: 'Repeated-binary' elicitation format (RB1, RB2, RB3, RB4) 'Single-binary' elicitation format (SB)
2.4	Comparison: 'Single-binary' elicitation format (SB) First questions of 'repeated-binary' elicitation format (FirstRB)
2.5	Inclusion of effects coded variables in econometric model representing strategy type (<i>strategy-type*cost</i>) ^a . Coding: First: -1,-1,-1 Min: 1,0,0 Max: 0,1,0 None: 0,0,1
2.6	Inclusion of interaction variables in econometric models representing the position of the choice set (<i>position*cost</i>) ^b . Coding: First: -1,-1,-1 Second: 1,0,0 Third: 0,1,0 Fourth: 0,0,1
2.6	Comparison of bid acceptance rates, bid acceptance curves, parameter vectors, scale factors, marginal WTP across P1, P2, P3, P4
2.5 & 2.6	Inclusion of interactions variables in econometric models representing both the strategy type (<i>minimum*cost</i> ; <i>maximum*cost</i>) and position in the choice set (<i>position*cost</i>) ^b . Coding strategy type: Coding choice set position: Min: 1, -1 Max: 1, -1 First: -1,-1,-1 Second: 1,0,0 Third: 0,1,0 Fourth: 0,0,1

^a One of the four *type*cost* variables (*first*cost*) was omitted from the model estimation. However, it was estimated in separate pRPL models with changed underlying coding.

^b One of the four *position*cost* variables (*position-1*cost*) was omitted from the model estimation. However, it was estimated in a separate pRPL model with changed underlying coding.

7.4 Conclusion

This Chapter has described the adjustments of the hypothesis based on the discussions in Chapter 5 and Chapter 6 and outlined the research design which is divided into two parts: the first part is concerned with investigating temporal value variability, whereas the second part is focused on analysing impacts of valuation variability across time on value estimates.

There are many possible drivers of value estimate variability. Distinguishing their individual impact on marginal WTP across time poses a significant and possibly

insuperable challenge. This thesis should be seen as a first attempt to provide insights into potential value variability across time. In addition, the reference study obtained only a few of those possible drivers on the demand and supply sides. Hence, changes in possibly omitted variables may have influenced value estimates. Furthermore, the tools used to elicit attitudinal variables are - from a current perspective - problematic with respect to producing valid information (see Section 6.3).

Moreover, the reference study determined the method for re-estimating the preservation value of Nadgee Nature Reserve. From a current perspective, the used open-ended CV is regarded problematic in many respects, including its incentive compatibility properties (see Section 4.1) and the potential effect of prominence in response (see Section 4.2.2). Hence, it will be very difficult – if not impossible – to rigorously analyse and explain the reasons for a change in value estimates given the temporal and budget constraints of this study. Irrespective of these limitations, this study is expected to contribute to a better understanding of value variability across extended time periods.

Finally, the investigation of values and valuation across time required stating the following maintained assumptions:

MA 7.1: The potential bias introduced by excluding the protest zero WTP bids from the analysis is the same for the reference sample and the replicated sample and thus does not confound the comparison.

The analysis presented in Chapter 9 is additionally based on the following assumption, which will be relaxed in Chapter 10:

AA 1: Potential bias introduced by the open-ended CV elicitation format did not confound the comparison across the reference and the replicated sample.

In summary, Chapter 7 has described the research design and hypotheses testing. Chapter 8 will describe the data collection process. Together, they provide the basis for analysing value and valuation variability across time as presented in Chapters 9 and 10, respectively.

8 DATA COLLECTION⁵⁶

8.1 Introduction

Chapter 7 discussed the research design and the methods to test the stated hypotheses. As outlined in Chapter 6, the stated hypotheses are grouped under two research questions. The first research question focuses on exploring potential value variability across time by analysing stated marginal WTP and explanatory variables of demand and supply. The methods used to explore the two research questions require a replication of the CV reference study (see Section 6.3) and undertaking a state-of-the-art DCE (see Section 4.3). The second research question focuses on potential impacts of the CV open-ended elicitation format on the stability of value estimates, the potential effects of progress in econometric modelling of open-ended CV data (see Section 4.2.3) on value estimates, the performance of DCE as an alternative to CV, and the influence of the evolution of elicitation formats on stated preference methods (see Sections 4.2.2 and 4.3.2).

This chapter supplies a detailed description of the data collection process of both the CV and DCE. Section 8.2 is concerned with the replication of the CV reference study. It describes the process of data sampling, the questionnaire design and the data coding. A comparison of the reference study with the replicated study requires that the respective surveys are as similar as possible. Otherwise, the comparison may be confounded. However, some changes were unavoidable or regarded as necessary. Since the reference study was described in detail in Section 6.3, this section is restricted to a discussion of the differences between the two studies as well as to a discussion of the reasons for these differences and their potential consequences. Section 8.2 concludes by providing the sample characteristics of both studies and by testing whether the two samples are representative of the population at each point in time. Section 8.3 outlines the process of data sampling, questionnaire design, data coding, and experimental design for all DCE split sample treatments (see Sections 7.3.1 and 7.3.3). It also describes the aims and results of the conducted focus groups and the pilot survey. This is followed by a

⁵⁶ The data collection was conducted under the protocol of the ANU Human Ethics Committee.

presentation of the sample characteristics and a test for sample representativeness. Chapter 8 closes with a summary and conclusion provided in Section 8.4.

8.2 Contingent valuation method

8.2.1 Sampling

Sampling theory suggests a range of alternative sampling strategies, including random sampling and stratified random sampling (for example, Levy and Lemeshow, 2008). A random sample is obtained if every possible sample of a population has the same probability of selection (Levy and Lemeshow, 2008). Obtaining a random sample is crucial for CV studies that use econometric models based on a random sample assumption (Wooldridge, 2010). Departing from the random sample assumption requires the use of econometric models capable of handling non-random samples. An example are sample selection models, which can be used to account for sample selection bias (see Section 4.2).

A random sample is often required to be representative of the population. Sample representativeness is a necessary condition for interpreting model results on a population level. However, random samples that are characterised by a small sample size are often not representative of the population. Stratified random sampling can be used to increase the chance of drawing a small random sample that is representative of the population. A stratified random sample is obtained by dividing the sampling frame into mutually exclusive groups (strata) and then drawing a random sample in each stratum (Levy and Lemeshow, 2008). The chance of obtaining a representative sample can be increased further through applying a target quota for particular sample characteristics.

The sample for the replicated study was obtained by multistage stratified random sampling applying a target quota for age and gender calculated on the basis of the sampling frame⁵⁷. The sampling frame was defined as the population of the Statistical Division of Canberra (SD), where only Australian citizens and permanent residents of

⁵⁷ The target quota is documented and available upon request.

Australia at the age of 18 years and older qualified. The SD Canberra is divided into seven Statistical Subdivisions (SSD) and 116 Statistical Local Areas (SLA) (Pink, 2008). Each SSD contains a certain number of SLAs but the number of SLAs varies across SSDs. The first stage involved the draw of a random sample of SLAs in each SSD. The second stage involved a random draw of starting points (street addresses) per SLA. Each starting point was assigned a certain number of households. The number of randomly selected SLAs per SSD, starting points per SLA, and households per starting point was determined by the number of inhabitants per SSD and SLA. This approach was chosen to account for differing population densities and income levels across SLAs. The interviewers used the last birthday method to randomly select one member of each household. The interviewers occasionally departed from this strict rule (using the second or third last birthday instead) to meet the target quota for age and gender. This approach increased the chance of obtaining a representative sample though it might have impacted on the random character of the sample. This thesis makes the maintained assumption that such a potential impact did not affect the results.

The target sample size of 380 respondents was set by the reference study. To approximate this sample size, 25 SLAs and 58 starting points were randomly drawn. The reference study also predefined the sampling mode. The required face-to-face in-home interviews were carried out by a marketing research company⁵⁸. All interviewers that were involved are Interviewer Quality Control Australia trained and were briefed for this particular survey by the marketing research company. The briefing informed the interviewers about the purpose of the survey and the survey material and familiarised them with the sampling protocol. The interviews were conducted from 27 July to 6 August 2009 from 8:00am – 8:30pm. Including weekends and after work hours increased the chance of obtaining a representative sample of the population of Canberra. The interviewers contacted 2627 individuals, of whom 558 refused to participate and 1682 did not qualify for the survey. Hence, the sampling for the replicated study returned 387 respondents implying a response rate of about 41%. Thirty-five questionnaires with contradictory answers regarding PTP and marginal WTP were listwise deleted, leaving 352 valid questionnaires.

The applied sampling strategy differed from that of the reference study (see Section 6.3). The sampling strategy of the reference study was based on a 1-in-100 systematic

⁵⁸ ekas (<http://www.ekas.com.au>)

random sample of block and section references from a Department of the Australian Capital Territory Rates File. Due to changes in data protection laws in Australia over the last 30 years this file is no longer available to the public.

Additionally, the interview process in the reference study as indicated by Bennett (1984) took longer (about 2 months)⁵⁹ than in the replicated study (11 days). Moreover, the sampling mode was not exactly the same for the two samples. The reference study used pen and pencil personal interviews (PPPI) whereas the replicated study used computer assisted personal interviews (CAPI). The CAPI format was used to limit the financial expenses of the data collection process. This difference in sampling mode, however, was thought to be irrelevant with respect to the survey results. Finally, it remains unknown whether the response rate differs across the two studies since this was not reported in the reference study. This thesis makes the maintained assumption that the described differences in data sampling did not impact on the survey results.

8.2.2 Questionnaire design

A comparison of the reference study with the replicated study requires that the respective surveys are based on the same questionnaire. Otherwise, the comparison may be confounded. This, however, comes at the cost of using a questionnaire design that – from today's perspective – is regarded as problematic. State-of-the-art CV questionnaire designs (for example, Alberini and Kahn, 2006; Carson, 2000) differ from the questionnaire used in the reference study in various aspects (see Section 6.3). Hence, using the same questionnaire implies the maintained assumption that the potential bias introduced by the old design is the same for both surveys and did not confound their comparison. The questionnaire of the replicated study is presented in Appendix B.

Even though using an identical questionnaire would have been the optimal approach some changes were unavoidable or regarded as necessary. Firstly, the reference study formed part of an omnibus survey (see Section 6.3). The limited budget available for this thesis did not allow for the extra expense of conducting an omnibus survey. However, the effect of the difference in survey formats was considered minimal since the major part of the questionnaire of the reference study was positioned at the

⁵⁹ The reference study did not report the exact length of the interview process.

beginning of the omnibus survey. Only some questions about socio-demographics were asked at the end of the survey. Moreover, the reference study was sponsored by a range of institutions related to the different parts of the omnibus questionnaire, whereas the replicated study was only sponsored by one institution⁶⁰. In each study, the respective sponsors were disclosed to the respondents. Furthermore, some questions were thought inappropriate or outmoded from today's perspective. This presents a dilemma. Both keeping and changing these questions might introduce a bias. The decision was to change these questions⁶¹.

Additionally, the reference survey neither specified explicitly whether personal or household income was asked nor did it allow for negative income. Hence, the income question in the repeated survey was divided into a question about personal gross income and household gross income using the Census format (Australian Bureau of Statistics, 2009). Based on implicit indications provided in the reference study, this thesis assumes that the reference study asked for the household income. Finally, in the reference study respondents were not told whether their monetary contribution would be a one-off or a continuous payment. Even though this question was not changed in the replicated questionnaire the interviewers were briefed to respond to potential questions from respondents in a standardised way⁶². Not specifying the payment type might have influenced marginal WTP bids. Respondents might or might not have taken the assumed payment type into account and thus might or might not have adjusted their stated marginal WTP accordingly. In summary, some or all of these differences in the questionnaire design might have affected the results. This thesis states the maintained assumption that this is not the case.

As outlined in Section 6.3, respondents were provided with background information about a preserved area of land located on the south coast of New South Wales. The described features included beaches, sand dunes, a salt water lake, river estuaries, areas of coastal heathland, dry and wet eucalypt forest, and upland heath. This description was supported by photographs showing different landscape types, native plants and

⁶⁰ The survey was funded by the Federal Government through the Environmental Economics Research Hub, based at the Crawford School of Economics and Government at the Australian National University.

⁶¹ Question 1 (relationship to head of household) was replaced by more general questions about household relationships. Question 7 (marital status) was extended by a 'de-facto married' category. In Question 4 (importance of national parks and nature reserves) the word 'fishing' was replaced by 'beach fishing'.

⁶² 'This has been left open to respondents. Please state your WTP as well as a payment type (one-off payment, annual payment, other)'.

native animals. The quality of these photos had deteriorated substantially over the last 30 years. This again presented a dilemma. Both using the qualitatively deteriorated old photographs and using new photographs reflecting changed weather and environmental conditions would have had the potential to influence the results of the survey. A test was conducted to analyse the extent of the potential bias of using new photographs⁶³. The test was carried out within a focus group that consisted of 13 participants (see Section 8.3.4). The test material is presented in Appendix C. Participants were shown two sequences of six pairs of photographs each. The first sequence exclusively included six pairs of new photographs. The second sequence presented six pairs consisting of one new and one old photograph. For each pair in each sequence respondents were asked to choose which of the photographed features they would like to visit and which they would like to preserve. A range of pre-tests were conducted to test for equality in choice proportions between pairs in the first and the second sequence. The differences were found to be statistically insignificant at the 1% level⁶⁴. These results suggest that the exchange of photographs did not have had a statistically significant effect on the survey results. It has to be noted, however, that due to the small number of participants the rigor of these tests is limited.

8.2.3 Data coding

In general, CV studies can produce nominal, ordinal and ratio data. Nominal and ordinal data need to be either dummy or effects coded to be usable for the estimation of non-linear effects in the variable levels through econometric modelling. Dummy coding involves the creation of $J - 1$ new variables that stand for the number of levels of the original variable, with each of these new variables being coded either 0 or 1 (for example, Gujarati, 2005). Hence, dummy coding implies that the base level will be represented by the constant term. Hensher et al. (2005) argue that dummy coding therefore confounds the base level with the mean value of the dependent variable. They recommend the use of effects coding, which avoids this problem by coding the base level as -1 rather than 0.

⁶³ The original photographs were replicated during a field trip to Nadgee Nature Reserve in December 2008. A substantial amount of effort was put towards the identification of the original locations. This was facilitated by the participation of the photographer of the original photographs.

⁶⁴ The results are documented and available upon request.

All non-numerical variables obtained in the replicated study used in the econometric analysis were effects coded with the exception of PTP (see Table 2)⁶⁵. The coding of the explanatory variables obtained in the reference study was adjusted accordingly to allow a comparison of the two data sets. The data of the reference study were preserved as Excel files by the Australian Social Science Data Achieves. As pointed out in Section 6.3, a codebook was available. Nevertheless, the coding remains partially unclear. Hence, potential coding errors might have biased the analysis. To test for such a potential bias OLS linear regression results of the reference study (Bennett, 1984) were compared with results obtained through an analysis of the data preserved in the Excel files using the same model and the assumed original coding. Despite a significant amount of time and effort spent on trials to reconstruct the data coding of the reference study it was not possible to reproduce exactly the parameter estimates (see Appendix D). The difference might be explained by errors introduced through the process of transferring the data from paper into an Excel file format. An alternative explanation is coding errors. Incomplete information about the original coding might have hindered its precise reconstruction. This thesis states the maintained assumption that these potential errors did not influence the results of the analysis.

Table 2: Coding of variables obtained in CV surveys

Variable	Coding
Marginal WTP	A\$
PTP	Yes 1, No 0
<i>Socio-demographics</i>	
Age	Years
Gender	Male 1; Female -1
Household income	AS\$ (midpoints of census range, lower limit for maximum category)
Non-school qualification	High ^a 1; Low ^b -1
Place of birth	Australia 1; Other -1
<i>Attitudinal and behavioural variables</i>	
Importance of conservation	Very important 1, Other -1
Visit of a protected natural area	Yes 1, No -1
Read about nature	Yes 1, No -1
Watch television shows featuring nature	Yes 1, No -1
Course in biology or related science	Yes 1, No -1
Subscriber to an environmental organisation	Yes 1, No -1

^a Defined as bachelor degree, graduate diploma and graduate certificate, postgraduate degree

^b Defined as advanced diploma and diploma, certificate, no non-school education

⁶⁵ The coding of all variables is documented in a data code book, which is available upon request.

Follow-up questions included in the questionnaire facilitated the categorisation of respondents into valid zero marginal WTP bids and protest zero marginal WTP bids. Valid zero marginal WTP bidders are classified as those respondents who stated a zero marginal WTP bid and either faced a binding budget constraint or were not interested in the preservation of the natural area of land. The remaining zero marginal WTP bids were classified as protest votes⁶⁶.

8.2.4 Sample characteristics

The characteristics of the two samples obtained by the reference and replicated study respectively are listed in Tables 3 – 6. A more detailed comparison of the characteristics of the two samples is given in Chapter 10. A series of chi-squared tests (goodness of fit) were conducted to check whether or not the two samples are representative of the population of Canberra. For this purpose the sample frequency distributions of age, gender, household income, and non-school education were compared to the census distribution from 1976 (Australian Bureau of Statistics, 1979a; b) and 2006 (Australian Bureau of Statistics, 2009), respectively. For the reference sample, the categorisation of household income and non-school education were harmonised with the categories used by the 1976 Census. Such a harmonisation was not necessary for the replicated study since the questionnaire used the 2006 Census income classifications.

In the reference sample, no statistically significant differences with respect to gender, age, and place of birth were discovered. However, household gross income and the level of non-school education⁶⁷ were statistically significantly different at the 1% level between the population and the sample. The reference sample was therefore not representative of the households of Canberra in 1976. In the replicated sample⁶⁸ the distributions of household income, age, and non-school education were statistically different at the 1% significance level from those of the population. The characteristics gender and place of birth were not statistically significantly different. The replicated

⁶⁶ Detailed information about the classification of protest zero WTP bids is documented in a data code book, which is available upon request.

⁶⁷ The 1976 census does provide non-school education and household gross income data for individuals 15 years and older, whereas the sample only includes individuals 18 years and older. This may have influenced the results of the comparison.

⁶⁸ The 2006 census does provide non-school education and household gross income data for individuals 15 years and older, whereas the sample only includes individuals 18 years and older. This may have influenced the results of the comparison.

sample was thus also unrepresentative of the population of Canberra in 2009. Hence care should be taken when interpreting the results of both samples at a population level. Additionally, the failure to obtain representative samples in either the reference or the replicated sample might influence the analysis of temporal value variability (see Chapter 9).

Table 3: Sample characteristics in CV surveys: age

Age (years)	Reference sample 1979	Replicated sample 2009
Distribution of age groups (frequency)		
18-19	4.86% (16)	5.41% (19)
20-24	13.98% (46)	6.55% (23)
25-34	31.91% (105)	12.54% (44)
35-44	24.62% (81)	18.52% (65)
45-54	13.07% (43)	13.68% (48)
55-64	6.08% (18)	19.09% (67)
65+	5.47% (18)	24.22% (85)
Sample size	329	351
Pearson's chi-squared test (goodness of fit)	$\chi^2 = 7.44$ df(6); p = 0.282	$\chi^2 = 66.93$ df(6); p <0.001

Table 4: Sample characteristics in CV surveys: household income

Household income (\$)		Reference sample 1979	Replicated sample 2009
Distribution of age groups (frequency)			
1979	2009		
Nil income	Nil income	1.93% (6)	8.19% (28)
1-1,499	1-7,799	0.32% (2)	1.46% (5)
1,500-1,999	7,800-12,999	0.64% (2)	1.46% (5)
2,000-2,999	13,000-18,199	0.96% (3)	3.22% (11)
3,000-3,999	18,200-25,999	0.96% (3)	5.26% (4)
4,000-4,999	26,000-33,799	1.29% (4)	4.97% (17)
5,000-5,999	33,800-41,599	1.61% (5)	4.39% (15)
6,000-6,999	41,600-51,999	0.64% (2)	7.89% (27)
7,000-7,999	52,000-62,399	1.29% (4)	5.26% (18)
8,000-8,999	62,400-72,799	0.96% (3)	6.73% (23)
9,000-11,999	72,800-88,399	9.00% (28)	7.31% (25)
12,000-14,999	88,400-103,999	11.58% (36)	9.65% (33)
15,000-17,999	104,000-129,999	9.65% (30)	11.11% (38)
>18000	130,000-155,999	59.61% (184)	8.19% (28)
	156,000-181,999		6.73% (23)
	182,000-207,999		2.63% (9)
	208,000 and more		5.56% (19)
Sample size		311	342
Pearson's chi-squared test (goodness of fit)		$\chi^2 = 114.86$ df(13); $p < 0.001$	$\chi^2 = 326.93$ df(16); $p < 0.001$

Table 5: Sample characteristics in CV surveys: gender and place of birth

Gender	Reference sample 1979	Replicated sample 2009
Male	45.59%	44.32% (156)
Female	54.41%	55.68% (196)
Sample size	329	352
Chi-squared test (goodness of fit)	$\chi^2 = 2.33$ df(1); $p = 0.127$	$\chi^2 = 2.71$ df(1); $p = 0.100$
Place of birth	Reference sample 1979	Replicated sample 2009
Australia	66.87% (220)	73.01% (257)
Other	33.13% (109)	26.99% (95)
Sample size	329	352
Pearson's chi-squared test (goodness of fit)	$\chi^2 = 0.70$ df(1); $p = 0.403$	$\chi^2 = 2.65$ df(1); $p = 0.104$

Table 6: Sample characteristics in CV surveys: non-school education

Non-school education (frequency)	Reference sample 1979	Replicated sample 2009
Postgraduate degree	2.13% (7)	11.71% (41)
Graduate diploma, graduate certificate	1.82% (6)	2.29% (8)
Bachelor	9.73% (32)	18.29% (64)
Certificate, advanced diploma, diploma	36.78% (121)	31.43% (110)
No non-school education	49.54% (163)	36.29% (127)
Sample size	329	350
Pearson's chi-squared test (goodness of fit)	$\chi^2 = 26.50$ df(4) $p < 0.001$	$\chi^2 = 24.48$ df(4); $p < 0.001$

8.3 Choice experiment method⁶⁹

8.3.1 Sampling

The split sample data sets of the pilot survey (see Section 8.3.5) and the main survey were derived from a random sample of the populations of Sydney and Canberra drawn from an internet panel. The internet panel was provided by a marketing research company⁷⁰. Even though subscribing to this internet panel is open to any volunteer, some individuals may be more inclined to sign up than others. In other words, using an internet panel may lead to sample selection. However, any sampling mode available to collect survey data has its own potential for sample selection. For example, an internet panel assumes people have a computer and a (probably fast) internet connection, and the time/ interest to participate in possibly many surveys. Consequently, one sampling mode is not superior to any other but may just produce a different sample selection bias.

The data were collected using an internet based survey. The sample frames were defined as the SD Sydney and the SD Canberra, with only Australian citizens or permanent residents of Australia 18 years and older qualified to participate. To increase the chance of selecting a representative sample of the populations of Sydney and Canberra, target

⁶⁹ Parts of the literature review presented in Section 4.3.2, parts of the research design presented in Section 7.3.3, parts of the data collection presented in Section 8.3, and parts of the findings presented in Section 10.4 were reported in one article that has been published in *Environmental and Resource Economics* (Scheufele and Bennett, 2012b) and another article that has been revised and re-submitted to *The Australian Journal of Agricultural Economics* (Scheufele and Bennett, 2012a). The candidate conducted all analysis and took the primary role in writing the article, with Bennett providing input at the research design stage and contributions during the writing of the article.

⁷⁰ Pureprofile (<http://pureprofile.com/au/en/home/welcome>)

quotas for age and gender based on the respective sample frames were applied. Screening questions at the beginning of the questionnaire facilitated the selection of qualifying respondents and supported the attainment of the target quotas.

The main survey was online from 20 July – 27 August 2009. Email invitations for the main survey were sent to 13,700 potential participants listed in the internet panel. Each of the respondents who participated in the survey collected points for answering questions, which could be cashed in through the internet panel provider to compensate for their interaction costs. 6595 invited panel subscribers refused to participate, 1618 participated but did not fulfil the screening criteria of being Australian citizens or permanent residents of Australia 18 years and older, 1092 participated but dropped out before completion, and 255 completed in less than five minutes. The time required to complete this questionnaire while providing information about preferences was found to be at least five minutes. Hence, questionnaires provided by participants who completed the survey in less than five minutes were regarded as invalid and were excluded. The sampling returned 4140 complete and valid questionnaires, which implies a response rate of about 34%. The distribution of completed questionnaires across all used split samples is as follows⁷¹: RB1: 367, RB2: 371, RB3: 369, RB4: 376, SB: 1444, RT: 347.

8.3.2 Questionnaire design and data coding

Developing the DCE questionnaire involved balancing between producing a state-of-the-art design (for example, Bennett and Birol, 2010; Birol and Koundouri, 2008) and describing the same environmental goods and services as presented in the CV questionnaire. In the reference CV questionnaire, the preserved area of land was described in terms of the features of Nadgee Nature Reserve without revealing its identity. This concealment was an attempt to isolate existence benefits from all other benefits provided by Nadgee Nature Reserve and maintain the credibility of the survey. Hence, to allow a comparison of CV and DCE results, the identity of the natural area of land could not be revealed in the DCE questionnaire. From today's perspective, however, failing to describe the environmental goods and services in detail may lead to insensitivities to scope or may produce 'warm glow effects' rather than information

⁷¹ The full data set included two more split samples with an additional sample size of 500 respondents from Sydney and Canberra, which were not used in this thesis.

about individual preferences (Chilton and Hutchinson, 2000). However, this thesis states the maintained assumption that respondents did express their preferences for particular environmental goods and services as described in the CV and DCE questionnaires.

A draft questionnaire was developed based on the literature review and expert consultations. The literature review included scientific publications, DCE studies, and planning documents about the Nadgee Nature Reserve (for example, Bennett and Birol, 2010; Birol and Koundouri, 2008; Brock and Xepapadeas, 2003; Christie et al., 2006; New South Wales National Parks and Wildlife Service, 2003). Expert consultations were undertaken with staff from Nadgee Nature Reserve park office in Merimbula in December 2008. Focus groups were used intensively to adjust the draft version of the questionnaire (see Section 8.3.4)⁷². A pilot survey (see Section 8.3.5) was conducted to test the adjusted questionnaire and the internet set-up. The pilot survey was also used to obtain parameter priors for the development of the experimental design (see Section 8.3.3) of the main survey questionnaire.

The final version of the survey was structured as follows. In the first part, respondents were asked some questions about their socio-demographic characteristics (age, gender, residence status, postcode), which facilitated the screening process required to restrict the survey to qualifying participants and to meet the target quotas. This was followed by questions about respondents' general experience of visiting protected areas in Australia and internationally. In the second part, respondents were provided with background information. The information consisted of explanations about Nadgee Nature Reserve, the way it is currently managed and proposed future management options. This verbal information was supported by photographs of the described natural features, which were the same as those used in the CV studies⁷³. In contrast to the CV studies, an incentive compatible payment vehicle was used. Respondents were told that money for the management of this area of land, including purchase and protection options, would have to be collected through additional government taxes and charges. Furthermore, a budget constraint reminder and a provision rule (majority vote) were included⁷⁴. The third part

⁷² Based on the results of the focus groups the questionnaire was adjusted as follows: inclusion of an explanation of the organisation of the questionnaire; clarification of background information.

⁷³ The replicated CV survey used print outs, whereas the DCE survey displayed the photographs on the display.

⁷⁴ 'The management option that receives the greatest support would be implemented and everyone would have to make the payment associated with that management option.'

of the survey asked respondents to make trade-offs between future management options including development and preservation alternatives. The management options were described by three attributes with five, four, and two levels, respectively (see Table 7). In general, respondents did not have the opportunity to click backwards. However, to increase the understanding of the choice task, respondents were presented with a verbal explanation of the outcome of their first choice and given the opportunity to revise it. Respondents could not revisit or revise their first choice once they proceeded to the second choice set^{75,76}. Both, the first and the revised choice were recorded. This was followed by questions designed to check the consequentiality, plausibility, credibility, and comprehensibility of the survey material. The final part of the survey asked additional socio-economic and attitudinal questions.

Table 7: Attributes and their levels in DCE survey

Attribute	Attribute level
Cost in \$A	A\$0 A\$50 A\$100 A\$200 A\$300
Area of land	30% (4,200ha) 50% (7,000 ha) 70% (9,800 ha) 100% (14,000 ha)
Access for minimum impact recreation	Yes No

Split sample treatments were used that differed only in the number of choice sets per respondent, the number of choice options per choice set, the choice set order, and the wording of some explanations and instructions as necessary. This study is based on the maintained assumption that the marginal differences in wording of the choice questions do not alter statistically significantly choice incentives across split samples. All split samples were based on two experimental designs (see Section 8.3.3) with a total of 16 choice sets each, which were divided into choice set blocks according to the split sample treatment (see Section 7.3). The choice sets contained one invariant zero cost choice option (‘do nothing option’) that was available in each choice set and, depending on the split sample treatment, one or two non-zero cost choice options that varied across

⁷⁵ About 7.5 % (RB), 8.3% (SB) of respondents revised their choices: 43% (RB), 43% (SB), changed their choice from a non-zero option to the zero cost option; 57% (RB), and 57% (SB) changed their choice vice versa. The research findings reported in Chapter 10 remain qualitatively unaltered if initial choices are used in the analysis.

⁷⁶ About 13.3% (RT) of respondents revised their choices.

choice sets. The final questionnaire of the RB split sample is presented as an example in Appendix E.

All non-numerical variables in all split sample treatments used in the econometric models were effects coded with the exception of the dummy coded generic constant (see Table 8)⁷⁷. Follow-up questions included in the questionnaire allowed the categorisation of respondents into valid zero marginal WTP bids and protest zero marginal WTP bids. Valid zero marginal WTP bidders are classified as those respondents who stated a zero marginal WTP bid and either faced a binding budget constraint or were not interested in the preservation of the natural area of land. The remaining zero marginal WTP bids were classified as protest votes⁷⁸.

Table 8: Coding of variables obtained in DCE survey

Variable	Coding
Generic constant	Included 1, Not included 0
Cost	A\$0 A\$50 A\$100 A\$200 A\$300
Area of land	30% (4,200ha) 50% (7,000 ha) 70% (9,800 ha) 100% (14,000 ha)
Access for minimum impact recreation	Yes 1, No -1
Age	Years
Gender	Male 1; Female -1
Household income	A\$ (midpoints of census range; lower limit for maximum category)
Non-school qualification	High ^a 1; Low ^b -1

^a Defined as bachelor degree, graduate diploma and graduate certificate, postgraduate degree

^b Defined as advanced diploma and diploma, certificate, no non-school education

8.3.3 Experimental design

Until recently, orthogonal designs were the dominant designs used as the basis for discrete choice experiments. Orthogonal designs are characterised by uncorrelated attribute levels, which is a desired statistical property. However, orthogonal designs

⁷⁷ The coding of all variables is documented in a data code book (see Appendix H).

⁷⁸ Detailed information about the classification of protest zero WTP bids is documented in a data code book, which is available upon request.

require equidistantly spaced attribute levels, are limited to linear effects, and prevent the exclusion of choice sets with dominant alternatives without reducing orthogonality (Rose and Bliemer, 2005). Most DCE do not meet these requirements. DCE commonly include dummy or effects coded binary attributes and attributes that are not equidistantly spaced.

Statistically efficient designs are an alternative to orthogonal designs. The aim of a statistically efficient design is to minimise the standard errors of the parameter estimates (Rose et al., 2008). The asymptotic variance-covariance matrix of a CL model is obtained by calculating the negative of the inverse of the expected Fisher information matrix, with the Fisher information matrix being the second derivative of the log-likelihood function (Train, 2009). As shown by McFadden (1974; 1980), this manipulation leads to the choice variable cancelling out. Thus, the statistical efficiency of a design depends only on the parameter values of the attributes, the design dimensions (such as number of attributes, number of attribute levels, number of choice options, number of choice sets) and the number of respondents (for example, Bliemer et al., 2008). Hence, given the design and number of respondents all that is required to calculate a statistically efficient design is information about the parameter values, whereas prior knowledge about respondents' choices is not needed. Since the true values of the parameters are not known, priors are used in the design process. Such priors are commonly retrieved by a pilot study (see Section 8.3.5).

A commonly used measure⁷⁹ to express the global level of efficiency is the D-error, which minimises the determinant of the variance-covariance matrix (for instance, Rose et al., 2008):

$$D_e = (\det \Omega)^{1/K} \quad [60],$$

where D_e is the D-error, Ω is the variance-covariance matrix, K is the total number of generic parameters to be estimated. The minimisation of the D-error will yield the design with the smallest possible standard errors around the estimated parameters. The smaller the D-error, the more statistically efficient the design is. Therefore, a statistically efficient design can be used to increase efficiency, holding the sample size fixed.

⁷⁹ Other suggested and used measures include the A-error and the C-error (for example, Scarpa and Rose, 2008).

The choice sets for all split samples were created using a Bayesian D-efficient design (Ferrini and Scarpa, 2007; Sándor and Wedel, 2001)⁸⁰ with 100 Halton draws. Bayesian D-efficient designs account for the fact that prior parameter values are only approximately known. That is, they assume that prior parameters are randomly distributed. They were developed based on the calculation of the Db-error of randomly selected designs (10,000 iterations) using the software package Excel. Two Db-efficient designs with 16 choice sets each were created. One design had split samples containing two choice options per choice set and the other had the split sample consisting of three choice options per choice set. Attribute levels were randomly assigned to each attribute in each choice set of the change options while accounting for attribute balance. The ‘do nothing option’ (zero cost option) was held constant but included in the design process. Priors were obtained from pilot studies targeting the populations of Sydney and Canberra (see Section 8.3.5). Following a suggestion of Rose and Bliemer (2005), the rows and columns related to the constant term were excluded from the calculation of the Db-error to avoid the dominance of the disproportionally large standard errors of the constant. Dominant and redundant choice sets were removed through restrictions and swapping of attribute levels. This procedure marginally reduced the Db-efficiency of both designs by 3% and 5%, respectively.

The Bayesian D-efficient designs are developed for CL models without accounting for covariate effects. Estimating different models may alter the design efficiency (Rose and Bliemer, 2005). A software package (Ngene 1.1) can be used to calculate designs for alternative model types and to account for covariate effects. Unfortunately, Ngene 1.1 was not yet available when the experimental design for this thesis was developed.

8.3.4 Focus groups

Focus groups were used to test the draft questionnaire, which was developed based on a literature review and expert consultations. A focus group is a semi-structured discussion with a small group of individuals facilitated by an experienced moderator. Creating an anonymous, tolerant and non-judgmental atmosphere is important for an open and constructive discussion process. Furthermore, the number of participants should be

⁸⁰ Sándor and Wedel originated Bayesian efficient designs (2001), while, to our knowledge, Ferrini and Scarpa (2007) were the first to apply Bayesian efficient designs to environmental economics.

small enough to allow all participants to engage but large enough to ensure heterogeneity of participants and opinions.

Two focus groups of 90 minutes with nine participants each were conducted on 24 March 2009 in Canberra. Participants were recruited through informal networks by sending an invitation per email to friends asking them to forward them to their friends and families who - if interested - stated their interest in participating in the focus group discussion. Participants of the focus groups were paid A\$40 as an incentive.

The focus groups were segmented in five parts. The first part was used to welcome the participants, provide information on logistics, and introduce the overall topic. The second part involved a series of open questions concerned with

- exploring participants' awareness with respect to protected areas,
- assessing participants' knowledge about management actions that could avoid or lessen development impacts,
- investigating participants' attitudes towards low impact recreational activities in protected areas,
- assessing opinions on relevant attributes and their levels, and
- exploring opinions on the payment vehicle and appropriate bounds of marginal WTP for alternative management options.

The draft questionnaire was tested in the third part of the focus group sessions. The draft version of the questionnaire included four choice sets per respondent with three choice options each. The choice sets were based on a random experimental design. Each participant received a printed copy of the draft questionnaire. Participants were asked to fill out the questionnaire. Additionally, the participants were encouraged to write down comments of any sort related to this questionnaire. Afterwards, participants were asked in a discussion to provide specific comments about the following aspects of the questionnaire:

- clarity and comprehensibility of the information,
- credibility of the scenario and the management options,

- possible contradictions or biases in the questionnaire,
- appropriateness of the information type, quantity, and presentation
- sufficiency of quality of the photographs,
- appropriateness of the question format,
- appropriateness of the number of choice sets,
- reactions to a single-binary choice format,
- usefulness of the choice task explanations/ possibility for revision,
- reactions to the fact that the name of the area is not disclosed,
- appropriateness of the attributes and attribute levels.

The fourth part of the focus group was concerned with the testing of photographs (see Section 8.2.2), which was followed by general remarks and a formal closure of the event. The results of the focus groups were used to adjust the draft questionnaire.

8.3.5 Pilot survey

A pilot survey was conducted to test the adjusted draft questionnaire, check the internet set-up, and obtain priors for the experimental designs of the main survey questionnaires.

The pilot survey consisted of three split sample treatments (Pilot1, Pilot2, Pilot3), which only differed in the choice format, the sampled population, and the sample size (see Table 9).

Table 9: Pilot survey split sample treatments in DCE survey

Split sample	Choice format	Sampled population	Sample size
Pilot1	repeated-binary	Sydney	52
Pilot2	single-binary	Sydney	96
Pilot3	repeated-triple	Canberra	48

All choice sets of the pilot study were created using a Bayesian D-efficient design with 10,000 iterations. Priors were obtained from DCE conducted during the focus groups. The same priors were used as best guesses for all split sample treatments in the pilot

survey. The pilot survey was online from 29 June to 2 July. The data was analysed using a simple CL model⁸¹. In Pilot1, the *cost* parameter estimate was statistically significantly different from zero at the 1% level, the *area* parameter estimate at the 10% level and the *access* parameter estimate was not statistically significantly different from zero. All parameter estimates of Pilot2 were statistically insignificantly different from zero. This indicated that the number of observations obtained in the single binary elicitation format might not have been large enough to produce parameter estimates that are statistically significantly different from zero. All of the parameter estimates of Pilot3 were statistically significantly different from zero at the 1% level. The parameter estimates obtained in Pilot1 and Pilot3 were used as priors for the calculation of the experimental designs for the main survey. Even though the *access* parameter estimate obtained in P1 is statistically insignificant it is still the best guess available to be used as a prior.

As outlined previously, both focus groups were conducted in Canberra. To ensure the applicability of the survey material for a sample of the population of Sydney the pilot surveys included four follow-up questions at the end of the questionnaire (for an example see Appendix F). Respondents were asked if they had any concerns, comments or suggestions with any part of the questionnaire. The information obtained was used to adjust the survey material accordingly.

8.3.6 Sample characteristics

The characteristics of split samples RB, SB, and RT are presented in Tables 10 – 13.

⁸¹ The complete model results are documented and available upon request.

Table 10: Sample characteristics in DCE survey: age

Age (years)	RB	SB	RT
Distribution of age groups (frequency)			
18-24	14.09% (209)	14.26% (206)	15.85% (55)
25-34	21.58% (324)	22.77% (329)	23.34% (81)
35-44	21.71% (322)	22.42% (324)	22.48% (78)
45-54	19.22% (285)	19.65% (284)	21.33% (74)
55-64	14.70% (218)	14.81% (214)	12.68% (44)
65+	8.43% (125)	6.09% (88)	4.32% (15)
Sample size	1,483	1,445	347
Pearson's chi-squared test (goodness of fit)	$\chi^2 = 64.35$ df(5); p <0.001	$\chi^2 = 107.03$ df(5); p <0.001	$\chi^2 = 24.17$ df(5); p <0.001

Table 11: Sample characteristics in DCE survey: household income

Household income (\$) ^a	RB	SB	RT
Distribution of age groups (frequency)			
Nil income	0.49% (6)	0.52% (6)	0.33% (1)
1-7,799	0.49% (6)	0.52% (6)	0.00% (0)
7,800-12,999	1.23% (15)	1.12% (13)	0.99% (3)
13,000-18,199	1.98% (24)	1.37% (16)	1.32% (4)
18,200-25,999	2.72% (33)	2.58% (30)	2.97% (9)
26,000-33,799	3.79% (46)	4.29% (50)	2.64% (8)
33,800-41,599	3.87% (47)	3.26% (38)	2.97% (9)
41,600-51,999	6.09% (74)	5.84% (68)	3.63% (11)
52,000-62,399	6.83% (83)	6.95% (81)	4.95% (15)
62,400-72,799	6.67% (81)	7.04% (82)	7.26% (22)
72,800-88,399	9.55% (116)	8.93% (104)	6.93% (21)
88,400-103,999	15.39% (187)	14.25% (166)	13.53% (41)
104,000-129,999	14.81% (180)	15.54% (181)	19.47% (49)
130,000-155,999	10.12% (123)	10.39% (121)	12.54% (38)
156,000-181,999	5.43% (66)	6.95% (81)	11.88% (36)
182,000-207,999	4.03% (49)	3.86% (45)	4.29% (13)
208,000 and more	6.50% (79)	6.61% (77)	4.29% (13)
Sample size	1,215	1,165	303
Pearson's chi-squared test (goodness of fit)	$\chi^2 = 422.03$ df(16); p <0.001	$\chi^2 = 434.42$ df(16); p <0.001	$\chi^2 = 107.72$ df(15); p <0.001 ^a

^a Category 2 (1-7,799) not populated. For goodness of fit test Category 2 was merged with Category 1 (nil income)

Table 12: Sample characteristics in DCE survey: gender

Gender	RB	SB	RT
Male	43.96% (652)	42.49% (614)	45.53% (158)
Female	56.04% (831)	51.51% (831)	54.47% (189)
Sample size	1,483	1,445	347
Chi-squared test (goodness of fit)	$\chi^2 = 12.75$ df(1); p < 0.001	$\chi^2 = 21.59$ df(1); p < 0.001	$\chi^2 = 1.39$ df(1); p = 0.238

Table 13: Sample characteristics in DCE survey: non-school education

Non-school education (frequency)	RB	SB	RT
Postgraduate degree	12.20% (181)	11.83% (171)	12.10% (42)
Graduate diploma, graduate certificate	6.88% (102)	7.68% (111)	9.22% (32)
Bachelor	25.02% (371)	25.33% (366)	21.90% (76)
Certificate, advanced diploma, diploma	41.07% (609)	42.49% (614)	42.36% (147)
No non-school education	14.83% (220)	12.66% (183)	14.41% (50)
Sample size	1,483	1,483	347
Pearson's chi-squared test (goodness of fit)	$\chi^2 = 931.92$ df(4); p < 0.001	$\chi^2 = 1045.34$ df(4); p < 0.001	$\chi^2 = 24.48$ df(4); p < 0.001

A series of chi-squared tests was conducted to test for equivalence between the population statistics using the 2006 census data (Australian Bureau of Statistics, 2009) and the split samples⁸². Despite the effort expended in the sampling procedure, all tested sample characteristics except gender in RT (p = 0.238) were statistically significantly different between the census data and the sample data at the 1% level. One reason for this is that the target quotas for gender and age (65 years and older) could not be met due to an insufficient number of respondents. Another reason may be sample selection introduced by the internet sampling mode, with effects on age, gender, income, and education. The samples are therefore not representative of the households of Sydney and Canberra. Hence, care should be taken when interpreting the results on a population level.

Follow-up questions were included in the questionnaires to check for consequentiality of the survey format, plausibility of the choice questions, credibility of the policy

⁸² The 2006 census does provide non-school education and household gross income data for individuals 15 years and older, whereas the sample only includes individuals 18 years and older. This may have influenced the results of the comparison.

scenario, and comprehensibility of the choice task⁸³. The results presented in Table 14 are based on adding the percentages of the categories ‘strongly agree’ and ‘agree’ chosen by respondents. On average, 68% of the Sydney respondents and 69% of the Canberra respondents were interested in the management of the natural area of land. The provided information was understandable for 74% and 79% of respondents. Seventy-seven percent and 84% of respondents understood the concept of making choices but 16% of respondents in both locations found making choices confusing. Thirty-eight percent and 36% of respondents did not believe that recreation – even if it is low impact - would cause only minor environmental changes. The management options made sense for 50% and 54%. Thirty-eight percent and 36% thought their choices would have an impact and 26% and 23% of respondents believed that the management plan would be implemented. These results indicate that the survey format may lack to some extent consequentiality and other associated properties that influence the incentive properties of the surveys.

Table 14: Consequentiality of DCE survey format

	RB (Sydney)	SB (Sydney)	RB&SB (Sydney)	TR (Canberra)
‘I am interested in the future management of the area of land.’	68%	69%	68%	69%
‘I understood all the information.’	74%	75%	74%	79%
‘I understood the concept of making choices between alternative management options.’	77%	79%	77%	84%
‘I found making choices between alternative management options confusing.’	17%	15%	16%	16%
‘I do not believe that recreation – even if it is low impact – would cause only minor environmental changes.’	37%	40%	38%	36%
‘I found that the available management options made sense.’	51%	49%	51%	54%
‘I believe that my choices will have an impact on how the area of land will be managed in the future.’	39%	40%	39%	36%
‘I believe that the chosen plan for management will be implemented.’	27%	25%	27%	23%

⁸³ All questions were based on a five-point Likert scale: ‘strongly disagree’, ‘disagree’, ‘neither disagree nor agree’, ‘agree’, ‘strongly agree’.

8.4 Conclusion

Chapter 8 has described the data collection process for both the CV and DEC. Section 8.2 described the differences between the reference study and the replicated study with respect to the process of data sampling, questionnaire design and data coding. Necessary changes in the data generation process compared to the original one (reference study) were detailed. Both reasons for and consequences of these differences with respect to the comparability of the two studies were evaluated.

Section 8.3 outlined the process of data sampling, questionnaire design, data coding, and experimental design for all DCE split sample treatments (see Section 7.2) and described the aims and results from the conducted focus groups and the pilot survey. This section stressed the challenge of dealing with the reference study not disclosing the location of the natural area of land. Failing to describe its location might have led to insensitivities to scope or might have produced ‘warm glow effects’ rather than information about individual preferences (Chilton and Hutchinson, 2000). Hence, keeping the location undisclosed necessitated the maintained assumption that respondents did express their preferences for Nadgee Nature Reserve as described in the CV and not natural areas of land in general. Section 8.3 further presented the results of the consequentiality analysis. The results suggest that the DCE survey format may somewhat lack consequentiality and other associated properties that influence the incentive properties of the surveys. This needs to be taken into account in the discussion in Chapters 9 and 10, which present the main analysis of this thesis.

Stating the following maintained assumptions, this chapter has highlighted the challenges and limits of such a comparison:

MA 8.1: Differences in sampling procedure between the reference and the replicated study did not affect the survey results.

MA 8.2: Differences in the questionnaire design between the reference and the replicated study did not affect the survey results.

MA 8.3: The potential bias introduced by the reference questionnaire design is the same for both the reference survey and the replicated survey, and thus did not confound their comparison.

MA 8.4: Bias introduced by potential data entry and coding errors of the reference data did not confound its comparison with the replicated data.

MA 8.5: The marginal differences in wording of the choice questions did not significantly alter choice incentives across the SB, RBs, and FirstRB split samples.

MA 8.6: Respondents disclosed their true preferences for preserving Nadgee Nature Reserve as described in the survey questionnaires and not for the preservation of natural areas in general.

9 ANALYSING VALUE VARIABILITY ACROSS TIME

9.1 Introduction

Chapter 7 specified the overall research design and discussed the research methods used to test the stated hypotheses, which were adjusted to the case study as outlined in Chapter 6. This was followed by Chapter 8, which described the data collection process. Data were collected by means of a CV study and a DCE study. A detailed description of the data collection process was given, associated challenges were debated, and maintained assumptions were stated. Hence, Chapters 6 to 8 provided the basis for investigating the two research questions stated in Chapter 5.

Chapter 9 focuses on exploring the first research question: Do non-market values attached to environmental goods and services change across extended time periods? The remainder of Chapter 9 is organised as follows. Section 9.2 examines whether stated marginal WTP for the preservation of Nadgee Nature Reserve obtained by open-ended CV has changed over a 30-year time period (see Section 7.2.1). This is followed by Section 9.3 which analyses variables of demand that may explain its temporal variability (see Section 7.2.2). The effect of supply side variables on the temporal variability on the preservation value of Nadgee Nature Reserve (see Section 7.2.3) is examined in Section 9.4. A conclusion provided in Section 9.5 discusses the results presented in this chapter against the economic theory discussed in Chapters 2 and 3.

The analysis of change in marginal WTP for the preservation of Nadgee Nature Reserve across time is based on a range of maintained assumptions, which were acknowledged in Chapters 5 to 8. The following are relevant for this chapter:

MA 5.1: The applied stated preference method is temporally reliable.

MA 5.2: Changes in the relative prices of composite goods do not influence marginal WTP.

MA 5.3: The effect of varying perceptions of reference points is negligible.

MA 6.1: Respondents were unaware that the size of the area of land presented in the questionnaire in 2009 was held constant at 1979 levels and thus differed from its current size.

MA 7.1: The potential bias introduced by excluding the protest votes from the analysis is the same for the reference sample and the replicated sample and thus does not confound the comparison.

MA 8.1: Differences in sampling procedure between the reference and the replicated study did not affect the survey results.

MA 8.2: Differences in the questionnaire design between the reference and the replicated study did not affect the survey results.

MA 8.3: The potential bias introduced by the reference questionnaire design is the same for both the reference survey and the replicated survey, and did thus not confound their comparison.

MA 8.4: Bias introduced by potential data entry and coding errors of the reference data did not confound its comparison with the replicated data.

MA 8.6: Respondents did express their preferences for the preservation of Nadgee Nature Reserve as described in the survey questionnaires and not for the preservation of natural areas in general.

The analysis in this chapter requires the following additional assumption (AA):

AA 1: Potential bias introduced by the open-ended CV elicitation format did not confound the comparison across the reference and the replicated sample.

The potential stabilising influence of the open-ended CV elicitation format on temporal variability of marginal WTP estimates will be investigated in Chapter 10.

9.2 Stated marginal willingness to pay

A range of comparisons across the reference sample and the replicated sample were conducted to investigate whether stated marginal WTP for the preservation of Nadgee Nature Reserve obtained by an open-ended CV in 1979 is the same as stated marginal

WTP obtained in 2009 (**Hypothesis 1.1**). A comparison of PTP was conducted using a pr-test of equal proportions. Protest zero marginal WTP bids were excluded from this comparison. As shown in Table 15, the p-value of the pr-test and the overlapping confidence intervals suggest that the proportions of PTP are not statistically significantly different across the two samples. Potential differences were further analysed by comparing the stated marginal WTP (including and excluding valid zero marginal WTP bids) across the two samples using the t-test for equality of means, the non-parametric K-sample test on the equality of medians, the Mann-Whitney test of equality of distributions (Mann and Whitney, 1947), and a test of quality of standard deviations robust to non-normality (Brown and Forsythe, 1974; Levene, 1960). The results are presented in Table 16. The marginal WTP bids obtained in 1979 were adjusted for inflation⁸⁴. Again, protest zero marginal WTP bids were excluded. The mean and distribution of stated marginal WTP (including and excluding valid zero marginal WTP bids) are not statistically significantly different across the two samples (0.058 and 0.070, respectively). This finding is supported by the overlapping confidence intervals of the mean marginal WTP. However, the medians and the standard deviations of stated marginal WTP obtained in the reference study are statistically different from those elicited in the replicated study at the 1% and 5% level, respectively, with smaller median and larger standard deviations observed in the reference sample.

The comparison of stated marginal WTP across the two samples was expanded by normalising the stated marginal WTP bids to the number of HM. As presented in Table 17, this alternative treatment changed the results. The means and the distributions of stated marginal WTP are still not statistically significantly different across the two samples (including and excluding valid zero marginal WTP bids). However, the standard deviations are just/barely statistically significantly different from zero at the 5% level (for the W_0 test statistic only) with smaller median in the reference sample (including and excluding valid zero marginal WTP bids). Only the K-sample test for medians applied to the sample exclusive of valid zero marginal WTP bids still suggests a statistical difference at the 1% significance level, with larger standard deviations observed in the reference sample.

⁸⁴ Consumer Price Index Canberra (1979-2009)

Table 15: Comparison of PTP

	PTP	
	1979	2009
Sample size	263	272
Yes (1)	231	233
No (0)	32	39
Proportion	0.88	0.86
Standard deviation ^a	0.33	0.35
95% confidence interval	0.84-0.92	0.82-0.90
pr-test	z = -0.7399; p = 0.4593	

^a Standard deviation of binominal distributed variable

Table 16: Comparison of stated marginal WTP

	Stated marginal WTP \$ (including valid zero marginal WTP bids)	
	1979*CPI	2009
Sample size	262	271
Mean	136.71	96.66
Median	38.22	50.00
Standard deviation	283.85	196.45
95% confidence interval ^a	102.18-171.24	73.17-120.16
t-test	t = -1.899 df(531); p = 0.0581	
Robust standard deviations ^b	W_0 = 10.028 df(1, 531); p = 0.002 W_50 = 4.100 df(1, 531); p = 0.043	
Mann-Whitney	z = -0.007; p = 0.9941	
K-sample test	$\chi^2 = 10.079$ df(1); p = 0.002	
	Stated marginal WTP \$ (excluding valid zero marginal WTP bids)	
	1979*CPI	2009
Sample size	230	233
Mean	155.73	112.43
Median	38.22	50.00
Standard deviation	298.07	207.68
95% confidence interval ^a	117.01-194.46	85.62-139.24
t-test	t = -1.816 df(461); p = 0.070	
Robust standard deviations ^b	W_0 = 8.326 df(1, 461); p = 0.004 W_50 = 4.210 df(1, 461); p = 0.041	
Mann-Whitney	z = 0.438; p = 0.661	
K-sample test	$\chi^2 = 7.106$ df(1); p = 0.008	

^a The confidence intervals were calculated by the Delta method

^b W_0 defined as Levene's robust test statistic for the equality of variances; W_50 defined as statistic proposed by Brown and Forsythe that replaces the mean in Levene's formula with median (StataCorp, 2007)

Table 17: Comparison of stated marginal WTP per HM

	Stated marginal WTP/HM \$ (including valid zero marginal WTP bids)	
	1979*CPI	2009
Sample size	262	271
Mean	52.60	42.23
Median	12.10	12.05
Standard deviation	121.99	101.94
95% confidence interval ^a	37.76 - 67.44	30.04 – 54.42
t-test	t = -1.0657 df(531); 0.287	
Robust standard deviations ^b	W_0 = 4.124 df(1, 531); p = 0.043 W_50 = 1.418 df(1, 531); p = 0.234	
Mann-Whitney	z = 0.504; p = 0.614	
K-sample test	$\chi^2 = 0.045$ df(1); p = 0.831	
	Stated marginal WTP/HM \$ (excluding valid zero marginal WTP bids)	
	1979*CPI	2009
Sample size	230	233
Mean	59.91	49.12
Median	16.82	20.00
Standard deviation	128.53	108.41
95% confidence interval ^a	43.21 – 76.61	35.12 – 63.11
t-test	t = -0.977 df(461); p=0.329	
Robust standard deviations ^b	W_0 = 3.832 df(1, 461); p = 0.048 W_50 = 1.443 df(1, 461); p = 0.228	
Mann-Whitney	z = 1.068; p = 0.286	
K-sample test	$\chi^2 = 6.604$ df(1); p = 0.010	

^a The confidence intervals were calculated by the Delta method

^b W_0 defined as Levene's robust test statistic for the equality of variances; W_50 defined as statistic proposed by Brown and Forsythe that replaces the mean in Levene's formula with median (StataCorp, 2007)

The comparisons across stated marginal WTP were supplemented by an analysis of protest zero marginal WTP bids. A pr-test of equal proportions was conducted to explore differences in the proportions of such bids. As illustrated in Table 18, both samples exhibit a large proportion of protest zero marginal WTP bids. The p-values of the pr-test and the overlapping confidence intervals suggest that these proportions are, however, not statistically significantly different across the two samples. A range of t-tests and pr-tests⁸⁵ suggest, however, that for the reference sample the proportion of respondents who are subscribers to a conservation organisation and those who watch

⁸⁵ T-tests and pr-tests were performed for the socio-demographic, attitudinal and behavioural variables as listed in Table 2.

television shows featuring nature are statistically significantly different among protest and non-protest voters ($p=0.028$; $p=0.007$), whereas no statistically significant differences between these variables were found in the replicated sample. The analysis of protest zero marginal WTP bids was extended by investigating respondents' characteristics and motivations. A Pearson's chi-squared test (independence of observations) was conducted to explore differences in the frequencies of the alternative motivations. The results are presented in Table 19.

Table 18: Comparison of protest zero marginal WTP bids

	Protest zero marginal WTP bids (whole sample)		Protest zero marginal WTP bids (zero WTP bids only)	
	1979	2009	1979	2009
Protest (1)	19% (60)	22% (77)	65% (60)	67% (77)
Valid (0)	81% (263)	78% (271)	35% (32)	33% (38)
Proportion	0.19	0.22	0.65	0.67
Standard deviation ^a	0.39	0.42	0.48	0.47
95% confidence interval	0.14-0.23	0.18-0.27	0.55–0.75	0.58–0.76
Pr-test	$z = 1.140$; $p = 0.254$		$z = 0.263$; $p = 0.793$	
Sample size	323	348	92	115

^a Standard deviation of binominal distributed variable

Table 19: Comparison of protest motivations

Protest motivation	Distribution (frequency)	
	1979	2009
Anti-taxes	17% (10)	29% (22)
Anti-doorknocks	2% (1)	0% (0)
Government's responsibility	47% (28)	23% (18)
Need access to area	22% (13)	22% (17)
She'll be right	3% (2)	0% (0)
Need more information	2% (1)	14% (11)
Already give to other donations	0% (0)	3% (5)
Other	0% (0)	6% (2)
No answer	8% (5)	3% (2)
Sample size	60	77
Pearson's chi-squared test (independence of observations)	$\chi^2 = 33.557$ df(11); $p < 0.001$	

The Pearson's square test suggests statistical differences at the 1% significance level in protest motivation across the two samples. Protest respondents in 2009 are more tax averse (29% versus 17%), need more information (14% versus 2%), and assign less

responsibility to the government (23% versus 47%) than protest respondents in 1979⁸⁶. Twenty-two percent of protest respondents at both points in time stated zero marginal WTP because of the restricted access to the area of land. These results indicate that even though the protest motivations might have shifted to some extent the proportion of protest voters is not statistically significantly different across the two samples.

Based on the overall results stated in Section 9.2, the null **Hypothesis 1.1** cannot be rejected.

9.3 Explaining marginal willingness to pay - demand side

The potentially changing impact of explanatory variables of demand on the preservation value of Nadgee Nature Reserve across time was investigated by testing two hypotheses. The first hypothesis examined changes in underlying relationships between marginal WTP estimates and explanatory variables of demand represented by parameters that were estimated using state-of-the-art econometric models (**Hypothesis 1.2**). This required a search for a state-of-the-art model that is appropriate for open-ended CV data. The choice of a state-of-the-art econometric model best suited to analyse the data of both the reference and the replicated sample was facilitated by an investigation of the data characteristics. As illustrated in Figure 5, the distributions of both samples exhibit a non-trivial proportion of corner solutions. Furthermore, both distributions consist of discrete values, of which some might be interval values due to the presence of prominence in response.

These findings suggest that linear and log-linear regression models are not the most appropriate choices. Hence, a censored interval model may achieve an improvement over a linear and log-linear regression model since it is capable of accommodating a non-trivial proportion of zero marginal WTP bids and interval data simultaneously. However, the distributions of the positive marginal WTP bids in both samples are non-normal. A skewness and kurtosis test⁸⁷ suggests that the skewness and kurtosis of the distributions of the marginal WTP bids are statistically different from a normal distribution at the 1% significance level.

⁸⁶ In this thesis it is assumed that ‘anti-taxes’ and ‘government responsibility’ are not substitutes.

⁸⁷ Test procedure suggested by D’Agostino et al. (1990); estimated in Stata using the `sktest` command (StataCorp, 2007). The test results are available upon request.

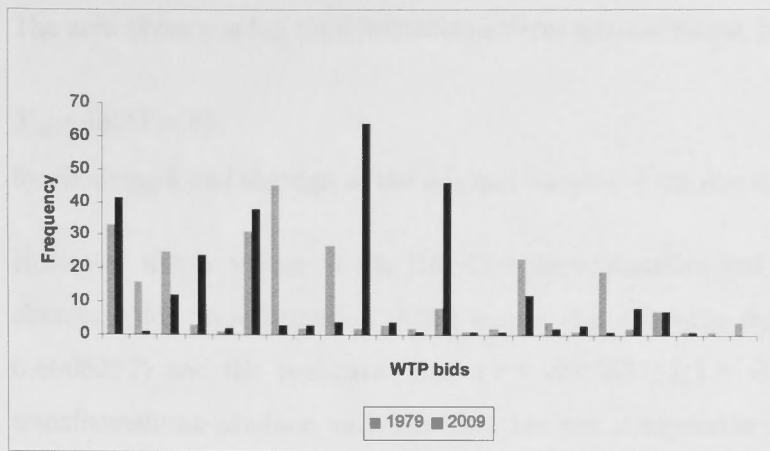


Figure 5: Data characteristics of reference and the replicated sample

Censored interval models strictly assume normally distributed error terms. Since corner solution models are a specific case of censored interval models, a conditional moment test conducted together with a corner solution model estimation provides an indication whether the normality assumption is violated beyond the skewness and kurtosis test as described above. Hence, corner solution models for marginal WTP per household and marginal WTP per HM using both the reference and the replicated data were estimated in STATA 10⁸⁸. Since a corner solution model does not allow a differentiation between valid and protesting zero marginal WTP bids, the latter were excluded. The conditional moment test conducted together with the model estimation suggests that the normality assumption is violated at the 1% significance level in both models and samples. Usually, data transformations are used to overcome this limitation. However, the occurrence of zero marginal WTP bids precludes a simple log transformation. More sophisticated transformations such as a Box-Cox power transformation or a zero skewness log transformation (for example, StataCorp, 2007) can accommodate zero marginal WTP bids.

The Box-Cox power transformation generates a transformed variable $Y_{Box-Cox}$,

$$Y_{Box-Cox} = (Y^r - 1)/r, \quad [52]$$

by choosing r so that the skewness of the original variable Y is zero.

⁸⁸ The selection of the explanatory variables used in these models was based on an analysis of correlations and potentially spurious relationships.

The zero skewness log transformation creates a transformed variable Y_{zsl} ,

$$Y_{zsl} = \ln(\pm Y - k), \quad [53]$$

by choosing k and the sign of the original variable Y so that the skewness of Y is zero.

However, the k values of the Box-Cox transformation and the r values of the zero skewness log transformation differ across the reference data ($r = -0.0449702; k = -0.6006253$) and the replicated data ($r = -0.0563757; k = -3.194117$)⁸⁹. Hence, these transformations produce variables that are not comparable across the two data sets. Consequently, corner solution models, censored interval models, and double hurdle models (based on truncated models) are inappropriate for either data set to analyse value change across time.

Count models and negative binominal models are capable of accommodating non-trivial proportions of corner solutions, non-normal distributions, and distributions that consist of discrete values. However, they require a monotonic distribution to produce a reasonable model fit. Clearly, as illustrated in Figure 5, the marginal WTP bid distributions of the two samples are non-monotonic. An alternative that can accommodate all these characteristics is the use of quantile regression models. Hence, quantile regression models with bootstrapped standard errors for marginal WTP per household and marginal WTP per HM were estimated using STATA 10⁹⁰. However, the unacceptably low pseudo R-squared values indicate that the model fits the data poorly. Only about 0.2%/ 3% (reference data) and 3%/ 5% (replicated data) of the variation in the data could be explained. Hence, it was concluded that the quantile regression model is not appropriate. There is thus no real alternative to a linear and log-linear regression model estimated for the positive marginal WTP bids even though such an approach is associated with problems as discussed in Section 4.2.3.

Log-linear models of the log of the stated positive marginal WTP bids obtained in the reference and the replicated surveys were estimated using STATA 10^{91,92}. The results

⁸⁹ Note that a negative k value requires a division by zero. Hence, in this sample zero WTP bids would have to be excluded from the transformation.

⁹⁰ Protest zero WTP bids were excluded from the analysis. The same explanatory variables as in the corner solution model were used.

⁹¹ A linear regression model was estimated for both samples but exhibited heteroscedasticity (Breusch-Pagan / Cook-Weisberg test for heteroscedasticity). The results are available upon request.

⁹² Protest zero WTP bids were excluded from the analysis.

for the log of the stated positive marginal WTP bids per household are presented in Table 20. The explanatory variables for the model estimation were selected based on a correlation analysis and subsequent investigation of potentially spurious relationships. The occurrence of spurious relationships was explored by examining correlation coefficients conditional on potentially confounding variables. That is, the correlation coefficient between the depended variable and a certain explanatory variable was examined while controlling for other potentially confounding explanatory variables.

For the reference sample, the theoretically suggested variables *household income*, *read about nature*, *place of birth*, and *subscriber to a conservation organisation* exhibited a correlation with the log of the stated marginal WTP bids that is statistically different from zero at least at the 10% significance level. However, the correlation coefficient of the variable *subscriber to conservation organisation* with the stated marginal WTP bids is statistically different from zero only at the 10% significance level, indicating a weak spurious relationship caused by *household income*. The parameter estimate of *subscriber to a conservation organisation* was excluded from the log-linear model since it was not statistically significantly different from zero and its removal did not have a significant impact on the parameter estimates of the other explanatory variables.

For the replicated sample, the theoretically suggested variables *gender*, *household income*, *non-school education*, *importance of conservation*, and *subscriber to a conservation organisation* demonstrated a correlation with the log of the stated marginal WTP bids that is statistically different from zero at least at the 10% significance level. However, the correlation coefficient between *household income* and stated marginal WTP bids is not statistically significantly different from zero if controlled by *gender*. This indicates that the relationship of *household income* and gender may be spurious. Hence, *household income* was excluded from the log-linear regression analysis.

The F-tests and the adjusted R-squared values of 0.085 and 0.157 suggest that both models have some power to explain variation in the log of stated marginal WTP. The ovtest conducted for both models did not discover indications for variable omission and the mean variance inflation factors (VIF) of 1.03 and 1.05 did not detect any serious multicollinearity. The t-tests conducted show that the predicted means of the log marginal WTP are not statistically significantly different from the observed means of the log marginal WTP bids in either sample. This indicates that the models predict the log marginal WTP rather well. In the reference sample, all three explanatory variables

are statistically significantly different from zero and have the expected sign. The results suggest that being born in Australia, reading books about nature and a higher household income have a positive effect on the log of marginal WTP per household. In the replicated sample the variables *gender*, *non-school education*, and *subscriber to conservation organisation* are statistically significantly different from zero and have the expected signs. The results indicate that being male, having a higher non-school education and being a subscriber to a conservation organisation results in higher log marginal WTP per household.

Together these results provide evidence that the underlying relationships between stated marginal WTP and explanatory variables of demand are not the same across the two samples. However, the *hettest* and the *imtest* indicate heteroscedasticity for the reference model. Furthermore, as shown in Figure 6, the residuals of both models are not randomly distributed. The observed pattern may be explained by the discrete nature of the stated marginal WTP bid distribution.

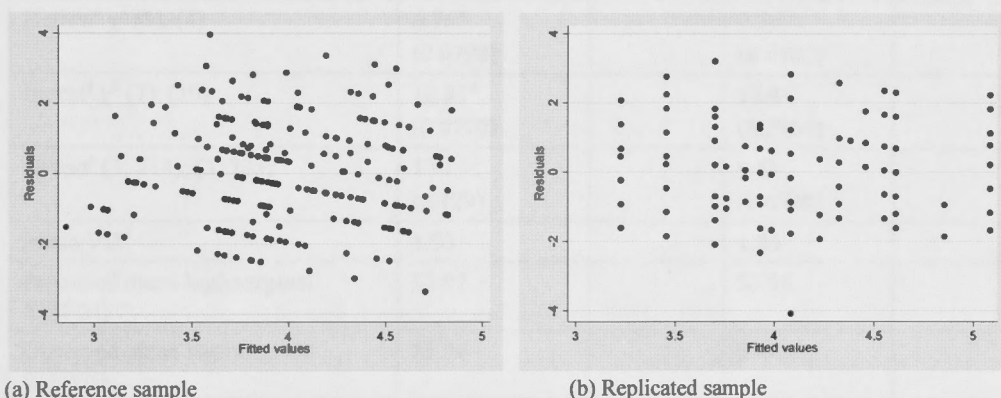


Figure 6: Residuals in log-linear regression model – marginal WTP per household

Table 20: Model results of log-linear regression model – marginal WTP per household

Variable	Reference sample		Replicated sample	
	Coefficient ^a	Standard error ^b	Coefficient ^a	Standard error ^b
Gender			0.23588*** (0.002)	0.075426
Household income ⁹³	5.06-e-06* (0.078)	2.86e-06		
Non-school qualifications			0.31278*** (0.000)	0.083606
Importance of conservation			0.11411 (0.129)	0.074988
Read about nature	0.31425*** (0.002)	0.10003		
Subscriber to a conservation organisation			0.264638** (0.012)	0.104303
Place of birth	0.30127*** (0.003)	0.10198		
Constant	3.43984*** (<0.001)	0.0.25162	4.15545*** (<0.001)	0.107882
F-test (3, 217); (2, 228)	7.36*** (<0.001)		10.82*** (<0.001)	
Adj. R-squared	0.085		0.157	
Hettest ^c χ^2 (3), (4)	6.76* (0.0798)		2.36 (0.6706)	
Imtest ^d χ^2 (7), (10)	12.81* (0.0770)		12.44 (0.2567)	
Ovtest ^e (3, 214); (3, 225)	1.91 (0.129)		0.51 (0.6740)	
Mean VIF ^f	1.03		1.05	
Predicted mean log(marginal WTP/H)	\$3.97		\$3.88	
Observed mean log(marginal WTP/H)	\$3.96		\$3.88	
t-test means (observed – predicted) (229), (232)	-0.0940 (0.925)		-0.0353 (0.9719)	
Number of observations	221		233	

^a ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

^b Variance-covariance matrix robust against robust to some types of mis-specification errors with respect to the characteristics of the postulated distribution for the error terms (StataCorp, 2007)

^c Breusch-Pagan/ Cook-Weisberg test for heteroscedasticity (StataCorp, 2007)

^d White's test for heteroscedasticity (StataCorp, 2007)

^e Ramsey Reset test using powers of the fitted values of y (StataCorp, 2007)

^f Variance inflation factor (indicator for multicollinearity)

⁹³ WTP bids and household income adjusted by the Consumer Price Index Canberra (1979-2009)

The observed heteroscedasticity, the non-randomly distributed residuals and the statistically insignificant household income variable raise concerns about the validity of the results. In an attempt to overcome these issues and to account for potentially confounding effects of differences in household sizes the log-linear regression models were re-estimated controlling the log of marginal WTP by the number of HM.

For the reference sample, the correlation analysis did produce the same explanatory variables of demand with the relationship between the log of stated marginal WTP per HM and the variable *subscriber to conservation organisation* being spurious. However, the correlation analysis produced different results for the replicated sample. Firstly, the correlation coefficient for the variables *age* and *read about nature* became statistically different from zero at the 5% significance level additionally to the variables *gender*, *household income per HM*, *non-school qualification*, *importance of conservation*, and *subscriber to conservation organisation*. Secondly, a spurious relationship between *non-school education* and *read about nature* required the exclusion of the variable *read about nature*. Weak spurious relationships were observed between *age* and *importance of conservation*, *household income per HM* and *subscriber to conservation organisation*, and *non-school education* and *subscriber to conservation organisation*. Hence, only the variables *household income per HM*, *age*, *gender*, *non-school qualifications*, and *subscriber to conservation organisation* were included into the log-linear OLS regression analysis.

The results of the OLS log-linear models are presented in Table 21. The R-squared values improved compared to the previous models. Additionally, the hypothesis of homoscedasticity cannot be rejected and the low mean VIF factors detect no sign of multicollinearity. Furthermore, the patterns previously observed in the distributions of the residuals are no longer present in the reference sample and reduced in the replicated sample (see Figure 7). As with the previous model specification, the predicted means of log marginal WTP per HM and the observed means of log stated marginal WTP bids are not statistically significantly different across the two samples. For the reference sample, the variables *household income per HM*, *read about nature*, and *place of birth* are statistically different from zero at the 1% significance level and have the expected sign. Higher household income per HM, reading books about nature and being born in Australia have a positive impact on the log of marginal WTP per HM. For the replicated sample, the variables *age*, *gender*, *household income per HM*, *non-school education*,

importance of conservation, and *subscriber to conservation organisation* are statistically significantly different from zero and have the expected signs.

These results provide evidence that, with the exception of *household income per HM*, the underlying relationships between the log of stated marginal WTP per HM and explanatory variables are not the same in the reference and the replicated sample. The value of the *household income per HM* parameter estimate is larger in the reference sample compared to the replicated sample. This is supported by the confidence intervals, which do not overlap. This indicates that *household income per HM* has a larger impact on the log of marginal WTP per HM in the reference than in the replicated sample.

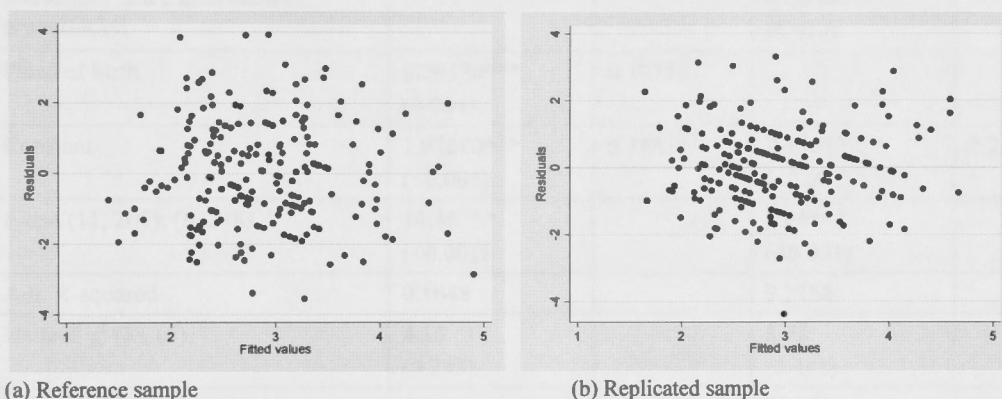


Figure 7: Residuals in the log-linear regression model – marginal WTP per HM

Table 21: Results of log-linear regression model – marginal WTP per HM

Variable	Reference sample		Replicated sample	
	Coefficient ^a	Standard error ^b	Coefficient ^a	Standard error ^b
Age			0.01218*** (0.005)	0.00425
Gender			0.19761** (0.023)	0.08021
Household income/ HM ⁹⁴	2.81e-5*** (<0.001)	6.05e-6	7.97e-6** (0.046)	4.03e-06
95% confidence interval	1.62e-5-4.01e-5		3.64e-8-1.59e-5	
Non-school qualifications			0.28081*** (0.002)	0.09213
Importance of conservation ⁹⁵			0.15422* (0.076)	0.07988
Read about nature	0.32394*** (0.002)	0.10162		
Subscriber to a conservation organisation			0.26064** (0.019)	0.11001
Place of birth	0.29576*** (0.004)	0.10255		
Constant	1.93680*** (<0.001)	0.18874	2.34070*** (<0.001)	0.29896
F-test (11, 206); (6, 218)	14.36*** (<0.001)		9.99*** (<0.001)	
Adj. R-squared	0.1648		0.2158	
Hettest ^c χ^2 (3), (6)	4.15 (0.245)		5.78 (0.448)	
Imtest ^d χ^2 (7), (23)	9.84 (0.198)		23.94 (0.407)	
Ovtest ^e (3, 214); (3, 215)	1.78 (0.151)		2.38* (0.070)	
Mean VIF ^f	1.02		1.10	
Predicted mean log(marginal WTP/HM)	\$2.85		\$2.93	
Observed mean log(marginal WTP/HM)	\$2.82		\$2.93	
t-test means (observed – predicted) (229), (232)	-0.156 (0.876)		-0.0046 (0.996)	
Number of observations	221		225	

^a ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

^b Variance-covariance matrix robust against robust to some types of mis-specification errors with respect to the characteristics of the postulated distribution for the error terms (StataCorp, 2007)

^c Breusch-Pagan/ Cook-Weisberg test for heteroscedasticity; STAT command `hettest` (StataCorp, 2007)

⁹⁴ WTP bids and household income adjusted by the Consumer Price Index Canberra (1979-2009)

⁹⁵ Excluding the variable *importance of conservation* from the model due to concerns about its validity (incentive compatibility of question) did not significantly change the results.

^d White's test for heteroscedasticity; STAT command `hettest` (StataCorp, 2007)

^e Ramsey Reset test using powers of the fitted values of y ; STATA command `ovtest` (StataCorp, 2007)

^f Variance inflation factor (indicator for multicollinearity)

An alternative to excluding valid zero marginal WTP from the analysis is the use of probit and logit models. However, these models are limited to binominal dependent variables. Hence, they can only be used to model PTP rather than stated marginal WTP bids. Logit models using the reference and the replicated data were estimated using STATA 10. The results are presented in Table 22. As with the log-linear regression models, correlation analysis was used to select the explanatory variables to be included in the logit models. For the reference sample, correlations with PTP that are statistically significantly different from zero at the 5 % level were observed for the variables *importance for conservation*, *watch television*, and *place of birth*. Spurious relationships could not be detected. For the replicated sample, correlations with PTP that are statistically significantly different from zero at the 5% level were found for *age*, *household income*, *importance of conservation*, *visit*, *read about nature*, *course*, and *subscriber to conservation organisation*. However, a number of spurious relationships were detected. The variable *course* is correlated to PTP only because of its correlation with all other variables. Hence, it was excluded from the logit model. Furthermore, the correlation of *subscriber to conservation organisations* with PTP is at least partially driven by the variables *read* and *importance of conservation*. Its correlation coefficient with PTP is statistically significant only at the 10% level if controlled by *read about nature* and *importance of conservation*. Finally, the correlation of the variable *read about nature* with PTP is at least partially driven by the variable *visit*. Its correlation coefficient with PTP is statistically significant only at the 10% level if controlled by *visit*. This, in turn, suggests that the correlation of *subscriber to conservation organisation* with PTP is at least partially a result of its correlation with *visit*.

Table 22 lists the variables that were included in the logit models. For both samples, the Wald test and the pseudo R-squared values suggest that the model has the power to explain some variation in the probability of PTP. The pr-test indicates that the observed and predicted probabilities of PTP are not statistically significantly different. Hence, the model predicts the dependent variable rather well. For the reference sample, all three explanatory variables included in the model are statistically significantly different from zero and have the expected signs. Assigning a high importance to conservation, watching television featuring nature and being born in Australia has a positive influence

on PTP. For the replicated sample, only *age*, *household income*, *importance of conservation*, and *read about nature* are statistically significantly different from zero. The only variable that is statistically significantly different from zero in both samples is *importance of conservation*. The largely overlapping confidence intervals suggest that the parameter estimates are not statistically significantly different from each other. As expected, the latter three have a positive impact on PTP. The sign of the *age* parameter estimate is negative suggesting that *age* negatively influences PTP. The results of the logit model indicate that most of the underlying relationships between explanatory variables of demand and PTP are not the same across the two samples.

Replicated sample		Non-replicated sample	
Read about nature	0.0000***	0.0000***	0.0000***
Watch television	0.0000***	0.0000***	0.0000***
Age	-0.0000***	-0.0000***	-0.0000***
Household income	0.0000***	0.0000***	0.0000***
Conservation	0.0000***	0.0000***	0.0000***
Probability of 1 purchase	0.0000***	0.0000***	0.0000***
Probability of 2 purchases	0.0000***	0.0000***	0.0000***
Age squared	0.0000***	0.0000***	0.0000***
Watch TV (h)	0.0000***	0.0000***	0.0000***
Watch TV squared	0.0000***	0.0000***	0.0000***
Age squared squared	0.0000***	0.0000***	0.0000***
Number of purchases	0.0000***	0.0000***	0.0000***
*** Significant at 100%, ** significant at 95%, * significant at 90% level			
* Variable is significant at 10% level of probability, but with additional testing, probability of significance is reduced to 5% or 10% level			

Table 9.1. Logit model results for the replicated and non-replicated samples

Table 9.1. Logit model results for the replicated and non-replicated samples. The table shows the results of the logit model for the replicated and non-replicated samples. The table includes the variable name, the coefficient estimate, the standard error, the z-value, and the p-value. The table is organized into two main sections: the replicated sample and the non-replicated sample. The replicated sample section includes variables such as Read about nature, Watch television, Age, Household income, Conservation, Probability of 1 purchase, Probability of 2 purchases, Age squared, Watch TV (h), Watch TV squared, Age squared squared, and Number of purchases. The non-replicated sample section includes the same variables. The table also includes a note about the significance levels: *** Significant at 100%, ** significant at 95%, * significant at 90% level. A second note states: * Variable is significant at 10% level of probability, but with additional testing, probability of significance is reduced to 5% or 10% level.

Table 22: Results of logit model⁹⁶

Variable	Reference sample		Replicated sample	
	Coefficient ^a	Standard error ^b	Coefficient ^a	Standard error ^b
Age			-0.03029*** (0.007)	0.01118
Household income ⁹⁷			8.44e-06** (0.022)	3.68e-06
Importance of conservation ⁹⁸	0.45740** (0.026)	0.20492	0.65387*** (0.002)	0.20824
95% confidence interval	0.05577-0.85904		0.24572-1.06201	
Visit of protected areas			0.18695 (0.366)	0.20696
Read about nature			0.63227*** (0.005)	0.22622
Watch television featuring nature	0.56938** (0.014)	0.23190		
Place of birth	0.57870*** (0.004)	0.20206		
Constant	1.57217*** (<0.001)	0.22882	2.87166*** (<0.001)	0.77176
Probability of y=1 predicted	0.879		0.855	
Probability of y=1 observed	0.878		0.857	
Pr-test z-value	-0.0435 (0.9653)		0.0518 (0.9587)	
Wald χ^2 (3), (5)	20.71*** (<0.001)		28.20*** (<0.001)	
Pseudo R-squared	0.1131		0.1912	
Log pseudo-likelihood	-84.139162		-87.727064	
Number of observations	259		262	

^a ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

^b Variance-covariance matrix robust against robust to some types of mis-specification errors with respect to the characteristics of the postulated distribution for the error terms (StataCorp, 2007).

Overall, these results justify the rejection of null **Hypothesis 1.2**.

⁹⁶ A probit model was estimated as comparison to the logit model. The conclusions drawn from the results of the probit model are the same as the conclusions drawn from the logit model. Hence the results of the probit model are not displayed but are available upon request.

⁹⁷ WTP bids and household income adjusted by the Consumer Price Index Canberra (1979-2009)

⁹⁸ Excluding the variable *importance of conservation* from the model due to concerns about its validity (incentive compatibility of question) did not significantly change the results.

Changes in the impact of explanatory variables of demand on the stated marginal WTP bids for the preservation of Nadgee Nature Reserve were further investigated by examining changes in the magnitude of explanatory variables that exhibit a statistically significant underlying relationship with the stated marginal WTP bids (**Hypothesis 1.3**). For the same reasons as stated previously, respondents who were classified as having stated a protest zero marginal WTP bid were excluded from this analysis. Additionally, to make this analysis compatible with the log-linear regression analysis, the valid zero marginal WTP bids were excluded⁹⁹.

A t-test on the equality of means, mean value confidence intervals, and a Mann-Whitney test on equality of distributions suggest that both the mean values and the distributions of the *age* variable are statistically different across the two samples at the 1% significance level (see Table 23). On average, respondents in the 2009 sample are older than respondents in the 1979 sample and the age distribution shifted from younger to older respondents.

Differences in the distributions of *non-school education* were investigated using Pearson's chi-squared tests (independence of observations). The p-values suggest that the observed differences are statistically different at the 1% significance level (see Table 24). The sample obtained in 2009 exhibits a higher percentage of respondents with a higher level of *non-school education* than the sample collected in 1979.

Furthermore, the p-values of the t-test and the Mann-Whitney test indicate that the differences in the mean and the distribution of *household income* and *household income per HM* are statistically significantly different from zero (see Table 25).

Contrary, the proportions of *gender* and *place of birth* are not statistically significantly different (see Table 26).

⁹⁹ The same analysis was conducted without excluding the valid zero marginal WTP bids to make it compatible with the logit analysis. The results are not displayed but are available upon request. The conclusions drawn from the results are the same with the exception of the variable *household income*, which was not statistically significantly different across the reference and reference samples.

Table 23: Comparison of age – CV sample 1979:2009

Age (years)	Reference sample 1979	Replicated sample 2009	Difference
Mean	35.91	47.00	+11.9
Standard deviation	13.10	18.85	
t-test	t = 7.416 df(468); p < 0.000		
95% confidence interval	34.23-37.58	44.56-49.83	
Mann-Whitney	z = 6.474; p < 0.001		
Distribution of age groups (frequency)			
18-19	4.64% (11)	7.73% (18)	+3.09%
20-24	15.61% (37)	7.73% (18)	-7.88%
25-34	33.33% (79)	14.59% (34)	-18.74%
35-44	23.63% (56)	17.60% (41)	-6.03%
45-54	12.24% (29)	13.30% (31)	+1.06%
55-64	6.33% (15)	17.60% (41)	+11.27%
65+	4.22% (10)	21.46% (50)	+17.24%
Sample size	237	233	

Table 24: Comparison of non-school education – CV sample 1979:2009

Non-school education (frequency)	Reference sample 1979	Replicated sample 2009	Difference
Postgraduate degree	2.53% (6)	11.97% (28)	+9.44%
Graduate diploma, graduate certificate	1.69% (4)	3.42% (8)	+1.73%
Bachelor	8.86% (21)	19.23% (45)	+10.37%
Certificate, advanced diploma, diploma	38.40% (91)	29.92% (70)	-8.48%
No non-school education	48.42% (115)	35.47% (83)	-12.95%
Sample size	237	234	
Pearson's chi-squared (independence of observations)	$\chi^2 = 39.050$ df(5); p < 0.001		
Non-school education (frequency) high level, low level ^a	Reference sample 1979	Replicated sample 2009	Difference
High level (1)	13.08% (31)	34.19% (80)	+21.11%
Low level (0)	86.92% (206)	65.81% (154)	
Proportion	0.13	0.34	
Standard deviation	0.34	0.48	
95% confidence interval	0.09-0.17	0.28-0.40	
Pr-test	z = 5.397; p < 0.001		
Sample size	237	234	

^a 'High' is defined as bachelor degree, graduate diploma and graduate certificate, postgraduate degree; 'low' is defined as advanced diploma and diploma, certificate, no non-school education

Table 25: Comparison of household income – CV sample 1979:2009

Household income (\$) ^a	Reference sample 1979	Replicated sample 2009	Difference
Mean	79,546	88,819	+9,273
Standard deviation	35,349	60,521	
t-test	t = 1.997 df(453); p =0.047		
95% confidence interval	74,933-84,159	80,904-96,735	
Mann-Whitney	z = 0.630; p = 0.528		
Distribution of income groups (frequency)			
Nil income	1.32% (3)	6.17% (14)	+4.85%
1-7,799	0.88% (2)	1.32% (3)	+0.44%
7,800-12,999	1.32% (3)	0.88% (2)	-0.44%
13,000-18,199	0.88% (2)	2.64% (6)	+1.76%
18,200-25,999	2.19% (5)	4.41% (10)	+2.22%
26,000-33,799	1.75% (4)	6.61% (15)	+4.86%
33,800-41,599	5.26% (12)	3.96% (9)	-1.30%
41,600-51,999	11.40% (26)	8.81% (20)	-2.59%
52,000-62,399	7.89% (18)	4.85% (11)	-3.04%
62,400-72,799	7.46% (17)	7.05% (16)	-0.41%
72,800-88,399	20.61% (47)	6.61% (15)	-14.00%
88,400-103,999	17.11% (39)	11.01% (25)	-6.10%
104,000-129,999	12.72% (29)	10.13% (23)	-2.59%
130,000-155,999	8.33% (19)	8.37% (19)	+0.04%
156,000 and more	0.88% (2)	17.18% (39) ^b	+16.30%
156,000-181,999	-	7.49% (17)	
182,000-207,999	-	3.08% (7)	
208,000 and more	-	6.61% (15)	
Sample size	228	227	
Household income/ HM(\$) ^a	Reference sample 1979	Replicated sample 2009	Difference
Mean	27,903	32,748	+4,845
Standard deviation	18,567	22,003	
t-test	t = 2.539 df(453); p =0.012		
95% confidence interval	25,481-30,327	29,871-35,626	
Mann-Whitney	z = 2.604; p = 0.009		
Sample size	228	227	

^a Household income was adjusted by the Consumer Price Index Canberra (1979-2009)^b The highest 3 income categories are bundled

Table 26: Comparison of gender and place of birth – CV sample 1979:2009

Gender	Reference sample 1979	Replicated sample 2009	Difference
Male (1)	43.88% (104)	43.16% (101)	-
Female (0)	56.12% (133)	56.84% (133)	
Proportion	0.44	0.43	
Standard deviation	0.50	0.50	
95% confidence interval	0.38-0.50	0.37-0.50	
Pr-test	z = -0.158; p =0.875		
Sample size	237	234	
Place of birth	Reference sample 1979	Replicated sample 2009	Difference
Australia (1)	69.20% (164)	74.36% (174)	-
Other (0)	30.80% (73)	25.64% (60)	
Proportion	0.69	0.74	
Standard deviation	0.46	0.44	
95% confidence interval	0.63-0.75	0.69-0.80	
Pr-test	z = 1.244; p = 0.214		
Sample size	237	234	

The analysis of the socio-demographic variables was followed by a comparison of the attitudinal and behavioural variables (see Table 27). The results of the pr-test on equality in proportions indicates that the proportions of respondents who *visited a protected natural area* during the last year, those who *watched television shows featuring nature*, and those who *did a course in biology or related science* are statistically significantly different from zero (1% level, 5% level, 1% level). The former two proportions are lower in the 2009 sample than they are in the 1979 sample. In contrast, the proportion of respondents who *did a course in biology or related science* is higher in 2009 than it was in 1979. In contrast, the proportions of respondents who *find conservation very important*, who *read about nature*, and who are *subscribers of a conservation organisation* are not statistically significantly different across the two samples.

Table 27: Comparison of behavioural and attitudinal variables – sample 1979:2009

Importance of conservation	Reference sample 1979	Replicated sample 2009	Difference
Very important (1)	63.95% (149)	64.96% (152)	-
Other (0)	36.05% (84)	35.04% (82)	
Proportion	0.64	0.65	
Standard deviation	0.48	0.48	
95% confidence interval	0.58-0.70	0.59-0.71	
Pr-test	$z = 0.228$; $p = 0.820$		
Sample size	233	234	
Visit of a protected natural area	Reference sample 1979		Difference
Yes (1)	76.79% (182)	66.24% (155)	-10.55%
No (0)	23.21% (55)	33.76% (79)	
Proportion	0.77	0.66	
Standard deviation	0.42	0.47	
95% confidence interval	0.71-0.82	0.60-0.72	
Pr-test	$z = -2.538$; $p = 0.011$		
Sample size	237	234	
Read about nature	Reference sample 1979	Replicated sample 2009	Difference
Yes (1)	53.16% (126)	50.64% (118)	-
No (0)	46.84% (111)	49.36% (115)	
Proportion	0.53	0.51	
Standard deviation	0.50	0.50	
95% confidence interval	0.47-0.60	0.44-0.57	
Pr-test	$z = -1.547$; $p = 0.585$		
Sample size	237	233	
Watch television shows featuring nature	Reference sample 1979	Replicated sample 2009	Difference
Yes (1)	86.50% (205)	76.92% (180)	-9.58%
No (0)	13.50% (32)	23.08% (54)	
Proportion	0.87	0.77	
Standard deviation	0.34	0.42	
95% confidence interval	0.82-0.91	0.72-0.82	
Pr-test	$z = -2.689$; $p = 0.007$		
Sample size	237	234	

Table 27: Comparison of behavioural and attitudinal variables – sample 1979:2009 - continued

Course in biology or related science	Reference sample 1979	Replicated sample 2009	Difference
Yes (1)	17.72% (42)	33.91% (79)	+16.19%
No (0)	82.28% (195)	66.09% (154)	
Proportion	0.18	0.34	
Standard deviation	0.38	0.47	
95% confidence interval	0.13-0.23	0.28-0.40	
Pr-test	z = 4.012; p = < 0.001		
Sample size	237	233	
Subscriber to an environmental organisation	Reference sample 1979	Replicated sample 2009	Difference
Yes (1)	18.99% (45)	16.67% (39)	-
No (0)	81.01% (192)	83.33% (195)	
Proportion	0.19	0.17	
Standard deviation	0.39	0.37	
95% confidence interval	0.14-0.24	0.12-0.21	
Pr-test	z = -0.658; p = 0.511		
Sample size	237	234	

In summary, the results suggest statistically significant differences in the following variables across the reference and replicated samples:

- *Household income* (→ richer for the analysis that excluded valid zero marginal WTP bids; not statistically significantly different for the analysis that included valid zero marginal WTP bids)
- *Household income per HM* (→ richer)
- *Age* (→ older),
- *Non-school education* (→ generally better educated),
- *Visit of a protected natural area* (→ fewer visits),
- *Watch television shows featuring nature* (→ less watching television),
- *Course in biology or related science* (→ better educated in biology/ related science).

A comparison of these variables across the census data from 1976 and 2006 allowed investigating whether the observed changes in explanatory variables of demand found in the two samples are representative of the respective populations of Canberra. Since the census did not record the attitudinal and behavioural variables included in the CV studies the comparison was limited to the socio-demographic variables. As shown in Tables 28 – 31, the population in 2009 is, on average, older, better educated, richer, and characterised by more people in lower and higher income categories, and fewer people in the middle income categories than in 1979. The distribution of gender and place of birth has not changed.

Table 28: Comparison of age – Census 1976:2006

Age (years)	Census 1979	Census 2009	Difference
Distribution of age (frequency)			
18-19	5.85% (7,309)	4.03% (9,983)	-1.82%
20-24	14.73% (18,404)	11.07% (27,437)	-3.66%
25-34	31.87% (39,809)	20.22% (50,104)	-11.64%
35-44	19.62% (24,507)	19.68% (48,742)	+0.06%
45-54	15.09% (18,847)	18.53% (45,902)	+3.44%
55-64	7.93% (9,905)	13.76% (34,089)	+5.83%
65+	4.92% (6,145)	12.72% (31,507)	+7.80%
Population size	124,926	247,774	

^a Adjusted by the Consumer Price Index between 1976-2009

^b Population not stated? excluded

^c The 1976 census did not include the nursing homes

^d The adjustment by the Consumer Price Index required merging the income brackets 21 - 26 to 21-26 and 31 - 36 to 31-36

^e The highest 7 income categories are combined

Table 29: Comparison of household income – Census 1976:2006

Household income (\$) ^{a,b} (15 years and older)	Census 1976	Census 2006	Difference
Mean	49,729	94,111	+44,382
Distribution of household income (frequency)			
Nil income or negative income ^c	0.81% (430)	0.68% (707)	-0.13%
1-7,799	1.24% (656)	0.99% (1,022)	-0.25%
7,800-12,999	2.10% (1106)	3.67% (3,779)	+1.58%
13,000-18,199	3.52% (1860)	4.62% (4,752)	+1.10%
18,200-25,999	5.76% (3042)	2.20% (2,266)	-3.56%
26,000-33,799	9.25% (4884)	6.93% (7,128)	-2.32%
33,800-41,599	14.36% (7580)	5.42% (5,572)	-8.95%
41,600-62,399 ^d	14.86% (7844)	6.79% (6,984)	-8.07%
62,400-72,799	15.29% (8070)	11.58% (11,914)	-3.71%
72,800 and more	32.79% (17,304)	57% (58,739) ^e	+25.00%
72,800-88,399	-	5.08% (5,229)	
88,400-103,999	-	10.51% (10,810)	
104,000-129,999	-	9.59% (9,869)	
130,000-155,999	-	11.28% (11,603)	
156,000-181,999	-	11.71% (12,044)	
182,000-207,999	-	4.49% (5,086)	
208,000 and more	-	3.89% (4,098)	
Population size	59,933	102,857	

^a Adjusted by the Consumer Price Index Canberra (1979-2009)
^b Responses ‘not stated’ excluded
^c The 1976 census did not asked for negative income.
^d The adjustment by the Consumer Price Index required merging the income brackets 41,600-51,999 and 52,000-62,399.
^e The highest 7 income categories are bundled

Table 30: Comparison of gender and place of birth – Census 1976:2006

Gender (18 years and older)	Census 1976	Census 2006	Difference
Distribution of gender (frequency)			
Male	50.20% (62,719)	48.73% (120,737)	-1.48%
Female	49.80% (62,207)	51.27% (127,037)	+1.48%
Population size	124,926	247,774	
Place of birth (15 years and older)	Census 1979	Census 2009	
Distribution of place of birth (frequency)			
Australia	70.26% (95,333)	68.87% (179,949)	-1.39%
Other	29.74% (40,357)	31.13% (81,339)	+1.39%
Population size	135,690	261,22	

Table 31: Comparison of non-school education – Census 1976:2006

Non-school education (15 years and older) ^a	Census 1976	Census 2006	Difference
Distribution of non-school education (frequency)			
Postgraduate degree	2.48% (3,005)	7.55% (17,816)	+5.07%
Graduate diploma, graduate certificate	1.03% (1,245)	3.88% (9,164)	+2.85%
Bachelor	9.02% (10,925)	21.81% (51,495)	+12.79%
Certificate, advanced diploma, diploma	22.16% (26,848)	23.91% (56,443)	+1.75%
No non-school education	65.32% (79,135)	42.86% (101,189)	-22.46%
Population size	121,158	236,107	

^a Unclassified responses excluded

Comparing the changes in the samples with those in the census data provides the following picture. No significant changes in *gender* and *place of birth* were observed in either the samples or the populations. This suggests that the distribution stabilities of *gender* and *place of birth* observed in the population are reflected by the distributions observed in the samples. As illustrated in Figures 8 and 9, changes in *non-school education* and *age* shifted largely in the same direction. These similar shifts indicate that the changes in *non-school education* and *age* observed in the population are represented at least to some extent by the changes observed in the samples. The mean of *household income* of the samples also shifted in the same direction as the census. The shifts by income categories, however, are less concordant with a shift in the highest income category being largely under-represented (see Figure 10). Census data on *household*

income per HM was not available due to missing information on the mean number of people per household in the 1976 Census. However, knowing that the mean *household income* of the population increased and assuming that the number of HM did not increase (more likely decreased) from 1976 to 2006, allows the conclusion that *household income per HM* also increased over the last 30 years. Hence, it can be assumed that the mean of *household income per HM* in the sample and the census shifted in the same direction indicating that changes in the mean of this variable observed in the population are represented at least to some extent by the changes observed in the samples. However, the positive shift observed in the sample may under-represent the positive shift in the population.

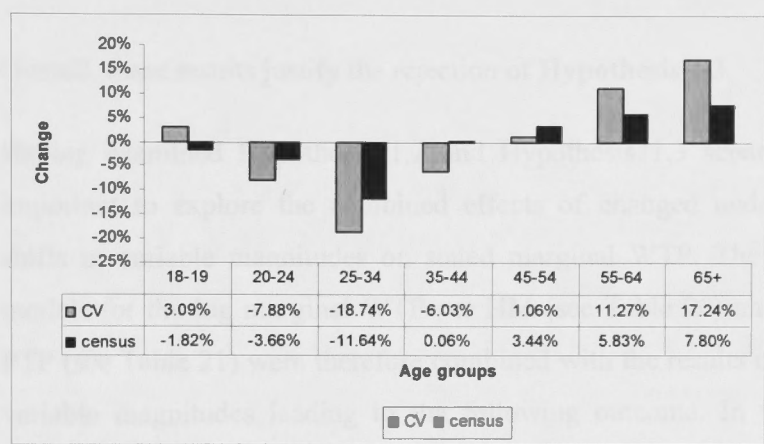


Figure 8: Comparison of changes in age between CV sample and Census data

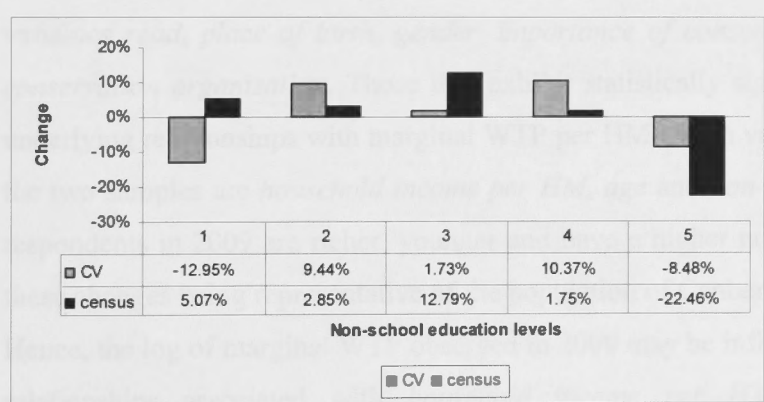


Figure 9: Comparison of changes in non-school education between CV sample and Census data

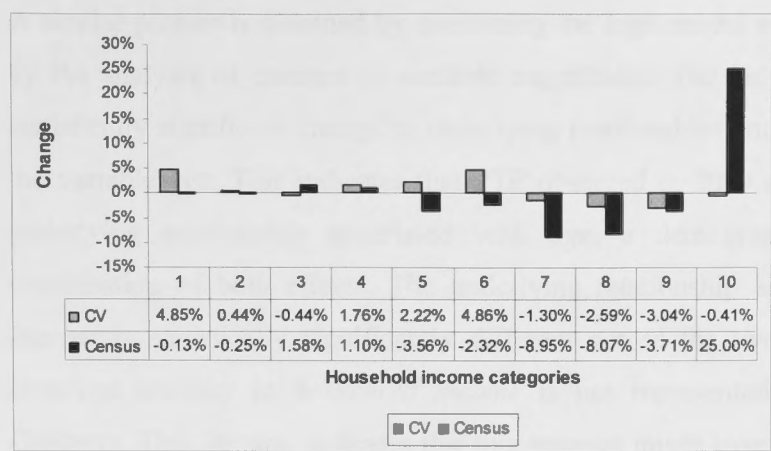


Figure 10: Comparison of household income between CV sample and Census data

Overall, these results justify the rejection of **Hypothesis 1.3**.

Having examined Hypothesis 1.2 and Hypothesis 1.3 separately, it was considered important to explore the combined effects of changed underlying relationships and shifts in variable magnitudes on stated marginal WTP. The results of the log-linear models for the log marginal WTP per HM (see Table 20) and the logit model for the PTP (see Table 21) were therefore combined with the results of the analysis of shifts in variable magnitudes leading to the following outcome. In the log-linear model the variables that exhibit statistically significant changes in their underlying relationships with marginal WTP per HM but without having changes in variable magnitudes are the variables *read*, *place of birth*, *gender*, *importance of conservation* and *subscriber to conservation organisation*. Those that exhibit statistically significant changes in their underlying relationships with marginal WTP per HM and in variable magnitudes across the two samples are *household income per HM*, *age* and *non-school qualification*. The respondents in 2009 are richer, younger and have a higher non-school education, with these changes being representative of the population of Canberra at least to some extent. Hence, the log of marginal WTP observed in 2009 may be influenced by the underlying relationships associated with *household income per HM*, *age* and *non-school qualification*, a demographic shift in *household income per HM*, *age* and *non-school qualification*, or a combination of both effects. The underlying relationship associated with *household income per HM* is statistically significantly different across the two samples, with the parameter estimate obtained in 1979 having a stronger positive influence on the log marginal WTP per HM than that obtained in 2009.

A similar picture is obtained by combining the logit model results with those obtained by the analysis of changes in variable magnitudes. The only variable that exhibits a statistically significant change in underlying relationships and changes in magnitude is the variable *age*. This indicates that PTP observed in 2009 may be influenced by the underlying relationship associated with *age*, a demographic shift in *age*, or a combination of both effects. The underlying relationship associated with *household income* is statistically significantly different across the two samples. However, the observed stability in *household income* is not representative of the population of Canberra. This, in turn, indicates that this analysis might have underestimated the effect of *household income* on the PTP in the 2009 sample. In other words, the sampling might have stabilised potential change in PTP across the two samples through failing to account for a population shift in *household income*.

This analysis was expanded further by examining changes in respondents' choice of payment vehicles across the two samples (**Hypothesis 1.4**). As illustrated in Table 32, the Pearson's chi-squared test suggests statistical differences at the 1% significance level. The majority of respondents in 1979 equally preferred either the government or the conservation organisation. This contrasts with the situation in 2009. The data indicates a shift towards preferences for a mixed payment vehicle. This may suggest that respondents in 2009 may not trust either the government or the conservation organisation by itself. This difference in payment vehicle might have influenced the marginal WTP. Rigorously testing such an effect, however, was not possible since only respondents with a positive marginal WTP were asked this question in the reference study.

Table 32: Comparison of payment vehicle choice – CV sample 1979:2009

Payment vehicle	Reference sample 1979	Replicated sample 2009	Difference
Government	40.69% (94)	27.35% (64)	-13.34%
Conservation organisation	47.19% (109)	13.25% (31)	-33.94%
Either	4.33% (10)	4.70% (11)	+0.37%
Both	7.79% (18)	54.70% (128)	+46.91%
Pearson’s chi-squared test (independence of observations)	$\chi^2 = 132.064$ df (3); $p < 0.001$		
Sample size	231	234	

These results warrant a rejection of the null **Hypothesis 1.4**.

9.4 Explaining marginal willingness to pay – supply side

Possible effects of changes in variables of supply on stated marginal WTP for the preservation of Nadgee Nature Reserve were explored by examining whether the availability of close substitutes for and associated relative scarcity of the preservation of Nadgee Nature Reserve are the same in 1979 and 2009 (**Hypothesis 1.5**). This hypothesis was tested by comparing the number and extent of protected terrestrial natural areas in New South Wales¹⁰⁰ in 1979 and 2009 assuming that for the respondents they are (nearly) perfect substitutes for the preservation of Nadgee Nature Reserve. The results of this comparison are displayed in Tables 33 and 34. The data shows that both the number and extent of protected terrestrial natural areas increased significantly. In 2009 there existed about four times more protected natural areas than in 1979. The extension of protected areas increased about three times from 1979 to 2009. This increase in supply of protected natural areas might indicate that the relative scarcity of protected natural areas decreased from 1979 to 2009, as did the preservation value of Nadgee Nature Reserve.

¹⁰⁰ Protected terrestrial areas include national parks, nature reserves and state conservation areas, as defined by the New South Wales Government.

Table 33: Number of protected areas in New South Wales

Number of protected areas in NSW ^a	1979	2009	Change
National Parks	45	186	+141
Nature Reserve	118	394	+276
State Conservation Areas	2	103	+101
Total	165	683	+518

^a Data obtained from NSW National Parks and Wildlife Service, New South Wales Government (2011)

Table 34: Extent of protected areas in New South Wales

Area (ha) of protected areas in NSW	1979 ^a	2009 ^b	Change
National Parks	1,750,518	4,983,811	+3,233,293
Nature Reserve	357,209	887,766	+530,557
State Conservation Areas	5,537 ^c	456,391	+450,854
Total	2,113,264	6,329,977	+4,216,713

^a Data obtained from Goldstein (1979)

^b Data obtained from NSW National Parks and Wildlife Service, New South Wales Government (2011)

^c Extent as stated in 2009; no data is available for extent as stated in 1979

However, this interpretation of results needs to be treated with caution. Firstly, testing this hypothesis is based on the maintained assumption that the actual change in the number and extent of protected terrestrial natural areas in New South Wales and the change perceived by respondents were the same. A perceived change that is smaller than the actual change might not have decreased or even increased the relative scarcity. Secondly, the analysis is based on the assumption that respondents restricted close substitutes to protected terrestrial natural areas. A wider definition including protected and unprotected terrestrial natural areas would imply a different interpretation. First of all, declaring a terrestrial natural area as being protected does not increase the supply of terrestrial natural areas. Together with the decreasing extent of terrestrial natural areas through ongoing development of land this implies a decrease in supply of close substitutes and thus an increase in relative scarcity. Hence, a wider definition of what constitutes close substitutes implies an increase in the preservation value of Nadgee Nature Reserve. Since respondents’ definitions of close substitutes for the preservation of Nadgee Nature Reserve remain unknown to the analyst no conclusive evidence could be collected to substantiate either of the interpretations.

Consequently, the lack of conclusive evidence does not warrant a decision on whether **Hypothesis 1.5** can be rejected.

9.5 Conclusion

Chapter 9 has analysed value variability across a 30-year time period using the preservation value of Nadgee Nature Reserve as an example. Section 9.2 investigated the temporal variability of stated marginal WTP obtained by an open-ended CV in 1979 and 2009 (**Hypothesis 1.1**). The results suggest that the mean marginal WTP per HM (including and excluding zero values) attached to the preservation of this pristine area of land in 1979 is not different from that in 2009. This is supported by stable PTP to preserve Nadgee Nature Reserve across the two points in time. In contrast, indications for temporal differences were observed in the variance of the marginal WTP per HM (including and excluding zero values) and in the median marginal WTP per HM (excluding zero values). This indicates that even though the mean marginal WTP remained unchanged across these two points in time underlying demand factors influencing marginal WTP did not remain unchanged. The results also suggest that the proportion of protest zero marginal WTP bids remained unchanged, whereas the underlying protest motivations are different for 1979 and 2009. That is, the reasons why respondents stated protest zero marginal WTP bids differ across the two points in time. This difference is concentrated on attitudes towards tax payments, government responsibility and the amount of information that was available to answer the marginal WTP question.

Section 9.3 explored the potentially changing impact of explanatory variables of demand on the preservation value of Nadgee Nature Reserve (**Hypothesis 1.2** and **Hypothesis 1.3**). Testing these two hypotheses required the choice of an appropriate econometric model. The data characteristics of both samples disqualified any model that could handle a non-trivial proportion of corner solutions, and that, in theory, would be superior to a linear and log-linear regression model using an OLS estimator. Since neither the reference nor the replicated data set is considered to be an outlier, it is concluded that analysing data obtained by open-ended CV generally still poses a significant challenge to the analyst and casts doubts on the explanatory power and usability of results obtained with this particular stated preference method.

The best-fitting model, a log-linear regression model estimating the relationship between log positive marginal WTP per HM and socio-economic, attitudinal and behavioural variables was used to explain marginal WTP from a demand side

perspective. The results suggest that in 1979 *household income per HM*, *read about nature* and *place of birth*, and in 2009 *gender*, *household income per HM*, *importance of conservation*, and *subscriber to conservation organisation* positively influenced preferences through their underlying relationships with log marginal WTP. Whether the observed change related to *household income per HM* affected marginal WTP through an impact on preferences or in its capacity as a budget constraint could not be distinguished. Moreover, the model outcomes provide evidence that in 2009 *household income per HM*, *non-school education* and *age* positively influenced preferences through a demographic shift and/ or their underlying relationship with log marginal WTP. *Household income per HM* was the only demand factor explaining log marginal WTP at both points in time. Whether these changes observed through the log-linear model resulted in changed preferences reflected in the slope of the indifference curves, and thus in a change in the marginal rate of substitution for the preservation of Nadgee Nature Reserve in 2009 compared to 1979, depends on the relative influences of preference changes in 1979 relative to 2009 and the effect of the budget constraint, *ceteris paribus*.

The results of analysing PTP to preserve Nadgee Nature Reserve using a logit model provided indications that in 1979 *importance of conservation*, *watch television*, and *place of birth*, and in 2009 *household income*, *importance of conservation*, and *read about nature* (2009) positively influenced preferences through their underlying relationships with PTP. The budget constraint represented as *household income* was found to be not statistically significantly different across the 30-year time period. Furthermore, the results suggest that in 2009 *age* negatively influenced preferences through a demographic shift and/ or its underlying relationship with PTP. None of the demand factors included in these models was significant in 1979 and 2009 simultaneously. Whether these findings suggest a changed slope of the indifference curves and thus changed preferences for the preservation of Nadgee Nature Reserve in 2009 compared to 1979 depends on the magnitude of the influence related to *age* in comparison to the other demand factors. However, the statistically insignificant *household income* parameter estimate casts doubt on the validity of the model results. It might be an indication of inconsequentiality in the surveys. Additionally, it needs to be noted that the stability in *household income* found in the samples was not representative of the population of Canberra, where the Census data indicates that residents in 2009, on average, are wealthier than those in 1979. Hence, the estimations may have

underestimated the positive impact of *household income* on PTP for the preservation of Nadgee Nature Reserve through its potential effect on preferences and in its capacity as budget constraint.

Section 9.3 also investigated whether the attitudes of respondents towards the payment vehicle presented in an open-ended CV survey associated with the preservation value of Nadgee Nature Reserve were the same for 1979 and 2009 (**Hypothesis 1.4**). The findings indicate that these attitudes were different in 1979 compared to those in 2009. However, whether this change affected stated marginal WTP could not be rigorously tested.

Section 9.4 explored the impact of changing availability of close substitutes for and associated relative scarcity of the preservation of Nadgee Nature Reserve across time (**Hypothesis 1.5**). The comparison of the number and extent of protected natural areas in New South Wales indicated that the availability of close substitutes and thus the relative scarcity increased from 1979 to 2009. This may have reduced the marginal rate of substitution though a downward movement along the indifference curve. This would indicate a decrease in marginal WTP for the preservation of Nadgee Nature Reserve, *ceteris paribus*. However, a wider definition of close substitutes may have resulted in an opposite scenario: an increase in relative scarcity may have increased the marginal rate of substitution in marginal WTP through an upward movement along the indifferent curve, *ceteris paribus*.

In summary, Chapter 9 did not find conclusive evidence that the marginal value for the preservation of Nadgee Nature Reserve is different for 1979 and 2009. The results indicate that mean stated marginal WTP per HM for the preservation of Nadgee Nature Reserve are the same between the two points in time. However, underlying demand factors influencing marginal WTP per HM were found to be different. This finding is supported by results indicating a change in the influence of demand side factors on marginal WTP per HM. This indicates that underlying demand side factors influencing preferences and budget constraint were different in 1979 compared to 2009. Whether these changes together altered preferences for the preservation of Nadgee Nature Reserve or cancelled each other out remains unclear. However, stable marginal WTP per HM together with a reduced relative scarcity (decreased marginal rate of substitution) would indicate a preference change (increased marginal rate of substitution). On the other hand, stable marginal WTP per HM together with an

increased relative scarcity (increased marginal rate of substitution) would still suggest a preference change but in the other direction (decreased marginal rate of substitution).

Other possible interpretations are related to the problems encountered with the representativeness of the sample and the associated changes in the sample characteristics across the two points in time. Even though *household income per HM* shifted in the same direction in the sample and the census data, the magnitude of this shift in the census data is unknown. Hence, representative samples with respect to income (or any omitted variables) might have yielded higher (or lower) stated marginal WTP per HM. That is, the sampling might have stabilised temporal variable value estimates. This discussion clearly shows that without additional information the results do not reveal the reasons that underpin the temporal stability of the preservation value of Nadgee Nature Reserve.

Chapter 9 analysed temporal value variability ignoring potential influences of changes in stated preference methods on value estimates. Chapter 10 will relax this assumption by analysing some of the effects of this methodological evolution.

10 ANALYSING IMPACTS OF METHODOLOGICAL EVOLUTION

10.1 Introduction

Chapter 9 investigated the first research question: Do non-market values attached to environmental goods and services change across extended time periods? Chapter 10 focuses on the exploration of the second research question: Has the evolution of stated preference methods impacted on marginal WTP estimates of non-market values attached to environmental goods and services? As with the first research question, the foundation for the analysis of the second research question was given by the case study introduced in Chapter 6, the overall research design discussed in Chapter 7, and the data collection process described in Chapter 8. The remainder of Chapter 10 is structured as follows. Section 10.2 examines the effect of the alternative stated preference methods CV and DCE (see Sections 4.2 and 4.3) on the equivalent surplus for the preservation of Nadgee Nature Reserve (see Section 7.3.1). This is followed by Section 10.3, which investigates prominence in response in open-ended CV (see Section 4.2.2) and its potential effects on temporal variability of stated marginal WTP for the preservation of Nadgee Nature Reserve (see Section 7.3.2). Section 10.4 assesses incentive compatibility properties of DCE (see Section 4.3.2) and investigates the occurrence of strategic response (see Section 7.3.3). A conclusion provided in Section 10.5 discusses the results presented in this Chapter against the background of the econometric theory discussed in Section 4.

The analysis of valuation across time presented in this chapter is based on the following maintained assumptions stated in Chapters 5 -8:

MA 5.1: The applied stated preference method is temporally reliable.

MA 5.2: Changes in the relative prices of composite goods do not influence marginal WTP.

MA 5.3: The effect of varying perceptions of reference points is negligible.

MA 6.1: Respondents were unaware that the size of the area of land presented in the questionnaire in 2009 was held constant at 1979 levels and thus differed from its current size.

MA 8.5: The marginal differences in wording of the choice questions did not significantly alter choice incentives across the SB, RBs, and FirstRB split samples.

MA 8.6: Respondents did express their preferences for the preservation of Nadgee Nature Reserve as described in the survey questionnaires, and not for the preservation of natural areas in general.

10.2 Discrete choice experiments as an alternative to contingent valuation

The effect of alternative valuation methods on the preservation value of Nadgee Nature Reserve (**Hypothesis 2.1**) was investigated by comparing equivalent surpluses obtained in 2009 through an open-ended CV study and a state-of-the-art DCE-TR. The open-ended CV equivalent surplus was calculated as the mean value of the stated marginal WTP bids obtained in the replicated study (see Table 35).

Table 35: Equivalent surplus obtained through DCE-RT and open-ended CV¹⁰¹

	Implicit price	95% confidence interval ^a (implicit price)	Equivalent surplus	95% confidence interval ^b (equivalent surplus)
pRPL(H-C)	\$29.43	\$23.41 - \$35.45	\$412.00	\$327.72 - \$496.28
pRPL(P-C)	\$28.29	\$21.37 - \$35.21	\$396.06	\$299.20 - \$492.92
CV	-	-	\$96.66	\$73.17 - \$120.16

^a The 95% confidence interval of the implicit prices was calculated by the Delta method (mean +/- 1.96*standard error)

^b The confidence interval of equivalent surplus was calculated using the confidence intervals of the implicit prices.

The calculation of the DCE-RT equivalent surplus is based on implicit prices estimated through pRPL models in WTP space, allowing for free correlations across the random parameters. All attributes in all models were defined as random parameters to account

¹⁰¹ Protest voters were excluded from the analysis.

for preference heterogeneity assuming normal distributions, whereas the constant was treated as a fixed parameter. The models were estimated with Biogeme 2.0 (Bierlaire, 2003) through maximum simulated likelihood with 500 Halton draws. The BIO (Conn et al., 2000) and SOLVOPT (Kuntsevich and Kappel, 1997) algorithms available in Biogeme 2.0 were used to maximise the log-likelihood function. As with all currently used optimisation algorithms, BIO and SOLVOPT are local optimisation algorithms. Hence, there is a chance that estimation results represent a local rather than a global optimum. However, the problem of local optima was reduced by estimating the same model using more than one local optimisation algorithm. Another approach to reducing this problem is to perform local optimisations from random starting points. This, though, is a highly time consuming exercise, which is beyond the scope of this thesis. Besides, neither the performance of optimisations from random starting points nor the use of alternative optimisation algorithms guarantees avoiding a local optimum. In this study, implicit prices obtained through CL model estimation in preference space were used as starting values for the parameter estimates related to the attributes *cost*, *area*, and *access* (see Appendix G).

The *area* parameter was presented to respondents as both percentage (P) and hectares (H) (see Section 8.3.2). To account for possible differences introduced through this dual communication, the pRPL models were estimated using both hectares and percentages. Separate models were estimated with and without allowing for freely correlated random parameters (C). This leads to four different model specifications: pRPL(H), pRPL(H-C), pRPL(P), pRPL(P-C). Additionally, socio-demographic variables (SD) were included into the pRPL(H-C) model, resulting in a model including *age*, *gender*, *household income*, and *non-school education*, denominated as pRPL(H-C-SD), and a model solely incorporating *household income*, labelled pRPL(H-C-HI). To test if the algorithms used impact on the results, the pRPL(P) model was estimated using the BIO and the SOLVOPT algorithms. The final log-likelihood functions for the pRPL(P) model are -1005.462 and -1004.996, respectively. Since the difference between the two log-likelihood functions is rendered negligible, the remaining three models were estimated using the BIO algorithm only.

The results of the pRPL(H) model and the pRPL (H-C) model are presented in Table 36. A likelihood ratio test¹⁰² and the McFadden pseudo rho-square adjusted¹⁰³ indicate a

¹⁰² Likelihood-ratio test: $\chi^2, df = -2 * (LL_{base\ model} - LL_{estimated\ model})$

reasonable model fit for both pRPL(H) and pRPL(H-C). In both models, the estimates for the implicit prices of *area* (1000 ha) and *access* are positive and statistically different from zero at the 1% significance level. Furthermore, the diagonal elements of the Cholesky matrix are statistically different from zero at the 1% significance level for both *area* and *access*. This suggests that the implicit prices for these attributes are heterogeneous across respondents. A comparison of the information criteria AIC, BIC, and HQIC suggests that the pRPL(H-C) model that allows for freely correlated random parameters is superior to the pRPL(H) model that assumes zero correlation. This is supported by the off-diagonal element in the Cholesky matrix (*area* – *access*), which is statistically different from zero at the 1% significance level.

¹⁰³ Rho-square adjusted test statistic: $\rho^2_{adj.} = 1 - ((LL_{estimated\ model} - K) / LL_{base\ model})$, with K estimated parameters

Table 36: Model results of pRPL(H) and pRPL(H-C) estimation in WTP space¹⁰⁴

Variable	pRPL(H)		pRPL(H-C)	
	Coefficient ^a	Standard error ^b	Coefficient ^a	Standard error ^b
<i>Nonrandom parameters</i>				
constant	190.62*** (<0.01)	31.40	199.89*** (<0.01)	28.10
<i>Random parameters</i>				
cost (\$)	-4.32*** (<0.01)	0.33	-4.24*** (<0.01)	0.25
area (hectares)	28.66*** (0.00)	4.31	28.29*** (<0.01)	3.53
access (yes 1, no -1)	73.15*** (<0.01)	12.40	76.86*** (<0.01)	10.0
<i>Diagonal elements in Cholesky matrix</i>				
cost (n)	0.87** (<0.01)	0.39	0.91*** (<0.01)	0.289
area (n)	45.51*** (<0.01)	7.59	44.54*** (<0.01)	6.91
access (n)	-105.39*** (<0.01)	14.9	-101.67*** (<0.01)	12.1
<i>Off-diagonal elements in Cholesky matrix</i>				
area - access			45.33*** (<0.01)	11.1
<i>Model statistics</i>				
n (observations)	1228		1228	
LL _{β}	-1004.996		-997.573	
$\chi^2_{\cdot 7}$ compared to LL _{$\beta=0$}	688.199			
$\chi^2_{\cdot 6}$ compared to LL _{base}	391.944			
$\chi^2_{\cdot 8}$ compared to LL _{$\beta=0$}			703.573	
$\chi^2_{\cdot 7}$ compared to LL _{base}			407.316	
McFadden pseudo ρ^2 adj.	0.250		0.255	
AIC ^c	1.6482		1.6373	
AIC ^c finite sample	1.7400		1.7554	
BIC ^c	1.6773		1.6706	
HQIC ^c	1.6592		1.6498	

^a ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

^b Variance-covariance matrix robust against non severe mis-specification errors with respect to the characteristics of the postulated distribution for the error terms (Bierlaire, 2009)

^c Normalised by sample size

The results of the pRPL (P) model and the pRPL(P-C) model are presented in Table 37. Both models exhibit a reasonable model fit as suggested by a likelihood ratio test¹⁰⁵ and the McFadden pseudo rho-square¹⁰⁶. In both models the implicit prices for *area* and *access* are positive as expected and statistically different from zero at the 1% significance level. The elements of the diagonal Cholesky matrix are statistically different from zero at the 1% significance level, indicating heterogenic implicit prices

¹⁰⁴ The results of the pRPL (H-C) model estimated by Biogeme were checked by calculating the variance-covariance matrix based on the diagonal and off-diagonal elements in the Cholesky matrix using STATA 10 (matrix cc=c*c').

¹⁰⁵ Likelihood-ratio test: $\chi^2, df=2*(LL_{base model} - LL_{estimated model})$

¹⁰⁶ Rho-square adjusted test statistic: $\rho^2 adj. = 1 - ((LL_{estimated model} - K) / LL_{base model})$, with K estimated parameters

across respondents. Again, the model that allows freely correlated random parameters is to be preferred over the constrained alternative as suggested by the information criteria and the off-diagonal element, which is statistically different from zero at the 1% significance level.

Table 37: Model results of pRPL(P) and pRPL(P-C) estimation in WTP space¹⁰⁷

Variable	<i>pRPL(P)</i>		<i>pRPL(P-C)</i>	
	Coefficient ^a	Standard error ^b	Coefficient ^a	Standard error ^b
<i>Nonrandom parameters</i>				
constant	188.57*** (<0.01)	30.30	197.30*** (<0.01)	29.60
<i>Random parameters</i>				
cost (\$)	-4.22*** (<0.01)	0.31	-4.22*** (<0.01)	0.31
area (percent)	4.14*** (0.00)	0.41	4.12*** (<0.01)	0.43
access (yes 1, no -1)	76.91*** (<0.01)	11.7	80.24*** (<0.01)	9.44
<i>Diagonal elements in Cholesky matrix</i>				
cost (n)	1.03*** (0.01)	0.38	0.95*** (0.01)	0.34
area (n)	-6.26*** (<0.01)	0.85	-6.27*** (<0.01)	0.79
access (n)	-105.56*** (<0.01)	13.1	-103.00*** (<0.01)	11.5
<i>Off-diagonal elements in Cholesky matrix</i>				
area - access			42.86*** (<0.01)	13.9
<i>Model statistics</i>				
n (observations)	1228		1228	
LL _{β}	-1005.462		-997.614	
$\chi^2_{\cdot 7}$ compared to LL _{$\beta=0$}	687.268			
$\chi^2_{\cdot 6}$ compared to LL _{Basic}	391.012			
$\chi^2_{\cdot 8}$ compared to LL _{$\beta=0$}			702.963	
$\chi^2_{\cdot 7}$ compared to LL _{Basic}			406.708	
McFadden pseudo ρ^2 adj.	0.250		0.255	
AIC ^c	1.6489		1.6378	
AIC ^c finite sample	1.7408		1.7400	
BIC ^c	1.6781		1.6711	
HQIC ^c	1.6599		1.6503	

^a ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

^b Variance-covariance matrix robust against non severe mis-specification errors with respect to the characteristics of the postulated distribution for the error terms (Bierlaire, 2009)

^c Normalised by sample size

The pRPL(H-C) was extended through the inclusion of socio-demographics. This resulted in a model that incorporates *age*, *gender*, *household income*, and *non-school education*, denominated as pRPL(H-C-SD) and a model that solely incorporates

¹⁰⁷ The results of the pRPL (H-C) model estimated by Biogeme were checked by calculating the variance-covariance matrix based on the diagonal and off-diagonal elements in the Cholesky matrix using STATA 10 (matrix cc=c*c').

household income, labelled pRPL(H-C-HI). As shown in Table 38, the inclusion of socio-demographics did not improve the model fit. The information criteria AIC finite sample, BIC, and HQIC of both pRPL(H-C-SD) and pRPL(H-C-HI) are larger than those of pRPL(H-C). Only the AIC suggests the opposite. However, the AIC finite sample is the better indicator since it punishes more severely an improvement in model fit through increased number of estimates variables. Additionally, none of the parameter estimates associated with the socio-demographics is statistically different from zero at the 5% significance level. In the pRPL(H-C-SD) model, *gender* and *non-school education* are statistically different from zero at the 10% significance level, indicating that male respondents have a higher marginal WTP than female respondents and respondents with a higher non-school education have a lower marginal WTP than respondents with no or a lower level of non-school education. In the pRPL(H-C-HI) model, *household income* is not statistically significantly different from zero ($p = 0.20$) but the standard error is lower than in the pRPL(H-C-SD) model suggesting potential correlation problems between *household income* and the other three socio-demographics. All attribute parameter estimates are statistically different from zero at the 1% level and have the expected signs in both models.

Table 38: Model results of pRPL(H-C-SD) and pRPL(H-C-HI) estimation in WTP space¹⁰⁸

Variable	pRPL(H-C-SD)		pRPL(H-C-HI)	
	Coefficient ^a	Standard error ^b	Coefficient ^a	Standard error ^b
<i>Nonrandom parameters</i>				
constant	87.55 (0.21)	69.60	243.81*** (<0.01)	48.00
age	-2.07 (0.92)	19.70		
gender (male 1, female -1)	2.57* (0.06)	1.39		
household income / 1000\$	0.01 (0.98)	0.32	-0.40 (0.20)	0.31
non-school education ^d (high 1, low -1)	-29.08* (0.08)	16.60		
<i>Random parameters</i>				
cost (\$)	-4.19*** (<0.01)	0.30	-4.23*** (<0.01)	0.27
area (1000 hectares)	28.56*** (<0.01)	3.76	28.24*** (0.00)	3.14
access (yes 1, no -1)	76.51*** (<0.01)	9.42	77.06*** (<0.01)	10.00
<i>Diagonal elements in Cholesky matrix</i>				
cost (n)	0.95*** (0.01)	0.34	0.93*** (<0.01)	0.31
area (n)	44.09*** (<0.01)	5.96	44.54*** (0.00)	6.60
access (n)	-100.44*** (<0.01)	12.10	-100.69*** (<0.01)	11.6
<i>Off-diagonal elements in Cholesky matrix</i>				
area - access	39.86*** (<0.01)	10.50	44.54*** (<0.01)	11.2
<i>Model statistics</i>				
n (observations)	1228		1228	
LL _β	-992.401		-996.492	
χ ² _{·12 compared to LL_β=0}	713.390		705.207	
χ ² _{·11 compared to LL_βbase}	417.134		408.952	
Mcfadden pseudo ρ ² adj.	0.256		0.255	
AIC ^c	1.6358		1.6376	
AIC ^c finite sample	1.8926		1.7854	
BIC ^c	1.6858		1.6751	
HQIC ^c	1.6546		1.6517	

^a ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

^b Variance-covariance matrix robust against non severe mis-specification errors with respect to the characteristics of the postulated distribution for the error terms (Bierlaire, 2009)

^c Normalised by sample size

^d 'High' is defined as bachelor degree, graduate diploma and graduate certificate, postgraduate degree; 'low' is defined as advanced diploma and diploma, certificate, no non-school education

Hence, the implicit prices of pRPL(H-C) and pRPL(P-C) were used to calculate the equivalent surplus¹⁰⁹. The equivalent surplus is defined as follows (see Section 4.3.3):

¹⁰⁸ The results of the pRPL (H-C-SD) model and the pRPL(H-C-HI) model estimated by Biogeme were checked by calculating the variance-covariance matrix based on the diagonal and off-diagonal elements in the Cholesky matrix using STATA 10 (matrix cc=c*c').

$$ES = -\frac{1}{\gamma}(V_0 - V_1).$$

Since the implicit prices rather than the attribute parameters were estimated this equation simplifies to

$$ES = (V_0 - V_1), \text{ with utility function } V = \text{con} + \beta_{\text{area}} * \text{area} + \beta_{\text{access}} * \text{access}.$$

$$V_1 = \text{con} + \beta_{\text{area}} * 14 + \beta_{\text{access}} * -1, \text{ assuming an area of 14000 hectares and no access, and}$$

$$V_0 = \text{con} + \beta_{\text{area}} * 0 + \beta_{\text{access}} * -1, \text{ assuming an area of 0 hectares and no access, hence}$$

$$ES = \beta_{\text{area}} * 14.$$

The equivalent surpluses based on implicit prices in hectares and percentages are presented in Table 35. To allow a direct comparison, the percentage estimates were converted into hectares. As suggested by the overlapping 95% confidence intervals the equivalent surpluses obtained based on pRPL(H-C) and pRPL(P-C) are not statistically significantly different from each other.

As shown in Table 35, the confidence intervals of the equivalent surplus obtained through DCE-TR do not overlap with the 95% confidence interval of the equivalent surplus obtained through open-ended CV. This suggests that the equivalent surpluses estimated at the mean values are statistically different from each other at the 5% significance level. The DCE-TR equivalent surplus is approximately four times larger. This discrepancy may be even larger if the assumption of a linear relationship between *area* and marginal WTP over the range from zero hectares to 4,200 hectares is violated, with respondents exhibiting a diminishing marginal benefit.

These results strongly suggest that the equivalent surplus for the preservation of Nadgee Nature Reserve in 2009 obtained by an open-ended CV is statistically significantly different (lower WTP) from that elicited by a state-of-the-art DCE-TR. These findings are supported by other studies comparing open-ended CV with DCE or dichotomous choice CV (for example, Balistreri et al., 2001; Bijlenga et al., 2011). As outlined in Section 7.3.1, testing for causes explaining this discrepancy in equivalent surplus is beyond the scope of this study. Potential drivers are too manifold as to be controllable in

¹⁰⁹ The equivalent surplus based on the pRPL(H-C-SD) and pRPL(H-C-HI) models was not statistically significantly different from that based on the pRPL(H-C) model.

this research design. However, these results do illustrate that using data that is based on outdated stated preference methods for current policy decisions may lead to inefficient resource allocations.

The discrepancy across the two methods may be introduced by differences in sampling, questionnaire design, survey mode, elicitation format, the definition of environmental goods and services, the econometric analysis, differences in protest votes, or a combination of these factors. Nevertheless, it is possible to investigate whether differences in sample characteristics as a result of some of these factors may have contributed to this discrepancy. The sample characteristics of the open-ended CV and the DCE-TR sample are presented in Tables 39 - 42. Pearson’s chi-squared tests, t-tests for equality in means, and pr-tests for equality in proportions suggest that the mean of *age*, mean of *household income*, and the proportions of *non-school education* are statistically different from each other at the 1% significance level, whereas the proportion of female and male respondents is equal. The DCE-TR sample is younger, better educated and has a higher income. Even though *age*, *household income*, and *non-school education* did not have a statistically significant impact on the DCE equivalent surplus, these socio-demographics may have affected the CV equivalent surplus (see Section 9.3). Additionally, using a pr-test, the proportions of protest votes across the open-ended CV replicated sample (0.20) and the DCE-TR sample (0.12) are found to be statistically significantly different from each other ($p<0.001$). This difference in protest vote proportion might have confounded the comparison.

Table 39: Comparison of age – CV(replicated sample):DCE-RT

Age (years)	CV	DCE	Difference
Mean	48.35	40.23	-8.12
Standard deviation	19.08	13.68	
t-test	t =5.928 df(576); p < 0.001		
95% confidence interval	46.07-50.63	38.69-41.78	
Distribution of age groups (frequency)			
18-24	13.65% (37)	15.64% (48)	
25-34	14.39% (39)	22.80% (70)	
35-44	17.71% (48)	22.80% (70)	
45-54	13.28% (36)	21.82% (67)	
55-64	16.61% (45)	12.38% (38)	
65+	24.35% (66)	4.56% (14)	
Sample size	271	307	

Table 40: Comparison of non-school education – CV(replicated sample):DCE-RT

Non-school education (frequency)	CV	DCE	Difference
Postgraduate degree	11.85% (32)	12.38% (38)	+0.53%
Graduate diploma, graduate certificate	2.96% (8)	9.45% (29)	+6.49%
Bachelor	18.89% (51)	21.50% (66)	+2.61%
Certificate, advanced diploma, diploma	29.66% (80)	42.32% (130)	+12.66%
No non-school education	36.67% (99)	14.33% (44)	-22.34%
Pearson's chi-squared (independence of observations)	$\chi^2=45.228$ df(4); $p < 0.001$		
Sample size	270	307	
Non-school education (frequency) high level, low level ^a	CV	DCE	Difference
High level (1)	33.33% (90)	43.97% (135)	+10.64%
Low level (0)	66.67% (180)	56.03% (172)	
Proportion	0.33	0.44	
95% confidence interval	0.27-0.39	0.38-0.50	
Pr-test	$z = -2.705$; $p = 0.007$		
Sample size	270	307	

^a 'High' is defined as bachelor degree, graduate diploma and graduate certificate, postgraduate degree; 'low' is defined as advanced diploma and diploma, certificate, no non-school education

Table 41: Comparison of household income – CV(replicated sample):DCE-RT

Household income (\$)	CV	DCE	Difference
Mean	84,216	109,340	
Standard deviation	59,876	47,470	
t-test	t = -5.361 df(529); p < 0.001		
95% confidence interval	76,960-91,472	103,620-115,060	
Distribution of income groups (frequency)			
Nil income	6.82% (18)	0.37% (1)	-6.45%
1-7,799	1.52% (4)	0.00% (0)	-1.52%
7,800-12,999	0.76% (2)	1.12% (3)	+0.36%
13,000-18,199	3.79% (10)	1.50% (4)	-2.29%
18,200-25,999	6.06% (16)	3.00% (8)	-3.06%
26,000-33,799	6.06% (16)	2.25% (6)	-3.81%
33,800-41,599	3.79% (10)	2.25% (6)	-1.54%
41,600-51,999	9.09% (24)	3.75% (10)	-5.34%
52,000-62,399	4.92% (13)	4.49% (12)	-0.43%
62,400-72,799	6.82% (18)	8.24% (22)	+1.42%
72,800-88,399	6.06% (16)	6.37% (17)	+0.31%
88,400-103,999	11.36% (30)	13.48% (36)	+2.12%
104,000-129,999	10.23% (27)	20.22% (54)	+9.99%
130,000-155,999	7.58% (20)	13.11% (35)	+5.53%
156,000-181,999	6.44% (17)	12.36% (33)	+5.92%
182,000-207,999	3.03% (8)	4.12% (11)	+1.09%
208,000 and more	5.68% (15)	3.37% (9)	-2.31%
Sample size	264	267	

Table 42: Comparison of gender– CV(replicated sample):DCE-RT

Gender	CV	DCE	Difference
Male (1)	42.28% (115)	43.32% (133)	-
Female (0)	57.72% (157)	56.68% (174)	
Proportion	0.42	0.43	
Pr-test	z = -0.243 ; p = 0.808		
95% confidence interval	0.36-0.48	0.37-0.49	
Sample size	272	307	

Based on these results, the null **Hypothesis 2.1** can be rejected. The causes for the discrepancy in equivalent surplus across the two methods remain, however, unknown.

10.3 Prominence in response in open-ended contingent valuation

Potential effects of prominence in response in open-ended CV on marginal WTP estimates for the preservation of Nadgee Nature Reserve (**Hypothesis 2.2**) were investigated by examining the marginal WTP distributions of both the reference sample and the replicated sample for prominent numbers. As illustrated in Table 43, the prominent numbers in both samples are the same with the exceptions of the \$2 and the \$1000 value. The \$2 value is only present in the reference sample whereas the \$1000 value is only found in the replicated sample. Furthermore, 86% of the marginal WTP bids in the reference sample and 89% of the marginal WTP bids in the replicated sample can be classified as prominent numbers (see Table 43). This suggests a strong presence of prominence in response in both samples. A test on equal proportions shows that the proportion of prominent numbers across the two samples is not statistically significantly different ($p = 0.358$). This result indicates that the extent of prominence in response in the reference sample is not different from that in the replicated sample.

Table 43: Frequency distribution of prominent numbers

Prominent number	Frequency			
	1979		2009	
\$1	16	8.1%	1	0.5%
\$2	25	12.7%	0	0.0%
\$5	31	15.7%	12	5.7%
\$10	45	22.8%	24	11.3%
\$20	27	13.7%	38	17.9%
\$50	23	11.7%	64	30.2%
\$100	19	9.6%	46	21.7%
\$200	7	3.6%	12	5.7%
\$500	4	2.0%	8	3.8%
\$1000	0	0.0%	7	3.3%
Total	197		212	
Percentage in sample	86%		89%	
pr-test	z = -0.918; p = 0.358			

Hypothesis 2.2 is further tested by exploring whether the distribution of prominent numbers is stable across the two samples (see Figure 11). A Mann-Whitney test of equality of distributions suggests that the distributions of the two samples are statistically significantly different at the 1% confidence level.

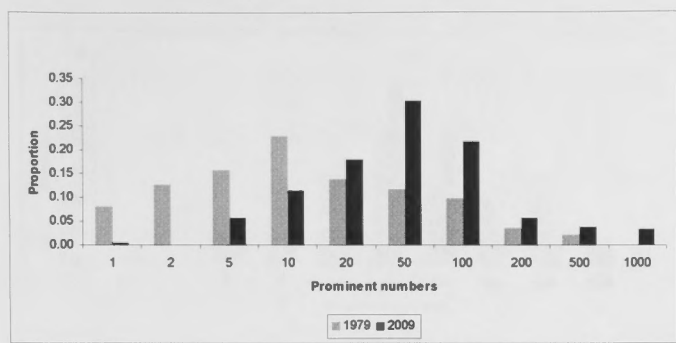


Figure 11: Distribution of prominent numbers inflation-unadjusted

These test results change, however, after adjusting the prominent numbers of the reference study by the inflation over the last 30 years and regrouping these adjusted numbers into prominent number categories by using the boundaries suggested by the theory of prominence (see Table 44, and Figure 12). A Mann-Whitney test of equality of distributions indicates that the distributions of the two samples are no longer statistically significantly different ($p = 0.269$). This result suggests that the distribution of prominent numbers found in the replicated study might have shifted relative to the reference study by the rate of inflation. That is, respondents in 2009 might have used the same prominent numbers to the same extent as 30 years ago but adjusted them by inflation. Hence, a change in marginal WTP would remain undetected if this change was too small to trigger a jump from one to another prominent number. A change in the true underlying distribution of respondents' marginal WTP within the interval bounds of a prominent number would be ignored in the analysis. This true underlying distribution, of course, is unknown to the analyst.

Table 44: Regrouped prominent number categories

Prominent number	Intervals ^a	Prominent number*CPI
5	3.5 - 7.5	4
10	7.5 - 15.0	8
20	15 - 35	19
50	35 - 75	38
100	75 - 150	76
200	150 - 350	191
500	350 - 750	382
1000	750 - 1500	764
2000	1500 - 3500	1911

^a Interval bounds are defined as $IB = (a_1 + a_2)/2$, with a_1, a_2 being subsequent prominent numbers following the sequence $a \times 10^i : a \in \{1,2,5\}$, with i being an integer.

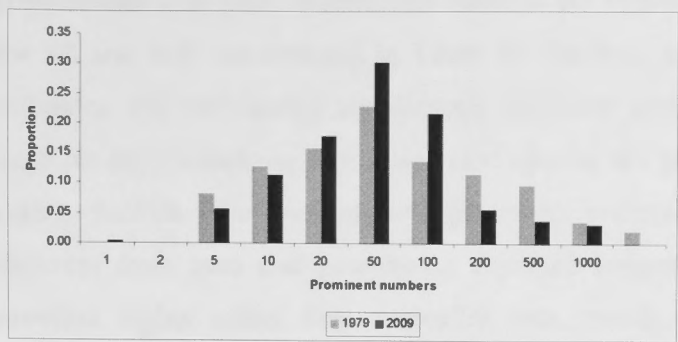


Figure 12: Distribution of prominent numbers inflation-adjusted¹¹⁰

These results suggest that prominence in response is present in and stable across both samples. Clearly, the theory of prominence implies that temporarily stable prominence in response might have stabilised temporally variable marginal WTP estimates for the preservation of Nadgee Nature Reserve. However, the extent of such a stabilising effect cannot be elicited. In fact, the stabilising effect may be zero in the case where the marginal WTP bids actually remained the same across the 30-year time period. Therefore, it is not possible to distinguish between genuinely stable marginal WTP estimates and those that are in fact variable but are stabilised by stable prominence in response.

Consequently, even though the results provide evidence that prominence in response is present in both samples and might have had a stabilising effect on the marginal WTP estimates of the preservation of Nadgee Nature Reserve, the null **Hypothesis 2.2** cannot be rejected.

10.4 Incentive compatibility properties of discrete choice experiments¹¹¹

Effects of repeated-binary choice relative to single-binary choice concerned with preservation and recreational values attached to Nadgee Nature Reserve

¹¹⁰ Consumer Price Index Canberra (1979-2009)

¹¹¹ Parts of the literature review presented in Section 4.3.2, parts of the research design presented in Section 7.3.3, parts of the data collection presented in Section 8.3, and parts of the findings presented in Section 10.4 were reported in one article that has been published in *Environmental and Resource Economics* (Scheufele and Bennett, 2012b) and another article that has been revised and re-submitted to *The Australian Journal of Agricultural Economics* (Scheufele and Bennett, 2012a). The candidate conducted all analysis and took the primary role in writing the article, with Bennett providing input at the research design stage and contributions during the writing of the article.

(Hypothesis 2.3) were investigated through the estimation of CL models. The results for SB and RB1 are reported in Table 45. For both split samples, the *cost* parameter estimates are statistically significantly different from zero and have the expected negative sign indicating that lower cost options are preferred to higher cost options, *ceteris paribus*. The *area of land* parameter estimates are statistically significantly different from zero and positive as expected suggesting that a larger area of land provides higher utility than a smaller area, *ceteris paribus*. The *access* parameter estimates, however, are not statistically significantly different from zero in either split sample.

Table 45: CL Model Results for RB1, SB and FirstRB

	‘Repeated-binary’ (RB1)		‘Single-binary’ (SB)		First of ‘repeated-binary’ (FirstRB)	
Variable	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Constant	0.39878*** (0.0039)	0.13803	0.18680 (0.1760)	0.13805	0.00666 (0.9611)	0.13641
Cost	-0.00453*** (<0.0001)	0.00057	-0.00218*** (0.0001)	0.00056	-0.00224*** (0.0001)	0.00056
Area of land	0.01342*** (<0.0001)	0.00214	0.01273*** (<0.0001)	0.00217	0.01291*** (<0.0001)	0.00206
Access	0.00059 (0.9914)	0.05430	-0.00537 (0.9211)	0.05426	0.05640 (0.2887)	0.05316
Model statistics						
N (observations)	1468		1445		1483	
LL _{ASC}	-1015.9650		-991.9259		-1279.8540	
LL _β	-962.2373		-965.8225		-995.7384	
χ ² ₃	107.46 (<0.0001)		52.21 (<0.0001)		568.23 (<0.0001)	
Adjusted ρ ²	0.05		0.02		0.22	
AIC	1.29595		1.34231		1.34826	
BIC	1.33560		1.35692		1.36257	

Note: Asterisk (*), double asterisk (**) and triple asterisk (***) denote variables significant at 10%, 5% and 1% respectively. P-values are displayed in parentheses.

The marginal WTP estimates for SB and all RB split samples are reported in Table 46. Equivalence of marginal WTP estimates was tested using a procedure suggested by Poe et al. (2005; 1994). The test results suggest a statistically significantly higher marginal WTP for SB than for RB1, RB2, RB3, and RB4.

Table 46: Marginal WTP estimates for RB1, RB2, RB3, RB4, and SB

	Marginal WTP _{area of land}	Standard error	CI(95%) _{marginal WTP}	p-value RB-SB (Test procedure by Poe)
RB1	\$3.00*** (<0.0001)	0.61584	\$1.88-4.34	0.0566*
RB2	\$2.41*** (<0.0001)	0.50173	\$1.48-\$3.46	0.0145**
RB3	\$2.51*** (<0.0001)	0.52867	\$1.62-\$3.58	0.01808**
RB4	\$2.99*** (<0.0001)	0.54493	\$2.03-\$4.17	0.04636**
SB	\$6.44** (0.0203)	2.7705	\$3.42-\$13.08	-

Note: Asterisk (*), double asterisk (**) and triple asterisk (***) denote variables significant at 10%, 5% and 1% respectively. P-values are displayed in parentheses. 95% confidence intervals in parentheses are based on the 2.5th and 97.5th percentile of the simulated marginal WTP distribution. In comparison to the delta method, this method does not imply a normal distribution.

Differences in the attributes and scale parameter estimates between SB and the RB split samples were explored using a test procedure suggested by Swait and Louviere (1993)^{112,113}. The results are displayed in Table 47. The tests results suggest statically significant differences in attribute parameter estimates comparing SB to RB2, RB3, and RB4 at the 1%-level.

A statistically significant difference in an attribute parameter estimate prevents a test for scale parameter estimate equality¹¹⁴. Hence, only RB1 was tested in this regard. The hypothesis of equal scales was rejected ($p_{scale}=0.0124$). The reduced relative scale parameter for SB suggests a less accurate choice since the scale parameter is inversely related to the variance of the error term. Smaller relative scale parameters for SB, as opposed to RB1, indicate that the difference in marginal WTP is mainly induced by differences in the variance of the error term. This finding indicates that repeated choices increase choice accuracy. However, it has to be noted that this hypothesis could not be tested for the RB2, RB3, and RB4 split samples.

¹¹² The relative scale parameter was estimated based on a heteroscedastic multinomial logit model in STATA 10 (for example, Hensher et al., 1999).

¹¹³ For a detailed discussion about the Swait-Louviere test and the confounding influence of the scale parameter in multinomial logit models see Louviere and Eagle (2006).

¹¹⁴ Parameter vector and scale parameter are confounded in MNL models.

Table 47: Test results for Attribute and Scale Factor Equality across RB split samples and SB

LL RB	LL SB	LL Pooled ^a	LR-test ^b (5d.f.)	Reject $H_0:\beta_i=\beta_j$	Scale ratio λ_i/λ_j	LL Pooled ^c	LR-test ^d (1d.f.)	Reject $H_0:\lambda_i=\lambda_j$
-962.2373 RB1	-965.823	-1931.128	0.2932	no	0.7795	-1934.252	0.0124	yes
-965.627 RB2	-965.823	-1942.752	0.0004	yes	0.6087	-1953.900	-	-
-952.739 RB3	-965.823	-1930.921	<0.0001	yes	0.4694	-1949.737	-	-
-961.4410 RB4	-965.823	-1937.9	0.0007	yes	0.4609	-1957.801	-	-

^a Pooled CL model allowing varying scale parameter

^b Log-likelihood ratio test; test statistics $-2(LL_{\text{pool}}-(LL_1+LL_2))$ with d.f. $k+1$, where k is the number of parameters including the constant; asymptotically chi-squared distributed

^c Pooled CL model assuming equal scale parameter in both split samples

^d Log-likelihood ratio test, test statistics $-2(LL_{\text{equalscale}}-(LL_{\text{varyingscale}}))$ with 1 d.f. is asymptotically chi-squared distributed

Overall, these findings suggest a rejection of the null **Hypothesis 2.3**.

The effect of awareness of having repeated choice concerned with preservation and recreational values attached to Nadgee Nature Reserve (**Hypothesis 2.4**) was tested by comparing the CL results of SB with those of FirstRB. The attribute parameter estimates for SB and FirstRB are reported in Table 45. For both split samples, the *cost* parameter estimates are statistically significantly different from zero and have the expected negative sign, indicating that lower cost options are preferred to higher cost options, *ceteris paribus*. The *area of land* parameter estimates is statistically significantly different from zero and positive as expected suggesting that a larger area of land provides higher utility than a smaller area, *ceteris paribus*. The *access* parameter estimates, however, are not statistically significantly different from zero in either split sample.

The test procedure suggested by Poe (2005) did not reveal statistically significant differences in marginal WTP ($p=0.99$) between SB (\$6.43) and FirstRB (\$6.20) and the 95% confident intervals are similar (SB:\$3.42-\$13.08; FirstRB:\$3.36-\$11.56)¹¹⁵. Furthermore, an investigation of differences in the attribute and scale parameter estimates between SB and FirstRB (see Table 48) did not reveal statistically significant differences in either ($p_{\text{attribute}}=0.5785$, $p_{\text{scale}}=0.5179$). These results indicate that the

¹¹⁵ 95% confidence intervals in parentheses based on the 2.5th and 97.5th percentile of the simulated WTP distribution. In comparison to the delta method, this method does not imply a normal distribution.

knowledge of the possibility of making multiple choices does not affect choice outcomes.

Table 48: Test results Attribute and Scale Factor Equality for SB and FirstRB

LL SB	LL FirstRB	LL Pooled ^a	LR-test ^b (5d.f.)	Reject $H_0:\beta_i=\beta_j$	Scale ratio λ_i/λ_j	LL Pooled ^c	LR-test ^d (1 d.f.)	Reject $H_0:\lambda_i=\lambda_j$
-965.823	-995.738	-1963.461	0.5785	no	1.1062	-1963.252	0.5179	no

^a Pooled CL model allowing varying scale parameter
^b Log-likelihood ratio test, test statistics $-2(LL_{\text{pool}} - (LL_1+LL_2))$ with d.f. $k+1$, where k is the number of parameters including the constant is asymptotically chi-squared distributed
^c Pooled CL model assuming equal scale parameter in both split samples
^d Log-likelihood ratio test, test statistics $-2(LL_{\text{equalscale}} - LL_{\text{varyingscale}})$ with 1 d.f. is asymptotically chi-squared distributed

These findings do not justify a rejection of the null **Hypothesis 2.4**, which contradicts the research of Bateman et al. (2008). No evidence that awareness of having repeated choices may induce strategic behaviour was found. One explanation for these contradicting results is the possible influence of the amount of information given to respondents before they made their choices. In the study of Bateman et al. (2008), respondents were informed about the number of choices, the type, and the levels of the attributes, whereas in this study only the number of choices and the attribute types but not the attribute levels were communicated. In the study performed for this thesis, respondents were not informed about the attribute levels before they engaged into the choice tasks. That is, the information given in this study might not have been sufficient to trigger strategic behaviour.

To explore whether strategic opportunities provided by the choice set order affect choice decisions concerned with preservation and recreational values attached to Nadgee Nature Reserve (**Hypothesis 2.5**), *strategy-type*con* variables were incorporated into a pRPL model estimation (see Table 49). The *cost*, *area of land*, and *access* parameter estimates are statistically significantly different from zero at the 1% level and have the expected signs. The model fit improved after the inclusion of these variables ($p<0.0001$)¹¹⁶. The parameter estimate *first*cost* is statistically significantly different from zero at the 1% level and positive as expected. The parameter estimates of *max*cost* and *none*cost* are statistically significantly different from zero at the 1% level

¹¹⁶ Log-likelihood ratio test ($-2[LL_r-LL_{ur}]$)

and negative as expected. The parameter estimate *min*cost* is not statistically significantly different from zero.

By comparing the parameter estimates of *min*cost*, *max*cost*, and *none*cost* to the assumed incentive compatible *first*cost* parameter estimate an increasing cost sensitivity and a decreasing marginal WTP can be observed: $first > min > none > max$. All parameter estimates are different from each other at the 5% significance level. This relative relationship further indicates that the *strategy-type* of the choice set may influence the cost sensitivity of respondents. However, contrary to expectations, offering a choice question in the *min* category compared to one in the incentive compatible *first* category does not decrease cost sensitivity. This suggests that respondents treat the first choice question differently because of its position at the beginning of the choice set compared to the following ones.

By comparing the parameter estimates of *min*cost*, *max*cost*, *none*cost*, and *first*cost* with the weighted average of the *strategy-type*cost* parameter estimates based on the sample averages of the *strategic-type*cost* variables (-0.00016)¹¹⁷, we find that respondents being presented with a choice set in the *first* category are less cost-sensitive and thus have a higher marginal WTP than they are on average at the 1% significance level, whereas respondents being presented with a choice set in the *none* and the *max* category are more cost-sensitive and thus have a lower marginal WTP than they have on average at the 1% significance level. Respondents being presented with a choice set in the *min* category are similarly cost-sensitive and thus have a similar marginal WTP to what they have on average ($p=0.1989$).

¹¹⁷ *Strategy-type*cost* weighted average = $(-0.01821*0.00187)+(0.03287*-0.00516*)+(-0.01467*-0.00265) = -0.00016$

Table 49: Model results of pRPL estimation – *strategy-type* and *positions* in the sequence separately

	<i>Strategy-type</i> ^a		<i>Position</i> ^b	
Variable	Coefficient ^c	Standard error	Coefficient ^c	Standard error
<i>Nonrandom parameters</i>				
constant	0.49984*** (0.0052)	0.17887	0.80284*** (<0.0001)	0.14436
<i>Random parameters</i>				
cost	-0.01893*** (<0.0001)	0.00193	-0.02161*** (<0.0001)	0.00178
area of land	0.05504*** (<0.0001)	0.00554	0.05372*** (<0.0001)	0.00547
access	0.19760*** (0.0044)	0.06943	0.20869*** (0.0029)	0.06995
<i>Standard deviations</i>				
cost (n)	0.02298*** (<0.0001)	0.00201	0.02172*** (<0.0001)	0.00188
area of land (n)	0.09364*** (<0.0001)	0.00830	0.09536*** (<0.0001)	0.00888
access (n)	1.26530*** (<0.0001)	0.17309	1.33832*** (<0.0001)	0.18433
<i>Heterogeneity in mean</i>				
min*cost	0.00187 (0.2363)	0.00158		
max*cost	-0.00516*** (<0.0001)	0.00111		
none*cost	-0.00265*** (0.0024)	0.00087		
position-2*cost			-0.00096 (0.2015)	0.00075
position-3*cost			-0.00360*** (<0.0001)	0.00079
position-4*cost			-0.00214*** (<0.0001)	0.00077
<i>Model statistics</i>				
N (observations)	5932		5932	
LLβ	-3178.505		-3181.972	
χ^2_{3} (inclusion vs. exclusion of <i>strategy-type</i> * <i>cost</i> variables)	95.610 (<0.0001)			
χ^2_{3} (inclusion vs. exclusion of <i>position</i> * <i>cost</i> variables)			88.766 (<0.0001)	
AIC	1.07502		1.07619	
BIC	1.08629		1.08746	

^a One of the four *strategy-type***cost* variables (*first***cost*) was omitted from the model estimation. However, it was estimated in a separate pRPL model with changed underlying coding: parameter estimate (0.00594***), p-value (0.0000), standard error (0.00091).

^b One of the four *position*cost* variables (*position-1*cost*) was omitted from the model estimation. However, it was estimated in a separate pRPL model with changed underlying coding: parameter estimate (0.00670***), p-value (0.0000), standard error (0.00086).

^c ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses

To test whether choice decisions concerned with preservation and recreational values attached to Nadgee Nature Reserve are stable across a sequence of DCE choice tasks (**Hypothesis 2.6**), *position*cost* variables were incorporated into the pRPL model estimation. The results are presented in Table 49. The *cost*, *area of land*, and *access* parameter estimates are statistically significantly different from zero at the 1% level and have the expected signs. The model fit improved after including the *position*cost* variables ($p < 0.0001$)¹¹⁸. The *position*cost* parameter estimates are statistically significantly different from zero for the first, third and fourth choice question at the 1% level. A comparison of the parameter estimates indicates that the first is statistically significantly different from the second (1% level), the second is statistically significantly different from the third (1% level), but the third is not statistically significantly different from the fourth. The magnitude of the parameter estimates decreases along the sequence of choice questions at a decreasing rate.

These findings suggest that the cost sensitivity increases (and thus the marginal WTP decreases) as respondents progress through the choice task, with the extent of increase (decrease) lessening as more choice questions are answered. These findings are supported by comparing the parameter estimates of *position-1*cost*, *position-2*cost*, *position-3*cost*, and *position-4*cost* with the weighted average of the *position*cost* parameter estimates based on the sample averages of the *position*cost* variables, which is zero¹¹⁹. The results suggest that respondents are less cost-sensitive and thus have a higher marginal WTP in the first choice question than on average but are more cost-sensitive, and thus have a lower marginal WTP in the third and fourth choice question than on average at the 1% level.

To test Hypothesis 2.6 further, the bid acceptance rates (proportion of chosen non-zero cost options) related to choice questions in the first, second, third, and fourth choice set position (P1, P2, P3, and P4) were investigated. The bid acceptance rate is 53% for P1, 46% for P2, 42% for P3, and 44% for P4. The bid acceptance rate of P1 is statistically

¹¹⁸ Log-likelihood ratio test ($-2[LLR-LLur]$)

¹¹⁹ Respondents answered the same number of choice questions in the first, second, third and fourth position. Hence, *position*cost* weighted average is zero.

significantly different from P2, P3, and P4 at the 1% level. The bid acceptance rate of P2 is not statistically significantly different from P3, and the bid acceptance rate of P3 is not statistically significantly different from P4. Bid acceptance curves (proportion of chosen non-zero cost options evaluated at different cost levels) were further explored.

As pointed out in Section 7.3.3, the research design ensures that the choice sets presented in P1, P2, P3, and P4 are essentially the same. Hence, there is no confounding influence of varying attribute levels. Bid acceptance curves for P1, P2, P3 and P4 are displayed in Figure 13. This figure shows choice sensitivity to the relative cost levels within P1, P2, P3, and P4, with bid acceptance rates declining with increasing cost levels. As expected, the bid acceptance curve of P1 lies above those of P2, P3, and P4. Statistically significant differences at the 1% level are observed between P1 and P2, P1 and P3, and P1 and P4 at the \$200 and \$300 cost levels¹²⁰. This indicates that non-zero cost options with higher cost levels were chosen more often in the first choice than in the following ones. That is, respondents may become more sensitive towards higher cost levels when moving from the first to the second choice task.

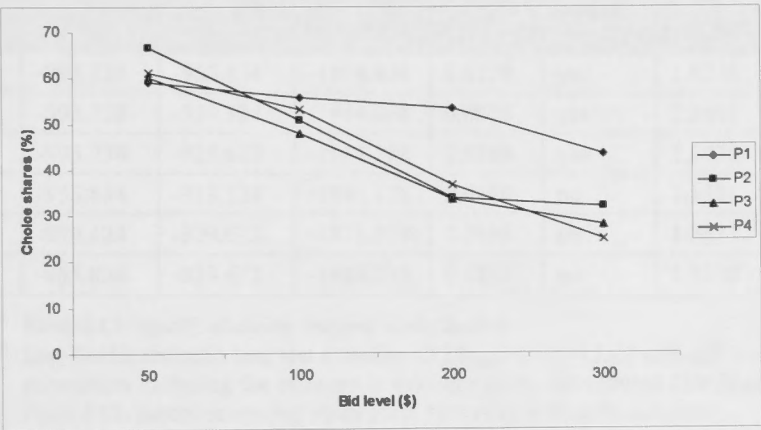


Figure 13: Bid acceptance curves

Additionally, parameter estimates and scale factors were compared across P1, P2, P3, and P4. This analysis is restricted to the CL models since for single choice data it is not possible to discover random parameters that are statistically significantly different from

¹²⁰ Statistically significant differences at the 5% level are also observed between P1 and P2 at the \$50 cost level, and between P1 and P3 at the \$100 cost level.

zero¹²¹. The *cost* parameter estimates are statistically significantly different from zero and have the expected negative signs. The *area of land* parameter estimates are statistically significantly different from zero and positive as expected in P1, P2, P3, and P4 at the 1% level. The *access* parameter estimates are not statistically significantly different from zero across all four choice set positions.

Differences in the *area of land* parameter estimates and scale factors across P1, P2, P3, and P4 were explored using the Swait-Louviere test (1993)¹²². The test results are reported in Table 50. A comparison of P1 with P2, P3, and P4 reveals statistically significant differences in the *area of land* parameter estimates after having made the first choice. A statistically significant difference in a parameter estimate prevents a test for scale factor equality¹²³. However, scale factors of 1.8278, 2.2491, and 2.2928, respectively, provide weak evidence of differences. A comparison of P2 with P3, P3 with P4, and P2 with P4 reveals no statistically significant difference either in the *area of land* parameter estimates or scale factors.

Table 50: Test results for equality for attribute and scale factors between P1, P2, P3 and P4

Position P	LL 1 st split	LL 2 nd split	LL Pooled ^a	LR-test ^b (5d.f.)	Reject H ₀ :β _i =β _j	Scale ratio λ _i /λ _j	LL Pooled ^c	LR-test ^d (1 d.f.)	Reject H ₀ :λ _i =λ _j
1 vs. 2	-995.738	-955.834	-1958.804	0.0129	yes	1.8278	-	-	-
1 vs. 3	-995.738	-939.124	-1944.068	0.0025	yes	2.2491	-	-	-
1 vs. 4	-995.738	-929.612	-1931.596	0.0286	yes	2.2928	-	-	-
2 vs. 3	-955.834	-939.124	-1897.178	0.4880	no	1.1421	-1898.107	0.1729	no
3 vs. 4	-939.124	-929.612	-1871.819	0.2905	no	1.0277	-1871.763	0.7379	no
2 vs. 4	-955.834	-929.612	-1889.235	0.1811	no	1.1770	-1890.164	0.1729	no

^a Pooled CL model allowing varying scale factors
^b Log-likelihood ratio test, test statistics $-2(LL_{\text{pool}} - (LL_1 + LL_2))$ with d.f. $k+1$, where k is the number of parameters including the constant is asymptotically chi-squared distributed
^c Pooled CL model assuming equal scale factors in both split samples
^d Log-likelihood ratio test, test statistics $-2(LL_{\text{equalscale}} - (LL_{\text{varyingscale}}))$ with 1 d.f. is asymptotically chi-squared distributed

A final test investigated whether marginal WTP estimates for *area of land* varied across choice set positions. Implicit prices were estimated using a bootstrapping procedure

¹²¹ A possible explanation is that in the absence of a very large sample it is theoretically impossible to disentangle the assumed distribution of random terms associated with preference parameters or alternatives from the assumed EV1 distribution of the remaining random term that is assumed to be IID across alternatives and individuals (Rose et al., 2009).

¹²² The access parameter being insignificant in P1, P3, and P4 prevents a comparison across choice set positions.

¹²³ Parameter vector and scale factor are confounded in MNL models.

with 1000 draws (Krinsky and Robb, 1986). The marginal WTP estimates for P1, P2, P3, and P4 are displayed in Table 51. The equivalence of marginal WTP estimates was tested using a procedure described in Poe et al. (2002; 2005). The marginal WTP for P1 is statistically significantly different from P2 at the 1% level. The confidence interval of P1 is wider than that for P2. A comparison of the marginal WTP estimates of P2 with P3, and P3 with P4, reveals no statistically significant differences at the 5% level. The confidence intervals are similar for P2, P3, and P4. The differences in confidence intervals and the relative scale factors between P1 and P2 are indications that respondents may make more precise choices after having answered the first one.

Table 51: Marginal WTP of P1, P2, P3, and P4

Position P	Marginal WTP ^a	95% confidence interval Marginal WTP
1	\$6.20*** (0.0075)	\$3.36-\$11.56
2	\$1.87*** (<0.0001)	\$1.12-\$2.72
3	\$2.75*** (<0.0001)	\$1.92-\$3.73
4	\$2.28*** (<0.0001)	\$1.55-\$3.09
<i>Poe tests</i>	<i>p-value</i>	
1 vs. 2	<0.01	
2 vs. 3	0.16	
3 vs. 4	1.56	

^a p-values in parentheses; 95% confidence intervals in parentheses based on the 2.5th and 97.5th percentile of the simulated marginal WTP distribution. In comparison to the delta method, this method does not imply a normal distribution.

Hypothesis 2.5 and 2.6 explore *strategic response* and *choice instability* separately. In order to justify a rejection of either hypothesis, two combined pRPL models were estimated to test whether these two effects coexist in the data. A combined model was estimated for which the *strategy-types* were re-categorised by a binary effects coded nominal *minimum*cost* variable, and a binary effects coded nominal *maximum*cost* variable to avoid perfect multicollinearity. The results of the combined model are represented in Table 52. The *position*cost* parameter estimates are statistically significantly different from zero for the first, third and fourth choice question at the 1% level. A comparison of the parameter estimates indicates that the first is statistically significantly different from the second (1% level), the second is statistically significantly different from the third (1% level), but the third is not statistically significantly different from the fourth. As in the separate model, the magnitude of the *position*cost* parameter estimates decreases along the sequence of choice questions

indicating an increasing cost sensitivity, (and thus a decreasing marginal WTP) at a decreasing rate. The *minimum*cost* and the *maximum*cost* parameter estimates are statistically significantly different from each other (1% level). The *maximum*cost* parameter estimate is statistically significantly different from zero (1% level) and negative as expected. The *minimum*cost* parameter estimate is not statistically significantly different from zero. These results support the findings of the separate model.

Intercept	11.2024*** (4.0104)	11.2024*** (4.0104)
distance	0.00000000*** (0.00000000)	0.00000000*** (0.00000000)
Distance squared (m)	-0.00000000*** (-0.00000000)	-0.00000000*** (-0.00000000)
area of field (m)	0.00000000*** (0.00000000)	0.00000000*** (0.00000000)
number of trees	1.22000000*** (0.4000)	1.22000000*** (0.4000)
Number squared (m)	-0.00000000*** (-0.00000000)	-0.00000000*** (-0.00000000)
position 1*cost	-0.00000000*** (0.00000000)	-0.00000000*** (0.00000000)
position 2*cost	-0.00000000*** (0.00000000)	-0.00000000*** (0.00000000)
position 3*cost	-0.00000000*** (0.00000000)	-0.00000000*** (0.00000000)
position 4*cost	-0.00000000*** (0.00000000)	-0.00000000*** (0.00000000)
minimum*cost	0.00000000 (0.0000)	0.00000000 (0.0000)
maximum*cost	-0.00000000*** (0.00000000)	-0.00000000*** (0.00000000)
Model statistics		
N (observations)	997	997
Chi2	3126.129	3126.129
R2	100.000	100.000
Adjusted R-squared (percentage)	100.000	100.000
Adjusted R-squared (percentage)	100.000	100.000
F(1)	1.00000	1.00000
F(1)	1.00000	1.00000

* One of the four position*cost variables (position 1*cost) was excluded from the model estimation. However, it was estimated in a separate PSM model with a good identifying power (p-value = 0.0067777, p-value = 0.00077, standard error = 0.00000000).

*** is significant at 1% level, ** is significant at 5% level, * is significant at 10% level, + is significant at 15% level.

Table 52: Model results of pRPL estimation – *strategy-type* and *positions* in the sequence combined

Variable	Combined model ^a	
	Coefficient ^b	Standard error
<i>Nonrandom parameters</i>		
constant	0.51478*** (0.0043)	0.18024
<i>Random parameters</i>		
cost	-0.01924*** (<0.0001)	0.00196
area of land	0.05511*** (<0.0001)	0.00555
access	0.19726*** (0.0045)	0.06952
<i>Standard deviations</i>		
cost (n)	0.02341*** (<0.0001)	0.00206
area of land (n)	0.09340*** (<0.0001)	0.00826
access (n)	1.26769*** (<0.0001)	0.17319
<i>Heterogeneity in mean</i>		
position-2*cost	-0.00050 (0.5292)	0.00079
position-3*cost	-0.00361*** (0.0014)	0.00113
position-4*cost	-0.00267*** (0.0266)	0.00120
minimum*cost	0.00120 (0.3039)	0.00118
maximum*cost	-0.00190** (0.0205)	0.00082
<i>Model statistics</i>		
N (observations)	5932	
LL β	-3176.124	
$\chi^2_{.5}$ (Inclusion vs. exclusion of strategy-type and position variables in Approach II)	100.372 (<0.0001)	
AIC	1.07489	
BIC	1.08842	

^a One of the four *position*cost* variables (*position-1*cost*) was omitted from the model estimation. However, it was estimated in a separate pRPL model with changed underlying coding: parameter estimate (0.00677***), p-value (0.0007), standard error (0.00200).

^b ***=significant at 1% level, **=significant at 5% level, *=significant at 10% level; p-values in parentheses;

Overall, these findings justify a rejection of the null **Hypothesis 2.5** and the null **Hypothesis 2.6**.

10.5 Conclusion

Chapter 10 has presented the results of the analysis concerned with exploring whether the evolution of stated preference methods has impacted on marginal WTP estimates of non-marketed benefits attached to environmental goods and services with public good character.

Section 10.2 investigated the impact of alternative stated preference methods on the preservation value of Nadgee Nature Reserve in 2009 by comparing equivalent surplus obtained by an open-ended CV and a state-of-the-art DCE (**Hypothesis 2.1**). The open-ended CV equivalent surplus was computed as the mean value of stated marginal WTP bids, whereas the calculation of the DCE equivalent surplus was based on implicit prices estimated in WTP space. The results suggest that the equivalent surplus estimated by means of the DCE is statistically significantly different from and about four times higher than that obtained through the open-ended CV. Potential drivers that may explain this discrepancy are numerous, including differences in sampling, questionnaire design, survey mode, elicitation format, definition of the environmental goods and services, econometric analysis, or a combination of these factors. It was beyond the capacity of this study to rigorously analyse the individual effects of these factors. This analysis found, however, evidence for a potential stabilising effect of the open-ended CV elicitation format (Section 10.3), and influences related to differences in the incentive compatibility properties of the two alternative elicitation formats (Section 10.4). Both observations may explain at least partially the disparity in equivalent surplus found across the two stated preference methods.

Section 10.3 focused on the analysis of prominence in response related to the open-ended CV elicitation format and its potential stabilising effect on the temporal variability of the preservation value of Nadgee Nature Reserve (**Hypothesis 2.2**). Even though prominence in response in open-ended CV data has been examined before, its potential impact on value variability across time has not been investigated. The results indicate a strong presence of such prominence in response in both the reference and the replicated sample. Furthermore, the proportion of prominence in response and the

inflation-adjusted distributions of prominent numbers suggest that this prominence in response is stable across the two samples, and thus across the two points in time. Hence, stable prominence in response might have stabilised temporally variable marginal WTP estimates for the preservation of Nadgee Nature Reserve. However, it was not possible to differentiate between genuinely stable marginal WTP and marginal WTP that is variable but stabilised by stable prominence in response. Hence, the extent of such a potentially stabilising effect was not identifiable. Furthermore, prominence in response might be confounded by effects associated with behaviour such as habits and impulsive emotions. As stated in Section 7.3.2, differentiating effects of prominence in response from those of non-rational behaviour further limits the explanatory power of this analysis. In an extreme case, stated marginal WTP might be neither genuinely stable nor seemingly stable but rather a result of habits and impulsive emotions completely unrelated to the preservation of Nadgee Nature Reserve. Nevertheless, the presented evidence demonstrates that the open-ended CV elicitation format has the potential to produce values that are seemingly stable across time disguising value changes that are too small to trigger and be represented by a jump from one prominent number to another.

Section 10.4 investigated the potential effect of methodological evolution on marginal WTP estimates by examining incentive compatibility properties of elicitation formats in DCE. The incentive comparability properties of open-ended CV elicitation formats are well known, whereas those of DCE are still not fully understood. Hence, the main objective of this analysis was to explore the effects of repeated choice on choice behaviour. This section firstly explored whether a repeated-binary elicitation format affects choice outcomes (**Hypothesis 2.3**). The results reveal statistically significant differences in attribute parameter estimates comparing choice data obtained by a single-binary with a repeated-binary elicitation format. The results also show that marginal WTP estimates obtained using a single-binary choice format are statistically higher compared to those estimated using a repeated-binary elicitation format.

Possible explanations suggested in the literature include value learning and strategic learning. Yet, smaller relative scale parameter estimates and larger confidence intervals derived from single-binary choice data indicate that the differences in marginal WTP may be mainly induced by differences in the variance of the error term; that is, repeated choices increased the choice accuracy. However, the results presented here are based on

restricted CL models due to the econometric limitations outlined earlier. Bateman et al. (2008) suggest that differences in scale parameters are at least partially a result of preference heterogeneity. That is, a difference in scale induced by preference heterogeneity will vanish if a model specification is used that allows for preference heterogeneity. Additionally, since the comparative analysis was restricted to the use of CL models, the panel character of the repeated-binary elicitation format could not be taken into account and may have influenced the results. Furthermore, the potentially incentive compatible baseline may have been compromised by the respondent perceived reduced consequentiality of the survey format. This may have confounded comparisons across split samples. On the other hand, the incentive properties to answer such follow-up questions truthfully are unknown. That is, the answers to the follow-up questions may themselves be strategically biased and may thus not reflect the actual opinions of respondents. Further testing of these issues is required.

Secondly, Section 10.4 investigated whether awareness of having repeated choices influences choice outcomes (**Hypothesis 2.4**). No evidence was found that awareness of having multiple choices affects choice outcomes. This result contrasts with the findings of Bateman et al. (2008) who found choice set number awareness to be significant. The differences may be explained by the different questionnaire designs. Bateman et al. (2008) provided respondents with information about the number of choices, the attributes, and all possible attribute levels, whereas in this study only information about the number of choices and the attributes (including a note explaining that choice options are based on different attribute levels) was given to respondents. One possible explanation is, therefore, that the information provided may not have been sufficiently detailed to trigger measurable strategic behaviour. Further testing is needed to explore this issue. However, this section provided evidence for effects on choice outcomes that may be explained by institutional learning and effects that may be explained by either strategic response or value learning. The choice between a repeated-binary and a single-binary elicitation format may imply a trade-off between decreased choice accuracy with the single-binary choice format and potentially increased strategic behaviour under the incentive incompatible repeated-binary choice format.

Thirdly, Section 10.4 explored whether strategic opportunities provided by the order in which choice sets are presented to respondents affect choice decisions (**Hypothesis 2.5**) and whether strategic response coexists with other effects discussed in the literature

such as strategic learning, value learning, and institutional learning (**Hypothesis 2.6**). This section provided evidence that strategic opportunities provided by the order in which choice sets are presented to respondents affect choice decisions (strategic response). Increasing cost sensitivity and decreasing marginal WTP were observed in the following order: first > min > none > max. The results suggest, however, that respondents may react more strongly to high cost rather than low cost inconsistencies: respondents are more cost-sensitive, and thus have a lower marginal WTP than on average, if the same or a similar level of provision was offered in a previous choice set at a lower cost. Yet, the cost sensitivity, and thus the marginal WTP, remains unaffected compared to the average marginal WTP if the same or a similar level of provision was offered in a previous choice set at a higher cost.

The results also propose that choice decisions are not stable across a sequence of choice tasks. The findings indicate that cost sensitivity increases (and thus marginal WTP decreases), when respondents progress through the choice task, with this increase (decrease) lessening as more choice questions are answered. Possible explanations suggested in the literature are value learning (respondents ‘discover’ their true underlying preferences through a learning process affecting parameter estimates) and strategic learning (respondents become increasingly familiar with the strategic opportunities provided by the choice tasks they face and adjust their choices accordingly without adjusting their preferences). More research is required to support either or both of these explanations.

The results further indicate that respondents might treat the first choice question differently because of its position at the beginning of the choice sequence compared to the following ones. Non-zero cost options, especially the ones with higher cost levels, were chosen more often in the first choice than in the following ones. This suggests that respondents may become more sensitive towards higher cost levels when moving from the first to the second choice task. This finding is supported by the observed decreasing marginal WTP estimates when moving from the first to the second choice question. Again, one possible explanation proposed in the literature is value learning. Additionally, our findings support previous results that participants of repeated-binary DCE may make more accurate choices as they progress from the first to the second choice question. One possible explanation for this effect suggested in the literature is institutional learning.

These findings provide a dilemma. If respondents choose strategically (strategic response) and increasingly become aware of and take advantage of strategic opportunities (strategic learning), we may conclude that the first response most closely represents respondents' real values. However, if respondents learn about their values and the choice task after having answered the first choice question, we may conclude that the first choice question should be excluded. Of course, it is possible that some respondents respond strategically while others learn about and adjust their values. Consequently, it cannot be determined which response most closely represents respondents' real values and it remains unclear which responses should be used. Further research is required to disentangle these effects.

In summary, Chapter 10 has presented evidence that the methodological evolution of CV and DEC has an effect on marginal WTP estimates for the preservation of Nadgee Nature Reserve. The discrepancy of equivalent surpluses obtained through the two alternative stated preference methods may be explained partially by effects of the evolution of elicitation formats related to changes in their incentive compatibility properties and prominence in response.

11 CONCLUSION

11.1 Discussion of research findings

The research findings presented in Chapter 9 indicate that the marginal value respondents from Canberra attached to the preservation of Nadgee Nature Reserve in 1979 was not significantly different from comparable estimates made in 2009. However, the results suggest multiple possible explanations for this apparent temporal value stability across a 30-year time period.

One interpretation is focused on shifts in the magnitude of demand side factors and changes in their underlying relationships with marginal WTP. Both of these types of changes might have affected either or both preferences and budget constraints. Since these factors may work in opposite directions a stable marginal WTP on the demand side could be explained by a cancellation of their impacts. This explanation, however, would imply that the observed changes in supply side factors were irrelevant. That is, neither the change in the availability of close substitutes and associated relative scarcity nor reference point relativity had any impact on marginal WTP.

A second possible interpretation of the results involves a proportional change in both demand and supply side factors, with the direction of these changes being dependent on the definition of close substitutes and their perceived relative scarcity. Possible definitions include protected terrestrial natural areas and terrestrial natural areas independent of their protection status. The numbers and extent of protected terrestrial natural areas has increased over the last 30 years. Hence, respondents who defined close substitutes as protected terrestrial natural areas might have perceived them as being relatively abundant. A perceived low relative scarcity of available close substitutes in 2009 compared to 1979 would have been reflected by a decrease in marginal WTP. Hence, a perceived increase in supply would imply an increase in demand. An increase in marginal WTP on the demand side might have been caused by either or both an expansion of the competitive budget set and a preference change. The latter might indicate that respondents were less concerned about the preservation of Nadgee Nature Reserve in 2009 than they were in 1979. On the contrary, respondents might have

defined close substitutes as terrestrial natural areas irrespective of their protection status. In this case, respondents might have perceived close substitutes as becoming relatively scarcer as indicated by ongoing developments over the last 30 years. A perceived high relative scarcity of available close substitutes in 2009 compared to 1979 would have been expressed by an increase in marginal WTP. Hence, a decrease in supply would imply a decrease in demand caused by a change in preferences that exceeds the effect of a possible expansion of the competitive budget set. A preference change might indicate that respondents were significantly more concerned about the loss of Nadgee Nature Reserve in 2009 than they were in 1979. Of course, a mixture of both definitions distributed over respondents might have – in the extreme case – nullified the effects associated with a change in the perceived availability of close substitutes and their perceived relative scarcity. This would support the first explanation, suggesting that supply side factors were irrelevant.

A third possible explanation for the observed temporal stability of the preservation value estimates of Nadgee Nature Reserve is related to the limited representativeness of the change in sample characteristics compared to the changes suggested by the census data across the 30-year time period. The magnitudes of those variables that were found in the regression analysis to have influenced marginal WTP changed in the direction as suggested by the census, but not to the same extent. Hence, sampling might have misrepresented the effects of demand side factors, and thus biased marginal WTP for the preservation of Nadgee Nature Reserve.

The analysis of temporal value variability presented in Chapter 9 also included an investigation of respondents' preferences with respect to the payment vehicle. The evidence shows that respondents' attitudes towards the payment vehicle changed across the 30-year time period. In 1979, the majority of respondents equally preferred exclusively either the government or the conservation organisation to collect the required payment. In contrast, the majority of respondents in 2009 preferred a mixed payment vehicle. This preference change might be interpreted as a changed risk spreading strategy. Respondents in 1979 trusted either the government or the conservation organisation by themselves. In 2009, respondents appeared to be reluctant to put their trust in just one institution. Whether or not this change influenced the temporal variability of marginal WTP for the preservation of Nadgee Nature Reserve could not be tested rigorously since only respondents who stated a positive marginal

WTP were asked the question about the payment vehicle. However, the fact that respondents were offered a choice might have reduced such an impact.

Finally, Chapter 9 explored state-of-the-art econometric models theoretically suitable to fit open-ended CV data. However, no superior model over a log-linear model using an OLS estimator and excluding zero values could be found, which disclosed the difficulties with analysing open-ended CV data. This raises the question of whether potential future econometric models will be able to overcome these difficulties and whether applying those models would yield different results, leading to different conclusions about value variability across time.

The research findings presented in Chapter 10 suggested that the evolution of stated preference methods and the associated evolution of elicitation formats over the last 30 years might have impacted on the temporal variability of value estimates related to the preservation of Nadgee Nature Reserve. The equivalent surplus elicited in 2009 by repeated-triple DCE was about four times higher than the equivalent surplus estimated using open-ended CV. Therefore, if the equivalent surplus in 1979 was estimated through repeated-triple DCE it is likely to have been about four times higher than the one estimated by open-ended CV. This, of course, is based on the assumption that the difference caused by the different methods was constant from 1979 to 2009. Hence, values estimated in the past might have to be adjusted to what is currently regarded as the state-of-the-art stated preference method before they can be used in CBA. Since the notion of state-of-the-art is temporally dynamic a continuous adjustment would be required.

Chapter 10 argued that it was beyond the scope of this thesis to differentiate the reasons for the observed differences in equivalent surplus across the two stated preference methods. Nevertheless, the results suggest that the evolution of the elicitation format associated with the evolution of stated preference methods might be an important driver for changes in value estimates. The results presented in Chapter 10 indicate that the open-ended CV data used in this study exhibited prominence in response that was stable across 1979 and 2009. That is, the open-ended CV elicitation format might have stabilised and thus disguised temporal value variability. This adds another possible explanation for the observed temporal stability of the preservation value of Nadgee Nature Reserve. The extent of such a potentially stabilising effect of the open-ended CV elicitation format was not differentiable from other factors driving temporal value

variability or stability. However, the evidence indicated that value estimates based on open-ended CV might be biased and could potentially confound value change across extended time periods.

Chapter 10 also provided evidence that the evolution of elicitation formats and associated changes in incentive compatibility properties that have accompanied the evolution of stated preference methods from open-ended CV to repeated-multiple DCE have affected marginal value estimates for the preservation of Nadgee Nature Reserve. The results suggest that marginal WTP estimates obtained by a single-binary DCE are higher than those elicited by a repeated-binary DCE. Additionally, the results indicate that the order in which the choice sets are presented to respondents affects choice decisions. Possible explanations for this observed value estimate difference supported by the research findings presented in detail in Chapter 10 are strategic response, strategic learning, institutional learning, and value learning. Together, the results suggest that repeated elicitation formats predominantly used in DCE exhibit incentive properties that are different from those of an open-ended or dichotomous CV. Consequently, one would expect that the methodological evolution from an open-ended CV towards a repeated DCE has had an impact on value estimates across time. Strategic response, strategic learning and institutional learning might have the potential to confound an analysis of temporal value variability.

In summary, this thesis has provided evidence that estimates of the preservation value of Nadgee Nature Reserve used as an example for a value related to non-marketed environmental goods and services were stable across a 30-year time period. Possible explanations for this temporal stability can be distinguished into conceptual and methodological causes as well as causes related to sampling. However, the discussion presented in Section 11.1 demonstrates that the interpretation of the research findings is speculative at this stage and firm conclusions cannot be drawn. This emphasises the importance of setting out ways in which these speculations can be confirmed or rejected as testable hypotheses.

Challenges and limitations that limit the explanatory power of the presented analysis include the lack of supply side information, challenges associated with sampling, difficulties regarding the econometric analysis of open-ended CV data, and problems with differentiating the effects of the evolution of elicitation formats from other methodological impacts. These challenges and limitations will be further elaborated in

Section 11.2. Nevertheless, this analysis has allowed first insights into a complex topic that will be supportive of further research (see Section 11.3.) and policy making concerned with decisions involving non-marketed environmental goods and services (see Section 11.4).

11.2 Challenges and limitations

The discussion of the research findings presented in Section 11.1 is based on a range of maintained assumptions. It was beyond the scope of this thesis to predict the probability and consequences of their violation. Nevertheless, this section attempts an assessment. In general, a violation of any of these assumptions might have confounded the analysis of values and valuation across time.

MA 5.1: The applied stated preference method is temporally reliable.

It cannot be excluded that MA 5.1 was violated and, as a consequence, influenced the research findings.

MA 5.2: Changes in the relative prices of composite goods do not influence marginal WTP.

Following Hanemann's separability condition (1988), a violation of MA 5.2 is expected to be unlikely and the consequences are considered negligible.

MA 5.3: The effect of varying perceptions of reference points is negligible.

It cannot be excluded that MA 5.3 was violated. A violation might have impacted on the interpretation of results related to comparing the preservation value of Nadgee Nature Reserve across time. Changes in the perceived availability of substitutes and associated perceived relative scarcity might have impacted on marginal WTP. An increase or decrease in the availability of substitutes and associated relative scarcity might not have been perceived as such if respondents in 2009 accepted a level that respondents in 1979 were concerned about.

MA 6.1: Respondents were unaware that the size of the area of land presented in the questionnaire in 2009 was held constant at 1979 levels and thus differed from its current size.

A violation of MA 6.1 is expected to be unlikely since the name of the area of land was undisclosed. Nevertheless, it cannot be excluded on principle that a violation of MA 6.1 influenced the research findings.

MA 7.1: The potential bias introduced by excluding the protest votes from the analysis is the same for the reference sample and the replicated sample and thus does not confound the comparison.

The proportion of protest votes across the reference and the replicated samples was found to be not statistically significantly different. However, evidence indicates that the proportion of respondents who are subscribers to a conservation organisation is statistically significantly different across protest and non-protest voters in the reference sample but not in the replicated sample. The variable *subscriber to conservation organisation* was statistically significantly different from zero in the 2009 log-linear regression model. Hence, it cannot be excluded that this difference in protest characteristics across the two samples influenced the research findings. Furthermore, the results suggest that the protest motivations are statistically significantly different across the two samples. Hence, if these differences in protest motivation were reflected in marginal WTP of protest respondents, a violation of MA 7.1 would have influenced the research findings.

MA 8.1: Differences in sampling procedure between the reference and the replicated study described did not affect the survey results.

MA 8.2: Differences in the questionnaire design between the reference and the replicated study did not affect the survey results.

MA 8.3: The potential bias introduced by the reference questionnaire design is the same for both the reference survey and the replicated survey, and thus did not confound their comparison.

Violations of MA 8.1, MA 8.2, and MA 8.3 are expected to be unlikely and the consequences are considered negligible since the stated differences in sampling procedure and questionnaire design were minor. Nevertheless, it cannot be excluded on principle that a violation of these assumptions influenced the research findings.

MA 8.4: Bias introduced by potential data entry and coding errors of the reference data did not confound its comparison with the replicated data.

The data from the reference study were only available as an Excel spreadsheet. The original survey questionnaires were no longer available. This might have resulted in data entry and coding errors. However, a violation of MA 8.4 is expected to be unlikely and the consequences are considered negligible since the observed differences of model outcomes between the original estimation and the reproduction were minor. Nevertheless, it cannot be excluded on principle that a violation of this assumption influenced the research findings.

MA 8.5: The marginal differences in wording of the choice questions did not significantly alter choice incentives across the SB, RBs, and FirstRB split samples.

A violation of MA 8.5 is expected to be unlikely and the consequences are considered negligible since the stated differences in wording were minor. Nevertheless, it cannot be excluded on principle that a violation of these assumptions influenced the research findings.

MA 8.6: Respondents did express their preferences for the preservation of Nadgee Nature Reserve as described in the survey questionnaires and not for the preservation of natural areas in general.

It cannot be excluded that MA 8.6 was violated since the area of land was presented to respondents without revealing its identity. A violation of this maintained assumption might have impacted on the research findings.

This assessment suggests that strong maintained assumptions were made regarding the reliability of the open-ended CV method (MA 5.1), reference point relativity (MA 6.1), protest votes (MA 7.1), and preference expression (MA 8.6). Possibilities to account for the latter three factors that would make these maintained assumptions dispensable will be discussed in Section 11.3.

These maintained assumptions were not the only challenges and limitations this study faced. One major challenge was the establishment of a baseline for the comparison across time. The ongoing methodological evolution implies that comparisons across time will always have to access data that were collected by stated preference methods that are not state-of-the-art from the current perspective. Hence, the research findings of

such comparisons will always be questionable. The range of consequences of this dilemma faced in this thesis is outlined in the following paragraphs.

The reference data were collected by an open-ended CV. This elicitation format might have confounded the analysis of temporal value variability mainly through prominence in response and challenges associated with the econometric analysis. A comparison across time based on data obtained by an alternative elicitation format might have produced different results.

Another challenge in the establishment of a baseline originated from the description of the area of land in the reference questionnaire. The area of land was presented to respondents without revealing its identity. This lent it a somewhat hypothetical character. This hypothetical character might have resulted in reduced credibility of the choice scenario and hence in a reduction of survey consequentiality. This reduction of survey consequentiality had to be balanced with concerns about confounding the comparison by changing the description in the replicated questionnaire. The consequences of keeping the original description were considered the lesser of two evils. Additionally, proposing protecting a natural area of land that is already protected would have further reduced survey credibility.

A further limitation was that neither the reference nor the replicated sample was completely representative of the population of Canberra. Hence, the shifts in magnitudes of demand side factors observed across the two samples did not perfectly reflect the magnitude shifts observed in the population such that they only shifted in the same direction. Consequently, the results of the sample based analysis might deviate from the temporal value variability of the population of Canberra.

An additional challenge originated from the limited data availability. The reference study did not collect any supply side information about respondents' definition and perception of close substitutes. However supply side information is crucial to interpret the research findings. Another limiting factor was the availability and validity of attitudinal variables. Even though socio-psychological research suggests the relevance of such variables in explaining marginal values their incentive compatibility properties question the validity of the obtained data. Are observed temporally stable attitudes a reflection of temporally stable attitudes or stable incentive compatibility properties? Similarly, variables capturing habits and impulsive emotions were not available even

though they might influence temporal value variability. That is, the observed value stability across time might have been influenced by stable habits and impulsive emotions that are unrelated to true underlying preferences. Of course, many other unidentified drivers of change not considered in this study might have impacted on the variability of the preservation value of Nadgee Nature Reserve.

Additionally, this study based the analysis of value variability across time on only two points in time. This limited the explanatory power of the research findings. Theory and empirical evidence suggest that extreme events such as large-scale environmental pollution, unusual weather events, or long-term threats may change marginal values people attach to environmental goods and services. Immediate short-term responses to an extreme event may change over time with increased availability and distribution of more information regarding the causes and consequences of that event. Consequently, an analysis based on just two points in time might reflect reactions to such events rather than a long-term change in marginal value. Even though no such extreme events were observed during either the reference or the replicated study, occurrences might have happened during the 30-year time period between them.

Another challenge was differentiating effects of the methodological evolution from drivers of value change on the demand and supply side. Even though this study provided evidence that elicitation formats might impact on the analysis of value variability across time, it was not possible to quantify these impacts. Consequently, it remains unclear to what extent prominence in response, strategic response, strategic learning, value learning, and institutional learning bias estimates of real value change across time.

Finally, this was the first attempt to analyse the value variability related to non-marketed environmental goods and services across extended time periods using the preservation value of Nadgee Nature Reserve as an example. Other examples and other types of non-marketed environmental goods and services might produce different results. Put simply, further investigation is required to gain a more generalised picture.

These challenges and limitations open a range of opportunities for future research, which are outlined in the Section 11.3.

11.3 Recommendations for future research

The research findings and experiences presented in this thesis have delineated prospects for future research to enhance our understanding of value variability related to non-marketed environmental goods and services across extended time periods.

First of all, the database for analysing temporal value variability needs to be expanded beyond the two points in time that were available in this study. Collecting time-series data on a frequent basis (for example, annual, biennial, or every 5 years) over an extended time period (for example, 30 years) would support levelling out 'extreme value' data points. Furthermore, collecting data on extreme events might be a way to differentiate between short-term responses to extreme events and long-term trends in value change across time.

Such a time-series approach could be combined with a research design customised to overcome the dilemma related to the ongoing evolution of stated preference methods. Clearly, the potentially stabilising effect of the open-ended CV elicitation format limits the potential of this method for analysing temporal value variability. Since most state-of-the-art stated preference methods have moved away from using open-ended elicitation formats towards other formats such as n-bound dichotomous choice CV, payment cards CV, or repeated DCE, current methods might be able to reveal that seemingly stable value estimates may in fact be variable. Therefore, the data within the time-series should be collected by both the stated preference method previously used and a current state-of-the-art stated preference method. The use of a state-of-the-art stated preference method also allows applying a state-of-the-art econometric model. Such a model may be capable of dealing with protest votes more appropriately, which would facilitate testing the maintained assumption MA 7.1.

In general, a comparison of results obtained with alternative stated preference methods would allow an analysis of the effect of the ongoing methodological evolution on value estimates. This approach would support the discrimination of impacts caused by methodological evolution on value estimates from value change across time driven by demand and supply side factors. Put simply, estimating overlapping change in value and change in valuation across time would provide the means for calibration. Expanding the overlapping approach through the inclusion of a range of state-of-the-art methods would facilitate distinguishing causal factors of differences in value estimates over time.

This, of course, would be supported by a quantification of impacts related to methodological evolution. Hence, further research in quantifying effects of, for example, alternative elicitation formats on value estimates such as prominence in response, strategic response, strategic learning, institutional learning, and value learning is required. More detailed recommendations in this respect were presented in Section 10.5 and are not repeated here.

Such a research approach could be enhanced further by collecting supply side data including information about how respondents define close substitutes, perceive their availability and assess associated relative scarcity. This would provide information about perceived relative scarcity as well as the reference point relativity across time needed to narrow down possible interpretations of research results. Hence, the maintained assumption MA 6.1 could be tested. Such information could be collected through follow-up questions included in the survey questionnaires. However, their incentive compatibility properties are unknown. Further research is required to understand their effect on respondents' answers.

Moreover, a better understanding of potentially confounding effects through impulsive emotions and habits on temporal value variability would be supportive. Expanding the research of Weinberg (1981), Kroeber-Riel and Weinberg (1996), and Fisher and Hanley (2007), data could be complemented by collecting information about decision-making types (cognitive-control, habits, impulsive emotions) to differentiate between economically meaningful marginal WTP estimates and those which do not reveal information about respondents' preferences. This would allow discovering potentially stabilising effects of habits and impulsive emotions on temporal value variability.

Finally, expanding the analysis beyond Nadgee Nature Reserve by investigating temporal value variability for other non-marketed environmental goods and services would empower an interpretation of research findings beyond the bounds of a single case study.

11.4 Policy implications

The insights into value variability related to the preservation of Nadgee Nature Reserve gained through this study will be supportive of environmental decision making that

involves long-term and irreversible consequences. The results presented in this thesis indicate that the estimated preservation value of Nadegee Nature Reserve was the same across a 30-year time period. However, the analysis also provided evidence that the open-ended CV elicitation format might have stabilised value estimates across time through prominence in response. This implies the possibility that the preservation value of Nadegee Nature Reserve might have changed from 1979 to 2009, but was not disclosed by the applied elicitation format.

The possibility of value change across extended time periods indicated by this study questions the current practice of using snap-shot value estimates adjusted by a discount rate to predict future values. This implies that asking about the right point in time for estimating a value misses the point. These snap-shot value estimates are frequently obtained by BT through time. That is, value estimates obtained in the past are used for current CBA. Such a practice might intensify an over- or underestimate of future values. Furthermore, values used in CBA might change over time, with this change not necessarily being monotonic. That is, value change might exhibit periods of value increase and value decrease.

This poses a dilemma. Time-series data for values associated with non-marketed environmental goods and services are currently not available. Establishing such a database would be a costly, long-term project. That is, such a database would not be available in the short term. Nevertheless, pilot projects to improve our understanding of temporal value variability related to non-marketed environmental goods and services would be supportive in deciding whether the benefits of such an investment in time-series analysis measured as having the potential to improve efficiency in resource allocation would outweigh its' cost.

The potentially stabilising effect of the open-ended CV elicitation format emphasises that using data obtained in the past that is based on open-ended CV for current policy decisions is not advisable. Nevertheless, the marginal WTP estimates obtained by the reference study are still listed in the Environmental Valuation Reference Inventory¹²⁴, a benefit transfer database for policy applications. More generally, the research findings presented in this thesis raise awareness that the ongoing evolution of elicitation formats used to estimate marginal WTP that has been triggered through an evolution of stated preference methods might have had a significant impact on value estimates. Hence, care

¹²⁴ <https://www.evri.ca/Global/Splash.aspx>

should be taken when using such values in CBA. Adjusting value estimates obtained through BT across time by the impact related to the ongoing methodological evolution should be considered. This, of course, requires the availability of information regarding the magnitudes of these impacts. Available comparisons across alternative stated preference methods are generally based on a specific case study and might not necessarily be applicable to other contexts.

In summary, this thesis has provided first insights into temporal variability of values related to non-marketed environmental goods and services. These insights warrant future research, which is required to draw further conclusions generally relevant to policy making.

APPENDICES

Appendix A: CV questionnaire of the reference study

See attached CD.

Appendix B: CV questionnaire of the replicated study

See attached CD.

Appendix C: Photo-test material

See attached CD.

Appendix D: Model result comparison (linear OLS regression)

Variable	Reference model	Reproduced model
Number of observations	283	278
Constant (t-value)	11.61 (-)	13.05 (2.05)
Payment vehicle (t-value)	16.34 (3.31)	15.66 (3.10)
Household income (t-value)	0.01 (1.96)	0.01 (1.86)
R-squared	0.05	0.04
F2,280/ F2,275	7.19	6.29

Appendix E: DCE questionnaire of RB split sample

See attached CD.

Appendix F: DCE questionnaire of Pilot 1 split sample

See attached CD.

Appendix G: Results of CL model

```
--> sample ;all$
--> reject; protest=1$
--> Nlogit
      ;lhs=choice,cset,altij
      ;choices=A,B,C
      ;Model:
      U(A) = cost*cost + areap*areap + ac*ac/
      U(B) = con + cost*cost + areap*areap + ac*ac/
      U(C) = con + cost*cost + areap*areap + ac*ac$
Normal exit: 6 iterations. Status=0. F= 1158.104
```

Discrete choice (multinomial logit) model
Dependent variable Choice
Log likelihood function -1158.104
Estimation based on N = 1228, K = 4
Information Criteria: Normalization=1/N

	Normalized	Unnormalized
AIC	1.89268	2324.20739
Fin.Smpl.AIC	1.89270	2324.24010
Bayes IC	1.90933	2344.65996
Hannan Quinn	1.89894	2331.90294

Model estimated: Jun 09, 2011, 14:57:13
R2=1-LogL/LogL* Log-L fncn R-sqrd R2Adj
Constants only must be computed directly
Use NLOGIT ;...; RHS=ONE\$
Response data are given as ind. choices
Number of obs.= 1228, skipped 0 obs

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]
COST	-.00433***	.00057	-7.640	.0000
AREAP	.01928***	.00229	8.425	.0000
AC	.42209***	.05130	8.228	.0000
CON	.98948***	.11450	8.642	.0000

Note: ***, **, * = Significance at 1%, 5%, 10% level.

Appendix H: Data codebook

See attached CD.

- AGRI, AND, D. M., CAPLAN, A. J. & PHILLIPS, O. R. (2007) A Bayesian transmission of information and uncertainty in contingent valuation. *Journal of Risk and Uncertainty*, 33, 149-175.
- AJZEN, I. (1985) From intentions to actions: A theory of planned behavior. IN KUSH, J. & BECKMAN, J. (Eds.) *Action-control: From cognition to behaviour*. Heidelberg, Springer.
- AJZEN, I. (1991) The theory of planned behavior. *Organizational Behaviour and Human Decision Processes*, 50, 179-211.
- AJZEN, I. (2001) Nature and operation of attitudes. *Annual Review of Psychology*, 52, 27-58.
- AJZEN, I. & DRIVER, R. L. (1993) Contingent value measurement: on the nature and meaning of willingness to pay. *Journal of Consumer Psychology*, 3, 197-316.
- AJZEN, I. & PETERSON, G. L. (1993) Contingent value measurement: The price of everything or the value of nothing? IN PETERSON, G. L., DRIVER, R. L. & GREGORY, R. (Eds.) *Surveying resource valuations: Integrating economics with other disciplines*. State College, PA, Venture Publishing.
- ALBERTI, A. & KALIN, J. E. (2006) *Handbook on contingent valuation*. Cheltenham, Edward Elgar Publishing Ltd.
- ALBERS, W. (1997) Foundations of a theory of prominence in the decimal system. Part I: numerical responses as a process: evidence, scales, and structure of scales. *Working paper*.
- ALBERS, W. (2001) Prominence theory as a rational model boundedly rational decision. IN KREMERISCHER, G. & SHLICHMAN, R. (Eds.) *Behavioral economics: The adaptive approach*. Cambridge, MA, MIT Press.

REFERENCES

- AADLAND, D. M., CAPLAN, A. J. & PHILLIPS, O. R. (2007) A Bayesian examination of information and uncertainty in contingent valuation. *Journal of Risk and Uncertainty*, 35, 149-178.
- AJZEN, I. (1985) From intentions to actions: A theory of planned behavior. IN KUHL, J. & BECKMAN, J. (Eds.) *Action-control: from cognition to behaviour*. Heidelberg, Springer.
- AJZEN, I. (1991) The theory of planned behavior. *Organizational Behaviour and Human Decision Processes*, 50, 179-211.
- AJZEN, I. (2001) Nature and operation of attitudes. *Annual Review of Psychology*, 52, 27-58.
- AJZEN, I. & DRIVER, B. L. (1992) Contingent value measurement: on the nature and meaning of willingness to pay. *Journal of Consumer Psychology*, 1, 297-316.
- AJZEN, I. & PETERSON, G. L. (1988) Contingent value measurement: The price of everything or the value of nothing? IN PETERSON, G. L., DRIVER, B. L. & GREGORY, R. (Eds.) *Amenity resource valuation: integrating economics with other disciplines*. State College, PA, Venture Publishing.
- ALBERINI, A. & KAHN, J. R. (2006) *Handbook on contingent valuation*, Cheltenham, Edward Elgar Publishing Ltd.
- ALBERS, W. (1997) Foundations of a theory of prominence in the decimal system. Part I: numerical response as a process, exactness, scales, and structure of scales. *Working papers*.
- ALBERS, W. (2001) Prominence theory as a tool to model boundedly rational decisions. IN GIGERENZER, G. & SELTEN, R. (Eds.) *Bounded rationality: the adaptive toolbox*. Cambridge, MA, MIT Press.

- AMEMIYA, T. (1985) *Advanced econometrics*, Cambridge, MA, Harvard University Press.
- ANDREONI, J. (1990) Impure altruism and donations to public goods: a theory of warm-glow giving. *Economic Journal*, 100, 464-477.
- APRAHAMIAN, F., CHANEL, O. & LUCHINI, S. (2008) Heterogeneous anchoring and the shift effect in iterative valuation questions. *Resource and Energy Economics*, 30, 12-20.
- ARROW, K., SOLOW, R., PORTNEY, P. R., LEAMER, E. E., RADNER, R. & SCHUMAN, H. (1993) Report of the NOAA Panel on Contingent Valuation.
- ARROW, K. J. & FISHER, A. C. (1974) Environmental preservation, uncertainty and irreversibility. *Quarterly Journal of Economics*, 88, 312-319.
- AUSTRALIAN BUREAU OF STATISTICS (1979a) 1976 Census: Characteristics of the Population and Dwellings in Local Government Areas.
- AUSTRALIAN BUREAU OF STATISTICS (1979b) 1976 Census: Population and Dwellings: Cross-classified Tables.
- AUSTRALIAN BUREAU OF STATISTICS (2009) 2006 Census data.
- BALISTRERI, E., MCCLELLAND, G., POE, G. & SCHULZE, W. (2001) Can hypothetical questions reveal true values? A laboratory comparison of dichotomous choice and open-ended contingent valuation with auction values. *Environmental and Resource Economics*, 18.
- BANDARA, R. & TISDELL, C. (2004) The net benefit of saving the Asian elephant: a policy and contingent valuation study. *Ecological Economics*, 48, 93-107.
- BARAL, N., STERN, M. J. & BHATTARAI, R. (2008) Contingent valuation of ecotourism in Annapurna conservation area, Nepal: Implications for sustainable park finance and local development. *Ecological Economics*, 66, 218-227.

- BATEMAN, I. J., BROUWER, R., FERRINI, S., SCHAAFSMA, M., BARTON, D. N., DUBGAARD, A., HASLER, B., HIME, S., LIEKENS, I., NAVRUD, S., NOCKER, L. D., ŠČEPONAVIČIŪTĖ, R. & SEMĖNIĖN, D. (2011) Making Benefit Transfers Work: Deriving and Testing Principles for Value Transfers for Similar and Dissimilar Sites Using a Case Study of the Non-Market Benefits of Water Quality Improvements Across Europe. *Environmental and Resource Economics*, Online first.
- BATEMAN, I. J., CARSON, R. T., DAY, B., DUPONT, D., LOUVIERE, J. J., MORIMOTO, S., SCARPA, R. & WANG, P. (2008) Choice set awareness and ordering effects in discrete choice experiments. *Working Paper EDM 08-01*.
- BATEMAN, I. J., CARSON, R. T., DAY, B., HANEMANN, W. M., HANLEY, N., HETT, T., MICHAEL, J.-L., LOOMES, G., MOURATO, S., ÖZDEMIROGLU, E., PEARCE OBE, D. W., SUGDEN, R. & SWANSON, J. (2002) *Economic valuation with stated preference techniques: a manual*, Cheltenham, UK, Northampton, MA, USA, Edward Elgar Publishing Ltd.
- BECK, L. & AJZEN, I. (1991) Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality*, 25, 285-301.
- BECKER, G. S. (1993) Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy*, 101, 385-409.
- BEN-AKIVA, M. & LERMAN, S. (1985) *Discrete choice analysis: theory and application to travel demand*, Cambridge, MIT Press.
- BENNETT, J. (1981) *Evaluation of the benefits of natural ecosystem preservation with special reference to existence values*, Doctor of Philosophy, Australian National University, Department of Economics, Canberra.
- BENNETT, J. (1984) Using direct questioning to value the existence benefits of preserved natural areas. *Australian Journal of Agricultural Economics*, 28, 136-152.
- BENNETT, J. & BIROL, E. (2010) *Choice experiments in developing countries: implementation, challenges and policy implications* Edward Elgar Publishing Ltd.
- BENNETT, J. & BLAMEY, R. (2001) *The choice modelling approach to environmental valuation*, Massachusetts, Edward Elgar Publishing Ltd.

- BERGSTROM, J. C., STOLL, J. R. & RANDALL, A. (1989) Information Effects in Contingent Markets. *American Journal of Agricultural Economics*, 71, 685-691.
- BERNATH, K. & ROSCHEWITZ, A. (2008) Recreational benefits of urban forests: explaining visitors' willingness to pay in the context of the theory of planned behavior. *Journal of Environmental Management*, 89, 155-166.
- BIERLAIRE, M. (2003) BIOGEME: A free package for the estimation of discrete choice models *Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland*.
- BIERLAIRE, M. (2009) *Estimation of discrete choice models with Biogeme 1.8*, Lausanne, Transport and Mobility Laboratory.
- BIJLENGA, D., BONSEL, G. J. & BIRNIE, E. (2011) Eliciting willingness to pay in obstetrics: comparing a direct and an indirect valuation method for complex health outcomes. *Health Economics*, 20, 1392-1406.
- BIROL, E. & KOUNDOURI, P. (2008) *Choice experiments informing environmental policy*, Cheltenham UK, Northampton USA, Edward Elgar Publishing Ltd.
- BLACK, J. S., STERN, P. C. & ELWORTH, J. T. (1985) Personal and contextual influences on household energy adaptations. *Journal of Applied Psychology*, 70, 3-21.
- BLIEMER, M. C. J., ROSE, J. M. & HESS, S. (2008) Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling*, 1, 98-127.
- BLUMENSCHIN, K., BLOMQUIST, G. C., JOHANNESSON, M., HORN, N. & FREEMAN, P. (2008) Eliciting willingness to pay without bias: evidence from a field experiment. *The Economic Journal*, 118, 114-137.
- BOCKSTAEL, N. & FREEMAN, A. M. (2005) Welfare theory and valuation. IN MÄLER, K.-G. & VINCENT, J. R. (Eds.) *Handbook of environmental economics - valuing environmental changes*. Amsterdam, Elsevier.
- BOYLE, K. J. & BERGSTROM, J. (1999) Doubts, doubt, and doubters: the genesis of a new research agenda? IN BATEMAN, I. J. & WILLIS, K. G. (Eds.) *Valuing environmental preferences*. Oxford, Oxford University Press.

- BRAGA, J. & STARMER, C. (2005) Preference anomalies, preference elicitation and the discovered preference hypothesis. *Environmental and Resource Economics*, 32, 55-89.
- BROCK, W. A. & XEPAPADEAS, A. (2003) Valuing biodiversity from an economic perspective: A unified economic, ecological, and genetic approach. *American Economic Review*, 93, 1597-1614.
- BROUWER, R. (2006) Do stated preference methods stand the test of time? A test of stability of contingent values and models for health risks when facing an extreme event. *Ecological Economics*, 60, 399-406.
- BROUWER, R. & BATEMAN, I. J. (2005) Temporal stability and transferability of models of willingness to pay for flood control and wetland conservation. *Water Resources Research*, 41, 1-6.
- BROUWER, R., BEUKERING, P. V. & SULTANIAN, E. (2008) The impact of the bird flu on public willingness to pay for the protection of migratory birds. *Ecological Economics*, 64, 575-585.
- BROWN, G. M. & HAMMACK, J. (1973) Dynamic economic management of migratory waterfowl. *The Review of Economics and Statistics*, 55, 73-82.
- BROWN, M. B. & FORSYTHE, A. B. (1974) Robust Tests for Equality of Variances. *Journal of the American Statistical Association*, 69, 364-367.
- BROWN, T. C. (1984) The concept of value in resource allocation. *Land Economics*, 60, 231-246.
- BROWNSTONE, D. & TRAIN, K. (1999) Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89, 109-129.
- CAMERON, T. (1988) A new paradigm for valuing non-market goods using referendum data: maximumlikelihood estimation by censored logistic regression. *Journal of Environmental Economics and Management*, 15, 355-379.

- CAMERON, T. & JAMES, M. (1987) Efficient estimation methods for closed-ended contingent valuation survey data. *Review of Economics and Statistics*, 69, 269-276.
- CAMERON, T. A. & TRIVERDI, P. K. (2005) *Microeconometrics: methods and applications*, New York, Cambridge University Press.
- CARLSSON, F. & MARTINSSON, P. (2001) Do hypothetical and actual marginal willingness to pay differ in choice experiments? *Journal of Environmental Economics and Management*, 41, 179-192.
- CARLSSON, F. & MARTINSSON, P. (2006a) Do experience and cheap talk influence willingness to pay in an open-ended contingent valuation survey? *Working Papers in Economics*. Göteborg, Department of Economics, School of Business, Economics and Law, Göteborg University.
- CARLSSON, F. & MARTINSSON, P. (2006b) How much is too much? An investigation of the effect of the number of choice sets, starting point and the choice of bid vectors in choice experiments. *Working Papers in Economics*. Göteborg, Göteborg University.
- CARLSSON, F. & MARTINSSON, P. (2007) Willingness to pay among Swedish households to avoid power outages: a random parameter Tobit model approach. *The Energy Journal*, 28, 75-89.
- CARSON, R., HANEMANN, M. & STEINBERG, D. (1990) A discrete choice contingent valuation estimate of the value of Kenai King salmon. *Journal of Behavioral Economics*, 19, 53-68.
- CARSON, R. T. (1998) Valuation of tropical rainforests: philosophical and practical issues in the use of contingent valuation. *Ecological Economics*, 24, 15-29.
- CARSON, R. T. (2000) Contingent valuation: A user's guide. *Environmental Science and Technology*, 34, 1413-1418.
- CARSON, R. T. & GROVES, T. (2007) Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37, 181-210.
- CARSON, R. T. & GROVES, T. (2011) Incentive and information properties of preference questions: commentary and extensions. IN BENNETT, J. (Ed.) *The*

international handbook on non-market environmental evaluation. Cheltenham, Edward Elgar Publishing Ltd.

CARSON, R. T., HANEMANN, W. M., KOPP, R. J., KROSNICK, J. A., MITCHELL, R. C., PRESSER, S., RUUD, P. A. & SMITH, V. K. (1995) Referendum design and contingent valuation: the NOAA panel's no-vote recommendation. *Discussion paper* 96-05.

CARSON, R. T., HANEMANN, W. M., KOPP, R. J., KROSNICK, J. A., MITCHELL, R. C., PRESSER, S., RUUD, P. A. & SMITH, V. K. W. C., MICHAEL AND MARTIN, KERRY (1997) Temporal reliability of estimates from contingent valuation. *Land Economics*, 73, 151-163.

CARSON, R. T. & MITCHELL, R. C. (1995) Sequencing and nesting in contingent valuation surveys. *Journal of Environmental Economics and Management*, 28, 155-173.

CASEY, P. J. & SCOTT, K. (2006) Environmental concern and behaviour in an Australian sample within an ecocentric – anthropocentric framework. *Australian Journal of Psychology*, 58, 57-67.

CAUSSADE, S. J., ORTUZAR, J., RIZZI, L. & HENSHER, D. (2005) Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B*, 39, 621-640.

CHAMP, P. A., ALBERINI, A. & CORREAS, I. (2005) Using contingent valuation to value a noxious weeds control program: the effects of including an unsure response category. *Ecological Economics*, 55, 47-60.

CHANG, J.-I., YOO, S.-H. & KWAK, S.-J. (2007) An investigation of preference uncertainty in the contingent valuation study. *Applied Economics Letters*, 14, 691-695.

CHILTON, S. M. & HUTCHINSON, W. G. (2000) A note on the warm glow of giving and scope sensitivity in contingent valuation studies. *Journal of Economic Psychology*, 21, 343-349.

CHRISTIE, M., HANLEY, N., WARREN, J., MURPHY, K., WRIGHT, R. & HYDE, T. (2006) Valuing the diversity of biodiversity. *Ecological Economics*, 58, 304-317.

- CIRIACY-WANTRUP, S. V. (1947) Capital returns from soil-conservation practices. *Journal of Farm Economics*, 29, 1181-1196.
- COLLINS, J. P. & VOSSLER, C. A. (2009) Incentive compatibility tests of choice experiment value elicitation questions. *Journal of Environmental Economics and Management*, 58, 226-235.
- CONN, A. R., GOULD, N. I. M. & TOINT, P. (2000) *Trust region methods*, Philadelphia, Society for Industrial and Applied Mathematics.
- COOPER, P., POE, G. L. & BATEMAN, I. J. (2004) The structure of motivation for contingent values: a case study of lake water quality improvement. *Ecological Economics*, 50, 69-82.
- COVEY, J., LOOMES, G. & BATEMAN, I. J. (2007) Valuing risk reductions: testing for range biases in payment card and random card sorting methods. *Journal of Environmental Planning and Management*, 50, 467-482.
- COVEY, J. & SMITH, R. D. (2006) How common is the 'prominence effect'? Additional evidence to Whynes et al. *Health Economics*, 15, 205-210.
- CRAGG, J. (1971)
Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39, 829-844.
- CUMMINGS, R. G., ELLIOT, S., HARRISON, G. W. & MURPHY, J. (1997) Are hypothetical referenda incentive compatible? *Journal of Political Economy*, 105, 609-621.
- CUMMINGS, R. G., HARRISON, G. W. & RUTSTROM, E. E. (1995) Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive compatible? *American Economic Review*, 85, 260-266.
- D'AGOSTINO, R. B., BALANGER, A. & JR., D. A. R. B. (1990) A suggestion for using powerful and informative tests of normality. *American Statistician*.
- DAVIS, R. K. (1963) *The value of outdoor recreation: an economic study of the Maine woods*, Ph.D., Harvard University, Department of Economics, Harvard.

- DAY, B. & PRADES, J.-L. P. (2010) Ordering anomalies in choice experiments. *Journal of Environmental Economics and Management*, 59, 271-285.
- DESHAZO, J. R. & FERMO, G. (2002) Designing choice sets for stated-preference methods: the effect of complexity on choice consistency. *Journal of Environmental Economics and Management*, 44, 123-143.
- DUNLAP, R. E. & VAN LIERE, K. D. (1978) The new environmental paradigm: a proposed measuring instrument and preliminary results. *Journal of Environmental Education*, 9, 10-19.
- DUNLAP, R. E., VAN LIERE, K. D., MERTIG, A. G. & JONES, R. E. (2000) New trends in measuring environmental attitudes: Measuring endorsement of the new ecological paradigm: a revised NEP scale *Journal of Social Issues*, 56, 425-442.
- ENVIRONMENTAL VALUATION REFERENCE INVENTORY (2011).
- FARMER, M. C. & LIPSCOMB, C. A. (2008) Conservative dichotomous choice responses in the active policy setting: DC rejections below WTP. *Environmental and Resource Economics*, 35, 223-246.
- FERRINI, S. & SCARPA, R. (2007) Designs with a-priori information for non-market valuation with choice experiments: a Monte Carlo study. *Journal of Environmental Economics and Management* 53, 342-363.
- FIN, T. & SCHMIDT, P. (1984) A test for the tobit specification versus an alternative suggested by Cragg. *Review of Economics and Statistics*, 66, 174-177.
- FISCHER, A. & HANLEY, N. (2007) Analysing decision behaviour in stated preference surveys: a consumer psychological approach. *Ecological Economics*, 61, 303-314.
- FISHBEIN, M. & AJZEN, I. (Eds.) (1975) *Belief, attitude, intention, and behavior: an introduction to theory and research*, Reading, Massachusetts, Addison-Wesley.
- FRANK, R. H. (2003) *Microeconomics and Behavior*, McGraw-Hill.

- FREEMAN, A. M. (1993) *The measurement of environmental and resource values*, Washington, DC., Resources for the future.
- FRÖR, O. (2008) Bounded rationality in contingent valuation: empirical evidence using cognitive psychology. *Ecological Economics*, 68, 570-581.
- GÄRLING, T., FUJII, S., GÄRLING, A. & JAKOBSSON, C. (2003) Moderating effects of social value orientation on determinants of pro-environmental behavior intention. *Journal of Environmental Psychology*, 23, 1-9.
- GARROD, G. & WILLIS, K. G. (1999) *Economic valuation of the environment*, Cheltenham, UK, Northampton, USA, Edward Elgar Publishing Ltd.
- GIBBARD, A. (1973) Manipulation of voting schemes. *Econometrica*, 41, 587-601.
- GOLDSTEIN, W. (1979) Australia's 100 Years of National Parks. *Parks and wildlife*, 2.
- GORUSCH, R. L. & ORTBERG, J. (1983) Moral obligation and attitudes: Their relation to behavioral intentions. *Journal of Personality and Social Psychology*, 44, 1025-1028.
- GRAVES, P. E. (2007) *Environmental Economics: a critique of benefit-cost analysis*, Lanham, Massachusetts, Rowman and Littlefield Publishers.
- GREEN, C. & TUNSTALL, S. (1999) A psychological perspective. IN J., B. I. & WILLIS, K. G. (Eds.) *Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries*. Oxford, Oxford University Press.
- GREEN, J. R. & LAFFONT, J. J. (1978) A sampling approach to the free rider problem. IN SANDMO, A. (Ed.) *Essays in public economics*. Lexington Books.
- GREENE, W. & HENSHER, D. (2007) Heteroscedastic control for random coefficients and error components in mixed logit. *Transportation Research Part E: Logistics and Transportation Review*, 43, 610-623.

- GREENE, W. H. (2002) *LIMDEP Version 8.0 econometric modeling guide*, Castle Hill, Econometric Software, Inc.
- GREENE, W. H. (2007) *Nlogit version 4.0 reference guide*, Castle Hill, Econometric Software, Inc.
- GREENE, W. H. (2008) *Econometric analysis*, Upper Saddle River, N.J., Pearson/Prentice Hall.
- GUAGNANO, G. A., STERN, P. C. & DIETZ, T. (1995) Influences on attitudes-behavior relationships: a natural experiment with curbside recycling. *Environment and Behavior*, 27, 699-718.
- GUJARATI, D. N. (2005) *Essentials of Econometrics*, Boston, McGrawHill.
- HADDAD, B. & HOWARTH, R. B. (2006) Protest bids, commensurability, and substitution: contingent valuation and ecological economics. IN ALBERINI, A. & KAHN, J., R. (Eds.) *Handbook on contingent valuation*. Cheltenham, UK, Edward Elgar Publishing Ltd.
- HAMMACK, J. & BROWN, G. M. (1974) *Waterfowl and wetlands: Toward bioeconomic analysis*, Baltimore, The John Hopkins University Press.
- HANEMANN, W. M. (1988) Three approaches to defining 'existence' or non-use values under certainty. *Working Paper*.
- HANEMANN, W. M. (1999) The economic theory of WTP and WTA. IN BATEMAN, I. J. & WILLIS, K. G. (Eds.) *Valuing environmental preferences*. Oxford, Oxford University Press.
- HANLEY, N. (2001) Cost-benefit analysis and environmental policy making. *Environment and Planning C: Government and Policy*, 19, 103-118.
- HANLEY, N. & BARBIER, E. B. (2009) *Pricing nature: cost-benefit analysis and environmental policy*, Cheltenham, UK; Northampton, MA, USA, Edward Elgar.

- HANSLA, A., GAMBLE, A., JULIUSSON, A. & GÄRLING, T. (2008) The relationships between awareness of consequences, environmental concern, and value orientations. *Journal of Environmental Psychology*, 28, 1-9.
- HECKMAN, J. J. (1976) The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Annals of Economics and Social Measurement*, 5, 475-492.
- HENSHER, D. & GREENE, W. H. (2011) Valuation of travel time savings in WTP space in the presence of taste and scale heterogeneity. *Journal of Transport Economics and Policy*, 45, 505-525.
- HENSHER, D., LOUVIERE, J. & SWAIT, J. (1999) Combining sources of preference data. *Journal of Econometrics*, 89, 197-221.
- HENSHER, D., ROSE, J. M. & GREENE, W. H. (2005) *Applied choice analysis : a primer* Cambridge, Cambridge University Press.
- HENSHER, D. A. & GREENE, W. H. (2003) The Mixed Logit model: The state of practice. *Transportation*, 30, 133-176.
- HICKS, J. (1939) *Value and capital*, Oxford, Clarendon Press.
- HICKS, J. R. (1945) The generalised theory of consumer's surplus. *Review of Economic Studies*, 13, 68-74.
- HILL, R. C., GRIFFITHS, W. E. & LIM, G. C. (2008) *Principles of econometrics*, Hoboken, NJ, Wiley.
- HODGSON, G. M. (1988) *Economics and institutions*, Cambridge, Policy Press.
- HOEHN, J. P. & RANDALL, A. (2002) The effect of resource quality information on resource injury perceptions and contingent values. *Resource and Energy Economics*, 24, 13-31.

- HOLMES, T. P. & BOYLE, K. J. (2005) Dynamic learning and context-dependence in sequential, attribute-based, stated preference valuation questions. *Land Economics*, 81, 114-126.
- HOLT, C. A. (1986) Preference reversals and the independence axiom. *The American Economic Review*, 76, 508-515.
- HOROWITZ, J. K. (2002) Preferences in the future. *Environmental and Resource Economics*, 21, 241-259.
- HURWICZ, L. (1960) Optimality and informational efficiency in resource allocation processes. IN ARROW, K., KARLIN, S. & SUPPES, P. (Eds.) *Mathematical Methods in the Social Sciences*. Stanford University Press.
- HURWICZ, L. (1972) On informationally decentralized systems. IN RADNER, R. & MCGUIRE, C. B. (Eds.) *Decision and Organisation*. Amsterdam.
- JOHANSSON, P.-O. (1993) *Cost-benefit analysis of environmental change*, Cambridge, Cambridge University Press.
- JOHNSTON, R. J. & ROSENBERGER, R. S. (2010) Methods, trends and controversies in contemporary benefit transfer. *Journal of Economic Surveys*, 24, 479-510.
- JOIREMAN, J. A., LASANE, T. P., BENNETT, J., RICHARDS, D. & SOLAIMANI, S. (2001) Integrating social value orientation and the consideration of future consequences within the extended norm activation model of proenvironmental behaviour. *British Journal of Social Psychology*, 40, 133-155.
- JORGENSEN, B. S. & SYME, G. J. (2000) Protest responses and willingness to pay: attitude toward paying for stormwater pollution abatement. *Ecological Economics*, 251-265.
- KAHNEMAN, D. & KNETSCH, J. L. (1992) Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22, 57-70.
- KAHNEMAN, D., RITOV, L., JACOWITZ, K. E. & GRANT, P. (1993) Stated willingness to pay for public goods: a psychological perspective. *Psychological Science*, 4, 310-315.

- KALBFLEISCH, J. & PRENTICE, R. (2002) *The statistical analysis of failure time data*, New York, Wiley.
- KOTCHEN, M. J. & REILING, S. D. (2000) Environmental attitudes, motivations, and contingent valuation of nonuse values: a case study involving endangered species. *Ecological Economics*, 32, 93-107.
- KRINSKY, I. & ROBB, A. L. (1986) On approximating the statistical properties of elasticities. *Review of Economics and Statistics*, 68, 715-719.
- KRISTRÖM, B. & RIERA, P. (1996) Is the income elasticity of environmental improvements less than one? *Environmental and Resource Economics*, 7, 45-55.
- KROEBER-RIEL, W. & WEINBERG, P. (1996) *Konsumentenverhalten*, München, Vahlen.
- KRUTILLA, J. V. (1967) Conservation reconsidered. *American Economic Review*, 57, 777-786.
- KRUTILLA, J. V. & CICCHETTI, C. J. (1972) Evaluating benefits of environmental resources with special application to the Hells Canyon. *Natural Resources Journal*, 12, 1-29.
- KUNTSEVICH, A. V. & KAPPEL, F. (1997) The solver for local non-linear optimisation problems. Graz, Karl-Franzens University of Graz.
- LANCASTER, K. (1966) A new approach to consumer theory. *Journal of Political Economics*, 74, 217-231.
- LARSON, K. L. & LACH, D. (2008) Participants and non-participants of place-based groups: an assessment of attitudes and implications for public participation in water resource management. *Journal of Environmental Management*, 88, 817-830.
- LEVENE, H. (1960) Robust tests for equality of variances. IN AL., I. O. E. (Ed.) *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford, Stanford University Press.

- LEVY, P. S. & LEMESHOW, S. (2008) *Sampling of populations: methods and applications*, Hoboken, New Jersey, John Wiley and Sons.
- LIST, J., SINAH, P. & TAYLOR, M. (2006) Using choice experiments to value non-market goods and services: evidence from field experiments. *Advances in Economic Analysis and Policy*, 6, 1-37.
- LONG, J. S. (1997) *Regression models for categorical and limited dependent variables* Thousand Oaks, Sage Publications.
- LONG, J. S. & FREESE, J. (2006) *Regression models for categorical dependent variables using stata*, Stata Press.
- LOOMIS, J. B. (1989) Test-retest reliability of the contingent valuation method: a comparison of general population and visitor responses. *American Journal of Agricultural Economics*, 71, 76-84.
- LOUVIERE, J. & EAGLE, T. C. (2006) Confound it! That Pesky Little Scale Constant Messes Up Our Convenient Assumptions! *CenSoC Working Paper No. 06-002*. Sydney, Centre for the Study of Choice, University of Technology.
- LOUVIERE, J. & HENSHER, D. (1982) On the design and analysis of simulated choice or allocation experiments in travel choice modelling. *Transportation Research Records*, 890, 11-17.
- LOUVIERE, J., HENSHER, D. & SWAIT, J. (2000) *Sated choice methods: analysis and applications*, Cambridge, UK ; New York, NY, USA, Cambridge University Press.
- LOUVIERE, J. & WOODWORTH, G. (1983) Design and analysis of simulated consumer choice experiments or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research*, 20, 350-367.
- MACMILLAN, D., HANLEY, N. & LIENHOOP, N. (2006) Contingent valuation: environmental polling or preference engine? *Ecological Economics*, 60, 299-307.

- MÄLER, K.-G. (1971) A method of estimating social benefits from pollution control. *Swedish Journal of Economics*, 73, 121-133.
- MÄLER, K.-G. (1994) *Environmental Economics: a theoretical enquiry*, Baltimore, John Hopkins University Press.
- MANN, H. B. & WHITNEY, D. R. (1947) On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18.
- MANSKI, J. (1977) The structure of random utility models. *Theory and Decision*, 8, 229-254.
- MARSHALL, A. (1920) *Principles of economics*, London, Macmillan.
- MAS-COLELL, A., WHINSTON, M. D. & GREEN, J. R. (1995) *Microeconomic theory*, New York Oxford, Oxford University Press.
- MASKIN, E. (1977) Nash equilibrium and welfare optimality. *Summer workshop of the Econometric Society Paris, Published 1999 in the Review of Economic Studies* 66, 23-38.
- MCCONNELL, K. E., STRAND, I. E. & VALDÉS, S. (1998) Testing temporal reliability and carry-over effect: the role of correlated responses in test-retest reliability studies. *Environmental and Resource Economics*, 12, 357-374.
- MCFADDEN, D. (1974) Conditional logit analysis of qualitative choice behaviour. IN ZAREMBKA, P. (Ed.) *Frontiers in econometrics*. New York, Academic Press.
- MCFADDEN, D. (1980) Econometric models for probabilistic choice among products. *Journal of Business*, 53, 13-29.
- MCFADDEN, D. (1994) Contingent valuation and social choice. *American Journal of Agricultural Economics*, 76, 689-708.
- MCFADDEN, D. & TRAIN, K. (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15, 447-470.

- MCNAIR, B., BENNETT, J. & HENSHER, D. (2011) A comparison of responses to single and repeated discrete choice questions. *Resource and Energy Economics*, 33, 554-571.
- MEYERHOFF, J. & LIEBE, U. (2006) Protest beliefs in contingent valuation: explaining their motivation. *Ecological Economics* 57, 583-594.
- MILLER, D. & PATASSINI, D. (2005) *Beyond benefit cost analysis: accounting for non-market values in planning evaluation*, Aldershot, England, Ashgate Publishing.
- MITCHELL, R. C. & CARSON, R. T. (1989) *Using surveys to value public goods: The contingent valuation method*, Washington DC Resources for the future.
- MORRISON, M. & BENNETT, J. (2004) Valuing New South Wales rivers for use in benefit transfer. *Australian Journal of Agricultural and Resource Economics*, 48, 591-611.
- MOULIN, H. (1994) Social choice. IN AUMANN, R. J. & S., H. (Eds.) *Handbook of game theory with economic applications*. Amsterdam.
- MOZUMDER, P. & BERRENS, R. P. (2007) Investigating hypothetical bias: induced-value tests of the referendum voting mechanism with uncertainty. *Applied Economic Letters*, 14, 705-709.
- MULLAINATHAN, S. & THALER, R. (2000) Behavioural economics. *NBER Working Paper Series*.
- MULLER, E. & SATTERTHWAITE, M. A. (1985) Strategy-proofness: the existence of dominant-strategy mechanisms. IN HURWICZ, L., SCHMIEDLER, D. & SONNENSCHN, H. (Eds.) *Social Goals and Social Organization*. Cambridge, Cambridge University Press.
- MUNRO, A. & HANLEY, N. (1999) Information, uncertainty and contingent valuation. IN BATEMAN, I. J. & WILLIS, K. G. (Eds.) *Contingent valuation of environmental preferences: assessing theory and practice in the USA, Europe, and developing countries*. Oxford, Oxford University Press.
- NEW SOUTH WALES GOVERNMENT (1987) Wilderness Act 1987.

NEW SOUTH WALES NATIONAL PARKS AND WILDLIFE SERVICE (2003)
Nadgee Nature Reserve: plan of management.

NEW SOUTH WALES NATIONAL PARKS AND WILDLIFE SERVICE (2011)
Reserve types in New South Wales.

NORDLUND, A. M. & GARVILL, J. (2003) Effects of values, problem awareness, and personal norm on willingness to reduce personal car use. *Journal of Environmental Psychology*, 23, 339-347.

NORTON, B., CONSTANZA, R. & BISHOP, R. C. (1998) The evolution of preference: why 'sovereign' preferences may not lead to sustainable policies and what to do about it. *Ecological Economics*, 24, 193-211.

OJEDA, M. I., MAYER, A. S. & SOLOMON, B. D. (2008) Economic valuation of environmental services sustained by water flows in the Yaqui River Delta. *Ecological Economics*, 65, 155-166.

PAGAN, A. R. & VELLA, F. (1989) Diagnostic tests for models based on individual data. *Journal of Applied Econometrics*, 4, 29-59.

PEARCE, D. (1993) *Economic values and the natural world*, London.

PEARCE, D. W., ATKINSON, G. & MOURATO, S. (2006) *Cost-benefit analysis and the environment: recent developments*, Paris, France, Organisation for Economic Cooperation and Development Publishing.

PINK, B. (2008) Australian Standard Geographical Classification: Australian Capital Territory maps. IN STATISTICS, A. B. O. (Ed.). Canberra.

PLOTT, C. R. (1996) Rational individual behavior in markets and social choice processes: the discovered preference hypothesis. IN ARROW, K., COLOMBATTO, E., PERLEMAN, M. & SCHMIDT, C. (Eds.) *Rational foundations of economic behavior*. London, Macmillan.

POE, G. L., CLARK, J., RONDEAU, D. & SCHULZE, W. D. (2002) Provision point mechanisms and field validity tests of contingent valuation. *Environmental and Resource Economics*, 23, 105-131.

- POE, G. L., GIRAUD, K. L. & LOOMIS, J. B. (2005) Computational methods for measuring the difference of empirical distributions. *American Journal of Agricultural Economics*, 87.
- POE, G. L., SEVERANCE-LOSSIN, E. & WELSH, M. P. (1994) Measuring the difference (X-Y) of simulated distributions: a convolutions approach. *American Journal of Agricultural Economics*, 76, 904-914.
- POMAZAL, R. J. & JACCARD, J. J. (1976) An informational approach to altruistic behaviour. *Journal of Personality and Social Psychology*, 33, 317-326.
- POUTA, E. & RECOLA, M. (2001) The theory of planned behavior in predicting willingness to pay for abatement of forest regeneration. *Society and Natural Resources*, 14, 93-106.
- POWELL, J. (1984) Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*, 25, 303-325.
- RACEVSKIS, L. A. & LUPI, F. (2008) Incentive compatibility in an attribute-based referendum model. *American Agricultural Economics Association Annual Meeting*. Orlando.
- RANDALL, A. (1974) Quantifying the Unquantifiable: Benefits from Abatement of Aesthetic Environmental Damage. *American Journal of Agricultural Economics*, 56, 1209.
- RANDALL, A., IVES, B. & EASTMAN, C. (1974) Bidding games for valuation of aesthetic environmental improvements. *Journal of Environmental Economics and Management*, 1, 132-149.
- RANDALL, A. & STOLL, J. R. (1982) Existence value in a total valuation framework. IN ROWE, R. D. & CHESTNUT, L. G. (Eds.) *Managing air quality and scenic resources at national parks and wilderness areas*. Boulder, Colorado, Westview Press.
- REILING, S. D., BOYLE, K. J., PHILLIPS, M. L. & ANERSON, M. W. (1990) Temporal reliability of contingent values. *Land Economics*, 66, 128-134.

- REVELT, D. & TRAIN, K. (1998) Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, 80, 647-657.
- ROKEACH, M. (1973) *The nature of human values*, New York, Free Press.
- ROLFE, J. & BENNETT, J. W. (1996) Respondents to contingent valuation surveys: consumers or citizens - a comment. *Australian Journal of Agricultural Economics*, 40, 129-133.
- ROLFE, J. & PRAYAGA, P. (2007) Estimating values for recreational fishing at freshwater dams in Queensland. *Australian Journal of Agricultural and Resource Economics*, 51, 157-174.
- ROLFE, J., WINDLE, J., BENNETT, J. & MAZUR, K. (2011) Calibration of values in benefit transfer to account for variations in geographic scale and scope: Comparing two choice modelling experiments. *Australian Journal of Agricultural and Resource Economics*, Forthcoming.
- ROLLINS, K. & LYKE, A. (1998) The case of diminishing marginal existence values. *Journal of Environmental Economics and Management*, 36, 324-344.
- ROSE, J. M. & BLIEMER, M. C. J. (2005) Designing efficient data for stated choice experiments: accounting for socio-demographic and contextual effects in designing stated choice experiments. *11th International Conference on Travel Behaviour Research*. Kyoto, Japan.
- ROSE, J. M., BLIEMER, M. C. J., HENSHER, D. A. & A., C. (2008) Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B*, 42, 395-406.
- ROSE, J. M., HESS, S., BLIEMER, M. C. J. & DALY, A. (2009) The impact of varying the number of repeated choice observations on mixed multinomial logit models. *European Transport Conference*. Leiden, The Netherlands.
- RUUD, P. A. (1996) Simulation of the multinomial probit model: An analysis of covariance matrix estimation. *Working Paper*.

- SAMUELSON, P. (1954) The pure theory of public expenditure. *Review of Economics and Statistics*, 36, 387-389.
- SÁNDOR, Z. & WEDEL, M. (2001) Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38, 430-444.
- SATTERTHWAITE, M. A. (1975) Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10, 187-217.
- SAVAGE, S. J. & WALDMAN, D. M. (2008) Learning and fatigue during choice experiments: a comparison of online and mail survey modes. *Journal of Applied Econometrics*, 23, 351-371.
- SCARPA, R. & ALBERINI, A. (2005) *Applications of simulation methods in environmental and resource economics*, Dordrecht, Springer.
- SCARPA, R. & ROSE, J. M. (2008) Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. *The Australian Journal of Agricultural and Resource Economics*, 52, 253-282.
- SCARPA, R., THIENE, M. & 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, 994-1010.
- SCHELLING, T. C. (1960) Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. IN SAUERMAN, H. (Ed.) *Beiträge zur experimentellen Wirtschaftsforschung*. Tübingen, J.C.B. Mohr (Paul Siebeck).
- SCHEUFELE, G. & BENNETT, J. (2012a) Effects of alternative elicitation formats in discrete choice experiments. *Revised and re-submitted to The Australian Journal of Agricultural and Resource Economics*.
- SCHEUFELE, G. & BENNETT, J. (2012b) Response strategies and learning in discrete choice experiments. *Environmental and Resource Economics*, Online first.
- SCHLÄPFER, F. & BRAEUER, I. (2007) Theoretical incentive properties of contingent valuation questions: Do they matter in the field? *Ecological Economics*, 62, 451-460.

- SCHULTZ, P. W. & ZELENZKY, L. (1999) Values as predictors of environmental attitudes: Evidence of consistency of across cultures. *Journal of Environmental Psychology*, 19, 255-265.
- SCHWARTZ, S. H. (1977) Normative influences on altruism. IN BERKOWITZ, L. (Ed.) *Advances in Experimental Social Psychology*. San Diego, Academic Press.
- SCHWARTZ, S. H. (1992) Universals in the content of and structure of values: theoretical advances and empirical tests in 20 countries. IN ZANNA, M. P. (Ed.) *Advances in Experimental Social Psychology*. Orlando, FL, Academic Press.
- SCHWARTZ, S. H. & TESSLER, R. C. (1972) A test for a model reducing measured attitude-behavior inconsistencies. *Journal of Personality and Social Psychology*, 24, 225-236.
- SHAIKH, S. L., SUN, L. & KOOTEN, G. C. V. (2007) Treating respondent uncertainty in contingent valuation: a comparison of empirical treatments. *Ecological Economics*, 62, 115-125.
- SHOGREN, J. F. & TAYLOR, L. O. (2008) On behavioral-environmental economics. *Review of Environmental Economics and Policy*, 2, 26-44.
- SKEELS, C. L. & VELLA, F. (1999) A Monte Carlo investigation of the sampling behavior of conditional moment tests in tobit and probit models. *Journal of Econometrics* 92, 275-294.
- SLIMAK, M. W. & DIETZ, T. (2006) Personal values, beliefs, and ecological risk perception. *Risk Analysis*, 26, 1689-1705.
- SMITH, V. K. & OSBORNE, L. L. (1996) Do contingent valuation estimates pass a "scope" test? A meta-analysis. *Journal of Environmental Economics and Management*, 31, 287-301.
- SMITH, V. L. (1998) The two faces of Adam Smith. *Southern Economic Journal*, 65, 1-19.

- SNELGAR, R. S. (2006) Egoistic, altruistic, and biospheric environmental concerns: measurement and structure. *Journal of Environmental Psychology*, 26, 87-99.
- SÖDERQUIST, T., MITSCH, W. J. & TURNER, R. K. (2000) Valuation of wetlands in a landscape and institutional perspective. *Ecological Economics*, 35, 1-6.
- SOLOMON, B. D. & JOHNSON, N. H. (2009) Valuing climate protection through willingness to pay for biomass ethanol. *Ecological Economics*, 68, 2137-2144.
- SONNIER, G., AINSLIE, A. & OTTER, T. (2007) Heterogeneity distributions of willingness-to-pay in choice models. *Quantitative Marketing and Economics*, 5, 313-331.
- SPASH, C. L. (2006) Non-economic motivation for contingent values: rights and attitudinal beliefs in the willingness to pay for environmental improvements. *Land Economics*, 82, 602-622.
- STATACORP, L. (2007) *Stata release 10*, Texas, Stata Press.
- STERN, P. C. (2000) Towards a coherent theory of environmentally significant behavior. *Journal of Social Issues*, 56, 407-424.
- STERN, P. C., DIETZ, T., ABEL, T., GUAGNANO, G. A. & KALOF, L. (1999) A value-belief-norm theory of support for social movements: the case of environmental concern. *Human Ecology Review*, 6, 81-97.
- STERN, P. C., DIETZ, T. & KALOF, L. (1993) Value orientations, gender, and environmental concern. *Environment and Behaviour*, 25, 322-348.
- STERN, P. C., DIETZ, T., KALOF, L. & GUAGNANO, G. A. (1995) Values, beliefs and pro-environmental action: attitude formation toward emergent attitude objects. *Journal of Applied Social Psychology*, 25, 1611-1636.
- STEVENS, T. H., ECHEVERRIA, J., GLASS, R. J., HAGER, T. & MORE, T. A. (1991) Measuring the existence value of wildlife: What do CVM estimates really show? *Land Economics*, 67, 390-400.

- STEVENS, T. H., MORE, T. A. & GLASS, R. J. (1994) Interpretation and temporal stability of CV bids for wildlife existence: a panel study. *Land Economics*, 70, 355-363.
- SUN, L. & VAN KOOTEN, G. C. (2009) Comparing fuzzy and probabilistic approaches to preference uncertainty in non-market valuation. *Environmental and Resource Economics*, 42, 471-489.
- SVEDSATER, H. (2007) Ambivalent statements in contingent valuation studies: inclusive response formats and giving respondents time to think. *The Australian Journal of Agricultural and Resource Economics*, 51, 91-107.
- SWAIT, J. & ADAMOWICZ, W. (2001) Choice environment, market complexity, and consumer behavior: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organizational Behavior and Human Decision Processes*, 86, 141-167.
- SWAIT, J. & LOUVIERE, J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30, 305-314.
- TEISL, M. F., BOYLE, K. J., MCCOLLUM, D. W. & REILING, S. D. (1995) Test-retest reliability of contingent valuation with independent sample pretest and posttest control groups. *American Journal of Agricultural Economics*, 77, 613-619.
- THURSTONE, L. (1927) A law of comparative judgment. *Psychological Review*, 4, 273-286.
- TIETENBERG, T. & LEWIS, L. (2009) *Environmental and natural resource economics*, Pearson, Addison Wesley.
- TISDELL, C., WILSON, C. & NANTHA, H. S. (2008) Contingent valuation as a dynamic process. *The Journal of Socio-Economics*, 37, 1443-1458.
- TOBIN, J. (1958) Estimation of relationships for limited dependent variables. *Econometrica* 26, 24-36.
- TRAIN, K. (Ed.) (2009) *Discrete choice methods with simulation*, Cambridge, Cambridge University Press.

- TRAIN, K. & WEEKS, M. (2005) Discrete choice models in preference space and willingness-to-pay space. IN SCARPA, R. & ALBERINI, A. (Eds.) *Applications of simulation methods in environmental and resource economics*. Dordrecht, The Netherlands, Springer Publisher.
- TURNER, R. K. (1999) The place of economic values in environmental valuation. IN BATEMAN, I. J. & WILLIS, K. G. (Eds.) *Valuing environmental preferences*. Oxford, University Press.
- TURNER, R. K. (2003) Valuing nature: lessons learned and future research directions. *Ecological Economics*, 46, 493-510.
- TURNER, R. K., BERGH, J. C. J. M. V. D., SÖDERQVIST, T., BARENDREGT, A., STRAATEN, J. V. D., MALTBY, E. & IERLAND, E. C. V. (2000) Ecological-economic analysis of wetlands: scientific integration for management and policy. *Ecological Economics*, 35, 7-23.
- TVERSKY, A. & KAHNEMAN, D. (1991) Loss aversion in riskless choices: A reference dependent model. *Quarterly Journal of Economics*, 106, 1039-1061.
- VARIAN, H. R. (2006) *Intermediate microeconomics : a modern approach*, New York, W.W. Norton & Co.
- VOSSLER, C. A., POE, G. L., WELSH, M. & ETHIER, R. G. (2004) Bid design effects in multiple bounded discrete choice contingent valuation. *Environmental and Resource Economics*, 29, 401-418.
- WEINBERG, P. (1981) *Das Entscheidungsverhalten der Konsumenten*, Paderborn, Schöningh.
- WERNER, P., SCHNAIDER-BEERI, M., AHARON, J. & DAVIDSON, M. (2002) Family caregivers' willingness to pay for drugs indicated for the treatment of Alzheimer's disease: An economic or psychological model? *Dementia*, 1, 59-74.
- WHITEHEAD, J. C. & AIKEN, R. (2007) Temporal reliability of willingness to pay from the national survey of fishing, hunting and wildlife-associated recreation. *Applied Economics*, 39, 777-786.

- WHITEHEAD, J. C. & HOBAN, T. J. (1999) Testing for temporal reliability in contingent valuation with time for changes in factors affecting demand. *Land Economics*, 75, 453-465.
- WHYNES, D. K., FREW, E. J., PHILIPS, Z. N., COVEY, J. & SMITH, R. D. (2007) On the numerical forms of contingent valuation responses. *Journal of Economic Psychology*, 28, 462-476.
- WHYNES, D. K., ZOËPHILIPS & FREW, E. (2005) Think of a number. . . any number? *Health Economics*, 14, 1191-1195.
- WIIDEGREN, Ö. (1998) The new environmental paradigm and personal norms *Environment and Behavior* 30, 75-100.
- WOOLDRIDGE, J. M. (2006) *Introductory econometrics : a modern approach*, Mason, Ohio, Thomson/South-Western.
- WOOLDRIDGE, J. M. (2010) *Econometric analysis of cross section and panel data*, Cambridge, Massachusetts, Massachusetts Institute of Technology.
- YOUNG, R. (1991) The economic significance of environmental resources: a review of the evidence. *Review of Marketing and Agricultural Economics*, 59, 229-254.
- ZANDERSEN, M., TERMANSEN, M. & JENSEN, F. S. (2005a) Benefit transfer over time of ecosystem values: the case of forest recreation. *Working papers*. Hamburg, Hamburg University and Centre for Marine and Atmospheric Science.
- ZANDERSEN, M., TERMANSEN, M. & JENSEN, F. S. (2005b) Valuing new forest sites over time: the case of afforestation and recreation in Denmark. *Working papers*. Hamburg University and Centre for Marine and Atmospheric Science.