

Lehr, Ulrike

Book — Digitized Version

Contingent Valuation Daten und Bayes'sche Verfahren: Ein Vorschlag zur Verbesserung von Umweltbewertung und Nutzentransfer

Hohenheimer volkswirtschaftliche Schriften, No. 55

Provided in Cooperation with:

Peter Lang International Academic Publishers

Suggested Citation: Lehr, Ulrike (2006) : Contingent Valuation Daten und Bayes'sche Verfahren: Ein Vorschlag zur Verbesserung von Umweltbewertung und Nutzentransfer, Hohenheimer volkswirtschaftliche Schriften, No. 55, ISBN 978-3-631-75404-7, Peter Lang International Academic Publishers, Berlin, <https://doi.org/10.3726/b13922>

This Version is available at:

<https://hdl.handle.net/10419/182958>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Ulrike Lehr

Contingent Valuation Daten und Bayes'sche Verfahren



Ulrike Lehr

Contingent Valuation Daten und Bayes'sche Verfahren

Mit Methoden zur Umweltbewertung, die – wie die interviewgestützte Contingent Valuation Methode – auf der Messung individueller Präferenzen basieren, lässt sich der Wert einer Umweltveränderung umfassend bestimmen, aber sie erweisen sich in der Praxis oft als zu teuer. Durch die Verwendung von Bayes'schen Ansätzen können die Kosten der Kosten-Nutzen-Analyse für explizite Bewertungsstudien sowie für die implizite Bewertung durch den Nutzentransfer deutlich gesenkt werden. Darüber hinaus wird durch diesen Ansatz der Nutzentransfer erheblich zuverlässiger, so dass dieses *Sparmodell der Umweltbewertung* mit weitaus geringeren Einschränkungen als bislang empfohlen werden kann. Die Arbeit illustriert diese Aussagen mit einer Vielzahl simulierter und empirischer Beispiele.

Ulrike Lehr, 1985–1990 Studium der Physik an der Universität-GHS Essen, 1990–1993 Studium der Wirtschaftswissenschaften an der Virginia Polytechnic Institute and State University in Blacksburg, VA, (USA); 1993–1999 Referentin am RWI Essen e.V. in der Forschungsgruppe Energiewirtschaft; 1999–2001 Wissenschaftliche Mitarbeiterin am Lehrstuhl für VWL, insbesondere Umweltökonomie an der Universität in Cottbus; 2001–2005 Wissenschaftliche Mitarbeiterin am Lehrstuhl für VWL, insbesondere Umweltökonomie sowie Ordnungs-, Struktur- und Verbraucherpolitik; 2005 Promotion; seit November 2005 Wissenschaftlerin und Projektleiterin am DLR-Institut für Technische Thermodynamik in der Abteilung für Systemanalyse und Technikbewertung.

Contingent Valuation Daten und Bayes'sche Verfahren

Hohenheimer Volkswirtschaftliche Schriften

Herausgegeben von

Prof. Dr. Michael Ahlheim, Prof. Dr. Ansgar Belke,
Prof. Dr. Rolf Caesar, Prof. Dr. Harald Hagemann, Prof. Dr. Klaus Herdzina,
Prof. Dr. Walter Plesch, Prof. Dr. Ingo Schmidt, Prof. Dr. Ulrich Schwalbe,
Prof. Dr. Peter Spahn, Prof. Dr. Gerhard Wagenhals,

Band 55



PETER LANG

Frankfurt am Main · Berlin · Bern · Bruxelles · New York · Oxford · Wien

Ulrike Lehr

**Contingent Valuation
Daten und
Bayes' sche Verfahren**

Ein Vorschlag zur Verbesserung
von Umweltbewertung
und Nutzentransfer



PETER LANG

Frankfurt am Main · Berlin · Bern · Bruxelles · New York · Oxford · Wien

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über <<http://dnb.ddb.de>> abrufbar.

Open Access: The online version of this publication is published
on www.peterlang.com and www.econstor.eu under the interna-
tional Creative Commons License CC-BY 4.0. Learn more on how
you can use and share this work: [http://creativecommons.org/
licenses/by/4.0](http://creativecommons.org/licenses/by/4.0).



This book is available Open Access thanks to the kind support of
ZBW – Leibniz-Informationszentrum Wirtschaft.

Zugl.: Hohenheim, Univ., Diss., 2005

**Gedruckt auf alterungsbeständigem,
säurefreiem Papier.**

D 100

ISSN 0721-3085

ISBN 3-631-55080-4

ISBN 978-3-631-75404-7 (eBook)

© Peter Lang GmbH

Europäischer Verlag der Wissenschaften
Frankfurt am Main 2006

Printed in Germany 1 2 3 4 5 7

www.peterlang.de

Für Elsa und Bruno

In God we trust - all others bring data.

W. Edwards Deming, 1900 - 1993, Amerikanischer Statistiker

Vorwort

Diese Arbeit hat Inspiration und Unterstützung aus einer Vielzahl von Quellen erfahren. So möchte ich an dieser Stelle meinem Doktorvater Professor Dr. Michael Ahlheim und meinem Zweitgutachter Professor Dr. Gerhard Wagenhals meinen Dank aussprechen, die diese Arbeit inspiriert und begleitet haben und immer für Gespräche zur Verfügung standen. Meinem Kollegen Oliver Frör gebührt Dank für eine Vielzahl von Diskussionen, bei denen er sich geduldig meinen allerneuesten Ideen aussetzte. In beinahe ebenso vielen Diskussionen hat Holger Willert die Höhen und Tiefen dieser Arbeit intensiv erfahren und mich so stets unterstützt. Dem Projektteam des DFG-geförderten Sonderforschungsbereichs 565, Teilprojekt D4, danke ich dafür, dass es mir die Daten zur Verfügung stellte, ebenso wie Thomas Bu Bjørner und Cliff Russell für die Daten aus dem Projekt "Tokkebob Hegn". Insbesondere Thomas hat große Mühe in Kauf genommen und mitten in einem beruflichen Ortswechsel und Umzug noch die Originaldaten der Studie zur Verfügung gestellt.

Die Kapitel zur Anwendung Bayes'scher Verfahren in dieser Arbeit profitierten in ihrer vorliegenden Fassung von fruchtbaren Gesprächen am Rande des Joint Statistical Meetings 2004 in Toronto, Kanada, besonders mit Victor de Oliviera von der University of Arkansas, Eric P. Smith von meiner Heimatuniversität Virginia Tech und Robert McCulloch von der University of Chicago. Mit Hilfe der großzügigen Unterstützung des Universitätsbunds Hohenheim hatte ich die Gelegenheit, auf dieser Konferenz Teile der Arbeit zu präsentieren.

Zwischen der Idee zur Arbeit und dem Buch liegt das Schreiben. Dafür, dass die Rechtschreibung in dieser Arbeit nicht mehr Gewalt erfährt, als ihr durch verschiedene Reformen ohnehin angetan wurde, bin ich Isabell Benignus und Helga Büchler-Gehring zu Dank verpflichtet. Für alle verbliebenen Fehler bin nur ich selbst verantwortlich.

Inhaltsverzeichnis

Vorwort	ix
Einleitung	1
I Studien zur Umweltbewertung	9
1 Umweltbewertung und CVM	11
1.1 Einführung zur Umweltbewertung	11
1.2 Theoretische Grundlagen der CVM	24
1.3 Contingent Valuation - die Praxis	32
1.3.1 Definition des Umweltguts	34
1.3.2 Der hypothetische Markt	38
1.3.3 Die Frage nach der Zahlungsbereitschaft	41
1.3.4 Soziodemographische und psychosoziale Daten	49
1.3.5 CVM-Interview bei Umweltschäden	51
2 Das statistische Modell	55
2.1 Ökonomische Fundierung	56
2.2 Klassische parametrische Modelle	58
2.2.1 Die Parameterschätzung	64
2.2.2 Zur Modellwahl	68
2.3 Weitere Ansätze der klassischen Statistik	72
2.4 Ein alternativer Ansatz: Das Bayes'sche Modell	76
2.4.1 Bayes'sches Modell für CVM Daten	78
2.4.2 Der Gibbs-Sampler	81
2.4.3 Über Konvergenz und Konfidenz	83
2.4.4 Vorteile des Bayes'schen Ansatzes	86
2.4.5 Die Wahl der A-Priori-Verteilung	87
3 Zahlungsbereitschaftsanalyse	91
3.1 Eine Monte Carlo Simulation	92

3.1.1	Simulation einer CVM-Befragung	92
3.1.2	Klassische Modellierung	98
3.1.3	Bayes'sche Modellierung	106
3.2	Bewertung des Cottbuser Ostsees	111
3.2.1	Ein neuer See in der Lausitz	112
3.2.2	Ergebnisse der Befragung	114
3.2.3	Berechnung der Zahlungsbereitschaft	118
4	Zwischenfazit	123
II	Nutzentransfer	127
5	Klassischer Nutzentransfer	129
5.1	Nutzentransfer und Umweltbewertung	129
5.2	Herkömmliche Verfahren	132
5.2.1	Value Transfer	133
5.2.2	Benefit Function Transfer	144
5.3	Validitätstests für den Nutzentransfer	155
5.4	Die Grenzen klassischer Verfahren	158
6	Bayes'scher Nutzentransfer	161
6.1	Bayes'sche Verfahren als Chance für den NT	161
6.2	Anwendung auf eine simulierte Befragung	164
6.3	Bayes'scher NT für Dichotomous Choice Daten	168
6.3.1	Einfacher Bayes-Transfer	172
6.3.2	Der Power Prior	178
6.4	Bayes'scher NT für den Cottbuser Ostsee	180
6.5	Ein neuer Test für den Nutzentransfer	187
7	Zusammenfassung und Ausblick	191

Abbildungsverzeichnis

2.1	Parametrische Modellierung von CVM Daten	63
2.2	Graphische Analyse der Konvergenz von Markov-Ketten .	84
2.3	Veränderung des Konfidenzintervalls	85
3.1	Nichtparametrische Analyse der Erstantworten	104
3.2	Häufigkeiten aller Antworten	105
3.3	Graphische Analyse	108
3.4	Textauszug der Szenariodarstellung im Fragebogen	113
6.1	A-Posteriori-Verteilung im Vergleich	189

Tabellenverzeichnis

3.1	Wahre mittlere WTP aus simulierten Befragungen	95
3.2	Zustimmungen zu Erst- und Zweitgebot	97
3.3	Vergleich der Modelle zur Schätzung der WTP	99
3.4	Einfluss des Stichprobenumfangs	101
3.5	Parameterschätzung im Bivariaten Probit-Modell	103
3.6	Bayes Schätzung mit nicht-informativer A-Priori-Verteilung	107
3.7	Bayes Schätzung mit informativer A-Priori-Verteilung . . .	110
3.8	Stichprobe und amtliche Statistik im Vergleich	115
3.9	Antworten und Erstgebote	116
3.10	Ankereffekte	117
3.11	Zahlungsbereitschaftsanalyse Cottbuser Ostsee	119
3.12	Modelle mit Kovariaten	120
5.1	Aufbau einer Suchabfrage in EVRI	135
5.2	Kenngrößen (2004)	141
5.3	Zahlungsbereitschaft für den Tokkekob Hegn	143
5.4	Unterschiedliche sozioökonomische Variablen und gleiche Präfe- renzen	146
5.5	Unterschiedliche Präferenzen	147
5.6	Modelle im Primär- und Sekundärstudiengebiet	149
5.7	Parameter Tokkekob Hegn und Bevölkerungsdaten Cottbus	151
6.1	Resultate des BT für eine offene Befragung	166
6.2	Stichproben aus dem Sekundärgebiet	173
6.3	Ergebnisse des Bayes'schen Transfers	174
6.4	Unterschiedliche Modelle	176
6.5	Einfluss der Stichprobengröße	177
6.6	Tokkekob simulierte Befragung	182
6.7	Einfacher Bayes-Transfer mit Kovariaten	183
6.8	Bayes-Transfer mit simulierter Befragung	184
6.9	Power Prior	185

6.10	Der Einfluss der Stichprobengröße auf a_0	186
6.11	Der Einfluss der A-Priori-Verteilung	186
6.12	Nutzentransfer Cottbus → Tokkekob	188

Einleitung

Methoden zur Umweltbewertung befinden sich im Spannungsfeld zwischen ökonomischer Theorie, statistischer Theorie und den pragmatisch-praktischen Anforderungen der wirtschaftswissenschaftlichen Politikberatung.

Insbesondere in Deutschland sind dies scheinbar widersprüchliche Anforderungen. Anders als in den USA, wo durch den Nobelpreis geadelte Ökonomen in der politischen Beratung oder in politischen Institutionen nicht nur gern gesehen, sondern auch anzutreffen sind, besteht in Deutschland häufig eine scharfe Trennung zwischen theoretischen und empirischen Arbeiten. Theoretische Arbeiten konzentrieren sich, zunehmend unter dem Druck des "publish or perish", auf eine immer komplexere Weiterentwicklung anspruchsvollster Modelle, und laufen dabei Gefahr, den Blick für die Realität zu verlieren. Arbeiten aus der praktischen Politikberatung hingegen stehen häufig unter dem Druck der Politik, möglichst schnell möglichst verständliche Ergebnisse zu liefern und verlieren dabei die wesentlichen neuen methodischen Entwicklungen in ihrem Gebiet aus den Augen. Diese Arbeit verfolgt unter anderem das Ziel, einen Beitrag zur Überbrückung dieses Gegensatzes zu leisten, indem anspruchsvolle statistisch-ökonomische Methoden zur Verbesserung der Antworten auf eine der Praxis entstammende Fragestellung herangezogen werden.

Anlass, sich mit der Umweltbewertung zu befassen, ist die Erkenntnis, dass der Zustand der Natur oder der Umwelt uns auf vielfältige Weise beeinflusst. Wir besuchen Luftkurorte und verbringen unseren Urlaub an Stränden; Naherholungsgebiete verbessern unsere Freizeitmöglichkeiten und der Erhalt der Artenvielfalt erfreut uns, sogar wenn wir diese Tiere und Pflanzen niemals aus der Nähe sahen und sehen werden. Besonders erfreulich scheint dabei zu sein, dass uns die meisten dieser Umweltgüter kostenlos zur Verfügung stehen, d.h. die Luft am Nordseestrand muss genauso wenig pro Atemzug bezahlt werden wie ein Blick auf das einzigartige Wattenmeer. Dieser scheinbare Vorteil wird jedoch schnell zu einem Problem, besonders wenn man bedenkt, dass die Umwelt auch zur Entsorgung von Abfällen und anderen Stoffen kostenlos zur Verfügung steht. So wird die kostenlose Entsorgung von Abgasen, die bei der Produktion von Gütern anfallen, zu einer Beeinträchtigung des kostenlosen Konsums von frischer Nordseeeluft. Was außerhalb des Interesses von Ökonomen liegen würde, weil es kostenlos und unendlich verfügbar ist, wird so zum Gegenstand von konkurrierenden Verwendungszwecken und somit auch im Sinne

der Ökonomie ein knappes Gut. Es fehlt jedoch bei diesem Gut die Information, die bei anderen Gütern die zunehmende Knappheit signalisiert: der sich auf dem Markt ergebende Preis. Da Umweltgüter gerade nicht auf Märkten gehandelt werden, besteht keine Möglichkeit, dass sich dieser Preis durch Angebot und Nachfrage ergibt. Ohne diese Information ist es für den Staat unmöglich, den Einfluss von Umweltverschlechterungen auf das Wohlbefinden seiner Bürger abzuschätzen, so wie es unmöglich ist, zwischen Maßnahmen zur Verbesserung der Umweltqualität und anderen staatlichen Maßnahmen zu entscheiden.

Will der Staat auf die Interessen seiner Bürger reagieren, so antwortet er auf Ängste vor gesundheitlichen Schäden durch Umweltkatastrophen oder auf die Forderung nach erweiterten Freizeitmöglichkeiten mit politischen Maßnahmen zum Schutz der Umwelt bis hin zu staatlichen Investitionen in den Umweltschutz und in die Bereitstellung von Umweltgütern wie Naturparks oder die Gestaltung von Seen und Landschaften. Die Ausgaben, die bei derartigen im Folgenden unter dem Begriff "Umweltprojekte" zusammengefassten Maßnahmen anfallen, müssen jedoch gegenüber den Bürgern spätestens an den Wahlurnen in einer Demokratie wieder verantwortet werden.

Daher sieht sich der Staat in der Pflicht, Rechenschaft darüber ablegen zu können, ob den Kosten eines Umweltprojekts auch ein Nutzen der Bevölkerung in entsprechender Höhe gegenübersteht. Des weiteren muss der Staat bei begrenztem Budget zwischen verschiedenen Einsatzmöglichkeiten von Geldern entscheiden, denn in den Umweltschutz investierte Mittel stehen anderen Zwecken nicht mehr zur Verfügung. Bei der Durchführung umweltqualitätsverbessernder Projekte müssen somit knappe volkswirtschaftliche Ressourcen aus anderen Verwendungen abgezogen und umgewidmet werden. Das resultiert auf der einen Seite in Konsumeinbußen, die sich nutzenvermindernd auf die betroffene Bevölkerung auswirken; andererseits entstehen jedoch durch die infolge des Projekts gestiegene Umweltqualität Nutzenerhöhungen. Hier sind daher Entscheidungshilfen gefragt, die die Rechtfertigung staatlicher Investitionen unterstützen.

Kosten-Nutzen-Analysen bieten genau diese Entscheidungshilfen an. Für eine Reihe von staatlichen Maßnahmen ist die Anfertigung von Nutzen-Kosten-Analysen daher gesetzlich verankert. Sie sind für die Gewährung von EU-Beihilfen für die Durchführung von Infrastrukturmaßnahmen und umweltrelevanten Projekten in EU-Beitrittsländern vorzulegen, wie aus der Council Regulation (EC)1267/ 1999 hervorgeht. Auch die Haushaltsord-

nungen¹ von Bund, Ländern und Gemeinden schreiben Kosten-Nutzen-Analysen vor, die über die reine betriebswirtschaftliche Investitionsentscheidung hinaus gehen und gesellschaftliche Kosten und Nutzen eines im Rahmen eines staatlichen Projekts bereitgestellten Guts mit dem "measuring rod of money" (eine Formulierung, die bereits auf Pigou (1932) [197] zurückgeht) messen.

Ein wesentliches Problem bei der Bewertung umweltverbessernder Projekte mithilfe von Nutzen-Kosten-Analysen besteht nun darin, dass zwar die Kosten eines solchen Projekts im Allgemeinen recht gut ermittelt werden können, da die eingesetzten Faktoren üblicherweise auf Märkten gehandelt werden und daher mit Marktpreisen bewertet werden können. Die empirische Bestimmung der Nutzenzuwaches bei den Bürgern, etwa durch einen neuen Naturpark, in den Kosten vergleichbaren Geldeinheiten scheint zunächst ungleich schwieriger. Ohne auf die genaue theoretische Herleitung einzugehen (vgl. Kapitel 1.2) leuchtet es jedoch unmittelbar ein, dass die Zahlungsbereitschaften bzw. die Entschädigungsforderungen der Betroffenen für eine Umweltverbesserung oder -verschlechterung eine Möglichkeit darstellen, auf die Nutzenänderung durch diese Verbesserung oder Verschlechterung zu schließen, denn steigende Umweltqualität wirkt sich nutzenerhöhend aus, so dass ein Teil des Einkommens nach einer Umweltverbesserung aufgegeben werden kann und man dennoch mindestens genauso gut gestellt ist wie zuvor. Umweltverschlechterungen hingegen können durch eine Einkommenserhöhung kompensiert werden. Umweltbewertungsverfahren messen die Höhe dieser Veränderungen des Einkommens, die sich in Zahlungsbereitschaften oder Entschädigungsforderungen seitens der von der Umweltveränderung betroffenen Individuen ausdrücken lassen. Eine Schätzung der Zahlungsbereitschaften lässt sich aus Beobachtungen gewinnen, wenn mit der Nutzung eines Umweltguts Markthandlungen verbunden sind. Von den Aufwendungen für eine Fahrt an die Nordsee lässt sich beispielsweise auf den Nutzen schließen, den jemand aus seinem Aufenthalt an der See zieht.

Es hat sich jedoch herausgestellt, dass ein Umweltgut auch dann von Individuen geschätzt wird, wenn sie es *nicht* für ihre Freizeit oder einen anderen Zweck nutzen. Eine Veränderung dieses Umweltgut oder der Umweltqualität, wie etwa das Robbensterben in der Nordsee, würde bei diesen Individuen zu einer Nutzeneinbuße führen, ohne dass der nächste Urlaub

¹§ 7 Abs. 2 Bundeshaushaltsordnung (BHO) für den Bund § 6 Abs. 2 Haushaltsgrundsätzegesetz für die Länder § 10 Abs. 2 Gemeindehaushaltsverordnung für die Gemeinden.

an der Nordsee geplant sein muss. Die Beobachtung von Markthandlungen reicht daher zum Erfassen dieser Nutzeneinbuße genauso wenig aus wie zum vollständigen Erfassen des Nutzenzuwachses infolge einer Umweltverbesserung. Hierfür ist die konkrete und gezielte Befragung der Bevölkerung die einzige zuverlässige Strategie. Diese Strategie verfolgt die Contingent Valuation Methode (CVM), die die Zahlungsbereitschaft der Bevölkerung "contingent", d.h. abhängig von einem bestimmten umweltverändernden Projekt - im Beispiel einem Programm zum Schutz der Robben - misst. Die Contingent Valuation Methode ist in den letzten zwanzig Jahren fester Bestandteil des Werkzeugkastens von Ökonomen, Theoretikern wie Praktikern, geworden, die sich mit der Bewertung von nicht marktgehandelten Gütern befassen. Die Auswertung des durch CVM Interviews gewonnenen Datenmaterials haben Ökonometriker in Theorie und Praxis über die gesamte wissenschaftliche Auseinandersetzung hinweg begleitet. Gemeinsam mit Umweltökonomern, Psychologen, Naturwissenschaftlern und Sozialwissenschaftlern wurden sowohl die Durchführung der Befragungen wie auch die Behandlung der Daten in den letzten mehr als 20 Jahren verbessert. Der ökonometrische Bezugsrahmen allerdings war und ist überwiegend die klassische Statistik. Etablierte Schätzverfahren wie die Maximum-Likelihood-Methode sind auch deshalb weit verbreitet, weil Routinen für diese Verfahren in vielen handelsüblichen Statistikprogrammen implementiert sind. Dies macht es den Praktikern wiederum leichter, auch komplexe Modelle zu schätzen.

Der Bayes'sche Ansatz hingegen, die zur klassischen Modellierung alternative statistische Denkschule, ist bislang nur vereinzelt auf die Auswertung von Contingent Valuation Daten angewendet worden. Ein Ziel dieser Arbeit ist es, diese Lücke zu schließen, denn Bayes'sche Verfahren haben in anderen Wissenschaftsbereichen zu einer Bereicherung der statistischen Auswertungen und zu einem tieferen Verständnis der Modelle beigetragen und dies scheint für Contingent Valuation Daten auch wünschenswert. Darüber hinaus kann die Tatsache, dass es der Bayes'sche Ansatz erlaubt, Vorabinformationen systematisch in die Schätzung eines Modells einzubeziehen, einen Beitrag zur Diskussion um die Möglichkeiten der Übertragung vorhandener Studienergebnisse auf neue Umweltveränderungen leisten.

Dies ist besonders dann wünschenswert, wenn man sich die Kosten einer Kosten-Nutzen-Analyse mithilfe der recht aufwändigen Contingent Valuation Methode vor Augen führt, denn die Zuverlässigkeit der Methode hängt kritisch von der Anzahl der durchgeführten Interviews ab und sie

erweist sich gerade für kleinere Projekte als schlichtweg zu teuer. Im Laufe der Jahrzehnte, die sich die Forschung mit der CVM beschäftigt, sind daher verschiedene Wege erdacht worden, wie man die Kosten der Kosten-Nutzen-Analyse senken kann. Eine Möglichkeit hierzu besteht im Übergang von den kostenintensiven persönlichen Interviews auf preiswertere Mail-Surveys, telefonische Befragungen oder - in jüngerer Zeit - Internetumfragen. Diese Vorgehensweise hat jedoch den Nachteil, dass die Kontrolle über die Reaktionen und das Verständnis der Befragten verloren geht. Das genaue Verständnis des zur Bewertung gestellten Umweltguts ist aber von zentraler Bedeutung, wie in Kapitel 1.3 noch ausführlich dargelegt wird.

Mit der Zunahme an durchgeführten Studien und dem damit einhergehenden Umfang des zur Verfügung stehenden Datenmaterials kam daher bei mit der Umweltbewertung befassten Praktikern der Vorschlag des sogenannten *Nutzentransfers* auf. Die Idee hierbei ist, dass sich vorhandene Schätzungen der Nutzenänderung aufgrund einer Umweltveränderung auf geplante Umweltprojekte übertragen lassen müssten, wenn die Grundannahmen über die individuellen Präferenzen für ein Umweltgut zutreffend sind. Die Übertragung kann dabei von einer einfachen Wiederverwendung bereits gemessener Werte für neue Projekte bis hin zu komplexen Anpassungen von bereits vorliegenden Studienergebnissen an eine durch ein neues Projekt entstehende Situation gehen. Es wird sich im weiteren Verlauf dieser Arbeit zeigen, dass bei sorgfältiger Anwendung statistischer Methoden die Abweichungen von einer eigens durchgeführten Studie bei geringem Kostenaufwand sich durchaus sehr klein halten lassen. Inwieweit sich jedoch bei dieser Maßnahme die Abweichungen von der theoretischen Fundierung auch gering halten lassen, wird im weiteren Verlauf der Arbeit noch intensiv zu diskutieren sein.

Die vorliegende Arbeit gliedert sich in zwei Teile. Dieser Einleitung folgt der erste Teil, der die Durchführung einer Studie zur Umweltbewertung unter Verwendung der Contingent Valuation Methode beschreibt. Hierzu wird im *ersten* Kapitel zur Umweltbewertung zunächst ein Überblick über die verschiedenen Anwendungsgebiete der Umweltbewertung und die sich daraus ableitenden Methoden gegeben und die Contingent Valuation Methode aufgrund ihrer Fähigkeit zur umfassenden Umweltbewertung hervorgehoben. Anschließend wird die der CVM zugrunde liegende Theorie der Messung individueller Wohlfahrtsänderungen beschrieben. Kapitel 1 schließt mit einem Überblick über die grundlegenden Schritte einer CVM Studie in der Praxis. Da das Ziel der Arbeit aber nicht die Ergänzung der Be-

fragungsmethode, sondern eine Bereicherung der statistischen Auswertung ist, wird im *zweiten* Kapitel aus den ökonomischen Grundlagen der CVM das statistische Modell entwickelt, welches mit den theoretischen Ansätzen kompatibel ist. Es zeigt sich, dass es eine Reihe von klassischen parametrischen und nicht-parametrischen Ausprägungen dieses Modells gibt, die in der Literatur ebenbürtig nebeneinander stehen. Das zweite Kapitel schließt mit der Darstellung eines alternativen statistischen Ansatzes und entwickelt ausführlich das statistische Modell in der Gedankenwelt der Bayes-Statistik. Da die Gemeinde der "Bayesianer" unter den mit der Umweltbewertung befassten Wissenschaftlern noch sehr klein ist, werden die Chancen, die die Verwendung Bayes'scher Statistik für die Umweltbewertung birgt, ausführlich erläutert und die Methoden beschrieben, die zur praktischen Anwendung beherrscht werden müssen.

Im *dritten* Kapitel der Arbeit werden die verschiedenen Methoden angewendet und verglichen. Zunächst werden die Daten einer Monte Carlo Simulation vorgestellt, um einen Eindruck zu vermitteln, wie die Ergebnisse aussehen, wenn "alles glatt läuft". Daher wird zuerst die klassische Auswertung durchgeführt, um danach zu zeigen, wie es im Bayes'schen Modell gelingt, einige in der Praxis auftretende Probleme der klassischen Auswertung zu überwinden. Simulationen sind immer dann hilfreich, wenn es darum geht, einen Überblick über die Leistungsfähigkeit verschiedener Ansätze zu gewinnen, denn das "wahre Modell" ist bei Simulationsmodellen durch den datengenerierenden Mechanismus bekannt. In der Realität liegt den gemessenen Zahlungsbereitschaften gerade kein solcher Mechanismus zugrunde, sondern bestenfalls die Präferenzen einer heterogenen Bevölkerung. Daher wird als zweites Anwendungsbeispiel eine im Sommer 2003 durchgeführte Befragung herangezogen, an der die klassische Auswertung von "real world data" gezeigt wird. Der erste Teil der Arbeit schließt mit einem Zwischenfazit im *vierten* Kapitel.

Der zweite Teil der Arbeit ist einer Methode gewidmet, die ursprünglich aus der Bewertungspraxis und der Not, kurzfristig und kostengünstig eine Bewertung durchzuführen, entstammt. Beim *Nutzentransfer* werden die für andere, ähnliche, Projekte in der Vergangenheit gemessenen Nutzen auf ein geplantes Projekt "übertragen" und so eine Abschätzung des zukünftigen Nutzens gewonnen. Das *fünfte* Kapitel führt in die klassischen Überlegungen zu diesem Verfahren ein. Es zeigt sich sehr schnell, dass die derzeitigen Verfahren zum Nutzentransfer sowohl theoretischen Mängeln unterliegen als auch in der Praxis zu unbefriedigenden Ergebnissen führen. Wieder-

um mithilfe simulierter Daten wird demonstriert, warum die beobachteten Mängel unvermeidlich sind. Eine derartige Überprüfung der herkömmlichen Ansätze zum Nutzentransfer unter Verwendung simulierter Daten ist bislang in der Literatur nicht durchgeführt worden. Anhand der Daten der durchgeführten Studie aus dem ersten Teil der Arbeit und einer Vergleichsstudie lässt sich zeigen, was geschieht, wenn die herkömmlichen Verfahren in der Realität zu einem Nutzentransfer verwendet werden. Dabei wird ausführlich auf das Problem des Praktikers eingegangen, der für einen Nutzentransfer aus der Fülle von bereits durchgeführten Studien diejenige auswählen muss, die für das ihm vorliegende Problem die geeignetste Studie ist. Zur Erleichterung dieser Wahl existiert seit einigen Jahren eine hilfreiche Systematik. Das *Environmental Valuation Research Inventory* (EVRI) ist eine Datenbank, die der wissenschaftlichen Gemeinschaft mit dem Ziel, den Nutzentransfer zu verbessern, zur Verfügung steht.

Das *sechste* Kapitel führt die Erkenntnisse der vorhergehenden Kapitel zusammen und schlägt die Möglichkeit eines Bayes'schen Nutzentransfers vor. Die vorgeschlagene Verbesserung besteht darin, im Gebiet des Projekts, auf welches der für frühere Projekte gemessene Nutzen "transferiert" werden soll, ebenfalls eine kleine, kostengünstige Erhebung durchzuführen und anschließend Informationen aus bereits durchgeführten Studien mithilfe der Ergebnisse der Vor-Ort-Studie nach dem Bayes'schen Ansatz aufzudatieren. Eine zentrale Bedeutung kommt im Bayes'schen Modell der Formalisierung von bereits vorliegenden Informationen zu - der Wahl der sogenannten A-Priori-Verteilung. Daher wird im sechsten Kapitel ausführlich auf verschiedene Modellierungen hierzu eingegangen und es kommen Verfahren zur Anwendung, die bislang weder für den Nutzentransfer noch für die Bewertung von Umweltgütern insgesamt eingesetzt wurden.

Der Bayes'sche Nutzentransfer wird im letzten Teil von Kapitel 6 wiederum auf simulierte Daten angewendet, um das beste Modell zu wählen. Mit diesem Modell wird zum Schluss ein Nutzentransfer durchgeführt und die Ergebnisse kommen den tatsächlich gemessenen Zahlungsbereitschaften erstaunlich nahe. Um diese Aussage auch zu belegen, wird in Kapitel 6 darüber hinaus ein neuer Test entwickelt, mit dem erstmals die Qualität eines in der Realität durchgeführten Nutzentransfers überprüft werden kann. Natürlich wirft die Beantwortung von Fragen immer auch neue interessante Aspekte auf und so beschließt das *siebte* Kapitel die Arbeit mit einem Ausblick auf weitere Forschungsfragen.

Teil I

**Studien zur
Umweltbewertung**

Kapitel 1

Umweltbewertung und Contingent Valuation

1.1 Einführung zur Umweltbewertung

Die Frage nach dem Wert der uns umgebenden Natur stellt sich in verschiedenen wirtschaftspolitischen Bereichen. Sie spielt immer dann eine Rolle, wenn Aufwendungen zum Erhalt der natürlichen Umwelt oder zur Verbesserung der Umweltqualität getätigt werden und diesen Investitionen der geschaffene Wert gegenübergestellt werden soll, oder wenn die Umwelt durch menschliche Aktivitäten geschädigt wurde und der daraus resultierende Wertverlust bestimmt werden soll. In beiden Fällen gilt es, den Gewinn oder Verlust an Umweltqualität in Geldeinheiten zu übersetzen, damit den entstandenen Kosten entsprechende Zuwächse gegenüberstehen, beziehungsweise die entstandenen Einbußen kompensiert werden können. Dies ist das Ziel monetärer Umweltbewertungsverfahren; sie messen die Auswirkungen einer Umweltveränderung auf das menschliche Wohlergehen in Geldeinheiten.

In dieser Einführung wird der Frage nachgegangen, zu welchen Bereichen die monetäre Umweltbewertung einen Beitrag leisten kann und welche Verfahren sie dabei anwendet. Hierfür wird zunächst erläutert, was man unter dem Wert der natürlichen Umwelt versteht, bevor auf die Verfahren die zur Bewertung eingesetzt werden, näher eingegangen wird. Es wird sich herausstellen, dass zur Kosten-Nutzen-Analyse eines umweltverbessernden Vorhabens Bewertungsverfahren von Vorteil sind, die auf umfangreichen Bevölkerungsbefragungen basieren. Im Verlauf dieser Arbeit werden Vor-

schläge entwickelt, wie sich der - kostenintensive - Umfang dieser Umfragen reduzieren lässt, ohne dass die Aussagefähigkeit der Ergebnisse gemindert wird. Um den Stellenwert dieses Verfahrens in einen größeren Zusammenhang einordnen zu können, wird in dieser Einführung kurz auf die weiteren möglichen Anwendungsbereiche monetärer Umweltbewertung und die dort jeweils angewendeten Verfahren eingegangen.

Der Wert der Natur

Wenn die Ökonomie nach dem Wert der Natur fragt, so ist dies immer der Wert der Natur für den Menschen. Zunächst tragen die Möglichkeiten der produktiven und konsumtiven Nutzung zu diesem Wert bei. Die natürliche Umwelt wird zum Beispiel in der Produktion von Nahrungsmitteln zum Anbau, zur Viehzucht und für die Fischerei genutzt und in der forstwirtschaftlichen Produktion wird Brenn- und Bauholz hergestellt. Die Bevölkerung konsumiert Luft und Wasser und nutzt die Natur zur Erholung und zur Freizeitgestaltung. Wenn sich die natürlichen Gegebenheiten verändern, so ändert sich auch der Wert der Natur für die Produktion oder den Konsum. Darüber hinaus scheinen die Menschen jenseits der Wertschätzung der Qualität eines Naherholungsgebiets oder des Ziels einer Urlaubsreise auch ein Interesse daran zu haben, dass die Umwelt in Gegenden unverändert intakt bleibt oder nach einer negativen Veränderung wiederhergestellt wird, in die sie selbst zu ihren Lebzeiten nicht mehr reisen werden.

Anders ausgedrückt heißt das, dass die Umwelt nicht nur einen Wert für Individuen hat, wenn sie eine ihrer Dienstleistungen tatsächlich nutzen, sondern scheinbar auch durch ihre bloße Existenz. Diese Erweiterung des Wertbegriffs wurde 1964 von Weisbrod [249] vorgeschlagen. Bis dahin galt aus Sicht der klassischen Wohlfahrtstheorie, dass ein Naturpark, der nicht aus Eintrittsgeldern finanzierbar ist, gewissermaßen seine Existenzberechtigung verliert, also nicht mehr mit öffentlichen Ausgaben bereitgestellt werden muss (vgl. Weikard (2002) [248]). Weisbrod (1964) argumentierte, dass die tatsächlich am Kassenhaus geäußerte Zahlungsbereitschaft den Wert des Parks unterschätzt, da es Individuen gibt, die den Park in der Zukunft besuchen wollen, und so seine Existenz durchaus schätzen. Diese Menschen möchten sich gewissermaßen die Option eines späteren Besuch erhalten. Krutilla (1967) [159] erweiterte diese Sichtweise um zwei zusätzliche Quellen der Wertschätzung, die sich ebenfalls nicht in einem

für den direkten Besuch (Gebrauch) bezahlten Entgelt niederschlagen, i.e. den Existenzwert, bei dem Menschen beispielsweise die Existenz von Naturschutzgebieten schätzen, ohne jeden einzelnen Naturpark besuchen zu wollen, und den Vererbungswert, der reflektiert, dass sie die unberührte Natur auf lange Sicht hin gewährleistet sehen möchten, damit nachfolgende Generationen dies auch noch genießen können. Am Kassenhaus eines Naturparks werden die Menschen, die die bloße Existenz des Parks schätzen, zwar niemals gesehen, nichtsdestotrotz haben sie eine Zahlungsbereitschaft für den Erhalt des Parks.

Heutige Literatur zur Umweltbewertung sieht den Wert eines Umweltguts als bestehend aus sogenannten Gebrauchs- und Nichtgebrauchswerten an. Vererbungs- und Existenzwert gehören zu den Nichtgebrauchswerten, denn falls sie in eine Zahlungsbereitschaft münden, so ist diese nicht an einen Gebrauch des Umweltguts geknüpft. Der Optionswert lässt sich beiden Kategorien zuordnen, da ihm wenn auch kein derzeitiger so doch ein zukünftig geplanter Gebrauch zugrunde liegt (vgl. Turner et al. (1994) [244], Randall (1991) [207], Shechter und Freeman (1994) [227] oder Hanley et al. (1997) [122] für die Zuordnung zum Gebrauchswert, anders hingegen Kolstad (2000) [157] oder Field (1997) [97]). Eine noch weitergehende, oftmals als Ökozentrismus bezeichnete Interpretation des Wertebegriffs (vgl. Ott (2003) [199] und die darin zitierte Literatur) schreibt der Natur einen "Wert an sich" zu. Wenn die Natur einen derartigen intrinsischen Wert hat, so bedeutet dies, dass sie sich durch kein "Geld der Welt" aufwiegen lässt. Diese Wertzuschreibung liegt jedoch außerhalb des Interessengebiets der Ökonomie. Der aus dem Options-, Existenz- und Vererbungsmotiv resultierende Wert hingegen sollte berücksichtigt werden, wenn der Staat die Kosten und den Nutzen einer Investition in die Bereitstellung eines Umweltguts gegeneinander abwägen will. Aus dieser Forderung leitet sich, wie nachfolgend gezeigt wird, ein wichtiges Entscheidungskriterium für die Wahl des zur Bewertung verwendeten Verfahrens ab.

Der Nutzen eines umweltverbessernden Projekts lässt sich durch die Zahlungsbereitschaft der von diesem Projekt betroffenen Bürger für diese Maßnahme abschätzen, wie in der Einleitung bereits erläutert wurde. Zur Ermittlung der individuellen Zahlungsbereitschaften oder auch der Entschädigungsforderungen im Falle einer Umweltverschlechterung kennt die Ökonomie eine Reihe von Verfahren, die nachfolgend kurz vorgestellt und diskutiert werden. Diese Verfahren lassen sich zunächst grob in *indirekte* und *direkte* Methoden der Umweltbewertung unterscheiden.

Indirekte Methoden der Umweltbewertung

Bei den indirekten Methoden der Umweltbewertung wird von beobachtbaren Markthandlungen der Individuen auf den Nutzen eines Umweltguts geschlossen. Die Voraussetzung hierfür ist die Existenz einer Komplementaritäts- oder Substitutionsbeziehung zwischen bestimmten Marktgütern und dem zu bewertenden Umweltgut. Liegt beispielsweise schwache Komplementarität vor, so lässt sich ein Marktgut finden, so dass aus Änderungen der Nachfrage nach diesem Gut auf die Nutzenänderung durch eine Umweltveränderung geschlossen werden kann. Wenn dieses schwach komplementäre Marktgut nicht konsumiert wird, so hat die Umweltveränderung auch keinen Einfluss auf den Nutzen des Individuums. Werden beispielsweise keine Angellizenzen erworben, weil sie eventuell zu teuer sind, so wird die Verbesserung oder die Verschlechterung der Wasserqualität in einem Angelgebiet den Nutzen des Individuums nicht beeinflussen. Auf einer Erweiterung dieses Arguments beruht die sogenannte *Reisekostenmethode* (Travel Cost Method), die die schwache Komplementaritätsbeziehung zwischen dem zu bewertenden Umweltgut und den Aufwendungen für die "Reise" zu diesem Umweltgut nutzt. Der Besucher beispielsweise eines Naturparks entscheidet sich, die Reisekosten aufzuwenden, um in den Genuss dieses Umweltguts zu gelangen.

Derartige Aufwendungen bestehen zum einen aus den echten Fahrtkosten, wie den Kosten für Treibstoff bei der Anreise im Auto, zum anderen aus den Kosten für Konsumgüter, die für die Nutzung des Umweltguts notwendig sind, und nicht zuletzt aus der für die Fahrt aufgewendeten Zeit. Der Haushalt setzt Benzin, Zeit, Gummistiefel und Feldstecher ein, um in den Genuss (zur "Produktion" (vgl. Lancaster (1971) [163])) des Guts "Vogelbeobachtung im Vogelschutzgebiet" zu kommen. Um tatsächlich einen monetären Wert des Umweltguts aus der Nachfrage nach den komplementären Gütern ableiten zu können, werden diese mit ihren Marktpreisen bewertet. Das ist bei den benötigten Konsumgütern wiederum kein Problem, erfordert jedoch auch eine Bewertung der aufgewendeten Zeit. Häufig wird die Fahrzeit mit den Löhnen oder Gehältern der Besucher bewertet; d.h. es werden die Opportunitätskosten einer entgangenen Arbeitsstunde angesetzt. Die wenigsten Besucher stehen jedoch als Arbeitnehmer vor der Wahl, entweder einige Stunden länger zu arbeiten oder diese für die Fahrt zu einem Naherholungsgebiet aufzuwenden. Die Literatur kennt eine lebhafte Diskussion zur Bewertung der Reisezeit in Travel Cost Studien. So führen Bockstael et al. (1987) [36] die Bewertung der Reisezeit in Abhängig-

keit von den Beschäftigungsverhältnissen der Besucher ein, d.h. die Bewertung mit einer Art Stundensatz wird tatsächlich nur bei selbständig tätigen Besuchern angewendet. McConnell und Strand (1981) [188] wählen einen einfacheren ad hoc Ansatz, der die Reisezeit mit einem festen Anteil des jeweiligen Lohnsatzes bewertet (zum Vergleich verschiedener Modellierungen vgl. Haab und McConnell (2002) [113].).

Trotz dieser letztlich ungeklärten Frage nach der besten Bewertung der eingesetzten Zeit wird die Reisekostenmethode wegen ihrer einfachen Durchführbarkeit häufig zur Bewertung von Freizeit- und Erholungsgebieten angewendet, sie hat jedoch gegenüber ihrem ursprünglichen Anwendungsgebiet einen Bedeutungswandel erfahren. Während sie von Hotelling (1947) (nach Garrod und Willis (1999) [98]) vorgeschlagen wurde, um die politische Abwägung zwischen der produktiven Landnutzung (etwa für den Ackerbau) und der Nutzung von Flächen für die Erholung der Bevölkerung zu erleichtern, ist sie heute im Einsatz zur Bewertung von Umweltqualität. Hierbei müssen die Reisekosten zu verschiedenen Zielen mit unterschiedlichen Qualitätsmerkmalen erhoben werden, was für eine aussagekräftige Studie mit einem ganz erheblichen Aufwand verbunden ist.

Die *Methode der Hedonischen Preise* leitet sich ebenfalls aus einer ursprünglich von Lancaster (1966) [162] entwickelten Idee zum Nachfrageverhalten der Haushalte ab. Sie basiert auf der Annahme, dass die Güterpreise sich als Funktionen verschiedener Charakteristika der Güter ausdrücken lassen. Auf die Bewertung von Umweltqualität übertragen heißt dies, dass die für bestimmte Marktgüter bezahlten Preise unter vielen anderen Charakteristika auch verschiedene Aspekte der mit diesen Gütern verbundenen Umweltqualität widerspiegeln sollten. Dies lässt sich, so die grundlegende Idee zu dieser Methode, die auf Ridker (1967) [214], Griliches (1971) [108] und Rosen (1974) [215] zurück geht, besonders gut auf dem Immobilienmarkt beobachten, wo sich die Haushalte beim Erwerb oder der Miete von Immobilien für ein Bündel verschiedener Merkmale entscheiden, unter denen sich die Umweltqualität befindet. Der Preis für diese Immobilien lässt nun als aus verschiedenen Teilpreisen für die einzelnen Merkmale zusammengesetzt denken. Aus der Beobachtung des Immobilienmarkts, der Immobilienpreise und der verschiedenen Charakteristika der nachgefragten Immobilien lässt sich dann die hedonische Preisfunktion ableiten, die die tatsächlich gezahlten Immobilienpreise als Funktion der Umweltqualität und der sonstigen Charakteristika eines Hauses oder einer Wohnung angibt. Hieraus wird die Wertschätzung für eine Änderung der Umwelt-

qualität bei ansonsten konstant gehaltenen anderen Ausprägungen der Immobilien abgeleitet. So würde sich beispielsweise eine Präferenz für "Ruhe" in einer erhöhten Zahlungsbereitschaft für Grundstücke beobachten lassen, die sich gegenüber anderen Grundstücken durch besonders wenig Lärm auszeichnet. Umgekehrt lässt sich beobachten, dass Grundstücke in lauten Gegenden, etwa der Einflugschneise eines Flughafens, häufig deutlich preiswerter sind als ähnliche Grundstücke außerhalb der Einflugschneise. Allerdings kann gerade auf dem Immobilienmarkt oftmals nicht wie auf einem vollkommenen Markt mit einem so hinreichend großen Angebot und einer genügend großen Käuferflexibilität gerechnet werden, dass jeder Käufer tatsächlich seinen optimalen Mix aus Charakteristika frei wählen kann. Darüber hinaus ist die Methode der hedonischen Preise nur für eine sehr begrenzte Auswahl von Umweltgütern einsetzbar, nämlich diejenigen, die auf den Immobilienmarkt überhaupt einen Einfluss ausüben. Da sich Naturparks oder Vogelschutzgebiete jedoch gerade in unbewohnten Gegenden befinden, müssen hier mit Sicherheit andere Bewertungsverfahren zum Einsatz kommen.

Besteht hingegen eine Substitutionalitätsbeziehung zwischen einem Marktgut und dem zu bewertenden Umweltgut, so kann aus der Nachfrage nach dem Marktgut ebenfalls indirekt auf die Wertschätzung des Umweltguts geschlossen werden. Derartige Zusammenhänge macht sich die *Vermeidungskostenmethode* (Averting Behavior Method) zunutze, bei der aus den Aufwendungen für ein Gut, mit dem schlechte Umweltqualität "ersetzt" wird, auf den Wert einer verbesserten Umweltqualität geschlossen wird. So lässt sich von den Aufwendungen für Lärmschutzfenster auf die Wertschätzung einer Lärminderung schließen. In Entwicklungsländern werden mithilfe dieser Methode Projekte zu Verbesserung der Trinkwasserqualität bewertet. In Ländern und Gebieten, in denen keine ausreichende Versorgung mit trinkbarem Leitungswasser gewährleistet ist, wird die Verbesserung der Versorgung mit den Aufwendungen für in Flaschen abgefülltes Wasser bewertet (vgl. Um et al. (2002) [245]).

Die Attraktivität der indirekten Methoden insgesamt liegt darin, dass sie sich auf beobachtbare Marktdaten stützen können, allerdings ist darin auch ein Nachteil begründet. Indem sich die Untersuchung auf Handlungen konzentriert, die auf der Nutzung der Ressource oder des Umweltguts beruhen, berücksichtigt sie im Wert des Umweltguts auch nur die sogenannten Gebrauchswerte. Umweltgüter sind aber, wie weiter oben ausführlicher erläutert, häufig durch einen hohen Anteil von Nichtgebrauchswerten ge-

kennzeichnet. So hat zum Beispiel der Schutz der Wale oder das Verbot des Walfangs für die wenigsten Menschen auf der Welt einen Wert, der sich aus der direkten Nutzung ableiten lässt, dennoch ist eine positive Zahlungsbereitschaft für den Schutz der Wale verbreitet, wie zahlreiche Spendenaktionen belegen.

Für die Umweltbewertung im Rahmen von Kosten-Nutzen-Analysen ist somit ein Bewertungsverfahren von Interesse, das in der Lage ist, möglichst alle Facetten des Werts zu ermitteln. Darüber hinaus gilt das Interesse des politischen Entscheidungsträgers häufig Projekten, die erst in Zukunft realisiert werden sollen, und über deren Realisierung mittels einer Kosten-Nutzen-Analyse entschieden werden soll. Zur Bewertung von in der Zukunft liegenden Projekten jedoch können die indirekten Methoden keinen Beitrag leisten, sondern man muss vielmehr zu anderen Methoden der Umweltbewertung übergehen, die auf direkten Befragungen der betroffenen Bevölkerung basieren.

Direkte Methoden der Umweltbewertung

Die bekanntesten und am weitesten verbreiteten direkten Methoden sind die *Contingent Valuation Methode* (CVM) und die sogenannten *Attribute Based Choice Modelling Methoden* (ABCM). Letztere umfassen verschiedene Techniken, deren Gemeinsamkeit darin besteht, dass sie auf die Bewertung einzelner Charakteristika eines Umweltguts abzielen und somit wiederum auf Lancasters bereits zitierten Ansätzen [162] [163] fußen.

Die definitorische Abgrenzung der einzelnen ABCM-Techniken ist oft nicht ganz klar in der Literatur; man findet Contingent Ranking Studien, bei denen verschiedene Kombinationen einzelner Merkmalsausprägungen von den Befragten in eine Rangfolge gebracht werden müssen, Contingent Rating Studien, in denen diese Reihenfolgen zahlenmäßig bewertet werden, sowie Choice Experimente, bei denen die verschiedenen Kombinationen jeweils mit einem Preis versehen sind. Letztlich sind diese Techniken Derivate der dem Marketing entstammenden *Conjoint Methode*, deren grundlegende Idee hier kurz skizziert wird.

Vor der Einführung komplexer neuer Produkte möchten die Hersteller wissen, in welcher Kombination bestimmte Eigenschaften ihres zukünftigen Produkts am Markt honoriert werden. So kann eine Fluglinie zur Planung ihrer zukünftigen Angebote ermitteln, in welchem Umfang ihre Gäste mehr

Beinfreiheit gegenüber einer besseren Verköstigung bevorzugen und inwiefern die einfache Handhabung einer Flugbuchung eine Rolle spielt. Diese Produkte befinden sich nicht in unterschiedlichen Ausprägungen auf dem Markt und es lassen sich somit keine Marktdaten aus der direkten Beobachtung ermitteln. Daher werden mögliche oder tatsächliche Kunden in einer direkten Befragung aufgefordert, sich zwischen auch im Preis unterschiedlichen Angeboten, in denen einzelne Merkmale des neuen Guts zusammengefasst sind (conjoint), zu entscheiden. Genauer gesagt werden in einer Conjoint-Studie die Kunden aufgefordert, sich zwischen verschiedenen paarweise vorgelegten Kombinationen verschiedener Ausprägungen der jeweiligen Charakteristika, die zu unterschiedlichen Preisen des Gesamtpakets führen, jeweils für diejenige Kombination zu entscheiden, die ihnen am stärksten zusagt. Übertragen auf die Bewertung einer Umweltverbesserung bedeutet dies, dass den Befragten in einem Choice Experiment verschiedene Ausprägungen dieser Verbesserung zur Auswahl vorgelegt werden. Soll beispielsweise ein neuer See in einem ehemaligen Braunkohletagebau gestaltet werden, so wären im Rahmen einer solchen Untersuchung eine Vielzahl von Ausgestaltungsmerkmalen dieses Sees - wie die Anlage und Größe eines Strands, eines Naturparks, eines Hafens mit verschiedenen Nutzungsmöglichkeiten - zu Auswahl gestellt, verbunden mit verschiedenen "Preisen", für die die jeweilige Kombination zu haben wäre.

Die Anforderungen an die Befragten in einer solchen Studie sind daher recht hoch. Sie müssen sich nicht nur auf die Beschreibung eines neuen Umweltzustands einlassen, sondern auch noch auf variierende Merkmale dieses neuen Zustands und auf die damit einhergehenden verschiedenen Preise für das jeweilige Angebot. Im Ergebnis trifft eine Choice-Studie Aussagen über die Teil-Zahlungsbereitschaften für die jeweiligen Charakteristika. Damit zu den verschiedenen Ausgestaltungsmerkmalen jeweils belastbare Aussagen über die Teil-Zahlungsbereitschaften gemacht werden können, sollte eine große Anzahl von Personen befragt werden. Der erhebliche Aufwand einer Choice-Studie und die damit einhergehenden statistischen Probleme (vgl. hierzu Ahlheim und Frör (2003) [5]) lassen sich allerdings nur dann rechtfertigen, wenn der Staat zwischen verschiedenen Ausprägungen eines Investitionsprojekts entscheiden muss und sich tatsächlich für die verschiedenen Teilwerte interessiert, aus denen sich die Zahlungsbereitschaft für eine bestimmte Ausgestaltung zusammen setzt. In diesem Fall können AB-CM Studien eine sinnvolle Entscheidungshilfe darstellen.

Ist das Ziel jedoch die Kosten-Nutzen-Analyse eines konkreten geplanten Umweltguts, so bietet sich die Verwendung der *Contingent Valuation Methode* (CVM) an, die sich auf ein bestimmtes Projekt bezieht, ohne Variationen in einzelnen Merkmalen zu betrachten. Diese von Ciriacy-Wantrup (1947) [66] erstmals vorgeschlagene Methode beruht auf der Ermittlung der individuellen Zahlungsbereitschaften für ein Umweltgut auf der Basis von Interviews mit Personen, denen dieses Umweltgut zum "Kauf" angeboten wird. Wegen der in der Realität fehlenden Märkte wird im Rahmen eines solchen Interviews für die Befragten ein hypothetischer Markt konstruiert, auf dem sie das Umweltgut erwerben können. Die Contingent Valuation Methode teilt also mit den vorher besprochenen ABCM Techniken die Tatsache, dass sie das zu bewertende Umweltgut zunächst für die Befragten entwerfen und beschreiben muss. Dies stellt einen wesentlichen Unterschied zu den indirekten Verfahren dar. Während dort das Umweltgut, wie etwa ein Naturpark, existiert und die Befragten die Entscheidung zum "Kauf" durch den Besuch fällen, muss hier das Umweltgut den Befragten erst nahe gebracht werden. Mit zunehmender Fähigkeit einer Methode, umfassende Werte zu bestimmen, steigen somit die Ansprüche an die Messinstrumente, hier an den Fragebogen, der in den Interviews eingesetzt wird. Im Unterschied zu den ABCM-Techniken werden jedoch die einzelnen Merkmale des zu bewertenden Sees oder Parks festgelegt, so dass den Befragten nur ein einziges sogenanntes Szenario zur Bewertung vorgelegt wird.

Die Einzelheiten eines CVM-Interviews werden im weiteren Verlauf dieser Arbeit ausführlich beschrieben und diskutiert, so dass hier nur auf die prinzipielle Funktionsweise eingegangen wurde. Das Ergebnis einer CVM-Studie sind Aussagen zu den Zahlungsbereitschaften einzelner Individuen aus einer Stichprobe der gesamten von der Umweltveränderung betroffenen Bevölkerung. Damit diese Aussagen für die Kosten-Nutzen-Analyse verwertbar sind, müssen sie zunächst zu einem Stichprobenmittelwert zusammengefasst werden und anschließend auf die gesamte Bevölkerung hochgerechnet werden (zu den damit verbundenen Schwierigkeiten vgl. Kap. 1.2). So einleuchtend der Ansatz erscheint, so viel Kritik und Widerspruch hat die Methode doch in den letzten 15 Jahren erfahren (vgl. Kap. 1.3 und die darin enthaltene Literatur). Dennoch sind zur Kosten-Nutzen-Analyse gerade bei Projekten in Entwicklungs- und Schwellenländern zahllose Studien für die praktische Politikberatung mittels der Contingent Valuation Methode durchgeführt worden (vgl. z. B. Ardila et al. (1998) [17] für eine Übersicht).

Zusammenfassend lässt sich festhalten, dass es eine Reihe von Verfahren gibt, mit denen sich der Nutzen eines staatlich bereitgestellten Umweltguts monetär bewerten lässt. Die auf der Hand liegende Frage, ob diese Verfahren nicht auch zum Einsatz kommen sollten, wenn es darum geht, Entschädigungen für eine Umweltverschlechterung, etwa nach einem Umweltunfall, festzulegen, wird im nachfolgenden Abschnitt beantwortet. Dabei wird sich herausstellen, dass der Vorteil gerade der Contingent Valuation Methode bei der umfassenden Bewertung eines Umweltguts bei der Schadensbemessung zu ihrem Nachteil gerät. Die im Rahmen von Entschädigungsforderungen zur Kasse gebetenen Verursacher eines Schadens haben natürlich ein eher geringes Interesse an einer Bewertung nicht nur des verlorenen Gebrauchs- sondern auch noch des Nichtgebrauchswerts.

Die Bewertung von Umweltschäden

Die wohl umfassendste Verankerung von Entschädigungsregelungen nach einem Umweltschaden finden sich in den Gesetzestexten der USA. So erließ bereits Präsident Carter 1980 in Reaktion auf einen großen Umweltskandal den Comprehensive Environmental Response, Compensation and Liability Act (CERCLA). Die wichtigste Botschaft von CERCLA war die Verankerung des Verursacherprinzips in der Gesetzgebung zum Umgang mit Gefahrstoffen. Jedoch enthielt das Gesetz eine Klausel, durch die die Umweltbewertung durch die Hintertür ins Spiel kam. Regierungsbehörden erhielten das Recht, auf Schadensersatz zu klagen bei Schäden, die durch Gefahrstoffe an natürlichen Ressourcen entstanden, die in ihrer treuhänderischen Verwaltung lagen (z. B. Seen, Flüsse, Wälder, Feuchtgebiete etc.). Daraus resultierte die Aufforderung an das Innenministerium, einen Kriterienkatalog zu entwickeln, der die kompensierbaren Schäden genau auflistet und die Bewertungsverfahren benennt, die zur Schadensbemessung zugelassen sind. Der 1986 in einer ersten Fassung verabschiedete Katalog enthält ein eindeutiges Votum für die Messung von geschädigten Gebrauchswerten und lässt die Bewertung von Nichtgebrauchswerten nur dann zu, wenn keine Gebrauchswerte festzustellen sind. Der Kriterienkatalog des Innenministeriums wurde in Folge mehrerer Prozesse in Frage gestellt, so dass es 1989 zu einer Revisionsgerichtsentscheidung kam, die das Innenministerium explizit anweist, Nichtgebrauchswerten denselben Stellenwert zuzuweisen wie Gebrauchswerten. In der Begründung heißt es: "(...) regulation prescribing hierarchy of methodologies by which lost use value of natural resources could be measured, that focuses exclusively on market values

for such resources when market values are available, is not a reasonable interpretation of CERCLA.” (vgl. Lehr (2002) [166]).

Seit 1995 wird eine Neuauflage, Fortschreibung oder Wiedereinführung von CERCLA mindestens einmal jährlich im Congress und auch im Senat diskutiert. Die teilweise erbittert geführte Auseinandersetzung konzentriert sich hierbei auf den Einschluss der Nichtgebrauchswerte in die Schadensbemessung. So hat die Senatsvorlage S.1285 (1995) explizit die Aufhebung der Verwendung von Nichtgebrauchswerten in Title VII zum Gegenstand. Hier heißt es in der Beschreibung dieser Gesetzesvorlage: "The bill writes a new definition of 'natural resource' which includes a 'commitment for use' provision requiring that the resource be 'subject to public use or to a planned public use, for which there is an authorized and documented legal, administrative, budgetary, or financial commitment.' It eliminates non-use damages (...)." Mit zunehmenden Fällen außergerichtlicher Regelungen, bei denen das Konzept der Nichtgebrauchswerte zum Tragen kam, nahmen auch die Widerstände gegen dieses Konzept zu.

In anderen Gesetzen in den USA sind Nichtgebrauchswerte durchaus ebenfalls Bestandteil der Schadensbemessung - so zum Beispiel beim Clean Water Act, dem Deepwater Port Act, dem Trans Alaska Pipeline Act oder dem Outer Continental Shelf Lands Act. Das lauteste Medienecho erhielten die Nichtgebrauchswerte jedoch infolge des Oil Pollution Acts, eines Gesetzes, das als Reaktion auf den Tankerunfall der Exxon Valdez verabschiedet wurde. Im März 1989 lief die Exxon Valdez im Prince William Sound vor Alaska auf Grund, fast 50 Millionen Liter Rohöl verunreinigten ca. 2000 Kilometer Küstenlinie, und ca. 250.000 Seevögel starben an den Folgen der Ölpest. Im folgenden Jahr verabschiedete der Kongress den Oil Pollution Act und beauftragte die National Oceanic Atmospheric Administration (NOAA), Kriterien zur Schadensbemessung aufzustellen. Vielleicht aufgrund der Öffentlichkeitswirksamkeit dieser Entscheidung und der hitzigen Debatte über den Einschluss von Nichtgebrauchswerten beauftragte die NOAA einen Expertenausschuss, dem die Nobelpreisträger Kenneth Arrow and Robert Solow vorstanden, mit der Beantwortung der Frage, wie zuverlässig die bekannteste Messmethode zur Bewertung von Nichtgebrauchswerten, die Contingent Valuation Methode, sei. Der Abschlussbericht dieses sogenannten NOAA Panels wurde 1993 veröffentlicht (Federal Registry, January 15, 1993) und empfiehlt die Anwendung der CVM zur Messung von Gebrauchs- und Nichtgebrauchswerten, legt allerdings einen "Best Practice"-Kriterienkatalog fest und empfiehlt die Durchführung von

Studien, die diesen erhöhten Ansprüchen genügen und als Referenzstudien dienen können.

Wenngleich es gerade zur Bewertung eines Umweltschadens in der Vergangenheit dringend geboten scheint, neben den Gebrauchswerten auch die Nichtgebrauchswerte zu erfassen und eine umfangreiche monetäre Bewertung des aufgetretenen Schadens durchzuführen, treten die heutigen Bestrebungen eher wieder hinter dieses Ziel zurück. So empfiehlt die mit der Bewertung von Schäden befasste amerikanische Umweltbehörde EPA heute, vom Einschluss von Nichtgebrauchswerten abzusehen, wenn es um die Bemessung eines Umweltschadens geht. Sie verweist darauf, dass sich die Ökonomen nicht einig seien, was eigentlich genau unter diesen Nichtgebrauchswerten zu verstehen ist (vgl. EPA (2005) [93]). Es sind wohl nicht so sehr die Ökonomen, die sich nicht einig sind bei der Bewertung von Nichtgebrauchswerten, sondern vielmehr die Interessenkonflikte zwischen den "verschmutzenden Industrien" und den Geschädigten, die dieser Einschätzung der EPA zugrunde liegen.

Sozialprodukt und Umweltbewertung

Abschließend soll ein kurzer Blick auf das wohl umfangreichste potenzielle Einsatzgebiet monetärer Umweltbewertungsverfahren, die Ergänzung der amtlichen Sozialproduktstatistik zu einem sogenannten Ökosozialprodukt, geworfen werden. Noch Mitte der neunziger Jahre wurde vom Club of Rome eine Bilanzierung der Gewinne aus der Produktion und der damit einhergehenden Verluste an Umweltqualität gefordert (vgl. Van Dieren (1995) [89]). Denn das ursprünglich zur Messung der ökonomischen Leistungsfähigkeit eines Landes konzipierte Sozialprodukt, das seinen Ursprung 1690 mit Thomas Petty und Gregory King in der Vorbereitung der Kriegswirtschaft in England (vgl. Hull (1899/ 1963) [135]) hat und in den 30er Jahren des letzten Jahrhundert angesichts der Weltwirtschaftskrise (vgl. Kendrick (1970) [153]) in seiner heutigen Form konzipiert wurde, ist im Laufe seiner Weiterentwicklung in vielen Ländern (vgl. Kuznets (1941) [161] und Leontief (1941) [171] für die USA, Meade und Stone (1944) [235] für Großbritannien und Tinbergen (1936) [241] und Derksen (1941) [83] für die Niederlande) nach dem zweiten Weltkrieg zunehmend als Indikator für das Wohlergehen eines Landes angesehen worden. In dieser Funktion erweist sich die Fokussierung des Sozialprodukts auf den Wert aller marktgehandelten Güter und Dienstleistungen als zu eng, denn das Wohlergehen der Bevölkerung hängt darüber hinaus von anderen Faktoren wie

dem Sicherheitsgefühl, dem Gesundheits-, Bildungs- und nicht zuletzt dem Umweltzustand ab. Cobb et al. (1995) stellen daher in ihrem vielbeachteten Artikel die Frage: "If the GDP is up, why is America down?".

Juster schlägt 1973 [145] vor, neben dem Sachkapital auch Veränderungen des Humankapitals und des Umweltkapitals in die Volkswirtschaftliche Gesamtrechnung einzubeziehen. Die Notwendigkeit einer Berücksichtigung des Umweltkapitals liegt für Juster in der Art, wie Umweltschutzinvestitionen in die Volkswirtschaftliche Gesamtrechnung eingehen und zu einem steigenden Sozialprodukt beitragen. Er stellt die Frage, ob es den Menschen durch eine Umweltschutzinvestition tatsächlich, wie vom steigenden Sozialprodukts angezeigt, besser geht im Vergleich zu einer Situation, in der diese Umweltschutzinvestition gar nicht notwendig wäre. Juster stellt fest, dass die große Lücke bei den derzeit bestehenden Systemen der Volkswirtschaftlichen Gesamtrechnung darin besteht, dass zwar die produktiven Nutzen des Umweltkapitals durchaus erfasst sind, die Schäden und der entgangene konsumtive Nutzen jedoch vernachlässigt werden. Im Licht der weiter oben geführten Diskussion um Gebrauchs- und Nichtgebrauchswerte müsste diese Einschätzung sogar noch um die entgangenen Existenz-, Options- und Vererbungswerte ergänzt werden.

Auf welche Größenordnung des Werts man dabei stoßen würde, zeigt vielleicht eine Studie von Constanza et al. (1987) [68], die den geldmäßigen Wert aller Naturdienstleistungen weltweit ermitteln. Sie finden ein durchschnittlichen Gesamtwert von 33 Billionen US\$ pro Jahr¹. Der anfängliche Enthusiasmus in den neunziger Jahren wich allerdings schnell der Erkenntnis, mit welchem immensen Erhebungsaufwand dies verbunden wäre. Aus dieser Erkenntnis heraus wurden in den letzten 10 Jahren alternative Konzepte erdacht², die den Zustand der natürlichen Umwelt und seine Veränderungen anzeigen, aber nicht versuchen, ihn monetär zu bewerten.

In dieser Einführung wurden verschiedene Gründe genannt, warum es sich lohnt, der Umwelt einen Wert zu geben. Es soll dabei jedoch nicht verschwiegen werden, dass es auch Stimmen gibt, die eine monetäre Bewertung der natürlichen Umwelt für moralisch verwerflich halten. Wenngleich die Natur zwar nicht im Mittelpunkt dieses Weltbilds steht, so scheint seinen Vertretern doch unmöglich, etwa die Schönheit eines Waldes in der "kal-

¹Zum Vergleich betrug das weltweite Sozialprodukt im Berechnungsjahr 18 Billionen US\$.

²Für eine Übersicht vgl. Lintott (1995) [175], Max-Neef (1995) [186], Radermacher (1998) [206] oder Schäfer (2000) [222]; zum Sustainable National Income Hueting (1974) [132] oder (1991a) [134], (1991b) [133]. Einen auf individuellen Wohlfahrtsänderungen basierenden Indikatorenansatz vertreten Ahlheim und Frör (2005) [6].

ten Maßeinheit Geld" auszudrücken (vgl. van Dieren (1995) [89]). Jedoch schärft die Verwendung dieser kalten Maßeinheit das Verständnis für den durch einen Umweltschaden ausgelösten Verlust oder den Gewinn durch eine Umweltverbesserung. So schreibt auch Schulz (1991) der Monetarisierung von Umweltnutzen verschiedene Funktionen zu. Neben dem Beitrag zum Verstehen des Ausmaßes einer Schädigung, kann die Diskussion versachlicht werden und die "Dosierung" umweltpolitischen Eingreifens lässt sich rational begründen (vgl. Schulz und Schulz (1991) [224]).

Gehen wir also davon aus, dass Umweltbewertung sinnvoll ist und einen Beitrag zu den Kosten-Nutzen-Rechnungen für staatliche Investitionsprojekte in eine verbesserte Umweltqualität leisten kann. Um den Nutzen eines geplanten Projekts erfassen zu können, sollte man sich dann nicht auf Teilaspekte dieses Nutzens zurückziehen, wie einige der weiter oben vorgestellten Bewertungsmethoden, sondern zu denjenigen Methoden greifen, die in der Lage sind, den gesamten Wert eines Projekts zu erfassen. Gilt es, konkrete Projekte, wie eine Rekultivierungsmaßnahme, den Investitionen in anderen Bereichen, wie etwa dem Straßenbau, vergleichbar zu machen, so bietet sich hier, wie in dieser Einführung beschrieben, die Contingent Valuation Methode an. Der Nachteil in der Praxis liegt allerdings darin, dass Contingent Valuation Studien mit einem erheblichen Zeit- und Geldaufwand verbunden sind, wenn sie zu belastbaren Aussagen führen sollen. Zunächst wendet sich der Abschnitt 1.2 aber den theoretischen Grundlagen der Contingent Valuation Methode zu, um dann in Abschnitt 1.3 auf die praktische Durchführung einer Contingent Valuation Studie und die Weiterentwicklungen der Methode in der Literatur einzugehen.

1.2 Wohlfahrtstheoretische Grundlagen der Contingent Valuation Methode

In den vorangegangenen Ausführungen wurde dafür plädiert, zur Bewertung einer Veränderung der Umweltqualität diejenigen Verfahren heranzuziehen, die in der Lage sind, alle Aspekte des zu ermittelnden Werts vollständig zu erfassen.

Der Wert einer Umweltverbesserung oder -verschlechterung lässt sich, zumindest im hier unterstellten anthropozentrischen Weltbild, mit der Wohlfahrtsänderung infolge der Umweltveränderung bemessen. Das Ziel bei den

Anwendungen von Umweltbewertungsverfahren ist somit, herauszubekommen, ob es der Gesellschaft durch die Umweltveränderung insgesamt besser geht oder nicht. Da die Änderung der gesellschaftlichen Wohlfahrt jedoch untrennbar mit der individuellen Wohlfahrt verbunden ist, besteht der erste Schritt darin, die individuelle Wohlfahrtsänderung zu bestimmen. Auf den zweiten Schritt - die Aggregation dieser individuellen Wohlfahrtsänderungen zu einer gesamtgesellschaftlichen Änderung - wird weiter unten detaillierter eingegangen.

Betrachten wir zunächst die individuelle Wohlfahrtsänderung. Die Wohlfahrt des i -ten Gesellschaftsmitglieds lässt sich durch den in einer Situation realisierten Nutzen U_i ausdrücken. Die Änderung der individuellen Wohlfahrt besteht daher in der Änderung des individuellen Nutzen zwischen einer Ausgangssituation und einer oder mehreren Endsituationen. Der Nutzen selbst oder der Unterschied zwischen in verschiedenen Situationen $k = 0, 1, \dots, K$ realisierten Nutzen U_i^k ist jedoch nicht messbar. Daher greift man zurück auf sogenannte Wohlfahrtsmaße, die in der Lage sind, uneindeutig eine Nutzenänderung anzuzeigen (Indikatorbedingung), d. h. es muss für ein zuverlässiges³ Wohlfahrtsmaß W_i^{01} für die Änderung der individuellen Wohlfahrt von Zustand 0 nach Zustand 1 gelten

$$W_i^{01} = W_i(U_i^0, U_i^1) \geq 0 \quad \Leftrightarrow \quad U_i^1 \geq U_i^0 \quad (1.1)$$

Will man die Nutzenänderung durch ein umweltveränderndes Projekt messen, so muss zunächst die Umweltqualität als Variable in die Nutzenfunktion einbezogen werden. Die Auswirkung einer Änderung der Umweltqualität auf das Individuum lässt sich dann wie in Gleichung (1.2) ausdrücken.

$$U_i^{01} = U_i^1 - U_i^0 = u_i(\mathbf{x}_i^1, \mathbf{z}^1) - u_i(\mathbf{x}_i^0, \mathbf{z}^0) \quad (1.2)$$

u_i ist dabei die direkte Nutzenfunktion des Individuums i , \mathbf{x}_i ist das von i konsumierte Güterbündel und mit 0 und 1 wird der Ausgangszustand sowie die Situation nach der Änderung gekennzeichnet. Der Nutzen des Individuums steigt monoton mit steigender Umweltqualität \mathbf{z} . Ein umweltveränderndes Projekt kann darüber hinaus auch Einfluss auf Marktgüterpreise und die Einkommen der betroffenen Individuen haben. So können in Folge eines neuen Naturparks neue Möglichkeiten zur Erwerbstätigkeit geschaffen werden, der Tourismus sich beleben oder die Nachfrage nach bestimmten Marktgütern steigen. Eine Umweltkatastrophe hingegen würde

³Eine weitere Bedingung an ein zuverlässiges Wohlfahrtsmaß ist die Messbarkeits- bzw. Operationalisierbarkeitsbedingung, die fordert, dass das Maß mit erheblichen Größen berechenbar sein muss.

die Erwerbsmöglichkeiten einer Region erheblich schwächen, sich nachteilig auf die Einkommen auswirken und auch die Marktgüternachfrage beeinflussen. Daher ist es nützlich, auf eine Formulierung mittels indirekter Nutzenfunktionen überzugehen. Aus Gleichung (1.2) wird dann

$$U_i^{01} = v_i(\mathbf{p}^1, y_i^1, \mathbf{z}^1) - v_i(\mathbf{p}^0, y_i^0, \mathbf{z}^0), \quad (1.3)$$

wobei mit y_i das Einkommen der Individuen bezeichnet wird und \mathbf{p} der Preisvektor für Marktgüter ist.

Um zu einem Wohlfahrtsmaß zu gelangen, kann man einer Idee von Hurwicz und Uzawa (1971) [136] folgen und dem nicht beobachtbaren Nutzen des Individuums, den es aus dem Konsum von Gütern und der Umweltqualität zieht, einen beobachtbaren Wert in Geldeinheiten zuordnen. Diese Größe übernimmt somit die Rolle einer "money-metric utility" (vgl. Samuelson (1974) [219]). Mithilfe der Ausgabenfunktion $e(\mathbf{p}^k, \mathbf{z}^k, U_i)$, die für das mindestens notwendige Einkommen steht, mit dem ein Individuum bei den Preisen \mathbf{p}^k und der Umweltqualität \mathbf{z}^k einen Nutzen U_i realisieren kann, lässt sich das Einkommensäquivalent des Nutzens bilden. Will man nun die Wohlfahrtsänderung messen, die mit einer Änderung der Umweltqualität einhergeht, betrachtet man zunächst die Nutzenänderung durch die Umweltqualitätsveränderung und überlegt sich dann, welche Änderung des Einkommens zu derselben Nutzenänderung bei Konstanz der Preise und der Umweltqualität führen würde. Die Preise für Konsumgüter und die Umweltqualität werden hierbei auf einem zu wählenden Niveau festgehalten. Hierfür bieten sich die Ausgangssituation oder die Endsituation als fixe Punkte an. Im Prinzip könnte jedoch auch jeder beliebige andere Punkt als Referenzpunkt gewählt werden (vgl. Ahlheim und Rose (1989, S. 59 ff.) [10]). Auf diese Art lassen sich die auf John Hicks zurückgehenden "Variationsmaße", die Äquivalente Variation und die Kompensierende Variation (vgl. Hicks (1939) und (1943) [124], [125]) für Veränderungen der Umweltqualität als Differenz zwischen zwei Werten der Ausgabenfunktion formulieren.

Zunächst betrachten wir das Maß, das entsteht, wenn man die *Anfangssituation* als Referenzpunkt wählt. Dann wird dieses Maß zur Differenz zwischen der Ausgabenfunktion bei den "alten" Preisen, der "alten" Umweltqualität und dem ursprünglichen Nutzenniveau und der Ausgabenfunktion bei "alten" Preisen und Umweltqualität und dem neuen Nutzenniveau.

$$EV_i^{01} = e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^1) - e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^0) = e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^1) - y_i^0, \quad (1.4)$$

wobei sich die letzte Umformung die Identität $e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^0) \equiv y_i^0$ zwischen der Ausgabenfunktion bei den Preisen, der Umweltqualität und dem Nutzenniveau in der Ausgangssituation und dem Einkommen in der Ausgangssituation zunutze macht. Diese sogenannte Äquivalente Variation misst für eine Umweltverbesserung, d.h. für ein nutzensteigerndes Vorhaben denjenigen Geldbetrag, den man einem Haushalt anstelle eines geplanten nutzenerhöhenden Vorhabens mindestens geben müsste, damit er auch ohne das Vorhaben das neue Nutzenniveau erreichen kann und somit den Verzicht auf das nutzenerhöhende Vorhaben akzeptiert ("willingness to accept"). Um die Äquivalente Variation für ein nutzenerhöhendes Projekt zu messen, muss man sich vorstellen, welches Nutzenniveau durch das Projekt erreichbar wäre, um dann den Geldbetrag zu ermitteln, der das Erreichen eben dieses Nutzenniveaus bei derzeitigen Preisen und der derzeitigen Umweltsituation ermöglicht. Für ein nutzenverminderndes Vorhaben misst die Äquivalente Variation denjenigen Geldbetrag, den das Individuum höchstens zu zahlen bereit wäre, damit das Vorhaben nicht durchgeführt wird ("willingness to pay").

Interpretatorisch einfacher scheint das Maß zu sein, das bei der Festlegung von Preisen und der Umweltqualität auf die *Endsituation* entsteht. Nimmt man die Endsituation nach der Durchführung eines Projekts bzw. nach einem geschehenen Umweltunfall als Referenzpunkt, so erhält man die sogenannte Kompensierende Variation

$$CV_i^{01} = e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^1) - e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^0) = y_i^1 - e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^0). \quad (1.5)$$

Für ein nutzenerhöhendes Projekt misst die (positive) Kompensierende Variation den Geldbetrag, den das Individuum bezahlen würde, um das nutzenerhöhende Projekt in der Zukunft zu bekommen. Für ein nutzenverminderndes Projekt wird der Geldbetrag errechnet, den das Individuum erhalten müsste (i.e. $CV \leq 0$), um nach dem Projekt nicht schlechter gestellt zu sein als zuvor ("willingness to accept"). Dies führt zur Interpretation der CV als minimale Entschädigungsforderung für eine Umweltverschlechterung.

Die Kompensierende Variation nach Gleichung (1.5) wird im weiteren Verlauf dieser Arbeit verwendet. Es scheint der Alltagserfahrung weit eher zu entsprechen, dass man für ein Gut, das man haben möchte, bezahlen muss und auch bereit ist, etwas zu bezahlen. Ebenso selbstverständlich scheint die Entschädigungsforderung für einen zugefügten Schaden zu sein. Der Mechanismus der Äquivalenten Variation wirkt zumindest für nutzenerhöhende, umweltverbessernde Projekte umständlich, da er danach fragt,

welchen Geldbetrag die Individuen erhalten müssten, wenn sie ein nutzerhöhendes Projekt nicht bekämen. Beide Maße werden in den meisten Fällen für dasselbe Projekt zu einem unterschiedlichen Wert führen, denn sie messen zwei unterschiedlich normierte money-metric utility-Funktionen. Sie sind nur in einzelnen Spezialfällen wie etwa linearen Indifferenzkurven oder quasi-linearen Nutzenfunktionen identisch (vgl. Ahlheim und Buchholz (2000) [4]).

Umweltveränderungen können, wie bereits erwähnt, mit Veränderungen des Einkommens und der Marktgüterpreise einhergehen und es ist nützlich, die Kompensierende Variation aus Gleichung (1.5) so zu zerlegen, dass sie diese Veränderungen widerspiegelt (vgl. Stephan und Ahlheim (1996) [234]).

$$\begin{aligned}
 CV_i^{01} &= e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^1) - e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^0) \\
 &= e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^1) - e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^0) \\
 &\quad + e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^0) - e_i(\mathbf{p}^1, \mathbf{z}^0, U_i^0) \\
 &\quad + e_i(\mathbf{p}^1, \mathbf{z}^0, U_i^0) - e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^0) \\
 &= y_i^1 - y_i^0 \\
 &\quad + e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^0) - e_i(\mathbf{p}^1, \mathbf{z}^0, U_i^0) \\
 &\quad + e_i(\mathbf{p}^1, \mathbf{z}^0, U_i^0) - e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^0)
 \end{aligned} \tag{1.6}$$

Die verschiedenen Terme in Gleichung (1.6) lassen sich zu einer anschaulichen Darstellung zusammenfassen.

$$CV_i^{01} = CVY_i^{01} + CVP_i^{01} + CVZ_i^{01} \tag{1.7}$$

Dies bedeutet, dass sich die Kompensierende Variation für die durch ein Umweltprojekt hervorgerufene Nutzenänderung als die Summe der kompensierenden Variation für die durch das Projekt hervorgerufenen Einkommensänderung $CVY_i^{01} = y_i^1 - y_i^0$, der kompensierenden Variation für die durch das Projekt induzierten Preisänderungen $CVP_i^{01} = e_i(\mathbf{p}^0, \mathbf{z}^0, U_i^0) - e_i(\mathbf{p}^1, \mathbf{z}^0, U_i^0)$, und der kompensierenden Variation für die durch das Projekt hervorgerufene Änderung der Umweltqualität $CVZ_i^{01} = e_i(\mathbf{p}^1, \mathbf{z}^0, U_i^0) - e_i(\mathbf{p}^1, \mathbf{z}^1, U_i^0)$ auffassen lässt. Dies ist eine besonders unter praktischen Erwägungen überaus nützliche Eigenschaft, da für die Berechnung der ersten beiden Terme in Gleichung (1.7) exakte Berechnungsverfahren existieren. Der mit CVY bezeichnete Term entspricht der beobachtbaren Einkommensveränderung selbst, für den mit CVP bezeichneten Term existieren exakte Verfahren, wie beispielsweise das Vartia (vgl. Vartia (1983) [246])

Verfahren. Diese Verfahren stützen sich auf die umfassende Erhebung von Nachfragedaten, wie Beobachtungen zu Preisen und nachgefragten Mengen, so dass die Berechnung lediglich eine Frage des Aufwands ist, den man zum Erstellen der jeweiligen Statistik betreiben will. Die Berechnung des dritten Terms ist hingegen ungleich schwieriger, da die "Nachfrage" nach Umweltgütern wegen der eingangs beschriebenen fehlenden Märkte im Allgemeinen nicht ohne Weiteres messbar ist. Hier müssen andere Verfahren, wie die in der Einführung beschriebenen Methoden, angewendet werden.

Wählt man aufgrund der weiter oben erläuterten Fähigkeit zur umfassenden Umweltbewertung eine direkte Bewertungsmethode, so wird die Zahlungsbereitschaft oder die Entschädigungsforderung von betroffenen Personen in einer direkten Befragung ermittelt. In einer Contingent Valuation Studie zu einer Umweltqualitätsverbesserung wird beispielsweise ermittelt, welchen Beitrag ein Haushalt höchstens zu zahlen bereit wäre, um nach Durchführung der Verbesserung mindestens genauso gut gestellt zu sein wie zuvor. Eine derartige Befragung ist in ihrer praktischen Umsetzung mit verschiedenen Schwierigkeiten behaftet, auf die in Kapitel 1.3 im Detail eingegangen wird. Wenn es trotz der Umsetzungsschwierigkeiten gelingt, den letzten Summanden in Gleichung (1.7) zu berechnen und somit den letzten Schritt zur Berechnung der individuellen Wohlfahrtsänderung zu vollenden, ist damit das eingangs formulierte Ziel noch nicht erreicht. Für die Kosten-Nutzen-Analyse einer staatlichen Maßnahme ist es notwendig, einen Wert zu ermitteln, der das Wohlergehen aller von der Maßnahme betroffenen Bürger umfasst. Das *Identifikationsproblem* der Umweltbewertung, d.h. die Bestimmung der durch eine Umweltveränderung herbeigeführten *individuellen* Wohlfahrtsänderung, ist zwar gelöst, aber die Berechnung einer gesamtgesellschaftlichen Wohlfahrtsänderung, die Lösung des *Aggregationsproblems*, steht noch aus. Das einfache paretianische Werturteil, wonach es einer Gesellschaft besser geht, wenn ein Individuum besser gestellt ist und kein anderes schlechter gestellt ist als zuvor, ist in vielen Anwendungen der Umweltbewertung nicht zielführend, unter anderem da die weiter oben beschriebenen Projekte oftmals Gewinner *und* Verlierer haben. Um die Wohlfahrtsmessung für derart umfassende Entscheidungen operationalisierbar zu machen, muss ein Kriterium für die gesamtgesellschaftliche Wohlfahrtsänderung gefunden werden.

Die naive Vorgehensweise würde auf einem einfachen Analogieschluss basieren, statt der individuellen Wohlfahrtsänderung suchen wir nun die gesamt-

gesellschaftliche soziale Wohlfahrtsänderung $SW^{01} = SW^{01}(U_1^{01}, \dots, U_H^{01})$, die sich als Funktion der individuellen Nutzenänderungen ausdrücken lässt. Dem steht jedoch der in der ordinalen Nutzentheorie angelegte Ausschluss von interpersonellen Nutzenvergleichen entgegen. Vor die Wahl zwischen zwei Güterbündeln gestellt, lässt sich auf der individuellen Ebene nur angeben, welches der Bündel mindestens so gut oder besser ist als ein anderes. Ist ein Bündel besser als ein anderes, so weiß das Individuum, das es mit diesem glücklicher wird, kann jedoch keine Angaben darüber machen, um wie viel sein Glück sich gesteigert hat. Ebenso wenig lässt sich eine Aussage darüber treffen, ob man mit dem gewählten nutzenmaximierenden Güterbündel glücklicher ist als ein anderer mit seiner persönlichen Wahl, der Nutzenvergleich zwischen verschiedenen Individuen ist daher unmöglich. Ohne interpersonelle Nutzenvergleiche lässt sich jedoch, wie Arrow bereits 1950 [19] und (1951) [20] zeigen konnte, keine gesellschaftliche Präferenzordnung definieren, die nicht diktatorisch, d.h. mit einer der individuellen Präferenzordnungen identisch ist. Im Rahmen einer streng ordinal fundierten Wohlfahrtstheorie kann dem Politiker zur Entscheidung somit nur eine Übersicht über die Nutzenänderungen der von einem Projekt betroffenen Individuen angeboten werden. Das reicht jedoch für die Bedarfe der praktischen Politikberatung häufig nicht aus. Daher wird an dieser Stelle von der "reinen Lehre" abgewichen und einer von Hicks (1939) [124] und Kaldor (1939) [149] vorgeschlagenen Idee gefolgt. Das Hicks-Kaldor Kriterium bewertet ein Projekt als positiv, wenn nach der theoretisch möglichen Kompensation der "Verlierer" durch die "Gewinner" des Projekts die resultierende Änderung des "Nettonutzens", ausgedrückt durch die Differenz zwischen allen Zahlungsbereitschaften und Entschädigungsforderungen, größer ist als Null. Wurde die Kompensierende Variation als Maß für die individuellen Nutzenänderungen herangezogen, so lässt sich diese Bedingung folgendermaßen ausdrücken

$$\sum_{i=1}^N CV_i^{01} \geq 0 \Rightarrow SW^{01} \geq 0. \quad (1.8)$$

Zur Umsetzung dieses Ansatzes in eine praktische Kosten-Nutzen-Analyse sind zusätzliche Näherungen notwendig, da die Erhebung der notwendigen Daten für den Ausdruck in (1.8) gerade bei kleineren Projekten unverhältnismäßig zeit- und kostenintensiv wäre. Zur Anwendung beispielsweise der exakten Berechnungsverfahren nach Vartia müssten für die Bestimmung der Preiseffekte CVP die gesamten Nachfragesysteme der befragten Haushalte ökonometrisch geschätzt werden, was einen erheblichen Aufwand be-

deutet. In der Praxis stützt man sich daher auf die folgende Hilfskonstruktion (für eine ausführliche formale Herleitung dieses Arguments, das auf der ersten Ordnung einer Taylorapproximation der CV beruht (vgl. Ahlheim und Rose (1989) [10]). Man unterstellt, dass bei der Bereitstellung des Umweltguts knappe volkswirtschaftliche Ressourcen genutzt werden, die vorher der Marktgüterproduktion zur Verfügung standen. Dies hat gerade jene Veränderung des Einkommens und des Marktgüterkonsums zur Folge, die die Terme CVI und CVP in dem Ausdruck (1.7) verursacht. Unterstellt man ferner, dass die zur Bereitstellung des Umweltguts notwendigen Faktoren \mathbf{r}^1 auch vor dem Projekt effizient eingesetzt und nach ihrem Wertgrenzprodukt mit \mathbf{q}^1 bezahlt wurden, so können diese beiden ersten Terme durch die bei dem Umweltprojekt entstehenden Kosten abgeschätzt werden. Akzeptiert man diese Näherung, so reduziert sich die Abschätzung des Wohlfahrtseffekts eines umweltverbessernden Projekts in der Tat auf die Erhebung der Kosten und die Ermittlung der Zahlungsbereitschaften der betroffenen Bevölkerung als Maß für den gestifteten Nutzen. Daher lautet die vielfach in der Praxis verwendete Entscheidungsregel der Kosten-Nutzen-Analyse

$$\mathbf{q}^1 \mathbf{r}^1 \leq \sum_{i=1}^N CV Z_i^{01} \Leftrightarrow SW^{01} \geq 0; \quad (1.9)$$

d.h., ein Projekt gilt dann als wohlfahrtserhöhend, wenn die Summe über die Zahlungsbereitschaften der betroffenen Bevölkerung die Kosten des Projekts übersteigt.

Es leuchtet jedoch ein, dass man ebenfalls aus Kostengründen nicht die gesamte vom Projekt betroffene Bevölkerung befragen kann, sondern darauf angewiesen ist, die Erhebung auf eine Stichprobe zu beschränken. Damit entfernt man sich einen Schritt weiter weg von der Theorie, denn nun sind nicht mehr die individuellen Präferenzen Grundlage des ermittelten Werts, sondern es wird von den Präferenzen einiger auf die Präferenzen aller betroffenen Haushalte oder Bürger geschlossen. Dabei unterstellt man implizit, dass die Präferenzen der betroffenen Bevölkerung dieselben sind wie die Präferenzen der befragten Individuen. Bei einer repräsentativen Stichprobe stellt dies statistisch gesehen kein Problem dar, von der strengen nutzen-theoretischen Fundierung, die auf individuellen Präferenzen aufbaut, ist eine derartige Vorgehensweise jedoch schon sehr entfernt.

In diesem Abschnitt wurde das Wohlfahrtsmaß definiert, das in seiner Interpretation als Zahlungsbereitschaft oder Entschädigungsforderung mit einer Befragung nach der Contingent Valuation Methode gemessen werden kann, und beschrieben, mit welchen Näherungen man in der Praxis aus diesen gemessenen individuellen Zahlungsbereitschaften eine Aussage über die gesellschaftliche Wohlfahrtsänderung im Rahmen einer Kosten-Nutzen-Analyse gewinnen kann. Diese Arbeit konzentriert sich auf eine Verbesserung der Methoden zur Berechnung von $\sum_{i=1}^N CVZ_i^{01}$, der rechten Seite der ersten Ungleichung in (1.9) und hierbei insbesondere auf die Zahlungsbereitschaft für ein umweltverbesserndes Projekt. Der Umsetzung der Frage nach der Zahlungsbereitschaft in einer Contingent Valuation Studie wird im folgenden Abschnitt nachgegangen, wobei der Aufbau eines CVM-Interviews ausführlich erläutert wird und auf diejenigen Elemente des Interviews eingegangen wird, die den Erfolg einer CVM-Studie zur Bewertung eines Umweltprojekts kritisch beeinflussen.

1.3 Contingent Valuation - die Praxis

Die Contingent Valuation Methode basiert auf einem einfachen Grundgedanken. Da Umweltgüter nicht auf Märkten gehandelt werden, konstruiert man im Rahmen eines Interviews hypothetische Märkte, auf denen die Befragten das Umweltgut "kaufen" können. In der Praxis wird dies dadurch umgesetzt, dass den Befragten ein umweltverbesserndes Projekt vorgestellt wird und sie nach ihrer Zahlungsbereitschaft für die Durchführung dieses Projekts gefragt werden. Als Urheber der CVM-Idee gilt Ciriacy-Wantrup (1947) [66], der in einem Artikel zur Messung des Nutzens eines Programms zur Vermeidung von Erosion vorschlug, Befragungen einzusetzen, um zu ermitteln, wie groß die Zahlungsbereitschaft der Bevölkerung für eine zusätzliche Einheit dieses öffentlichen Guts ist. Als erster praktischer Versuch, die Zahlungsbereitschaft einzelner Haushalte durch Befragung zu ermitteln, wird die Dissertation von Robert Davis (1963) [82] angesehen, der den Wert eines Erholungsgebiets für Jäger und Naturschützer maß. Seitdem sind die unterschiedlichsten Umweltgüter, sowohl staatliche Investitionsvorhaben, durch die ein neues Umweltgut geschaffen wird, als auch Projekte zur Vermeidung von Umweltschäden oder Gesundheitsschäden, mit dieser Methode bewertet worden.

Im Rahmen eines CVM-Interviews sind die Befragten mit einer Reihe von ungewohnten Dingen konfrontiert. Zunächst müssen sie verstehen, was das

zu bewertende Umweltgut ist, danach müssen sie einen hypothetischen Markt für dieses Gut akzeptieren und schließlich ihre Zahlungsbereitschaften für das Umweltgut angeben, was für die Befragten oftmals eine nicht alltägliche Aufgabe ist. Um sie allmählich an diese Aufgabe heranzuführen, gliedert sich der Aufbau eines Fragebogens für eine Contingent Valuation Studie im Wesentlichen in fünf Teile (zum grundlegenden Aufbau eines CVM-Interviews vgl. Mitchell und Carson (1989) [193], die relevante Literatur zu den einzelnen Teilen findet sich weiter unten in den entsprechenden Abschnitten). Nach einer kurzen Einleitung und einem sogenannten Warm-up Teil, in dem den Befragten die Scheu vor der häufig als Test empfundenen Interviewsituation dadurch genommen wird, dass ihnen zunächst leicht beantwortbare Fragen nach ihren Lebensgewohnheiten oder ihren Hobbys gestellt werden, folgt

- die Definition des zu bewertenden Umweltguts,
- die Definition des hypothetischen Marktes,
- die Frage nach der Zahlungsbereitschaft und
- die Erhebung zusätzlicher sozioökonomischer Daten.

Ein solches CVM-Interview kann grundsätzlich als persönliches Interview, als Telefonbefragung, als Briefbefragung oder als Internetbefragung durchgeführt werden. Dabei erlaubt das persönliche Interview die stärkste Kontrolle über den Verlauf, ist jedoch auch die teuerste Interviewform. Die Vorteile sind zum einen, dass das Interview in einer den Befragten vertrauten Umgebung, ihrer eigenen Wohnung, durchgeführt werden kann, und zum anderen, dass die verbale Darstellung des zu bewertenden Guts und des hypothetischen Marktes mit einer Vielzahl von optischen Hilfsmitteln, wie Karten, Fotos, Skizzen und Videos ergänzt werden kann. Der Interviewer kann die Rezeption dieser Darstellung beobachten und für spätere Plausibilitätsüberprüfungen der gegebenen Antworten Besonderheiten und Auffälligkeiten schriftlich festhalten (vgl. Mitchell und Carson (1989) [193] oder Carson (1997) [58]). Darüber hinaus können die interviewten Personen Rückfragen stellen, wodurch das Verständnis des Umweltguts und des hypothetischen Marktes besser gewährleistet werden kann (zu den Einzelheiten vgl. Abschnitt 1.3.1 und 1.3.2).

In einer Telefonbefragung ist zwar auch die Kontrolle über den Ablauf des Interviews gegeben und es besteht, ebenso wie im persönlichen Interview,

die Möglichkeit für Rückfragen, aber die Verwendung optischer Hilfsmittel entfällt völlig. Da die Befragten häufig keine Erfahrung mit dem zur Bewertung gestellten Umweltgut haben, können allerdings Schwierigkeiten auftreten, wenn sie sich die konkreten Einzelheiten des geplanten Projekts nur aufgrund einer verbalen Darstellung vorstellen sollen. So kann es zu Missverständnissen über die Art und besonders den Umfang des Projekts kommen, die durch Kartenmaterial und Illustrationen aufgeklärt werden könnten. Darüber hinaus sind Telefoninterviews in der Regel kürzer und somit weniger ausführlich, da es schwieriger ist, die Aufmerksamkeit des Gegenübers am Telefon über eine längere Zeitspanne aufrecht zu erhalten (vgl. Carson (1999) [59]).

Bei der preiswertesten Interviewform, der Briefbefragung, entfällt auch die Kontrolle über den Ablauf des Interviews. Es besteht keinerlei Einfluss darauf, in welcher Reihenfolge und über welchen Zeitraum hinweg die Befragten sich dem Material widmen; es bleibt sogar offen, wer letztlich den Fragebogen beantwortet hat. Rückfragen sind nahezu ausgeschlossen und das Verständnis des Umweltguts ist nur durch Kontrollfragen, deren wahrheitsgemäße Beantwortung im Dunklen bleibt, zu überprüfen. Ähnliche Kritikpunkte gelten auch für die Internetbefragung, allerdings lässt sich hier ein breites Spektrum zusätzlicher Informationen in den Fragebogen integrieren, indem weiteres Material durch Hyperlinks mit dem Fragebogen verbunden wird (vgl. Behrens et al. (2004) [31]). Die Erfahrungen mit Internetbefragungen sind derzeit noch recht gering, so dass ein endgültiges Votum hier nicht abgegeben werden kann. Zwischen persönlichem Interview, Telefon- und Briefbefragung ist die Reihenfolge jedoch eindeutig: wenn die Zeit und die Mittel für eine persönliche Befragung zur Verfügung stehen, so ist dieser Methode unzweifelhaft der Vorzug zu geben. Daher beziehen sich die nachfolgenden Ausführungen zu den einzelnen Bestandteilen eines CVM-Fragebogens auch auf den Einsatz in persönlichen Interviews.

1.3.1 Definition des Umweltguts

Ausgangspunkt einer jeden CVM-Untersuchung ist die Definition des zu bewertenden Umweltguts. Anders als bei den in der Einführung beschriebenen indirekten Methoden, bei denen sich abgeleitete Nachfragen nach bestehenden Umweltgütern beobachten lassen, muss beim Einsatz der Contingent Valuation Methode das zu bewertende Umweltgut für den Befragten gewissermaßen erst erschaffen werden. Die Befragten müssen den Zu-

stand der Umwelt *ohne* das vorgeschlagene Projekt und die Veränderung durch das Projekt kennen, damit sie genau verstehen, was sie für die von ihnen geäußerte Zahlungsbereitschaft erhalten werden. Zu einer derartigen Beschreibung - dem Szenario - gehört zum einen eine verbale Darstellung, zum anderen werden jedoch, wie weiter oben geschildert, bei persönlich durchgeführten Interviews auch visuelle Hilfsmittel wie Landkarten, Grafiken und Fotos und teilweise sogar Videos eingesetzt. Die Entwicklung eines glaubwürdigen Szenarios stellt bei der Vorbereitung einer CVM-Studie eine der wesentlichen Aufgaben dar, denn von dem Verständnis des zu bewertenden Umweltguts hängt die Zuverlässigkeit der Methode kritisch ab (zur Bedeutung des Szenarios und für Richtlinien zur Szenarioerstellung vgl. NOAA (1993) [18]). In der Praxis wird die Entwicklung eines solchen Szenarios von Experten unterstützt, die technisches Fachwissen aus der Ökologie, der Biologie oder anderen naturwissenschaftlichen Disziplinen zu den jeweiligen Umweltgegebenheiten einspeisen können.

Die verständliche und eindeutige Formulierung des Szenarios ist notwendig, um den Befragten eine genaue Vorstellung von Art und Umfang des zu bewertenden Umweltguts zu geben. Dies kann bei Umweltgütern zu einer vielschichtigen Anforderung werden, denn die Zusammenhänge zwischen einem Umweltgut und dem im Interview zur Bewertung vorgeschlagenen Projekt können recht komplex sein. Ein Programm zum Schutz der Regenwälder besteht selten aus einem einfachen Zaun um den Regenwald, sondern aus vielfältigen Maßnahmen, die beispielsweise Anreize zur Wiederaufforstung und zur Eindämmung des Holzeinschlags durch Subventionen umfassen. Die Ausdehnung des durch solche Maßnahmen geschützten Regenwalds sollte zudem durch Vergleiche mit anderen Flächen illustriert werden, denn es ist für die meisten Menschen ungleich schwieriger, sich eine Fläche in Quadratkilometern vorzustellen als etwa die Fläche einer ihnen bekannten großen Stadt. Darüber hinaus ist es für die Befragten hilfreich, die Fläche des geplanten Projekts in einen Bezug zu setzen mit der gesamten von Regenwald bedeckten Fläche, denn nur so ist gewährleistet, dass sie nicht intuitiv ein viel größeres Projekt bewerten - wie etwa den gesamten brasilianischen Regenwald.

Die Bewertung eines "größeren" Projekts findet man nicht nur als flächenmäßige Fehleinschätzung, die noch leicht auszumerzen wäre. Weitaus schlimmer im Sinne der klaren Abgrenzung des zu bewertenden Umweltguts wird die Fehleinschätzung, wenn anstelle eines konkreten Projekts unbewusst ein übergeordnetes Ideal bewertet wird. Dies kann geschehen, da die Befrag-

ten die Informationen, die sie über das geplante Umweltprojekt bekommen, häufig unbewusst mit Informationen in Zusammenhang bringen, die sie aus anderen Quellen gewonnen haben. Wenn in den Medien kurz vor der Befragung zahlreiche Dokumentationen zur Funktion der Regenwälder als CO₂-Senke oder zur Artenvielfalt in Regenwäldern erscheinen, wird ein Szenario zum Erhalt des Regenwalds in diesen persönlichen Informationskontext des Befragten eingebettet werden. Dies birgt die Gefahr, dass ein wesentlicher Teil der geäußerten Zahlungsbereitschaften der Wertschätzung der "Natur an sich" oder der "Rettung des globalen Klimas und der Artenvielfalt" und nicht eines konkreten Regenwaldprojekts entspringt.

Beide Arten der Fehleinschätzung werden in der Literatur als "Part-Whole-Bias" oder als Embedding-Effekt diskutiert, wobei die Abgrenzung dieser Begriffe nicht trennscharf ist. Einige Autoren (vgl. z.B. Kahneman und Knetsch (1992) [146]) wenden den Begriff des Part-Whole-Bias auf diejenigen Situationen an, in denen es um offensichtliche Fehleinschätzungen der räumlichen oder mengenmäßigen Abgrenzung geht, wie zum Beispiel in einer von Kritikern der Methode vielzitierten Studie von Boyle et al. (1994) [42], bei der die Zahlungsbereitschaften für die Rettung von entweder 2000, 20.000 oder 200.000 Exemplaren von durch ölverschmutzte Gewässer geschädigten Wasservögeln gemessen wurden. Es ergaben sich keine wesentlichen Unterschiede für diese doch auf den ersten Blick recht unterschiedlichen Mengen. Die Vermutung ist, dass eine Zahlungsbereitschaft zur Rettung aller betroffenen Vögel geäußert wurde. Für Bewertungen, die die Wertschätzung eines bestimmten Ideals anstelle eines konkreten Projekts enthalten, verwendet Knetsch (1994) [155] den Begriff des Embedding. Wieder andere Autoren verwenden den Begriff des Embedding synonym zu dem ansonsten als Warm-Glow Effekt bekannten Phänomen (vgl. Nunes et al. (2001) [198]), bei dem die Befragten einen Wohlfühleffekt (Warm Glow vgl. Andreoni (1990) [15]) erleben, weil sie etwas für einen (beliebigen) guten Zweck gespendet haben, ganz gleich, ob der Gegenstand dieser guten Tat ein Kind, der Regenwald oder ein Küstenstreifen ist.

Mitchell und Carson (1989) [193] beschreiben als Part-Whole-Bias "major amenity misspecifications, and [are] also a result of the tendency of respondents to respond to public goods as global symbols without paying sufficient attention to the specific description offered in a CV survey. [...] The dimensions of a good that are particularly prone to this misperception are its geographic distribution, its benefit composition, and the package of policies of which it is a part.", und schließen in diese Definition die andern-

orts als Embedding bezeichneten Fehleinschätzungen mit ein. Loomis et al. (1993) [176] definieren den Embedding-Effekt genau wie Whitehead et al. (1998) [251]) aus der Beobachtung, dass dasselbe Gut einen geringeren Wert zugemessen bekommt, wenn es als Teil eines größeren Guts bewertet wird, als wenn es alleine steht. Wenn also die Befragten dezidiert auf den Umfang des Umweltguts hingewiesen werden, reduzieren sie die Angaben zu ihrer Zahlungsbereitschaft. Umgekehrt bedeutet dies, dass die geäußerte Zahlungsbereitschaft ohne den expliziten Hinweis zu hoch sein wird.

Die etwas unklare begriffliche Abgrenzung rührt vor allem daher, dass die beschriebenen Effekte eines gemeinsam haben: man findet in Studien zu unterschiedlich großen Projekten ähnliche Zahlungsbereitschaften, obwohl die Zahlungsbereitschaften die Größe des Projekts - den Scope - reflektieren müssten, wenn sie die Nutzenänderung infolge des Projekts widerspiegeln. Der Nutzen nimmt mit den Mengen eines "konsumierten Gutes" monoton zu, somit auch mit den Mengen des zur Verfügung gestellten Umweltguts, so lange keine Sättigung unterstellt wird. Daher muss die Zahlungsbereitschaft signifikant messbar ansteigen, wenn das auf dem hypothetischen Markt angebotene Umweltgut deutlich vergrößert wird, d.h. wenn beispielsweise eine ganze Seenplatte anstatt eines einzelnen Sees zur Bewertung gestellt wird. In zahlreichen Fällen ließ sich jedoch beobachten, dass Contingent Valuation Studien den Scope-Test nicht bestehen. Insbesondere in den frühen neunziger Jahren, als die Contingent Valuation Methode mit großem Eifer öffentlich diskutiert wurde, da sie zur Wertfeststellung bei Umwelthaftungsprozessen zu hohen Entschädigungsfordernungen zugelassen wurde (vgl. Lehr (2002) [166]), war das Nicht-Bestehen des Scope-Tests einer der Hauptvorwürfe von Kritikern der Methode (vgl. Diamond und Hausman (1994) [87], Diamond und Hausman (1993) [86]).

Die Auseinandersetzung in der Literatur mit dem Embedding-Effekt und dem Part-Whole-Bias (vgl. Carson und Mitchell (1995) [62] und Carson et al. (2001) [60] oder zur Entwicklung neuer Testverfahren zum Beispiel Smith und Osborne (1996) [230]) hat sich bis in die heutige Zeit gehalten. So findet Svedsater (2000) [239] die Embedding Hypothese bestätigt, Giraud et al. (1999) [106] hingegen untersuchen eine Reihe von Studien und stellen fest, dass ihre Tests die Hypothese überwiegend zurückweisen. Bennett et al. (1998) [29] wiederum kommen zu unterschiedlichen Ergebnissen beim Vergleich verschiedener australischer Studien.

Einigkeit besteht jedoch über die zentrale Bedeutung, die die sorgfältige und unmissverständliche Beschreibung des Szenarios für das Gelingen

eines CVM-Interviews hat. Die derzeit gängige Praxis besteht darin, die Befragten explizit darauf hinzuweisen, dass nur ein Teil eines möglicherweise größeren Ganzen zu bewerten ist (vgl. z.B. Schulze (1998) [225] für ein interessantes Fragebogendesign). Zu einer fundierten Szenariodarstellung gehören diese Hinweise auf den genauen Umfang des zu bewertenden Umweltguts genauso wie eine naturwissenschaftlich fundierte, jedoch allgemein verständliche Beschreibung des geplanten Projekts.

Das Verständnis des Szenarios lässt sich, wie letztlich der gesamte Fragebogen, nur in der Praxis überprüfen. Daher werden üblicherweise Tests in Form von Intensiv-Interviews und sogenannten Pretests durchgeführt. In Intensivinterviews werden beispielsweise die Ausgestaltung des Szenarios und die zur Illustration geplanten Materialien mit Experten diskutiert. Dies können Experten für verschiedene Teilaspekte des Szenarios sein, wie Naturwissenschaftler oder auch Stadt- und Regionalplaner, sowie Repräsentanten der relevanten Bevölkerungsgruppe, die das sprachliche Niveau und die Allgemeinverständlichkeit prüfen. In Pretests wird anschließend die Bevölkerung selbst einbezogen. Mit zufällig ausgewählten Haushalten werden Interviews durchgeführt, die auf dem zu prüfenden Fragebogen basieren. Diese Interviews werden, bei Einverständnis der Befragten, typischerweise aufgezeichnet, so dass sich ergebende Rückfragen, Unklarheiten und Probleme in den Pretests hinterher so genau wie möglich in eine Weiterentwicklung zum Beispiel des Szenarios einfließen können. Die Ergebnisse der Pretests und die daraus gewonnenen Informationen werden in dieser Vorbereitung des endgültigen Fragebogens sehr intensiv genutzt, zur späteren Auswertung der Hauptbefragung jedoch nicht mehr herangezogen.

1.3.2 Der hypothetische Markt

In einem zweiten Schritt wird den Befragten der sogenannte Zahlungsmechanismus erläutert. Zum einen muss geklärt werden, in welcher Form die Zahlung für das Umweltgut stattfinden soll. Hierbei gilt es, ein *Zahlungsvehikel* zu finden, das bei den Befragten nicht mit negativen Emotionen belastet ist und instinktiv abgelehnt wird. Schlägt man zum Beispiel die Finanzierung des Projektes über eine Steuer vor, so kann es geschehen, dass die Befragten eine Zahlungsbereitschaft von Null äußern, obwohl sie das Projekt selbst schätzen. Sie lehnen vielmehr die Zahlung einer zusätzlichen Steuer ab, weil sie der Auffassung sind, es gäbe doch schon zu viele Steuern und der Staat würde sich bei Steuern ohnehin nur bereichern. Die Gefahr

des "Payment Vehicle Bias", einer Verzerrung der Ergebnisse durch die Wahl eines misstrauten oder missverständlichen Zahlungsvermittlers, wurde in der CVM-Literatur schon recht früh erkannt. Studien von Brookshire, Randall und Stoll (1980) [47], Rowe, d'Arge und Brookshire (1980) [218] oder Daubert und Young (1981) [81] weisen auf den Zusammenhang zwischen dem Zahlungsvermittlungsmechanismus und den mittleren Zahlungsbereitschaften sowie der Anzahl der Protestantworten hin. Derartige Verzerrungen lassen sich im Vorfeld der Hauptbefragung durch intensives Testen des Fragebogens in vorgeschalteten Pretests ausschließen. Darüber hinaus kann auch in der Hauptbefragung noch kontrolliert werden, ob das Zahlungsinstrument von den Befragten verstanden und angenommen wurde, indem sogenannte Debriefing-Fragen mit aufgenommen werden, die das Verständnis und die Rezeption bestimmter Fragen überprüfen. Mithilfe der Debriefing-Fragen lassen sich beispielsweise Antworten, die doch einen Payment Vehicle Bias aufweisen, umkodieren (vgl. Morrison und Blamey (2000) [196]).

Zum anderen muss für die Befragten eine eindeutige Beziehung zwischen ihrer geäußerten Zahlungsbereitschaft und dem Zustandekommen des Projektes sowie der später zu leistenden Zahlung erkennbar sein. Diese Beziehung stellt man durch die Festlegung zweier Regeln her:

- in der *Implementation-Rule* wird festgesetzt, wie die geäußerten Zahlungsbereitschaften das Zustandekommen des Projekts beeinflussen und
- in der *Payment-Rule* wird festgelegt, wie die tatsächliche Zahlung bei Bereitstellung des Umweltguts erfolgt.

Eine nahe liegende Implementierungsregel ist, dass das geplante Projekt nur zustande kommt, wenn die Summe der geäußerten Zahlungsbereitschaften die Kosten des Projekts deckt. Dies ist vor allem dann plausibel, wenn es sich um einmalige Zahlungen handelt und das Umweltgut keine Kosten zum Erhalt und zum zukünftigen Betrieb verursacht. Anderenfalls sollte vermittelt werden, wie diese Kosten aus periodischen, zum Beispiel monatlichen oder jährlichen Zahlungen der Befragten gedeckt werden können.

Die plausibelste Zahlungsregel besteht auf den ersten Blick sicherlich darin, dass jeder der Befragten in dem Umfang seiner geäußerten Zahlungsbereitschaft "zur Kasse gebeten wird". Es ergibt sich jedoch dann schnell bei den Interviewpartnern die Frage, in welchem Umfang nicht interviewte Nutznießer zu dem Umweltgut beitragen. Eine alternative Zahlungsregel besteht

darin, die Kosten des Projekts auf alle betroffenen Haushalte umzulegen, so dass kein direkter Zusammenhang zwischen der im Interview geäußerten Zahlungsbereitschaft und den späteren Zahlungen besteht. Beide Varianten der Zahlungsregel bergen Anreize zu strategischem Verhalten seitens der Befragten.

Wenn sich die spätere Zahlung an der Äußerung der Zahlungsbereitschaft orientiert, man also tatsächlich in dem Umfang zur Kasse gebeten wird, den man selbst angegeben hat, so besteht ein starker Anreiz, die eigene Zahlungsbereitschaft zu untertreiben, in der Hoffnung, dass die anderen eine zur Bereitstellung des Umweltguts hinreichend hohe Zahlungsbereitschaft äußern. Der Befragte wäre Nutznießer des Umweltguts, würde jedoch selbst einen seine Wertschätzung deutlich unterschreitenden Beitrag bezahlen. Dieses Verhalten findet sich unter dem Begriff "Trittbrettfahrerverhalten" oder "Free Riding" in der Literatur. In experimentelle Studien lässt sich allerdings eine weitaus geringere Tendenz zum Trittbrettfahrer beobachten, als die Theorie vermuten lässt (vgl. Schneider und Pommerheine (1981) [223] oder Milon (1989) [192] auch Andreoni (1995) [16] oder Fehr und Gächter (2000) [94]). Der Blick in die Spendenbereitschaft der Bevölkerung für eine Vielzahl gemeinnütziger Zwecke belegt, dass die Free Rider Mentalität bei weiten Teilen der Bevölkerung nicht überwiegt.

Können die Befragten dagegen davon ausgehen, dass die tatsächliche Zahlung nicht mit der Angabe der Zahlungsbereitschaft verbunden ist, so besteht ein deutlicher Anreiz für eine Übertreibung der Zahlungsbereitschaft. Wenn die Implementierungsregel besagt, dass das Umweltgut nur bereit gestellt wird, wenn die Summe der Zahlungsbereitschaften einen bestimmten Betrag übersteigt, so kann durch Äußerung einer höheren Zahlungsbereitschaft diese Bereitstellung gesichert werden, ohne den angegebenen Betrag hinterher bezahlen zu müssen. Diese Strategie ist für den Befragten sinnvoll, wenn die Zahlungsregel beispielsweise eine gleichmäßige Aufteilung der Kosten des Projekts auf alle betroffenen Personen beinhaltet.

Unabhängig von der gewählten Ausgestaltung der Zahlungsregel muss im Interview den Befragten glaubhaft versichert werden, dass es überhaupt zu einer späteren Zahlung kommt. Beschränkt man sich auf Absichtserklärungen über spätere freiwillige Zahlungen, liegen wiederum Anreize zur Äußerung einer höheren Spendenbereitschaft vor. Der Befragte täuscht eine höhere spätere Spendenbereitschaft vor, damit das Gut zunächst überhaupt hergestellt wird (Carson et al. (2001)[60]). Mittels einer speziellen Interviewtechnik, die im nächsten Abschnitt ausführlicher dargestellt wird,

gelingt es jedoch, die Zahlungsbereitschaft für umweltverbesserndes Projekt zu ermitteln, ohne dass die Befragten Anreize haben, falsche Angaben zu machen.

1.3.3 Die Frage nach der Zahlungsbereitschaft

Die offene Fragestellung

Im dritten Schritt erfolgt die eigentliche Frage nach der Zahlungsbereitschaft. Die zunächst einfach anmutende direkte Frage "Was wären Sie für das beschriebene Projekt maximal zu zahlen bereit?", bei der die Befragten einen beliebigen Wert angeben können, hat sich aus einer Reihe von Gründen als problematisch herausgestellt. Die Befragten sind mit der Aufforderung, ihre Zahlungsbereitschaft für ein Umweltgut zu äußern, vor eine ihnen völlig neue Situation gestellt. Schließlich ist der "Kauf" von Umweltgütern keine alltägliche Erfahrung, und beim Kauf von Konsumgütern entscheidet man bei einem gegebenen Preis, ob man das Gut haben möchte. Hier soll der Befragte bei einem gegebenen Gut entscheiden, was er zu zahlen bereit ist. Eine natürliche Reaktion des Befragten wäre somit der Versuch, zu feilschen und somit zunächst einen niedrigen Preis anzugeben, um zu sehen, wie der "Verkäufer" reagiert.

Selbst wenn die Befragten ehrlich antworten möchten, sind sie jedoch häufig mit der Aufgabe, für ein Umweltgut einen "Preis" zu finden, schlichtweg überfordert und unsicher, welche Zahlung realistisch und plausibel ist. Obwohl die offene Frage somit zwar eine direkte Umsetzung des Forschungsanliegens zu sein scheint, sind die Ergebnisse von Studien mit offener Fragestellung oftmals enttäuschend gewesen (vgl. z. Bsp. Bateman et al. (2001) [27]). Die Ergebnisse zeigen eine Häufung der Antworten bei einer Zahlungsbereitschaft von Null, einige Angaben von sehr kleinen Zahlungsbereitschaften und wiederum etliche Angaben von sehr großen Zahlungsbereitschaften. Dies liegt laut Carson, Groves und Machina (1999) [61] daran, dass die Befragten auf der Suche nach einem Anhaltspunkt für den Wert des beschriebenen Umweltguts sich an einer *eigenen* Abschätzung der zu erwartenden Kosten des Projekts orientieren. Diese erwarteten Kosten leiten sie oftmals aus der Erfahrung mit ähnlichen Projekten ab (Eintrittspreise für einen Park o.ä.). Wenn die eigene Zahlungsbereitschaft nun geringer ist als die erwarteten Kosten, so neigen viele Befragte dazu, eine Zahlungsbereitschaft von Null zu äußern. Ist die eigene Zahlungsbereitschaft größer als

die erwarteten Kosten, so tendieren die Befragten dazu, eine Zahlungsbereitschaft zwischen den erwarteten Kosten und der für sie selbst maximal tragbaren Zahlung zu äußern. So kommt die häufig beobachtete Verteilung der Antworten zustande.

Dichotomous Choice

Um ein der Alltagserfahrung mit Konsumgütern ähnlicheres Anreizmuster zu schaffen, wurde das sogenannte Dichotomous Choice Format (Bishop und Heberlein (1979) [33]) entwickelt, das auch als Discrete Response Format (Carson (1991) [57]) oder als Take-it-or-Leave-it-Format (Hoevenagel (1994) [126]) bezeichnet wird. Die Grundidee besteht darin, die Entscheidung über den "Erwerb" des Umweltguts mit einem zu entrichtenden gegebenen "Preis" zu verbinden, und somit die Entscheidungssituation beim Erwerb von Konsumgütern auf die Situation des hypothetischen Marktes zu übertragen. Dies wird in der einfachsten Form als Single Binary Discrete Choice (SBDC) Format umgesetzt, indem den Befragten das Umweltgut beschrieben und ein sogenannter Zahlungsvorschlag unterbreitet wird. Die anschließende Frage lautet beispielsweise: "Würden Sie der Durchführung des Projekts zustimmen, wenn damit eine Zahlung von X verbunden ist?" und die Befragten entscheiden, ob sie das Projekt zu diesem Preisvorschlag haben möchten oder nicht. Da diese Art der Fragestellung an Referenden oder Volksabstimmungen erinnert, hat sich hierfür auch der Begriff "Referendumsformat" etabliert.

Bei Akzeptanz des Programms erhält man Auskunft über die untere Schranke für die Zahlungsbereitschaften, denn die eigentliche Zahlungsbereitschaft kann den vorgeschlagenen Wert weit übersteigen; bei Ablehnung des Programms erhält man eine obere Schranke, denn der vorgeschlagene Betrag übersteigt die wirkliche Zahlungsbereitschaft. Bei der Durchführung einer CVM-Studie im Referendumsformat ist man nun an der Anzahl der Ablehnungen und Zustimmungen zum Projekt bei unterschiedlichen Zahlungsvorschlägen interessiert. Daher werden Untergruppen der insgesamt befragten Haushalte jeweils mit einem unterschiedlichen Zahlungsvorschlag konfrontiert und man erhält aus der gesamten Befragung die Häufigkeit der Zustimmung zu dem Projekt in Abhängigkeit von den Zahlungsvorschlägen.

Carson et al. (2001) [60] und (1999) [61] sehen das Referendumsformat als die einzige mögliche anreizkompatible Form, in der strategisches Verhalten

bei der Beantwortung der Zahlungsbereitschaftsfrage zumindest vermieden werden kann. Im Referendumsformat wird der Befragte aufgefordert, sich zu entscheiden zwischen der Situation "alles bleibt, wie es ist" und einer neuen Situation, in der der Staat ein genau beschriebenes Projekt durchführt und genau definierte Kosten auf den Einzelnen zukommen. Versichert man darüber hinaus glaubwürdig, dass dies die einzige Möglichkeit ist, dass das Projekt zustande kommen wird, und stellt sicher, dass die spätere Zahlung unausweichlich ist, so beinhaltet die Methode erhebliche Anreize, die wahren Präferenzen zu äußern. Wenn die Befragten eine Zahlungsbereitschaft haben, die die vorgeschlagene Zahlung unterschreitet, so lohnt es sich nicht, dem Projekt zuzustimmen, denn zu dem aufgerufenen Preis möchten sie es nicht haben. Wenn sie dagegen eine höhere Zahlungsbereitschaft haben, so werden sie dem Projekt zu dem aufgerufenen Preis auf jeden Fall zustimmen. Zusammenfassend lässt sich festhalten, dass das Referendumsformat für die interviewten Haushalte leichter zu verstehen ist, ihnen die Unsicherheit, die bei einer offenen Fragestellung auftritt, nimmt und wenig Anreize für eine strategisch motivierte Unwahrheit liefert.

Bei aller Überzeugungskraft und Eleganz, mit der das Referendumsformat so die Probleme des offenen Formats ausmerzt, hat es jedoch Nachteile, die sich bei der Berechnung und Auswertung einer Studie zeigen. Die Daten, die man bei einer Befragung im offenen Format erhält, unterscheiden sich erheblich von den Daten einer Befragung im Referendumsformat. Während im ersten Fall die Zahlungsbereitschaften und eine Reihe von erklärenden Variablen direkt gemessen werden können, erhält man aus dem Referendumsformat lediglich Zustimmungen und Ablehnung in Abhängigkeit von den jeweiligen Zahlungsvorschlägen, die ebenfalls durch erklärende Variablen ergänzt werden.

Die genaue Herleitung des im jeweiligen Fall verwendeten statistischen Modells erfolgt in Kapitel 3, jedoch hängt ein Teil der Diskussion um die zu verwendenden Frageformate eng mit der statistischen Modellierung zusammen, so dass an dieser Stelle dem nachfolgenden Kapitel geringfügig vorgegriffen wird. Es leuchtet intuitiv ein, dass die Berechnung von $\sum_{i=1}^N CV Z_i^{01}$ für die Stichprobe aus Antworten auf die offene Fragestellung denkbar einfach ist, schließlich muss man nur die genannten Zahlungsbereitschaften aufsummieren. Die Berechnung der Größe für die gesamte betroffene Bevölkerung erfordert dann, dass man eine durchschnittliche Zahlungsbereitschaft \overline{CVZ}^{01} aus der Stichprobe bestimmt und diese mit der Anzahl der betroffenen Haushalte multipliziert. Für diese Berechnungen ist es so-

gar nicht notwendig, ein statistisches Modell aufzustellen, dieses wird nur benötigt, um die Abhängigkeit der geäußerten Zahlungsbereitschaften von Variablen wie dem Alter, dem Einkommen oder dem Bildungsstand der Betroffenen abzubilden.

Anders ist die Situation für die Berechnung der erwarteten Zahlungsbereitschaft aus SBDC-Daten. Hier muss ein statistisches Modell aufgestellt werden, um von der Häufigkeit der Zustimmungen und Ablehnungen auf \overline{CVZ}^{01} schließen zu können. Da sich nur Informationen über die unteren oder oberen Grenzen der individuellen Zahlungsbereitschaften gewinnen lassen, geht zwangsläufig etwas an Schärfe der Aussage über die mittlere Zahlungsbereitschaft der befragten Bevölkerung verloren. Technisch bedeutet dies, dass die Konfidenzintervalle um die Zahlungsbereitschaften, die aus dem mit dem Referendumsformat gewonnenen Daten ermittelt werden, bei kleinen Stichproben eher groß sind. Die Zuverlässigkeit, mit der die Aussage über die mittlere Zahlungsbereitschaft getroffen werden kann, geht somit zurück. Dies führt im Umkehrschluss zu der Forderung nach großen Stichproben, wenn die Konfidenzintervalle enger sein sollen - und folglich zu erheblich teureren Studien. Der Preis für eine Verbesserung der Anreizkompatibilität und besseren Kommunizierbarkeit der Zahlungsbereitschaftsfrage besteht somit in der notwendigen Vergrößerung des Stichprobenumfangs um ähnlich statistisch zuverlässige Aussagen treffen zu können wie zuvor.

Doppeltes und anderthalbfaches Referendumsformat

Zur Verbesserung dieser statistischen Eigenschaften des Referendumsformats entwickelten Hanemann et al. (1991) [117] und Kanninen (1993) [150] das sogenannte doppelte Referendumsformat. Die Idee dabei ist, die Aussagefähigkeit der unteren sowie der oberen Grenze zu schärfen, oder Informationen über das Intervall, in dem die Zahlungsbereitschaft der Befragten liegt, zu bekommen. Dies lässt sich erreichen, indem man statt einer zwei Fragen stellt. Antworten die Befragten mit "Nein" auf die erste Frage ("Wenn die Durchführung des Projekts mit einer Belastung von $X \text{ €}$ einhergeht, würden Sie dann dem Projekt zustimmen?"), wird ihnen im nächsten Schritt ein niedrigeres Gebot unterbreitet, antworten sie auf die erste Frage mit "Ja", wird der Zahlungsvorschlag im zweiten Schritt erhöht. Falls die Befragten auf beide Fragen mit "Ja" antworten, liegt die Zahlungsbereitschaft in dem offenen Intervall, das durch das zweite Gebot

und ∞ begrenzt ist, bei einer ablehnenden Antwort auf beide Fragen liegt die Zahlungsbereitschaft in dem Intervall, welches durch 0 und das zweite Gebot begrenzt ist. Bei gemischten Antworten liegt die Zahlungsbereitschaft jeweils in dem Intervall zwischen den Geboten aus der ersten und der zweiten Frage.

Das doppelte Referendumsformat schien zunächst eine elegante Lösung zu den Problemen anzubieten, die das offene Format auf der einen Seite und das einfache Referendumsformat auf der anderen Seite aufwarfen. Es bietet scheinbar eine Formulierung der Frage nach der Zahlungsbereitschaft, die den Befragten wenig Anreiz zu strategischem Antwortverhalten gibt und doch in der Voraussage der Zielgröße auch bei kleineren Stichprobengrößen noch verlässliche Werte lieferte, d.h. in diesem Kontext auch noch "vernünftige" Konfidenzintervalle um den ermittelten Wert herum liefert.

Die Praxis hat allerdings gezeigt, dass das doppelte Referendumsformat ebenfalls Anlass zu einer Reihe von Missverständnissen bei den Befragten geben kann. Studien, die das doppelt begrenzte Referendumsformat verwenden, zeigen insofern eine Anomalie in den Daten, als dass die Anzahl der zustimmenden Antworten in einem Intervall davon abzuhängen scheint, ob zunächst einem Gebot zugestimmt wurde, und das höhere dann abgelehnt wurde oder umgekehrt (vgl. auch die Analyse der Daten in Kapitel 3.2.2). So ist in demselben Intervall von Zahlungsbereitschaften zu beobachten, dass die Ablehnung eines Zahlungsvorschlags (an der oberen Grenze) nach vorheriger Zustimmung zu einem geringeren Zahlungsvorschlag (an der unteren Grenze) deutlich häufiger auftritt als die Zustimmung zu einem niedrigeren Zahlungsvorschlag nach Ablehnung eines höheren. Im ersten Fall erhält man eine "Ja-Nein"-Antwort, die bedeutet, dass die Zahlungsbereitschaft die untere Grenze überschreitet und geringer als die obere Grenze des Intervalls ausfällt. Im zweiten Fall ist die "Nein-Ja"-Antwortkombination so zu interpretieren, dass die Zahlungsbereitschaft geringer ist als die obere Grenze und die untere Grenze überschreitet. In beiden Fällen handelt es sich somit um dieselbe Zahlungsbereitschaft, die sich in den Antworten widerspiegelt. Geht man nun davon aus, dass die Fragereihenfolge zufällig auf die Stichprobe verteilt wurde, so ist es verblüffend, dass sich stark unterscheidende Häufigkeiten gemessen werden. Dies lässt sich nur dahingehend interpretieren, dass die Zahlungsbereitschaft von der Fragereihenfolge abhängt.

Es lassen sich verschiedene Erklärungsansätze diskutieren, die alle darauf beruhen, dass die Befragten offensichtlich die Absicht des Interviewers,

durch den zweiten Zahlungsvorschlag die Angabe der Zahlungsbereitschaft zu präzisieren, nicht erkennen oder sogar völlig missverstehen. Eine mögliche Erklärung besteht darin, dass die Befragten die zweite Antwort als einen Versuch des Interviewers interpretieren, mit ihnen zu "feilschen" und den Preis in die Höhe zu treiben (vgl. DeShazo (2002) [84]). Sie reagieren darauf, laut dieser Theorie, verstimmt und lehnen den zweiten Zahlungsvorschlag auch dann ab, wenn sie eine höhere Zahlungsbereitschaft besitzen. So wäre ein Teil dieser "Ja-Nein"-Antworten als "Ja-Ja"-Antworten zu interpretieren, und die Zahl der "echten Ja-Nein"-Antworten würde eher der Zahl der "Nein-Ja"-Antworten entsprechen. Alternativ lässt sich die zu niedrige Zahl originärer "Nein-Ja"-Antworten dadurch erklären, dass die Befragten nach einer Ablehnung des höheren Zahlungsvorschlags auch den niedrigeren ablehnen, weil sie befürchten, für weniger Geld auch nur eine "Sparversion" des Umweltguts zu erhalten (vgl. Cooper et al. (2002) [72]). Sie lehnen somit den Zahlungsvorschlag ab, obwohl ihre Zahlungsbereitschaft größer wäre. Diese Antworten finden sich als "Nein-Nein"-Antworten festgehalten, lassen sich aber eigentlich als "Nein-Ja"-Antworten interpretieren.

Einen möglichen Ausweg aus diesen Verzerrungen bieten Vorschläge für ein anderthalbfach begrenztes Format (vgl. Cooper et al. (2002) [72] und DeShazo (2002) [84]). Cooper et al. schlagen vor, die Befragten zunächst über das Intervall zu informieren, innerhalb dessen die Zahlungsvorschläge liegen werden. Dann wird zufällig der Zahlungsvorschlag an der oberen oder an der unteren Grenze unterbreitet. Wird der untere Zahlungsvorschlag abgelehnt, erfolgt kein weiterer Vorschlag, ebenso wenig, wenn dem oberen Vorschlag zugestimmt wird. In den beiden anderen möglichen Situationen hingegen erfolgt ein weiterer Vorschlag; i.e. nach Zustimmung zum unteren Vorschlag oder bei Ablehnung des oberen Vorschlags. Aus dieser Einschränkung der Möglichkeit für zweite Zahlungsvorschläge stammt die Bezeichnung als anderthalbfaches Referendumsformat. DeShazo (2002) hingegen schlägt vor, nur fallende Zahlungsvorschläge, die einer ablehnenden Antwort folgen, zu unterbreiten. Allerdings opfern diese Formate wiederum einen Teil des Effizienzvorteils der doppelten Formate und die bisherigen Anwendungen geben wenig Anlass zur Hoffnung auf eine tatsächliche Verbesserung der Ergebnisse der Befragungen.

In theoretischer Hinsicht bereitet die Diskussion um die Missverständnisse beim doppelten Referendumsformat ein leichtes Unwohlsein, muss man doch akzeptieren, dass die der ersten Antwort zugrunde liegende Zahlungs-

bereitschaft sich von derjenigen unterscheidet, die der zweiten Antwort zugrunde liegt - eine Annahme, die unterstellt, dass sich die Präferenzen der Befragten zwischen den Antworten ändern, oder dass sie in der zweiten Frage einen anderen Gegenstand bewerten. Die genauen Vorgänge bei der Entscheidung der Befragten für oder gegen einen Zahlungsvorschlag sollten gründlicher erforscht werden, bevor man sich für eine der vorgeschlagenen Änderungen des Frageformats entscheidet. Letztlich erscheinen die implizierten Verzerrungen als ein zu hoher Preis für die größere Genauigkeit der Messung der Zahlungsbereitschaft, zumal sich die Einengung von Konfidenzintervallen auch mittels einer anderen Herangehensweise erreichen lässt, wie weiter unten noch ausführlich gezeigt wird.

Die Höhe der geäußerten Zahlungsbereitschaft

Abschließend werden die Einflüsse der geschilderten Frageformaten auf die Höhe der geäußerten Zahlungsbereitschaften betrachtet. Für den Einsatz der Contingent Valuation Methode bei praktischen Kosten-Nutzen-Analysen ist es wichtig, zu überprüfen, ob Schätzungen der Zahlungsbereitschaft, die aus einer offenen Befragung stammen, systematisch von Schätzungen der Zahlungsbereitschaft abweichen, die auf einer Befragung nach der Referendumsmethode basieren, denn die gemessene Wohlfahrtsänderung sollte nach Möglichkeit nicht vom Messverfahren abhängen. Es hat sich gezeigt, dass die ermittelten Zahlungsbereitschaften beim Referendumsformat in der Regel die Ergebnisse von offenen Befragungen überschreiten. Dies hat in der Literatur zu einer kontroversen Diskussion und verschiedenen Erklärungsansätzen geführt; nicht zuletzt sind beim Einsatz des Verfahrens zur Schadensbemessung die Verursacher natürlich an einer möglichst konservativen, d.h. niedrigen Schätzung interessiert.

Mitchell und Carson (1989) [193] führen den niedrigeren Wert, der in offenen Formaten gefunden wird, wie weiter oben erwähnt, auf strategisches Verhalten bei den Befragten zurück. Diese versuchen ihre tatsächlichen Zahlungen zu "drücken", da sie davon ausgehen, mit großer Wahrscheinlichkeit tatsächlich zahlen zu müssen. Zusätzlich vermuten die Befragten, ihr eigener Beitrag zum Zustandekommen des Projektes sei klein, und gehen daher davon aus, dass die Angabe einer untertriebenen Zahlungsbereitschaft keinen großen Schaden anrichte. Kanninen und Kriström (1993) [151] untersuchen, ob das Referendumsformat sogenannten Ankereffekten unterliegt, d.h. ob der Zahlungsvorschlag den Befragten den Eindruck ver-

mittelt, dieses sei der Preis, zu dem das Gut zu haben ist, obwohl gemessen werden soll, ob es der Preis ist, zu dem die Befragten das Gut haben wollen. Im Vergleich von offenen und geschlossenen Frageformulierungen finden Kanninen und Kriström zwar ebenfalls eine höhere geäußerte Zahlungsbereitschaft auf die Referendumsfrage, Ankereffekte lassen sich jedoch nicht nachweisen.

Im Vergleich beider Frageformate bei der Bewertung von Konsumgütern und öffentlichen Gütern stellt sich heraus, dass die Unterbewertung im offenen Format nur bei öffentlichen Gütern gilt und somit die Unsicherheit der Befragten bei der Bewertung eines öffentlichen Guts reflektiert (vgl. Kealy und Turner (1993) [152]). Halvorsen und Saelensminde (1998) [115] hingegen führen die unterschiedlichen Ergebnisse auf die unterschiedlichen Anforderungen an eine korrekte Modellspezifizierung zurück. Im Falle des Referendumsformats ergibt sich der Erwartungswert der Zahlungsbereitschaft in der Regel, wie weiter unten gezeigt wird, als Quotient aus zwei ökonometrisch geschätzten Parametern, die ihrerseits einer Fülle von möglichen Modellspezifizierungen entspringen können. Welsh und Poe (1998) testeten ein Referendumsformat, bei welchem die Befragten zusätzlich angeben können, wie sicher sie sich ihrer Entscheidung sind. Es stellt sich heraus, dass ein großer Teil der Befragten zur Zustimmung zu einem Projektvorschlag und dem Gebot neigt, selbst wenn sie nicht sehr sicher sind, diesen Betrag auch tatsächlich zahlen zu wollen.

Obwohl das 1993 zur Klärung der Validität und Reliabilität von Contingent Valuation Ergebnissen eingesetzte Blue Ribbon Panel um Arrow und andere [18] sich explizit für die Verwendung geschlossener Formate ausgesprochen hat, ist die Debatte um das Frageformat bis heute nicht abgeklungen. Jüngere Studien versuchen, mithilfe von Experimenten die Anreizstrukturen der Formate zu überprüfen (vgl. Lunander (1998) [180] sowie Bohara et al. (1998) [37] oder Botelho und Pinto (2001) [39]). Allerdings sind die experimentellen Situationen häufig wiederum so weit von der Realität einer Bewertungsstudie entfernt, dass man keine endgültige Klärung der Frage erwarten darf. Bislang verwenden sehr wenige Studien das anderthalbfache Referendumsformat, so dass sich hier keine abschließende Bewertung aus der Praxis ergibt.

1.3.4 Soziodemographische und psychosoziale Daten

Im Prinzip ist ein Contingent Valuation Interview mit den bis hierher aufgeführten Schritten beendet. Man erhält entweder eine Reihe von offen ausgedrückten Zahlungsbereitschaften und bildet hieraus einen einfachen Mittelwert oder wendet auf die Antworten aus der geschlossenen Fragestellung die weiter unten ausführlich erläuterten statistischen Methoden an. Allerdings hat es sich aus einer Vielzahl von Gründen als sinnvoll erwiesen, weitere Daten zu erheben.

So kann es im Interesse des politischen Entscheidungsträgers liegen, sich ein Bild von den Zusammenhängen zwischen bestimmten Charakteristika seiner Bürger (wie dem Einkommen, dem Alter oder dem Ausbildungsgrad) und den geäußerten Zahlungsbereitschaften bzw. Zustimmungen und Ablehnungen eines Zahlungsvorschlags zu machen. Bedient das geplante Projekt vielleicht nur den jungen oder den reichen Teil der Bevölkerung? Wie ist die Akzeptanz des Projekts in einer ausschlaggebenden Wählerschicht? Derartige Frage lassen sich teilweise beantworten, wenn in einer Contingent Valuation Studie ein relativ breites Bild der Einstellungen und Charakteristika der betroffenen Bevölkerung aufgenommen wird.

Darüber hinaus lässt sich anhand soziodemographischer Daten die Repräsentativität einer Stichprobe durch den Vergleich der Verteilung verschiedener Merkmale mit einer amtlichen Statistik überprüfen. Tabelle 3.8 in Kapitel 3 zeigt einen derartigen Vergleich für eine Studie zur Bewertung eines Rekultivierungsprojekts in der Nähe von Cottbus in Brandenburg. Darüber hinaus werden sozioökonomische Daten für Plausibilitätstests des Ergebnisses genutzt. Aus dem ökonomischen Modell erwarten wir, dass die Zahlungsbereitschaft für ein Umweltgut nicht nur mit steigendem Preis fällt, sondern auch mit steigendem Pro-Kopf-Einkommen steigt und mit steigendem Alter fällt, da der Gebrauchsnutzen eines Umweltguts mit steigendem Alter sinkt. Bei einer Untersuchung, die diese Zusammenhänge mit den entgegengesetzten Vorzeichen zeigt, wäre man auf jeden Fall zunächst skeptisch gegenüber den Ergebnissen.

Eine Größe, die oftmals wegen ihres Einflusses auf die Zahlungsbereitschaft für Verwunderung sorgt, ist die Haushaltsgröße. Es ist verbreitet, in CVM-Interviews anstelle der individuellen die Zahlungsbereitschaften eines gesamten Haushalts zu ermitteln, da eine Umweltveränderung für alle Haushaltsmitglieder, auch für nichterwerbsfähige Kinder, nutzenändernd wirkt. Es wird hierbei unterstellt, dass der Haushaltsvorstand die Ent-

scheidung über die Zahlung für eine Umweltveränderung "zum Wohle aller Haushaltsmitglieder" ebenso fällen kann, wie auch andere alltägliche Anschaffungen entschieden werden. Zunächst würde man annehmen, dass der *Haushaltsnutzen* mit zunehmender Größe eines Haushalts steigt. In der Praxis lässt sich jedoch ebenso oft das Gegenteil beobachten. Eine mögliche Erklärung hierfür lautet (vgl. Ahlheim und Lehr (2001) [8]), dass bei umfangreichen Projekten, die für den Haushalt eine erhebliche Bedeutung haben, die *Zahlungsbereitschaft* größer ist als die *Zahlungsfähigkeit* des Haushalts. Denkbar ist dies zum Beispiel bei kinderreichen jungen Familien, deren verfügbares Haushaltseinkommen nur sehr geringe Spielräume aufweist. Anstelle der den Nutzenzuwachs anzeigenden Zahlungsbereitschaft würde eine solche Familie in einer Befragung oftmals nur die ihre Einkommenssituation anzeigende Zahlungsfähigkeit angeben.

Aber nicht nur die "harten Fakten" werden abgefragt, sondern auch Einstellungen zur Umwelt, zum Staat und zur Einschätzung der eigenen wirtschaftlichen Situation. Dies dient zum einen ebenfalls der Verifikation des Ergebnisses, zum anderen jedoch auch dazu, dass besondere soziokulturelle Effekte überprüft werden können. So stellt sich zum Beispiel die Frage, ob eine Contingent Valuation Untersuchung in Ostdeutschland auch zu spezifischen ostdeutschen Ergebnissen führt, da dort die Wahrnehmung der Bereitstellung öffentlicher Güter durch den Staat aufgrund der Erfahrungen in einem sozialistischen Staat anders ausgeprägt sein kann (vgl. Ahlheim et al. (2004) [7]). Um dies überprüfen zu können, müssen die Einstellungen zu verschiedenen öffentlichen Gütern und verschiedenen Formen der Bereitstellung abgefragt werden.

Schließlich enthält dieser Teil des Fragebogens die bereits angesprochenen Debriefing-Fragen. Hier wird eine zusätzliche Klarstellung der bereits gegebenen Antworten versucht. Debriefing-Fragen werden zum einen eingesetzt, um sicher zu stellen, dass das zu bewertende Gut und der hypothetische Markt richtig verstanden worden sind. Darüber hinaus werden Debriefing-Fragen verwendet, um sicher zu stellen, dass die Befragten tatsächlich vor dem Hintergrund ihrer eigenen Präferenzen beantwortet haben. Blamey und Common [35] stellen 1995 fest, dass die Befragten die Höhe ihrer Zahlungsbereitschaft danach richten, was ihrer Meinung nach von ihnen als "guten Bürgern" erwartet wird (citizen versus consumer preferences). Während sie selbst dem vorgeschlagenen Projekt vielleicht indifferent oder sogar ablehnend gegenüberstehen, stimmen sie dennoch einem Projekt zu. Die Ursache hierfür kann sein, dass die Befragten ihren Bürgerpflichten

auch dadurch genügen wollen, dass sie ein eventuell dem Gemeinwohl förderliches Projekt unterstützen. Eine derartige Haltung, so die Kritiker der Methode, führt dazu, dass die geäußerten Zahlungsbereitschaften nicht mehr die individuellen Wohlfahrtsänderungen reflektieren.

Einen ähnlichen Vorwurf erfährt die CVM, wenn die Befragten einem Zahlungsvorschlag zugestimmt haben, weil sie dadurch den bereits kurz erwähnten Warm Glow of Giving (Andreoni (1990) [15] erfahren. Diese Nutzensteigerung entspringt einfach aus der guten Tat des Gebens und nicht einer Nutzenänderung aufgrund eines bestimmten umweltverbessern- den Projekts. Daher sollten auch Zahlungsbereitschaften, die aus diesem Motiv geäußert werden, nicht in die Kalkulation einbezogen werden. Grundsätzlich gilt, dass die Motivation für eine bestimmte Zahlungsbereitschaft für die Bereitstellung eines Umweltguts dem Forscher eigentlich gleichgültig sein kann. Ebenso wie bei Marktgütern, bei denen es auch irrelevant ist, ob man den Porsche kauft, weil man gerne schnelle Autos fährt oder weil es ein Prestigeobjekt ist, muss auch hier gelten, dass die Kaufwilligkeit zu einem bestimmten Preis gerade die Präferenzen für das Umweltgut reflektiert. Schließlich interessieren wir uns vorrangig für die Nutzenänderung, und nicht für die Motive eines Individuums. Es gilt jedoch auch, dass die Ergebnisse einer CVM Studie so belastbar wie möglich sein sollten. Daher werden Debriefing-Fragen eingesetzt, die sicherstellen sollen, aus welchen Gründen die Befragten einem Zahlungsvorschlag und dem Projekt zugestimmt haben.

1.3.5 CVM-Interview bei Umweltschäden

Da in der Einführung zur Umweltbewertung auch der Einsatz von CVM-Studien zur Bewertung von Umweltschäden diskutiert wurde, soll an dieser Stelle kurz auf die Fragebogengestaltung für einen solchen Einsatz eingegangen werden. Die Schritte zur Bewertung einer positiven und einer negativen Umweltveränderung sind im Wesentlichen gleich. Zur Bewertung eines Schadens erscheint es zunächst natürlich, nach dem Geldbetrag zu fragen, den der Haushalt als Entschädigung fordert, um nicht schlechter gestellt zu sein als vor Eintritt des Schadens. In der Praxis stellt sich allerdings heraus, dass dies zu (politisch unerwünschten) Übertreibungen der Schadensersatzforderung führt. Daher geht man typischerweise dazu über, die Zahlungsbereitschaft für ein den Umweltschaden vermeidendes Projekt zu ermitteln. Diese Technik lässt sich am Besten anhand der Bewertung

der Schäden infolge des Exxon Valdez Unfalls verdeutlichen. Anstatt die Bevölkerung zu fragen, wie groß ihre Entschädigungsforderung für die verseuchte Küste, die getöteten Seevögel und alle anderen Effekte ist, wurde die Zahlungsbereitschaft für ein Projekt erhoben, das derartige Schäden in Zukunft verhindert. Ganz konkret wurde die Begleitung von Tankern durch kleinere Schiffe vorgeschlagen, die verhindern könnten, dass ein Tanker in gefährlichen Gewässern auf Grund läuft. Die so ermittelte Zahlungsbereitschaft zur Vermeidung ähnlicher Unfälle im Meer vor der Küste von Alaska wurde dann als Entschädigungsforderung für den Exxon Valdez Unfall eingesetzt.

Diese Technik wurde vor allem wegen der erbitterten Debatte entwickelt, die die Bewertung des Exxon Valdez Unfalls und seiner Folgen ausgelöst hatten und die nicht durch "überhöhte Forderungen" intensiviert werden sollte. Wirtschaftssubjekte geben als Entschädigungsforderungen für den Verlust eines Gutes im Durchschnitt höhere Beträge an, als sie in der umgekehrten Situation für das Gut zu zahlen bereit wären. Die Ursache wird einer Art "Endowment"-Effekt zugeschrieben, wobei es hier auch einen moralischen Endowment-Effekt zu geben scheint. So haben Boyce et al. (1992) [40] einen Teil der amerikanischen Bevölkerung gefragt, was sie als Entschädigung von den Japanern dafür verlangen würden, dass diese den Walfang weiterhin fortsetzen dürften. Umgekehrt wurde ein Teil der Bevölkerung gefragt, was sie bereit wären, den Japanern für eine Beendigung des Walfangs zu zahlen. Natürlich war dieser Betrag deutlich niedriger, da die Bevölkerung sich moralisch im Recht fühlte als Walschützer und den "Bösen" nicht auch noch Geld geben wollte für etwas, das sich gewissermaßen moralisch verbietet. Notorsche Kritiker der Methode werten den Unterschied zwischen der Willingness to Accept und der Willingness to Pay als nur ein weiteres Indiz für die Launenhaftigkeit der Contingent Valuation und die Literatur kennt eine lebhafte Diskussion hierzu (vgl. Diamond und Hausman (1994) [87] oder Hanemann (1991) [120]).

Zusammenfassend lässt sich festhalten, dass der zunächst so einfach anmutende Grundgedanke der Contingent Valuation Methode, die Zahlungsbereitschaft der Bevölkerung zu ermitteln, indem man sie befragt, bei der Umsetzung in die Realität erhebliche Vorarbeiten verlangt. Die "naive" Formulierung dieser Frage hat in der Anfangszeit der Contingent Valuation Methode zu einer Vielzahl von Verzerrungen der Ergebnisse geführt, die die Methode zeitweise in Misskredit brachten. Inzwischen sind für eine Reihe von problematischen Fragen Techniken entwickelt worden, die helfen, diese

ursprünglichen Fehler zu vermeiden. In modernen Contingent Valuation-Fragebögen wird das Szenario so formuliert, dass ein Part-Whole-Bias nahezu ausgeschlossen ist und Debriefing-Fragen sichern das Verständnis des Szenarios bei den Befragten und geben Aufschluss über eventuell vorliegende "falsche" Motive. So können gewissermaßen als zusätzliche Absicherung missverstandene Fragen auch noch im Nachhinein identifiziert werden. Darüber hinaus wird der Fragebogen mehrfach im Vorfeld der eigentlichen Hauptbefragung im Rahmen von Pretests überprüft und intensiv mit der Haushalten aus der betroffenen Bevölkerung sowie mit Experten diskutiert.

Zur tatsächlichen Berechnung und Erklärung der Zahlungsbereitschaft aus den in einer repräsentativen Befragung gewonnen Daten ist im nächsten Schritt die Formulierung eines statistischen Modells notwendig. Die Herleitung eines solchen Modells aus der ökonomischen Theorie sowie der Einfluss der Modellwahl auf die Ergebnisse einer Studie sind Gegenstand des nächsten Kapitels.

Kapitel 2

Das statistische Modell

Nachdem im ersten Kapitel dieser Arbeit die in einer Umweltbewertungsstudie zu messende Größe - die individuelle Zahlungsbereitschaft für ein umweltverbesserndes Projekt - aus der Wohlfahrtstheorie hergeleitet wurde und als Messmethode die CVM wegen der geschilderten Vorteile festgelegt wurde, sind die nächsten Kapitel der Auswertung der bei Durchführung einer CVM-Studie gemessenen Daten gewidmet. Eine einfache Auswertung liegt zunächst in der Beschreibung der innerhalb der Stichprobe gemessenen Daten, beispielsweise durch ihren Mittelwert, die mittlere Abweichung von diesem Mittelwert oder durch die höchsten und niedrigsten Werte für eine gemessene Größe. So kann der Mittelwert aller auf eine offene Fragestellung geäußerten Zahlungsbereitschaften gebildet werden oder mittels der Berechnung des Medians die Frage nach demjenigen Betrag beantwortet werden, den die Hälfte aller Befragten zu zahlen bereit wären. Im Falle einer geschlossenen Befragung lässt sich die Anzahl der Zustimmungen und Ablehnungen zu einem bestimmten Zahlungsvorschlag ermitteln.

Weitaus interessanter jedoch ist die Analyse von möglichen Zusammenhängen zwischen einzelnen individuellen Charakteristika der Befragten und der Höhe ihrer Zahlungsbereitschaftsäußerung, bzw. die Ermittlung einer mittleren Zahlungsbereitschaft aus der Anzahl der Zustimmungen und Ablehnungen nach Höhe der Zahlungsvorschläge. Eine derartige Analyse erfordert zum einen die Aufstellung von Hypothesen, die sich aus der ökonomischen Theorie ableiten lassen und zum anderen die Aufstellung eines statistischen Modells, mit dessen Hilfe die Hypothesen getestet werden können. Aus dem großen Vorrat möglicher statistischer Modelle werden nachfolgend die in der umweltökonomischen Bewertungsliteratur überwiegend verwendeten vorgestellt, um dann in einem zweiten Schritt zu zeigen,

was geschieht, wenn der herkömmliche statistische Bezugsrahmen verlassen wird. Ziel dieses Kapitels ist es herzuleiten, wie die Analyse der Daten aus einer Bewertungsstudie durch diesen Schritt verbessert werden kann.

2.1 Ökonomische Fundierung

In der praktischen Anwendung der CVM wird häufig "vergessen", dass die verwendeten Modelle letztlich mit der Nutzentheorie vereinbar sein müssen. Die Darstellung praktischer Anwendungen beginnt vielfach mit einer Darstellung der individuellen Zahlungsbereitschaftsfunktion, die Zusammenhänge zwischen der Zahlungsbereitschaft und dem Alter der Befragten oder ihrem Einkommen, abbildet. Allerdings unterliegen die erklärenden Variablen, die in eine derartige Gleichung aufgenommen werden können, sowohl in ihrer funktionalen Form als auch in ihrer Zulässigkeit als erklärende Variable genau genommen einer Reihe von Einschränkungen, die von der genauen Spezifikation des der Berechnung zugrunde liegenden Nutzenmodells abhängen. Die Antworten auf die Frage nach der Zahlungsbereitschaft für ein bestimmtes umweltverbesserndes Projekt müssen ableitbar sein aus der Entscheidung der Befragten zwischen ihrem Nutzen ohne das vorgeschlagene Projekt und dem Nutzen mit dem Projekt und nach Zahlung eines bestimmten Betrags.

Ein Individuum entscheidet sich für ein Projekt und eine Zahlung A_i , wenn es hierdurch mindestens ebenso gut gestellt ist wie zuvor; d. h., es muss gelten

$$v_i(y_i, \mathbf{s}_i, \mathbf{z}^0) \leq v_i(y_i - A_i, \mathbf{s}_i, \mathbf{z}^1), \quad (2.1)$$

wobei y wie zuvor das Einkommen bezeichnet, \mathbf{z} den Zustand der Umwelt charakterisiert und der Vektor \mathbf{s} der individuellen Charakteristika wie Alter, Ausbildungsgrad und Einstellungen des Befragten explizit in die Nutzenfunktion aufgenommen wurde. Im Folgenden wird unterstellt, dass es sich bei y_i um das verfügbare Einkommen handelt. Darüber hinaus werden die Marktgüterpreise als konstant betrachtet und in der Darstellung unterdrückt. In einer *offenen* Befragung, wie sie zu Beginn von Kapitel 1.3.3 vorgestellt wurde, sind die Befragten aufgefordert, ihre Zahlungsbereitschaften für das zuvor beschriebene Projekt anzugeben. Bei $A_i = WTP_i$ sind die Befragten gerade indifferent, d.h. in (2.1) gilt das Gleichheitszeichen. Diesen Beitrag geben die Befragten als Zahlungsbereitschaft für das Projekt an, wenn sie wahrheitsgemäß antworten. In einer *geschlossenen* Befragung

hingegen stimmen die Befragten dem Projekt und dem Zahlungsvorschlag (Bid) zu, wenn gilt $\text{Bid}_i \leq \text{WTP}_i$, d.h. wenn der Zahlungsvorschlag höchstens so groß ist wie ihre Zahlungsbereitschaft (vgl. McFadden und Leonard (1993) [190] S.171).

Allerdings ist nur den Befragten selbst der dieser Abwägung zugrunde liegende Nutzen bekannt. Dem Beobachter hingegen sind die verschiedenen Charakteristika \mathbf{s}_i der Befragten, der Zustand der Umwelt \mathbf{z}^k und das Antwortverhalten der Befragten zugänglich. Der Beobachter kann eine Funktion definieren, die diese Beobachtungen mit dem Nutzen der Befragten und somit ihrer Zahlungsbereitschaft in Verbindung bringt. Diese Funktion $\tilde{v}_i(y_i^k, \mathbf{s}_i, \mathbf{z}^k)$ wird häufig als *repräsentative Nutzenfunktion* bezeichnet (vgl. Train (2001) [242]); y_i^k steht für das verfügbare Einkommen mit bzw. ohne Projekt und Zahlung. Für die - unbekanntenen - individuellen Nutzenfunktionen lässt sich somit eine Zerlegung in einen beobachtbaren Teil und einen unbeobachtbaren Teil ε_i denken, der als zufällig aufgefasst wird

$$v_i(y_i^k, \mathbf{s}_i, \mathbf{z}^k) = \tilde{v}_i(y_i^k, \mathbf{s}_i, \mathbf{z}^k) + \varepsilon_i^k. \quad (2.2)$$

Train (2001) [242] diskutiert die möglichen Interpretationen dieser Zerlegung. Die Verteilung der ε_i wird häufig als die Verteilung des unbeobachtbaren Teils der individuellen Nutzen in der Bevölkerung interpretiert. Alternativ lässt sich diese Verteilung als die vom Beobachter subjektiv angenommene Verteilung dieses unbeobachtbaren Nutzens annehmen. Eine dritte Interpretation besteht darin, die Verteilung der ε_i als das Ergebnis von Faktoren und Einflüssen zu betrachten, die sich dem Individuum selbst entziehen und somit dem Individuum ebenfalls als zufällig erscheinen.

Modelle, die auf einer solchen Zerlegung der Nutzenfunktion basieren, werden als Random Utility Modelle (RUM) oder Zufallsnutzenmodelle bezeichnet, die ersten Ursprünge finden sich bei Thurstone (1927) [240]. Marschak (1960) [184] entwickelte Thurstones Modell der Reaktion auf psychologische Stimuli für Entscheidungssituationen unter Nutzenmaximierung weiter. Einführende Texte zur Modellierung von Contingent Valuation Daten (vgl. Haab und McConnell (2002) [113], Hanemann (1984) [119] oder (2002) [118] und Mitchell und Carson (1989) [193]) verweisen häufig auf die Formulierung des Random Utility Modells bei McFadden z. Bsp. (1974) [189].

Die in einer offenen Befragung direkt ermittelten Zahlungsbereitschaften sind im Random Utility Modell ebenfalls Zufallsgrößen, denn sie folgen aus der Auflösung von (2.1) nach A_i , wenn das Gleichheitszeichen gilt.

Will man die Einflüsse der unabhängigen Variablen auf die Zahlungsbereitschaft untersuchen, so verwendet man zur Schätzung der Zahlungsbereitschaftsfunktion überwiegend einfache lineare Regressionen. Es lässt sich jedoch diskutieren, ob die möglichen Zahlungsbereitschaften nicht begrenzt sind, sei es durch die Obergrenze des verfügbaren Einkommens oder einfach durch die Tatsache, dass in einer Befragung nur positive Werte als gültige Antworten auf die Frage nach der Zahlungsbereitschaft zugelassen wurden. Um dies abzubilden, bietet es sich an, auf Modelle für gestutzte oder zensierte Daten zurückzugreifen (für eine Tobit-Analyse von Daten aus einer Befragung, die auf die Äußerung positiver Werte beschränkt war, vgl. Halstead et al. (1991) [114]).

Wenn die Befragung im Referendumsformat durchgeführt wurde, so sind nicht mehr einzelne Angaben der Zahlungsbereitschaft beobachtbar, sondern die Zustimmungen der Befragten zu einem zuvor beschriebenen Projekt und einem Zahlungsvorschlag oder die Ablehnung dieser Kombination. Die Wahrscheinlichkeit für eine Zustimmung ergibt sich als

$$\begin{aligned} Pr(yes_i) &= Pr(\tilde{v}_i(y_i^1, \mathbf{s}_i, \mathbf{z}^1) + \varepsilon_i^1 \geq \tilde{v}_i(y_i^0, \mathbf{s}_i, \mathbf{z}^0) + \varepsilon_i^0) \\ &= Pr(\varepsilon_i^1 - \varepsilon_i^0 \geq \tilde{v}_i(y_i^0, \mathbf{s}_i, \mathbf{z}^0) - \tilde{v}_i(y_i^1, \mathbf{s}_i, \mathbf{z}^1)). \end{aligned} \quad (2.3)$$

Die indirekte Nutzenfunktion $v_i(y_i^k, \mathbf{s}_i, \mathbf{z}^1)$ ist zunehmend im Einkommen, nicht zunehmend in den Marktpreisen, die hier unterdrückt sind, homogen vom Grad 0 in Preisen und Einkommen und für umweltverbessernde Projekte zunehmend in der Umweltqualität. An dieser Stelle endet bereits fast alles, was die Ökonomie vorgibt.

2.2 Klassische parametrische Modelle

Die klassische Vorgehensweise besteht in der Übersetzung des Entscheidungsmodells in ein parametrisches Modell und der anschließenden ökonomischen Schätzung der Parameter. Hierfür gilt es, die Modellierung der stochastischen Komponente, das Zusammenspiel zwischen der deterministischen und der stochastischen Komponente sowie eine Modellierung der deterministischen Komponente zu wählen. Aus den vielen möglichen Modellen haben sich einige in der Durchführung praktischer CVM-Studien etabliert, die nachfolgend vorgestellt werden. Das Ziel dieses Kapitels besteht letztlich nicht in einer umfassenden Darstellung möglicher Modellierungen, sondern darin, die Grundlagen für die Ergänzung der häufig verwendeten Modelle durch einen Wechsel des statistischen Bezugsrahmens zu

legen und eine Basis für die Diskussion des Nutzentransfers zur Bewertung zukünftiger Projekte im zweiten Teil dieser Arbeit zu schaffen.

Führt man in Gleichung (2.3) die Notationen $v_i(\cdot, \mathbf{z}^k) = v_i^k$ mit $k = 0, 1$ für den Ausgangs- und Endzustand der Umwelt, $\eta = \varepsilon_i^0 - \varepsilon_i^1$ und $\Delta\tilde{v} = \tilde{v}_i^1 - \tilde{v}_i^0$ ein, und definiert ferner durch $F_\eta(\bullet)$ die Verteilungsfunktion von η , so lässt sich die Zustimmungswahrscheinlichkeit schreiben als

$$Pr(yes_i) = Pr(\eta \leq \Delta\tilde{v}) = F_\eta(\Delta\tilde{v}) \quad (2.4)$$

und das Modell lässt sich durch die Wahl einer funktionalen Form für $\Delta\tilde{v}$ und die Festlegung einer Verteilung von η spezifizieren. Für eine korrekte Identifizierung und Spezifizierung eines Modells, das auf der Entscheidungssituation der Individuen basiert, muss berücksichtigt werden, dass es sich beim Nutzen um ein ordinales Konzept handelt (vgl. Kapitel 1.2). Dies drückt sich darin aus, dass nur Nutzdifferenzen für die Entscheidung eines Individuums von Bedeutung sind und das jeweilige absolute Nutzenniveau als Wert bedeutungslos ist. Wenn jedem Nutzenniveau, das in den zur Wahl stehenden Alternativen erreicht wird, dieselbe Konstante hinzugefügt wird, ändert sich die Entscheidung des Individuums nicht. Die Zustimmungswahrscheinlichkeit in Gleichung (2.4) hängt nur von Nutzdifferenzen ab. Daher können in einem parametrischen Modell nur solche Parameter identifiziert und somit geschätzt werden, die Unterschiede zwischen den Alternativen reflektieren (vgl. Train (2001) S. 24 ff. [242]). Dies bedeutet, dass sich alternativenspezifische Konstante und Parameter, die den Einfluss sozioökonomischer Variablen beschreiben, nicht als einzelne Größen schätzen lassen. Vielmehr lässt sich die Differenz der Konstanten und der genannten Parameter zwischen den jeweiligen Alternativen bestimmen. Dies ist gleichbedeutend mit einer Normierung der Ergebnisse auf - im Falle des einfachen Referendumsformats - entweder die Ausgangs- oder die Endsituation. Der Einfluss beispielsweise des Alters auf die Zahlungsbereitschaft bei Projektdurchführung wird relativ zum Einfluss des Alters auf den Nutzen in der Anfangssituation geschätzt. Ändert sich dieser Einfluss in den beiden Alternativen bei einem in den Parametern linearen Modell nicht, so kann die Variable nicht in das Modell aufgenommen werden. Diese Eigenschaft des linearen Modells wird weiter unten ausführlicher diskutiert.

Da das absolute Nutzenniveau irrelevant ist, muss zur Spezifizierung eines Modells eine weitere Normierung durchgeführt werden. Diese Skalierung lässt sich durch die Normierung der Varianz der ε_i^k erreichen (vgl. Train (2001) [242]), denn Multiplikation des Nutzens $v_i^k = \lambda\tilde{v}_i^k + \lambda\varepsilon_i^k$ führt zu

$\text{Var}(\lambda \varepsilon_i^k) = \lambda^2 \text{Var}(\varepsilon_i^k)$. Der zahlenmäßige Wert, auf den die Varianz festgelegt wird, ist nur dann von Bedeutung, wenn die Parameter der Einflussvariablen über verschiedene Modelle hinweg verglichen werden. Beim weiter unten beschriebenen Probit-Modell wird die Varianz typischerweise auf den Wert 1 normiert, in Logit-Modellen hat sich eine Normierung auf $\pi^2/6$ etabliert. Will man die geschätzten Parameter eines Logit- und eines Probit-Modells vergleichen, so muss man berücksichtigen, dass letztere um den Faktor $\sqrt{1,6}$ größer sein werden. Die Ordinalität des Nutzens hat folglich einen Einfluss auf die Modellierung sowohl der deterministischen als auch der stochastischen Komponente.

Modellierung der stochastischen Komponente

Überraschenderweise werden im größten Teil aller in der Praxis durchgeführten Umweltbewertungsstudien zur Modellierung der stochastischen Komponente symmetrische Verteilungen in Probit- und Logit-Modellen verwendet, Recherchen in den Datenbanken EVRI[©] [1] und Envalue [2] ergeben, dass das Logit-Modell von 60-80% der dort aufgeführten Studien angewendet wurde. Dies ist deshalb verblüffend, weil bei umweltverbessernden Projekten davon ausgegangen werden muss, dass die Nutzendifferenz, die aus der Umweltverbesserung bei konstanten Preisen folgt, positiv ist, und eine symmetrische Verteilung diese Tatsache nur unzureichend abbildet. Das Logit-Modell erhält man, wenn die ε_i^k als i.i.d. extremwertverteilt für alle k angenommen werden, die Verteilungsfunktion $F_\eta(\bullet)$ ist standard-logistisch mit

$$F_\eta(\Delta\tilde{v}) = (1 + e^{-\Delta\tilde{v}})^{-1}. \quad (2.5)$$

Das Probit-Modell folgt aus der Annahme, dass die ε^k gemeinsam normalverteilt sind

$$F_\eta(\Delta\tilde{v}) = \Phi(\Delta\tilde{v}). \quad (2.6)$$

Logit und Probit-Modell unterscheiden sich vor allem durch die Form der Verteilungen in den äußeren "Tails". Die Logit Verteilung misst diesen außen liegenden Bereichen mehr Gewicht bei.

Will man hingegen die Tatsache, dass $\tilde{v} \geq 0$ sein muss, auch explizit modellieren, so bieten sich die sogenannten Survivalmodelle an. Diese wurden ursprünglich, wie der Bezeichnung zu entnehmen ist, aus der Modellierung der "Überlebenszeiten" entweder bei biologischen und medizinischen Studien oder technischen Materialprüfungsuntersuchungen entwickelt (zur ausführlichen Darstellung von Intervalldaten und den geeigneten Modellansätzen

vgl. Kalbfleisch und Prentice (1980) [148]). Die Verteilungsfunktionen der in der Umweltbewertung am häufigsten verwendeten Spezifikationen sind

$$\begin{aligned} F_{\eta}(\Delta\tilde{v}) &= \Phi(\ln(\Delta\tilde{v})) && \text{für die Lognormale Verteilung,} \\ &= \exp(-e^{\Delta\tilde{v}}) && \text{für die Weibull-Verteilung,} \\ &= (1 + e^{-\ln(\Delta\tilde{v})})^{-1} && \text{für die Loglogistische Verteilung.} \end{aligned} \quad (2.7)$$

Survivalmodelle wurden mit dem Aufkommen des *doppelten* Referendumsformats überaus populär in der Contingent Valuation Literatur (vgl. Alberini (1995) [12], Cameron (1991) [53], Fernandez et al. (2001) [96] oder Zhou (2002) [255] um nur einige zu nennen). Wie in Kapitel 1.3.3 beschrieben, erhält man aus dem doppelten Referendumsformat Aussagen über endliche Intervalle, in denen die Zahlungsbereitschaften der Befragten liegen, wenn die Befragten einer Zahlung zugestimmt und einen höheren Zahlungsvorschlag abgelehnt haben. Zur Schätzung von Intervalldaten sind in verschiedenen Statistikpaketen (Limdep, Systat etc.) bereits Routinen implementiert, die auf Survivalmodellen basieren. So ist die Beliebtheit dieser Modelle auch aus rein pragmatischen und anwendungsorientierten Gründen nachvollziehbar. Aufgrund dieser Popularität werden in Kapitel 3 unter anderem auch Survivalmodelle simuliert.

Modellierung des deterministischen Teils

Für die funktionale Form des deterministischen Teils der Nutzenfunktion haben sich ebenfalls unterschiedlichen Modelle in der Literatur durchgesetzt. Das am häufigsten verwendete Modell ist linear im Einkommen.

$$\tilde{v}_i^k = \alpha_0^k + \beta^k \cdot y_i^k + \alpha_s^k s_i \quad (2.8)$$

wobei $(\alpha_0^k, \beta^k, \alpha_s^k)$ der Vektor der Parameter für den Zustand vor und nach dem Projekt bezeichnet ($k = 0, 1$). In den meisten Anwendungen dieses Modells wird darüber hinaus unterstellt, dass der Grenznutzen des Einkommens nicht vom jeweiligen Umweltzustand abhängt (vgl. Hanemann (1984) [119], McFadden und Leonard (1993) [190] oder Haab und McConnell (2002) [113]), so dass $\beta^0 = \beta^1 = \beta$ gelten muss. Da dem Entscheidungsmodell die Nutzenänderung durch das Projekt zugrunde liegt,

interessieren wir uns für die Differenz

$$\begin{aligned} \Delta \tilde{v}_i &= \alpha_0 - \beta \cdot \text{Bid}_i + \alpha \mathbf{s}_i \\ &\text{mit} \\ \alpha_0^1 - \alpha_0^0 &= \alpha_0, \\ \alpha_s^1 - \alpha_s^0 &= \alpha. \end{aligned} \quad (2.9)$$

Allerdings hat dieses Modell den Nachteil, dass es eine Einkommenselastizität der Zahlungsbereitschaft von 0 impliziert, da die Nutzendifferenz, und damit des Wohlfahrtsmaß nicht mehr vom Einkommen abhängt. Hahnemann und Kanninen (2001) [116] schlagen daher die Verwendung der flexibleren Box-Cox-Transformation des Einkommens vor, die das lineare Modell als Spezialfall für $\lambda = 1$ enthält

$$\tilde{v}_i^k = \alpha^k + \beta \left(\frac{y_i^\lambda - 1}{\lambda} \right). \quad (2.10)$$

Für $\lambda = 0$ enthält dieser Ansatz darüber hinaus ein Modell für die repräsentative Nutzenfunktion, das logarithmisch im Einkommen und linear in den sozioökonomischen Variablen ist.

Bishop und Heberlein (1979) [33], die als erste das geschlossene Format verwendeten, entwickelten zur Schätzung der Zahlungsbereitschaften ein Modell, bei dem die Wahrscheinlichkeit einer zustimmenden Antwort gegeben ist durch

$$Pr(yes_i) = (1 + e^{-\alpha + \beta \ln A_i})^{-1}, \quad (2.11)$$

das sich aus keinem der bislang beschriebenen Modelle für den deterministischen Teil herleiten lässt. Nachdem dieser Ansatz in der Literatur lange als nicht kompatibel mit einem Random Utility Modell diskutiert wurde, stellte sich heraus (vgl. Hanemann und Kanninen (2001) [116]), dass der Ansatz

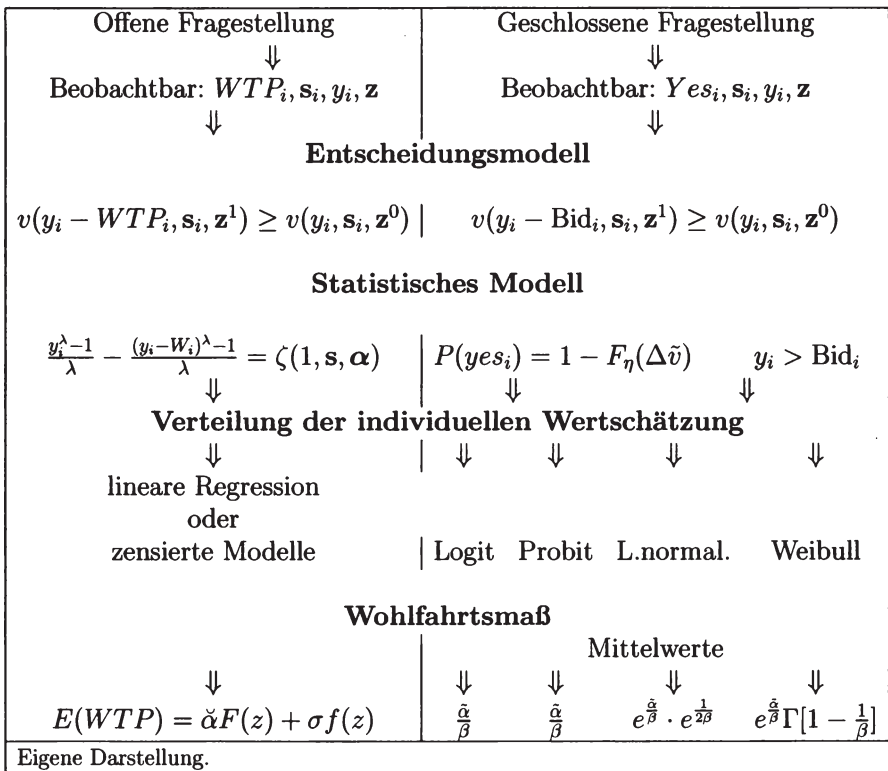
$$\begin{aligned} v_i^0 &= y_i^0 + \delta \\ v_i^1 &= y_i^1 + \delta + e^{\frac{\alpha_0 + \alpha_s \mathbf{s}_i + \epsilon_i}{\beta}} \end{aligned} \quad (2.12)$$

in Kombination mit einer logistischen Spezifizierung der stochastischen Komponente gerade zu den Ergebnissen von Bishop und Heberlein führt. Im Unterschied zu den weiter oben diskutierten Modellen, die additiv separabel in der deterministischen und der stochastischen Komponente sind, weist das Modell von Bishop und Heberlein eine multiplikative Struktur auf.

Wohlfahrtsmaße

Abbildung 2.1 fasst die Entscheidungssituation, die Verteilungen und die Wohlfahrtsmaße der gängigsten Modellen zusammen. Für jede Modellspezifizierung ergibt sich die Zahlungsbereitschaft dadurch, dass in (2.1) das Gleichheitszeichen gilt und nach A_i aufgelöst wird. Diese Zahlungsbereitschaft ist folglich wiederum eine Zufallsgröße und die Definition des Wohlfahrtsmaßes kann auf verschiedenen Wegen erfolgen (vgl. Hanemann (1984) [119]). Nimmt man den Mittelwert dieser Zufallsgröße, so bildet man

Abbildung 2.1: Parametrische Modellierung von CVM Daten



gewissermaßen den Erwartungswert des Beobachters bezüglich der maximalen Zahlungsbereitschaft der Individuen. Eine alternative Berechnung des Wohlfahrtsmaßes folgt daraus, dass man sich überlegt, bei welchem

Betrag gerade eine 50%ige Wahrscheinlichkeit besteht, dass das Individuum eine Zahlung leisten und dem Projekt zustimmen wird. Dieser Betrag A' erfüllt somit

$$Pr(v_i^1 \geq v_i^0) = 0,5. \quad (2.13)$$

Der sich hieraus ergebende Wert der Zahlungsbereitschaft ist der Median der Verteilung $F_\eta(\Delta\tilde{v})$ wegen

$$Pr(v_i^1 \geq v_i^0) = F_\eta(\Delta\tilde{v}). \quad (2.14)$$

Für die symmetrischen Verteilungen fallen diese beiden Maße zusammen, für das Logit- und das Probit-Modell ohne Kovariate ergibt sich die Zahlungsbereitschaft als negativer Quotient aus der Konstanten und dem Wert des Gebotsparameters. Für die nicht-symmetrischen Verteilungen jedoch ist der Median kleiner als der Mittelwert, so dass sich in der Literatur die Verwendung des Medians als sogenanntes konservatives Maß durchgesetzt hat (vgl. Mitchell und Carson (1989) [193]). Diese Empfehlung stammt insbesondere aus dem Einsatz der Contingent Valuation Methode zu Schadensbemessung, wird aber auch in der Kosten-Nutzen-Analyse umweltverbessernder Projekte verwendet.

2.2.1 Die Parameterschätzung

Zur Berechnung des mit dem jeweiligen Modell verbundenen Wohlfahrtsmaßes müssen nun die Parameter der gewählten Verteilung geschätzt werden. Hierzu bedient man sich meistens¹ der Maximum-Likelihood-Methode. Die Parameter eines Modells werden hierbei so bestimmt, dass sie den Abstand zu einem zwar unbekanntem aber als existent vorausgesetzten "wahren Modell" minimieren. Sei $f_0(\mathbf{x}|\mathbf{y})$ die "wahre" (durch den Index 0 angezeigt) bedingte Dichte eines Zufallsvektors \mathbf{x}_i bei gegebenen $\mathbf{y}_i = \mathbf{y}$. $\mathbb{X} \subset \mathbb{R}^K$ und $\mathbb{Y} \subset \mathbb{R}^K$ seien die Werte, die \mathbf{x} und \mathbf{y} annehmen können.

Den "Abstand" zwischen zwei Modellen bzw. den Abstand eines Modells zur "Realität" kann man mithilfe eines Maßes aus der Informationstheorie messen, dem Kullback-Leibler-Abstand (K-L-Abstand)

$$\mathcal{KL}(f; \mathbf{y}) \equiv \int_{\mathbb{X}} \ln \left(\frac{f_0(\mathbf{x}|\mathbf{y})}{f(\mathbf{x}|\mathbf{y})} \right) f_0(\mathbf{x}|\mathbf{y}) \, d\mathbf{x}. \quad (2.15)$$

¹Andere Möglichkeiten der Parameterschätzung umfassen die Minimum Chi-Quadrat Schätzung, verallgemeinerte lineare Modelle oder die Quasi-Maximum-Likelihood Schätzung (vgl. Hanemann und Kaninen (2001) [116]).

Dieser Abstand ist immer nicht-negativ, d.h. es gilt die KL-Informationsungleichung (vgl. für die folgende Argumentation Wooldridge (2002) [254], hier insbesondere Anhang 13A.)

$$\mathcal{KL}(f; \mathbf{y}) \geq 0. \quad (2.16)$$

Da dieser Ausdruck nur für $f = f_0$ zu 0 wird, muss er für alle Modelle, die nicht dem wahren Modell entsprechen, streng positiv sein.

Ein solches Modell, dessen Parameter θ noch zu bestimmen sind, lässt sich beschreiben durch die angenommene Verteilung, den Parametervektor und den Definitionsbereich des Parametervektors $\{f(\cdot|\mathbf{y}; \theta), \theta \in \Theta, \Theta \subset \mathbb{R}^p\}$. Die KL-Ungleichung (2.16) muss für dieses Modell für alle Werte von \mathbf{y} und θ gelten. Ist dieses Modell richtig spezifiziert, so muss es Werte θ_0 geben, für die gilt

$$f(\cdot|\mathbf{y}; \theta_0) = f_0(\cdot|\mathbf{y}) \quad \forall \mathbf{y} \in \mathbb{Y}, \quad (2.17)$$

so dass sich aus (2.16) ergibt

$$\begin{aligned} \int_{\mathbf{x}} \{\ln[f_0(\mathbf{x}|\mathbf{y})] - \ln[f(\mathbf{x}|\mathbf{y})]\} f_0(\mathbf{x}|\mathbf{y}) d\mathbf{x} &= E\{\ln[f_0(\mathbf{x}_i|\mathbf{y}_i)]|\mathbf{y}_i = \mathbf{y}\} \\ &- E\{\ln[f(\mathbf{x}_i|\mathbf{y}_i)]|\mathbf{y}_i = \mathbf{y}\} \geq 0. \end{aligned} \quad (2.18)$$

Für ein richtig spezifiziertes Modell folgt mit (2.17)

$$E\{\ln[f(\mathbf{x}_i|\mathbf{y}_i; \theta_0)]|\mathbf{y}_i\} \geq E\{\ln[f(\mathbf{x}_i|\mathbf{y}_i; \theta)]|\mathbf{y}_i\}. \quad (2.19)$$

Mit der Definition $\ell_i(\theta) \equiv \ln f(\mathbf{x}|\mathbf{y}_i; \theta)$ lässt sich schlussfolgern, dass ein Parametervektor θ_0 existiert, der den Erwartungswert $E(\ell_i(\theta))$ maximiert. Da $\ell_i(\theta)$ wiederum eine Zufallszahl ist, lässt sich der Erwartungswert als Stichprobenfunktion ausdrücken und das Maximierungsproblem lautet

$$\max_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n \ln f(\mathbf{x}_i|\mathbf{y}_i; \theta). \quad (2.20)$$

In diesem Problem wird $\ell_i(\theta)$ als die Log-Likelihoodfunktion bezeichnet; die Likelihoodfunktion ist $f(\mathbf{x}|\mathbf{y}_i; \theta)$. Übertragen auf die Schätzung des Parametervektors für die weiter oben beschriebenen Modelle lässt sich die Likelihoodfunktion schreiben als

$$L = \prod_{i=1}^N P(Yes_i = 1)^{Yes_i} \cdot (1 - P(Yes_i = 0))^{1 - Yes_i}. \quad (2.21)$$

Diese Funktion soll als Funktion der Parameter maximiert werden. Einsetzen der Zustimmungswahrscheinlichkeiten und Ablehnungswahrscheinlichkeiten führt zur Definition der Log-Likelihoodfunktion für das jeweils gewählte Modell

$$\begin{aligned} \ell(\boldsymbol{\theta}|\hat{\mathbf{x}}_i, \text{Bid}_i) &= \ln(L(\boldsymbol{\theta}|\hat{\mathbf{x}}_i, \text{Bid}_i)) \\ &= \sum_{i=1}^N \ln(1 - (F_\eta(\boldsymbol{\theta}, \hat{\mathbf{x}}_i, \text{Bid}_i))) \cdot \text{Yes}_i \\ &\quad + \ln(F_\eta(\boldsymbol{\theta}, \hat{\mathbf{x}}_i, \text{Bid}_i)) \cdot (1 - \text{Yes}_i). \end{aligned} \quad (2.22)$$

Die Bedingungen erster Ordnung lauten:

$$\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\hat{\mathbf{x}}_i, \text{Bid}_i) = 0, \quad (2.23)$$

wobei $\hat{\mathbf{x}}_i$ hier den Vektor aller beobachteten Größen inklusive der Konstanten darstellt (mit $\hat{x}_{0i} = 1$), die Abhängigkeit vom Zahlungsvorschlag jedoch explizit aufgenommen wurde, um die Darstellung der doppelt-gebundenen Fragestellung zu erleichtern; und $\boldsymbol{\theta}$ der Vektor aller zu schätzender Parameter ist. Die Parameterschätzung für Daten aus einer doppelt-gebundenen Fragestellung verläuft analog mittels der Log-Likelihoodfunktion

$$\begin{aligned} \ell &= \sum_{i=1}^N [\ln(P((\text{YesYes})_i = 1)) \cdot (\text{YesYes})_i + \\ &\quad \ln(P((\text{YesNo})_i = 1)) \cdot (\text{YesNo})_i + \\ &\quad \ln(P((\text{NoYes})_i = 1)) \cdot (\text{NoYes})_i + \\ &\quad \ln(P((\text{NoNo})_i = 1)) \cdot (\text{NoNo})_i]. \end{aligned} \quad (2.24)$$

Für das Probit-Modell unter Annahme eines normalverteilten stochastischen Terms ergibt sich zum Beispiel

$$\begin{aligned} Pr(\text{NoNo}) &= \Phi(\text{Bid}_{low}) \\ Pr(\text{NoYes}) &= \Phi(\text{Bid}_1) - \Phi(\text{Bid}_{low}) \\ Pr(\text{YesNo}) &= \Phi(\text{Bid}_{up}) - \Phi(\text{Bid}_1) \\ Pr(\text{YesYes}) &= 1 - \Phi(\text{Bid}_{up}). \end{aligned} \quad (2.25)$$

Je nach Zahlungsbereitschaft antworten die Befragten mit einer von vier Antwortmöglichkeiten. Wenn ihre Zahlungsbereitschaft das höchste Gebot überschreitet, antworten sie beispielsweise mit "YesYes". Obwohl die Gebotshöhe der zweiten Frage von der Antwort auf die erste Frage abhängt, darf die Antwort auf die zweite Frage nicht von der ersten Antwort abhängen, wenn man beiden Fragen dieselbe zugrunde liegende Zahlungsbereitschaft unterstellt. Die in 1.3.3 beschriebenen Widersprüche in Daten, die

mittels der doppelten Referendumsmethode erhoben wurden, gaben Anlass, diese Annahme zu überprüfen und Modelle zu entwickeln, in denen der Antwort auf die zweite Frage eine andere Zahlungsbereitschaft zugrunde liegen darf als der ersten Antwort. Das bivariate Probit-Modell (vgl. Cameron und Quiggin (1994) [55] oder Haab (1997) [109] und (1998) [110]) ist die allgemeine Form einer solchen Modellierung, wenn man für den stochastischen Term einen normalverteilten Fehler unterstellt. Die Zustimmungswahrscheinlichkeiten lassen sich in diesem Modell ausdrücken als

$$\begin{aligned} Pr(NoNo) &= \Phi(\text{Bid}_1, \text{Bid}_{low}, \rho) \\ Pr(NoYes) &= \Phi(\text{Bid}_1) - \Phi(\text{Bid}_1, \text{Bid}_{low}, \rho) \\ Pr(YesNo) &= \Phi(\text{Bid}_{up}) - \Phi(\text{Bid}_1, \text{Bid}_{up}, \rho) \\ Pr(YesYes) &= 1 - \Phi(\text{Bid}_1) - \Phi(\text{Bid}_{up}) + \Phi(\text{Bid}_1, \text{Bid}_{low}, \rho). \end{aligned} \quad (2.26)$$

$\Phi(\text{Bid}_K)$ ist die univariate Normalverteilung, deren Argument von Bid_K abhängt (neben den sonstigen Kovariaten); $\Phi(\text{Bid}_K, \text{Bid}_M, \rho)$ ist die bivariate Normalverteilung, die darüber hinaus vom Gebot auf die zweite Frage abhängt und von einem Kopplungsparameter ρ , der den Zusammenhang zwischen der Antwort auf die erste und die zweite Frage nach der Zahlungsbereitschaft misst. Ist $\rho = 1$ so erhält man die ursprüngliche Vermutung zurück, dass den beiden Fragen dieselbe Zahlungsbereitschaft zugrunde liegt. Es ist leicht zu sehen, dass sich Gleichung (2.25) als restringierte Fassung des Ansatzes in Gleichung (2.26) gewinnen lässt, indem an die Stelle der bivariaten Terme jeweils ein univariater Term eingesetzt wird, der das zuletzt genannte Gebot als Argument hat. Anders ausgedrückt ist die Verwendung von Gleichung (2.25) synonym zur Schätzung eines bivariaten Modells, bei dem die Parameter der beiden unterstellten Zahlungsbereitschaften identisch gewählt werden. Im bivariaten Modell, bei dem alle Parameter "frei" bestimmt werden, erhält man als Ergebnis schlimmstenfalls zwei sehr unterschiedliche Zahlungsbereitschaften. Bislang wurde kein Modell entwickelt, das hilft zu entscheiden, welcher Mechanismus bei den Befragten gewirkt hat, d.h. welche die "gültige" Zahlungsbereitschaft für das vorgeschlagene Projekt ist. Darüber hinaus gibt man den Vorteil der Verringerung der Konfidenzintervalle auf, der das eigentliche Ziel der Entwicklung des doppelten Referendummodells ist. Alberini (1995) [12] konnte zeigen, dass insgesamt die Vorteile des restringierten Modells überwiegen und dass bei den von ihr getesteten Studien die Verbesserung der Schätzung mit dem bivariaten Modell die Nachteile der weniger engen Konfidenzintervalle nicht aufwiegt.

Das Ziel der Analyse schließlich ist die Berechnung des Erwartungswerts der Zahlungsbereitschaft in der Stichprobe der betroffenen Bevölkerung bzw. die Bestimmung der Wertschätzung eines umweltverändernden Projekts durch die Bevölkerung mit einer Genauigkeit, auf der sich eine politische Entscheidung basieren lässt. Während ein erster Plausibilitätstest des gewählten Modells darin besteht, dass man den Einfluss verschiedener Kovariate auf die Zielgröße untersucht, lässt sich dieses Ziel bei einer Fragestellung im Referendumsformat erreichen, indem nur die Gebotshöhe als erklärende Variable im Modell auftritt, und eine Konstante zugelassen wird. Wie auch McFadden und Leonard (1993) [190] (S.178) bemerken, "(...) it is unnecessary to introduce covariates in order to estimate mean WTP, because the *unconditional* average of WTP over a representative sample takes the effects of covariates into account and is a consistent estimate of mean WTP of the population." Dies ist die ökonometrische Entsprechung der Argumentation auf Seite 51, bei der schon darauf hingewiesen wurde, dass bei der Entscheidung des Konsumenten zwischen einer Situation ohne das Projekt und ohne Zahlung oder für das Projekt bei einer bestimmten Zahlung die *Motivation* des Befragten keine Rolle spielen darf.

2.2.2 Zur Modellwahl

Für den Politiker, der eine Entscheidung auf den Ergebnissen einer CVM Studie basieren will, ist nun vor allem die Frage wichtig, welches Modell zu wählen ist und warum. Zur Modellwahl existieren eine Reihe von Gütekriterien², die im Wesentlichen alle von derselben Grundidee ausgehen. Um aus verschiedenen möglichen und mit der ökonomischen Theorie vereinbaren Modellen das geeignetste auszuwählen, muss die Abweichung der Modelle von einem "wahren Modell" gemessen werden. Dieses wahre Modell wird entweder implizit angenommen, wie bei der Idee des Akaike Information Criterion (AIC) (vgl. Burnham und Anderson (1998) [51] S. 66 ff.), oder die Daten werden als wahr angenommen und das Modell, das die Daten am besten annähert, wird als bestes Modell ausgewählt, wie beim Devianzkriterium. Hierbei schwingt implizit mit, dass die gewählte Messmethode der Fragestellung auch am besten entspricht - hierzu kann das statistische Modell, das auf den Daten beruht, keine Aussage treffen.

²Bei der offenen Fragestellung, die in eine lineare Regression zur Berechnung der erwarteten Zahlungsbereitschaft mündet, bedient man sich der Standardtechniken und berechnet die R^2 des jeweiligen Modells.

Das AIC lässt sich wiederum aus dem Kullback-Leibler-Abstand herleiten. Während die Herleitung der Likelihoodfunktion aus diesem Maß von einem gewählten Modell und zu bestimmenden Parametern ausging, sei nun der Parametervektor bereits für verschiedene Modelle bestimmt und es soll das beste Modell gefunden werden. Angenommen, es existiere ein "wahres Modell" beschrieben durch f und eine Reihe von verschiedenen Modellen $g(\mathbf{x}|\theta)$ zur Beschreibung dieses wahren Modells. Der K-L-Abstand zwischen f und g ist dann gegeben durch (für eine ausführliche und sehr intuitive Herleitung vgl. Burnham und Anderson (1998) [51])

$$I(f, g) = \int f(\mathbf{x}) \ln \left(\frac{f(\mathbf{x})}{g(\mathbf{x}|\theta)} \right) d\mathbf{x}. \quad (2.27)$$

$I(f, g)$ beschreibt die Information, die verloren geht, wenn das wahre Modell f durch das Modell g angenähert wird. Zur genauen Berechnung von $I(f, g)$ müssten die Parameter von f und g bekannt sein, was in der Realität nicht gegeben ist. Durch Umformung kann man ein relatives Abstandsmaß definieren, das den Anspruch, das beste Modell zu kennzeichnen, genauso erfüllt

$$\begin{aligned} I(f, g) &= \int f(\mathbf{x}) \ln(f(\mathbf{x})) d\mathbf{x} - \int f(\mathbf{x}) \ln(g(\mathbf{x}|\theta)) d\mathbf{x} \\ &= E_f[\ln(f(\mathbf{x}))] - E_f[\ln(g(\mathbf{x}|\theta))] \\ &= C - E_f[\ln(g(\mathbf{x}|\theta))]. \end{aligned} \quad (2.28)$$

Bei der Modellwahl müssen nun die unbekannt Parameter θ durch ihre Schätzer ersetzt werden, d.h. es lässt sich nur ein Wert $\hat{I}(f, g)$ tatsächlich zur Modellwahl heranziehen. Der Erwartungswert dieser Größe lässt sich ausdrücken durch

$$\begin{aligned} E_{\hat{\theta}}[\hat{I}(f, g)] &= \int f(\mathbf{y}) \left[\int f(\mathbf{x}) \ln(f(\mathbf{x})) d\mathbf{x} \right] d\mathbf{y} \\ &\quad - \int f(\mathbf{y}) \left[\int f(\mathbf{x}) \ln(g(\mathbf{x}|\hat{\theta}(\mathbf{y}))) d\mathbf{x} \right] d\mathbf{y} \\ &= \hat{C} - E_y E_x[\ln(g(\mathbf{x}|\hat{\theta}))]. \end{aligned} \quad (2.29)$$

Akaike (1973) [11] konnte zeigen, dass die zentrale Größe, die es infolge der Ersetzung durch die Schätzer zu berechnen gilt, die Schätzung von $E_y E_x[\ln(g(\mathbf{x}|\hat{\theta}))]$ ist. Diesen Ausdruck kann man durch das Maximum der Likelihoodfunktion des jeweiligen Modells ersetzen, einen Wert, der bei

der Parameterschätzung automatisch berechnet wird. Als ein erstes Kriterium für den Modellvergleich wird der Wert der Likelihoodfunktion häufig herangezogen. Das Modell mit dem höchsten Wert ist das beste Modell. Allerdings lässt sich zeigen, dass das Maximum der Likelihoodfunktion zu einer Überschätzung führt, wenn bei einem Modell sehr viele Parameter einbezogen wurden und es mit einem anderen Modell verglichen werden soll, das deutlich weniger Parameter enthält. Akaike schlägt zur Korrektur dieser Überschätzung vor, die Anzahl der im Modell einbezogenen Parameter K abzuziehen. So ergibt sich das in der Literatur bekannte Akaike Information Criterion (AIC) zu

$$AIC = -2\ln(L(\hat{\theta}|\mathbf{y})) + 2K. \quad (2.30)$$

Das Modell mit dem kleinsten Wert für AIC ist nach diesem Kriterium das Beste. Diese von Akaike vorgeschlagene Korrektur geht von einer festen und typischerweise niedrigen Anzahl der einbezogenen Parameter aus (vgl. Burnham und Anderson (1998) [51]). Es lässt sich fragen, ob die Stichprobengröße n ebenfalls einen Einfluss auf die Korrektur haben sollte. Aus dieser Fragestellungen sind eine Reihe von Informationskriterien hervorgegangen, die den Korrekturterm auf andere Weise als von Akaike vorgeschlagen modellieren. Das bekannteste ist das BIC (Bayesian Information Criterion) das auch unter dem Begriff Schwartz'sches Informationskriterium (SIC) zu finden ist, da es von Schwartz (1978) [226] vorgeschlagen wurde

$$BIC = -2\ln(L(\hat{\theta}|\mathbf{y})) + K\ln(n), \quad (2.31)$$

bei dem die Abhängigkeit des Korrekturterms von der Stichprobengröße n durch $\ln(n)$ explizit berücksichtigt wird. Eine Übersicht über weitere Modifizierungen und ihre Verwendungsgebiete findet sich bei Burnham und Anderson (1998) [51].

Wählt man das weniger ehrgeizige Ziel, dass ein Modell die beobachteten Daten am besten repräsentiert, lässt sich der Abstand zwischen Modell und Daten messen durch die Devianz (vgl. Cox und Snell (1999) [74])

$$\begin{aligned} D &= -2(\ell(\hat{\theta}|\mathbf{y}) - \ell_{freq}) \\ &= -2\sum_{j=1}^k \ln(L_j(\hat{\theta}|\mathbf{y})) - \sum_{j=1}^k yes_j \cdot \ln(\pi_j) + no_j \cdot \ln(1 - \pi_j), \end{aligned} \quad (2.32)$$

wobei über die einzelnen Gebotsklassen summiert wird, d.h. π_j gibt die Häufigkeit der positiven Antworten auf das j -te Gebot an, yes_j bzw. no_j die

Anzahl der Zustimmungen bzw. Ablehnungen der jeweiligen Gebote. Wenn das Modell richtig spezifiziert ist, ist die Devianz χ^2 -verteilt mit $(k - p)$ Freiheitsgraden, wobei k die Anzahl der Gebotsklassen beschreibt und p die Anzahl der in der Schätzung verwendeten Parameter. Auch hier gilt, dass das Modell mit der kleinsten Devianz das Beste ist, darüber hinaus lassen sich aus den kritischen Werte der χ^2 -Verteilung Aussagen über die Güte des einzelnen Modells treffen (vgl. Alberini et al. (1997) [13]). Bei näherem Hinsehen wird die Verwandtschaft zwischen dem Devianzkriterium und dem Ausdruck für AIC deutlich.

Abgesehen von einer Hilfe bei der Wahl des besten Modells will der Entscheidungsträger aber auch eine Aussage über die Zuverlässigkeit des ermittelten Erwartungswerts der Zahlungsbereitschaft. Ein statistisches Maß für die Zuverlässigkeit eines geschätzten Parameters ist das Konfidenzintervall. Es gibt den Wertebereich an, in dem man den interessierenden Parameter in der Grundgesamtheit mit einer festgelegten Wahrscheinlichkeit erwartet. Genau genommen gibt das Konfidenzintervall an, in welchem Bereich der geschätzte Parameter mit $(1 - \alpha)\%$ iger Sicherheit läge, wenn man unendlich viele Stichproben ziehen könnte und aus diesen unendlichen vielen Stichproben jeweils den Parameter schätzen könnte. Bei einer linearen Regression - der Lösung zur Parameterschätzung einer offenen Befragung - ergibt sich unter der Annahme normalverteilter Störterme das Intervall, in welchem der geschätzte Parameter mit 95%iger Wahrscheinlichkeit liegt als Lösung der Gleichung $Prob(b_k - t_{0,95}se_{b_k} \leq \beta_k \leq b_k + t_{0,95}se_{b_k}) = 0,95$. Zur Bestimmung des Konfidenzintervalls sind der Standardfehler se_{b_k} des geschätzten Parameters und der Wert $t_{0,95}$ der t -Verteilung mit $(n - k)$ Freiheitsgraden, der mit 95%iger Wahrscheinlichkeit überschritten wird, notwendig. In den meisten Statistikpaketen ist eine Routine zur Bestimmung von Konfidenzintervallen für lineare Regressionen implementiert.

Die Zielgröße ist bei Referendumsdaten jedoch als Quotient aus zwei Parametern - Zufallsgrößen - definiert. Um Konfidenzintervalle nach dem obigen Muster zu bestimmen, müssten sich Aussagen treffen lassen über die Verteilung dieses Quotienten, was je nach Modell schwierig werden kann³. In der Umweltbewertung werden zur Bestimmung von Konfidenzintervallen Verfahren herangezogen, die auf die ursprüngliche Interpretation des Konfidenzintervalls zurückgreifen und eine Vielzahl von Stichproben aus den vorhandenen Daten simulieren. Aus den simulierten Stichproben werden

³Für eine asymptotische Bestimmung von Konfidenzintervallen für Funktionen von Parametern nach der Delta-Methode vgl. Wooldridge (2002) [254] S. 43 ff.

dann die Parameter geschätzt und der Bereich gebildet, in dem der Schätzwert mit $(1 - \alpha)\%$ iger Wahrscheinlichkeit liegt. Cooper (1994) [71] stellt die wichtigsten Methoden zur Berechnung von simulierten Konfidenzintervallen zusammen. Die wesentliche Idee bei der Berechnung von Konfidenzintervallen mithilfe von Simulationsverfahren ist die Erzeugung einer Verteilung von Erwartungswerten der Zahlungsbereitschaft mit großem Umfang, indem nach dem Bootstrap-Verfahren verschiedene Stichproben (mit Zurücklegen) aus den vorhandenen Ergebnissen gezogen werden. Aus einer solchen künstlich erzeugten Verteilung lassen sich dann durch Sortieren der Werte nach der Größe die 2,5 oder 5 Prozent der niedrigsten und höchsten Werte ermitteln (Efron'sche Perzentilmethode). Diese unteren und oberen Ausläufer der Verteilung werden ignoriert und die Werte dazwischen stehen für das Konfidenzintervall. Zur Erzeugung der Verteilung von Zahlungsbereitschaften aus Bootstrap-Stichproben sind zwei Verfahren aus der Literatur bekannt. Park, Loomis und Creel, (1991) [200] übertragen ein von Krinsky und Robb (1989) entwickeltes Verfahren auf die Berechnung von Konfidenzintervallen bei Referendumsdaten. Hierbei wird unterstellt, dass die aus der Maximum-Likelihood-Schätzung resultierenden Parameterschätzer einer multivariaten Normalverteilung unterliegen, die als Mittelwert den geschätzten Parametervektor hat und als Kovarianzmatrix die geschätzte Kovarianzmatrix. Aus dieser Verteilung werden nun Werte gezogen, aus denen jeweils die Zahlungsbereitschaft neu berechnet wird. So entsteht eine Verteilung der Zahlungsbereitschaften. Alternativ hierzu lassen sich Bootstrap-Stichproben aus den Daten ziehen (vgl. Cooper (1994) [71]) und sich aus jeder Bootstrap-Stichprobe in einem der oben beschriebenen Modellansätze die Parameter neu schätzen und die Zahlungsbereitschaft berechnen. Dieser Ansatz kommt ohne zusätzliche Verteilungsannahmen aus und wird in Kapitel 3 verwendet werden.

2.3 Weitere Ansätze aus dem Bereich der klassischen Statistik

Es ist ein weiter Weg von der Definition der Kompensierenden Variation als Abstand zwischen zwei individuellen Nutzenfunktionen, über die Interpretation dieses Abstands als Zahlungsbereitschaft bei einem nutzerhöhenden Projekt hin zur Berechnung der Zahlungsbereitschaft als Erwartungswert einer Zufallsvariablen mit einem statistischen Modell aus einer Stichprobe von begrenztem Umfang.

Die Vielzahl der Anwendungen auf unterschiedliche Umweltgüter mit ihren jeweiligen Eigenschaften spiegelt sich in der Vielfalt ökonometrischer Modellierungen wider (vgl. Stumborg et al. (2001) [237] für die Modellierung von zeitabhängigen Erwartungswerten, oder Lewbel and McFadden (2002) [173] für eine Analyse der Eigenschaften verschiedener Verteilungsfunktionen und ihrer Einsatzmöglichkeiten). Creel (1998) [76] geht der Frage nach, wie man mit fehlspezifizierten Einflussvariablen umgehen kann, Geweke und Keane (1997) [104] experimentieren mit flexibleren funktionalen Formen, denn die Annahme der Normalverteilung in den häufig verwendeten Probit-Modellen erscheint häufig als zu vereinfachend. Die Autoren schlagen statt dessen einen "Mixture" Ansatz vor, bei dem die Verteilung aus einer Überlagerung verschiedener Normalverteilungen gebildet wird. Mixed *Logit*-Ansätze sind in der Literatur zur Modellierung von Choice Daten inzwischen verbreitet (vgl. Train (2001) [242] S. 138 ff., und die dort zitierte Literatur) und werden mit dem Aufkommen von Simulationstechniken zur Berechnung von Integralen, auf die im Zusammenhang mit der Bayes'schen Sichtweise in dieser Arbeit ebenfalls eingegangen wird, zunehmend verwendet. In der Umweltbewertungsliteratur findet sich dieser Ansatz bislang selten.

Eine Besonderheit der Daten aus einer CVM-Studie hat die Forschung stark beschäftigt (vgl. Ready und Hu (1995) [209], Reiser und Shechter (1999) [212], Werner (1999) [250] oder Haab (1999) [111]); man findet häufig eine hohe Zahl von Zahlungsablehnungen. Um dies modellieren zu können, wurden Modelle mit einem sogenannten Spike bei $WTP = 0$ entwickelt, die die befragte Bevölkerung in zwei Gruppen einteilen; eine Gruppe, die keine oder eine sehr geringe Zahlungsbereitschaft hat und eine zweite Gruppe, die dem Projekt einen positiven Wert beimisst⁴. Auf die Ursachen derartiger Antworten und mögliche Modellierungen gehen Jorgensen und Syme (1999) [144] sowie Messonnier und Bergstrom (2000) [191] ein.

Darüber hinaus wird in der Literatur zunehmend der Frage nachgegangen, ob die parametrische Schätzung mit ihren festen Verteilungsannahmen geeignet ist, Verteilungen von Zahlungsbereitschaften zu modellieren (vgl. Li (1996) [174], Boman et al. (1999) [38], An (2000) [14] und Crooker und Kling (2000) [79]). Die ökonomische Theorie trifft keinerlei Aussagen über die *Verteilung von Präferenzen* in der Bevölkerung, und gibt somit auch keinerlei Hinweise, wie diese Verteilungen zu wählen sind. Die Entscheidung

⁴Zur Diskussion von Projekten, die sowohl positive als auch negative Zahlungsbereitschaften hervorufen, vgl. Berrens et al. (1998) [32].

für eine bestimmte Verteilungsfunktion wird entweder aus pragmatischen Erwägungen getroffen, oder sie folgt vorab vorliegenden Erfahrungen und Kenntnissen über die Bevölkerung, die außerhalb der eigentlichen Studie liegen. Will man sich diesen Einschränkungen durch die Wahl einer festen Form der Verteilung entziehen, so kann man zu sogenannten nichtparametrischen oder semiparametrischen Verfahren übergehen. Erstere kommen völlig ohne Verteilungsannahmen aus, letztere modellieren den deterministischen Teil der Zahlungsbereitschaft mit aus der ökonomischen Theorie ableitbaren funktionalen (parametrischen) Zusammenhängen und den stochastischen Teil ohne feste Verteilungsannahme, indem sie zum Beispiel eine der gängigen aus der Mathematik bekannten Reihenzerlegungen und Näherungsverfahren (Fourierreihen, Besselfunktionen, Hermite Polynome etc.) wählen. Kriström (1990) [158] schlägt eine elegante und einfach zu berechnende Methode für die nichtparametrische Modellierung vor, bei welcher die Häufigkeiten der jeweiligen Ja-Antworten als (nichtparametrischer) Schätzer für die Wahrscheinlichkeit einer Ja-Antwort eingesetzt wird. Trägt man diese Wahrscheinlichkeiten in eine Grafik ein, so erhält man - vorausgesetzt die Wahrscheinlichkeiten sind monoton nicht-steigend - eine nichtparametrische Verteilung der Zahlungsbereitschaften, aus der sich der Mittelwert oder der Median leicht ablesen lässt. Falls die Monotonieannahme verletzt ist, greift Kriström auf das Ayer-Verfahren (vgl. Ayer et al. (1955) [22]) zurück, mit dessen Hilfe die Verteilung entsprechend geglättet wird (für eine Anwendung dieses Verfahrens siehe Kapitel 3). Der Nachteil dieses Ansatzes besteht darin, dass es nicht möglich ist, Kovariate einzubeziehen, und somit auch die Möglichkeit von Plausibilitätstests durch zusätzliche sozioökonomische Variablen entfällt. Darüber hinaus ist die Verwendung des Verfahrens bei Daten aus einem doppelten Referendum nur dann möglich, wenn die Gebotsintervalle sich nicht überlappen.

Andere Ansätze für eine Modellierung völlig ohne Verteilungsannahmen basieren auf Funktionsklassen, die die ökonomischen Monotonie- und Homogenitätseigenschaften besitzen und maximieren die Likelihoodfunktion direkt über diese Klassen (vgl. Manski (1987) [183] oder Matzkin (1991) [185]). Allerdings sind zur Berechnung der Likelihoodfunktion teilweise lineare Interpolationen notwendig, die nur bei einer großen Zahl von Beobachtungen sinnvoll erscheinen. Manski (1975) [181] und (1985) [182] schlägt den Maximum Score Schätzer vor, bei dem der Parametervektor so gewählt wird, dass die Anzahl der korrekten Vorhersagen an Ja-Antworten maximiert wird. Obwohl dieser Schätzer konsistent ist, sind seine asymptotischen Eigenschaften unhandlich (vgl. Horowitz und Neumann (1987) [130])

und er konvergiert sehr langsam mit der dritten Wurzel aus der Stichprobengröße. Daher schlug Horowitz (1992) [129] (vgl. auch Horowitz und Savin (2001) [131]) einen geglätteten Maximum Score Schätzer vor, dessen asymptotische Eigenschaften erheblich besser sind. An die Stelle der Stufenfunktion im Maximum Score Schätzer wird eine glattere Funktion gesetzt, die zweimal stetig differenzierbar ist. Im Resultat ähnelt dieses Vorgehen einer anderen semiparametrischen Strategie, bei der die Verteilungsfunktion durch eine Summe von glatten Funktionen angenähert wird, den sogenannten Kernschätzern. Die wesentliche Idee eines Kernschätzers besteht darin, Wahrscheinlichkeitsverteilung durch Funktionen, sogenannte "Kerne" anzunähern, deren funktionale Form jedoch wiederum festgelegt werden muss (vgl. Chay und Powell (2001) [64], Ichimura und Linton (2001) [141] sowie Lewbel und Linton (2000) [172] oder für frühere Anwendungen Lee (1992) [165] und Creel und Loomis (1997) [77]). Creel (1997a) [75] kann zeigen, dass diese Verfahren dem einfachen Logit-Modell in einer Monte Carlo Simulationen überlegen sind. Crooker und Herriges (2004) [78] vergleichen zwei parametrische Ansätze, drei nichtparametrische und zwei semiparametrische ebenfalls unter Verwendung von simulierten Daten. Ihre Ergebnisse sind so nachvollziehbar wie niederschmetternd für die alternativen Ansätze. Eben gerade die scheinbare Stärke der Ansätze, der Verzicht auf eine Verteilungsannahme, lässt sie bei kleinen Stichproben oder einer geringen Anzahl von Geboten deutlich schlechter abschneiden als das herkömmliche und überaus etablierte Probit-Modell. Die Schätzung von Modellen mit wenigen Beobachtungen muss eine Schwachstelle der nichtparametrischen Verfahren sein, denn sie haben keine Verteilungsannahme. Wo das Probit-Modell leichter interpolieren kann, da es auf einer vorgegebenen funktionalen Form beruht, ist das nichtparametrische Modell allein auf die Daten angewiesen. Ähnlich lässt sich bei einer zu geringen Anzahl von Geboten argumentieren. Nichtparametrische Modelle haben bei wenigen Geboten viel weniger "Stützstellen" aus denen sie ihre Information beziehen. Obwohl Crooker und Herriges Verteilungen der Zahlungsbereitschaft modellieren, die weit von der Normalverteilung abweichen, weist das Probit-Modell im Vergleich überwiegend den niedrigsten Fehler auf. Aus der Sicht des Praktikers ist es das Hauptanliegen, eine widerstandsfähige Schätzung mit einem engen Vertrauensintervall aus einer Contingent Valuation Studie zu erhalten. Auch wenn die nichtparametrischen Verfahren ökonometrisch sehr interessant sind, wird ihnen gerade ihre inhärente fehlende Zuverlässigkeit bei kleinen Stichproben zum Verhängnis.

Der Ausgangspunkt der geschilderten Weiterentwicklungen des ursprüng-

lichen klassischen Modells entstammt der Anwendung auf spezielle in der Praxis vorgefundene Charakteristika der Daten. Es ist häufig der Eindruck, dass man etwas in einem Modell gefunden hat, "was so nicht sein kann", weil es entweder der ökonomischen Fundierung oder einfach nur der ökonomischen Intuition widerspricht. Diese Intuition, bzw. dieses Wissen hat jedoch in der klassischen Modellierung keinen Platz. Somit liegt es auf der Hand, nach einem Modellrahmen zu suchen, in welchem für eben diese vorab vorliegenden Kenntnisse die Möglichkeit einer systematischen Einbindung besteht. Ein möglicher Ansatz besteht im Wechsel zur Bayes'schen Sichtweise, die im nächsten Abschnitt vorgestellt wird.

2.4 Ein alternativer Ansatz: Das Bayes'sche Modell

Vorab vorliegende Kenntnisse über eine statistische Größe sind das Herzstück des Bayes'schen Ansatzes. Ergebnisse eines ökonomischen oder naturwissenschaftlichen Experiments werden dabei nicht isoliert betrachtet, wie bei den klassischen Verfahren, sondern gelten immer als neue Information, mit deren Hilfe bereits bestehende Erkenntnisse aus ähnlichen Experimenten aufdatiert werden. Leonhardt (2001) [170] hat dies in einem Artikel in der *New York Times* unter dem Titel "Adding Art to the Rigor of Statistical Science" sehr anschaulich ausgedrückt. Er schreibt: "If a study, even a statistical significant one, suggests that pigs can fly, Bayes's theorem allows researchers to combine the study's results mathematically with hundreds of years of knowledge about the travel habits of swine." Was sich aufgrund der alltäglichen Erfahrung nach einer sinnvollen Vorgehensweise anhört, war lange Zeit die Ursache heftiger Auseinandersetzungen, akademischer Debatten und intensiver Diskussionen, die das jeweilige andere Lager der Engstirnigkeit oder schlimmerer Verbrechen anklagten. Erst in jüngerer Zeit werden Lehrbücher verfasst, die Ansätze der klassischen und der Bayes'schen Statistik für unterschiedliche Probleme beinhalten. Der wesentliche philosophische Unterschied der beiden Ansätze besteht darin, dass die klassischen (parametrischen) Verfahren die Daten als eine Realisierung einer Verteilung mit festen Parametern betrachten, im Bayes'schen Modell hingegen sind die Parameter Zufallsvariablen und die Daten werden als jeweils neue Information angesehen, die dabei hilft, die Verteilung dieser Zufallsvariablen besser beschreiben zu können. Streng genommen

kann die klassische Statistik somit bei einer unendlichen Wiederholung des Experiments die Schätzung des allen Experimenten zugrunde liegenden "wahren" Parameters verbessern. So ist auch - streng genommen - die Definition eines Konfidenzintervalls zu verstehen: Bei einer unendlich häufigen Wiederholung des Experiments, wäre der wahre Wert des Parameters in 95% aller Schätzungen eben dieses Konfidenzintervalls enthalten. Diese wenig intuitive Interpretation des Konfidenzintervalls ist auch der Hauptvorwurf der "Bayesianer" gegenüber der klassischen Statistik. Ein Bayes'sches Vertrauensintervall hingegen ist ganz intuitiv formuliert, es ist die 95%ige Abdeckung der Verteilung der Parameter. In der Ausprägung der Vertrauensintervalle liegt einer der Vorteile des Bayes'schen Ansatzes, denn mit der Möglichkeit, vorhandene Kenntnisse systematisch einzubauen, können sich Bayes-Modelle gewissermaßen Information bei anderen Experimenten "ausleihen" - und so die Zuverlässigkeit der Schätzung verbessern, d.h. ein weniger weites Vertrauensintervall angeben. Gerade in Situationen, in denen der Stichprobenumfang durch finanzielle oder andere Beschränkungen begrenzt ist, lohnt sich daher ein Blick auf diese alternative Modellierung.

Lange Zeit waren die Einsatzmöglichkeiten durch die Komplexität der resultierenden Modelle - man muss nicht nur die Likelihoodfunktion spezifizieren, sondern auch noch die Vorabkenntnisse - begrenzt. Neuere Simulationstechniken, wie etwa die Monte Carlo Markov Ketten Simulation und Algorithmen wie der Gibbs Sampler haben jedoch diese Lücke geschlossen. Im Folgenden werden diese Techniken kurz beschrieben, da sie bislang noch nicht häufig in der Contingent Valuation Literatur zu finden sind.

2.4.1 Bayes'sches Modell für Contingent Valuation Daten

Ausgangspunkt der Bayes'schen Sichtweise ist der Satz von der totalen Wahrscheinlichkeit, nach dem sich ein sicheres Ereignis E zerlegen lässt in die Summe von paarweise unvereinbaren Ereignissen, die E ausschöpfen, $E = A_1 \cup A_2 \cup \dots \cup A_n$ mit $A_j \cup A_k = \emptyset$ für $j \neq k$. Für ein beliebiges Ereignis B gilt $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_n \cap B)$. Nach dem Additionsaxiom für Wahrscheinlichkeiten und unter Verwendung der Multiplikationsregel folgt (vgl. Bronstein und Semendjajew (1981) [44])

$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i). \quad (2.33)$$

Fragt man nun nach der A-Posteriori-Wahrscheinlichkeit, dass A_i eintritt, wenn B eingetreten ist, so erhält man die Formel von Bayes

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_j)P(B|A_j)}. \quad (2.34)$$

Dieser Zusammenhang lässt sich zur Schätzung von Parametern in einem statistischen Modell nutzbar machen, indem man die Parameter als ein Ereignis im Parameterraum auffasst und die Daten als gegebene Menge von Informationen, die innerhalb einer bestimmten durchgeführten Untersuchung konstant sind. Somit lässt sich Gleichung (2.34) schreiben als⁵

$$P(\text{Parameter}|\text{Daten}) \propto P(\text{Daten}|\text{Parameter}) * P(\text{Parameter}). \quad (2.35)$$

Hierbei werden die Daten als eine zusätzliche Informationsquelle aufgefasst, mit deren Hilfe unser Wissen über die Parameter auf den neusten Stand gebracht werden kann. $P(\text{Daten}|\text{Parameter})$ lässt sich unschwer als die bereits bekannte Likelihoodfunktion aus Gleichung (2.21) identifizieren. Der zweite Term $P(\text{Parameter})$ wird als A-Priori-Dichte der Parameter bezeichnet und beinhaltet den wesentlichen Unterschied zur klassischen

⁵Aus der Tatsache, dass links vom Proportionalitätszeichen wiederum eine Wahrscheinlichkeit steht, ergeben sich Bedingungen an das rechts stehende Produkt, und somit an die Konstruktion der A-Priori-Verteilung, da man an der Likelihoodfunktion wenig ändern kann. Eine A-Priori-Verteilung ist dann *zulässig* ("proper"), wenn sie nicht von denselben Daten abhängt, wie die Likelihoodfunktion, und die aus ihr resultierende A-Posteriori-Verteilung sich zu 1 integrieren lässt. Wenn sich die A-Posteriori-Verteilung zu einer Konstanten integrieren lässt, ist das Problem natürlich auch fast gelöst und lässt sich durch eine einfache Normierung beseitigen.

Sichtweise. In diesem Ausdruck werden die a priori vorhandenen Kenntnisse über die Parameter zusammengefasst. Das Produkt aus A-Priori-Dichte und Likelihoodfunktion ist proportional zur A-Posteriori-Dichte der Parameter $P(\text{Parameter}|\text{Daten})$. In dieser A-Posteriori-Dichte sind die Informationen über die Parameter aus der Likelihoodfunktion und aus den A-Priori-Dichten enthalten. Allerdings sind Dichten für eine weitergehende Datenanalyse oftmals zu unhandlich, daher werden die Parameter als Lagemaße der A-Posteriori Funktion gewonnen. Wie bereits im Falle der klassischen Analyse wird entweder der Mittelwert oder der Medianwert der Verteilung als Schätzer des Parameters genommen.

Die a priori vorhandenen Kenntnisse über die Parameter waren über eine lange Zeit der Gegenstand erbitterter Diskussionen und hitziger Debatten unter den Vertretern der klassischen und der Bayes'schen Sichtweise. Während die Anhänger von Bayes'schen Verfahren den Standpunkt vertreten, im Grunde sei doch jeder Bayesianer, denn irgendeine Vorabinformation oder Vermutung habe man schließlich vor Beginn eines Experiments immer, waren die Vertreter der klassischen Theorie der Auffassung, man habe bereits genug Möglichkeiten, ein Modell zu missspezifizieren, ohne noch zusätzliche Freiheitsgrade über die A-Priori-Dichte in ein Modell einbauen zu müssen. Die Gegner der Bayes'schen Sichtweise konnten zudem lange Zeit anführen, dass bei aller philosophischen Schönheit die Modelle sehr schnell extrem komplex und schlecht handhabbar wurden, der Gewinn beim Einsatz Bayes'scher Ansätze in der Praxis somit recht gering ist. Tatsächlich lassen sich wenige Modelle in einer geschlossenen Darstellung schreiben und exakt berechnen.

Zur Verdeutlichung betrachten wir das einfachste Modell zur Analyse von Contingent Valuation Daten aus einer offenen Befragung, das in Kapitel 1.4 kurz erläutert wurde, die lineare Regression. Beobachtet wurden individuelle Äußerungen von Zahlungsbereitschaften WTP_i und Kovariate \mathbf{x}_i . Im klassischen Bild entsteht das lineare Regressionsmodell aus der bedingten multivariaten Normalverteilung $N(\mathbf{x}\boldsymbol{\alpha}, \sigma^2\mathbf{I}|\mathbf{x})$, mit $\boldsymbol{\alpha}$ als Parametervektor und dem Vektor der Kovariate \mathbf{x} ($x_{0i} = 1$) und diese Verteilung beinhaltet die Information über die beobachteten Daten bei gegebener Verteilung und gegebenen Parametern (Greene (2003) [107]). In der Bayes'schen Welt beschreibt die Likelihoodfunktion, was eine bestimmte Realisierung der Daten an Informationen über die angenommene Verteilung der Parameter enthält, wobei sich formal wieder die Likelihoodfunktion schreiben lässt als

$$L(\boldsymbol{\alpha}, \sigma^2 | \mathbf{WTP}, \mathbf{x}) = (2\pi\sigma^2)^{-n/2} e^{-(1/2\sigma^2) \cdot (\mathbf{WTP} - \mathbf{x}\boldsymbol{\alpha})'(\mathbf{WTP} - \mathbf{x}\boldsymbol{\alpha})}. \quad (2.36)$$

Während im klassischen Modell nun die Likelihoodfunktion dadurch maximiert wird, dass der Exponent minimiert wird (und somit das Quadrat der Abweichungen minimiert wird), muss bei der Bayes-Analyse die A-Priori-Verteilung der Parameter spezifiziert werden. Dies sollte möglichst so geschehen, dass der daraus entstehende Ausdruck von einer Form ist, die es anschließend ermöglicht, aus der entstehenden A-Posteriori-Dichte die jeweiligen A-Posteriori-Dichten für α und σ durch Integration zu erhalten. Hier liegt genau das Problem der Bayes-Analyse. Einen Ausweg stellt hier für manche Probleme die Verwendung sogenannter *konjugierter* Verteilungen dar. Konjugierte Verteilungen sind gerade solche, die dieselbe funktionale Form haben, wie die Likelihoodfunktion des Problems; in diesem Falle nimmt man also für die Verteilung der Parameter α wiederum eine multivariate Normalverteilung $N(\beta_0, \Sigma_0^2)$ an. Daraus ergeben sich die Erwartungswerte der Parameter in einer intuitiven Form

$$\alpha_{\text{bayes}} = [\Sigma_0^{-1} + (\sigma^2(\mathbf{x}'\mathbf{x})^{-1})^{-1}]^{-1}(\Sigma_0^{-1}\beta_0 + (\sigma^2(\mathbf{x}'\mathbf{x})^{-1})^{-1}\alpha). \quad (2.37)$$

Der Bayes'sche Erwartungswert der Parameter α ist eine Art gewichtetes Mittel aus dem Mittelwert der A-Priori-Verteilung und dem aus der Likelihoodfunktion abgeleiteten Erwartungswert. Als Gewichtung dienen jeweils die Inversen der Varianzmatrizen Σ_0 und der "bedingten" Varianz $\sigma^2(\mathbf{x}'\mathbf{x})^{-1}$. Je größer die Varianz der A-Priori-Verteilung ist, d.h. je unsicherer unser a priori vorliegendes Wissen ist, desto geringer wird der Einfluss der A-Priori-Verteilung auf den Parameter α . Im Falle völliger Unkenntnis über die A-Priori-Verteilung nennt man diese eine nicht-informative Verteilung und erhält dieselben Ergebnisse wie in der klassischen Sichtweise. Für Discrete Choice Daten hingegen ist die Likelihoodfunktion gegeben als

$$L = \prod_{i=1}^N P(Y_{es_i} = 1)^{Y_{es_i}} \cdot (1 - P(Y_{es_i} = 1))^{1-Y_{es_i}}, \quad (2.38)$$

so dass das Bayes-Modell für Daten aus einem einfachen Referendum die folgende Form annimmt

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\theta) \cdot \prod_{i=1}^n (1 - F(\theta, \mathbf{x}, \text{Bid}))^{y_{es_i}} \cdot F(\theta, \mathbf{x}, \text{Bid})^{1-y_{es_i}}}{\int \pi(\theta) \prod_{i=1}^n (1 - F(\theta, \mathbf{x}, \text{Bid}))^{y_{es_i}} \cdot F(\theta, \mathbf{x}, \text{Bid})^{1-y_{es_i}} d\theta} \quad (2.39)$$

Je nach der funktionalen Form der A-Priori-Verteilung wird Gleichung (2.39) mehr oder weniger komplex. Obwohl ein Teil der entstehenden Komplexität an der in Gleichung (2.39) modellierten Datenstruktur liegt, weisen

auch schon einfache Bayes-Modelle einen Hang zu komplizierten funktionalen Formen auf. Wenn die Erwartung über die Verteilung der Parameter keine geschlossen konjugierte Form zulässt, müssen andere Methoden zur Bestimmung der A-Posteriori-Verteilung herangezogen werden. Die interpretatorisch ansprechende Tatsache, dass das Resultat einer Bayes-Analyse die A-Posteriori-Verteilung der Parameter ist⁶, war lange Zeit genau ihr Problem. Das Zusammenspiel von A-Priori-Dichte und Likelihoodfunktion führt in der Regel nicht zu einfachen funktionalen Formen bei der A-Posteriori-Dichte, wie in dem oben beschriebenen Beispiel. Sogar für die Modellierung der Daten aus einer Befragung nach der einfachen Referendumsmethode wie in Gleichung (2.39) scheint eine konjugierte Verteilung schwer vorstellbar. Daten aus der doppelten Referendumsmethode sind mit diesem Ansatz fast nicht mehr handhabbar, insbesondere wenn man, wie weiter unten vorgeschlagen wird, Vorabinformationen in einer informativen A-Priori-Verteilung modellieren will. Was, wie bereits erwähnt, lange Zeit als Nachteil des Bayes-Verfahrens galt, da die handhabbaren Modelle erheblichen Einschränkungen in Bezug auf die wählbaren funktionalen Formen unterlagen, kann heute als weitestgehend gelöst gelten durch simulationsgestützte Verfahren, die die Integrationen über komplexe Dichten erheblich vereinfachen. Da diese Verfahren bislang in der Auswertung von Contingent Valuation Daten selten angewendet wurden, wird nachfolgend kurz das in dieser Arbeit des weiteren verwendete Verfahren dargestellt.

2.4.2 Der Gibbs-Sampler

Der Erwartungswert der Zahlungsbereitschaft für ein umweltverbesserndes Projekt lässt sich aus Contingent Valuation Daten als Funktion ökonomisch geschätzter Parameter berechnen. Der A-posteriori-Erwartungswert einer Funktion $f(\theta)$ ist im Bayes'schen Modell gegeben durch die Integration der Funktion über die A-Posteriori-Dichte $\pi(\theta|x)$

$$E[f(\theta)] = \frac{\int f(\theta)\pi(\theta|x)d\theta}{\int \pi(\theta|x)d\theta}. \quad (2.40)$$

Die Berechnung eines solchen Erwartungswerts mithilfe von Monte Carlo Simulationen nimmt eine gänzlich andere Perspektive ein (vgl. Gilks et al. (1996) [105]). Angenommen, wir kennen die Dichtefunktion, über die

⁶Aus dieser Tatsache folgt vor allem die Einengung der Vertrauensintervalle sogar bei kleinen Stichproben durch geeignete Wahl der A-Priori-Verteilung, vgl. die Diskussion weiter unten.

integriert wird, und können Stichproben $(X_t, t = 1 \dots N)$ erzeugen, bei denen die X_t eben dieser Verteilung unterliegen. Der Erwartungswert der Funktion $f(X)$ lässt sich immer annähern durch

$$E[f(X)] \approx \frac{1}{N} \sum_{t=1}^N f(X_t), \quad (2.41)$$

i. e. durch die Summe über die Funktionswerte, die berechnet wurden an den Werten der Stichprobe. Diese Annäherung wird umso genauer, je größer der Stichprobenumfang ist. Der Stichprobenumfang liegt hier jedoch in der Hand des Statistikers und lässt sich (durch Rechenzeit begrenzt) beliebig erhöhen.

Zunächst hat sich das Problem nun nur verlagert von der Berechnung eines sehr schwierigen Integrals auf die Ziehung einer Stichprobe aus einer komplizierten Verteilung. Für dieses Problem sind jedoch etablierte Verfahren bekannt, die auf dem Konzept der Markov-Ketten basieren. Eine Markov-Kette ist eine Folge von Zufallsvariablen $\{X_0, X_1, X_2, \dots\}$, bei denen ausgehend von einem X_t der nächste Wert X_{t+1} aus einer Verteilung $P(X_{t+1}|X_t)$ stammt, die nur vom momentanen Wert der Variablen X_t abhängt. Das bekannteste Beispiel für eine Zufallsvariable ist der Random Walk (oftmals der "betrunkene Seemann" genannt), bei dem an jedem Punkt auf einem zweidimensionalen Gitter mit gleicher Wahrscheinlichkeit ein Schritt vorwärts, seitwärts (nach rechts oder links) oder rückwärts getan wird. Wichtig ist hierbei, dass jeder Schritt nur vom Ausgangspunkt und den gegebenen Übergangswahrscheinlichkeiten abhängt und unabhängig ist von allen Schritten vorher. Darüber hinaus wird auch der Anfangswert der Kette - die Bar, um im Beispiel zu bleiben - im Verlauf "vergessen". Bei einer großen Anzahl von Schritten entsprechen die einzelnen Werte von $\{X_t\}$ einer Stichprobe aus der Verteilung der Übergangswahrscheinlichkeiten.

Auf das vorliegende Problem übertragen heißt das, dass sich Markov-Ketten konstruieren lassen, die nach einer großen Anzahl von Schritten gegen eine bestimmte stationäre Verteilung konvergieren. Die Anfangsphase, in der die Kette den Startwert noch "in Erinnerung" hat, wird gemeinhin als "burn-in"-Phase von m Schritten bezeichnet. Stichprobenmittelwerte wie in Gleichung (2.41) werden dann über den Rest der Kette berechnet.

$$\bar{f} = \frac{1}{N - m} \sum_{t=m+1}^N f(X_t) \quad (2.42)$$

Somit ist das ursprüngliche Problem "Wie gewinne ich die Erwartungswerte über Funktionen der geschätzten Parameter?" ersetzt worden durch die Frage "Wie generiere ich Stichproben aus einer komplizierten Verteilung?" und mündet hier in die Frage "Wie schaffe ich eine Markov-Kette mit einer bestimmten stationären Verteilung als Ziel?". Die letzte Frage - und damit auch die erste und zweite Frage - ist erstaunlich einfach zu beantworten (vgl. Gilks et al. (1996) [105]). Während die ersten Algorithmen zur Erzeugung einer solchen Kette bereits auf Metropolis et al. (1953) zurückgehen und von Hastings (1970) weiterentwickelt wurden, gelang es Geman und Geman (1984) [103] einen Spezialfall des Metropolis-Hastings-Algorithmus unter dem Namen "Gibbs-Sampler" einzuführen, der durch die Arbeiten von Gelfand und Smith (1990) [100] oder Gelfand et al. (1990) [99] äußerst populär wurde und vielen Programmpaketen zugrunde liegt. Hierbei wird jeweils eine Komponente des Parametervektors θ gezogen und alle anderen Komponenten festgehalten (vgl. Casella and George (1992) [63]), d.h. eine Komponente $\alpha \in \theta$ wird aufdatiert aus den bedingten Wahrscheinlichkeiten $\pi(\alpha | \theta_{\text{not } \alpha}^{t-1})$, $\theta_{\text{not } \alpha}^{t-1}$ repräsentiert alle Komponenten von θ außer α . Bei einem Parametervektor mit zwei Komponenten $\theta = (\alpha, \beta)$ durchläuft jede Iteration somit die Schritte:

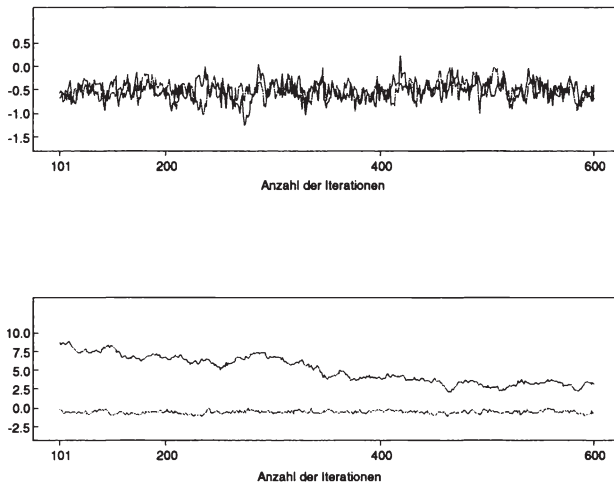
- 1. Wähle einen Startwert α_0 und ziehe β_0 aus $\pi(\beta | \alpha_0)$.
- 2. Ziehe α_1 aus $\pi(\alpha | \beta_0)$.
- 3. Wiederholung für $t = 2 \dots N$.

Der Gibbs-Sampler hat aufgrund seiner einfachen Struktur auch Eingang in viele Anwendungsbeispiele gefunden (vgl. Racine et al. (1986) [205], Gelfand et al. (1990) [99] oder auch in neuere Lehrbücher wie Greene (2003) [107], Train (2003) [243] und Congdon (2001) [67]). Zwei Fragen bleiben allerdings immer noch zu klären: Woher weiß ich, wann die burn-in Phase abgeschlossen ist? Wann ist meine Kette "lang genug"?

2.4.3 Über Konvergenz und Konfidenz

Während man bei Maximum-Likelihood-Verfahren ein eindeutiges Abbruchkriterium hat - der Gradient der Funktion ist gleich Null - ist dies bei den stichprobenbasierten Verfahren nicht der Fall. Es gibt kein eindeutiges Kriterium dafür, wann die durch Markov-Ketten erzeugte Stichprobe

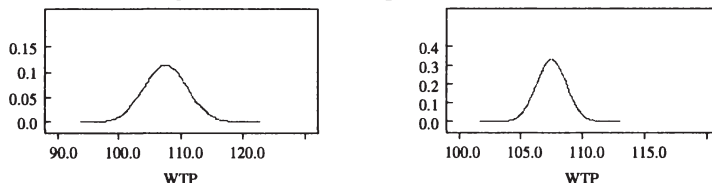
Abbildung 2.2: Graphische Analyse der Konvergenz von Markov-Ketten



tatsächlich eine Stichprobe aus der gewünschten Verteilung ist. Allerdings sind verschiedene Vorschläge in der Literatur gemacht worden, wie man zumindest sehr sicher sein kann, dass man einen "genügenden" Stichprobenumfang erzeugt hat (vgl. Casella und George (1992) für eine kurze Übersicht). Eine Stichprobe ist dann groß genug, wenn die Markov Kette ihren Startwert "vergessen" hat, d.h. wenn Ketten unterschiedlicher Startwerte konvergieren. Gelman und Rubin (1992) [102] schlagen ein Konvergenzkriterium vor, das von Brooks und Gelman (1998) [45] zu einem graphischen Analysetool weiterentwickelt wurde. Die Grundidee besteht darin, dass Stichprobenmittelwerte aus einer Kette und aus einem Pool aller Ketten miteinander verglichen werden. Wenn die Ketten auf einen gemeinsamen Wert konvergieren, so muss das Verhältnis der Varianz innerhalb einer Kette und der Varianz aus gepoolten Werten nahe bei 1 liegen. Eine Anwendung findet sich in Kapitel 3.

Einen ersten Eindruck der Konvergenz verschiedener Ketten gibt der Vergleich von Plots von Markov-Ketten mit unterschiedlichen Startwerten. Abbildung 2.2 zeigt ein Beispiel aus Spiegelhalter et al. (1999) [232] mit zwei Markovketten, deren Verlauf graphisch dargestellt ist. Der Verlauf in der oberen Abbildung lässt Konvergenz vermuten, wohingegen die untere Abbildung keine Konvergenz vermuten lassen.

Abbildung 2.3: Veränderung des Konfidenzintervalls



Wenn man auf diese Art festgestellt hat, dass Konvergenz erreicht ist, muss als nächstes die Anzahl der Iterationen entschieden werden, die man für die eigentliche Berechnung der Stichprobenstatistik nutzen möchte. Je größer diese Anzahl ist, desto besser ist natürlich das Ergebnis. Spiegelhalter et al. (1999) [232] schlagen als Faustregel vor, die Iterationen dann abzubrechen, wenn der Monte Carlo Fehler weniger als 5% der Stichprobenstandardabweichung beträgt. Der Monte Carlo Fehler der Simulation ist eine Schätzung des Unterschieds zwischen den Stichprobenwerten und den wirklichen Werten. Die meisten Programmpakete weisen die benötigten Größen für dieses Kriterium mit aus.

Aussagen zur Konvergenz der Markov-Ketten und die Bestimmung des Stichprobenumfangs gehören zu den eher technischen Aspekten der *Durchführung* einer Bayes-Analyse mithilfe von Monte Carlo Markov Chain Methoden. Unabhängig von der jeweiligen Technik der Berechnung der A-Posteriori-Dichten stellt sich natürlich auch bei der Anwendung von Bayes-Verfahren die Frage nach der Zuverlässigkeit des Ergebnisses, und somit die Frage nach einem Vertrauensintervall. Da mit $\pi(\theta|x)$ eine Aussage über die Verteilung der Parameter vorliegt, lassen sich Entsprechungen zu den aus der klassischen Statistik bekannten Konfidenzintervallen finden, indem die entsprechenden Perzentile der A-Posteriori-Dichte berechnet werden. Ein $100(1 - 2p)\%$ iges Vertrauensintervall für eine Komponente des Parametervektors entspricht somit dem p -ten Perzentil an der unteren Grenze und dem $(1 - p)$ ten Perzentil an der oberen Grenze. Die Durchführung einer Messung führt in der Bayes'schen Logik zu einer Verbesserung einer A-Priori-Vermutung über den Erwartungswert einer beispielsweise normalverteilten Variablen. Abbildung 2.3 zeigt die Simulation einer Veränderung der A-Priori-Verteilung durch zusätzliche Daten. Die A-Posteriori-Verteilung zeichnet sich durch erheblich engere Konfidenzintervalle aus.

Diese Logik lässt sich jedoch auch umkehren: im Falle einer Messung mit geringem Datenumfang, die in der Regel auch weite Konfidenzintervalle

zur Folge hat, lässt sich mit einer spezifischeren A-Priori-Dichte eine Verringerung der A-Posteriori-Intervalle erreichen. Bei nicht-informativen A-Priori-Verteilungen hingegen ergeben sich auch für die Konfidenzintervalle dieselben Größenordnungen wie für die klassische Maximum-Likelihood-Schätzung (vgl. Gelman et al. (2000) [101]). Die Verkleinerung der Konfidenzintervalle durch die Wahl einer spezifischeren Dichte ist eine grundlegende Eigenschaft der Bayes'schen Verfahren. In Kapitel 3 wird diese Eigenschaft mittels simulierter Daten illustriert.

2.4.4 Vorteile des Bayes'schen Ansatzes

Wenn also gilt, dass Bayes-Modelle unter einer nicht-informativen A-Priori-Verteilung zu denselben Ergebnissen führen wie die klassischen Modelle, und der Rechenaufwand ungleich höher ist, stellt sich die Frage: Warum Bayes? Die möglichen Antworten auf diese Frage basieren alle auf der Tatsache, dass Bayes-Verfahren es auch erlauben, eine *informative* A-Priori-Verteilung zu verwenden. An dieser Stelle wird es interessant, denn die Verwendung informativer A-Priori-Verteilungen hat einen direkten Einfluss auf die A-Posteriori-Dichte der Parameter - sie führt wie oben gezeigt zu einer schmaleren Verteilung, d.h. zu engeren Konfidenzintervallen. Wenn also systematisierbare Vorerfahrungen oder Informationen vorliegen, kann es mithilfe Bayes'scher Verfahren gelingen, die "Sicherheit", mit der Aussagen über den Erwartungswert der Zahlungsbereitschaft getroffen werden, zu erhöhen, ohne etwas an der grundlegenden Struktur der Modelle verändert zu haben, oder ohne die Verbindung zwischen theoretischem Modell und empirischer Schätzung gestört oder verändert zu haben. Die Unsicherheit über die geschätzten Werte hat ihren Ursprung, wie in Kapitel 2.1 gezeigt, in der Modellierung des stochastischen Terms zum einen und in der Größe der Stichprobe zum anderen. Beide Unsicherheiten werden direkt in Angriff genommen, wenn es Vorabinformationen gibt. Die Abnahme einer Unsicherheit über die Zahlungsbereitschaften lässt sich direkt in das Modell integrieren - eine Möglichkeit, wie sie bei der Verwendung klassischer Verfahren nicht existiert. Die Stichprobengröße wird "erhöht" indem beispielsweise die Ergebnisse von Vorabinterviews mit Experten zu Rate gezogen werden. Die Forschungsergebnisse in diesem Bereich sind noch wenig umfangreich. Erste Versuche wurden von Leon et al. (2003) [167] unternommen, bei denen Experten das statistische Modell zur Schätzung der Zahlungsbereitschaft vorgelegt wurde und sie gebeten wurden, Erwartungen bezüglich der Parameter zu formulieren. Diese Methode setzt natürlich

voraus, dass die befragten Experten die grundlegende Wirkungsweise des Modells verstehen und somit mit statistischer Modellierung vertraut sind. Das grenzt die Auswahl der möglichen Kandidaten nicht unerheblich ein. Dennoch sind die Möglichkeiten der Verbesserung einer auf einer kleineren (kostensparenden) Stichprobe basierenden Schätzung durch gezielte Erhebung von Vorabinformationen und die Verwendung Bayes'scher Verfahren noch nicht völlig ausgelotet worden. Es wäre beispielsweise denkbar eine offene Befragung mit einer Experteneinschätzung des "wahrscheinlichsten" geäußerten Werts und der Einschätzung eines 95%igen "Sicherheitsintervalls" zu kombinieren und dann das Verfahren aus Gleichung (2.37) anzuwenden. Eine Anwendung dieser Vorgehensweise auf simulierte Daten wird in Kapitel 3 gezeigt.

2.4.5 Die Wahl der A-Priori-Verteilung

Über die Wahl einer "richtig" spezifizierten A-Priori-Verteilung gibt es eine intensive Diskussion sowohl innerhalb der Anwender Bayes'scher Methoden als auch von Seiten der Kritiker des Ansatzes. Im Allgemeinen ist diese Wahl auch gewiss eine schwierige und von vielen Möglichkeiten der willkürlichen Spezifikationen gekennzeichnet. Drei grundlegend verschiedenen Ansätze sind bei der Spezifizierung der A-Priori-Verteilung denkbar.

- Es liegen wenige, vage oder sehr unterschiedliche Vorabinformationen vor.
- Es liegen konkretere Vorabinformationen und dezidierte Vermutungen über die funktionale Form der A-Priori-Verteilung vor.
- Es liegen gemessene Daten aus anderen Bewertungsstudien zu einem ähnlichen Umweltgut vor (vgl. Kapitel 6).

Der erste Fall führt zur Spezifikation einer vagen oder einer nicht-informativen A-Priori-Verteilung, die, wie weiter oben beschrieben, bei großen Stichproben die Ergebnisse der klassischen Schätzer repliziert. Die Spezifikation einer nicht-informativen A-Priori-Verteilung stellt in diesem Fall einfach eine alternative Schätzmethode für eine herkömmliche Contingent Valuation Studie dar, die einen konsistenten Referenzrahmen bieten kann, wenn beispielsweise Vergleiche zwischen unterschiedlichen Vorabinformationen gezogen werden sollen.

Der zweite Fall bietet die Möglichkeit Vorabinformationen in einer nachvollziehbaren und geschlossenen Darstellung einzubeziehen. Bayes-Verfahren mit einer *informativen* A-Priori-Verteilung können somit zur Verbesserung der Schätzung aus einer herkömmlichen CVM-Studie gewonnener Daten verwendet werden. In diesem Fall können Vorabinformationen zum Beispiel aus den Pretests und Experteninterviews genutzt werden, die sich bei Verwendung der klassischen Ansätzen nicht einbeziehen lassen. Insbesondere die Integration von Pretestergebnissen kann jedoch wertvolle Beiträge zur Modellierung einer informativen A-Priori-Verteilung leisten. In diesem Falle datiert man das in verschiedenen Pretestwellen gewonnene Wissen iterativ auf. Ziel der Bewertungsstudie ist letztlich die Messung der individuellen Nutzenänderungen in dem Gebiet, in dem die Umweltänderung vorgeschlagen wird. Daher sollten in die endgültige Auswertung und Schätzung eines statistischen Modells, das die Berechnung des Erwartungswerts der Zahlungsbereitschaft ermöglicht auch nach Möglichkeit alle Daten eingehen, die in dem zu bewertenden Gebiet und von der bewertenden Bevölkerung gewonnen werden.

Leon et al. (2003) [167] schlagen ein Verfahren vor (das allerdings irreführenderweise unter dem Titel "Nutzentransfer" publiziert wurde), bei dem die befragten Experten Perzentile der zu erwartenden Verteilung schätzen. Einfacher wäre beispielsweise das folgende Vorgehen:

- Bei einer geplanten offenen Befragung können der Mittelwert der Zahlungsbereitschaft und Konfidenzintervalle von den Experten geschätzt werden, mit deren Hilfe sich einfach auf eine Abschätzung der Standardabweichung zugreifen lässt, wenn man von einer Normalverteilung der Zahlungsbereitschaft in der Bevölkerung ausgeht.
- Bei einer geplanten geschlossenen Befragung ist die Erhebung der Expertenmeinungen etwas schwieriger. Hier müssten die Experten genügend Vertrautheit mit der statistischen Modellierung besitzen um eine Abschätzung der Parameter geben zu können, das ist jedoch fast unwahrscheinlich. Alternativ könnte man die Zustimmungsraten zu den einzelnen Geboten schätzen lassen und so mit einer nicht-parametrischen A-Priori-Verteilung die endgültige Schätzung verbessern.

Wenn hier von "Verbesserung" der Schätzung die Rede ist, muss man sich wieder in Erinnerung rufen, was das Qualitätskriterium sein kann. Die bes-

te Schätzung findet auf der Basis einer großen Stichprobe mit einem theoriekompatiblen und der Realität in vielen Zügen entsprechenden Modell statt. Somit sollte man an dieser Stelle präziser von einer Verbesserung der Aussagefähigkeit der Schätzergebnisse sprechen, denn diese gründet sich auf die Vertrauensintervalle für das berechnete Ergebnis. An die Stelle einer kombinierten Aussage von Experten und erhobenen Daten würde eine kombinierte Auswertung dieser beiden Quellen bezogen auf den Bewertungsgegenstand der Studie treten. Die Verbesserung von Vertrauensintervallen lässt sich ganz anschaulich dadurch begründen, dass insgesamt mehr Information vorliegt, wenn Expertenmeinungen und Pretests einbezogen werden. Dieser Informationsgewinn spiegelt sich beim Einsatz von Bayes-Verfahren in einer schmaleren A-Posteriori-Verteilung wider und wirkt sich auf die Breite der Vertrauensintervalle aus.

Kapitel 3

Die Ermittlung der Zahlungsbereitschaft

In diesem Kapitel werden die Modelle mit Daten gefüllt. Ausgehend von den Ergebnissen einer Contingent Valuation Befragung wird mit den beschriebenen Modellspezifikationen die jeweilige Zahlungsbereitschaft berechnet. Im ersten Schritt konzentriert sich die Analyse auf simulierte Daten. Diese haben den Vorteil, dass die Wahrheit bekannt ist, d.h. man kann die Ergebnisse der Zahlungsbereitschaftsanalyse mit der tatsächlichen mittleren Zahlungsbereitschaft vergleichen und somit ein sehr einfaches Kriterium für die Modellgüte verwenden. In der Simulation lässt sich prüfen, wie oft ein Modell dem wahren Wert nahe kommt, oder zumindest der wahre Wert im Konfidenzintervall eines Schätzers enthalten ist. Darüber hinaus lassen sich bewusst Modelle auf Daten anwenden, die mit einem vom Modell abweichenden Prozess erzeugt werden, so dass sich die Konsequenzen einer solchen Missspezifizierung beobachten lassen, wie sie bei realen Daten vorkommen wird, da die wirkliche Verteilung unbekannt ist. Im zweiten Teil dieses Abschnitts werden erhobene Daten verwendet, die aus einer im Rahmen des DFG-geförderten Sonderforschungsbereichs 565 durchgeführten Bewertungsstudie zur Rekultivierung in Brandenburg stammen. Diese Studie wird im einleitenden Teil des zweiten Abschnitts dieses Kapitels zum besseren Verständnis der Daten kurz vorgestellt.

3.1 Eine Monte Carlo Simulation

3.1.1 Simulation einer CVM-Befragung

Sinn und Zweck der Monte Carlo Simulation ist es, anhand eines erzeugten Datensatzes, der "well-behaved" ist, Strukturen zu illustrieren und nachvollziehbar zu machen, die beim Umgang mit gemessenen Daten leicht in Messfehlern oder Besonderheiten eines realen Datensatzes untergehen. Monte Carlo Simulationen werden in der Contingent Valuation Literatur häufig verwendet, um die Effekte der Modellwahl auf die Ergebnisse der Schätzung zu zeigen (vgl. Creel (1997a) [75] oder Calia und Strazzeria (1998) [52]), wenn der Einfluss der Wahl der Gebote auf die Resultate der Schätzungen diskutiert wird (vgl. Jordan und Elnagheeb (1994) [143] und Elnagheeb und Jordan (1995) [92], Cooper (1993) [70] oder Kanninen (1993) [150]), zum Vergleich verschiedener Frageformate (vgl. Cameron und Huppert (1991) [54] sowie zur Entwicklung von Methoden zur Bestimmung von oberen und unteren Schranken der Zahlungsbereitschaft aus Contingent Valuation Daten (vgl. Crooker und Kling (2000) [79]).

Simulierte Daten haben zwei wesentliche Vorteile: Zum einen ist die Verteilung der Zahlungsbereitschaften in der Bevölkerung bekannt, da sie durch den datengenerierenden Mechanismus vorgegeben ist. Man kann daher einfach überprüfen, ob das "richtige" Modell auch das richtige Ergebnis liefert. Der fast noch wichtigere Punkt ist jedoch zum anderen, dass sich die Daten für eine Stichprobe von beliebiger Größe simulieren lassen. Während der Stichprobenumfang bei einer tatsächlich durchgeführten Studie durch die zur Verfügung stehenden Gelder teilweise erheblich begrenzt ist, können in einer Simulation ohne weiteres "Befragungen" an mehreren tausend Haushalten durchgeführt werden. Dies ermöglicht es, die Effekte einer unterschiedlichen Stichprobengröße zu simulieren und zu studieren.

Während Simulationen für den Modellvergleich bereits des öfteren eingesetzt wurden, sind sie für die Analyse von Nutzentransferverfahren bislang nicht verwendet worden. Somit erfüllt dieser Abschnitt einen doppelten Zweck: er dient zum einen der Illustration der im vorherigen Kapitel eingeführten Modelle, zum anderen jedoch soll an dieser Stelle die Simulation eingeführt werden, die für die Untersuchung und auch die Kritik der bislang verwendeten Nutzentransferverfahren in Kapitel 5 und zur Verdeutlichung des Bayes'schen Nutzentransfers im Kapitel 6 benötigt wird.

Im Folgenden werden für die Ergebnisse der einfachen und der doppelten Referendumsmethode die verschiedenen klassischen und Bayes'schen parametrischen Schätzansätze verglichen. Die Monte Carlo Simulation verläuft in vier Schritten. Zunächst wird die Verteilung der Zahlungsbereitschaft in der Bevölkerung durch einen datengenerierenden Mechanismus erzeugt. Ein solcher Mechanismus kann jede zulässige funktionale Form aufweisen, die im letzten Kapitel aus dem nutzentheoretischen Entscheidungsmodell hergeleitet wurde. Darüber hinaus umfasst der datengenerierende Prozess einen stochastischen Prozess, der die Verteilung aller Teile des Entscheidungsproblems umfasst, die nicht modelliert werden sollen. In den hier verwendeten Prozessen werden drei sozioökonomische Variablen und das Einkommen verwendet. Für diese zusätzlichen Variablen muss man ebenfalls eine Verteilung in der Bevölkerung annehmen, allerdings werden hier meistens recht einfache Verteilungen wie die Normalverteilung oder die Gleichverteilung auf einem bestimmten Intervall angenommen.

Anschließend wird eine Stichprobe der Bevölkerung generiert. Der Stichprobenumfang kann variiert werden, um die Effekte größerer und kleinerer Stichproben zu studieren, die größte Stichprobe in diesem Beispiel umfasst $N = 2000$ simulierte Individuen bzw. Zahlungsbereitschaften. Im dritten Schritt wird die eigentliche Befragung simuliert. Dazu müssen die Gebote vorgegeben werden und eine Regel, nach der der Vektor aus Ja- und Nein-Antworten erzeugt werden kann. Die Bestimmung der Zahlungsbereitschaften aus dem Ergebnis der simulierten Befragung erfolgt durch Schätzung aus den simulierten Daten mit verschiedenen Modellspezifikationen.

Zunächst werden die datengenerierenden Prozesse gewählt. Sie folgen den zulässigen Modellen, die mit der Nutzentheorie vereinbar sind, d.h. der deterministische Teil folgt der Box-Cox-Transformation des Einkommens und für die stochastische Komponente werden verschiedene Verteilungen gewählt. Es werden fünf verschiedene datengenerierende Prozesse betrachtet.

- Das lineare Nutzenmodell ($\lambda = 1$), das zu einer Zahlungsbereitschaft führt, die das Einkommen nicht als Variable enthält. Für den stochastischen Term ϵ_{model} werden die Normalverteilung, die Log-Normalverteilung und die Weibull-Verteilung eingesetzt, so dass sich drei Modelle aus Gleichung (3.1) ergeben

$$WTP_{lin,i} = 10 + 2 \cdot s_{1,i} - 0,05 \cdot s_{2,i} + 0,8 \cdot s_{3,i} + \epsilon_{model,i}. \quad (3.1)$$

- Das logarithmische Modell mit ($\lambda = 0$), bei dem die Nutzenfunktion logarithmisch vom Einkommen und linear von den sozioökonomischen Variablen s abhängt. Der stochastische Term wird normalverteilt simuliert.
- Das Box-Cox-Modell mit $\lambda = 2$ hat ebenfalls eine normalverteilte stochastische Komponente.

Mit diesen Prozessen werden Monte Carlo Simulationen von fünf verschiedenen Bevölkerungen durchgeführt. Die sozioökonomischen Variablen $s_{1,i}$ und $s_{3,i}$ sind gleichverteilte Zufallszahlen aus den beiden Intervallen $[5, 60]$ und $[1, 73]$ und $s_{2,i}$ ist aus einer Normalverteilung gezogen:

$$\begin{aligned} s_{1,i} &\in [5, 60] \\ s_{2,i} &\sim N(4, 5) \\ s_{3,i} &\in [1, 73] \end{aligned} \tag{3.2}$$

Für die Modelle 4 und 5 muss darüber hinaus das Einkommen definiert werden, es wird auf dem Intervall $y \in [1000, 1400]$ bestimmt. Die Simulation folgt einem Vorschlag von Jordan und Elnagheeb (1994) [143], die sich ihrerseits auf die Studie von Cooper und Loomis (1992) [69] stützen, bei der eine Befragung im Referendumsformat zum Wert von Jagdgebieten durchgeführt wurde. Zu jeder Bevölkerung werden 500 Stichproben mit $N = 2000$ simuliert. Tabelle 3.1 gibt eine Übersicht über die mittleren Zahlungsbereitschaften und die Mediane der jeweiligen Stichproben als Mittelwert über 500 Simulationsdurchläufe. In jedem Durchlauf werden der Mittelwert und der Median der Stichprobe für die fünf Modelle berechnet und das Verfahren wird 500 Mal wiederholt. Es werden beide Maße für die Zahlungsbereitschaft berechnet, um sich für das konservativere und somit niedrigere Maß entscheiden zu können. Bei einigen der Modelle finden sich große Unterschiede zwischen den beiden Maßen. Die in der Tabelle ausgewiesenen Werte sind die über alle Durchläufe gemittelten Größen. Die ausgewiesene Standardabweichung ist die Standardabweichung der Monte Carlo Simulation. Die Parameter des jeweiligen datengenerierenden Prozesses wurden so angepasst, dass sich aus allen Modellen mittlere Zahlungsbereitschaften in ungefähr gleicher Höhe ergeben.

Mittels dieser Daten wird im nächsten Schritt die Befragung simuliert. Wie bei der Konstruktion eines Fragebogens für eine tatsächliche Befragung, müssen auch hier zunächst die Gebote festgelegt werden. Wünschenswert

Tabelle 3.1: Wahre mittlere WTP aus simulierten Befragungen

	Monte Carlo Mittelwert	Standard Abweichung
Box-Cox, $\lambda = 2$, $\epsilon \sim N(0, 2)$		
Mittelwert	102,34	0,57
Median	99,74	0,72
Lineares Modell, $\epsilon \sim N(0, 2)$		
Mittelwert	109,49	0,93
Median	109,50	1,39
Lineares Modell, $\epsilon^\epsilon \sim N(0, 2)$		
Mittelwert	111,14	0,93
Median	111,14	1,37
Lineares Modell, $\epsilon \sim Wei(3, 4)$		
Mittelwert	109,72	0,93
Median	109,70	1,37
Logarithmisches Modell, $\lambda = 0$, $\epsilon \sim N(0, 2)$		
Mittelwert	101,78	1,65
Median	98,06	2,68
Kovariate (Mittelwerte)		
s_1	35,01	0,40
s_2	4,00	0,11
s_3	36,97	0,53
y	1199,68	3,53

$N = 2000$, 500 Wiederholungen. Eigene Berechnungen.

ist ein Gebotsdesign, das um den Median oder den Mittelwert einer Verteilung herum konstruiert wird. Unglücklicherweise müsste man für ein derartiges Gebotsdesign die Verteilung der Zahlungsbereitschaften bereits kennen, und die eigentliche Studie wäre überflüssig. Bei der Konstruktion einer simulierten Befragung kennt man jedoch die Zahlungsbereitschaften genau und kann somit die Gebote entsprechend konstruieren. Hanemann und Kanninen (2001) [116] schlagen die Verwendung von 4 - 6 Geboten vor. Für den vorliegenden Fall orientieren wir uns an den Perzentilen der wahren Zahlungsbereitschaften und wählen als Vektor der Erstgebote $Bid_1 = (40, 80, 120, 160, 200)$. Um den in Kapitel 2 dargestellten nichtparametrischen Schätzer nach Kriström (1990) [158] anwenden zu können, werden die Zweitgebote so bestimmt, dass sich die jeweiligen entstehenden Intervalle nicht überlappen. Der Vektor der Zweitgebote wird festgelegt

mit $Bid_2^{low} = (20, 60, 100, 140, 180)$, falls die Antwort auf die erste Frage ablehnend ist, und mit $Bid_2^{up} = (60, 100, 140, 180, 220)$, wenn die Antwort auf die erste Frage zustimmend ist. Die vollständigen Antworten auf zwei aufeinander folgende Fragen nach der Zahlungsbereitschaft lassen sich aus der Bedingung (3.3) gewinnen.

$$\begin{aligned}
 Yes &= \begin{cases} 1 & \text{wenn gilt: } WTP_{wahr} \geq Bid_1 \\ 0 & \text{sonst} \end{cases} \\
 YesYes &= 1 \quad \text{wenn gilt: } WTP_{wahr} \geq Bid_2^{\uparrow} \quad \text{und } Yes = 1 \\
 YesNo &= 1 \quad \text{wenn gilt: } WTP_{wahr} \leq Bid_2^{\uparrow} \quad \text{und } Yes = 1 \\
 NoYes &= 1 \quad \text{wenn gilt: } WTP_{wahr} \geq Bid_2^{\downarrow} \quad \text{und } Yes = 0 \\
 NoNo &= 1 \quad \text{wenn gilt: } WTP_{wahr} \leq Bid_2^{\downarrow} \quad \text{und } Yes = 0
 \end{aligned} \tag{3.3}$$

Die Gebote werden gleichmäßig und zufällig auf die simulierte Bevölkerung verteilt. Die Anzahl der Zustimmungen je Gebot unterscheidet sich in den fünf verschiedenen Modellen. Die drei vom linearen Nutzenmodell abgeleiteten Modelle zeigen insgesamt eine recht ähnliche Struktur, wobei die Kombination mit der Normalverteilung die höchste Zustimmungsrates beim kleinsten Gebot unter diesen drei Modellen hat. Die Zustimmungen im logarithmischen Nutzenmodell fallen zunächst sehr stark ab, aber selbst auf das höchste Gebot entfallen noch einige Zustimmungen. Im Box-Cox-Modell mit $\lambda = 2$ hingegen entfällt mit 58,15 % der höchste Anteil über alle Modelle auf das kleinste Gebot und die Zustimmungen fallen sehr schnell ab, so dass schon auf das vorletzte Gebot nur noch eine sehr geringe Zustimmungsrates zu beobachten ist. Eine detaillierte Auskunft über die Zustimmungsrates zu Erst- und Zweitgeboten gibt Tabelle 3.2, die sich auf eine ausgewählte Stichprobe aus den 500 simulierten Stichproben bezieht.

Die Anzahl der *Zweit*antworten zeigt eine scheinbare Anomalie, denn sie ist in jedem Modell bei einem Gebot von 60 € größer als bei einem Gebot von 40 €. Dies liegt natürlich daran, dass das Gebot in Höhe von 60 € mehr Teilnehmern unterbreitet wird als das Erstgebot von 40 €. 60 € wird denjenigen Befragten als höheres Zweitgebot vorgeschlagen, die auf das Erstgebot 40 € ihre Zustimmung geäußert haben *und* denjenigen, die auf das Erstgebot 80 € ihre Ablehnung geäußert haben. Bei höheren Zweitgeboten nimmt sowohl die Häufigkeit als auch die Anzahl der Zustimmungen mit steigenden Geboten stark ab (vgl. auch die nichtparametrische Analyse der Zahlungsbereitschaften weiter unten).

Tabelle 3.2: Zustimmungen zu Erst- und Zweitgebot

	Lineare Nutzenfunktion, Normalverteilung				
	Erstgebot				
Yes	40	80	120	160	200
	392	301	173	52	0
	Zweitgebot				
Yes	60	100	140	180	220
	421	303	151	19	0
	Lineare Nutzenfunktion, Weibull-Verteilung				
	Erstgebot				
Yes	40	80	120	160	200
	391	304	173	53	0
	Zweitgebot				
Yes	60	100	140	180	220
	416	305	149	19	0
	Lineare Nutzenfunktion, Log-Normalverteilung				
	Erstgebot				
Yes	40	80	120	160	200
	392	306	178	59	0
	Zweitgebot				
Yes	60	100	140	180	220
	421	311	151	23	0
	Logarithmische Nutzenfunktion, Normalverteilung				
	Erstgebot				
Yes	40	80	120	160	200
	301	243	150	94	39
	Zweitgebot				
Yes	60	100	140	180	220
	320	250	158	96	25
	Box-Cox ($\lambda = 2$) Nutzenfunktion, Normalverteilung				
	Erstgebot				
Yes	40	80	120	160	200
	398	333	104	8	0
	Zweitgebot				
Yes	60	100	140	180	220
	472	296	77	0	0
Eigene Berechnungen.					

3.1.2 Klassische Modellierung

Damit ist die Simulation der Befragung abgeschlossen und wir können uns der Anwendung der verschiedenen Modelle zuwenden. Da die wahren Mittelwerte der Zahlungsbereitschaften für jede Stichprobe bekannt sind, lassen sich die Modelle über ihre jeweilige Abweichung des geschätzten Mittelwerts von der wahren Zahlungsbereitschaft vergleichen. Darüber hinaus werden die Abhängigkeit des geschätzten Werts von der Stichprobengröße sowie die Verbesserung von Konfidenzintervallen durch die Anwendung des doppelten Referendumsformats untersucht. Daher werden zunächst fünf Modelle zur Schätzung der Zahlungsbereitschaft auf das einfache Referendumsformat (die Erstantworten) angewendet.

Tabelle 3.3 fasst die Ergebnisse der verschiedenen Ansätze zusammen. Es wurde in allen Modellen zunächst nur das Gebot als Variable in die Schätzung eingeführt. Die Schätzungen wurden an jeweils einem simulierten Datensatz aus jedem datengenerierenden Prozess mit allen fünf Modellen durchgeführt. Die Konfidenzintervalle wurden mit der in Kapitel 3 auf Seite 72 beschriebenen Bootstrap-Methode mit 500 Wiederholungen berechnet. Die wahre Zahlungsbereitschaft der Stichprobe ist in der Tabelle als Mittelwert mit der Standardabweichung in Klammern vor den eigentlichen Schätzergebnissen aufgeführt. Die letzten beiden Spalten enthalten den Likelihood-Endwert und die Devianz, zwei Werte, aufgrund derer man zusätzliche Aussagen über die Güte des Modells gewinnen kann. Zur Berechnung der Devianz muss man zunächst die Funktion L_{freq} für die verschiedenen Modelle berechnen. Die letzte Spalte in Tabelle 3.3 zeigt die Werte dieser Funktion und die Devianz der verschiedenen Modelle. Bei einem kritischen Wert für die χ^2 -Verteilung mit $(5-2=3)$ Freiheitsgraden von 11,34 wird das Probit-Modell in allen Spezifikationen der Modelle akzeptiert, das Logit-Modell in fast allen Fällen, alle anderen Modelle werden fast immer zurückgewiesen. Einzig für das Box-Cox-Modell mit $\lambda = 2$ erweist sich das lognormale Modell als die beste Modellierung der Daten. Vergleicht man jedoch den Mittelwert, der mit diesem Modell berechnet wird, mit dem "wahren" Mittelwert, so zeigen sich hier erhebliche Abweichungen, der wahre Wert ist sogar in den Konfidenzintervallen des geschätzten Wertes nicht enthalten und umgekehrt.

Das Probit-Modell hingegen kommt bei der Anwendung auf die verschiedenen Grundmodelle immer dem wahren Mittelwert der Zahlungsbereitschaft am nächsten. In den meisten Fällen führt es zu Abweichungen der geschätzten Zahlungsbereitschaft von der wahren Zahlungsbereitschaft um

Tabelle 3.3: Vergleich der Modelle zur Schätzung der WTP

Modell	Mittelwert	95 % Konfidenz	L.Likelihood	Devianz
	Lin. Nutzenmodell, Normalvert.; WTP = 109,83			
Logit	107,97	[105,16 - 110,95]	-681,96	16,74
Log-Logit	109,64	[106,86 - 112,44]	-695,25	43,32
Probit	108,22	[105,52 - 110,96]	-678,29	9,40
Log-Normal	121,28	[118,32 - 124,85]	-698,23	49,28
Weibull ^{a)}	92,20	[89,61 - 95,31]	-753,22	159,26
	Lin. Nutzenmodell, Weibull-Vert.; WTP = 110,07			
Logit	108,38	[105,61 - 111,66]	-689,69	15,56
Log-Logit	110,22	[107,65 - 113,48]	-706,78	49,74
Probit	108,52	[105,66 - 111,60]	-686,14	8,46
Log-Normal	121,83	[118,62 - 125,29]	-710,64	57,46
Weibull ^{a)}	92,12	[89,27 - 95,02]	-767,57	171,32
	Lin. Nutzenmodell, Log-Normalvert.; WTP = 111,51			
Logit	109,70	[106,58 - 112,16]	-688,89	11,19
Log-Logit	111,48	[108,93 - 114,41]	-703,36	40,13
Probit	109,89	[106,65 - 112,42]	-685,63	4,67
Log-Normal	123,38	[119,97 - 126,99]	-707,14	47,69
Weibull ^{a)}	93,81	[91,10 - 96,75]	-763,17	159,75
	Log. Nutzenmodell, Normalvert.; WTP = 103,02			
Logit	97,75	[91,72 - 103,35]	-1148,51	6,02
Log-Logit	161,38	[143,65 - 187,83]	-1167,67	44,34
Probit	97,89	[92,17 - 103,37]	-1147,95	4,90
Log-Normal	135,57	[128,64 - 143,61]	-1167,08	43,16
Weibull ^{a)}	78,43	[74,65 - 82,74]	-1178,22	65,44
	Box-Cox $\lambda = 2$, Normalvert.; WTP=102,31			
Logit	102,33	[99,73 - 104,56]	-452,65	11,21
Log-Logit	102,12	[100,33 - 104,18]	-451,19	8,29
Probit	102,71	[100,57 - 105,39]	-451,45	8,81
Log-Normal	111,80	[109,69 - 113,99]	-447,80	1,51
Weibull ^{a)}	94,70	[93,18 - 96,31]	-256,65	*

N = 500, ^a Medianwert, *nicht aussagefähig. Eigene Berechnungen.

1 Prozent. Eine Ausnahme stellt die Anwendung des Probit-Modells auf das lognormale Modell dar, hier beträgt die Abweichung fast 5 Prozent. Die Likelihoodfunktion nimmt immer beim Probit-Modell den größten Wert an, ebenso wie die Devianz immer den kleinsten Wert annimmt. Dies spricht tatsächlich für die Wahl des einfachsten Schätzmodells, zumindest für Strukturen, die den hier simulierten Daten ähneln. Probit- und Logit-Modelle unterscheiden sich wenig in ihren Voraussagen. Auch die Werte der Log-Likelihoodfunktion liegt immer sehr nahe beieinander. Das Devianzkriterium bevorzugt allerdings das Probit-Modell.

Unter den Modellen mit nicht-symmetrischen Verteilungen überschätzt das Log-Normal-Modell die Zahlungsbereitschaft systematisch, in keinem Fall liegt die "wahre" Zahlungsbereitschaft auch nur im Konfidenzintervall des Modells. An der Log-Likelihood-Funktion lässt sich das allerdings nicht erkennen. Das Weibull-Modell dagegen unterschätzt die "wahre" Zahlungsbereitschaft in allen Fällen. Für die Box-Cox- Transformation mit $\lambda = 2$ und für das logarithmische Modell führt es allerdings zu einer ebenso guten Anpassung wie Probit und Logit.

Wenn man wirklich auf der Suche nach dem besten Modell für die vorliegenden Daten ist, lässt sich keine eindeutige Aussage treffen. Um spezielle Anpassungen vorzunehmen, müsste das zugrunde liegende Modell im Detail bekannt sein, was bei Daten in der Realität nicht gegeben sein wird. Wie bereits in Kapitel 2 diskutiert, zeigen sich das Probit- und das Logit-Modell in den dort zitierten praktischen Anwendungen und auch in dieser Simulation überlegen. Das Logit-Modell hat darüber hinaus den Vorteil, dass die Verteilung sich in einer geschlossenen Form darstellen lässt, so dass Berechnungen der Zustimmungswahrscheinlichkeiten ohne großen Aufwand möglich sind. Dieser Vorteil ist es, der weiter unten in den Überlegungen zu informativen Verteilungen beim Bayes'schen Ansatz eine Rolle spielen wird. Nun ist ein Stichprobenumfang von 2000 Befragten in der Realität selten zu realisieren. Daher wird als nächstes der Einfluss des Stichprobenumfangs auf die Qualität des Resultats simuliert. Dieser Einfluss ist bei den unterschiedlichen Modellen ähnlich, daher zeigt die nachstehende Tabelle die Ergebnisse für das Probit-Modell, das bislang im Modellvergleich gut abgeschnitten hat. Aus der großen Stichprobe werden Unterstichproben gezogen mit der Größe $n'=(200, 400, 1000)$ und Monte Carlo Simulationen durchgeführt. Dabei ist es bei den kleinen Stichproben umso wahrscheinlicher, dass man eine Unterstichprobe zieht, die nahezu keine Zahlungsablehnungen oder zu wenige Zahlungszustimmungen enthält.

Tabelle 3.4: Einfluss des Stichprobenumfangs

N	Linear, Normalverteilung			
	Einfaches Format		Doppeltes Format	
	Mittelwert	95 % Konfidenz	Mittelwert	95 % Konfidenz
200	110,15	[100,54 - 119,75]	110,16	[102,60 - 117,73]
400	109,99	[103,29 - 116,98]	109,73	[104,55 - 114,91]
1000	109,69	[105,61 - 113,78]	110,77	[107,41 - 111,13]
N	Linear, Log-Normalverteilung			
	Mittelwert	95 % Konfidenz	Mittelwert	95 % Konfidenz
200	111,18	[101,62 - 120,75]	111,85	[104,55 - 119,16]
400	111,70	[105,27 - 118,14]	111,93	[106,58 - 117,30]
1000	111,63	[107,46 - 115,80]	111,79	[108,44 - 114,14]
N	Linear, Weibull-Verteilung			
	Mittelwert	95 % Konfidenz	Mittelwert	95 % Konfidenz
200	109,89	[100,73 - 119,06]	110,87	[103,37 - 118,38]
400	110,46	[103,84 - 117,08]	110,81	[105,45 - 116,16]
1000	109,98	[105,88 - 114,09]	110,80	[107,45 - 112,16]
N	Logarithmisch, Normalverteilung			
	Mittelwert	95 % Konfidenz	Mittelwert	95 % Konfidenz
200	101,08	[83,69 - 118,46]	99,39	[83,36 - 115,41]
400	101,59	[90,28 - 112,90]	98,73	[87,51 - 109,95]
1000	100,48	[93,01 - 107,96]	98,64	[92,87 - 105,42]
N	Box - Cox, Normalverteilung			
	Mittelwert	95 % Konfidenz	Mittelwert	95 % Konfidenz
200	105,84	[99,34 - 112,34]	102,32	[97,52 - 107,11]
400	102,71	[97,65 - 107,77]	102,28	[98,91 - 105,64]
1000	102,93	[99,73 - 106,13]	102,38	[100,29 - 104,46]

Eigene Berechnungen.

Daher wird die Anzahl der Wiederholungen auf 50 herabgesetzt. Ein anderer rechenaufwändigerer Weg wäre, kleine einzelne Stichproben der entsprechenden Größe zu simulieren und diese Simulationen mit allen folgenden Berechnungen zu wiederholen. Die Ergebnisse verändern sich jedoch bei dieser Vorgehensweise nur geringfügig und die Tendenz zeigt sich recht eindrucksvoll bereits in Tabelle 3.4. Zur Vergleichbarkeit der Ergebnisse werden alle Simulationen in diesem Block mit derselben Anzahl von Replikationen durchgeführt. In allen Modellen nehmen die Konfidenzintervalle mit zunehmendem Stichprobenumfang ganz erheblich ab, bei den ersten

vier Modellen um mehr als die Hälfte. Am stärksten wirkt sich ein kleiner Stichprobenumfang auf das logarithmische Modell aus, hier überdeckt das Konfidenzintervall mehr als 30 €.

Zum Vergleich wird im nächsten Schritt die Verbesserung der Konfidenzintervalle durch die Verwendung des doppelt gebundenen Formats untersucht. Die Verbesserung der Schätzung ist bei allen Modellen gut zu erkennen. Die Konfidenzintervalle der Stichproben aus 200 Befragten sind bei Verwendung des doppelten Referendumsformats fast genauso weit, wie die Konfidenzintervalle der doppelt so großen Stichprobe im einfachen Format.

Dies demonstriert recht eindrucksvoll die Verbesserung der Effizienz durch das doppelte Format und somit eine Erhöhung der Sicherheit bezüglich der gemessenen Zahlungsbereitschaft bei den Politikern, die eine Entscheidung aufgrund derartige Zahlen fällen werden. Bei realen Daten müsste sich diese Effizienzsteigerung ebenfalls zeigen, denn zu diesem Zweck wurde das doppelte Referendumsformat entwickelt. Allerdings muss, wie im vorherigen Abschnitt erläutert wurde, gewährleistet sein, dass die Befragten beide Antworten auf der Basis ein und derselben Zahlungsbereitschaft formulieren. Zur Überprüfung dieser Tatsache kann man das bivariate Modell aus Gleichung (2.26) verwenden, was bei simulierten Daten natürlich bestätigt, dass beiden Antworten dieselbe Zahlungsbereitschaft zugrunde liegt (vgl. Tabelle 3.5). Das bivariate Probit-Modell unterstellt zunächst, dass der Erstantwort eine andere Zahlungsbereitschaft zugrunde liegt als der Zweitantwort. Es lässt eine Korrelation zwischen diesen beiden Zahlungsbereitschaften zu, ausgedrückt durch den Parameter ρ . Ist $\rho \simeq 1$ so kann man von einer gemeinsamen Zahlungsbereitschaft ausgehen, ansonsten muss das Modell zwei Zahlungsbereitschaften zulassen. Für die Überprüfung der Hypothese $\rho = 0$ reicht ein einfacher Likelihood-Ratio-Test, die Testergebnisse sind in der Tabelle angegeben. Für die große Stichprobe mit $N = 2000$ weichen die Parameter im nicht restringierten Modell, bei dem die Parameter α_0 und β nicht gleichgesetzt werden, nur geringfügig von einander ab. Der Likelihood-Ratio-Test weist die Hypothese $\rho = 0$ zurück. Für eine kleine Stichprobe mit $N = 200$ hingegen sehen die Ergebnisse ganz anders aus. Hier zeigt sich nur eine sehr geringen Korrelation zwischen den Zahlungsbereitschaften und die Parameter sind sehr unterschiedlich im nicht restringierten Fall. Hier wird eine Gefahr des bivariaten Probit-Modells deutlich.

Tabelle 3.5: Parameterschätzung im Bivariaten Probit-Modell

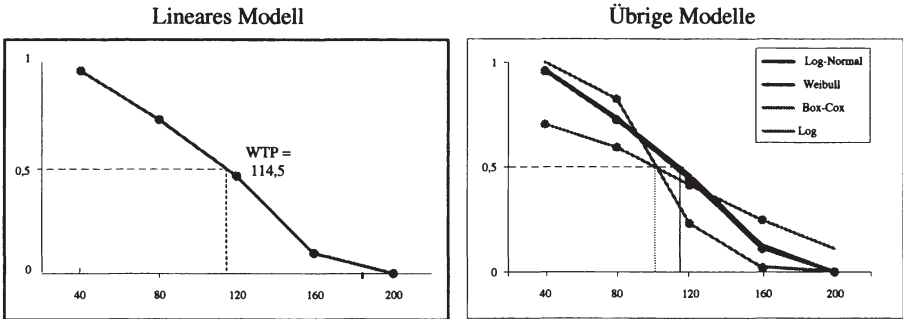
	α_0	β
$N = 2000$		
Nicht restringiertes Modell		
Erstantwort	2,70 [2,50 - 2,90]	-0,025 [-0,026 - -0,023]
Zweitantwort	2,65 [2,50 - 2,84]	-0,024 [-0,025 - -0,022]
	$\rho = 0,85$ LR-Test auf $\rho = 0$	$P > \chi^2 = 0,000$
Restringiertes Modell, α_0		
Erstantwort	2,67 [2,50 - 2,84]	-0,024 [-0,026 - -0,023]
Zweitantwort	2,67 [2,50 - 2,84]	-0,024 [-0,025 - -0,022]
Restringiertes Modell, α_0, β		
Erstantwort	2,69 [2,52 - 2,86]	-0,024 [-0,026 - -0,023]
Zweitantwort	2,69 [2,52 - 2,86]	-0,024 [-0,026 - -0,023]
$N = 200$		
Nicht restringiertes Modell		
Erstantwort	2,27 [1,34 - 3,19]	-0,045 [-0,060 - -0,030]
Zweitantwort	9,84 [5,34 - 16,58]	-0,168 [-0,172 - -0,163]
	$\rho = 0,22$ LR-Test auf $\rho = 0$	$P > \chi^2 = 0,6135$
Restringiertes Modell, α_0		
Erstantwort	2,67 [1,90 - 3,45]	-0,051 [-0,064 - -0,037]
Zweitantwort	2,67 [1,90 - 3,45]	-0,048 [-0,061 - -0,036]
Restringiertes Modell, α_0, β		
Erstantwort	2,68 [1,89 - 3,46]	-0,049 [-0,062 - -0,037]
Zweitantwort	2,68 [1,89 - 3,46]	-0,049 [-0,062 - -0,037]

Eigene Berechnungen.

Obwohl die Zahlungsbereitschaften ganz eindeutig aus derselben simulierten Verteilung kommen, wird $\rho = 0$ auf einem hohen Niveau mit über 50%iger Wahrscheinlichkeit vom Likelihood-Ratio-Test für möglich gehalten. Dies liegt daran, dass das bivariate Modell aus wenigen Daten *zwei* Zahlungsbereitschaften anpassen muss, während das restringierte Modell nur die Beobachtungen einem Zahlungsbereitschaftsmodell zuordnen muss. Die Erhöhung der Freiheitsgrade im Modell kann bei kleinen Stichproben zu erheblichen Verzerrungen führen.

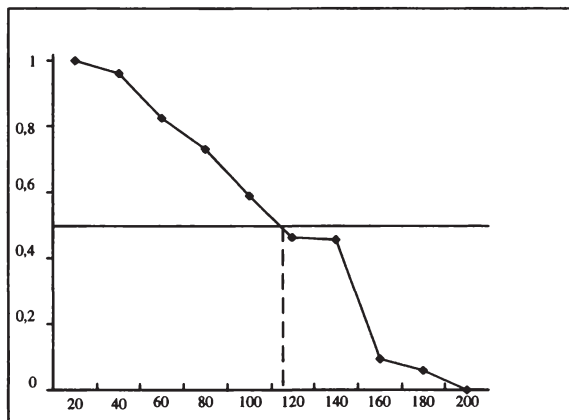
Die bislang genutzten Modelle sind im Falle des einfachen Formats in vielen

Abbildung 3.1: Nichtparametrische Analyse der Erstantworten



Statistikpaketen implementiert, für das doppelte Format müssen meistens kleinere Programme geschrieben werden. Völlig ohne Computeraufwand, ja sogar mit einem Taschenrechner lässt sich das von Kriström (1990) vorgeschlagene nichtparametrische Modell berechnen. Hierzu wird wie in Abbildung 3.1 die Verteilung der Häufigkeiten der Zustimmungen zur ersten Frage aufgetragen. In der Abbildung ist im oberen Teil die Häufigkeitsverteilung für das lineare Modell mit einer Normalverteilung herausgegriffen. Die Verteilung der Häufigkeiten kann als eine nicht-parametrische Annäherung an die Wahrscheinlichkeit einer Zustimmung und somit an die Verteilung der Zahlungsbereitschaften aufgefasst werden. Daher lässt sich der Medianwert der erwarteten Zahlungsbereitschaft aus Abbildung 3.1 ablesen, indem man den Schnittpunkt der 50%-Linie mit der Verteilung der Häufigkeiten bestimmt. Den Mittelwert der Zahlungsbereitschaft würde man in dieser Darstellung errechnen, indem man die Fläche unter der Kurve aufsummiert. Allerdings muss man dazu die Schnittpunkte mit den Achsen bestimmen, man müsste also den kleinsten Betrag, bei dem jeder zustimmt und den höchsten Betrag, bei dem keiner mehr zustimmt, festlegen. Dies wird in den seltensten Fällen möglich sein, und der Median ist von dieser Wahl unabhängig, daher ist er das richtige Maß. Es ergibt sich für die lineare Nutzenfunktion mit normalverteiltem stochastischem Term eine erwartete Zahlungsbereitschaft von 114,48 €. Die Abbildung zeigt die unterschiedlichen Zustimmungsraten, zu denen die weiteren Modellierungen der Nutzenfunktion führen. Die Modelle, die das Einkommen als Variable einschließen (Logarithmisch und Box-Cox mit $\lambda = 2$) liegen zumindest vom Medianwert der Zahlungsbereitschaft eng beieinander, ebenso

Abbildung 3.2: Häufigkeiten aller Antworten



wie die Modelle, die auf der linearen Nutzenfunktion (ohne Einkommen) basieren. Letztere weisen auch einen sehr ähnlichen Kurvenverlauf über die ganze Breite aller Gebote auf.

Zur Analyse der Zweitantworten müssen die Häufigkeiten der Zustimmung zu einem bestimmten Zweitgebot aus den Zustimmungen nach einer *Ablehnung* und nach einer *Annahme* des Erstgebots berechnet werden. Die Zustimmung zu einem Gebot von 60 € errechnet sich somit als

$$\pi_{60} = \frac{335 + 68}{377 + 111} = 0,826. \quad (3.4)$$

Abbildung 3.2 zeigt den Verlauf der sich daraus ergebenden Kurve für die Häufigkeiten der positiven Antworten nach Geboten mit Auswertung der Zweitantworten. Mit derselben Vorgehensweise wie bei der Analyse der Erstantworten ergibt sich der Median der Zahlungsbereitschaft zu 114,17 €, und ist somit geringfügig niedriger als der aus der Analyse der Erstantworten gewonnene Wert. In Abbildung 3.2 deutet sich eine Struktur in den Daten an, die man bei tatsächlich gemessenen Daten ausgeprägter und häufiger findet. Die Kurve der Häufigkeiten ist bei gemessenen Daten oftmals an einigen Stellen nicht monoton fallend, so dass man mit einem entsprechenden Verfahren eine monoton fallende Folge von Häufigkeiten erzwingen muss. Ayer (1955) konnte zeigen, dass man $\pi_i \leq \pi_{i+1}$ ersetzen kann durch $\bar{\pi}_j = \bar{\pi}_{j+1}$, wobei sich $\bar{\pi}$ berechnen lässt aus $\frac{\text{Summe der Zustimmungen}_{j,j+1}}{\text{Summe der Befragten}_{j,j+1}}$ (vgl.

Kriström (1990) [158] S. 89). Im vorliegenden Fall ist diese Glättung nicht notwendig, da die Zustimmungsraten geringfügig fallen und der Schnittpunkt mit der Medianlinie ohnehin links von dieser Stelle liegt. Bei einem anderen datengenerierenden Prozess als dem gewählten muss dies aber nicht immer der Fall sein, und das Ayer-Verfahren wird angewendet. In Daten aus einer durchgeführten Befragung findet man derartige Unregelmäßigkeiten häufig, insbesondere bei Daten aus dem doppelten Referendumsformat. Wenn also die doppelte Referendumsmethode zwar zum versprochenen Effizienzgewinn führt, es aber nicht als gesichert gelten kann, dass sie auch von den Befragten richtig verstanden wird, gibt es Anreize, sich dem oben beschriebenen alternativen statistischen Ansatz zuzuwenden, und zu untersuchen, ob im Rahmen von Bayes'schen Verfahren sich die Vorteile des doppelten Formats mit anderen Mitteln erzielen lassen.

Da diese Vorgehensweise bislang in der Literatur nicht verwendet wird, geht Kapitel 3.1.3 ausführlich zunächst auf den Einsatz einer nicht-informativen A-Priori-Verteilung im Bayes'schen Modell ein und illustriert diese Vorgehensweise wiederum mithilfe der Simulationsdaten. Anschließend wird gezeigt, wie sich Vorabinformationen gewinnen und verwenden lassen, ebenfalls anhand der simulierten Daten, da zur Konstruktion von informativen A-Priori-Verteilungen für die in Kapitel 3.2 geschilderte Bewertungsstudie keine Daten zur Verfügung stehen.

3.1.3 Bayes'sche Modellierung

Bayes'scher Ansatz mit nicht-informativer A-Priori-Verteilung

Der wesentliche Unterschied eines Bayes'schen Modells zu einem klassischen Modell besteht in der Notwendigkeit, eine A-Priori-Verteilung zu spezifizieren. In dieser A-Priori-Verteilung wird das vor der eigentlichen Durchführung einer Untersuchung vorhandene Wissen abgebildet. Wenn nur wenig Vorabwissen vorliegt, so greift man im Bayes'schen Modell zu einer nicht-informativen A-Priori-Verteilung. Da alle A-Priori-Verteilungen im Bayes'schen Modellrahmen der in Kapitel 2 definierten Bedingung der "Zulässigkeit" genügen müssen, wird für die vorliegende Modellierung der nicht-informativen Verteilung eine Normalverteilung mit einer sehr kleinen Präzision, bzw. mit einer sehr großen Varianz gewählt.

Tabelle 3.6: Bayes Schätzung mit nicht-informativer A-Priori-Verteilung

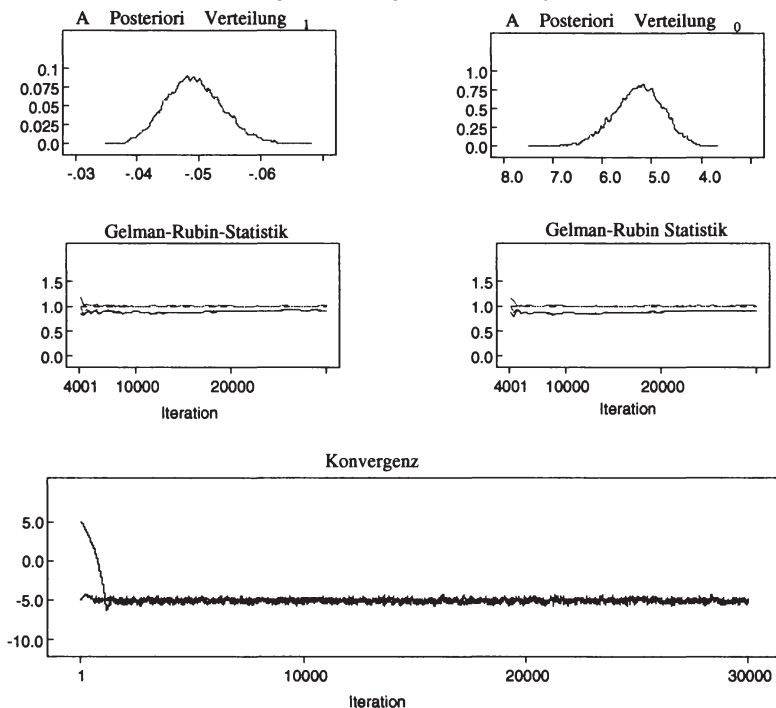
		Linear, Normalverteilung	
Stichprobenumfang		Mittelwert	95 % Konfidenz
	200	110,0	[101,4 - 119,0]
	400	107,1	[100,7 - 113,7]
	2000	107,9	[105,0 - 110,9]
		Linear, Log-Normalverteilung	
Stichprobenumfang		Mittelwert	95 % Konfidenz
	200	112,2	[103,5 - 121,1]
	400	109,3	[102,7 - 116,0]
	2000	109,7	[106,8 - 112,6]
		Linear, Weibull-Verteilung	
Stichprobenumfang		Mittelwert	95 % Konfidenz
	200	111,2	[102,2 - 120,5]
	400	108,3	[101,7 - 115,1]
	2000	108,4	[105,5 - 111,3]
		Logarithmisch, Normalverteilung	
Stichprobenumfang		Mittelwert	95 % Konfidenz
	200	97,05	[79,95 - 113,2]
	400	96,69	[83,54 - 109,0]
	2000	97,72	[92,09 - 103,2]
		Box - Cox, Normalverteilung	
Stichprobenumfang		Mittelwert	95 % Konfidenz
	200	106,1	[99,06 - 113,2]
	400	104,8	[99,44 - 110,3]
	2000	102,3	[99,99 - 104,7]

Eigene Berechnungen.

Die Ergebnisse der Bayes'schen Schätzung mit einer nicht-informativen A-Priori-Verteilung entsprechen den Ergebnissen der klassischen Schätzung für die untersuchten Logit- und Probit-Modelle. Tabelle 3.6 weist die Resultate der Logit-Schätzung aus¹. Das Modell umfasst die Konstante und einen Parameter zur Gebotsvariablen (α_0 und α_{bid}), beide A-Priori-Verteilungen werden mit der Normalverteilung mit Mittelwert 0 und einer Präzision von $\tau = 10^{-6}$ angesetzt, also mit einer Normalverteilung mit sehr großer Vari-

¹Die Werte in Tabelle 3.6, sowie alle Bayes'schen Modelle in dieser Arbeit sind mit der Software BUGS (Bayesian Inference Using Gibbs Sampling) berechnet, die auch eine Reihe von Analyseinstrumenten zur Bestimmung der Konvergenz der Modelle bietet (www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml).

Abbildung 3.3: Graphische Analyse



anz (zur Verwendung dieser A-Priori-Verteilung vgl. beispielsweise Spiegelhalter et al. (1999) [232]). Diese Modelle laufen mit simulierten Daten ausgesprochen robust und schnell und die Markov Ketten konvergieren rasch gegen die endgültige Verteilung. Es wurden 30000 Iterationen nach einer Burn-In Phase von jeweils 5000 Iterationen benutzt um die Verteilung der Parameter zu bestimmen. Abbildung 3.3 zeigt die graphischen Ergebnisse des Linearen Nutzenmodells mit normalverteilterm stochastischem Term. Im oberen Teil ist die Konvergenz zweier Ketten mit unterschiedlichen Startwerten gezeigt. Im unteren Teil ist die A-Posteriori-Verteilung der Parameter α_0 und α_{bid} gezeigt. Um das Konvergenzverhalten von Markov Ketten zu untersuchen, kann man, wie in Abbildung 3.3 unten, sich zunächst, wie in Kapitel 2 beschrieben, verschiedene Ketten mit unterschiedlichen Startwerten anschauen. Man sieht die unterschiedlichen Startwerte links im Bild und erkennt, dass die Ketten sehr schnell anfangen, um einen ge-

meinsamen Wert zu schwanken. Das auf der Gelman-Rubin Statistik (vgl. Kapitel 2, S. 84) basierende graphische Gütekriterium wird in Abbildung 3.3 in der Mitte gezeigt. Die obere Linie in den Abbildungen zur Gelman-Rubin Statistik repräsentiert das Verhältnis der Varianz innerhalb einer Kette zur Varianz einer aus einem Pool gebildeten Kette. Es liegt sehr nahe bei 1, was auf Konvergenz hindeutet.

Bayes'scher Ansatz mit informativer A-Priori-Verteilung

Die eigentlichen Vorteile des Bayes'schen Modells lassen sich jedoch nur erkennen, wenn man bereits Informationen über die Zahlungsbereitschaft hat und versucht, diese in das Modell einfließen zu lassen. Hierzu bietet die klassische Statistik kein systematisches Verfahren an. Zur Vorbereitung einer CVM Befragung werden jedoch immer systematische Pretests und Expertenbefragungen gemacht, wie in Kapitel 1.3 ausführlich dargestellt wurde. Resultate aus diesen Voruntersuchungen fließen bei Verwendung der klassischen Methoden bislang in unsystematischer Weise in die Fragebogengestaltung oder das Design der Gebote ein. Ein alternatives Vorgehen bietet sich im Rahmen der Bayes-Analyse an, es wurde in Kapitel 2.4.5 vorgeschlagen. Angenommen, die Experten werden gefragt, bei welchem Betrag sie 50% Zustimmung erwarten und bei welchem Betrag sie eine Zustimmung von 90% aller Befragten erwarten. Eine derartige Frage benötigt keine vertieften Statistikkenntnisse zu ihrer Beantwortung, zumal die Experten häufig aus Berufsfeldern kommen, in denen sie mit derartigen Aussagen häufig konfrontiert werden, beispielsweise aus planerischen Berufen. Die Aussagen zu den erwarteten Zustimmungsraten lassen sich in A-Priori-Verteilungen zu den Modellparametern übersetzen, für das Logit-Modell wird dies mit den simulierten Daten aus unserem Beispiel kurz gezeigt. Dieser recht einfach wirkende Ansatz stellt eine Möglichkeit dar, in den zur Szenarioentwicklung ohnehin durchgeführten Expertengesprächen eine Abschätzung für die informative A-Priori-Verteilung zu gewinnen. Die grundsätzliche Wirkungsweise lässt sich besten am linearen Nutzenmodell mit einer Normalverteilung und der Anwendung des Logit-Modells zeigen. Im Logit-Modell ist die Wahrscheinlichkeit der Zustimmung zu einem Gebot gegeben durch

$$Pr(yes) = \frac{1}{1 + e^{-(\alpha_0 + \beta \cdot bid_1)}}. \quad (3.5)$$

Angenommen, die Experten hätten den Median der Verteilung bei 90 € gesehen und wären sich sicher, dass bei 50 € mindestens neunzig Prozent aller

Tabelle 3.7: Bayes Schätzung mit informativer A-Priori-Verteilung

Stichprobenumfang	Mittelwert	95 % Konfidenz
Eher unsicher $\alpha_0 \propto N(4, 3; 4)$, $\beta \propto N(-0, 04; 0, 01)$		
200	110,2	[101,4 - 120,0]
400	107,6	[102,7 - 114,7]
Weniger unsicher $\alpha_0 \propto N(4, 3; 40)$, $\beta \propto N(-0, 04; 0, 1)$		
200	108,7	[102,5 - 114,9]
400	107,6	[102,7 - 113,0]
Sehr sicher $\alpha_0 \propto N(4, 3; 400)$, $\beta \propto N(-0, 04; 1)$		
200	107,7	[105,2 - 110,9]
400	107,5	[105,5 - 109,9]
Eigene Berechnungen.		

Befragten zustimmen. Setzt man die Wahrscheinlichkeit der Zustimmung in Gleichung (3.5) gleich 0,5 bzw. 0,9 bei den entsprechenden Gebotswerten, so kann man nach α_0 und β auflösen. Es ergibt sich der Parametervektor $(\alpha_0, \beta) = (4, 3, -0, 04)$. Diese Werte lassen sich als Mittelwerte einer informativen A-Priori-Verteilung interpretieren. Die Varianz der A-Priori-Verteilung ergibt sich aus der "Sicherheit", mit welcher die Experten ihre Meinung abgegeben haben. Tabelle 3.7 zeigt die Ergebnisse für unterschiedlich große Stichproben und unterschiedliche Niveaus an Sicherheit bei den Experten. Vergleicht man die Konfidenzintervalle des Modells mit der informativen A-Priori-Verteilung, so kommen sie nahe an die Resultate der Berechnungen mittels der klassischen Modells bei Verwendung des *doppelten Formats*. Bei allen bereits skizzierten Schwierigkeiten des doppelten Formats und der Tatsache, dass die Rechtfertigung für seine Verwendung gerade in der Reduzierung von Konfidenzintervallen besteht, gibt diese Aussage zumindest Anlass zu neuen Überlegungen. Es stellt sich die Frage, ob man eine schlecht kommunizierbare Frageform eintauschen möchte gegen eine eventuell willkürliche und von Anreizen zur Übertreibung geprägte Expertenmeinung. Diese Frage lässt sich nur in einer Reihe von neuen Experimenten klären. Da eine Bayes'sche A-Posteriori-Verteilung wiederum eine A-Priori-Verteilung eines neuen Experiments werden kann, lässt sich ein sequentielles Design denken, ähnlich dem, welches von Hanemann und Kanninen (2001) [116] zur Bestimmung des optimalen Gebotsdesigns vorgeschlagen wurde. Hierbei würde im ersten Schritt die Expertenmeinung eingeholt, anschließend würden die ohnehin notwendigen

Pretests zur Überprüfung der Fragebögen durchgeführt und ein Modell mit den Pretest-Resultaten und der A-Priori-Verteilung aus den Expertenmeinungen geschätzt. Die so entstehende A-Posteriori-Verteilung kann als A-Priori-Verteilung in eine zweite Staffel von Pretests eingehen. Das Verfahren bietet eine Möglichkeit, die ohnehin in mehreren Pretests anfallende Information systematisch in die endgültige Schätzung der Ergebnisse der Hauptbefragung einzubeziehen, ohne dass aus diesem Design Mehrkosten entstehen. Der Vorteil besteht vor allem darin, dass die Befragungen jeweils im einfachen Referendumsformat durchgeführt werden können, einem Format, das sich als anreizkompatibel und gut kommunizierbar herausgestellt hat, ohne dass man zu sehr großen Stichproben greifen muss, um die fehlende Effizienz der Schätzung auszugleichen. Der Einsatz von Bayes'schen Verfahren erlaubt es, sich zusätzliche Informationen aus anderen Quellen zu "borgen".

Im Zusammenhang mit simulierten Daten ist die Reduktion der Konfidenzintervalle vielleicht noch nicht so beeindruckend, denn diese sind so konstruiert, dass die Befragten immer die Wahrheit sagen, und es wurden positive Zahlungsbereitschaften für das zu bewertende Umweltgut angenommen und simuliert. In Daten aus einer tatsächlich durchgeführten Befragung hat man es teilweise mit höheren Ablehnungen des Projekts zu tun, und erhält so eine zusätzliche Unsicherheit der Ergebnisse. Daher werden im Folgenden die bis hierher beschriebenen Modelle auf die Daten einer im Sommer 2003 durchgeführten Befragung zur Zahlungsbereitschaft für ein Rekultivierungsprojekt in der Lausitz angewendet.

3.2 Die Bewertung des Cottbuser Ostsees

Die im nachstehenden Kapitel verwendeten Daten entstammen dem Forschungsprojekt "Ökonomische Bewertung von Änderungen der Umweltqualität mit Hilfe der Kontingenten Evaluierungsmethode" des Lehrstuhls für Volkswirtschaftslehre, insbesondere Umweltökonomie sowie Ordnungs-, Verbraucher- und Strukturpolitik der Universität Hohenheim, im Rahmen des an der BTU Cottbus angesiedelten Sonderforschungsbereichs 565 "Entwicklung und Bewertung gestörter Kulturlandschaften". Die Textauszüge zur Szenariobeschreibung entstammen den Originalfragebögen, die in dieser Studie verwendet wurden.

3.2.1 Ein neuer See in der Lausitz

Südlich von Berlin erstreckt sich über Brandenburg bis nach Sachsen mit der Lausitz das zweitgrößte Braunkohlefördergebiet Deutschlands. Braunkohle - immer noch zweitwichtigster Primärenergieträger - wird im offenen Tagebau gewonnen und ihre Förderung geht mit einer gewaltigen Landschaftszerstörung einher. Nach Beendigung des Kohleabbaus bleiben veränderte Flächen zurück, die bearbeitet und gestaltet werden müssen, damit sie anschließend verschiedenen Nutzungen zugeführt werden können. So müssen zunächst Flächen befestigt und gesichert werden, um dann darauf forstwirtschaftliche oder landwirtschaftliche Nutzungsgebiete wiederherzustellen sowie Erholungsräume zu schaffen. Dieser auch als Rekultivierung bezeichnete Prozess ist natürlich mit hohen Kosten verbunden, von denen nur die bergbaurechtliche Sicherung direkt den Bergbauunternehmen angelastet werden kann, den Rest tragen die umliegenden Gemeinden.

In der Vergangenheit, unter der DDR-Regierung, ließ man oftmals der Natur freien Lauf, d.h. die sogenannten Restlöcher füllten sich zum einen durch Regenwasser und zum anderen durch das für den Tagebau ständig abgepumpte und nun zurück drängende Grundwasser und es entstanden unweigerlich Seen. Die Wasserqualität dieser Seen ist allerdings durch die im Lausitzer Tagebau verbreitete Fördermethode, bei welcher buchstäblich das Unterste zuoberst gekehrt wird, stark beeinträchtigt. Bei der Braunkohleförderung mithilfe von großen Abbaubrücken werden die unten liegenden tertiären Sände emporgehoben und bilden anschließend die Schicht, die mit dem einströmenden Wasser in Kontakt kommt. Durch die in diesen Sänden enthaltenen chemischen Verbindungen sind die entstehenden Seen einer starken Gefahr ausgesetzt zu versauern. Alternativ werden heute anstelle dieser "minderwertigen" ungestalteten Seen für vielfältige Freizeitaktivitäten nutzbare gestaltete Seen geplant.

In der Studie wurde die geplante Rekultivierung eines derzeit noch in Betrieb befindlichen Tagebaus im Osten der brandenburgischen Stadt Cottbus zur Bewertung gestellt. Die genaue Beschreibung des Szenarios findet sich als Auszug aus dem verwendeten Fragebogen in Abbildung 3.4. Die in der Studie zur Bewertung gestellte Umweltveränderung besteht in der Anlage des Cottbuser Ostsees. Dieses Projekt beinhaltet zum einen Nichtgebrauchswerte in Form des geplanten Naturschutzgebiets, denn dieses wird in weiten Teilen den Befragten nicht zugänglich werden. Zum anderen werden durch die Anlage von Freizeit- und Bademöglichkeiten und durch die Gestaltung eines Strands und eines Bootshafens klassische Gebrauchswerte

Abbildung 3.4: Textauszug der Szenariodarstellung im Fragebogen

Östlich von Cottbus befindet sich das Tagebaugelände Cottbus-Nord, in dem noch bis 2015 Braunkohle abgebaut wird. Nach 2015 werden die 2700 Hektar Tagebaugelände rekultiviert. Dabei wird auch das Abpumpen des Grundwassers beendet, das zu einer Absenkung des Grundwasserspiegels stellenweise um bis zu 35 Meter geführt hat. Zweidrittel der Tagebaufläche sollen bis 2027 in einen See, den Cottbuser Ostsee, umgewandelt werden. Mit einer Fläche von 1900 Hektar wird der Cottbuser Ostsee der größte See in der Region sein. In einem ersten Schritt werden bei der Rekultivierung der See und die Randbereiche so gesichert, dass für Mensch und Tier keine Gefahr mehr besteht. Die Kosten für die Grundsicherung und die Flutung des Sees werden von LAUBAG und LMBV getragen, die auch für die Sicherung der Wasserqualität verantwortlich sind. LAUBAG und LMBV sind gesetzlich nur zu diesen Mindestmaßnahmen verpflichtet. Über die Mindestmaßnahmen hinaus wird eine weitergehende Gestaltung des Sees diskutiert. Für den Freizeitbereich sind Badestrände mit Duschen, sanitären Anlagen, Sportmöglichkeiten, Gaststätten und Kiosken, Grillplätzen, Campingplätzen u.s.w. am Cottbuser Ufer geplant. Am Strand bei Dissenchen und Schlichow soll eine begehbare Dünenlandschaft entstehen. Die Cottbuser Innenstadt soll über eine Seepromenade mit dem See verbunden werden. Der See selbst wird für wassersportliche Aktivitäten wie Baden, Segeln, Surfen, Rudern, Angeln und Tauchen nutzbar sein. Am Cottbuser Ufer soll außerdem ein Freizeithafen angelegt werden. Um den See herum wird ein Wander- und Fahrradweg führen, der in das bestehende Wegenetz einbezogen wird. Der See wird sowohl mit öffentlichen Verkehrsmitteln als auch mit privaten PKWs und mit dem Fahrrad gut erreichbar sein. Im östlichen Teil soll ein Naturschutzgebiet entstehen, das teilweise als Ruhezone für Tiere und Pflanzen dienen soll. In diesem Gebiet wird es zwar Wander- und Fahrradwege geben, Autos werden aber nicht zugelassen sein. Im Uferbereich sollen Inseln, Flachwasserzonen und ruhige Buchten entstehen, in denen sich Lebensräume für unterschiedliche Tiere und Pflanzen ausbilden können. Man hofft, dass sich dort Wasservögel, z.B. Flusseeeschwalben und Flussregenpfeifer, ansiedeln. Flach Buchten werden verschilfen, so dass ein Lebensraum für seltene Schilfrüher wie z.B. die Zwergrohrdommel entsteht. In den sumpfigen Uferwäldern dürften sich weitere Arten wie z.B. die Beutelmeise und der Laubfrosch ansiedeln. Viele dieser Tierarten sind gefährdet oder vom Aussterben bedroht. Mit dieser Gestaltung wird einerseits ein Freizeit- und Erholungsgebiet für den Menschen und andererseits eine Schutz- und Ruhezone für Tiere und Pflanzen angelegt.

geschaffen. Ziel der Studie war es neben methodischen Zielen zur Weiterentwicklung der Contingent Valuation Methode, den Wert zu messen, den die Bevölkerung dem angelegten See beimisst, wobei das alternative Szenario gewissermaßen im "Business-as-usual" Modell besteht, bei welchem ein umgestalteter See mit minderer Wasserqualität entstünde.

3.2.2 Ergebnisse der Befragung

Im Sommer 2003 wurden im Einzugsgebiet Cottbus sowie in den umliegenden Gemeinden 1012 Haushaltsvorstände befragt. Die Interviews dauerten ca. 40 Minuten und es wurde neben der eigentlichen Zahlungsbereitschaft eine Fülle von soziodemographischen Daten erhoben, sowie zahlreiche Fragen zur Einstellung gegenüber Umweltveränderungen, zur Einschätzung der eigenen Zufriedenheit und zur Lebenssituation der Bevölkerung gestellt.

Die Grundgesamtheit besteht aus 50.889 Haushalten, was 106.911 Einwohnern entspricht. Das in Berlin ansässige Meinungsforschungsinstitut USUMA wählte 209 Sample-Points, davon 202 in Cottbus. Ausgehend von diesen Punkten wurden im Random Route Verfahren je 8 Haushalte zufällig ausgewählt. Die Interviews wurden von speziell auf den Fragebogen geschulten Personen durchgeführt.

Tabelle 3.8 gibt eine Übersicht über die Verteilung der wichtigsten sozioökonomischen Daten in der Stichprobe im Vergleich zu der Struktur in Cottbus und in Brandenburg. Eine erste Sichtung der Daten vermittelt einen Eindruck von der Repräsentativität der Stichprobe. Eines der hauptsächlichen Probleme hierbei besteht allerdings darin, dass amtliche Statistiken nur selten in der Detailliertheit geführt werden, die notwendig wäre, um diese Daten für die von der Umweltveränderung betroffenen Bevölkerung zum Vergleich heranziehen zu können. Der "relevante Markt" für das Rekultivierungsprojekt "Cottbuser Ostsee" besteht in einem Gebiet um Cottbus herum, von dem aus die Anreise zum Cottbuser Ostsee noch sinnvoll ist. Es ist bei weitem nicht deckungsgleich mit dem gesamten Flächenland Brandenburg. Trotzdem lässt sich durch den Vergleich mit der Statistik des Landes Brandenburg ein Anhaltspunkt für die Repräsentativität der Befragung gewinnen. Die beobachtbaren Abweichungen lassen sich vor allem dadurch erklären, dass die Wahrscheinlichkeit nichtberufstätige Personen zu Hause anzutreffen natürlich höher ist, und somit die männliche Bevölkerung überrepräsentiert ist. Unterrepräsentiert sind

Tabelle 3.8: Stichprobe und amtliche Statistik im Vergleich

	Brandenburg (2003)	Stichprobe
	in 1000	
Anzahl	2 597,6	1,012
	in Prozent	
Männlich	49,4	53,6
Weiblich	50,6	46,3
Haushaltsgröße		
Einpersonenhaushalt	31,7	18,4
Mehrpersonenhaushalt		
mit 2 Personen	34,8	46,3
mit 3 Personen	18,9	20,3
mit > 3 Personen	14,5	15,0
Alter		
20 - 40	32,8	29,5
40 - 65	46,1	50,2
> 65	21,2	30,3
Erwerbstätigkeit		
Erwerbstätig	44,0	47,0
Ausbildung	9,2	8,0
Arbeitslos	18,8	11,0
Rentner	30,0	32,0

<http://www.lds-bb.de/sixcms/list.php/lds>, Landesbetrieb für Datenverarbeitung und Statistik Land Brandenburg.

die Einpersonenhaushalte, die entweder weniger bereit zu sein scheinen, sich einem Interview zu stellen oder auch weniger häufig zu Hause anzutreffen waren. Insgesamt kann man jedoch bei den meisten Eigenschaften von Repräsentativität im Bereich eines 10 prozentigen Konfidenzintervalls bei der vorliegenden Stichprobengröße sprechen. Die Befragung wurde im doppelten Referendumsformat durchgeführt.

Die erste Frage lautete: "Die erwähnten Mindestmaßnahmen, die ja von LAUBAG und LMBV finanziert werden, werden auf jeden Fall durchgeführt. Die beschriebene weitergehende Gestaltung des Sees würde jedoch vermutlich steigende Lebenshaltungskosten für die Haushalte mit sich bringen. Würden Sie dieser weitergehenden Gestaltung zustimmen, wenn dadurch die Lebenshaltungskosten Ihres eigenen Haushalts ab sofort um

monatlich 5 (10, 15, 20) € steigen würden? ... *ja*□ *nein*□

Bei einer Ja-Antwort wurde die folgende Nachfrage gestellt:

Würden Sie auch dann zustimmen, wenn die Lebenshaltungskosten Ihres Haushalts ab sofort um monatlich 7,50 (15, 20, 25) € steigen würden? *ja*□ ... *nein*□,

während die Nachfrage bei einer anfänglichen Nein-Antwort lautete:

Würden Sie ihr denn zustimmen, wenn die Lebenshaltungskosten Ihres Haushalts ab sofort nur um monatlich 2,50 (5, 10, 15) € steigen würden? ... *ja*□ *nein*□

Durch die Aufforderung "auch dann" zuzustimmen, wenn der Preis höher ist, wurde versucht, den Eindruck, dass der Staat hier versucht zu feilschen, zu vermindern. Inwieweit das gelungen ist, lässt sich aus den Antworthäufigkeiten zu den verschiedenen Geboten ablesen. Tabelle 3.9 fasst die Häufigkeiten der Antwortkombinationen nach Erstgeboten zusammen. Erwartungsgemäß fällt die Häufigkeit einer zustimmenden Antwort mit der Gebotshöhe. Im Vergleich mit den Simulationsdaten des letzten Abschnitts ist die Anzahl der ablehnenden Antworten deutlich höher, selbst bei geringen Geboten. In jeder empirischen Untersuchung ist die Zahl der ablehnenden Antworten typischerweise höher als in einer Simulation, die durchweg positive Zahlungsbereitschaften simuliert. Die Anzahl der zustimmenden Antworten in den Intervallen [10,15] und [15, 20] scheint von der Reihenfolge der Frage abzuhängen. Methodisch bleibt somit als Fazit auch aus dieser Studie festzuhalten, dass es trotz der aufgebotenen Sorgfalt bei der Formulierung der Zahlungsbereitschaftsfrage im doppelten Referendumsformat zu Verzerrungen bei den Antworten gekommen ist. Tabelle 3.10 zeigt diesen Effekt ganz deutlich. Die Wahrscheinlichkeit, eines von vier möglichen Erstgeboten präsentiert zu bekommen, war für alle Befragten annähernd gleich groß, wie die Zeilensumme über die Antworten in der

Tabelle 3.9: Antworten und Erstgebote

Gebote	YesYes	YesNo	NoYes	NoNo
5,00	53	54	26	120
10,00	31	51	41	126
15,00	32	31	24	168
20,00	24	24	15	192

Eigene Berechnungen.

Tabelle 3.10: Ankereffekte

	NoNo	NoYes	YesNo	YesYes	Insgesamt
WTP Interval	0 - 2,5	2,5 - 5	5 - 7,5	7,5 - ∞	253 Antw.
Häufigkeit	47,4%	10,3%	21,3%	21%	100%
WTP Interval	0 - 5	5 -10	10 -15	15 - ∞	249 Antw.
Häufigkeit	50,8%	16,4%	20,4%	12,4%	100%
WTP Interval	0 - 10	10 -15	15 -20	20 - ∞	255 Antw.
Häufigkeit	65,9%	9,4%	12,2%	12,5%	100%
WTP Interval	0 - 15	15 - 20	20 - 25	25 - ∞	255 Antw.
Häufigkeit	75,3%	5,9%	9,4%	9,4%	100%
Insgesamt					
Antworten	606	106	160	140	1012
Häufigkeit	59,9%	10,5%	15,8%	13,8%	100%

Eigene Darstellung.

Tabelle zeigt. Somit muss auch die Anzahl der Haushalte, deren Zahlungsbereitschaft zwischen 10 und 15 € liegt annähernd gleich groß sein. Dies ist jedoch nicht der Fall, wie man sieht, wenn man die Häufigkeiten in der zweiten Zeile und dritten Spalte mit denen in der dritten Zeile und zweiten Spalte vergleicht. Ein ähnliches Muster findet sich im Intervall zwischen 15 € und 20 €. Eine derartige Struktur in den Daten wird in der Literatur häufig als Ankereffekt bezeichnet, wie in Kapitel 1.3 ausführlich diskutiert wurde. Das Beispiel zeigt, dass das doppelt gebundene Format für die Befragten schwierig zu verstehen ist, auch wenn die Frageformulierung sehr vorsichtig und vermeintlich eindeutig gewählt wird. Dieser Ankereffekt wirkt sich auch auf die Qualität der Schätzung aus, wie man vor allem bei den Intervalldatenmodellen in Tabelle 3.11 beobachten kann. Es bleibt also fraglich, ob das doppelte Referendumsformat nicht zugunsten des verständlicheren einfachen Referendumsformat aufgegeben werden sollte.

3.2.3 Zahlungsbereitschaft für den Cottbuser Ostsee

Aus den Ergebnissen der Berechnung von Zahlungsbereitschaften mit simulierten Daten im vorangegangenen Abschnitt könnte man schließen, dass es ohnehin nur lohnend ist, sich auf das Probit-Modell oder das Logit-Modell in der Analyse zu konzentrieren. Allerdings ist die Datenstruktur hier eine andere, denn nun liegen den Daten nicht die wohlgezogenen Zahlungsbereitschaften einer simulierten Bevölkerung zu Grunde, sondern wir haben es mit einer realen Bevölkerung und einer Reihe von möglichen Verzerrungen zu tun. Zwei kritische Punkte haben sich bereits bei der ersten Analyse der Daten erschlossen: die hohe Anzahl von Äußerungen einer Zahlungsbereitschaft von Null und das vermutliche Vorliegen von Ankereffekten. Die Null-Antworten wurden auf das Vorliegen von Protestantworten hin untersucht, indem Korrelationen mit einer Ablehnung des Zahlungsinstruments gebildet wurden. Das Zahlungsinstrument scheint jedoch von allen Befragten akzeptiert und für plausibel gehalten worden zu sein. So liegt die Ursache der Null-Antworten entweder in einer sehr geringen Wertschätzung des Projekts oder in einer trotz der geringen Zahlungen als zu einschränkend empfundenen Budgetrestriktion, so dass bei vielen Befragten eher eine Zahlungs(un)fähigkeit als eine Zahlungsbereitschaft angegeben wurde. Diesem Effekt kann man bei der Modellierung Rechnung tragen, indem ein Modell gewählt wird, das ausschließlich positive Zahlungsbereitschaften zulässt. Wenn mit einer derartigen Datenstruktur ein symmetrisches Modell geschätzt wird, so kann es zu einer künstlichen Verringerung des Erwartungswerts der Zahlungsbereitschaft kommen.

Tabelle 3.11 zeigt die Ergebnisse der Berechnung der Zahlungsbereitschaft mit den bereits für simulierte Daten diskutierten Modellen. Es überrascht beinahe, dass die Ergebnisse - jedenfalls für den Median - sehr nahe bei einander liegen. Der Mittelwert des lognormalen Modells weicht erheblich von allen übrigen Schätzwerten ab, so dass er, trotz der anfänglichen Überlegungen, für unzutreffend gehalten wird. Die Tabelle gibt zum einen die geschätzten Parameter an, zum anderen die daraus berechenbaren Wohlfahrtsmaße als Mittelwert und als Median der Verteilung. Der Median der Verteilung gilt allgemein als das konservativere Maß, denn bei denjenigen Verteilungen, bei denen sich beide Größen berechnen lassen und sich voneinander unterscheiden, ist der Median immer niedriger als der Mittelwert. Gemessen am Endwert der Likelihood-Funktion, einem der Gütekriterien zum Modellvergleich (vgl. Kapitel 2.2.2), schneidet das Weibull-Modell am besten ab, es weist den niedrigsten Wert für die erwartete Zahlungs-

Tabelle 3.11: Zahlungsbereitschaftsanalyse Cottbuser Ostsee

	Logit	Probit	Loglogistisch	Lognormal	Weibull
α	0,6233	0,3465	1,835	1,1060	0,2218
β	-0,1419	-0,0816	-1,242	-0,7521	-0,1322
	WTP[€]				
Mittelwert	4,39	4,24	19,08	8,46	- ^a
	[3,66 - 5,23] ^b	[3,25 - 5,17]	[14,29 - 32,64]	[7,89 - 9,26]	- ^a
Median	4,39	4,24	4,44	4,35	3,71
	[3,66 - 5,23] ^b	[3,25 - 5,17]	[4,14 - 4,94]	[4,01 - 4,817]	[1,81 - 8,36]
Likelihood	-1140.8	-1144.4	-1161.8	-1156.3	-996.1

^a Berechnung des Mittelwerts aus der Weibull-Verteilung für $\beta > 1$ nicht möglich.
^b Konfidenzintervalle nach der Bootstrap-Methode.
 Eigene Berechnungen.

bereitschaft auf. Allerdings hat dieser Schätzer ein extrem breites Konfidenzintervall, so dass er für die politische Beratung nahezu aussagelos ist. Schließlich würde man wohl Erstaunen ernten, wenn das Ergebnis einer Studie besagt, dass die Zahlungsbereitschaft für den Cottbuser Ostsee bei 50.889 Haushalten zwischen 92.100 und 425.400 € pro Monat beträgt. Die Unsicherheit, mit der der Wert behaftet ist, lässt sich aus der Modellstruktur des Weibull-Modells erklären. Als eines der Survivalmodelle basiert es auf der Intervallstruktur der Daten und reagiert auf die durch den Ankereffekt hervorgerufene "unlogische" Besetzungszahl in den Intervallen [10, 15] und [15, 20]. Das Probit-Modell hingegen ist, wie wir im vorigen Abschnitt gesehen haben, eine restringierte Form des bivariaten Probit-Modells², bei dem die Möglichkeit unterschiedlicher Zahlungsbereitschaften bei der ersten und der zweiten Frage gar nicht zugelassen wird.

Natürlich ist man in einer solchen Studie auch immer daran interessiert, welche der erhobenen Kovariate einen positiven oder negativen Einfluss auf die Zahlungsbereitschaft ausüben. Hierfür werden verschiedene Modellspezifikationen geschätzt und die Signifikanz des Einflusses verschiedener Größen untersucht.

²Das bivariate Modell führt zumindest für die lineare Nutzenspezifikation zu sinnlosen Resultaten. Für eine weitere Analyse vgl. Ahlheim et al. (2004) [7].

Tabelle 3.12: Modelle mit Kovariaten

	Koeffizient	Std. Abw.	$P[Z > z]$
Modell I			
Konstante	-0,17	0,28	0,5592
Gebot	-146,55	7,22	0,0000
Einkommen	158,06	26,63	0,0000
Alter	-1,04	4,16	0,8022
Bildung	23,58	25,42	0,3535
WTP	4,45	[3,56 - 5,28]	
Likelihood		-1108,71	
Modell II			
Konstante	-0,61	0,50	0,1768
Gebot	-148,32	7,35	0,0000
Einkommen	148,52	25,05	0,0000
Staat	-0,39	0,16	0,0187
Wirtschaft/ Sorge	-100,35	52,25	0,0548
Heimat	0,25	0,64	0,0001
WTP	4,41	[3,51 - 5,25]	
Likelihood		-1098,04	
Modell III			
Konstante	-1,12	0,45	0,0124
Gebot	-148,30	7,34	0,0000
Einkommen	145,91	24,99	0,0000
Staat	-0,41	0,16	0,0124
Umwelt/ Sorge	0,46	0,35	0,1807
Heimat	0,19	0,65	0,0026
WTP	4,38	[3,53 - 5,18]	
Likelihood		-1098,83	

Eigene Berechnungen.

Bei der Fülle der erhobenen Einflussvariablen kann man die Antworten auf einige Fragen zunächst mithilfe der Faktoranalyse zu bedeutungstragenden Faktoren aggregieren. Eine ausführliche Darstellung würde an dieser Stelle zu weit führen und ist auch für das Anliegen dieser Arbeit nicht relevant (für ein Modell, das auf der Faktoranalyse beruht, vgl. Ahlheim et al. (2004) [7]). Daher werden nachstehend drei Modelle aufgeführt, die verschiedene nicht-aggregierte Variablen als Kovariate enthalten. Modell I verwendet die "klassischen" Einflussgrößen: Einkommen, Alter und Bildungsstand der

Befragten. In diesem, so wie in anderen, hier nicht aufgeführten Modellen stellt sich das Alter der Befragten als nicht-signifikante erklärende Variable heraus, was zunächst verblüfft, da das Projekt ja doch einen erheblichen Zeitraum umfasst und für die Befragten, die älter sind als 65 Jahre, fast gar nicht mehr erlebbar ist. Es lässt sich vermuten, dass der Nichtgebrauchswert an die Stelle des Gebrauchswerts getreten und der Cottbuser Ostsee aus einem Vererbungsmotiv heraus bewertet worden ist. Diese Hypothese lässt sich jedoch aus den erhobenen Daten nicht belegen.

Modell II untersucht die Auswirkungen der Einstellung der Befragten zur Bereitstellung öffentlicher Güter, ihrer Sorgen zur allgemeinen wirtschaftlichen Lage und ihrer Einstellung zur Notwendigkeit des Erhalts der Heimat ("Bitte sagen Sie uns, in welchem Maß Sie die folgenden Aussagen ablehnen oder befürworten. (...) Die landschaftliche Schönheit und Eigenart unserer Heimat sollte erhalten und geschützt werden."). Modell III bezieht die Sorge um die Situation der Umwelt in das Modell ein. Während das Einkommen erwartungsgemäß mit einem positiven Vorzeichen und signifikant auf allen Niveaus zur Erklärung des Erwartungswerts der Zahlungsbereitschaft beiträgt, wirkt sich die Einstellung zur Bereitstellung öffentlicher Güter durch den Staat negativ aus, d.h. je stärker die Befragten befürworten, dass bestimmte Güter durch den Staat bereitgestellt werden (hier wurde die Befürwortung staatlicher Universitäten modelliert), desto geringer ist ihre Zahlungsbereitschaft. In Modell II und III ist dieser Einfluss signifikant auf dem 2% Niveau. Die Sorge um die allgemeine wirtschaftliche Situation führt ebenfalls zu einer geringeren Zahlungsbereitschaft, sie ist jedoch nur signifikant auf dem 10% Niveau. Die Sorge um den Zustand der Umwelt erhöht zwar die Zahlungsbereitschaft, ist jedoch in 18% der Fälle zurückweisbar. Die Wertschätzung der Heimat hingegen ist in beiden Modellen signifikant und führt zu einer Erhöhung der Zahlungsbereitschaft. Alle Modelle sind sehr ähnlich in der Schätzung der erwarteten Zahlungsbereitschaft, den Konfidenzintervallen und auch im Likelihood-Endwert. Die gesamte jährliche Zahlungsbereitschaft für den Cottbuser Ostsee berechnet sich aus dem Erwartungswert der Haushaltszahlungsbereitschaften multipliziert mit 12 Monaten und der Anzahl aller Haushalte im Einzugsgebiet zu rund 223.000 €/ Monat oder 2,7 Mio. €/ Jahr.

Kapitel 4

Zwischenfazit

Es ist ein weiter Weg von dem Wunsch, der "Natur einen Wert" zu geben, bis hin zur tatsächlichen Berechnung dieses Werts mithilfe eines der in den vorangegangenen Kapiteln dargestellten Verfahren. Wählt man ein mit der Nutzentheorie vereinbares Maß für die individuelle Nutzenänderung, das die Richtung der Nutzenänderung zuverlässig anzeigt (Indikatorbedingung) und operationalisierbar ist, so sollte man sich für eines der Hicks'schen Maße entscheiden. Will man darüber hinaus neben den Gebrauchswerten eines Umweltguts auch die Nichtgebrauchswerte in die Messung einbeziehen, so kommt man um die Anwendung der Contingent Valuation Methode nicht herum. Aber auch wenn diese Entscheidung gefallen ist, steht man einer Vielzahl von verwendbaren Modellen gegenüber, mit deren Hilfe versucht wird, vielen der in der Realität beobachteten Besonderheiten gerecht zu werden. Aus Sicht des eigentlichen Nutzers von Kosten-Nutzen-Analysen ergibt sich die Frage, wie die Resultate der einzelnen Modelle zu bewerten sind und wie belastbar die Ergebnisse sind. Wohl auch aus dieser Unsicherheit heraus beobachtet man in der praktischen Umweltpolitik einen pluralistischen Ansatz gegenüber dem Einsatz von Bewertungsmethoden. Neben dem Einsatz von nicht-monetären Methoden, die die Umweltveränderungen eines geplanten Projekts auflisten und beschreiben, werden eine Vielzahl von ad hoc Ansätzen verwendet, die in dieser Arbeit nicht diskutiert wurden (für eine Übersicht vgl. Garrod und Willis (1999) [98].)

In den europäischen Ländern findet man einen unsystematischen ad hoc Einsatz von Bewertungsmethoden und bislang keine Rahmenbedingungen für den gezielten Einsatz von Umweltbewertung als Instrument zur Unterstützung der Entscheidungsfindung. So wurden in der Schweiz etliche Studien durchgeführt, die neben der Bewertung des Umweltguts ein metho-

disches Erkenntnisziel haben (vgl. Garrod und Willis (1999) [98]) und versuchen, die Contingent Valuation Methode zu verbessern. Demgegenüber wurden in Deutschland eher politikorientierte Studien durchgeführt (vgl. CVM@Network (2005) [3]). Von ca. 40 seit 1985 im deutschsprachigen Raum durchgeführten Studien ist etwa ein Drittel im direkten politischen Auftrag entstanden, weitere 20 haben als Ziel der Untersuchung die Bestimmung des Werts einer bestimmten politischen Maßnahme und nicht die methodische Weiterentwicklung der CVM benannt. Die norwegische Regierung hat etliche CVM-Studien in Auftrag gegeben und scheint auch an einem weiteren Einsatz der Methode für Kosten-Nutzen-Analysen interessiert zu sein (vgl. Navrud und Strand (1992) [236]). Die britische Regierung verfolgt einen ähnlichen Ansatz, jedoch ist ein systematischer Einfluss der Umweltbewertung auf die politische Entscheidungsfindung, bei dem die Ergebnisse von Bewertungsstudien genutzt werden, auch hier nicht sichtbar (Garrod und Willis (1999)[98]). Die Asian Development Bank, die Weltbank und die Interamerican Development Bank hingegen führen systematisch CVM Studien zur Bewertung von geplanten Projekten durch und Forscher an den verschiedenen Einrichtungen haben erheblich zur Weiterentwicklung der Methode beigetragen (vgl. Whittington et al. (1992) [252] oder Ardila et al. (1998) [17]). Der einzige systematische und institutionalisierte Fall einer Vorschrift, die Verfahren zur Kosten-Nutzen-Analyse erfordert und explizit die Verwendung von Reisekostenmethode, hedonischen Preisen oder der Contingent Valuation Methode vorschreibt, ist die Bewertung von Umweltschäden unter CERCLA in den USA; wobei auch diese Bestimmung, wie in Kapitel 1 beschrieben, inzwischen etwas zurückgenommen wurde (vgl. EPA (2005) [93]).

Ein wesentliches Hindernis für eine häufigere Anwendung der CVM scheint die Tatsache zu sein, dass man für eine Studie, die glaubwürdige Ergebnisse liefert, eine große Stichprobe braucht. Während die scheinbar natürlichste Umsetzung der Tatsache, dass man die Zahlungsbereitschaft von Individuen mit einer Interviewmethode ermitteln möchte, - die offene Befragung - die Befragten scheinbar vor unlösbare Probleme stellt, führt der Ansatz, die Befragten mit einem Referendum zu konfrontieren, zu statistischen Effizienzverlusten. Die Lösung des statistischen Problems wiederum, die Verwendung eines Formats, bei dem man Intervalle der Zahlungsbereitschaft abfragt, verwirrt die Befragten und führt zu widersprüchlichen Antworten. Modelle, die versuchen, die Abhängigkeit der Antworten von der Art der Fragestellung abzubilden, lassen sich schwer interpretieren, denn welche Zahlungsbereitschaft soll der Politiker letztendlich als die "wirkli-

che Zahlungsbereitschaft" in seine Kosten-Nutzen-Analyse einbeziehen? So bleibt das einfache Referendumsformat als einziges anreizkompatibles und verzerrungsfreies Format bestehen, und die statistischen Effizienzverluste lassen sich scheinbar nur durch eine große Stichprobe ausgleichen. Auch die Abkehr von den teilweise strengen Vorgaben der Verteilungsannahmen in den nicht-parametrischen Verfahren hilft hier nicht weiter. Im Gegenteil, durch den Verzicht auf eine (stützende) Verteilungsannahme, sind die nicht-parametrischen Verfahren noch auf eine größere Datenmenge und somit auf größere Stichproben angewiesen. In den Augen des Politikers ist Stichprobengröße jedoch gleichbedeutend mit "zu teuer", was dazu führt, dass häufig Studien durchgeführt werden, die mit einem geringen Stichprobenumfang zu wenig belastbaren Aussagen führen. An dieser Stelle muss dann wirklich mit Diamond und Hausman (1994) [87] gefragt werden "Is some number better than no number?".

Eine alternative Lösungsstrategie besteht darin, den Rahmen der klassischen Statistik zu verlassen und Bayes'sche Methoden zu verwenden. Durch die Möglichkeit im Bayes'schen Modell, sich Informationen aus anderen Quellen zu "borgen", kann die Aussagefähigkeit einer kleineren Menge von Beobachtungen erhöht werden. Informationen lassen sich beispielsweise aus Expertengesprächen oder Pretests gewinnen, ohne dass hierfür zusätzliche Kosten entstehen, denn diese Gespräche werden ohnehin bei der Szenarioentwicklung im Vorfeld einer CVM Studie geführt. Diese Vorgehensweise hat den Vorteil, dass sie den Rahmen der Theorie nicht verlässt, denn für das eigentliche Modell werden nach wie vor nur solche Verteilungen verwendet, die mit der Nutzentheorie vereinbar sind. Der "Bayes'sche Bonus" kommt bei der Überlegung ins Spiel, ob man nicht vor der Erhebung der Daten schon etwas über die Parameter weiß - und wie man dieses Wissen in das Modell integrieren kann. Erfahrungen mit dem Bayes'schen Ansatz zur Auswertung von Contingent Valuation Daten sind bislang nicht vorhanden, in dieser Arbeit wurden erste Versuche anhand simulierter Daten durchgeführt. Im letzten Kapitel wurde ein Vorschlag entwickelt, wie die Ergebnisse einer Expertenbefragung in die endgültige Schätzung einfließen können. Für simulierte Daten wurde dieses Verfahren angewendet, aber letztlich kann man dem Beispiel Willkür unterstellen, denn auch die Experten sind natürlich simulierte Experten. Vorbehaltlich dieser Zweifel sind die Ergebnisse der Simulation vielversprechend. Die Aussagekraft einer CVM Befragung konnte erhöht werden gegenüber der Auswertung derselben Befragung mit klassischen Methoden.

Teil II

Nutzentransfer

Kapitel 5

Klassischer Nutzentransfer

5.1 Der Nutzentransfer als Instrument der Umweltbewertung

Angesichts knapper staatlicher Budgets auf der einen Seite und einer inzwischen großen Zahl bereits durchgeführter Bewertungsstudien auf der anderen Seite, liegt der Wunsch auf der Hand, den an einem Ort ermittelten "Nutzen" eines umweltverändernden Projekts zur Bestimmung des Werts einer an einem anderen Ort geplanten Umweltveränderung zu verwenden. Als Nutzentransfer wird im Folgenden jegliches Verfahren bezeichnet, mit dem es gelingt, Informationen aus bereits durchgeführten Studien (Primärstudien) zur Bewertung einer geplanten Umweltveränderung (Sekundärstudie) heranzuziehen. Es wird sich zeigen, dass unter dieser Überschrift verschiedene Verfahren zu finden sind. Als Nutzentransfer lässt sich die einfache Übertragung ermittelter Werte auf ähnliche Projekte verstehen; aber auch die gezielte Verarbeitung von Informationen aus dem Gebiet der Primärstudie *und* der Sekundärstudie wird als Nutzentransfer bezeichnet. Heutige Verfahren sind von einer zunehmenden Berücksichtigung von zusätzlichen Informationen über die Besonderheiten der geplanten Umweltveränderung und der davon betroffenen Bevölkerung gekennzeichnet.

Der ursprüngliche Anstoß zum Nutzentransfer kommt aus der Praxis und wurde vorangetrieben vom Kostensenkungsargument. Lange Zeit wurde zu den Möglichkeiten und Grenzen des Nutzentransfers auch wenig geforscht

und veröffentlicht; er wurde vielmehr als "educated guess" einfach durchgeführt. Dies hat dazu geführt, dass dem Nutzentransfer oftmals noch ein Hauch des Unseriösen anhaftet. Noch heute werden Nutzentransfermethoden häufig als "Quick and Dirty" bezeichnet¹, obwohl die wissenschaftliche Auseinandersetzung mit dem Nutzentransfer sich bereits über mehrere Jahre erstreckt. In einer frühen wissenschaftlichen Auseinandersetzung beobachten Brookshire und Neill (1992) [46] drei Kategorien von Nutzentransferanwendungen:

- auf Erfahrung basierende Schätzungen,
- auf beobachtetem Verhalten basierende Schätzungen und
- auf Erhebungen basierte Schätzungen.

Nutzentransfers der ersten Kategorie bezeichnet Smith (1992) [229] als "economic alchemy (i.e., a set of judgemental adaptations passed from "master" to "apprentice" that are influenced by who is in control)." Ein erster umfassender Versuch, den Nutzentransfer aus der Alchemistenwelt zu befreien, und somit eine der ersten systematischen Auseinandersetzungen mit den Möglichkeiten und Grenzen des Nutzentransfers wurde 1992 im Rahmen einer Artikelserie in *Water Resources Research* unternommen. Hier wurde der aktuelle Stand der Technik dargestellt (wie beispielsweise durch Luken et al. (1992) [179] in einer Studie zur Abwasserkontrolle in der Papierindustrie) sowie eine Systematisierung bestehender Ansätze vorgenommen. Darüber hinaus wurden Designvorschläge für Überprüfungsstudien zur Gültigkeit von Nutzentransfers entwickelt. So stellen Boyle und Bergstrom (1992) [41] eine Forschungsagenda für zukünftige Nutzentransferstudien auf und führen gleichzeitig Richtlinien für die Durchführung von Transferstudien ein. Diese Richtlinien stellen erhebliche Anforderungen an die Vergleichbarkeit des Primär- und Sekundärstudiengiebts. So müssen die Charakteristika der Bevölkerung in beiden Gebieten ähnlich sein und die Umweltgüter sollten in Größe und Art vergleichbar sein. Falls diese Ähnlichkeiten gegeben sind, so die Vermutung, lässt sich auch eine ähnliche Zahlungsbereitschaft in beiden Gebieten erwarten und ein Nutzentransfer ist gerechtfertigt. Desvousges et al. (1992) [85] entwickeln in ihrem Beitrag etwas weniger rigorose Auswahlkriterien für zum Nutzentransfer geeignete Studien, finden jedoch ebenfalls heraus, dass die zu diesem Zeitpunkt für

¹So lautet zum Beispiel der Titel der Arbeit von Barton (1999) [24] zum Nutzentransfer im ersten Teil "The Quick, the Cheap and the Dirty".

das von ihnen gewählte Beispiel (Bewertung der Wasserqualität) vorliegenden Primärstudien aus mehreren Gründen für einen Transfer ungeeignet sind. Zum Transfer geeignete Studien sollen

- auf adäquaten Daten basieren, d.h. die Untersuchung ist methodisch einwandfrei und folgt beispielsweise den Richtlinien des NOAA Panels (1993) [18]), und zuverlässige statistische Techniken benutzen,
- ähnliche Umweltveränderungen bewerten wie die Transferstudie,
- Regressionen enthalten, die einen Zusammenhang herstellen zwischen der Zahlungsbereitschaft und den sozioökonomischen Charakteristika der untersuchten Bevölkerung,
- Regressionen enthalten, die einen Zusammenhang herstellen zwischen den Charakteristika des Umweltguts und der ermittelten Zahlungsbereitschaft und
- insgesamt ähnliche "Märkte" bewerten.

Der Sonderband in *Water Resource Research* hat die Diskussion um den Nutzentransfer befruchtet und die Idee einer zentralen Bibliothek für Umweltbewertungsstudien geboren. Auf der Jahreskonferenz der Association of Environmental and Resource Economists im folgenden Jahr wurden bereits die Grundzüge einer Nutzentransferdatenbank diskutiert. Zwischen 1994 und 1997 wurde diese Datenbank unter dem Einfluss etlicher Experten² auf dem Gebiet der Umweltbewertung eingerichtet und fortentwickelt. Inzwischen stellt das Environmental Valuation Reference Inventory[©] (EVRI[©]) eine Fülle von Studien bereit, explizit mit dem Ziel "(...) a tool to help policy analysts use the benefits transfer approach (...)" zu sein. "Using the EVRI[©] to do a benefits transfer is an alternative to doing new valuation research (...)" [1].

Darüber hinaus wurde der Nutzentransfer ein wenig aus seiner anrüchigen Schmutzdecke befreit und ist seitdem Gegenstand zahlreicher Publikationen gewesen. Allerdings stagnierte die methodische Weiterentwicklung über zehn Jahre und man widmete sich einer Verbesserung der Überprüfungsstudien sowie der Fortentwicklung von Testverfahren für bereits durchgeführte Transfers. Die Standardverfahren der Übertragung eines Punktschätzers

²Unter anderem wirkten Richard Bishop, Kerry Smith, Wiktor Adamowicz, Richard Carson, William Desvousges, John Loomis, Robert Rowe, Michael Welsh, Rick Freeman, Alan Krupnick, W. Desvousges, M. Jay, V. Kibler, J. Loomis, A. Olivier und Jason Shogren an der Entwicklung mit.

oder einer Bewertungsfunktion schienen lange Zeit die einzige Möglichkeit zu sein, einen Nutzentransfer durchzuführen. Beide Verfahren sind jedoch dadurch gekennzeichnet, dass sie keine bis wenig Informationen von der geplanten Umweltveränderung und vor allem von der davon betroffenen Bevölkerung einbeziehen können, wie nachfolgend gezeigt wird. Erst in den letzten zwei Jahren wurden - abgesehen von einer in der Literatur weitgehend ignorierten Studie von Parsons und Kealy (1994) [201] - durch den Übergang auf Bayes-Verfahren die Grenzen dieser herkömmlichen Modelle durchbrochen.

5.2 Herkömmliche Verfahren

Bevor auf die herkömmlichen Verfahren des Nutzentransfers im Einzelnen eingegangen wird, muss an dieser Stelle ein allen Verfahren gemeinsames Caveat aus theoretischer Sicht ausgesprochen werden. Unabhängig davon, welches der nachfolgend beschriebenen Verfahren eingesetzt wird, findet bei der Durchführung von Nutzentransfers eine Verschiebung von der "Nutzenmessung" schon fast zu einer "Nutzenvermutung" statt. In den bis hierher beschriebenen Verfahren zur Umweltbewertung, besonders aber bei der Contingent Valuation Methode, werden Zahlungsbereitschaften von Individuen aufgenommen, die bei einer sauber durchgeführten Studie zumindest einen messbaren Indikator für die individuelle Nutzenänderung darstellen. Damit wäre im Prinzip das Ende theoriebasierter Aussagen erreicht, denn schon die Aggregation nach dem Hicks-Kaldor Kriterium führt, wie weiter oben gezeigt, zu einer mit der ordinalen Nutzentheorie nicht vereinbaren Aussage. Selbst die Auswertung einer Primärstudie geht jedoch noch einen Schritt weiter und vermutet, dass der aus dieser Aggregation berechnete mittlere Wert sich auf alle betroffenen und nicht befragten Individuen verallgemeinern lässt. Schon hier wird unterstellt, dass die Präferenzen aller zumindest so ausgeprägt sind, dass sie bei Befragung im Mittel eine ähnliche Zahlungsbereitschaft geäußert hätten. In Nutzentransferstudien wird diese "Ähnlichkeit" noch mindestens eine Stufe weiter getrieben, denn man überträgt nun den Wert von ähnlichen Projekten auf eine neues Projekt - und impliziert damit, dass Individuen, die sich einer Umweltveränderung an anderer Stelle (und zu einer anderen Zeit) gegenüber sehen, auch ähnlich reagieren.

Eine theoretisch saubere Möglichkeit solcher Transfers ist also nicht gegeben, andererseits lässt die Wirklichkeit die Nutzenmessung bei allen Indivi-

duen und zu allen Umweltveränderungen nicht zu. Rosenberger und Loomis (2001) [217] beschreiben den Nutzentransfer als "second-best"-Strategie und merken an, dass die "worst-best"-Strategie darin besteht, den Wert einer Umweltveränderung überhaupt nicht zu ermitteln, da dies impliziert, der Wert sei Null. So kann die Lösung nur darin bestehen, die Schritte einer Nutzentransferstudie so nachvollziehbar und so kontrollierbar zu gestalten, dass ihre Ergebnisse sich zutreffend einschätzen lassen. Diese Kriterien sollen im Folgenden auch die Messlatte zur Qualitätsbeurteilung der einzelnen Verfahren darstellen.

5.2.1 Value Transfer

Der Value Transfer, d.h. die Übertragung des Zielwerts einer Primärstudie auf ein neues Projekt, stellt die einfachste Form des Nutzentransfers dar. Man bezeichnet diese Übertragung auch als Punktschätzerübertragung, denn der in der Primärstudie geschätzte Median oder Mittelwert wird auf die Bewertung einer neuen Umweltveränderung an einer anderen Stelle übertragen. Zunächst muss demzufolge die Umweltveränderung in studienübergreifend vergleichbaren Einheiten formuliert werden. Danach gilt es, eine Primärstudie zu finden, bei der eine ähnliche Umweltveränderung bewertet wurde. Aus dieser Primärstudie muss sich ein Wert je Einheit Umweltveränderung ablesen lassen. Eine derartige Einheit kann zum Beispiel bei Erholungswerten die Anzahl der in einem bestimmten Gebiet verbrachten Tage sein. Der der Primärstudie entnommene Wert wird dann mit der Größe der vermuteten Änderungen im Transfergebiet multipliziert. Die zentrale Bedeutung fällt dabei der Messung von "Ähnlichkeiten" zwischen zwei zu bewertenden Umweltveränderungen und zwischen den hiervon betroffenen Bevölkerungen zu. Um nicht völlig willkürlichen Expertenmeinungen ausgeliefert zu sein, bedarf es hierzu einer standardisierten Zusammenschau möglichst aller bereits vorgenommenen Bewertungsstudien (vgl. Kapitel 5.2.2 für eine Diskussion der Problematik bei derartigen Meta-Analysen). Die bereits erwähnte und weiter unten genauer beschriebene Datenbank EVRI[©] stellt Ansätze einer solchen Zusammenschau dar. Bei Rosenberger und Loomis (2001) findet sich (als Fortführung der Untersuchung von Walsh et al. (1992) [247]) eine weitere beeindruckende Übersicht über 163 Studien mit insgesamt 760 Maßen zu einzelnen Erholungsaktivitäten.

Angenommen, die Befragung zur Bewertung des Cottbuser Ostsees sei nicht durchgeführt worden, welchen Wert hätte man mithilfe des Nut-

zentransfers als Value Transfer ermitteln können? In einem ersten Schritt muss hierfür eine möglichst ähnliche Studie gefunden werden. In der Internetplattform und Datenbank EVRI[©] wurden mehr als 1400 Studien bislang zusammengetragen und systematisch erfasst (Stand Mai 2005). Die in EVRI[©] vorhandenen Studien sind nach verschiedenen Kenngrößen geordnet, darüber hinaus besteht die Möglichkeit zur Freitextsuche in den Titeln und Abstracts der erfassten Studien.

Da der Ausgangspunkt eines jeden Nutzentransfers die Wahl einer oder mehrerer Primärstudien ist, wird auf die Wahl der Primärstudie zum Nutzentransfer für den Cottbuser Ostsee an dieser Stelle ausführlich eingegangen. Zunächst muss man sich darüber klar werden, nach welchen Schlüsselbegriffen die "Ähnlichkeit" zwischen Primär- und Sekundärstudienggebiet definiert ist. Tabelle 5.1 gibt eine Übersicht über die vordefinierten Suchbegriffe in EVRI[©]. Die Rubriken "Environmental Goods and Services, and Environmental Assets" und "Environmental Stressor" sowie die geographischen Charakteristiken lassen sich darüber hinaus durch selbst definierte Begriffe eingrenzen. Eine geographische Eingrenzung auf Deutschland führt allerdings zu enttäuschenden Ergebnissen: es sind nur 10 deutsche Studien vertreten, von denen sich keine auch nur annähernd mit einem ähnlichen Gut befasst. Insgesamt dominieren die in Nordamerika (insgesamt 809 erfasste Studien) und darunter die in den USA durchgeführten Studien, was mit dem Eindruck übereinstimmt, den man auch aus der Literatur gewinnt. Somit wird zu einem Value Transfer auch eine Studie aus einem anderen Land herangezogen werden, wobei die dort ermittelten Werte auch durch Wechselkurse, oder besser durch Kaufkraftparitäten umgerechnet werden müssen³. Da sich in geographischer Hinsicht keine Übereinstimmung mit der Bewertung des Rekultivierungsprojekts in Cottbus finden lässt, müssen in anderer Hinsicht ähnliche Studien herangezogen werden⁴. Zwei wichtige Merkmale der Studie, an denen sich die gewünschte Ähnlichkeit verankern lässt, sind das zu bewertende Gut an sich und die gewählte Bewertungsmethodik. Da die Reisekostenmethode, die Methode der hedonischen Preise und die Contingent Valuation Methode sich in ihren Eigenschaften unterscheiden, vor allem in ihren Fähigkeiten Gebrauchswerte und Nichtgebrauchswerte zu messen, sind auch die mit den jeweiligen Methoden ermittelten Werte nur schlecht zu vergleichen oder sogar in ein-

³Zur Verwendung von Kaufkraftparitäten und Inflationsratenausgleich beim Nutzentransfer vgl. Patanayak et al. (2002) [202] oder Holm-Müller und Muthke (2004) [128].

⁴Die einzige Studie, die sich auch mit Rekultivierung befasst, scheidet aufgrund ihres Alters aus (vgl. Randall et al. (1978) [208]).

Tabelle 5.1: Aufbau einer Suchabfrage in EVRI

Type of Environmental Good or Service	<ul style="list-style-type: none"> ● Built environment, ● ecological functions, ● extractive uses, ● human, 	<ul style="list-style-type: none"> ● health, ● non extractive uses, ● passive uses.
Environmental stressor:	<ul style="list-style-type: none"> ● Non-toxic substance, ● persistent substance, ● anthropogenic substance, ● resource extraction, ● toxic substance, ● biotechnologies, 	<ul style="list-style-type: none"> ● micro-organisms, ● congestion / crowding, ● infrastructure development ● habitat conversion, ● solid waste, ● noise.
General Environmental Asset:	<ul style="list-style-type: none"> ● Air general, ● land general, ● water general, ● infrastructure / man-made, 	<ul style="list-style-type: none"> ● animals, ● micro-organisms, ● plants.
Geographic Characteristics:	<ul style="list-style-type: none"> ● Continent, ● country, 	<ul style="list-style-type: none"> ● state of development.
Economic Measure:	<ul style="list-style-type: none"> ● Price, ● compensating surplus, ● cost of injury/replacement, ● equivalent surplus, 	<ul style="list-style-type: none"> ● equivalent variation, ● compensating variation, ● willingness to accept, ● willingness to pay.
Valuation technique:	<ul style="list-style-type: none"> ● Count data models, ● market price of output, ● averting behavior , ● change in productivity, ● choice experiment, ● revealed + stated preferences, ● conjoint analysis, ● contingent ranking, ● CVM - dichotomous choice, ● CVM - open ended, 	<ul style="list-style-type: none"> ● CVM - payment card, ● CVM - iterative, ● demand analysis, ● hedonic property, ● hedonic wage, ● replacement costs, ● TCM multi-site, ● TCM - single site ● TCM - regional, ● TCM - RUM.
Year of Data		
Type of Study:	<ul style="list-style-type: none"> ● Primary, ● secondary/ benefit transfers, 	<ul style="list-style-type: none"> ● meta analysis / synthesis analysis.
Eigene Zusammenstellung.		

ander umzurechnen. Das Cottbuser Rekultivierungsprojekt enthält sowohl Gebrauchswerte (Freizeitaktivitäten am See, Wassersport) als auch Nichtgebrauchswerte (Naturschutzgebiet), so dass sich die Suche im Folgenden auf Studien konzentriert, bei denen die Contingent Valuation Methode angewendet wurde.

Eine Suche in EVRI[©] mittels der Eingrenzung "passive use values", "fresh water", "endangered species" und "contingent valuation" führt zu 89 Treffern, wobei die Beschreibung "fresh water" auch Studien zu Flussgebieten enthält. Aktivitäten an Flüssen und die damit verbundenen Problematiken sind so verschieden vom Cottbuser Ostsee, dass alle diese Studien ausscheiden. Ferner wurden Studien entfernt, die vor 1990 erschienen sind bzw. durchgeführt wurden, denn die Bewertungsmethode hat sich weiterentwickelt, das Bewusstsein für Umweltschäden sich verändert und die statistischen Auswertungsmethoden haben sich auch seit ca. 1990 erheblich verbessert⁵. Nach diesen Überlegungen bleiben noch 26 Studien in der engeren Wahl. Der Einschluss "endangered species", der bewusst gewählt wurde, um den Naturschutzaspekt des geplanten Sees auch in den Transferstudien wieder zu finden, führt zu einer Vielzahl von Studien, die aufgrund der dort bewerteten Tierart ausgeschlossen werden müssen⁶. Nachdem auch Studien zum Einfluss der Luftverschmutzung auf die Wasserqualität, dem Erhalt der Regenwälder, der Bewertung ganzer Bündel von schützenswerten Gütern⁷ aus der engeren Wahl entfernt wurden, bleiben 8 Studien zu näherer Betrachtung übrig. Die Ausschlusskriterien bei diesen 8 Studien werden im Folgenden etwas detaillierter dargestellt.

Bateman et al. (2001) [27] bewerten Maßnahmen zur Reduktion des Säuregehalts in schottischen Seen. Die Studie hat den Vorteil, dass sie erst 2001 durchgeführt wurde, also sehr zeitnah ist. Der Hauptfokus der Studie liegt auf Tests auf Größeneffekte (Scope effects), daher werden verschiedene Kombinationen einzelner Seen untersucht. Der ökonomische Wert der Seen enthält einen hohen Anteil an Nichtgebrauchswerten, da die Seen so tief im schottischen Bergland liegen, dass sie wenig genutzt werden. Darüber hinaus werden zwei verschiedene Szenarien abgefragt, eines, das zur Reduktion einer zukünftigen stärkeren Verschmutzung führt und eines, das

⁵Beispielsweise werden in Studien vor 1990 keine Konfidenzintervalle für die errechneten Zahlungsbereitschaften genannt.

⁶So wurden Karibus, Timberwölfe oder Marder in Neufundland bewertet, was für den geplanten Nutzentransfer von geringem Wert ist.

⁷Kahneman und Ritov (1994) [147] bewerten eine Reihe sehr unterschiedlicher Güter: Baudenkmäler, Korallen und verschiedene Tierarten.

zur Reduktion des bestehenden Niveaus auf ein niedrigeres führt. Die Studie nutzt die Contingent Valuation Methode als Mail Survey mit offenem Frageformat, gezahlt wird ein fester Aufschlag auf die monatliche Elektrizitätsrechnung. Insgesamt scheint das Umweltgut jedoch so stark von dem geplanten Cottbuser Ostsee abzuweichen, dass die Studie wenig geeignet ist zum Nutzentransfer auf dieses Gebiet.

ECOTEC (1993) [213] bewertet ein ökologisch ähnliches Problem wie Bateman et al. (2001). Auch hier geht es um die Reduzierung des Säuregehalts in Seen. Allerdings liegt in dieser Studie das Hauptaugenmerk auf der Unterscheidung zwischen Nutzern der Ressource - hier konzentriert auf Angler - und dem Teil der Bevölkerung, die die Seen nicht nutzen. Für die Bewertung des Sees durch Nicht-Angler wird eine persönliche Befragung durchgeführt. Technik und Stichprobenumfang sind ähnlich, gezahlt wird über einen Aufschlag auf die Wasserrechnung. Allerdings sollte der Cottbuser See auch neue Nutzungsmöglichkeiten schaffen, und es ist im Vorhinein unklar, ob Angeln, das zwar eine in Ostdeutschland weit verbreitete Freizeitbeschäftigung ist, hier die hauptsächliche Nutzungsmöglichkeit darstellen wird. Die in dieser Studie gemessene Zahlungsbereitschaft beträgt 25,33£ (1993 Britische Pfund) für den Teil der Bevölkerung, der den See nicht nutzt bzw. bei 38,69£ für Angler. Um diese Beträge in Euro umzurechnen, müsste man Britische Pfund auf 2003 hochrechnen und dann in Euro konvertieren. Der große zeitliche Abstand erfordert bei Verwendung der Studie auf jeden Fall eine Bewertung der Zahlungsbereitschaft mit der Inflationsrate, zusätzlich müssten Kaufkraftparitäten einbezogen werden.

Herriges und Shogren (1996) [123] untersuchen die Verbesserung der Wasserqualität hinsichtlich der Sedimentablagerung in einem See in Iowa. Die Studie wurde 1993 durchgeführt. Es sollte eine einmalige Zahlung stattfinden, die über 5 Jahre abgezahlt werden konnte. Die Autoren finden Werte zwischen \$ 148,85 (in 1993 US Dollar) bei Anliegern und \$ 43,10 bei Besuchern. Dies ergibt Zahlungsbereitschaften zwischen 2,05 € und ca. 0,40€ auf Monate während des Zahlungszeitraums umgerechnet. Schwierigkeiten macht hier allerdings die nicht mitveröffentlichte Diskontrate über den Zahlungszeitraum. Es werden keine Aussagen darüber getroffen, ob die Befragten bei einer Abzahlung über einen Zeitraum von 5 Jahren eventuell Zinsen einrechnen, wie sie es bei Ratenverträgen von Konsumgütern gewohnt sind. Eine einmalige Zahlung ist zudem schlecht vergleichbar mit den Modalitäten für den Cottbuser Ostsee.

Hoevenagel und van der Linden (1993) [127] betrachten hingegen eine sehr umfassende Umweltverbesserung. Das Interessante an dieser Studie ist jedoch, dass sie einen dem Rekultivierungsprojekt in Cottbus ähnlichen Zeithorizont in ihrem Szenario unterstellen. Die Zahlungsbereitschaften für eine Verbesserung der Umweltqualität bis 2015, i. e. eine Minderung der Verschmutzung von Gewässern, des Abbaus der Ozonschicht, der Entwaldung und des Nitrateintrags durch Nutztiere, wurde insgesamt mit einer monatlichen Zahlungsbereitschaft von zwischen 30,22 HFL (1993 Niederländische Gulden) und 52,77 HFL bedacht. Dies entspricht einer Zahlungsbereitschaft von 13,75 € bis 24 € ohne die Berücksichtigung von Inflationsraten oder Kaufkraftparitäten.

Willis und Saunders (1995) [253] untersuchen sogenannte "Environmentally Sensitive Areas". Diese Bezeichnung ist eine Kategorie zur landschaftsplanerischen Einordnung von Gebieten in Großbritannien, denen eine besondere Schutzwürdigkeit zugestanden wird. Willis und Saunders bewerten die Downs in Südengland. Diese Landschaft ist durch das Aufkommen seltener Tierarten, gekennzeichnet, beinhaltet jedoch auch landwirtschaftliche Nutzung. Die Studie schildert die Befragung von 250 Besuchern und 279 Einwohnerhaushalten. Es wurden Zahlkarten benutzt und als Zahlungsmittel wurde ein Aufschlag auf Steuern angegeben. Eine jährliche Zahlungsbereitschaft von zwischen 13,51 £ (1993 brit. Pfund) und 17,53 £ wurde ermittelt. Dies entspricht einer monatlichen Zahlungsbereitschaft von zwischen 1,44 € und ca. 1,90 € ohne die Berücksichtigung von Inflationsraten oder Kaufkraftparitäten.

Reaves, Kramer und Holmes [211] gehen 1999 der Frage nach, in wieweit sich das Frageformat auf die ermittelte Zahlungsbereitschaft auswirkt. Zu diesem Zweck analysieren sie die Zahlungsbereitschaft der Bevölkerung von South Carolina und der restlichen USA für die Wiederherstellung eines Naturparks, der durch den Hurrikan Hugo 1989 zerstört wurde und den Lebensraum für eine bestimmte Spechtart bildet. Obwohl die Studie zeitnah und sehr ausführlich dokumentiert ist, was immer ein wesentlicher Vorteil für die Anwendung einer Studie zum Nutzentransfer ist, ist der Bewertungsgegenstand zu unterschiedlich zum Cottbuser Gebiet, als dass die Studie in die engere Auswahl käme.

Bjørner et al. (2000) [34] untersuchen in Dänemark die Zahlungsbereitschaft der umliegenden Bevölkerung für den Tokkekob Hegn (dän. Tokkekob (Ortsname), Hegn = Hain oder Gehölz), ein walddreiches Gebiet mit einem See nördlich von Kopenhagen. Die Befragung ist ein Mail Survey,

gezahlt wird für "Eintrittskarten", die jeweils einem Haushalt den Zutritt für ein Jahr gewähren. Die Studie wurde im Rahmen einer internationalen Kooperation mit den USA und Großbritannien durchgeführt. Es wurden jährliche Zahlungsbereitschaften von 174 DK (1999 Dänische Kronen) bis 190 DK ermittelt. Dies entspricht einer monatlichen Zahlungsbereitschaft von zwischen 1,95 € und ca. 2,14 €. Die Zeitnähe, die ausführliche Dokumentation, sowie die Ähnlichkeit des Gebiets lässt diese Studie in die engere Wahl fallen. Darüber hinaus ist es auch günstig, dass es sich um eine europäische Studie handelt, denn die Umwelteinstellungen gerade in einem Land wie Dänemark, das auch in gewissen Umweltbereichen, wie etwa der Ökosteuer, eine Vorreiterrolle in der Umweltgesetzgebung eingenommen hat, in der Bevölkerung können als den deutschen ähnlich angenommen werden.

Die zweite Studie von Bateman et al. (2004) [26] hat die Erforschung von Scope Effekten als wesentliches Erkenntnisziel. Zu diesem Zweck analysieren sie die Verbesserung der Wasserqualität eines Sees durch verschiedenen Maßnahmen bei der Wasserentsorgung durch die Universität East Anglia und die Zahlungsbereitschaft von Studenten für verschiedene unterschiedliche Wasserqualitäten umfassende Szenarien. Die Konzentration auf einen Verursacher und die Tatsache, dass die Stichproben zum einen sehr klein sind und eine nicht für die Bevölkerung repräsentative Gruppe (Studenten) umfasst, machen die Studie zu einem weniger geeigneten Kandidaten für den geplanten Nutzentransfer.

Auffällig ist bei der Zusammenschau der Studien, dass sich allein durch die Eingrenzung auf die oben genannten Indikatoren eine Auswahl mit vielen europäischen Studien ergeben hat. Da es beim Value Transfer auch darauf ankommt, dass die zugrunde liegende Bevölkerungsstruktur möglichst ähnlich ist, kann dies jedoch nur von Vorteil sein⁸.

Diese kurze Übersicht hat gezeigt, dass im Wesentlichen zwei Studien in die engere Auswahl kommen: ECOTEC (1993) und Bjørner et al. (2000). Vor allem die größere Aktualität lässt die Studie von Bjørner et al. geeignet erscheinen. Um die Zahlungsbereitschaften von Dänischen Kronen im Jahr 2000 auf den Euro und seine Kaufkraft in Deutschland umzurechnen, benötigt man Daten zur Kaufkraftparität, zum Wechselkurs und zum Preisindex. Kaufkraftparitäten finden sich für viele Länder ausgedrückt in

⁸Eine interessante Studie bezieht sich auf die USA. Cordell und Bergstrom (1993) [73] untersuchen den Einfluss von Wassermanagement in Talsperren auf den Wert, den die Bevölkerung einer jährlich längeren Nutzungsperiode dieser Talsperren zumisst.

US\$ bei UNDP (United Nation Development Program). Der Preisindex für Dänemark ist den amtlichen Statistiken Dänemarks entnommen. Die Umrechnung von dänischen Kronen auf "deutsche Euro", d.h. gewichtet mit den Kaufkraftparitäten, folgt Gleichung (5.1). Alle nachfolgenden Angaben unterliegen dieser Umrechnung.

$$X \cdot DK_{2000} * CPI_{2004} = \frac{PPP_{DK-US}}{PPP_{BRD-US}} \cdot \text{Wechselkurs}_{DK-Euro}, \quad (5.1)$$

dabei bezeichnet DK_{2000} Dänische Kronen im Jahr 2000, CPI_{2004} kennzeichnet den Verbraucherpreisindex im Jahr 2004, PPP_{A-B} bezeichnet die Kaufkraftparitäten zwischen Land A und B .

Ähneln sich Dänemark und Deutschland stark genug, dass die dänische Studie zum Nutzentransfer herangezogen werden kann? Während die durchschnittliche Lebenserwartung in Deutschland etwas höher ist (78,2 Jahre gegenüber 76,6 Jahre) hat Dänemark ein höheres Pro-Kopf-Inlandsprodukt (30.940 \$ gegenüber 27.100 \$). Die öffentlichen Ausgaben werden anders verteilt: 4,6 Prozent gingen in Deutschland zwischen 1999 und 2001 in den Bildungssektor gegenüber 8,3 Prozent des Bruttoinlandsprodukts in Dänemark. Für das Gesundheitswesen wurden in demselben Zeitraum in Deutschland 8,1 Prozent des Bruttoinlandsprodukts aufgewendet gegenüber 7 Prozent in Dänemark. Die dänische Produktion ist energieeffizienter, was als ein Resultat der politischen Anstrengungen für die Verbesserung der Energieeffizienz gewertet werden kann, aber auch den Stellenwert von Umweltfragen andeutet. Aus einem Kilogramm Erdöläquivalent entstanden 2001 in Dänemark 7,3 US\$ Beitrag zum Bruttoinlandsprodukt, Deutschland bringt es hier auf 5,6 US\$. Auch bei den Pro-Kopf- CO_2 -Emissionen führt Deutschland 2000 auf unrühmliche Art; sie betragen in Deutschland 9,6 t gegenüber 8,4 t in Dänemark. Hinsichtlich der Arbeitslosenquote in 2002 liegt Deutschland mit 8,1 Prozent weit vor Dänemark mit 4,2 Prozent (alle Angaben aus UNDP).

Um die Übertragbarkeit der dänischen Studienergebnisse auf den Cottbuser Ostsee abschätzen zu können, sollte aber vor allem die Bevölkerung im Studiengebiet in Dänemark dem Einzugsgebiets des Sees ähnlich sein. Tabelle 5.2 zeigt eine Übersicht über die wichtigsten Kenngrößen. Die dänische Gruppe ist im Durchschnitt etwas jünger als die brandenburgische, allerdings sind die Unterschiede eher gering. Die Einwohnerzahl im Umland des dänischen Gebiets ist deutlich größer als im Einzugsgebiet des geplanten Cottbuser Ostsees.

Tabelle 5.2: Kenngrößen (2004)

	Tokkekob Hegn	Cottbuser Ostsee
Soziodemographische Kenngrößen		
Einwohner (1000)	612,05	106,91
Alter (%) < 15	19,70	15,18
15-24	9,76	10,86
25-44	28,31	29,64
45-65	28,40	27,41
> 65	13,82	16,91
Einkommen ^a (€/Monat)	ca. 2600	2118
Charakteristiken des Umweltguts		
Gebrauchswerte	Spazieren gehen, Rad fahren	
	Schwimmen, Fischen, Boot fahren	
Nicht-Gebrauchswerte	"Artenvielfalt" Naturschutzgebiet	
	Laubfrosch	
	Flussseeschwalbe	
	Flussregenpfeifer	
	Zwergrohrdommel	
	Beutelmeise	
See	2 km ²	19 km ²
Wald	ca. 25 km ²	ca. 8 km ²
Substitute	Waldgebiete	
	Küstennähe	mehrere Seen

^a Durchschnittseinkommen
<http://www.dst.dk/HomeUK.aspx>, Danmarks Statistik (entspricht dem Statistischen Bundesamt), <http://www.lds-bb.de/sixcms/list.php/lds>, Landesbetrieb für Datenverarbeitung und Statistik Land Brandenburg, eigene Umrechnung.

Ländervergleiche im durchschnittlichen Einkommen sind oft schwierig, da die Einkommens- und Verbrauchsstichproben oft erhebliche nationale Unterschiede aufweisen. Daher zeigt Tabelle 5.2 lediglich einen durchschnittlichen Wert gemittelt über die Primäreinkommen aller Arbeitnehmer für Dänemark und zum Vergleich das durchschnittliche Bruttoeinkommen für Brandenburg. Wegen der sehr unterschiedlichen Steuerzahlungen und staatlichen Leistungen in Deutschland und Dänemark sind diese Zahlen jedoch ohnehin schlecht miteinander vergleichbar. Bjørner et al. entnehmen die Einkommensgrößen direkt den Steuerstatistiken und weisen somit keine eigenen Einkommenstabellen aus.

Das eigentlich zu bewertende Gut ist zwar in seinen Nutzungsmöglichkeiten dem geplanten Cottbuser See sehr ähnlich, weist aber zwei wesentliche Unterschiede auf. Zum einen ist im Tokkekob Hegn der Waldanteil größer und der See von geringerer Fläche, zum anderen handelt es sich hier um einen bereits existierenden See, während die Cottbuser auf ihren See noch über zwanzig Jahre warten müssen. Über die "innere Diskontrakte" bei der Bewertung eines solchen Projekts ist auch aus der Literatur bislang noch wenig bekannt. Die Substitute unterscheiden sich bei beiden Gebieten. Während die Nutzung des Sees im Tokkekob Hegn aufgrund der Nähe zum Meer alternative Wassersportmöglichkeiten entgegenstehen, sind die umliegenden Waldgebiete eher kleiner. Der Besuch des Cottbuser Sees hingegen kann zwar durch den Besuch anderer Seen ähnlicher Größe ersetzt werden, allerdings sind die Nutzungsmöglichkeiten an diesen Seen nicht so entwickelt (beispielsweise gibt es keine ausgebauten Strände, etc.).

Der von Bjørner et al. durchgeführte Mailsurvey basiert auf 3700 versandten Fragebögen, von denen 2581 ausgefüllt und zurückgesandt wurden. Nachdem ungültige Antworten eliminiert wurden, blieben 2474 beantwortete Fragebögen als Stichprobe übrig. Die Frage nach der Zahlungsbereitschaft wurde offen gestellt, ihr ging die Frage nach den Präferenzen für 5 verschiedene Ausgestaltungsmöglichkeiten des Tokkekob Hegn voraus. Dabei stellt sich heraus, dass die Bevölkerung eine Erschließung eines Wald- und Seengebiets mit der Möglichkeit zu schwimmen und zu angeln an ausgesuchten Stellen und Zugängen zum Wasser durch Wander- und Radwege am stärksten bevorzugt. Diese Ausgestaltung ist weitgehend deckungsgleich mit dem Szenario für den Cottbuser Ostsee. Für dieses Szenario wurde die Zahlungsbereitschaft erfragt⁹. Gut 58% der Befragten äußerte eine Zahlungsbereitschaft zwischen 0 und 2000 DK, letzteres entsprach zum Erhebungszeitraum knapp 270 €. Wenn alle geäußerten Zahlungsbereitschaften in Höhe von 0 in das Modell zur Schätzung der Zahlungsbereitschaft einbezogen werden ergibt sich ein Mittelwert von 190 DK (25,6 €), ein Median von 100 DK (13,5 €) mit einer Standardabweichung von 633 DK (entspricht 85,2 €). Je nach Behandlung der Protestantworten und von Geboten über 2000 DK ergeben sich die in Tabelle 5.3 aufgeführten Werte. Außerdem analysieren Bjørner et al. in der Studie den Einfluss verschiedener Frageformulierungen auf die Höhe der angegebenen Zahlungsbereitschaft. So weisen sie in einem Teil der Fragebögen explizit darauf hin, dass man die Interessen anderer Personen mit in seine Antwort einbezie-

⁹Mein Dank gilt Thomas Bu Bjørner und Cliff Russell für die Daten aus dem Projekt "Tokkekob Hegn".

Tabelle 5.3: Zahlungsbereitschaft für den Tokkekob Hegn

		DK	€
Inkl. Nullantworten	Mittelwert	190	25,6
	Median	100	13,5
	Standardabweichung	633	85,2
Exkl. Nullantworten	Mittelwert	284	38,2
	Median	100	13,5
	Standardabweichung	756	101,7
Inkl. Nullantworten, exkl. Protestantworten, exkl. WTP > 2000 DK	Mittelwert	215	28,9
	Median	100	13,5
	Standardabweichung	250	33,6
Eigene Zusammenstellung nach Bjørner et al. (2000).			

hen soll, ein anderer Teil der Stichprobe wird aufgefordert, sich seiner Verpflichtung gegenüber der Natur bewusst zu werden. Diese Formulierungen bringen die Befragten offensichtlich in ihre Einschätzung der wünschenswerten Gestaltung des Tokkekob Hegn ein und richten auch die geäußerten Zahlungsbereitschaften danach aus. Das aus dieser Diskussion ableitbare konservativste Maß wäre somit der Median mit 100 DK im Jahr oder einer monatlichen Zahlung von 1,13 €.

Im Vergleich mit der gemessenen Zahlungsbereitschaft für den Cottbuser Ostsee von 4,39 € liegt dieser Wert so deutlich darunter, dass sich jeder statistische Test erübrigt. Die Zahlungsbereitschaft in Cottbus und Umgebung läge bei einer Hochrechnung bei knapp 700.000 €/ Jahr, gegenüber dem in Kapitel 3 ermittelten Wert von rund 2,7 Mio. €/ Jahr.

Angesichts der Tatsache, dass der Cottbuser See erst ein geplantes Projekt ist und das durchschnittliche Alter der Cottbuser Bevölkerung höher ist als in Dänemark, stellt sich die Frage, ob sogar diese konservative Schätzung zu einer Überbewertung führt. Hierzu wäre zum Beispiel interessant, den Einfluss der Altersvariable sowie anderer charakteristischer Variablen auf die Zahlungsbereitschaft zu untersuchen. Diesem Ansatz geht die Idee der Nutzentransferfunktion nach.

5.2.2 Benefit Function Transfer

Einige der offensichtlichen Mängel des Transfers von Punktschätzern lassen sich, so wird einschlägig gehofft, durch die Anwendung der sogenannten Transferfunktion überwinden. Bei diesem Ansatz wird der aus einer Primärstudie gewonnene statistische Zusammenhang zwischen der Zahlungsbereitschaft und soziodemographischen Daten wie dem Einkommen, dem Alter, dem Bildungsstand und Informationen über Einstellungen und Kenntnisse der Befragten auf die Bevölkerung im Gebiet der Sekundärstudie übertragen und so die Zahlungsbereitschaft im Sekundärstudiengebiet errechnet. Obwohl diese Idee auf den ersten Blick völlig einleuchtend erscheint, muss man sich doch die Details einer solchen Nutzentransferfunktion vor dem Hintergrund der gewählten Messverfahren erneut anschauen.

Der Vorschlag einer Nutzentransferfunktion geht auf Loomis (1992) [178] zurück, der im Wesentlichen eine Verbesserung des bis dato verbreiteten direkten Value Transfers anstrebt. Obwohl Contingent Valuation Daten im Beitrag von Loomis zusammen mit Reisekostendaten erwähnt werden, beziehen sich die von Loomis genannten Anwendungsbeispiele doch ausschließlich auf Reisekostenmodelle. Auch die späteren Anwendungen der Nutzentransferfunktion bei anderen Autoren beziehen sich zu einem großen Teil auf die Reisekostenmethode, jedoch soll auf diesen Teil der Literatur in dieser Arbeit nicht eingegangen werden.

Für Contingent Valuation Daten lässt sich die Grundidee der Nutzentransferfunktion folgendermaßen darstellen. Angenommen, es liege eine Bewertungstudie vor, die zu einem dem geplanten Projekt im Transfergebiet ähnlichen Umweltprojekt durchgeführt wurde. Die Ergebnisse dieser Studie liegen typischerweise - wie weiter oben ausführlich erläutert - als kontinuierliche Daten, als Ja/Nein - Antworten auf eine einfache Fragestellung oder als Ja/Nein-Antwortkombinationen aus einer doppelten Fragestellung vor. Der Erwartungswert der Zahlungsbereitschaft lässt sich als Funktion in Abhängigkeit von den erhobenen sozioökonomischen Charakteristika der Befragten und der Höhe des jeweiligen "Gebots" darstellen. Im Falle des linearen Nutzenmodells aus Gleichung (3.1) erhält man unter Verwendung des Probit-Modells einen in den Mittelwerten der unabhängigen Variablen \bar{s} linearen Ausdruck für diesen Erwartungswert

$$E(WTP) = -\frac{\theta \bar{s}}{\beta}, \quad (5.2)$$

wobei θ den Vektor der geschätzten Parameter darstellt und \bar{s} der Vektor

der Mittelwerte der sozioökonomischen Variablen ist mit $s_0 = 1$. Wenn Gleichung (5.2) tatsächlich alle wesentlichen Kovariate erfasst *und* im Gebiet der Sekundärstudie sich Informationen über alle diese Kovariate erhalten lassen *und* keine weitere Größe für die Zahlungsbereitschaft der Bevölkerung eine wichtige Rolle spielt, *dann* muss dieses Verfahren auch einwandfrei funktionieren; d.h. die Zahlungsbereitschaft für das geplante Umweltprojekt lässt sich mit dieser Abschätzung ebenso gut ermitteln wie mit einer Studie, die dasselbe Modell verwendet. Die Bevölkerung im Transfergebiet (Sekundär-Site) darf eine andere Altersstruktur, eine andere Verteilung der verfügbaren Einkommen und einen anderen Ausbildungsstand aufweisen, sich also in allen erklärenden Variablen von der Bevölkerung im Sekundärstudiengebiet unterscheiden. Wichtig für das Funktionieren eines Nutzentransfers nach diesem Ansatz ist nur, dass die Präferenzen der Bevölkerung in beiden Gebieten gleich sind. Formal ausgedrückt bedeutet dies, dass der Parametervektor θ in beiden Gebieten gleich ist.

Studien zum Nutzentransfer überprüfen daher auch typischerweise diese Parametergleichheit mit den in Kapitel 5.3 dargestellten Testverfahren. Die Überprüfung des Nutzentransfers ist insofern schwierig, als dass für das Transfergebiet gerade keine Daten vorliegen. Um den Nutzentransfer zu überprüfen, muss man zwei Studien mit großem Umfang und erheblichem Aufwand durchführen. Eine Alternative stellt hier die Simulation dar, bei der sich mit zwei "wohlerzogenen" simulierten Bevölkerungen eine Vielfalt an möglichen Übereinstimmungen und Abweichungen studieren lassen. So können die verschiedenen Effekte, die unterschiedliche Bevölkerungszusammensetzungen, unterschiedliche Präferenzen im Primär- und Sekundärgebiet und Kombinationen aus diesen Effekten untersucht werden. Bislang wurden Simulationsrechnungen in der Umweltbewertung für den Modellvergleich, den Einfluss des Frageformats und die Ausgestaltung der Gebote genutzt. Im Folgenden wird die Befragung der simulierten Bevölkerung aus Kapitel 3.1 als Primärstudie betrachtet und eine Bevölkerung im Gebiet der Sekundärstudie simuliert. Diese unterscheiden sich zunächst nur in der Ausprägung ihrer sozioökonomischen Variablen von der Bevölkerung im Gebiet der Primärstudie. Tabelle 5.4 fasst die sozioökonomischen Variablen aus beiden Gebieten zusammen und zeigt das Ergebnis der Anwendung der Transferfunktion. Zu jeder Bevölkerung lässt sich die wahre Zahlungsbereitschaft schätzen, für das Transfergebiet gilt diese wahre Zahlungsbereitschaft als Referenzpunkt. Das Probit-Modell wird für die Daten des Primärstudiengebiets mit allen erklärenden Variablen geschätzt. Man erhält einen Vektor mit geschätzten Parametern, die in Gleichung (5.3)

Tabelle 5.4: Unterschiedliche sozioökonomische Variablen und gleiche Präferenzen

Variable	Mittelwert	Std.Abw.
Primärstudie		
s_1^p	35,31	17,57
s_2^p	3,80	5,04
s_3^p	36,55	22,90
WTP_{wahr}^p	109,25	39,43
Sekundärstudie		
s_1^s	17,66	8,85
s_2^s	3,80	5,04
s_3^s	109,64	68,70
WTP_{wahr}^s	132,70	57,94
$WTP_{Transfer}^s$	133,66	1,72

$N = 500, 500$ Wiederholungen. Eigene Berechnungen.

eingesetzt werden können.

$$E(WTP)^s = \frac{-(\alpha_0^p + \alpha_1^p \cdot \bar{s}_1^s + \alpha_2^p \cdot \bar{s}_2^s + \alpha_3^s \cdot \bar{s}_3^p)}{\beta} \quad (5.3)$$

Diesen Wert kann man mit dem wahren Mittelwert der Zahlungsbereitschaft vergleichen. Es zeigt sich, dass der Transfer auf das Sekundärstudiengbiet nahezu exakt ist, wenn die Voraussetzungen der Nutzentransferfunktion erfüllt sind, wenn also die Präferenzen der Bevölkerung überall gleich sind und sich die Bevölkerungen nur durch ihre sozioökonomischen Variablen unterscheiden.

Wenn aber die Präferenzen sich in den beiden Gebieten unterscheiden, werden die Ergebnisse erheblich verzerrt, selbst wenn die sozioökonomischen Variablen in beiden Gebieten dieselben sind, die Bevölkerungen auf den ersten Blick also einen hinreichend ähnlichen Eindruck vermitteln. Die Nutzentransferfunktion schätzt dann natürlich denselben Wert, den die Zahlungsbereitschaft auch im Primärstudiengbiet angenommen hatte (vgl. Tabelle 5.5). Im nächsten Schritt wird gezeigt, welche Abweichungen zwischen dem wahren Wert der Zahlungsbereitschaft und dem transferierten Wert entstehen können, wenn die Präferenzen und die sozioökonomischen Variablen im Sekundärstudiengbiet anders sind als im Primärstudiengbiet.

Tabelle 5.5: Unterschiedliche Präferenzen

	Mittelwert	Std.Abw.
$s_1^p = s_1^s$	35,31	17,57
$s_2^p = s_2^s$	3,80	5,04
$s_3^p = s_3^s$	36,55	22,90
WTP^p	109,25	39,43
WTP^s	151,97	58,11
Transfer	109,39	0,96

N = 500, 500 Wiederholungen. Eigene Berechnungen.

Jetzt sei die wahre Zahlungsbereitschaft im Sekundärstudiengbiet durch Gleichung (5.4) gegeben

$$WTP_{wahr} = 20 + 2x_1 - 0,05x_2 + 0,8x_3 + \epsilon \quad (5.4)$$

mit:

$$x_1 \in [2, 5; 32, 5]$$

$$x_2 \sim N(4, 5)$$

$$x_3 \in [3, 219]$$

$$\epsilon \sim N(0, 2).$$

Um die Veränderung der Mittelwerte und der "wahren" Zahlungsbereitschaft zu zeigen, wurde auch hier eine Monte Carlo Simulation mit 500 Wiederholungen durchgeführt; die mittlere Zahlungsbereitschaft ergibt sich zu $E(WTP_{wahr}) = 143,67$. Durch den positiven Zusammenhang zwischen der Zahlungsbereitschaft und der Variablen x_3 und die Tatsache, dass diese Variable an der Sekundär-Site deutlicher ausgeprägt ist, ist auch die mittlere wahre Zahlungsbereitschaft an der Sekundär-Site erheblich höher als an der Primär-Site. Darüber hinaus wurde eine größere Konstante angenommen, die die Zahlungsbereitschaft an der Sekundär-Site zusätzlich erhöht. Eine mögliche Interpretation wäre, dass die Variable x_1 das Bildungsniveau darstellt, das an der Sekundär-Site geringer ausfällt. Die Konstante ließe sich so interpretieren, dass die Sekundär-Site andererseits in einem Land liegt, in welchem die Spendentätigkeit grundsätzlich höher ausfällt, also eine allen Individuen zu eigene grundlegend höhere Zahlungsbereitschaft für die Bereitstellung des Umweltguts vorliegt. Das Problembewusstsein für dieses spezielle zu bewertende Umweltgut aufgrund mag darüber hinaus aufgrund vorheriger Erfahrungen im Durchschnitt höher ausgeprägt sein, was sich in der Variablen x_3 äußert.

Tabelle 5.6 zeigt eine Übersicht über die Ergebnisse von verschiedenen Modellspezifikationen für die simulierten Daten in den beiden Gebieten. Es werden die Ergebnisse einer offenen Befragung und einer anschließenden Regressionsanalyse und die Ergebnisse einer geschlossenen Befragung mit verschiedenen Probit-Modellen gezeigt, die unterschiedlich viele erklärende Variablen beinhalten. In beiden Gebieten führen die jeweiligen Ansätze zu nahezu identischen Ergebnissen. Die Nullhypothese gleicher Zahlungsbereitschaften wurde in allen Kombinationen getestet und konnte nicht zurückgewiesen werden, allerdings sind die voll spezifizierten Modelle zu bevorzugen, denn sie haben bei weitem den größeren Wert der Likelihoodfunktion. Bei der Übertragung von Modell I weicht der transferierte Wert bereits erheblich ab mit rund 133 € gegenüber knapp 142 €. Das offene Modell zeigt eine ähnliche Abweichung. Besonders gravierend wird der Fehler jedoch, wenn die Primärstudie mit einer Variablen weniger geschätzt wurde, hier simuliert durch die Variable x_3 . Während die jeweiligen Schätzungen innerhalb der Primärstudie und innerhalb der Sekundärstudie hiervon nicht so stark betroffen sind - der Erwartungswert der Zahlungsbereitschaft ändert sich kaum - ist das Transferergebnis schlichtweg falsch mit 74,48 €. Durch den Nutzentransfer wird das Omitted Variable Problem, i.e. die Verzerrung der Schätzung durch Nichtberücksichtigung relevanter erklärender unabhängiger Größen, dadurch wichtig, dass Größen, die in der Primärstudie keine Bedeutung besitzen, für die Zahlungsbereitschaft im Transfergebiet als erklärende Variablen berücksichtigt werden müssten.

Das Beispiel wurde derart konstruiert, dass der Fehler durch eine vergessene Variable besonders deutlich hervortritt. Der nachfolgende Blick in bisher durchgeführte Studien mit der Nutzentransferfunktion wird aber zeigen, dass die bisherigen Erfahrungen nicht besonders gut sind. Der wesentliche Designfehler bei diesem Ansatz besteht darin, dass man einerseits zur Übertragung bisheriger Werte auf ein neues Projekt eine möglichst detaillierte Zerlegung der Determinanten der Zahlungsbereitschaft benötigt, andererseits gerade darin die Gefahr besteht, wesentliche Determinanten zu vernachlässigen und somit große Fehler zu machen. Anders ausgedrückt geht der Nutzentransfer mittels Transferfunktion von der Annahme aus, dass sich die Präferenzen der Bürger repräsentieren lassen durch die Abhängigkeit der Zahlungsbereitschaften von den Kovariaten, i.e. die Parameter in Gleichung (5.3). Über die gesamte Bevölkerung eines Landes hinweg ist es auch zutreffend, dass die Reaktionen auf beispielsweise eine Einkommensverminderung hinreichend ähnlich sind. Auch der Einfluss des Alters auf die Einschätzung von in der fernen Zukunft liegenden Projekten

Tabelle 5.6: Modelle im Primär- und Sekundärstudienggebiet

	WTP	Parameter	Std.Abw.
Primärstudie			
Probit/Modell I	109,38	β	-0,84 (0,13)
		α_1	1,67 (0,25)
		α_2	-0,08 (0,05)
		α_3	0,66 (0,10)
		α_0	8,35 (1,40)
Probit/Modell II	109,10	β	-0,05 (0,00)
		α_1	0,10 (0,01)
		α_2	-0,02 (0,01)
		α_0	2,23 (0,17)
Offen/ Modell II	109,82	α_1	1,97 (0,02)
		α_2	-0,07 (0,08)
		α_0	40,76 (0,97)
Offen/ Modell III	109,82	α_1	2,00 (0,00)
		α_2	-0,04 (0,01)
		α_3	0,80 (0,00)
		α_0	9,92 (0,13)
Sekundärstudie			
Probit/Modell I	141,50	β	-1,13 (0,21)
		α_1	2,29 (0,44)
		α_2	-0,01 (0,07)
		α_3	0,90 (0,17)
		α_0	21,14 (4,13)
Transfer	133,50		
Probit/Modell II	143,38	β	-0,02 (0,00)
		α_1	0,04 (0,01)
		α_2	-0,001 (0,01)
		α_0	1,88 (0,17)
Transfer	74,48		
Offen/ Modell II	141,96	α_1	2,02 (0,14)
		α_2	-0,12 (0,24)
		α_0	107,36 (2,88)
Transfer	74,48		
Offen/ Modell III	141,99	α_1	2,01 (0,01)
		α_2	-0,04 (0,01)
		α_3	0,80 (0,00)
		α_0	19,75 (0,13)
Transfer	133,80		
Eigene Berechnungen.			

verhält sich plausiblerweise ähnlich. Da Umweltprojekte aber häufig auch einen emotionalen Aspekt aufweisen, können gerade diese Aspekte an unterschiedlichen Orten zu sehr unterschiedlichen Einschätzungen führen. So wird zum Beispiel die Neugestaltung eines Sees in Brandenburg (als einer eher strukturschwachen Gegend) vielleicht mit anderen Hoffnungen, Ängsten und Wünschen begleitet als der Erhalt eines Waldgebiets mit See in der Nähe des eher wohlhabenden Kopenhagen.

Nutzentransferfunktion für den Cottbuser Ostsee

Trotz der teilweise entmutigenden Erfahrungen mit dem Nutzentransfer mittels einer Transferfunktion in der Literatur soll nachfolgend die Anwendung dieses Ansatzes auf die Bewertung des Cottbuser Ostsees und ihr Ergebnis kurz dargestellt werden. Als Referenzstudie dient die dänische Studie zum Tokkekob Hegn, die schon den Ausgangspunkt für den Nutzentransfer mittels Punktschätzern lieferte. Tabelle 5.2.2 gibt eine Übersicht über die Regressionsergebnisse der Studie im Einzelnen und die in Gleichung (5.3) einzusetzenden Mittelwerte der jeweiligen sozioökonomischen Variablen.

Allerdings sind diese Variablen bezogen auf das Einzugsgebiet des Cottbuser Ostsees nicht erhältlich; ein mehr als typisches Problem für derartige Nutzentransferstudien. So wurde das mittlere Einkommen des Bundeslands Brandenburg angesetzt, da für die betreffende Region keine Einkommenserhebungen vorliegen. Ebenso wurden bei den Variablen "Durchschnittsalter" und "Single", was dem Anteil von Singlehaushalten an der Gesamtzahl aller Haushalte entspricht, die brandenburgischen Durchschnittswerte angesetzt. Für die Variablen "walkoft" und "NGO", die das Freizeitverhalten mit häufigen Spaziergängen sowie die Mitgliedschaft in einem Umweltverband, oder anderen Non Government Organizations beschreiben, mussten die Durchschnittswerte für Ostdeutschland angesetzt werden (SOEP). Einsetzen dieser Werte und Umrechnung der Beträge von dänischen Kronen (1999) in Euro (2003) führt zu einer monatlichen Zahlungsbereitschaft von 2,24 €. Diese Abschätzung ist zwar schon deutlich näher als der Value Transfer an dem in der Contingent Valuation Studie ermittelten "wirklichen Wert" von 4,39 €, liegt jedoch immer noch erheblich darunter. Darüber hinaus ist sie mit einer Vielzahl an Unsicherheiten behaftet. Alleine die statistische Ermittlung der bei der Berechnung der Zahlungsbereitschaft wichtigen Einkommensgröße lässt viele Interpretationen offen.

Tabelle 5.7: Parameter Tokkekob Hegn und Bevölkerungsdaten Cottbus

	Parameter	t-Wert	Mittelwert Cottbus
age	-2.11	-3.61	41.53
income	0.355	3.97	2118
income*income	-0.00011	-3.10	-
single	71.61	3.16	0.53
NGO	61.01	3.23	0.05
walkoft	73.16	4.47	0.63
intercept	134.58	2.65	-
Eigene Berechnungen.			

Einkommen ist in der Bundesrepublik keine Größe, die erfasst wird, wie etwa Geburten, Sterbefälle etc. Einkommen wird vielmehr direkt im Rahmen der Einkommens- und Verbrauchsstichprobe oder des Mikrozensus des Statistischen Bundesamts erhoben.

Die Abgrenzung des Begriffs "verfügbares" Einkommen kann in einer weniger umfangreichen Befragung, wie etwa dem in Cottbus verwendeten Fragebogen ungenauer sein, denn die Befragten wissen nicht, ob sie ihre Nettoeinkommen angeben sollen, oder ihr Nettoeinkommen abzüglich aller Kosten wie Miete, Versicherungen und ähnliche Aufwendungen¹⁰. Darüber hinaus reflektiert das Einkommen in Dänemark, einem Land mit vielen Transferleistungen des Staates und hohen Steuern, eine andere Ausgangsbasis als etwa in der Bundesrepublik. Somit könnte man davon ausgehen, dass bei der Verwendung einer Nutzentransferfunktion nicht nur die Mittelwerte der sozioökonomischen Größen ersetzt werden müssen, sondern auch die Parameter von Land zu Land unterschiedlich ausfallen¹¹.

¹⁰Statistisches Bundesamt (2003) [50]. Was ist die Einkommens- und Verbrauchsstichprobe? "Die EVS wird alle fünf Jahre durchgeführt. Eine gesetzliche Verpflichtung zur Teilnahme besteht nicht, d.h. alle Haushalte nehmen auf freiwilliger Basis an der EVS teil. (...) Das sind insgesamt rund 75 000 Haushalte, darunter etwa 15 000 Haushalte in den neuen Ländern und Berlin-Ost. Die EVS ist damit die größte Erhebung dieser Art innerhalb der Europäischen Union. Im früheren Bundesgebiet findet die EVS seit 1962/63 statt, in den neuen Ländern und Berlin-Ost seit 1993. "

¹¹Dieser Idee geht ein Ansatz nach, der als "chirurgischer Transfer" in der Literatur zu finden ist. Die zentrale Idee hierbei ist, eine kleine Studie vor Ort durchzuführen und die Parameter der Zahlungsbeurteilungsfunktion der ursprünglichen Studie durch Parameterschätzungen aus dieser kleinen Studie zu ersetzen, sofern dies notwendig erscheint (vgl. Poulos (2000) [204]). Eine ähnliche Idee zur Anwendung bei Travel Cost Modellen verfolgen Smith et al. (2000) [228] mit dem sogenannten Kalibrierungsansatz.

Erfahrungen mit der Nutzentransferfunktion

Zahlreiche Anwendungs- und Überprüfungsstudien wurden durchgeführt seit dem Vorschlag der Nutzentransferfunktion durch Loomis (1992) [178]¹². Dabei ist die überwiegend anzutreffende Vorgehensweise, dass Studien explizit zur Überprüfung eines Nutzentransfers durchgeführt werden, die a priori als ähnlich eingeschätzte Umweltgüter bewerten und in einem zweiten Schritt vergleichen, ob sich die Resultate auch in einem Nutzentransfer hätten erreichen lassen¹³. Downing und Ozuna (1996) [90] führten beispielsweise an verschiedenen Parks in Texas Befragungen durch und errechnen die jeweiligen Zahlungsbereitschaften. Sie testen die Vergleichbarkeit der Parameter und entwickeln so ein Kriterium für Studien, die sich überhaupt übertragen lassen. Anschließend werden die Studien, die diesen Test bestehen, zu einem Funktionstransfer genutzt und die Vergleichbarkeit der so ermittelten Zahlungsbereitschaften wird überprüft. Die Ergebnisse sind eher enttäuschend und die Autoren kommen zu dem Schluss, dass das Verfahren des Nutzentransfers unzuverlässig sei. Beim näheren Hinsehen lassen sich allerdings einige Kritikpunkte an dem von den Autoren gewählten Ansatz vermerken, auf die weiter unten im Detail eingegangen wird. Ebenfalls zu einem enttäuschenden Ergebnis kommen Kirchoff et al. (1997) [154], die sowohl den direkten Nutzentransfer als auch den Transfer mittels der Transferfunktion einer kritischen Untersuchung unterziehen. Die Untersuchungen wurden mithilfe zweier Studien durchgeführt, bei denen die erste Studie die Erhaltung des Lebensraums des Graubussards in Arizona bewertet und die zweite Studie, ebenfalls in Arizona, die Erhaltung eines Flussauengebiets ohne Konzentration auf eine spezifische Tierart bewert-

Obwohl die Idee der Aufdatierung von Parametern grundsätzlich richtig erscheint und auch im Folgenden weiterverfolgt wird, ist diese Vorgehensweise doch zu einfach, denn allein aufgrund der unterschiedlichen Stichprobengrößen sind die Parameterschätzer mit sehr unterschiedlichen Fehlern behaftet und die stochastischen und die deterministischen Komponenten von zwei statistischen Modellen sind nicht zu trennen.

¹²Auch die Nutzentransferansätze folgen dem insgesamt in der Literatur zu verzeichnenden "Trend" zur Kombination verschiedener Bewertungsverfahren. So führte ein britisches Beratungsinstitut (EF-TEC (2002) [91]) einen Nutzentransfer zwischen England und Wales durch, der auf Choice Experiment Studien aufbaut. Ebenso vergleichen Hanley et al. (2003) [121] die Bewertung der Verbesserung von Küstengewässern (und der damit verbundenen Badequalität) indem sie Contingent Valuation Daten und reale Verhaltensweisen und Zahlungen beobachten.

¹³Navrud und Kristofersson (2001) [197] fragen allerdings zu Recht, ob dies die richtige Vorgehensweise ist, um ein Instrument zu testen, dessen Gebrauch vor allem von der praktischen politischen Anwendbarkeit diktiert ist. Sie schlagen alternative Tests vor, die statistische Signifikanz mit der politischen Relevanz verbinden (vgl. die Darstellung von Testverfahren, insbesondere Seite 157 f.)

tet. Die Zahlung wird in einer einmaligen Zahlung an eine Naturschutzorganisation geleistet. Obwohl auch in der zweiten Studie ein großer Teil der Befragten Vogelbeobachtungen für ein wichtiges Element des zu erhaltenden Gebiets ansah, lassen doch die Unterschiede in den Gebieten die Abweichungen im durchgeführten Nutzentransfer nicht allzu überraschend erscheinen. Brouwer und Spanninks (1999) [49] versuchen den Effekten, die in die Nutzenfunktion über die "klassischen" Größen wie Einkommen etc. hinaus eingehen, nachzugehen und anhand einer Isolierung dieser Effekte den Funktionstransfer zu verbessern. Darüber hinaus findet sich bei Brouwer und Spanninks eine Übersicht über die in der Literatur verwendeten Validitätstest für den Nutzentransfer, die in Kapitel 5.3 ausführlicher diskutiert werden. Brouwer und Spanninks schließen bei der Bewertung zweier Torflandschaften in den Niederlanden auch eine Reihe von Einstellungsfragen mit ein, in der Hoffnung, dass die Ermittlung des Einflusses derartiger Einstellungen, beispielsweise gegenüber der Umwelt oder der Bereitstellung öffentlicher Güter, den Nutzentransfer verbessern wird. Die empirischen Befunde belegen dies allerdings nicht. Zwar kann ein signifikanter (positiver) Einfluss sowohl der Einstellung zu dem bewerteten Gut im Speziellen als auch zur Umwelt im Allgemeinen auf die Zahlungsbereitschaft in beiden Studien nachgewiesen werden, allerdings unterscheiden sich die Parameter in beiden Studien erheblich. Einzig der Einfluss des Einkommens war in beiden Studien gleich und signifikant. Pearce (2000) [203] untersucht die Stabilität der Bewertung von Risiken für Leben und Gesundheit in verschiedenen europäischen Ländern, indem er in jedem Land ähnliche Studien durchführt.

Thiele und Wronka [80] hingegen kommen 2001 zu sehr kleinen Abweichungen bei einem Nutzentransfer zwischen zwei Gemeinden in Hessen. Allerdings sind die Studien auch in jeglicher Hinsicht völlig ähnlich, so dass ein Fehlschlagen des Nutzentransfers an dieser Stelle ernsthafte Verwunderung ausgelöst hätte. Mit recht kleinen Stichproben finden sie einen Transferfehler von acht Prozent, was angesichts des Stichprobenumfangs von jeweils ca. 150 eine andere Formulierung für "nichts" ist¹⁴.

Die Ergebnisse in der Literatur sind somit recht heterogen. In einem wirklichen Nutzentransfer ist man darüber hinaus auf das Vorliegen von bereits durchgeführten Studien angewiesen und muss sich zudem für eine Studie entscheiden, deren Parameter dann für den Transfer genutzt werden.

¹⁴Zum Vergleich: selbst MC Simulationen mit absolut identischen Funktionen weichen bei einer Stichprobe von 150 um ca. 6% voneinander ab.

Zumindest diese Schwierigkeit versuchen meta-analytische Ansätze zu umgehen, indem sie aus der Zusammenschau bereits durchgeführter Studien den Transfer auf ein neues Projekt ableiten. In vielen Bereichen ist die Zahl der wissenschaftlichen Artikel und Publikationen in den letzten Jahren auf ein nachgerade frustrierendes Niveau angestiegen. Dem gegenüber stehen politische Entscheidungsträger, die zum Beispiel staatliche Maßnahmen auf eine wissenschaftliche Kosten-Nutzen Analyse stützen wollen, sich jedoch einer Flut von eher unsystematischen Informationen ausgesetzt sehen. Als eine Methode der Systematisierung von vorhandenen Studien hat sich die sogenannte Meta-Analyse etabliert. Mit Hilfe dieses Verfahrens, so lässt sich vermuten, können typische Einflussfaktoren auf die Zahlungsbereitschaft für Umweltgüter isoliert werden, die sich zur Konstruktion einer Nutzentransferfunktion verwenden lassen. Grundlage dieser Nutzentransferfunktion ist nun nicht mehr eine ausgewählte Primärstudie, sondern die Zusammenschau von Studien die ein ähnliches Umweltgut bewerten.

Der erste Schritt einer Meta-Analyse liegt, analog zur Anwendung der Nutzentransferfunktion, in der Auswahl der zu verwendenden Studien. Die Unterschiede zwischen einzelnen in die Meta-Analyse einbezogenen Studien lassen sich zum einen mit Stichprobenfehlern erklären, d.h. die wirkliche zugrunde liegende Zahlungsbereitschaft ist in allen Studien dieselbe, und die Abweichungen sind auf Messfehler zurückzuführen; es kann aber auch sein, dass die zugrunde liegenden Zahlungsbereitschaften in den verschiedenen Studien tatsächlich unterschiedlich sind, dass man es somit mit tatsächlich heterogenen Ergebnissen zu tun hat. Ein einfacher Test auf Heterogenität ist der χ^2 -Test, der im Wesentlichen misst, ob verschiedenen Resultate Stichproben aus derselben Grundgesamtheit entsprechen können. Je nach Anzahl der zugrunde liegenden Studien ist die Stärke dieses Tests natürlich erheblich eingeschränkt, denn seine Aussagefähigkeit nimmt mit der Anzahl der in der Meta-Analyse eingeschlossenen Studien zu. Ziel einer Meta-Analyse ist es, eine Schätzung der Zahlungsbereitschaft aus verschiedenen Studien zu gewinnen. Daher ist die Gewichtung der Ergebnisse der Einzelstudien, die in die Meta-Analyse einfließen, ein zentrales Thema bei der Diskussion meta-analytischer Verfahren. Unterstellt man, dass alle Studien, die einer Meta-Analyse zugrunde liegen, ein und denselben Effekt messen, kann man sogenannte Fixed Effects Methoden benutzen (vgl. Sutton et al. (2002) [238]). Dabei wird allen Studien ein zu ihrer Varianz umgekehrt proportionales Gewicht (die Präzision) zugeordnet, d.h. die Studien mit der größeren Varianz werden geringer gewichtet; während in Random Effects Modellen der Verschiedenheit der einzelnen Studien zugrunde liegenden

Gesamtheiten Rechnung getragen und die Varianz zwischen den Studien explizit modelliert wird.

Rosenberger und Loomis (2000) [216] stellen eine beeindruckende Meta-Analyse für die Berechnung des Werts von Freizeitaktivitäten (Outdoor) in den USA vor, die auf der Auswertung von 163 kategorisierten Studien beruht. Dieses System enthält sowohl Charakteristika der gewählten Methode in jeder vergangenen Studie als auch Einzelheiten zu den zu bewertenden Freizeitaktivitäten und der betroffenen Bevölkerung. Allerdings liegen für Europa oder gar für Deutschland bei weitem nicht so viele Studien vor, so dass es für eine derartige Analyse an Daten mangelt.

5.3 Validitätstests für den Nutzentransfer

Kriterien zur Qualitätsbeurteilung eines Nutzentransfers haben die Literatur zum Value Transfer und zur Nutzentransferfunktion spätestens seit der bereits erwähnten Artikelserie in *Water Resources Research* 1992 beschäftigt. Die verwendeten Testverfahren unterscheiden sich nicht von den üblicherweise verwendeten, allerdings ist die Anwendbarkeit dieser Tests auf den Nutzentransfer begrenzt. Der reale Nutzentransfer nach den bislang beschriebenen Verfahren; i.e. die Übertragung des in der Primärstudie gemessenen Werts eines Umweltguts auf ein neues, entzieht sich jeglichem Testverfahren. Das Problem hierbei ist, dass es nicht, wie bei einem statistischen Modell zur Schätzung eines bestimmten Werts, eine Datenstruktur gibt, an die das beste Modell nach bestimmten Kriterien am besten anpasst wird. Beim Nutzentransfer herkömmlicher Art suchen wir nach einem Modell, mit dem sich die Zahlungsbereitschaft für ein neues Umweltgut schätzen lässt, ohne dass direkte Daten zum Verhalten der Bevölkerung vorliegen. Daher müssen für den Nutzentransfer immer explizite Überprüfungsstudien durchgeführt werden, bei denen dann das Modell mit der - bekannten - Schätzung verglichen werden kann. Zu diesem Zweck werden Studien gleichen Designs zur Bewertung ähnlicher Umweltgüter mit ähnlichen Bevölkerungen durchgeführt. Aus jeder Studie lässt sich nun der Wert des Umweltguts zum einen direkt und zum anderen durch Übertragung des Schätzwerts einer anderen Studie schätzen. An den Abweichungen zwischen dem direkt ermittelten und dem übertragenen Wert lässt sich die Gültigkeit des Transfers überprüfen.

Der einfachste Test ist der Test auf Gleichheit der Ergebnisse, d.h. die

durch Nutzentransfer ermittelte Zahlungsbereitschaft ist gleich der vor Ort gemessenen Zahlungsbereitschaft. Ein valider Transfer läge auf jeden Fall vor, wenn gilt $\overline{WTP}_{trans} = \overline{WTP}_s$ und dies lässt sich mit parametrischen oder nicht-parametrischen Tests überprüfen. Für den parametrischen t -Test nimmt man an, dass \overline{WTP}_{trans} und \overline{WTP}_s aus einer Normalverteilung gezogen wurden. Der Mann-Whitney oder Wilcoxon Rangsummentest ist ein nicht-parametrischer Test, der die Mittelwert oder Mediane zweier Verteilungen miteinander vergleicht (vgl. Diamond et al. 1993) [88]). Die Nullhypothese lautet $\overline{WTP}_{trans} = \overline{WTP}_s$ und eine zurückgewiesene Nullhypothese bedeutet, dass sich die Werte signifikant unterscheiden¹⁵. Der Test auf Gleichheit der Mittelwerte wird bei der dänischen Studie zum Tokkekob Hegn und der Cottbuser Studie zurückgewiesen, die Mittelwerte der Zahlungsbereitschaft sind nicht gleich.

Mit diesen Tests lässt sich ebenfalls untersuchen, ob die in der Primärstudie gemessene Zahlungsbereitschaft gleich der in der Sekundärstudie gemessenen ist, dann lautet die Nullhypothese $\overline{WTP}_p = \overline{WTP}_s$. Bergland et al. (1998) [30] schlagen darüber hinaus vor, zu testen ob die *Verteilung* der Zahlungsbereitschaften gleich ist ($H_0 : f(WTP_p) = g(WTP_s)$) und verwenden zur Überprüfung den Kolmogorov-Smirnov-Test. Weitere mögliche Tests beziehen sich auf die in der Primär- und der Sekundärstudie gemessenen Parameter und Varianzen. Zur Überprüfung von $\sigma_p^2 = \sigma_s^2$ schlagen Loomis (1992) [178] und Rosenberger und Loomis (2000) [216] den Chow-Test vor, Kirchhoff et al. (1997) [154] verwenden den Likelihood-Ratio-Test (LR-Test), zur Überprüfung von $\alpha_p = \alpha_s$ verwenden Bergland et al. (1995) den Lagrange-Multiplier-Test. Diese Tests sind jedoch globale Tests für die Modellgüte und Chow-Test und LR-Test überprüfen bei einer Anwendung auf die Validität des Nutzentransfers eigentlich die Gleichheit der in den Modellen erklärten Varianzen. Dies bedeutet, dass dieser Test keine Aussagen über die Gleichheit der Parameter macht, sondern nur besagt, dass die zur Erklärung der Zahlungsbereitschaft einbezogenen Variablen gleichermaßen die Variabilität erklären. Zur Überprüfung der Gleichheit der Parameter bevorzugen Brouwer und Spanninks [49] aus diesen Gründen den Wald-Test.

¹⁵Holm-Müller und Muthke (2004)[128] weisen darauf hin, dass ein nicht zurückgewiesener Test auf Gleichheit der Mittelwerte diese Gleichheit streng genommen noch nicht zeigt. Sie schlagen vor, auch die andere Richtung zu testen und die Hypothese der Gleichheit der Mittelwerte erst anzunehmen, wenn die Ungleichheit der Mittelwerte abgelehnt wurde. Allerdings scheitern viele der in der Literatur dokumentierten Studien schon an der ersten Stufe dieses Tests und eine Verschärfung der Tests erscheint in diesem Stadium fast unnötig.

Navrud und Kristófersson (2001) [197] gehen der Frage nach, ob diese Testverfahren nicht am eigentlichen Zweck einer Nutzentransferstudie vorbeigehen. Sie schlagen statt dessen sogenannte Äquivalenztests vor, bei denen statistische Gleichheit in Kombination mit politischer Bedeutsamkeit getestet werden kann. Die Inspiration für diesen Vorschlag beziehen die Autoren aus der pharmazeutischen Forschung, wo derartige Tests bereits angewendet werden. In einem Äquivalenztest wird überprüft, ob der gemessene Wert um einen im Voraus festzulegenden Betrag von einem anderen abweicht. Die Nullhypothese lautet somit $H_0 : \overline{WTP}_p - \overline{WTP}_s \leq \Delta$, mit dem festzulegenden Niveau Δ . Diese Nullhypothese lässt sich entweder durch $H_1 : \Delta + \overline{WTP}_s > \overline{WTP}_p \geq \overline{WTP}_s$ oder durch $H_1 : \overline{WTP}_p \leq \overline{WTP}_s$ zurückweisen. Mit zwei einseitigen t-Tests lässt sich dies überprüfen. Der Vorteil dieses Verfahrens besteht darin, dass die Akzeptanz oder Ablehnung eines Nutzentransferergebnisses nun auch von dem Parameter Δ abhängt, der extern festgelegt wird. Je nach Verwendungszweck der Nutzentransferstudie wird dieses Δ unterschiedlich ausfallen. Navrud und Kristófersson stellen eine Hierarchie verschiedener Verwendungszwecke auf. Der stärksten Einschränkung müsste die Abweichung sich bei der Verwendung der Studie für haftungsrechtliche Zwecke unterwerfen. Die Vorteile der Methode liegen darin, dass der Entscheidungsträger, der letztlich die Ergebnisse eines Nutzentransfers verwenden will und muss, eine Art Toleranzbereich der Abweichungen vom wahren Wert festlegen kann.

Zusammenfassend lässt sich festhalten, dass die bislang in der Literatur zu den herkömmlichen Verfahren des Nutzentransfers durchgeführten Tests den Schluss nahelegen, dass diese Verfahren nicht zu gleichen Werten für die Sekundär- und die Primärstudien führen. Die Tests haben ferner gezeigt, dass in vielen Überprüfungsstudien auch die Zahlungsbereitschaften sich in beiden Gebieten nicht auf dieselbe Art, d.h. mit demselben Einfluss derselben Variablen erklären lassen.

5.4 Möglichkeiten und Grenzen klassischer Verfahren

Ahlheim und Lehr kommen 2002 [9] zu dem Schluss, dass "es durchaus nicht einfach ist, sich eine abschließende Meinung zum Konzept des Nutzentransfers zu bilden. Einerseits möchte man sich als theoretisch arbeitender Ökonom indigniert abwenden von so vielen Verletzungen heiliger Prinzipien der reinen Lehre. Andererseits ist der Preis solcher Keuschheit recht hoch. Der Weg von der Theorie zum "wirklichen Leben" ist mit Kompromissen gepflastert, und wenn wir nicht bereit sind, diese Kompromisse einzugehen, bleibt uns der Weg zur Wirklichkeit verschlossen. Das aber kann sich weder die Ökonomie noch die Wirklichkeit leisten. Die Ökonomie kann es sich nicht leisten, weil sie so jeglichen Einfluss auf die Gestaltung der Realität und damit zugleich einen Teil ihrer Daseinsberechtigung verlöre, und die Wirklichkeit kann es sich nicht leisten, weil Umweltpolitik ohne ökonomische Leitlinien angesichts knapper Ressourcen und angespannter staatlicher Budgets zum Lotteriespiel wird."

Folgt man diesem recht klaren Votum *für* den Nutzentransfer, so stellt sich die Frage, wie sich die Verfahren weiterentwickeln lassen, um die "Verletzung der heiligen Prinzipien" nicht ganz so schmerzlich werden zu lassen. Die in diesem Kapitel beschriebenen Verfahren scheinen noch einige Verbesserungen vertragen zu können, denn die Fehlerhaftigkeit lässt sich in einfachen Simulationen zeigen und ist in vielfältiger Weise in der Literatur dokumentiert. Erstaunlicherweise hat sich die Scientific Community allerdings über fast zehn Jahre hinweg auf die Fortentwicklung der Testverfahren konzentriert (vgl. Mogas und Riera (2003) [195]) und den eigentlichen Nutzentransfer unberührt gelassen (vgl. Brouwer (2000) [48]). Das grundlegende Problem der Nutzentransferfunktion liegt jedoch darin, dass sie die Charakteristika eines für die Primärstudie gültigen Modells auf ein neues Projekt, ein neues Umweltgut und ein anders Gebiet oder sogar ein anderes Land überträgt (vgl. Ready et al. (1999) [210]). Wie McConnell (1992) [187] schon festhält, besteht jedoch ein Zusammenhang zwischen der Modellspezifizierung, der Güte der Schätzer und der zu erwartenden Güte eines Nutzentransfers.

Jeder Praktiker jedoch, der eine Bewertungsstudie durchgeführt hat, weiß, dass auch Variablen, die bislang in vielen Studien zur Erklärung der Zahlungsbereitschaft signifikant beigetragen haben, in einer anderen Studie keinen nennenswerten Einfluss auf die aktuelle Zahlungsbereitschaft ausüben.

Dafür ist die Bevölkerung plötzlich sensibel für einen bestimmten Aspekt des Umweltguts, der charakteristisch nur für diesen See oder diese Tierart ist. Wie in den vorangegangenen Beispielen gezeigt werden konnte, führt die Vernachlässigung einer Variablen beim Transfer zu einer erheblichen Verzerrung. In den Modellen zur Bewertung des Cottbuser Ostsees hat sich zum Beispiel die Variable "Alter" als nicht signifikant erwiesen, in der Studie zum Tokkekob Hegn war diese Variable signifikant. Auch die Beteiligung an NGOs, die in der dänischen Studie eine wichtige Variable war, trägt in Cottbus zu keinem Erkenntnisgewinn bei, weil dort die Mitgliedschaften in NGOs äußerst gering sind. Für einen Nutzentransfer der herkömmlichen Art benötigt man jedoch gerade die Zusammenhänge zwischen den Kovariaten und der Zahlungsbereitschaft, um einen möglichst guten Transfer zu erreichen.

Hat man also eine gute Primärstudie mit den weiter oben beschriebenen Verfahren isoliert, so wäre es nützlich zu wissen, wie die Bevölkerung im Gebiet des geplanten Umweltprojekts auf die Frage nach der Zahlungsbereitschaft reagiert, und in welchem Umfang die vorgefundenen Variablen eine Rolle spielen. Ob man dazu jedoch tatsächlich eine umfangreiche Studie, die etwa den NOAA-Kriterien genügen würde, benötigt, ist die Frage, die im nächsten Kapitel ausführlich beantwortet wird.

Eine Möglichkeit, die "Kompromisse, die zur Versöhnung von ökonomischer Theorie und Realität notwendig sind, so wenig schmerzlich wie möglich zu gestalten" (Ahlheim und Lehr (2002) [9]), besteht darin, zumindest in geringem Umfang direkt *erhobene* Daten vom Sekundärstudiengebiet zu nutzen, um die Präferenzen der betroffenen Bevölkerung nicht völlig zu ignorieren. Die Kombination dieser Daten mit den Vorabinformationen aus der gewählten Primärstudie muss dann nachvollziehbar und transparent durchgeführt werden. Eine schlüssige Möglichkeit zur Realisierung dieses Ansinnens stellt wiederum der Bayes'sche Ansatz zur Verfügung.

Kapitel 6

Bayes'scher Nutzentransfer

6.1 Bayes'sche Verfahren als Chance für den Nutzentransfer

Einen neuen Impuls bekam die Forschung zum Nutzentransfer durch Arbeiten, in denen der augenfällige Isomorphismus zwischen dem Anliegen "Ergebnisse bereits durchgeführter Studien auf die Bewertung eines geplanten Projekts zu übertragen" und dem Bayes'schen Ansatz der Nutzung von *informativen A-Priori-Verteilungen* zur Verbesserung der Schätzung von Daten aus einer kleinen Stichprobe umgesetzt wurde. In anderen Disziplinen¹ ist die Verwendung von Daten aus früheren Studien für die Verbesserung von Berechnungen, die auf kleinen Stichproben basieren, bereits seit langem Stand der Technik. In der Literatur wurde die Möglichkeit, Bayes'sche Methoden für den Nutzentransfer einzusetzen, zwar bereits seit längerem angedeutet, jedoch überwiegend im Rahmen des Ausblicks auf zukünftige Forschungsfragen. Die ersten Arbeiten zum *Einsatz* von Bayes'schen Verfahren für den Nutzentransfer verwenden weniger rechenaufwändige Empi-

¹Für Anwendungen aus der pharmazeutischen Forschung vgl. Racine et al. (1986) [205], für Ansätze aus der Gesundheitsökonomie Kochi et al. (2002) [156], die Meta-Analyse wird durch die Anwendung von Bayes'schen Verfahren in medizinischen Studien vielfach genutzt (vgl. Carlin (1992) [56]) und für die detaillierte Analyse von parametrischen Bayes-Modellen mit Discrete Choice Daten vgl. Chen und Kuo (1999) [160].

rical Bayes Verfahren² (vgl. Atkinson et al. (1992) [21]). Parsons und Kealy (1994) [201] stellen eine frühe und in der nachfolgenden Literatur weitgehend ignorierte Form des Nutzentransfers mittels Bayes'scher Methoden dar, auf den weiter unten noch detailliert eingegangen wird. Zu Beginn dieses Jahrhunderts erlebte der Bayes'sche Nutzentransfer ausgehend von einer Gruppe um León eine Belebung (vgl. León und Vazquez-Polo (1998) [168], León et al. (2003) [167], León et al. (2002) [169], Fernandez et al. (2001) [96] oder Barton (2002) [25]).

Grundsätzliches Anliegen eines jeden Nutzentransfers ist die Übertragung bereits vorliegender Informationen aus einer oder mehreren Primärstudien auf das in der sogenannten Sekundärstudie neu zu bewertende Objekt. Die im letzten Kapitel vorgestellten herkömmlichen Verfahren des Nutzentransfers stützen sich in diesem Prozess entweder ausschließlich auf die Daten der Primärstudie (Value Transfer) oder auf allgemein verfügbare Daten für die Bevölkerung im Bereich der Sekundärstudie (Function Transfer). Über die Präferenzen der betroffenen Bevölkerung in diesem Gebiet können beide Ansätze jedoch keine Aussagen treffen, die gemeinsame Annahme ist, dass die Präferenzen denen der in der Primärstudie untersuchten Bevölkerung entsprechen. Will man hingegen zumindest in geringem Umfang direkt *erhobene* Daten vom Sekundärstudiengebiet nutzen, um die Präferenzen der betroffenen Bevölkerung nicht völlig zu ignorieren, so braucht man einen Mechanismus, der die kleine Vor-Ort-Erhebung verbessert, denn diese lässt alleine statistisch zuverlässige Aussagen nur in geringem Umfang zu. Die vorgeschlagene Verbesserung des Nutzentransfers besteht somit darin, eine *kleine* Bewertungsstudie im Gebiet der Sekundärstudie durchzuführen und die daraus gewonnenen Erkenntnisse mit bereits durchgeführten Studien zu festigen. Im Kapitel 2 wurde der Blickwechsel vom klassischen statistischen Modell auf ein Bayes'sches Modell bereits ausführlich beschrieben. Erhobene Daten dienen in der Bayes'schen Perspektive dazu, bereits vorhandene Kenntnisse über die Parameter eines Modells "aufzudatieren", man erinnere sich an den Zusammenhang in (2.35) in Kapitel 2

$$P(\text{Parameter}|\text{Daten}) \propto P(\text{Daten}|\text{Parameter}) * P(\text{Parameter}).$$

$P(\text{Daten}|\text{Parameter})$ lässt sich in Analogie zum klassischen Modell als die Likelihoodfunktion der Daten bei gegebenen Parametern identifizieren, $P(\text{Parameter})$ ist die A-Priori-Verteilung, die das vorab vorhandene Wissen über die Parameter inkorporiert.

²Für eine detaillierte Einführung in die Empirical Bayes Verfahren vgl. Carlin und Louis (1996).

In der Anwendung auf die *Schätzung einer Einzelstudie* in Kapitel 3.1.3 wurde das Bayes'sche Modell mit einer *nicht-informativen* A-Priori-Verteilung angewendet und es konnte gezeigt werden, dass die Ergebnisse identisch sind mit den Ergebnissen der klassischen Verfahren. Die Modellierung einer *informativen* A-Priori-Verteilung im nächsten Schritt in Kapitel 3.1.4 stützte sich auf Vorabinformationen aus Expertengesprächen und Pretests und die Schätzergebnisse wiesen schmalere Konfidenzintervalle und somit eine größere Zuverlässigkeit der möglichen Aussagen auf. Will man Gleichung (2.35) für den *Nutzentransfer* verwenden, kommt der A-Priori-Verteilung eine noch wichtigere Rolle zu als zuvor. Aus der Primärstudie liegen nun explizit Informationen über die Verteilung der Parameter vor, die durch eine große Studie und eine sorgfältige Schätzung - und nur solche Studien kommen für einen Nutzentransfer überhaupt in Frage - abgesichert sind. Wie aber werden diese Informationen modelliert? Welchen Einfluss hat die Modellierung auf die Ergebnisse des Nutzentransfers? Und letztlich, wie groß muss die "kleine" Stichprobe sein? Diesen Fragen geht der Rest dieses Kapitels nach.

Hierfür wird zunächst der analytisch lösbare Ansatz von Parsons und Kealy (1994) [201] vorgestellt, mit dessen Hilfe sich ein Bayes'scher Nutzentransfer für die Ergebnisse einer offenen Befragung durchführen lässt³. Gerade weil dieser Ansatz keiner Simulationstechniken zu seiner Berechnung bedarf, lassen sich einige grundlegenden Eigenschaften des Bayes'schen Transfers an ihm verdeutlichen. Allerdings wissen wir spätestens seit Kapitel 1.3.3 um die Nachteile des offenen Frageformats bei der Durchführung einer CVM-Studie.

Daher wird im zweiten Abschnitt dieses Kapitels das Bayes'sche Nutzentransfermodell für Dichotomous Choice Daten, wie man sie aus einer Befragung im Referendumsformat enthält, vorgestellt. Anhand von simulierten Daten werden verschiedene Wege zur Modellierung der Vorabinformationen illustriert und der Einfluss der Modellierungen auf das Ergebnis des Transfers untersucht. Es stellt sich heraus, dass die Ergebnisse der Primärstudie oftmals ein zu großes Gewicht in der Berechnung des transferierten Werts zu bekommen scheinen. Dies ermutigt die Suche nach einem Modell, bei dem eine Gewichtung der "historischen Daten" der Primärstudie vorge-

³Parsons und Kealy wollen in der Studie die Möglichkeiten verschiedener Ansätze des Nutzentransfers untersuchen. Zu diesem Zweck führten sie eine Befragung zum Wert verschiedener Seen in Wisconsin bei der Bevölkerung von Milwaukee und bei Haushalten durch, die außerhalb von Milwaukee leben. Zwischen diesen beiden Gruppen wurde der Nutzentransfer durchgeführt; der Bayes'sche Ansatz schneidet mindestens so gut, teilweise erheblich besser ab.

nommen wird und mündet in einer Formulierung der A-Priori-Verteilung in Analogie zu Ibrahim und Chen (2000) [137], dem Power Prior. Der Einsatz des Power Priors ist neu in der ökonomischen Umweltbewertungsforschung und es kann gezeigt werden, dass er viele ansprechende Eigenschaften besitzt und in der Anwendung auf die simulierten Daten sowie auf die Cottbuser Daten den bislang vorgestellten Verfahren überlegen ist.

6.2 Anwendung auf eine simulierte offene Befragung

Bei der Anwendung auf den Nutzentransfer wird die Wahl der A-Priori-Verteilung erheblich vereinfacht, denn zur Beschreibung der Daten aus einer bestimmten, bereits durchgeführten Contingent Valuation Studie wird innerhalb der ausgewählten Primärstudie bereits ein Modell sorgfältig gewählt worden sein, aus dem sich die Form der A-Priori-Verteilung ganz natürlich ergibt. Andernfalls wäre die Wahl der Primärstudie nicht vereinbar mit den von Desvousges et al. (1992) [85] aufgestellten Kriterien (vgl. S. 131 dieser Arbeit).

Wurden Primär- und Sekundärstudie als offene Befragung durchgeführt, so lässt sich das in Kapitel 2 auf S.80 ff. beschriebene Modell verwenden. Seien α_{prim} mit der Verteilung $N(\mathbf{x}\alpha_{prim}, \sigma^2\mathbf{I}|\mathbf{x})$ und α_{sek} mit $N(\mathbf{x}\alpha_{sek}, \Sigma_0|\mathbf{x})$ die Parametervektoren aus der Primär- und der Sekundärstudie. Parsons und Kealy (1994) verwenden den Ansatz (2.37) für einen Nutzentransfer, der hier erneut aufgeführt wird

$$\alpha_{bayes} = [\Sigma_0^{-1} + (\sigma^2(\mathbf{x}'\mathbf{x})^{-1})^{-1}]^{-1}(\Sigma_0^{-1}\alpha_{prim} + (\sigma^2(\mathbf{x}'\mathbf{x})^{-1})^{-1}\alpha_{sek}). \quad (6.1)$$

Da dieses Modell das einzige analytisch lösbare Modell im Zusammenhang mit dem Nutzentransfer darstellt, wird der Nutzentransfer mit diesem Bayes'schen Ansatz anhand von simulierten Daten nachstehend kurz gezeigt. Die größere Stichprobe dient in dieser Studie zur Erzeugung der A-Priori-Verteilung, die Parameter α_{prim} verschiedener sozioökonomischer Variablen sowie die Varianz-Kovarianz-Matrix werden vorab in Regressionen (OLS) geschätzt. Verschiedene erheblich kleinere Stichproben der zweiten Bevölkerungsgruppe werden zur Schätzung der Parameter α_{sek} und ebenfalls der Varianz-Kovarianz-Matrix verwendet. Dieses Verfahren lässt zu, dass an verschiedenen Orten verschiedene Präferenzen vorliegen, da vor

Ort für die Sekundärstudie zumindest eine kleine Stichprobe zur Messung der Zahlungsbereitschaft verwendet wird.

Die simulierten Daten für eine offene Befragung im Gebiet der Primär- und Sekundärstudie aus dem letzten Kapitel werden nun zur Konstruktion der A-Priori-Verteilung und zur Simulation einer kleinen Vor-Ort-Befragung genutzt. Aus der kompletten Stichprobe ($N = 2000$) ergibt sich im Modell für die offene Befragung die Zahlungsbereitschaft der Sekundärstudie zu 141,99 €. Um die Vor-Ort-Befragung zu simulieren werden zunächst Unterstichproben der Größe $n_{sek} = 100, 200, 300, 400$ gezogen, aus der Primärstudie werden 1000 Beobachtungen verwendet. Tabelle 6.1 weist in der zweiten Spalte die Erwartungswerte der Zahlungsbereitschaften der jeweiligen Unterstichprobe aus. Anschließend wird auch die Größe der Unterstichprobe der Primärstudie variiert.

Es sei die Sekundärstudie durchgeführt worden und Schätzer für die Parametervektoren α_{sek} sind mit einer linearen Regression ermittelt worden. Aus der Primärstudie liegen bereits Schätzer für α_{prim} vor. Tabelle 6.1 zeigt eine Übersicht der Resultate für die Simulation für verschiedene Kombinationen der jeweiligen Stichprobengrößen. Es zeigt sich dass die jeweilige Stichprobengröße über die zunehmenden Varianzen bei abnehmenden Stichprobengrößen das Ergebnis beeinflusst. Insgesamt liegt das Ergebnis des Transfers, wenn alle Kovariate eingeschlossen sind, konstruktionsbedingt immer zwischen den Zahlungsbereitschaften der Primär- und der Sekundärstudie. Die Tabelle gibt immer drei Zeilen an, die zu einem Transfer gehören. Die erste Zeile enthält die geschätzte Zahlungsbereitschaft für das Primärgebiet, die dazugehörigen Konfidenzintervalle und die Anzahl der Beobachtungen im Primärgebiet; die jeweils zweite Zeile enthält eben diese Werte für das Sekundärgebiet. In der dritten Zeile wird das Transferresultat mit den Konfidenzintervallen für dieses Ergebnis gezeigt. Die in der Tabelle ausgewiesenen Zahlungsbereitschaften beziehen sich auf OLS Schätzungen innerhalb einer Stichprobe der Größe $N =$ Anzahl der in der letzten Spalte aufgeführten Beobachtungen. Die Bootstrap-Konfidenzintervalle wurden folgendermaßen gewonnen: Aus der jeweiligen Stichprobe werden Bootstrap-Stichproben gezogen und die Parameter in einer Regression (OLS) auf diesen Bootstrap-Stichproben geschätzt. Dieses Verfahren wird 1000 Mal wiederholt und aus der so erzeugten Verteilung der Parameter lassen sich die Konfidenzintervalle ablesen.

Die ersten vier Modelle beziehen sich jeweils auf eine Stichprobengröße von 1000 Beobachtungen in der Primärstudie, mit einem entsprechend kleinen

Tabelle 6.1: Resultate des BT für eine offene Befragung

	WTP	95% Konf.	Beobachtungen
Primär	109,42	[108,79 - 110,10]	1000
Sekundär	148,77	[146,89 - 151,11]	100
Transfer	143,97	[142,03 - 145,89]	
Primär	108,81	[108,03 - 109,57]	1000
Sekundär	143,48	[141,72 - 145,21]	200
Transfer	134,44	[133,27 - 135,64]	
Primär	110,17	[108,99 - 111,02]	1000
Sekundär	137,09	[136,01 - 138,79]	300
Transfer	133,04	[132,11 - 133,14]	
Primär	108,73	[107,57 - 110,11]	1000
Sekundär	147,15	[145,07 - 148,24]	400
Transfer	143,61	[142,43 - 145,31]	
Primär	108,67	[107,29 - 109,68]	500
Sekundär	142,69	[141,35 - 144,29]	400
Transfer	139,72	[138,14 - 140,44]	
Primär	106,57	[104,99 - 107,23]	500
Sekundär	125,44	[122,54 - 130,71]	40
Transfer	122,13	[119,02 - 124,18]	
	Modell ohne x_2		
Primär	112,99	[111,00 - 114,15]	500
Sekundär	138,77	[136,02 - 142,43]	50
Transfer	133,98	[130,63 - 135,21]	
	Modell ohne x_3		
Primär	111,03	[106,72 - 115,26]	1000
Sekundär	146,71	[132,65 - 156,48]	100
Transfer	89,16	[84,04 - 93,45]	

Eigene Berechnungen.

Konfidenzintervall, respektive kleinen Varianzen der Parameter. Demzufolge ist der Einfluss der Primärstudie größer, je kleiner die Stichprobe der Sekundärstudie ist. Dieser inverse Zusammenhang zwischen der Varianz der jeweiligen Parameter und ihrem Einfluss auf das Endergebnis führt dazu, dass beim einfachen Aufdatierungsschema berücksichtigt wird, mit welcher Unsicherheit die aufdatierten und die zur Aufdatierung benutzten Größen behaftet sind. Der Reiz der Methode besteht darin, dass sich recht unaufwändig Daten aus der Sekundärstudie mit den Daten aus der Primärstudie kombinieren lassen.

Obwohl der transferierte Wert in allen Fällen sehr nahe am tatsächlichen Wert zu liegen scheint, liegt er doch nirgends innerhalb des ursprünglichen Konfidenzintervalls, das jeweils in der zweiten Zeile eines Modells angegeben ist. Bei den Modellen, die eine Variable in der Regression ausgelassen haben, verschlechtert sich der Transfer, wie im Falle der Nutzentransferfunktion, allerdings ist der Transfer im letzten Model, bei dem x_3 ausgelassen wurde etwas besser als im klassischen Fall. Wird dagegen bei der Studie vor Ort festgestellt, dass die Variable x_3 relevant ist, so kann dieser Term auch ohne Bayes'sche Aufdatierung in die Berechnung übernommen werden. Allgemeiner formuliert ergibt sich für $m = 1, \dots, M$ erklärende Variablen, die in Primär- und Sekundärstudie berücksichtigt wurden, und für die mit $k = M, M + 1, \dots, M + K$ bezeichneten Variablen, die nur in der Sekundärstudie einen Einfluss haben, der Erwartungswert der Zahlungsbereitschaft im offenen Modell zu

$$E(WTP) = \sum_{m=1}^M \alpha_{Bayes,m} x_m^s + \sum_{k=M+1}^{M+K} \alpha_k^s x_k^s, \quad (6.2)$$

wobei α_k^s die für die Sekundärstudie geschätzten Parameter bezeichnet. Der transferierte Wert beträgt mit diesem Vorgehen in der Simulation wiederum 133,55, bei einer Sekundärstichprobengröße von 300. Dies verwundert nicht, denn in der Simulation ist x_3 gerade die fehlende Variable zum datengenerierenden Mechanismus, daher muss das Transferergebnis dem ursprünglichen kompletten Modell ähneln. Aufschlussreicher wäre hier eine Untersuchung mit Daten aus der Realität, bei denen auch die Konfidenzintervalle einer Schätzung, die auf einem Stichprobenumfang von $n = 100$ basiert, erheblich größer sind. Da jedoch die offene Befragung wie in Kapitel 1 beschrieben, den Nachteil hat, dass sie gemeinhin als nicht anreizkompatibel angesehen wird und bei den Befragten zu großen Unsicherheiten führt, sind immer weniger Studien in diesem Format durchgeführt wor-

den und die Untersuchung realer Daten in dieser Arbeit wird sich auf eine Befragung im Referendumsformat konzentrieren.

Obwohl sich also die Einflüsse der Stichprobengröße an dem einfachen Modell gut beobachten lassen, ist es für den Bayes'schen Nutzentransfer wünschenswert, Modelle zu entwickeln, die sich auf die geschlossenen Fragestellungen übertragen lassen. Diese Modelle werden überwiegend nicht mehr analytisch zu berechnen sein, so dass man meistens auf die in Kapitel 2.4 vorgestellten Verfahren zurückgreifen muss.

6.3 Bayes'scher Nutzentransfer für Dichotomous Choice Daten

Das Bayes'sche Modell für Dichotomous Choice Daten für das einfache Referendum aus Kapitel 3 ist wiederum das Produkt aus der A-Priori-Verteilung und der Likelihoodfunktion, letztere ist jetzt ausgedrückt in Ablehnungs- und Zustimmungswahrscheinlichkeiten.

$$\pi(\boldsymbol{\theta}|\mathbf{x}) \propto \pi(\boldsymbol{\theta}) \cdot \prod_{i=1}^n (1 - F(\boldsymbol{\theta}, \mathbf{x}_i, \text{Bid}_i))^{y_{es_i}} \cdot F(\boldsymbol{\theta}, \mathbf{x}_i, \text{Bid}_i)^{1-y_{es_i}}, \quad (6.3)$$

wobei $\boldsymbol{\theta}$ der Parametervektor angibt, und $F(\bullet)$ wie in Kapitel 2 die Verteilung des stochastischen Terms des gewählten Modells; Bid_i bezeichnet die Zahlungsvorschläge oder Gebote und \mathbf{x}_i ist der Vektor aller erklärender Variablen mit $x_{0i} = 1$. Die vorab vorhandene Information wird in der A-Priori-Verteilung ausgedrückt und es muss, im Gegensatz zum eben beschriebenen Modell für die offene Befragung, das insgesamt auf der Normalverteilung basiert, eine *funktionale Form* für diese Verteilung gefunden werden. Das Problem der Bestimmung von A-Priori-Verteilungen aus vorhandenen Informationen hat Statistiker lange beschäftigt und die frühe Diskussion überspannt eine Reihe von philosophischen Ansätzen zu diesem Problem; von Savage (1954, 1972) [221] als Begründer der "persönlichen" Denkschule bis hin zur Ablehnung der Spezifizierung einer A-Priori-Verteilung, da diese immer willkürlich sei, und somit auch zur Ablehnung des Bayes'schen Ansatzes.

Jaynes (1968) [142] formuliert zur Lösung dieses Problems zunächst Bedingungen an die *Qualität* der vorab vorliegenden *Information*, aus denen

sich überhaupt eine sinnvolle A-Priori-Verteilung konstruieren lässt. Demnach wird die Information über einen Parameter *testbar* genannt, wenn sich bei einem gegebenen Vorschlag einer A-Priori-Wahrscheinlichkeit eindeutig sagen lässt, ob dieser Vorschlag mit der Information vereinbar ist oder nicht. Anders ausgedrückt, muss die Information formalisierbar sein. Zur Formalisierung schlägt Jaynes den Maximum Entropy Prior vor, d.h. die Verteilung, die die höchste Entropie hat und vereinbar ist mit der a priori vorliegenden Information I . Dies ist gleichzeitig diejenige Verteilung, die keine zusätzlichen Annahmen trifft, sondern nur die vorliegende Information verwendet. Für diskrete Verteilungen lässt sich eine Form direkt aus der Maximierung der Entropie

$$H = - \sum_{i=1}^N p(i|I) \ln(p(i|I)) \quad (6.4)$$

unter den Nebenbedingungen

$$\sum_{i=1}^N p(i|I) = 1; \quad p(i|I) \geq 0; \quad \sum_{i=1}^N p(i|I) f(x_i) = const \quad (6.5)$$

gewinnen. Die erste Nebenbedingung reflektiert die Tatsache, dass sich die Wahrscheinlichkeiten zu eins addieren müssen, die zweite Nebenbedingung bedeutet, dass Wahrscheinlichkeiten ≥ 0 sein müssen und die dritte Nebenbedingung enthält gerade die aus einer Messung zur Verfügung stehende Information. Auf die Umweltbewertung übertragen wäre dies zum Beispiel der gemessene Wert für die Zahlungsbereitschaft aus einer früheren Studie.

Für kontinuierliche Verteilungen müssen zusätzliche Annahmen getroffen werden (vgl. Jaynes (1968) [142] S. 14 ff.), denn die Integraldarstellung von Gleichung (6.4) ist ohne eine zusätzliche Normierung nicht transformationsinvariant. Die Festlegung eines Definitionsbereichs für die A-Priori-Verteilung ist eine mögliche Normierung. Dann ergibt sich die Maximum Entropy A-Priori-Verteilung für ein Problem, bei dem Informationen über die (als kontinuierlich anzunehmende) Zahlungsbereitschaft $a \leq WTP \leq b$ vorliegen, als (vgl. León und Vázquez-Polo (1998) [168])

$$\pi(WTP) = \frac{\kappa e^{-\kappa \cdot WTP}}{e^{-\kappa \cdot a} - e^{-\kappa \cdot b}}. \quad (6.6)$$

Zur Bestimmung von κ benötigt man den Erwartungswert der Zahlungsbereitschaft aus einer früheren Studie \overline{WTP}_0 und löst

$$\frac{\kappa(ae^{-\kappa a} - be^{-\kappa b}) + e^{-\kappa a} - e^{-\kappa b}}{\kappa(e^{-\kappa a} - e^{-\kappa b})} = \overline{WTP}_0. \quad (6.7)$$

Ein Beispiel für diese Modellierung findet sich bei León et al. (2002) [169]. Die Autoren verwenden zur Konstruktion der A-Priori-Verteilung eine Art Meta-Analyse, bei der aus der Zusammenschau bereits durchgeführter Bewertungsstudien einzelne Studien und ihre Ergebnisse ausgewählt werden. In diesem Modell ergibt sich die A-Priori-Verteilung aus den A-Priori-Verteilungen der Zahlungsbereitschaft λ (zur Unterscheidung zu den in dieser Arbeit des weiteren verwendeten Modellen wird an dieser Stelle die Notation von León et al. aufgegriffen) verschiedener Studien $\pi^j(\lambda)$, die mit den Gewichten w_j versehen und aufsummiert werden

$$\pi(\lambda) = \sum_{j=1}^m w_j \cdot \pi^j(\lambda). \quad (6.8)$$

Die Gewichte ergeben sich durch Faktoranalyse aus den Ähnlichkeiten, die die verwendeten historischen Studien mit dem zur Bewertung stehenden Gebiet haben. Abweichend zu den in Kapitel 2 vorgestellten Modellen verwenden die Autoren für die Likelihood-Verteilung die allgemeine Gamma-Verteilung mit (vgl. Stacy (1962) [233])

$$F(\ln(\text{Bid} - \alpha)) = \frac{|\delta|}{\Gamma(\delta-2)} (\delta-2)^{\delta-2} e^{[\delta-2(\delta \ln(\text{Bid}-\alpha)) - e^{\delta \ln(\text{Bid}-\alpha)}]}, \quad (6.9)$$

mit δ als sogenanntem Shape-Parameter, der die Gamma-Verteilung kennzeichnet. Zu diesem Zweck wird eine Variablentransformation durchgeführt, die es erlaubt, die Verteilung der Zahlungsbereitschaft λ in der Bevölkerung direkt zu modellieren. In den in Kapitel 2 beschriebenen Modellen hingegen wurde die Zahlungsbereitschaft als Funktion der geschätzten Parameter berechnet. Letzteres hat den Vorteil, dass sich zum Beispiel der Einfluss des Gebots auf die Zahlungsbereitschaft direkt abbilden lässt. Die Modellierung nach León et al. hingegen ist leichter zu handhaben, wenn tatsächlich nur Vorabinformationen über die Mittelwerte der Zahlungsbereitschaften aus vorherigen Studien vorliegen. Darüber hinaus lässt sich die A-Posteriori-Verteilung in dieser Spezifikation analytisch berechnen, was jedoch bei den heutzutage verfügbaren Programmpaketen keinen erheblichen Vorteil darstellt. León et al. verwenden, wie die in Kapitel 5.3 zitierten

Überprüfungsstudien, die explizit für einen Nutzentransfer durchgeführt werden, eigene Daten sowohl vom Primärgebiet als auch vom Sekundärgebiet. Dies führt zu einer - für Forschungszwecke durchaus sinnvollen - Kontrollierbarkeit der Ergebnisse, die aber in der Realität des Nutzentransfers kaum gegeben sein wird. Hier wird man vielmehr auf einzelne Studien oder die Zusammenschau einer Reihe geeigneter Studien zurückgreifen und die A-Priori-Verteilung entsprechend konstruieren. Die Anwendung des Modells auf die in verschiedenen spanischen Nationalparks erhobenen Daten zeigt, dass die Maximum-Entropy-Verteilung auch bei kleinen Stichproben aus dem Sekundärstudienggebiet die Schätzergebnisse für die Sekundärstudie sehr wenig verzerrt, allerdings zu recht großen Konfidenzintervallen führt. Diesen Ergebnissen stellen León et al. ein Modell mit der Beta-Verteilung gegenüber, bei welcher die Parameter aus dem Mittelwert der Zahlungsbereitschaft der vorab vorliegenden Studie und dem Modus der Verteilung berechnet werden. Die Beta-Verteilung ist eine viel informativere A-Priori-Verteilung als der Maximum Entropy Prior, da hier konkrete funktionale Formen der A-Priori Verteilung festgelegt werden. Die Ergebnisse mit der Beta-Verteilung zeigen jedoch eine starke Abhängigkeit von der A-Priori-Verteilung, beziehungsweise davon, welche Funktion aus der Familie der Beta-Verteilungen spezifiziert wurde.

Liegen aus einer bekannten Studie mehr Informationen über die einzelnen Parameter vor, so kann man zu einer "informativeren" A-Priori-Verteilung einzelner Parameter übergehen. Die Modellierung einer informativen A-Priori-Verteilung sollte weder die Datenstruktur aus dem Primärgebiet noch das Material aus dem Sekundärgebiet in ein Modell zwingen, das beispielsweise den Definitionsbereich der Parameter stark beschränkt. Daher kann es nicht ratsam sein, dass eine zu große Detailverliebtheit auf die genaue Anpassung an ein sehr spezielles vermutetes Modell, das für das Sekundärgebiet gilt, ausgelebt wird. Da eine Normalverteilung für alle der in Kapitel 3 vorgestellten Modelle zulässig ist, wird sie im Folgenden für den einfachen Bayes'schen Transfer für Dichotomous Choice (DC) Daten verwendet.

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{x}) &\propto \pi(\boldsymbol{\theta}) \cdot \prod_{i=1}^n (1 - F(\beta_0^s - \beta_1^s bid))^{y_{es_i}} \cdot (F(\beta_0^s - \beta_1^s bid))^{1-y_{es_i}} \\ \boldsymbol{\theta} &= \boldsymbol{\alpha}^p \\ \boldsymbol{\alpha}_i^p &\sim N(\bar{\alpha}_i^p, \tau^p) \end{aligned} \tag{6.10}$$

Für die Wahrscheinlichkeit der Ja-Antworten $1 - F(\beta_0^s - \beta_1^s bid)$ werden im Folgenden wiederum die logistische Verteilung sowie die Weibull-Verteilung und die lognormale Verteilung angenommen. Gleichung (6.10) ist somit die Basis für einen einfachen Bayes'schen Transfer, wenn für die Primärstudie die Konstante und der Gebotsparameter bekannt sind. Sie lässt sich trivial um den Einfluss der Kovariate erweitern. An diesem einfachen Bayes-Transfer lassen sich nun die Auswirkungen der Modellwahl und der Stichprobengröße vor Ort zunächst mithilfe von simulierten Daten studieren.

6.3.1 Einfacher Bayes-Transfer

Zunächst werden die Ergebnisse eines Bayes'schen Nutzentransfers für simulierte Daten gezeigt. Hierfür werden die simulierten Populationen aus dem fünften Kapitel herangezogen und Populationen für die zusätzlichen Studien "vor Ort" durch Ziehungen kleinerer Stichproben aus der großen Stichprobe im Sekundärgebiet generiert. Es gilt mehrere Effekte zu untersuchen. Zum einen soll der Einfluss der Stichprobe vor Ort auf den Nutzentransfer analysiert werden. Hier steht die Frage im Vordergrund, ob eine bestimmte Stichprobe sich verzerrend auf das Ergebnis auswirkt. Als nächster Schritt wird der Einfluss der A-Priori-Verteilung auf das Ergebnis des Nutzentransfers in verschiedenen Modellen untersucht, um einen häufig geäußerten Vorwurf gegen den Einsatz von Bayes'schen Verfahren zu berücksichtigen. Wenn die Ergebnisse eines Bayes'schen Nutzentransfers tatsächlich kritisch von der Wahl der verwendeten A-Priori-Verteilung abhängen, wäre der Transfer zu willkürlich, um wirklich in Betracht gezogen zu werden. Was allerdings zu erwarten ist, ist die Abhängigkeit der Breite der Konfidenzintervalle von der Spezifität der verwendeten Information, d.h. von der Breite der Konfidenzintervalle der A-Priori-Verteilungen. Dies liegt jedoch in der Definition des Ansatzes und stellt keine Verzerrung der Ergebnisse dar. Abschließend wird mit dem robustesten Modell der beiden vorangegangenen Schritte der Einfluss der Stichprobengröße im Sekundärgebiet auf das Ergebnis beschrieben, denn nur falls das Verfahren auch wirklich bei kleinen Stichproben funktioniert, ist es in der Praxis umsetzbar.

Zunächst werden aus der Gesamtpopulation im Sekundärgebiet verschiedene Stichproben gezogen, die alle einer simulierten (einfach gebundenen) Befragung unterzogen werden. Tabelle 6.2 zeigt eine Zusammenfassung der Antworthäufigkeiten der Ja-Antworten der simulierten Stichproben. Die

Tabelle 6.2: Stichproben aus dem Sekundärgebiet

Yes	WTP	95% Konfidenz	Yes	WTP	95% Konfidenz
<i>N</i> = 100			<i>N</i> = 200		
1	56	135,6 [119,1 - 152,4]	1	117	142,3 [128,9 - 156,8]
2	58	131,1 [116,1 - 146,7]	2	113	136,5 [120,2 - 152,6]
3	60	144,7 [124,2 - 168,9]	3	113	133,3 [120,7 - 146,2]
4	61	147,9 [131,4 - 166,5]	4	130	150,2 [132,6 - 170,5]
5	61	146,3 [128,5 - 165,7]	5	128	149,0 [135,6 - 163,1]
<i>N</i> = 500			<i>N</i> = 500		
6	64	151,9 [134,2 - 170,8]	1	294	139,5 [132,4 - 146,7]
7	61	138,5 [123,4 - 155,0]	2	302	143,9 [136,5 - 151,5]
8	60	138,2 [122,0 - 155,0]	3	294	139,7 [131,7 - 148,2]
9	62	147,2 [130,5 - 165,3]	4	298	139,8 [132,3 - 147,4]
10	64	156,2 [135,3 - 181,9]	5	309	145,8 [138,1 - 153,9]

Eigene Berechnungen.

dritte und die sechste Spalte der Tabelle weisen die jeweiligen Erwartungswerte der Zahlungsbereitschaften in der Stichprobe aus. Man sieht recht deutlich, dass die Mittelwerte bei den kleinsten Stichproben mit $N = 100$ sehr stark streuen und die Konfidenzintervalle auch sehr breit sind. Angenommen, diese Daten (gemessen in €/Jahr) seien die einzigen, die für eine politische Entscheidung vorliegen, dann läge die Einschätzung des Werts eines geplanten Projekts beispielsweise bei einer betroffenen Bevölkerung von 500.000 jährlich bei mindestens 59,5 Mio. € für die Stichprobe 1 und für die Stichprobe 10 bei bis zu 90 Mio. €, ein Unterschied, der sich auch auf die politische Entscheidung auswirken dürfte.

Je größer die Stichprobe wird, desto stärker nähert sich der Mittelwert dem "wahren" Wert an und desto geringer fallen die Schwankungen der Zahlungsbereitschaften aus. Der Mittelwert der kompletten simulierten Bevölkerung beträgt 141,96 mit einem Bootstrap-Konfidenzintervall von [139,12 - 147,74].

Auf dieser Datenbasis wird nun ein Nutzentransfer simuliert. $F(\beta_0^s - \beta_1^s bid)$ aus Gleichung (6.10) wird zunächst als logistische Verteilung modelliert. Die A-Priori-Verteilungen für die Konstante und den Gebotsparameter

Tabelle 6.3: Ergebnisse des Bayes'schen Transfers

	α_0	α_1	WTP	95% Konf.
1	4,72	-35,91	131,7	[118,3 - 146,3]
2	4,73	-37,22	127,3	[114,5 - 141,6]
3	4,63	-33,41	138,9	[124,7 - 154,7]
4	4,79	-33,71	142,4	[128,5 - 157,5]
5	4,75	-33,91	140,4	[125,9 - 156,0]
6	4,83	-33,37	145,0	[130,0 - 161,5]
7	4,77	-35,67	133,9	[120,9 - 148,2]
8	4,76	-35,76	133,3	[119,5 - 148,1]
9	4,79	-33,86	141,6	[127,6 - 156,6]
10	4,71	-31,89	148,0	[132,6 - 165,1]

$N = 100$

Eigene Berechnungen.

werden aus den geschätzten Parametern der Primärstudie gebildet. In diesem Schritt wird eine normale A-Priori-Verteilung gewählt, die als Mittelwert den Schätzwert aus der Primärstudie bekommt und die halbe Präzision ($\alpha_k \sim N(\alpha_k^p, \tau_k^p/2)$). Die Präzision wird so gewählt, dass eine Parameteranpassung durch die Information aus dem Sekundärgebiet stattfinden kann. Legt man die A-Priori-Verteilung zu eng fest, so erhält man im Prinzip dieselben Ergebnisse, wie bei der Verwendung der klassischen Nutzentransferfunktion, denn die Parameter können sich nicht aufgrund der neuen in den Daten enthaltenen Informationen anpassen. Unser "Anfangsverdacht" für die Größe der verschiedenen Parameter des statistischen Modells basiert somit auf gemessenen Ergebnissen der Primärstudie und genügt somit dem Jaynes-Kriterium.

Tabelle 6.3 zeigt eine Übersicht der Berechnungen für 10 verschiedene Stichproben mit Umfang $N = 100$. Die A-Priori-Verteilungen für die Parameter sind $\alpha_0 \sim N(4, 50; 8)$ und $\alpha_1 \sim N(41, 8; 0, 05)$. Es wurden 20.000 Iterationen gerechnet, nach einer Burn-in-Phase von 4000 Iterationen. Alle Durchläufe zeigen eine Gelman-Rubin Statistik nahe bei eins. Die jeweilige Stichprobe beeinflusst den Bayes'schen Nutzentransfer. Wenn der

Mittelwert der Zahlungsbereitschaft der Stichprobe erheblich vom wahren Wert von 141,96 abweichen, wie beispielsweise bei Stichprobe 2, die erheblich nach unten abweicht, oder bei Stichprobe 10, die deutlich nach oben abweicht, so schlägt dieses Verhalten auch auf den Nutzentransfer durch. Im Mittel sind die Bayes'schen Transferwerte deutlich näher am wahren Wert, als die durch den herkömmlichen Ansatz, der die Transferfunktion einsetzt. Dort liegt der transferierte Wert bei rund 133, während der Mittelwert der obigen Beispiele 138,5 beträgt. Im Vergleich zu den aus den Unterstichproben errechneten Mittelwerten sind die transferierten systematisch kleiner, denn die Primärstudie weist einen geringeren Erwartungswert der Zahlungsbereitschaft auf. Im Durchschnitt überschätzen die Einzelschätzungen den Mittelwert mit 147,5. Daher ist der Bayes'sche Transfer auch für den Planer eine gute Nachricht im Sinne einer möglichst konservativen Schätzung. Die Konfidenzintervalle werden durch die Verwendung der informativen A-Priori-Verteilung verringert, besonders an der oberen Grenze. Das in Kapitel 2 bereits hergeleitete Prinzip der Verringerung von Konfidenzintervallen durch die Verwendung einer informativen A-Priori-Verteilung, wird an dieser Stelle durch die Simulation illustriert. Im Vergleich zum herkömmlichen Nutzentransfer ist die Möglichkeit, überhaupt Konfidenzintervalle anzugeben, ohnehin eine erhebliche Verbesserung.

Welchen Fehler aber macht man, wenn man eine andere A-Priori-Verteilung, beispielsweise eine informativere Verteilung annimmt? Ist dieser Fehler von dem gewählten Modell für die Daten der Sekundärstudie abhängig? Diesen Fragen wird nachgegangen, indem die Daten aus Stichprobe 3 der Sekundärstudie mit dem Weibull-Modell, dem lognormalen Modell und dem Logit-Modell geschätzt werden, und diese Ansätze mit verschiedenen A-Priori-Verteilungen kombiniert werden. Tabelle 6.4 zeigt eine Übersicht über die Ergebnisse.

In den drei Modellen wurde jeweils zunächst eine informative A-Priori-Verteilung gewählt, deren Mittelwert die Schätzwerte des jeweiligen Modells in der Primärstudie sind. Zu allen Modellen "passt" die Normalverteilung als zulässige A-Priori-Verteilung. Wenn man beim Weibull-Modell dazu übergehen würde, die Zahlungsbereitschaft direkt als zu schätzende Größe zu verwenden, wie es Leon et al. (2002) [169] vorschlagen, kann es hilfreich sein, als A-Priori-Verteilung eine Gamma-Verteilung anzunehmen (vgl. auch Ibrahim et al. (2003)[140]). Die Modelle werden durch diesen Schritt aber anfälliger für die Wahl der Breite der A-Priori-Verteilung. Im Prinzip lässt sich ein Befund aus den vorherigen Kapiteln bestäti-

Tabelle 6.4: Unterschiedliche Modelle

	π_0	Wert	95% Konfidenz
Logit-Modell, $WTP_{full} = 140, 1$			
α_0	$N(0, 10^{-6})$	3,4	[2,2 - 5,0]
α_1	$N(0, 10^{-6})$	-23,6	[-35,0 - -14,7]
WTP	-	144,9	[124,5 - 169,4]
α_0	$N(4, 9; 0, 1)$	6,2	[5,6 - 6,9]
α_1	$N(-45, 2; 1, 0)$	-44,64	[-46,5 - -42,8]
WTP	-	139,0	[125,9 - 152,0]
α_0	$N(4, 9; 1)$	6,2	[5,6 - 6,7]
α_1	$N(-45, 2; 10)$	-45,1	[-45,7 - -44,47]
WTP	-	136,8	[124,4 - 148,8]
α_0	$N(4, 9; 10)$	5,64	[5,21 - 6,05]
α_1	$N(-45, 2; 50)$	-45,1	[-45,4 - -44,9]
WTP	-	124,9	[115,1 - 134,0]
Log-Logit-Modell, $WTP_{full} = 135, 8$			
α_0	$N(0, 10^{-6})$	14,6	[9,2 - 19,5]
α_1	$N(0, 10^{-6})$	-297,7	[-398,2 - -185,4]
WTP	-	136,8	[115,7 - 163,2]
α_0	$N(21, 6; 1, 0)$	22,7	[22,1 - 23,2]
α_1	$N(-466, 6; 0, 1)$	-465,7	[-471,7 - -459,8]
WTP	-	132,3	[118,1 - 147,1]
α_0	$N(21, 6; 10)$	22,3	[21,9 - 22,7]
α_1	$N(-466, 6; 1)$	-466,2	[-468,2 - -464,3]
WTP	-	119,8	[109,8 - 130,6]
Weibull-Modell, $WTP_{full} = 130, 4$			
α_0	$N(0, 10^{-6})$	9,2	[5,9 - 13,4]
α_1	$N(0, 10^{-6})$	-1,9	[-2,8 - -1,3]
WTP	-	130,7	[110,5 - 157,4]
α_0	$N(10, 5; 3)$	11,6	[9,4 - 13,7]
α_1	$N(-2, 5; 0, 5)$	-2,5	[-2,9 - -2,0]
WTP	-	119,9	[104,5 - 137,4]
α_0	$N(10, 5; 6)$	11,9	[10,1 - 13,6]
α_1	$N(-2, 5; 1)$	-2,6	[-2,9 - -2,2]
WTP	-	112,6	[97,7 - 129,1]

Eigene Berechnungen.

Tabelle 6.5: Einfluss der Stichprobengröße

50	100	150	200	500
138,5	139,0	141,6	137,1	138,9
[121 - 155, 8]	[125,9 - 152,0]	[131,0 - 152,3]	[128,1 - 145,9]	[133,1 - 144,7]
Eigene Berechnungen.				

gen: das Logit-Modell ist am wenigsten anfällig für Verzerrungen. Allerdings ist auch im Logit-Modell bei einer zu spezifischen Festlegung der A-Priori-Verteilung der Parameter das Ergebnis stark in die Richtung der Primärstudie verzerrt. Eine moderat gewählte A-Priori-Verteilung hingegen schätzt die Zahlungsbereitschaft besser als der einfache Nutzentransfer und gibt darüber hinaus noch eine Verbesserung der Konfidenzintervalle gegenüber der Schätzung auf Basis der mit $N = 100$ recht kleinen Stichprobe.

Abschließend wird in diesem Abschnitt genau dieser Einfluss untersucht. Wie wirkt sich die Stichprobengröße auf den Bayes'schen Nutzentransfer aus? Natürlich wird sich die Schätzung mit zunehmender Datenmenge aus dem Sekundärgebiet verbessern und letztlich muss der politische Entscheidungsträger abwägen, bei welcher Stichprobengröße die durch Verzicht auf weitere Interviews eingesparten Kosten den Informationsverlust rechtfertigen. Tabelle 6.5 fasst die Schätzung verschiedener Stichproben ($N = 50, 100, 150, 200, 500$) mit dem Logit-Modell und den informativen A-Priori-Verteilungen $\alpha_0 \sim N(4, 9; 0, 1)$ und $\alpha_1 \sim N(-45, 2; 1, 0)$ zusammen. Es zeigt sich, dass der Stichprobenumfang sich insbesondere auf die Breite des Konfidenzintervalls auswirkt. Die Erwartungswerte der Zahlungsbereitschaft hingegen unterscheiden sich in diesem simulierten Beispiel nicht erheblich. Allerdings kann eine kleine Stichprobe realer Daten in stärkerem Umfang von Ausreißern gekennzeichnet sein, sodass hier noch ein erheblicher Forschungsbedarf an real durchgeführten Studien besteht. Versucht man den Einfluss auf die Konfidenzintervalle durch informativere A-Priori-Verteilungen zu mindern, würde man allerdings den Wert zu stark in Richtung der Primärstudie "drücken". Dies wirft insgesamt die Frage auf, ob der Primärstudie tatsächlich das gleiche Gewicht beigemessen werden sollte, oder ob es Möglichkeiten gibt, den Einfluss der Primärstudie auf die Schätzung der Daten aus dem Sekundärgebiet zu kontrollieren. Dieser Frage sind auch Ibrahim et al. (1999)[138] nachgegangen und schlagen zur

Lösung des Problems eine besondere Formulierung der A-Priori-Verteilung vor, den sogenannten Power Prior.

6.3.2 Der Power Prior

Der Power Prior biete eine Möglichkeit, "historische Daten" in die Schätzung der Ergebnisse eines neuen Experiments einzubeziehen und zu gewichten. Diese von Ibrahim und Chen (2000) [137] oder Ibrahim et al. (2003) [140] genutzte Form der A-Priori-Verteilung bietet die Möglichkeit, historische Datensätze in die Berechnung einzubeziehen, und liefert gleichzeitig die Möglichkeit, den Einfluss der historischen Daten auf die in der aktuellen Studie gemessenen Werte zu kontrollieren. Da der Power Prior in der Bewertungsliteratur bislang nicht wahrgenommen wurde, werden seine Konstruktion und seine Eigenschaften nachfolgend etwas ausführlicher erläutert.

Angenommen, es liegen Daten $D_0 = (n_0, \mathbf{y}_0, \mathbf{s}_0)$ aus einer Studie vor, wobei n_0 die Anzahl der Beobachtungen angibt, \mathbf{y}_0 den $n_0 \times 1$ Antwortvektor und \mathbf{s}_0 die $N_0 \times p$ Matrix der Kovariate (vgl. Ibrahim et al. (1999) [138]). Diese Daten lassen sich zur Konstruktion einer A-Priori-Verteilung für die Schätzung von Erwartungswerten mit Daten $D = (n, \mathbf{y}, \mathbf{s})$ aus einer aktuell durchgeführten Befragung verwenden. Das Modell lässt sich aus der Likelihoodfunktion $L(\boldsymbol{\theta}|D)$ und eine A-Priori-Verteilung der Parameter $\pi_0(\boldsymbol{\theta}|c_0)$ konstruieren. Der Power Prior fügt diesem statistischen Modell den Einfluss einer bereits durchgeführten Studie hinzu und ist definiert als

$$\pi(\boldsymbol{\theta}|D_0, a_0) \sim L(\boldsymbol{\theta}|D_0)^{a_0} \pi_0(\boldsymbol{\theta}|c_0). \quad (6.11)$$

Hierbei ist c_0 ein Hyperparameter für die Verteilung π_0 und a_0 ist eine Zahl, die den Einfluss der Primärstudie auf die Likelihoodfunktion der Sekundärstudie ausdrückt. Sinnvollerweise wird diese Zahl auf $0 \leq a_0 \leq 1$ beschränkt. Wenn $a_0 = 0$ ist, beschreibt dies den Fall, in dem der Primärstudie kein Einfluss zugewiesen wird und $a_0 = 1$ beschreibt den Fall, in dem die Daten der Primärstudie und die Daten der Sekundärstudie gleich gewichtet werden. Im Wesentlichen kontrolliert a_0 die Dicke der Enden der A-Priori-Verteilung. Gleichung (6.11) kann somit als Verallgemeinerung des üblichen Bayes'schen Aufdatierungsschemas angesehen werden, die es erlaubt, den Einfluss der Primärstudie auf die A-Posteriori-Verteilung der Parameter der Sekundärstudie zu kontrollieren. Dies ist besonders angezeigt, wenn Sekundärstudie und Primärstudie Unterschiede aufweisen, wie

es bei der Wahl von Studien zum Nutzentransfer beinahe immer der Fall ist.

Bezieht man den Power Prior in die Definition des statistischen Bayes-Modells mit ein, so muss dieses Modell als hierarchisches Modell spezifiziert werden. Zu jedem neu in das Modell eingeführten Parameter muss wiederum eine A-Priori-Verteilung eingeführt werden, die die Grundannahmen bezüglich dieses Parameters widerspiegelt. Daher muss für eine vollständige Spezifikation des statistischen Modells eine A-Priori-Verteilung für α_0 definiert werden. Eine natürliche Wahl für diese Verteilung ist die Beta-Verteilung, da sie auf dem Intervall zwischen 0 und 1 definiert ist, aber eine Fülle flexibler Formen zulässt. Alternative Spezifikationen sind die gestutzte Normalverteilung oder die gestutzte Gamma-Verteilung (vgl. Ibrahim et al. (2001)[139]), die sehr ähnliche Eigenschaften haben. Wegen dieser ähnlichen Eigenschaften konzentriert sich die Darstellung im Folgenden auf die Verwendung der Beta-Verteilung.

Obwohl die A-Priori-Verteilung in Gleichung (6.11) in der Regel nicht in einer geschlossenen Form darstellbar ist, lässt sich zeigen (vgl. Chen et al. (1999) [65], Ibrahim und Chen (2000) [137] und Ibrahim et al. (2001)[139]), dass für zulässige A-Priori-Verteilungen für α_0 auch (6.11) eine zulässige A-Priori-Verteilung darstellt. Die A-Posteriori-Verteilung ergibt sich dann als

$$\pi(\boldsymbol{\theta}|x) \propto L(\boldsymbol{\theta}|\mathbf{x}_0)^{\alpha_0} \cdot L(\boldsymbol{\theta}|\mathbf{x}) \cdot \pi_0(\boldsymbol{\theta}|c_0) \cdot \pi(\alpha_0|\gamma_0). \quad (6.12)$$

Der große Vorteil des Power Priors ist, dass der Einfluss der Primärstudien- daten auf die Sekundärstudie kontrolliert werden kann, indem man unterschiedliche A-Priori-Verteilungen für α_0 annimmt und Sensibilitätstests mit verschiedenen Werten von α_0 für die ermittelten Parameter durchführt. Der Ansatz hat seinen Ursprung in der medizinischen Forschung, bei der Experimente häufig unter sehr ähnlichen Bedingungen durchgeführt werden und das einzelne Experiment häufig eher kleine Stichproben aufweist, so dass die Ergebnisse früherer Experimente die Schätzungen verbessern können. Der Stichprobenumfang bei der medizinischen Forschung ist nicht nur durch Sparsamkeit begrenzt, sondern häufig auch durch ethische Überlegungen, beispielsweise bei Experimenten mit tödlichem Ausgang.

Obwohl die Tatsache, dass sie befragt werden, für die Einwohner der Stadt Cottbus hoffentlich nicht solch schreckliche Nebenwirkungen hat, kann der Power Prior auch einen interessanten Ansatz für Nutzentransferstudien bieten. Gerade die Kontrolle über den Einfluss der Vorabinformationen auf die Schätzung, macht ihn auch in den Gebieten attraktiv, in denen man

keine exakte Übereinstimmung zwischen dem "historischen Experiment" und dem "neuen Experiment" erwarten kann.

6.4 Bayes'scher Nutzentransfer für den Cottbuser Ostsee

In der Simulation lässt sich die Verbesserung des Nutzentransfers durch die Verwendung von Bayes'schen Verfahren nachvollziehbar darstellen, wie aber sieht es bei der Verwendung "echter" Daten aus? Im letzten Abschnitt dieser Arbeit wird das Beispiel aus Kapitel 1 und 2 erneut aufgegriffen und es wird der Nutzentransfer für die Bewertung des Reaktivierungsprojekts "Cottbuser Ostsee" durchgeführt, wobei die Ergebnisse der kompletten Studie als Referenzfall dienen.

Aus der dänischen Studie zum Tokkekob Hegn werden im Folgenden die A-Priori-Verteilungen gewonnen. Dabei stößt man zunächst auf ein Hindernis: die dänische Studie wurde im offenen Befragungsformat durchgeführt, die Ergebnisse der Cottbuser Studie hingegen liegen im doppelt gebundenen Format vor. Um diese beiden Formate zu "vereinen" gibt es zwei grundsätzlich verschiedene Vorgehensweisen, von denen jede ihre Vor- und Nachteile hat. Eine Möglichkeit besteht darin, dass man bei der Cottbuser Studie mit einem Probit-Modell arbeitet und die mit einem Regressionsmodell geschätzten Parameterwerte der dänischen Studie direkt als Mittelwerte normalverteilter A-Priori-Verteilungen einsetzt.

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{x}) &\propto \pi(\boldsymbol{\theta}) \cdot \prod_{i=1}^n (1 - \Phi(\beta_0^s - \beta_1^s bid + \boldsymbol{\beta}^s \mathbf{x}))^{yes_i} \\ &\quad \cdot (\Phi(\beta_0^s - \beta_1^s bid + \boldsymbol{\beta}^s \mathbf{x}))^{1-yes_i} \\ \boldsymbol{\theta}_i &= \alpha_i^p \cdot \beta_1^s \\ \beta_1^s &\sim N(0, 10^{-6}) \\ \alpha_{i \neq 1}^p &\sim N(\bar{\alpha}_{i \neq 1}^p, \tau^p) \end{aligned} \tag{6.13}$$

Dieser Ansatz hat den Nachteil, dass er wiederum auf Kovariate angewiesen ist, wie der einfache Bayes-Ansatz von Parsons und Kealy (1994) [201]. Der Grund liegt darin, dass das offene Format den Preis direkt abfragt und somit keine Informationen über die Reaktion *auf einen gegebenen Preis* beinhaltet. Möchte man diese Eigenschaft der Dichotomous Choice

Daten erhalten, bleibt als weitere Möglichkeit nur der Griff zur Simulation. Aus den offen geäußerten Zahlungsbereitschaften werden sinnvolle Gebote konstruiert und es wird die Reaktion der Bevölkerung auf diese Gebote simuliert, indem die einfache Entscheidungsregel aus Kapitel 1.2 verwendet wird⁴. So lässt sich hier eine Folge von Ja- und Nein-Antworten konstruieren, die es erlaubt, eines der im vorherigen Abschnitt vorgestellten Modelle zu verwenden. Dieses Verfahren findet für den Nutzentransfer an dieser Stelle eine erstmalige Anwendung. Es wird eine einfach gebundene Befragung simuliert und bei der Cottbuser Studie die Annahme beibehalten, dass die Antworten auf das doppelte Format der gleichen Zahlungsbereitschaften der Befragten entstammen (vgl. Kapitel 1.3.3). So sind die Parameter der Gleichung für die Zahlungsbereitschaften für beide Studien vergleichbar. Der hier verwendete einfache Schätzansatz lautet dann in allgemeiner Form

$$\begin{aligned}\pi(\boldsymbol{\theta}|\mathbf{x}) &\propto \pi(\boldsymbol{\theta}) \cdot \prod_{i=1}^n (1 - \Phi(\beta_0^s - \beta_1^s \text{bid} + \boldsymbol{\beta}^s \mathbf{x}))^{y_{es_i}} \\ &\quad \cdot (\Phi(\beta_0^s - \beta_1^s \text{bid} + \boldsymbol{\beta}^s \mathbf{x}))^{1-y_{es_i}} \\ \boldsymbol{\theta} &= \boldsymbol{\alpha}^p \\ \alpha_i^p &\sim N(\bar{\alpha}_i^p, \tau^p).\end{aligned}\tag{6.14}$$

Für den Power Prior wird der Ansatz zu

$$\begin{aligned}\pi(\boldsymbol{\theta}|\mathbf{x}) &\propto L(\boldsymbol{\alpha}|\text{bid}^p, \mathbf{x}^p)^{a_0} \prod_{i=1}^n (1 - \Phi(\beta_0^s - \beta_1^s \text{bid}^s + \boldsymbol{\beta}^s \mathbf{x}^s))^{y_{es_i}} \\ &\quad \cdot (\Phi(\beta_0^s - \beta_1^s \text{bid}^s + \boldsymbol{\beta}^s \mathbf{x}^s))^{1-y_{es_i}} \\ \alpha_i^p &\sim N(\bar{\alpha}_i^p, \tau^p) \\ a_0 &\sim \text{beta}(400, 1).\end{aligned}\tag{6.15}$$

Tabelle 6.6 zeigt eine Übersicht über die erzeugten "Befragungsergebnisse". Der Erwartungswert der Zahlungsbereitschaft aus einem einfachen Logit-Modell mit dem Gebot als einziger Variablen liegt für die Simulation bei 181,95 DKr, somit zwischen den in der Originalstudie ermittelten Extremwerten, die aus der jeweiligen Behandlung der Nein-Antworten als Protestantworten resultieren. Im Folgenden werden mit den Originaldaten der dänischen Studie sowie mit der aufgrund der Originaldaten simulierten geschlossenen Befragung die verschiedenen Bayes'schen Nutzen-

⁴Für ein ähnliches Verfahren vgl. Boyle et al. (1996) [43].

Tabelle 6.6: Tokkekob simulierte Befragung

Gebot ^p	yes	no	Insgesamt
50	86	256	351
100	126	165	291
400	60	220	280
300	165	67	232
500	216	28	244
Insgesamt	585	813	1398

Eigene Berechnungen.

transferansätze angewendet und der Erwartungswert der Zahlungsbereitschaft für den Cottbuser Ostsee aufgrund einer kleinen Stichprobe mit $n = 100$ ermittelt. Dabei wird zunächst wiederum die Abhängigkeit des transferierten Werts von der jeweiligen Stichprobe überprüft und die Ergebnisse der verschiedenen Ansätze werden miteinander verglichen. Tabelle 6.7 zeigt zunächst die Ergebnisse des einfachen Bayes-Transfers für ein Modell mit Kovariaten. Die Wahl der Kovariate orientiert sich in dieser Rechnung an dem Modell, das in der dänischen Studie verwendet wurde. Die dänische Studie wurde nach Gleichung (6.13) dazu verwendet, die A-Priori-Verteilungen um die dort geschätzten Parameterwerte als Normalverteilung zu konstruieren, für den Gebotsparameter wird eine nicht-informative Verteilung angenommen, da hierzu aus der offenen Befragung keine Informationen vorliegen. Alle zusätzlichen Variablen sind direkt vor Ort erhoben worden, im Gegensatz zur Verwendung von Mittelwerten beim Function Transfer, bei denen auf Mittelwerte aus dem Land Brandenburg zurückgegriffen wird. Die transferierten Werte in Tabelle 6.7 sind systematisch niedriger und stellen kaum eine bis gar keine Verbesserung gegenüber den herkömmlichen Verfahren dar. Die Einflüsse der erklärenden Variablen sind sehr unterschiedlich zwischen den Stichproben, sie weisen zum Teil sogar ein unterschiedliches Vorzeichen auf. Dies ist zwar kein Widerspruch zur ökonomischen Theorie, denn die Variablen, bei denen uns die Theorie ein bestimmtes Vorzeichen erwarten lässt, weisen dieses auch auf. Hierzu gehört das positive Vorzeichen von α_{income} . Welcher Art auf der anderen Seite der Einfluss einer aktiven oder passiven Mitgliedschaft bei Non-Governmental Organizations (NGOs) tatsächlich ist, dazu lässt sich zunächst keine Erwartungshaltung formulieren.

Tabelle 6.7: Einfacher Bayes-Transfer mit Kovariaten

	Stichprobe 1	Stichprobe 2	Stichprobe 3
<i>WTP</i>	1.26 [-3.78 - 7.56]	2.01 [-2.97 - 8.10]	2.89 [-1.3 - 6.87]
α_{bid}	-0.0896 [-0.99 - -0.68]	-0.0915 [-0.99 - -0.73]	-0.0937 [-0.99 - -0.79]
α_0	50.05 [4.5 - 109.0]	11.75 [-19.32 - 29.22]	63.4 [10.63 - 145.7]
α_{income}	0.012 [0.004 - 0.02]	0.010 [-0.01 - 0.012]	0.002 [0.005 - 0.009]
α_{age}	-0.03 [-0.036 - 0.0]	-0.40 [-0.72 - -0.07]	-0.03 [-0.3 - 0.0]
$\alpha_{walkoft}$	0.52 [-0.91 - 1.75]	3.10 [0.59 - 5.53]	0.85 [0.0 - 2.89]
α_{single}	1.8 [0.0 - 24.71]	16.76 [7.45 - 52.35]	1.03 [-10.59 - 12.73]
α_{ngo}	7.71 [5.32 - 24.87]	-10.64 [-32.71 - 0.86]	21.33 [-1.86 - 47.05]

10000 Iterationen mit 3000 Burn-in. Eigene Berechnungen.

Man könnte sich auf den Standpunkt stellen, wer engagiert in NGOs mitmacht, hat auch eine höhere Zahlungsbereitschaft oder die gegenläufige Argumentation vertreten, dass die aktive Mitgliedschaft dazu führt, dass sich die Leute für diesen speziellen Zweck besonders stark engagieren und ihre Ressourcen gewissermaßen gebunden sehen. Darüber hinaus ist die Mitgliedschaft in derartigen Organisationen in der Cottbuser Stichprobe sehr gering, so dass sie vermutlich keinen systematischen Einfluss entfaltet. Um ohne Annahmen über den möglichen Einfluss von zusätzlichen Variablen auszukommen, werden im nächsten Schritt die Daten aus der simulierten Befragung aus der dänischen Studie für den Bayes-Transfer nach Gleichung (6.14) ohne Kovariate herangezogen. Tabelle 6.8 zeigt eine Übersicht über die Ergebnisse dieses Transfers wiederum für drei Stichproben vom Umfang $N = 100$ im Vergleich. Die transferierten Zahlungsbereitschaften sind im Durchschnitt und im Einzelnen viel näher an den in der umfangreichen

Tabelle 6.8: Bayes-Transfer mit simulierter Befragung

	Stichprobe A	Stichprobe B	Stichprobe C
<i>WTP</i>	3.86 [3.55 - 4.18]	3.90 [3.59 - 4.21]	4.22 [3.89 - 4.55]
α_{bid}	-0.86 [-0.89 - -0.83]	-0.86 [-0.89 - -0.83]	-0.85 [-0.87 - -0.82]
α_0	3.31 [3.05 - 3.57]	3.36 [3.10 - 3.62]	3.57 [3.31 - 3.83]

35000 Iterationen mit 3000 Burn-in. Eigene Berechnungen.

Studie gemessenen Zahlungsbereitschaften, allerdings schließen die recht engen Konfidenzintervalle die wahre Zahlungsbereitschaft in zwei Fällen knapp nicht ein. Im Modellvergleich mittels des Devianzkriteriums (vgl. Spiegelhalter et al. 2002 [231]) erweist sich das Modell mit den simulierten Daten leicht überlegen. Andererseits misst dieses Kriterium die Anpassung des Modells an die kleine Stichprobe, ein auch in der Realität gültiges Kriterium für die Anpassung an unbekannte Daten liegt nicht vor.

Abschließend wird nun die Veränderung des Nutzentransfers durch die Verwendung des Power Priors untersucht. Zur Konstruktion des Power Priors wird die Likelihoodfunktion der dänischen Studie mit dem Exponenten $0 \leq a_0 \leq 1$ potenziert, der den Einfluss dieser Studie auf die Schätzung der Zahlungsbereitschaft für den Cottbuser Ostsee beinhaltet (vgl. Gleichung (6.12)). Zunächst wird der Nutzentransfer mit dem Power Prior für die drei Stichproben A, B und C durchgeführt, damit sich die Veränderung des Transferwerts beobachten lässt. Die A-Priori-Verteilung von a_0 ist eine Beta-Verteilung ($\text{beta}(400,1)$), die der dänischen Studie ein vergleichsweise großes Gewicht beimisst. Die Hyperparameter aus Gleichung (6.12) erhalten eine nicht-informative A-Priori-Verteilung. Tabelle 6.9 zeigt die Ergebnisse dieses Transfers für eine Stichprobengröße von $N = 1000$ bei der dänischen Studie. Die Erwartungswerte der Zahlungsbereitschaften liegen deutlich näher bei dem wahren Wert von 4,39 € der kompletten Cottbuser Studie, allerdings sind die Konfidenzintervalle ähnlich. Das entspricht den Erwartungen, wenn man sich die Funktionsweise des Power Priors vor Augen führt. Es kann kaum zu einer starken Verengung der Konfidenzintervalle kommen, da die Daten der Primärstudie in die Likelihoodfunktion eingesetzt wurden und nur die Gewichtung dieser Daten verändert wur-

de. Die Verschiebung in Richtung des tatsächlichen Mittelwerts lässt sich dadurch verstehen, dass die (niedrigere) mittlere Zahlungsbereitschaft der dänischen Studie ein kleineres Gewicht bekommt als beim einfachen Transfer, bei dem beide Studien gleichermaßen gewichtet sind. Der Power Prior nimmt genau den Wert an, der den Kullback-Leibler-Abstand zwischen der A-Posteriori-Verteilung ohne jeglichen Einbezug der Vorabinformationen und der A-Posteriori-Verteilung, bei welcher die Vorabinformationen mit demselben Gewicht eingehen wie die Daten, minimiert (vgl. Ibrahim et al. (2003) [140]).

Im Folgenden sind zwei Dinge interessant: Wie verändert sich dieses Gewicht, wenn mehr Daten von der Cottbuser Studie zur Verfügung stehen? Und wie verändert sich der geschätzte Wert, wenn die A-Priori-Verteilung des Exponenten verändert wird? Zur Beantwortung der ersten Frage werden vier verschiedene Unterstichproben aus den Cottbuser Daten mit $N = 100, 200, 300$ und 400 gezogen. Wie bereits für das simulierte Beispiel gezeigt wurde, verringert sich auch hier der Einfluss der Primärstudie auf die Schätzung mit zunehmender Größe der Sekundärstudie. Dies ist beim Einsatz des Power Priors genau das erwünschte Ergebnis, denn mit zunehmender direkter Information über die Zahlungsbereitschaft der tatsächlich betroffenen Bevölkerung sollte die Primärstudie an Gewicht verlieren.

Während der Einfluss der Primärstudie sich bei einem einmal gewählten Modell automatisch ergibt, muss abschließend überlegt werden, wie sich gerade die Modellwahl auf die Ergebnisse auswirkt. Dies ist die zweite der eingangs formulierten Fragen. In dem in Gleichung (6.12) aufgestellten Modell muss über drei A-Priori-Verteilungen entschieden werden. Zunächst gilt es, die A-Priori-Verteilungen zu den Parametern α_0 und α_1 zu modellieren. Zu diesen Parametern liegen in diesem Fall keine zusätzlichen Informationen vor. Tabelle 6.11 zeigt dennoch, was passiert, wenn man die A-Priori-Verteilung zu den Parametern als informative Verteilung model-

Tabelle 6.9: Power Prior

	Stichprobe A	Stichprobe B	Stichprobe C
WTP	4,29	4,18	4,22
	[3,87 - 4,77]	[3,79 - 4,59]	[4,07 - 5,03]
α_0	0,746	0,786	0,741

35000 Iterationen mit 3000 Burn-in. Eigene Berechnungen.

Tabelle 6.10: Der Einfluss der Stichprobengröße auf a_0

	100	200	300	400
a_0	0,76	0,65	0,65	0,63

35000 Iterationen mit 3000 Burn-in. Eigene Berechnungen.

liert. Als Beispiel wird wiederum die Stichprobe mit $N = 100$ herangezogen und die Modellierung ist $\alpha_i \sim N(\alpha_{tokke}, \tau_{tokke})$. Diese Modellierung drängt das System stärker als zuvor in die Richtung der Primärstudie, was sich am niedrigen Wert für die Zahlungsbereitschaft und am eher hohen Wert für a_0 zeigt. Es wäre interessant, beispielsweise aus Expertenbefragungen eine zusätzliche Information über die Parameter zu erhalten und dann ähnlich vorzugehen, wie es im zweiten Kapitel anhand simulierter Daten gezeigt wurde. Als nächstes werden verschiedene Modellierungen für die A-Priori-Verteilung des Exponenten a_0 durchgeführt. Die Beta-Verteilung $\text{beta}(300,1)$ in Tabelle 6.11 legt a priori ein geringeres Gewicht auf die Primärstudie, daher ist auch das Ergebnis für den Exponenten geringer, die höhere Zahlungsbereitschaft der Stichprobe erhält mehr Gewicht und beeinflusst den Schätzwert.

Die zweite Beta-Verteilung verhält sich gerade umgekehrt und gibt der Primärstudie ein größeres Gewicht. Hier fällt die Zahlungsbereitschaft geringer aus und der Einfluss der Primärstudie liegt nahe bei eins. Insgesamt ist die Beta-Verteilung $\text{beta}(400,1)$ in der Literatur für Choice Daten vielfach eingesetzt worden (vgl. Imbrahim et al. (2001) [139] für eine Fülle von Beispielen), und führt auch für den Cottbuser Nutzentransfer zu zufriedenstellenden Ergebnissen.

Tabelle 6.11: Der Einfluss der A-Priori-Verteilung

	$\alpha_i \sim N(\alpha_{tokke}, \tau_{tokke})$	$a_0 \sim \text{beta}(300, 1)$	$a_0 \sim \text{beta}(500, 1)$
WTP	3,48	4,66	4,01
	[3,29 - 3,67]	[4,10 - 5,31]	[3,71 - 4,35]
a_0	0,89	0,54	0,96

35000 Iterationen mit 3000 Burn-in. Eigene Berechnungen.

6.5 Konfidenzintervalle und ein neuer Test für den Nutzentransfer

In diesem Kapitel wurden verschiedene Anwendungen des Bayes'schen Nutzentransfers anhand von simulierten Daten und anhand der erhobenen Daten des Cottbuser Ostsees und des Tokkekob Hegn gezeigt. In diesem letzten Abschnitt wird ein neuer Test zur Überprüfung von Nutzentransferergebnissen vorgestellt.

In Kapitel 5.3 wurde auf die Testverfahren eingegangen, die in der Literatur zur Überprüfung der herkömmlichen Nutzentransferverfahren vorgeschlagen werden. Diese Tests setzen voraus, dass der wahre Wert im Sekundärgebiet bekannt ist, d.h. sie sind nur anwendbar in Überprüfungsstudien zum Nutzentransfer. In der "rauen Wirklichkeit" des Nutzentransfers hingegen lassen sich bereits die Konfidenzintervalle für den Transferwert nicht berechnen. Bei Anwendung des Valuetransfer wird der in einer Primärstudie ermittelte Wert übertragen, Aussagen zum Konfidenzintervall lassen sich somit nur bezüglich der Primärstudie treffen, nicht jedoch bezüglich des Transferwerts, der Nutzentransfer mit der Transferfunktion aus Kapitel 5 ermöglicht ebenso wenig die Berechnung eines Konfidenzintervalls des Transferwerts, denn die Parameter einer Primärstudie werden mit den Mittelwerten der sozioökonomischen Daten für die Sekundärstudie multipliziert. Als Konfidenzintervall des Transferwerts kann sich nichts ergeben, was über die ursprüngliche Parameterschätzung hinausgeht.

Anders hingegen bei den Bayes'schen Verfahren. Sie erlauben nicht nur die Angabe von Konfidenzintervallen, wie im vorigen Abschnitt gezeigt, sondern ermöglichen durch die unmittelbare Erhebung von Daten im Sekundärgebiet einen völlig neuartigen Test für die Transferwerte. Während herkömmliche Nutzentransferverfahren implizit gleiche Erwartungswerte für die Sekundär- und die Primärstudie voraussetzen, ist dies bei den Bayes'schen Verfahren nicht der Fall. Die Gleichheit der Erwartungswerte von Primär- und Sekundärstudie ist keine notwendige Voraussetzung für einen gelungenen Transfer, denn die Stärke des Bayes'schen Transfers liegt gerade darin, dass unterschiedliche Präferenzen an verschiedenen Orten zulässig sind. Hier bietet sich die Überprüfung der Umkehrbarkeit des Transfers als ein möglicher Test an. Auf unser Beispiel übertragen heißt dies, dass der Transfer ebenso zu vernünftigen (und richtigen) Ergebnissen führen muss, wenn man die Cottbuser Studie als Primärstudie heranzieht und aus der dänischen Studie eine kleine Stichprobe zieht, die hypothe-

Tabelle 6.12: Nutzentransfer Cottbus \rightarrow Tokkekob

	Mittelwert	95% Konfidenz
$WTP_{true} = 181,95$		
Stichprobe ₁	183,7	[171,3 - 195,9]
Stichprobe ₂	179,8	[167,3 - 192,1]
Stichprobe ₃	183,0	[170,5 - 195,1]
Stichprobe ₄	183,6	[171,0 - 196,0]
Stichprobe ₅	186,8	[174,2 - 199,3]
Stichprobe ₆	187,1	[174,6 - 199,3]
Stichprobe ₇	181,5	[169,2 - 193,8]
Stichprobe ₈	186,1	[173,8 - 198,4]
Stichprobe ₉	183,9	[171,4 - 196,3]

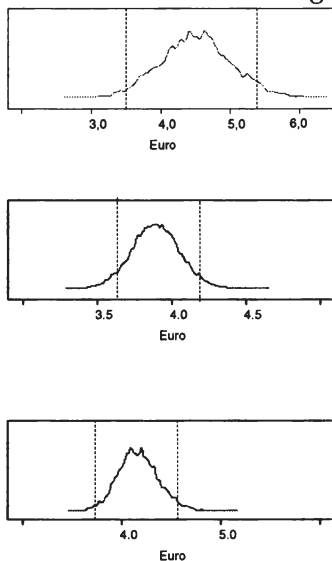
$N_0 = 100$
Eigene Berechnungen.

tisch dem Sekundärgebiet entstammt. Tabelle 6.12 zeigt das Ergebnis dieses Vorgehens für den einfachen Bayes'schen Transfer. Aus der dänischen Stichprobe werden 9 Unterstichproben gezogen, die für 9 Sekundärstichproben vom Umfang 100 stehen. Die Ergebnisse sind sehr homogen. Der Mittelwert der Zahlungsbereitschaft wird in fast allen Stichproben leicht überschätzt, darin zeigt sich der Einfluss der Cottbuser Studie, bei welcher der Erwartungswert der Zahlungsbereitschaft deutlich höher liegt, als bei der dänischen Studie. Da die komplette Cottbuser Studie vorliegt, sind die A-Priori-Verteilungen der Parameter recht schmal gewählt, bei einer tatsächlichen Sekundärstudie liegt ein kleinerer Stichprobenumfang vor, und die A-Priori-Verteilungen müssen dementsprechend weiter ausfallen. Man kann eine Überlappung aller A-Posteriori-Verteilungen feststellen und alle Mittelwerte sind in allen Konfidenzintervallen enthalten. Ferner sind alle Mittelwerte in einem Intervall von 10% des tatsächlichen Werts enthalten, so dass der Nutzentransfer nach dem Kriterium der Rückübertragbarkeit als gelungen angesehen werden kann.

Darüber hinaus ermöglicht die Anwendung von Bayes'schen Verfahren überhaupt erst Aussagen über die Konfidenzintervalle des Transferwerts. Hier liegt nicht nur eine neue - a posteriori - Verteilung der Parameter vor, die natürlich zu Vertrauensintervallen führt, sondern es gibt auch eine

leicht verständliche und kommunizierbare Darstellung der Ergebnisse. Racine et al. führen bereits 1986 [205] als einen der wesentlichen Vorteile der Bayes'schen Methoden in der angewandten pharmazeutischen Forschung an, dass die Darstellung von A-Posteriori-Verteilungen die Kommunikation zwischen Forschern und Anwendern erheblich erleichtert habe ("So far as inference summaries are concerned, we have had almost universally positive responses both to the idea of summarizing post-experimental uncertainties in the graphical form of posterior densities (...), and also to the idea of displaying a range of such pictures (...)."). Somit schließt dieses Kapitel, indem

Abbildung 6.1: A-Posteriori-Verteilung im Vergleich



es ebenfalls für die abschließende Bewertung den Vergleich der graphischen Ergebnisse der verschiedenen Ansätze für den Nutzentransfer zur Bewertung des Cottbuser Ostsees vergleicht (vgl. Abbildung 6.1). Im oberen Teil der Abbildung wird die A-Posteriori-Verteilung der Schätzung für die Stichprobe mit einer nicht-informativen A-Priori-Verteilung gezeigt. Wegen des geringen Stichprobenumfangs von $N_0 = 100$ ist die Verteilung recht breit. Im mittleren Teil ist der einfache Bayes'sche Transfer abgebildet, bei dem die A-Priori-Verteilung die Normalverteilung mit $\alpha_k \sim N(\alpha_k^p, \tau_k^p)$ ist. Die Verteilung ist deutlich schmalere, aber da die Daten der dänischen Studie gleichermaßen gewichtet sind, wie die in Cottbus erhobene Stichprobe, ist

der Mittelwert nach links verschoben, der Erwartungswert der Zahlungsbereitschaft wird in die Richtung des niedrigeren Werts der dänischen Studie beeinflusst. Die letzte Grafik in Abbildung 6.1 zeigt das Ergebnis der Formulierung der A-Priori-Verteilung als Power Prior. Der Erwartungswert der Zahlungsbereitschaft ist größer und die Verteilung zeigt eine geringfügig größere Breite als bei der mittleren Grafik, ist jedoch schmaler als in der ersten Grafik.

Zusammenfassend lässt sich festhalten, dass der Bayes'sche Nutzentransfer sowohl in der Simulation als auch bei gemessenen Daten eine erhebliche Verbesserung gegenüber der Nutzentransferfunktion darstellt. Die berechneten Werte sind deutlich näher an den tatsächlichen Werten, die Abweichungen betragen um die zehn Prozent. Ob diese Abweichung für eine Kosten-Nutzen-Analyse hinreichend ist, bleibt letztlich dem Nutzer der Transferergebnisse überlassen. Welche A-Priori-Verteilung letztlich gewählt wird, kann aufgrund der durchgeführten Experimente noch nicht abschließend beantwortet werden. Der einfache Bayes'sche Transfer neigt zum Wert der Primärstudie und somit - in diesem Fall - zur Unterschätzung des wahren Ergebnisses. In anderen Situationen kann er natürlich auch zur Überbewertung führen, wenn die Primärstudie einen höheren Wert vorgibt als die Sekundärstudie. Die Verwendung des Power Priors ist in sofern hilfreich, als sie den Einfluss der Primärstudie explizit sichtbar macht. Diese Eigenschaft könnte vom Entscheidungsträger durchaus geschätzt werden, denn sie macht die Berechnung des Transferwerts noch etwas transparenter.

Kapitel 7

Zusammenfassung und Ausblick

Es ist ein langer Weg von dem Modell von Sir Thomas Bayes (1702 - 1761), das nach seinem Tod unter dem Titel "Essay Towards Solving a Problem in the Doctrine of Chances" (1763) [28], in der Zeitschrift "Philosophical Transactions of the Royal Society of London" veröffentlicht wurde bis zum Einsatz seines Vorschlags bei der Bewertung von Umweltgütern. In dieser Arbeit wurde aufgezeigt, wie die Bayes'sche Sichtweise zu nützlichen Ergänzungen von derzeit in der ökonomischen Umweltbewertung verwendeten Verfahren führen kann. Hierzu musste zunächst in einem kurzen Überblick geklärt werden, was das Ziel der Umweltbewertung ist und welche Methode am geeignetsten ist, dieses Ziel zu verwirklichen. Der Einsatz einer Umweltbewertungsmethode in der Kosten-Nutzen-Analyse für umweltverbessernde Projekt lässt es wünschenswert erscheinen, dass die individuellen Nutzenänderungen durch die in diesem Projekt geschaffenen Gebrauchs- und Nichtgebrauchswerte gemessen werden können. Die Contingent Valuation Methode vermag dies zu leisten, denn sie leitet den Wert der Umweltverbesserung nicht indirekt aus beobachtbaren Markthandlungen ab, sondern basiert auf der direkten Frage nach der Zahlungsbereitschaft der von der Umweltveränderung betroffenen Bevölkerung. Daher beziehen sich die in dieser Arbeit gewählten Anwendungen und Beispiele auf CVM-Studien; die grundsätzlichen Überlegungen lassen sich jedoch unaufwändig auf andere Bewertungsverfahren übertragen.

Um dem Ziel einer Contingent Valuation Studie gerecht zu werden und zu einer Aussage über den Nutzen einer Umweltverbesserung zu gelangen, müssen verschiedene methodische und theoretische Anforderungen erfüllt sein. Die *methodischen* Anforderungen richten sich an die Qualität der eigentlichen Befragung (vgl. Arrow et al. (1993) [18] oder Mitchell und Carson (1989) [193]), die von einem professionellen Meinungsforschungsinstitut an einer repräsentativen Zufallsstichprobe in persönlichen Interviews durchgeführt werden sollte. Ferner sollten alle in der Befragung verwendeten Materialien in umfangreichen Pretests vor der Durchführung der Hauptbefragung auf Eindeutigkeit und Verständlichkeit getestet werden.

Die *theoretischen* Anforderungen sind erfüllt, wenn die Ergebnisse einer CVM Studie mit den Grundlagen der individuellen Wohlfahrtsmessung vereinbar sind. Dies führt zu einer Reihe von Implikationen für die Gestaltung eines CVM-Fragebogens. Neben einer sorgfältigen Beschreibung des Umweltguts ist insbesondere die Frage nach der Zahlungsbereitschaft so zu stellen, dass die Befragten auf dem hypothetischen Markt tatsächlich zu einer Kaufentscheidung kommen können, die ihren Präferenzen entspricht. Dies bedeutet, dass die Zahlungsbereitschaftsfrage die Befragten nicht zu falschen Angaben inspirieren sollte und gleichzeitig für die Befragten einfach zu beantworten ist.

Wie dies zu erreichen ist, ist bis heute stark umstritten. Die Literatur kennt unterschiedliche Formulierungen der Frage nach der Zahlungsbereitschaft. Die Formate unterscheiden sich in ihrer Anreizkompatibilität und damit auch der Vermittelbarkeit der Fragen gegenüber der interviewten Personen einerseits und der Zuverlässigkeit der aus den Antworten mittels statistischer Methoden ableitbaren Aussagen andererseits. Die natürliche Umsetzung der grundlegenden Idee scheint zunächst die Formulierung einer offenen Frage zu sein, wie sie in frühen Contingent Valuation Studien verwendet wurde (vgl. Davis (1963) [82] oder Sanders et al. (1990) [220]), jedoch auch in jüngerer Zeit noch Anwendung findet (vgl. Haab und McConnell (1998) [112] oder Botelho und Pinto (2001) [39]). Zur Berechnung der erwarteten Zahlungsbereitschaft innerhalb der Stichprobe wird entweder der Mittelwert aller geäußerten Zahlungsbereitschaften gebildet oder als Median derjenige Betrag berechnet, den die Hälfte aller Befragten zu zahlen bereit wären. Das offene Format scheint somit auf den ersten Blick die Vorteile einer einfachen statistischen Handhabung, der direkten Umsetzung des Forschungsanliegens und der leichten Interpretierbarkeit der Ergebnisse zu kombinieren. Letzteres ist insbesondere für den Auftragge-

ber einer solchen Contingent Valuation Studie ein wichtiges Argument, denn der politische Entscheidungsträger möchte die Ergebnisse einer Studie nicht nur glauben, sondern auch verstehen.

Bei näherem Hinsehen allerdings ergeben sich Zweifel an der Tauglichkeit der offenen Fragestellung (vgl. Lunander (1998) [180]). Offenbar widerspricht es der Alltagserfahrung eines Konsumenten, selbst Preise für ein Gut festzulegen, und führt zu großen Unsicherheiten bei den Befragten. Daher wird in heutigen CVM-Studien zumeist einem Vorschlag von Bishop und Heberlein (1979) [33] gefolgt und die Zahlungsbereitschaftsfrage wird im sogenannten Referendumsformat gestellt. Um die Alltagserfahrung der Haushalte mit dem Kauf von Konsumgütern auch auf den hypothetischen Markt für Umweltgüter zu übertragen, nennt man einen Preis für das Umweltgut und fragt, ob das Umweltgut zu diesem Preis gekauft wird oder nicht. Damit sind die Haushalte vor eine ähnliche Entscheidungssituation gestellt wie im Supermarkt, wo sie die Ware zu einem bestimmten Preis sehen und sich überlegen, ob sie diese Ware zu diesem Preis kaufen möchten. Diese Formulierung der Zahlungsbereitschaftsfrage wird besser verstanden und es gibt keinen Anreiz, nicht die wahre Zahlungsbereitschaft zu äußern.

In der Praxis hat sich das Referendumsformat in zahlreichen Studien als dem offenen Format überlegen erwiesen (vgl. Loomis (1990) [177], Kealy und Turner (1993) [152], Halvorsen und Saelesminde (1998) [115], Langford und Kontogianni (1998) [164], Balistreri und McClellan (2001) [23] oder Bateman et al. (2001) [27]). Vergleicht man die Ergebnisse von Befragungen im jeweiligen Format zu demselben Umweltgut, so scheint sich die Unsicherheit der Befragten bezüglich des Werts eines Umweltguts in einer häufigen Überschätzung dieses Werts bei Verwendung des offenen Formats widerzuspiegeln (vgl. Kealy und Turner (1993) [152]).

Der Zuwachs an Anreizkompatibilität und Verständlichkeit bei der Zahlungsbereitschaftsfrage geht jedoch einher mit einem Verlust an Aussagefähigkeit des Resultats. Anstelle von Aussagen über die Zahlungsbereitschaft erhält man hier bei einer zustimmenden Antwort lediglich Informationen über die Untergrenze der Zahlungsbereitschaft und bei einer ablehnenden Antwort über die Obergrenze der Zahlungsbereitschaft. Der Effizienzverlust äußert sich in Verbreiterung der Konfidenzintervalle bei der Berechnung der Willingness-to-Pay (vgl. Hanemann et al. (1991) [117]), d.h. die Aussage des Modells zum Erwartungswert der Zahlungsbereitschaft wird weniger zuverlässig. Um ein ähnliches Konfidenzintervall und somit eine ähnliche Belastbarkeit der Aussagen wie bei den Ergebnissen einer

offenen Befragung zu erreichen, muss die Stichprobe bei einer Befragung im Referendumsformat erheblich größer sein; die Befragung wird dadurch auch erheblich teurer.

Eine Möglichkeit zur Verbesserung der statistischen Effizienz (vgl. Hanemann et al. (1991) [117]) ist es, anstelle einer zwei Fragen nach der Zahlungsbereitschaft zu stellen, in der Hoffnung, auf eine obere *und* eine untere Grenze zu treffen (vgl. Alberini (1995) [12]). Bei einer ablehnenden Reaktion auf das erste Angebot aus Preis und Umweltgut wird ein zweites Angebot für dasselbe Umweltgut zu einem niedrigeren Preis unterbreitet, bei einer zustimmenden Reaktion wird nachgefragt, ob die Zustimmung für dasselbe Umweltgut und einen höheren Preis aufrechterhalten wird. Dieses sogenannte doppelte Referendumsformat erhöht die Effizienz der Auswertung, d.h. es lässt sich mit einem geringeren Stichprobenumfang eine zuverlässige Aussage treffen.

Dieses Frageformat erweist sich jedoch als schlecht vermittelbar. Die Befragten übersetzen die "Preissenkung" in eine Qualitätsminderung und bleiben bei ihrer ablehnenden Haltung oder erleben die Kombination mit dem höheren Preis schlichtweg als Versuch, sie zu schröpfen. Der überwiegende Teil aller Studien, die das sogenannte doppelte Referendumsformat verwenden, weist eine aufschlussreiche Anomalie in den Daten auf: die Anzahl der Zustimmungen in einem bestimmten Intervall von Zahlungsbereitschaften hängt davon ab, ob zunächst ein niedriger Preis vorgeschlagen wurde und dann erhöht wurde, oder ob der zuerst vorgeschlagene Preis der höhere war (vgl. Cooper (2002) [72] oder DeShazo (2002) [84]). Da in der Regel nicht davon ausgegangen werden kann, dass Personen ihre Präferenzen für ein Gut in 30 Sekunden ändern, muss bei der Antwort auf die zweite Frage ein anderer Faktor als die tatsächliche Zahlungsbereitschaft für die Antwort ausschlaggebend gewesen sein.

Einen Ausweg aus diesem Dilemma stellt der in dieser Arbeit vorgeschlagene Wechsel des statistischen Bezugsrahmens dar. Während die Vielzahl an vorgeschlagenen Modellen zur Auswertung von Contingent Valuation Daten sich auf Verfahren der klassischen Statistik stützt, wurde gelegentlich die Verwendung von Bayes'schen Ansätzen vorgeschlagen (vgl. Kriström (1990) [158]), ohne dass diese Vorschläge weiter verfolgt wurden. Erst in jüngerer Zeit existieren von einer Forschergruppe um C. J. Leon Ansätze (vgl. Leon und Vasquez-Polo (1998) [168]), den in anderen Disziplinen außerordentlich erfolgreichen Bayes'schen Bezugsrahmen auch in der Umweltbewertung einzusetzen. Diese Arbeit schloss sich den Vorschlägen an und

erkundete, welche Vorteile Bayes'sche Verfahren bei der Auswertung von Contingent Valuation Daten bergen. Die Stärke Bayes'scher Verfahren liegt darin, dass sie es in einer systematischen Weise ermöglichen, zur Schätzung eines Modells auch zusätzliche Informationen außerhalb der direkt erhobenen Daten einzubeziehen. Damit gelingt es im Bayes'schen Bezugsrahmen, die Vertrauensintervalle einer Schätzung zum Teil deutlich zu verbessern.

Daher wurde mit Blick auf den Zielkonflikt zwischen Effizienz der Schätzung und Verständlichkeit der Zahlungsbereitschaftsfrage vorgeschlagen, eine Befragung im einfachen Referendumsformat durchzuführen und diese mit Informationen aus Expertenbefragungen und Pretests zu ergänzen. Damit kann Datenmaterial, das in Umweltbewertungsstudien ohnehin zur Erfüllung der methodischen Anforderungen in Pretests erhoben wird, in die Analyse der Ergebnisse der Hauptbefragung einbezogen werden. Erste Vorschläge hierzu finden sich bei Leon et al. (2003) [167], die sich jedoch auf analytisch lösbare Bayes-Modelle konzentrieren, die zu für die befragten Experten schlecht nachvollziehbaren komplizierten funktionalen Formen führen. In dieser Arbeit wurde ein Modell entwickelt, das einfacher in der Handhabung ist. Experten werden aufgefordert, den Median und mindestens ein weiteres Perzentil der erwarteten Zahlungsbereitschaften in der Bevölkerung zu schätzen und anzugeben, wie sicher sie sich ihrer Schätzung sind. Aus diesen Informationen wurde eine Bayes'sche A-Priori-Verteilung der Modellparameter konstruiert, die in die Schätzung des Modells für die Zahlungsbereitschaft einging. Der vorgeschlagene Ansatz erweist sich durch die Verwendung von Monte Carlo Markov Ketten Simulationstechniken als flexibler in der Formulierung der funktionalen Verteilungen der Zahlungsbereitschaft. In einem Beispiel mit simulierten Daten wurde gezeigt, wie dieses Verfahren die Aussagen der praktische Umweltbewertung verbessert, denn es basiert auf dem anreizkompatiblen einfachen Referendumsformat und kommt dennoch mit einem geringeren - und somit kostengünstigeren - Stichprobenumfang aus.

Erweisen sich die Bayes'schen Verfahren somit schon bei der Auswertung einer großen Umweltbewertungsstudie als hilfreich, so sind sie noch nützlicher, wenn für eine derartige Studie keine Mittel vorhanden sind und eine *Nutzentransferstudie* durchgeführt werden soll. Als Nutzentransfer werden alle diejenigen Verfahren bezeichnet, die den für ein bestimmtes Umweltprojekt in einer umfangreichen und zuverlässigen Studie (Primärstudie) gemessenen Wert nach einem wohldefinierten Verfahren auf ein neues, in Planung befindliches Umweltprojekt übertragen (Sekundärstudie), um die

Kosten für eine vollständige neue Bewertungsstudie einsparen zu können. Als Ausgangspunkt der wissenschaftlichen Auseinandersetzung mit den Möglichkeiten des Nutzentransfers lässt sich ein Sonderband des Water Resources Research 1992 auffassen, der die zum damaligen Zeitpunkt bekannten Verfahren analysiert und bewertet. Die beiden klassischen Verfahren des Nutzentransfers unterscheiden sich vor allem durch den Umfang der Informationen über die Bevölkerung, die von dem neuen geplanten Umweltprojekt betroffen ist. Beim sogenannten *Value Transfer* wird der Erwartungswert der Zahlungsbereitschaft einfach auf die neue Studie übertragen, bzw. mithilfe von ad hoc Verfahren eher willkürlich angepasst; der Value-Transfer entspricht somit einer Punktschätzerübertragung (vgl. Smith (1992) [229] oder Boyle und Bergstrom (1992) [41]).

Bei Verwendung der *Nutzentransferfunktion* (vgl. Loomis (1992) [178]) wird der in der Primärstudie berechnete funktionale Zusammenhang zwischen dem Erwartungswert der Zahlungsbereitschaft und den in dieser Studie erhobenen erklärenden Variablen auf das neue Projekt übertragen, indem die erklärenden Variablen der Primärstudie ersetzt werden durch die Mittelwerte dieser Variablen aus der von dem neuen Projekt betroffenen Bevölkerung. Gibt die Primärstudie beispielsweise einen signifikanten statistischen Zusammenhang zwischen dem Erwartungswert der Zahlungsbereitschaft, dem Alter, dem Einkommen und dem Geschlecht eines Individuums in der Primärstudie an, so würden diese Größen ersetzt durch den Mittelwert des Alters etc. der Bevölkerung, die von dem geplanten Projekt betroffen ist. Mit Hilfe dieser Größen wird die Zahlungsbereitschaft neu geschätzt. Diese Vorgehensweise ist einleuchtend, wenn man davon ausgeht, dass eine ähnliche Bevölkerung für ein ähnliches Projekt auch eine ähnliche Zahlungsbereitschaft aufweisen muss.

Was aber ist eine ähnliche Bevölkerung? Kann es nicht trotz unterschiedlicher Beweggründe zu einer ähnlichen Bewertung eines ähnlichen Guts kommen? Kann es trotz einer ähnlichen Bevölkerungszusammensetzung zu einer unterschiedlichen Bewertung eines Guts kommen? Bereits beim Kaufverhalten bei Marktgütern beobachtet man unterschiedliche Nachfrage in verschiedenen Ländern, obwohl diese Länder vom durchschnittlichen Einkommen und ihrer Bevölkerungszusammensetzung her ähnlich sind. Gibt es also eine Art "Geschmack" für Umweltgüter? Diese Fragen zeigen, dass ein Modell, das unterschiedliche Präferenzen ausschließt, nicht vollständig richtig sein kann. Bei den herkömmlichen Verfahren des Nutzentransfers ist es jedoch nicht möglich, dass eine sozioökonomische Variable auf verschie-

dene Projekte einen unterschiedlichen Einfluss ausübt, denn die Transferfunktion stützt sich auf eine möglichst genaue Abbildung der Präferenzen in den geschätzten Parametern, die übertragen werden. Somit wäre es wünschenswert, wenn Informationen nicht nur über die erklärenden Variablen sondern auch über die Präferenzen der Bevölkerung "vor Ort" in ein Nutzentransfermodell eingeschlossen werden könnten.

Dies ist die Grundidee des Bayes'schen Nutzentransfers. Erste Ansätze zur Verwirklichung dieser Idee finden sich bei Parsons und Kealy (1994) [201], Barton (2002) [25] und bei Leon et al. (2002) [169]. In der vorliegenden Arbeit wurden diese Ansätze systematisiert und verallgemeinert. Die vorgeschlagene Verbesserung des Nutzentransfers besteht darin, dass eine kleine - kostengünstige - Studie vor Ort durchgeführt wird und diese Daten dazu genutzt werden, die Ergebnisse der Primärstudie an die Gegebenheiten der Sekundärstudie anzupassen. Die Primärstudie fungiert damit als informative A-Priori-Verteilung der Parameter im Bayes'schen Sinne. Somit kommt der Formulierung dieser A-Priori-Verteilung für den Bayes'schen Nutzentransfer eine zentrale Bedeutung zu. Je nach Qualität der Daten der Primärstudie lassen sich unterschiedliche Ansätze verfolgen. Eine intuitive und für viele Modelle für Contingent Valuation Daten zulässige Wahl der A-Priori-Verteilung ist die Normalverteilung, die in Kombination mit einem Probit- oder Logit-Modell für die Daten der Sekundärstudie zu einem Modell führt, das mit Monte Carlo Markov Ketten Technik rasch konvergiert und sich leicht handhaben lässt.

Ausschlaggebend für den politischen Entscheidungsträger als Nutzer der Nutzentransferstudie ist die Zuverlässigkeit und Qualität des Transfers, ein bislang weitgehend ungelöstes Problem bei Nutzentransferstudien. Natürlich kennt die Statistik beinahe beliebig viele Testverfahren zur Überprüfung von Hypothesen. Wie aber lautet die in einer Nutzentransferstudie zu formulierende Hypothese? Bei diesem Verfahren steht die Frage, ob der transferierte Wert gleich dem wahren Wert für das geplante Projekt ist, im Mittelpunkt des Interesses.

Daher gibt es beim Nutzentransfer - zumindest wenn er auf die herkömmliche Art durchgeführt wird - keine vor Ort erhobenen Daten, d.h. es entfällt die Basis für den klassischen Hypothesentest. Vorschläge zum Test des Nutzentransfers werden aus diesem Grund in explizit dazu angelegten Studien durchgeführt, bei denen die Ergebnisse zweier Studien auf ihre Transferierbarkeit hin überprüft werden (vgl. Brouwer und Spanninks (1999) [49] oder Holm-Müller und Muthke (2004) [128]). Wenn zwei komplette Stu-

die vorliegen, lassen sich Tests auf Gleichheit der Mittelwerte oder auf die Gleichheit der geschätzten Parameter durchführen (vgl. Thiele und Wronka (2001) [80]). In der "real world" Situation jedoch, in der zum geplanten Projekt nur sozioökonomische Bevölkerungsdaten vorliegen, lässt sich kein derartiger Vergleich durchführen. Dies bedeutet, dass die herkömmlichen Verfahren zum Nutzentransfer sich nur generell testen lassen, die Qualität einer speziellen Anwendung jedoch unüberprüfbar bleibt.

Der Bayes'sche Nutzentransfer eröffnet hier neue Möglichkeiten, denn nun liegen erhobene Daten für die Sekundärstudie vor. In dieser Arbeit wurde daher erstmals ein Testverfahren vorgeschlagen und angewendet, das sich auch für einen realen Nutzentransfer durchführen lässt. Ausgehend von der Überlegung, dass der Charme der Bayes'schen Transferverfahren darin besteht, dass es auch unterschiedliche Präferenzen in Primär- und Sekundärstudiengbiet zulässt, erwies sich ein Test auf Gleichheit der Mittelwert als sinnlos. Eine testbare Aussage hingegen besteht in der *Umkehrbarkeit* des Transfers. Die Testidee besteht somit darin, dass wenn der Transfer von der Primärstudie auf die Sekundärstudie valide sein soll, der Transfer in umgekehrter Richtung auch zu validen Ergebnissen führen muss. Die Qualität des Rücktransfers ist in der Realität zwar eingeschränkt durch den geringen Stichprobenumfang der Sekundärstudie, jedoch wirkt sich dieser im Wesentlichen auf die Konfidenzintervalle der Schätzung aus. Der Mittelwert hingegen sollte auf einem zuvor spezifizierten Niveau für verschiedene Stichproben gleich sein, d.h. er muss von der jeweiligen Stichprobe weitgehend unabhängig sein. Eine Anwendung dieses neuen Tests bestätigte den gelungenen Transfer. Als Anwendungsbeispiel diente eine im Rahmen des DFG-geförderten Sonderforschungsbereichs 565 durchgeführte Studie zur ökonomischen Bewertung eines Rekultivierungsvorhabens in einer Braunkohlefolgelandschaft.

Durch den Wechsel des statistischen Bezugsrahmens konnte dieser Arbeit somit zu drei wichtigen Fragen in der Umweltbewertungsliteratur beitragen. Die Frage nach dem zu verwendenden Format lässt sich im Licht Bayes'scher Verfahren mit einem Votum für das einfache Referendumsformat in Kombination mit Ergebnissen aus Expertenbefragungen und Pretestdaten beantworten. Bei Verwendung der vorgeschlagenen Ansätze für den Nutzentransfer konnte die für die herkömmlichen Modelle notwendigen Annahme völlig identischer Präferenzen bei der in Primär- und Sekundärstudie betrachteten Bevölkerung fallen gelassen werden und es wurde erstmals ein Test für einen in der Realität durchgeführten Nutzentrans-

fer vorgeschlagen. Für die Praxis stellt sich allerdings die Frage: Spart der Nutzentransfer nach diesem Ansatz auch Geld?

Selbstverständlich nehmen die Kosten einer Studie mit den für diese Studie benötigten Informationen zu. Ein Nutzentransfer, der sich auf Literaturrecherchen und Expertenwissen stützt, liegt in der Kostenhierarchie mit Sicherheit weit unten. Für den nächsten Schritt, die Nutzentransferfunktion, werden schon umfangreichere Daten benötigt, die bei den jeweiligen Städten und Gemeinden recherchiert und überwiegend bezahlt werden müssen. Der weitere Schritt hin zum Bayes'schen Nutzentransfer erfordert nun schon Befragungen vor Ort, wenn auch in geringem Umfang. Barton (1999) [24] gibt eine Kostenübersicht über diese verschiedenen Stufen des Nutzentransfers für ein Projekt, in dem Wasserqualität in Ecuador bewertet wird. In der Studie sind die Befragungen vor Ort mit erheblichen Reisekosten verbunden, während die spezifischen Kosten pro Interview in Entwicklungsländern eher gering sind. Dies ist zum Beispiel in Deutschland genau umgekehrt. Die Kosten je Interview betragen rund 60 € und zwar weitestgehend unabhängig von der Gesamtzahl der durchgeführten Interviews. Somit kann man von diesen Kosten 90% sparen, wenn der Stichprobenumfang auf 10% verringert wird, für die in der Arbeit als Beispiel genannte Studie wären das rund 55.000 €.

Folgt man den Vorschlägen dieser Arbeit und setzt Bayes'sche Verfahren für den Nutzentransfer ein, so ist die Formulierung der A-Priori-Verteilung von zentraler Bedeutung. Um den Einfluss der Primärstudie auf das Ergebnis des Nutzentransfers kontrollieren zu können, wurde in dieser Arbeit erstmals die Verwendung des sogenannten Power Priors (vgl. Ibrahim et al. (2003) [140]) für Nutzentransferstudien vorgeschlagen. Diese Bestimmung der A-Priori-Verteilung erlaubt es, den Einfluss der Primärstudien Daten explizit in das Modell aufzunehmen und zu beobachten.

Dieser Vorschlag ersetzt natürlich nicht die fundierte Einschätzung eines Transfers durch den Wissenschaftler, durch die alle Nutzentransferverfahren unweigerlich gekennzeichnet sind. Letztlich hat man mit der Entscheidung für einen Nutzentransfer den ohnehin leicht schwankenden Boden der theoriebasierten Umweltbewertungsverfahren ein Stück weit verlassen. In Zeiten knapper öffentlicher Kassen aber nicht abnehmender Umweltveränderungen ist jedoch zu fragen, ob dem Verzicht auf jegliches Bewertungsverfahren, wodurch entweder die Umwelt einen Wert von Null erhält, oder der Subjektivität Tür und Tor geöffnet werden, tatsächlich der Vorzug zu geben ist. Die Alternative besteht vielmehr darin, die bestehenden

Verfahren zum Nutzentransfer so transparent und gültig wie möglich zu gestalten. Wie gut die hier vorgeschlagenen Verfahren sich bewähren, kann nur in weiteren Studien in der Praxis geklärt werden (für eine Anwendung der entwickelten Verfahren auf Choice Daten zur Bewertung von Programmen zur Vermeidung von Bodenerosion vgl. Hanley, Colombo und Lehr (in Vorbereitung)).

Auch ist die Palette der Bayes'schen Verfahren mit dieser Arbeit nicht ausgereizt. In der Bayes'schen wie in der klassischen Statistik haben sich in der jüngeren Literatur Ansätze zu semiparametrischen Verfahren etabliert, deren Vorteil es ist, dass sie mit weniger restriktiven Verteilungsannahmen auskommen, als die parametrischen Verfahren. Bislang ist die Anwendung der semiparametrischen Verfahren in praktischen Studien zur Umweltbewertung dadurch begrenzt, dass der Verzicht auf eine Verteilungsannahme bei vielen Verfahren durch eine Erhöhung der Stichprobengröße kompensiert werden muss, was der angestrebten Kostensenkung entgegensteht. Allerdings ist die Abkehr von der Abhängigkeit von festen funktionalen Formen in den Verteilungsannahmen auch in der Bayes'schen Statistik zu beobachten. Obwohl sich eine nichtparametrische Bayes-Statistik zunächst wie ein absoluter Widerspruch anhört, findet sich hohe wissenschaftliche Aktivität auf diesem Gebiet (für eine gute Übersicht vgl. Müller und Quintana (2004) [194]. Ein möglicher Ansatz, der auch schon für Survival-Daten erfolgreich genutzt wurde (vgl. Ibrahim et al. (2001) [139], besteht in der Verwendung des sogenannten Dirichlet Prozesses (vgl. Ferguson (1973) [95]). Vielleicht eröffnet der Einsatz nicht- oder semiparametrischer Bayes'scher Verfahren eine neue Möglichkeit, die Abhängigkeit von festen funktionalen Formen ein wenig zu lockern.

Literaturverzeichnis

- [1] *The Environmental Valuation Reference Inventory*[®]. www. evri. ca. 1992
- [2] *A Searchable Environmental Valuation Database*. Datenbank zur Verfügung gestellt von der australischen Umweltbehörde, <http://www.environment.nsw.gov.au/envalue>. 1995
- [3] *CVM@Network, Zusammenstellung von CVM Studien im deutschsprachigen Raum*. <http://www.econ.uni-klu.ac.at/cvm/>. Zugriff am 24. Februar 2005
- [4] AHLHEIM, M. ; BUCHHOLZ, W.: WTP or WTA - Is that the Question? In: *Zeitschrift für Umweltpolitik & Umweltrecht* (2000), Nr. 23, S. 253–271
- [5] AHLHEIM, M. ; FRÖR, O.: Valuing the Non-Market Production of Agriculture. In: *Agrarwirtschaft* (2003), Nr. 52, S. 356–369
- [6] AHLHEIM, M. ; FRÖR, O.: Environmental indicators from an economic perspective – significance and political relevance. In: HUTTER, C.-P. (Hrsg.) ; BÖCKER, R. (Hrsg.): *Umweltindikatoren - Mythos oder Wirklichkeit? Was wissen wir wirklich über unsere Umwelt?* Wissenschaftliche Verlagsgesellschaft mbH Stuttgart, 2005 (Beiträge der Akademie für Natur- und Umweltschutz Baden-Württemberg Band 35)
- [7] AHLHEIM, M. ; FRÖR, O. ; LEHR, U. ; WAGENHALS, G. ; WOLF, U.: Contingent Valuation of Mining Land Reclamation. In: *IAW-Report* (2004), Nr. 2, S. 33–61
- [8] AHLHEIM, M. ; LEHR, U.: Household Equivalence Scales and the Assessment of Environmental Benefits. In: *Diskussionsschriftenreihe der Fakultät für Umweltwissenschaften und Verfahrenstechnik der BTU Cottbus* (2001), Nr. 1/2001
- [9] AHLHEIM, M. ; LEHR, U.: Nutzentransfer: Das Sparmodell der Umweltbewertung. In: *Perspektiven der Wirtschaftspolitik* (2002), Nr. 3/1, S. 85–104
- [10] AHLHEIM, M. ; ROSE, M.: *Messung individueller Wohlfahrt*. Heidelberg : Springer, 1989
- [11] AKAIKE, H.: *Information Theory and an Extension of the Maximum Likelihood Principle*. Budapest, Akademia Kiado, 1973
- [12] ALBERINI, A.: Efficiency vs. Bias of Willingness-to-Pay Estimates: Bivariate and Interval-Data Models. In: *Journal of Environmental Economics and Management* (1995), Nr. 29, S. 169–180
- [13] ALBERINI, A. ; KANNINEN, B. ; CARSON, R. T.: Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data / University of California, San Diego. 1997. – Forschungsbericht

- [14] AN, M. Y.: A Semiparametric Distribution for Willingness to Pay and Statistical Inference with Dichotomous Choice Contingent Valuation Data. In: *American Journal of Agricultural Economics* , Nr. 82
- [15] ANDREONI, J.: Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow-Giving. In: *The Economic Journal* (1990), Nr. 100, S. 464–477
- [16] ANDREONI, J.: Co-operation in Public Goods Experiments: Kindness or Confusion? In: *American Economic Review* (1995), Nr. 85, S. 891–904
- [17] ARDILA, S. ; QUIROGA, R. ; VAUGHAN, W. J.: A Review of the Use of Contingent Valuation Methods in Project Analysis at the Inter-American Development Bank / Inter-American Development Bank. 1998. – Forschungsbericht
- [18] ARROW, K. ; SOLOW, R. ; PORTNEY, P. ; LEAMER, E. ; RADNER, R. ; SCHUMAN, H.: Report on the NOAA Panel on contingent valuation: Natural resource damage assessment under the oil pollution act of 1990. In: *Federal Register 58/10* (1993), S. 4602–4614
- [19] ARROW, K. J.: A Difficulty in the Concept of Social Welfare. In: *Journal of Political Economy* (1950), Nr. 58, S. 328–346
- [20] ARROW, K. J.: *Social Choice and Individual Values*. New York : Wiley, 1951
- [21] ATKINSON, S. E. ; CROCKER, T. D. ; SHOGREN, J. F.: Bayesian Exchangeability, Benefit Transfer and Research Efficiency. In: *Water Resource Research* (1992), Nr. 28(3), S. 715–722
- [22] AYER, M. ; BRUNK, H. D. ; EWING, G. M. ; SILVERMAN, E.: An Empirical Distribution Function for Sampling with Incomplete Information. In: *Annals of Mathematical Statistics* (1955), Nr. 26, S. 64– 647
- [23] BALISTRERI, E. ; MCCLELLAN, G.: Can Hypothetical Questions Reveal True Values? Laboratory Comparison of Dichotomous Choice and Open-Ended Contingent Values with Auction Values. In: *Environmental and Resource Economics* (2001), Nr. 18, S. 275–292
- [24] BARTON, D. N.: *The Quick, the Cheap and the Dirty: Benefit Transfer Approaches to the Non-market Valuation of Coastal Water Quality in Costa Rica*, Department of Economics and Social Sciences, Agricultural University of Norway, Diss., 1999
- [25] BARTON, D. N.: The Transferability of Benefit Transfer: Contingent Valuation of Water Quality Improvements in Costa Rica. In: *Ecological Economics* (2002), Nr. 42, S. 147–164
- [26] BATEMAN, I. J. ; COOPER, P. ; GEORGIU, S. ; NAVRUD, S. ; POE, G. L. ; READY, R. C. ; RYAN, M. ; VOSSLER, C. A.: Scope Sensitivity Tests for Preference Robustness: An Empirical Examination of Economic Expectations Regarding the Economic Valuation of Policies for Reducing the Acidity in Remote Mountain Lakes. In: *CSEERGE Working Paper* (2004)
- [27] BATEMAN, I.J. ; LANGFORD, I.H. ; JONES, A.P. ; KERR, G.N.: Bound and path effects in double and triple bounded dichotomous choice contingent valuation. In: *Resource and Energy Economics 23(3)* (2001), S. 191–213
- [28] BAYES, T.: Essay Towards Solving a Problem in the Doctrine of Chances. In: *Philosophical Transactions of the Royal Society of London* (1763). – <http://www.stat.ucla.edu/history/essay.pdf>

- [29] BENNET, J. ; MORRISON, M. ; BLAMEY, R.: Testing the Validity of Responses to Contingent Valuation Questioning. In: *Australian Journal of Agricultural and Resource Economics* (1998), Nr. 42(2), S. 131–148
- [30] BERGLAND, O. ; MAGNUSSEN, K. ; NAVRUD, S.: Benefit Transfer: Testing for Accuracy and Reliability / Discussion paper 9503, Agricultural University of Norway. 1998. – Forschungsbericht
- [31] BERRENS, R. P. ; BOHARA, A. K. ; JENKINS-SMITH, H. C. ; SILVA, C. L. ; WEIMER, D. L.: Information and Effort in Contingent Valuation Surveys: Application to Global Climate Change Using National Internet Samples. In: *Journal of Environmental Economics and Management* (2004), Nr. 47(2), S. 331–363
- [32] BERRENS, R. P. ; BROOKSHIRE, D. ; GANDERTON, P. ; MCKEE, M.: Exploring Nonmarket Values for Social Impacts of Environmental Policy Change. In: *Resource and Energy Economics* (1998), Nr. 20, S. 117–137
- [33] BISHOP, T. A.: Measuring Values of Extra-Market Goods: Are Indirect Measures Biased? In: *American Journal of Agricultural Economics* (1979), Nr. 61, S. 926–930
- [34] BJØRNER, T. B. ; RUSSELL, C. S. ; DUBGAARD, A. ; DAMGAARD, C. ; ANDERSON, L. M.: Public and Private Preference for Environmental Quality in Denmark. In: *SOM publikation nr. 39, AKF Forlaget* (2000)
- [35] BLAMEY, R. ; COMMON, M.: Respondents to Contingent Valuation Surveys: Consumers or Citizens? In: *Australian Journal of Agricultural Economics* (1995), Nr. 39(3), S. 263–288
- [36] BOCKSTAEL, N. E. ; STRAND, I. ; HANEMANN, M. W.: Time and the Recreational Demand Model. In: *American Journal of Agricultural Economics* (1987), Nr. 69(2), S. 293–302
- [37] BOHARA, A. K. ; MCKEE, M. ; BERRENS, R. P. ; JENKINS-SMITH, H. ; SILVA, C. L. ; BROOKSHIRE, D.: Effects of Total Cost and Group-Size Information on Willingness to Pay Responses: Open Ended vs. Dichotomous Choice. In: *Journal of Environmental Economics and Management* (1998), Nr. 35, S. 142–163
- [38] BOMAN, M. ; BOSTEDT, G. ; KRISTRÖM, B.: Obtaining Welfare Bounds in Discrete-Response Valuation Studies: A Nonparametric Approach. In: *Land Economics* (1999), Nr. 75(2), S. 284–294
- [39] BOTELHO, A. ; PINTO, L.: Hypothetical, Real and Predicted Real Willingness to Pay in Open-Ended Surveys: Experimental Results. In: *Universidade do Minho, Portugal, Working Paper Series, No 12* (2001)
- [40] BOYCE, R. . ; BROWN, T. ; MCCLELLAND, G. ; PETERSON, G. ; SCHULZE, W.: An Experimental Examination of Intrinsic Values as a Source of the WTP-WTA Disparity. In: *The American Economic Review* (1992), Nr. 82 (5), S. 1366–1373
- [41] BOYLE, K. J. ; BERGSTROM, J. C.: Benefit Transfer Studies: Myths, Pragmatism, and Idealism. In: *Water Resources Research* (1992), Nr. 28, S. 657–663
- [42] BOYLE, K. J. ; DESVOUGES, W. H.: An Investigation of Part - Whole Biases in Contingent-Valuation Studies. In: *Journal of Environmental Economics and Management* (1994), Nr. 27, S. 64 – 83

- [43] BOYLE, K. J. ; JOHNSON, F. R. ; MCCOLLUM, D. W. ; DESVOUSGES, W. H. ; DUNFORD, R. W. ; HUDSON, Sara P.: Valuing Public Goods: Discrete Versus Continuous Contingent Valuation Responses. In: *Land Economics* (1996), S. 72(3)
- [44] BRONSTEIN, I. N. ; SEMENDJAJEW, K. A.: *Taschenbuch der Mathematik*. Thun : Harry Deutsch, 1981
- [45] BROOKS, S. P. ; GELMAN, A.: General Methods for Monitoring Convergence of Iterative Simulations. In: *Journal of Computational and Graphical Statistics* (1998), Nr. 7, S. 434-455
- [46] BROOKSHIRE, D. S. ; NEILL, H. R.: Benefit Transfers: Conceptual and Empirical Issues. In: *Water Resources Research* (1992), Nr. 28, S. 651-655
- [47] BROOKSHIRE, D.S. ; RANDALL, A. ; STOLL, J.R.: Valuing Increments and Decrements in Natural Resource Service Flows. In: *American Journal of Agricultural Economics* (1980), Nr. 62, S. 478-88
- [48] BROUWER, R.: Environmental Value Transfer: State of the Art and Future Prospects. In: *Ecological Economics* (2000), Nr. 32, S. 137-152
- [49] BROUWER, R. ; SPANNINKS, F. A.: The Validity of Environmental Benefits Transfer: Further Empirical Testing. In: *Environmental and Resource Economics* (1999), Nr. 14, S. 95-117
- [50] BUNDESAMT, Statistisches. *Einkommens- und Verbrauchsstichprobe*. <http://www.destatis.de>. zugegriffen am 8. Oktober 2004
- [51] BURNHAM, K. P. ; ANDERSON, D. R.: *Model Selection and Inference*. Heidelberg, New York, Berlin, Paris : Springer, 1998
- [52] CALIA, P. ; STRAZZERA, E.: Bias and efficiency of single vs. double bound models for contingent valuation studies: a Monte Carlo Analysis. In: *Working Paper CRENoS 199801* (1998)
- [53] CAMERON, T. A.: Interval Estimates of Non-Market Resource Values from Referendum Contingent Valuation Surveys. In: *Land Economics* (1991), Nr. 67(4), S. 413- 421
- [54] CAMERON, T. A. ; HUPPERT, D.: Referendum Contingent Valuation Estimates: Sensitivity to the Assignment of Offered Values. In: *Journal of the American Statistical Association* (1991), Nr. 86, S. 910 - 918
- [55] CAMERON, T. A. ; QUIGGIN, J.: Estimation using Contingent Valuation Data from a Dichotomous Choice with Follow-Up Questionnaire. In: *Journal of Environmental Economics and Management* (1994), Nr. 27, S. 218-234
- [56] CARLIN, J. B.: Meta-Analysis for 2x tables: a Bayesian approach. In: *Statistics in Medicine* (1992), Nr. 11, S. 141-158
- [57] CARSON, R. T.: Constructed Markets. In: BRADEN, J. B. (Hrsg.) ; KOLSTAD, C. D. (Hrsg.): *Measuring the Demand for Environmental Quality*. Amsterdam, 1991, S. 120-162
- [58] CARSON, R. T.: Contingent Valuation: Theoretical Advances and Empirical Tests since the NOAA Panel. In: *American Journal of Agricultural Economics* (1997), Nr. 79(5), S. 1501-1507

- [59] CARSON, R. T.: Contingent Valuation: A User's Guide / University of California, San Diego, Department of Economics; Discussion Paper. 1999 (99-26). – Forschungsbericht. – 1-19 S
- [60] CARSON, R. T. ; FLORES, N. E. ; MEADE, N. E.: Contingent Valuation: Controversies and Evidence. In: *Environmental and Resource Economics* (2001), Nr. 19, S. 173-210
- [61] CARSON, R. T. ; GROVES, T. ; MACHINA, M. J. *Incentive and Informational Properties of Preference Questions*. Plenary Address, European Association of Environmental and Resource Economists, Oslo, Norway. 1999
- [62] CARSON, R. T. ; MITCHELL, R. C.: Sequencing and Nesting in Contingent Valuation Surveys. In: *Journal of Environmental Economics and Management* (1995), Nr. 28, S. 1155- 1173
- [63] CASELLA, G. ; GEORGE, E. I.: Explaining the Gibbs Sampler. In: *Journal of the American Statistical Association* (1992), Nr. 46(3), S. 167-174
- [64] CHAY, K. Y. ; POWELL, J. L.: Semiparametric Censored Regression Models. In: *Journal of Economic Perspectives* (2001), Nr. 15(4), S. 29-42
- [65] CHEN, M.-H. ; IBRAHIM, J. G. ; YIANNOUTSOS, C.: Prior Elicitation and Bayesian Computation for Logistic Regression Models with Applications to Variable Selection. In: *Journal of the Royal Statistical Society, Series B* 61 (1999), S. 223-242
- [66] CIRIACY-WANTRUP, S. V.: Capital Returns from Soil-Conservation Practices. In: *Journal of Farm Economics* (1947), Nr. 29, S. 1188-1190
- [67] CONGDON, P.: *Bayesian Statistical Modelling*. New York : Wiley Series in Probability and Statistics, 2001
- [68] CONSTANZA, R. ; D'ARGE, R. ; GROOT, R. de ; FARBER, S. ; GRASSO, M. ; HANNON, B. ; LIMBURG, K. ; NAEEM, S.: The value of the world's ecosystem services and natural capital. In: *Nature* (1987)
- [69] COOPER, J. ; LOOMIS, J.: Sensitivity of Willingness-to-Pay Estimates to Bid Design in Dichotomous Choice Contingent Valuation Models. In: *Land Economics* (1992), Nr. 68(2), S. 217- 224
- [70] COOPER, J. C.: Optimal Bid Selection for Dichotomous Choice Contingent Valuation Surveys. In: *Journal of Environmental Economics and Management* (1993), Nr. 24(1), S. 2 - 40
- [71] COOPER, J. C.: A Comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys. In: *Land Economics* (1994), Nr. 70(1), S. 111-122
- [72] COOPER, J. C. ; HANEMANN, M. ; SIGNORELLO, G.: One-and-One-Half-Bound Dichotomous Choice Contingent Valuation. In: *The Reviews of Economics and Statistics* (2002), Nr. 84(4), S. 742-750
- [73] CORDELL, H. K. ; BERGSTROM, J. C.: Comparison of Recreation Use Value Among Alternative Reservoir Water Level Management Scenarios. In: *Water Resources Research* (1993), Nr. 29(2), S. 247-258

- [74] COX, D. R. ; SNELL, E. J.: *Analysis of Binary Data*. 2nd. Washington : Chapman and Hall, 1999
- [75] CREEL, M.: A Monte Carlo Comparison of Two Nonparametric Methods for Referendum Contingent Valuation Data. In: *Working Paper Series of the University of Barcelona* (1997)
- [76] CREEL, M.: A Note on Consistent Estimation of Mean WTP Using a Misspecified Logit Contingent Valuation Model. In: *Journal of Environmental Economics and Management* (1998), Nr. 35(3), S. 277–284
- [77] CREEL, M. ; LOOMIS, J.: Semi-Nonparametric Distribution-Free Dichotomous Choice Contingent Valuation. In: *Journal of Environmental Economics and Management* (1997), Nr. 32, S. 347– 358
- [78] CROOKER, J. ; HERRIGES, J. A.: Parametric and Semi-nonparametric Estimation of Willingness-to-Pay in a Contingent Valuation Framework. In: *Environmental and Resource Economics* (2004), Nr. 27(4), S. 451–480
- [79] CROOKER, J. ; KLING, C.: Nonparametric Bounds on Welfare Measures: A New Tool for Nonmarket Valuation. In: *Journal of Environmental Economics and Management* (2000), Nr. 39(2), S. 145–161
- [80] D., Thiele H. ; WRONKA, T. C.: Umweltgüter und ihre Bewertung: Möglichkeiten und Grenzen des Benefit Transfers. In: *FE-Working Papers , University of Kiel* (2001), Nr. EWP 0106
- [81] DAUBERT, J. T. ; YOUNG, R.A.: Recreational Demands for Maintaining Instream Flows: A Contingent Valuation Approach. In: *American Journal of Agricultural Economics* (1981), Nr. 63, S. 667–676
- [82] DAVIS, R.: *The Value of Outdoor Recreation: An Economic Study of the Maine Woods*. Harvard, 1963
- [83] DERKSEN, J.B.D.: *Berekeningen over het nationale inkomen van Nederland voor de periode 1900- 1920 [Computations of the National Income of the Netherlands for the period 1900-1920]. Speciale Onderzoekingen van de Nederlandse Conjunvtuur*. 1941
- [84] DESHAZO, J. R.: Designing Transactions without Framing Effects in Iterative Question Formats. In: *Journal of Environmental Economics and Management* (2002), Nr. 43(3), S. 360–385
- [85] DESVOUSGES, W. H. ; NAUGHTON, M. C. ; PARSONS, G. R.: Benefit Transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies. In: *Water Resource Research* (1992), Nr. 28(3), S. 675–683
- [86] DIAMOND, P. A. ; HAUSMAN, J. A.: On Contingent Valuation Measurement of Nonuse Values. In: HAUSMAN, J. A. (Hrsg.): *Contingent Valuation: A critical assessment*. Amsterdam : North Holland, 1993, S. 3–38
- [87] DIAMOND, P. A. ; HAUSMAN, J. A.: Contingent Valuation: Is Some Number Better than No Number? In: *Journal of Economic Perspectives* (1994), Nr. 8, S. 45–64
- [88] DIAMOND, P. A. ; HAUSMAN, J. A. ; LEONHARD, G. ; DENNING, M. A.: Does Contingent Valuation Measure Preferences? Experimental Evidence. In: HAUSMAN, J. A. (Hrsg.): *Contingent Valuation: A critical assessment*. Amsterdam : North Holland, 1993, S. 3–38

- [89] DIEREN, W. van: *Mit der Natur rechnen*. Basel : Birkhäuser Verlag, 1995
- [90] DOWNING, M. ; T. OZUNA, Jr.: Testing the Reliability of the Benefit Function Transfer Approach. In: *Journal of Environmental Economics and Management* (1996), Nr. 30, S. 316–322
- [91] EFTEC: Conjunctive Use of the Weir Wood Reservoir and River Rother, Sussex: A Small-Scale Valuation Study. In: *Report to Southern Water, EFTEC Ltd* (2000)
- [92] ELNAGHEEB, A. H. ; JORDAN, J. L.: Comparing Three Approaches That Generate Bids for the Referendum Contingent Valuation Method. In: *Journal of Environmental Economics and Management* (1995), Nr. 29, S. 92 – 104
- [93] ENVIRONMENTAL-PROTECTION-AGENCY. *What are "Use" and "Non-Use" Values?* [http : //www.epa.gov/superfund/programs/nrd/faqs.htm](http://www.epa.gov/superfund/programs/nrd/faqs.htm). Zugriff am 3. Februar 2005
- [94] FEHR, E. ; GÄCHTER, S.: Cooperation and Punishment in Public Goods Experiments. In: *American Economic Review* (2000), Nr. 90, S. 980–994
- [95] FERGUSON, T. S.: A Bayesian Analysis of Some Nonparametric Problems. In: *Annals of Statistics* (1973), Nr. 1, S. 209–230
- [96] FERNANDEZ, C. ; LEON, C. J. ; STEEL, M. F. J. ; VAZQUEZ-POLO, F. J.: Bayesian Analysis of Interval Data Contingent Valuation Models / University of Kent. 2001 (UKC/IMS/01/29). – Forschungsbericht
- [97] FIELD, B. ; FIELD, M.: *Environmental Economics: an Introduction. 3rd edition*. New York : McGraw Hill, 1997
- [98] GARROD, G. ; WILLIS, K.: *Economic Valuation of the Environment*. Cheltenham : Edward Elgar, 1999
- [99] GELFAND, A. E. ; HILLS, S. E. ; RACINE-POON, A. ; SMITH, A. F. M.: Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. In: *Journal of the American Statistical Association* (1990), Nr. 85, S. 972–985
- [100] GELFAND, A. E. ; SMITH, A. F. M.: Sampling-based Approaches to Calculating Marginal Densities. In: *Journal of the American Statistical Association* (1990), Nr. 85, S. 398–409
- [101] GELMAN, A. ; CARLIN, J. B. ; STERN, H. S. ; RUBIN, D. B.: *Bayesian Data Analysis*. Washington : Chapman and Hall, 2000 (Texts in Statistical Science)
- [102] GELMAN, A. ; RUBIN, D. B.: Inference from Iterative Simulation Using Multiple Sequences. In: *Statistical Science* (1992), S. 457–511
- [103] GEMAN, S. ; GEMAN, D.: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. In: *IEEE Trans. Pattern Anal. Machine Intel.* 6 (1984), S. 721–741
- [104] GEWEKE, J. ; KEANE, M.: Mixture of Normals Probit Model / Federal Reserve Bank of Minneapolis. 1997. – Forschungsbericht
- [105] GILKS, W. ; RICHARDSON, S. ; SPIEGELHALTER, D.: *Markov Chain Monte Carlo in Practice*. Washington : Chapman and Hall, 1996

- [106] GIRAUD, K. ; LOOMIS, J. B. ; JOHNSON, R. L.: Internal and External Scope in Willingness-to-Pay Estimates for Threatened and Endangered Wildlife. In: *Journal of Environmental Management* (1999), Nr. 56(3), S. 221–229
- [107] GREENE, W. H.: *Econometric Analysis*. London : Prentice Hall, 2003
- [108] GRILICHES, Z.: *Price Indexes and Quality Change*. Cambridge University Press, 1971
- [109] HAAB, T. C.: Analyzing Multiple Question Contingent Valuation Survey: A Reconsideration of the Bivariate Probit. In: *ECU971 East Carolina University, Working Papers* (1997), S. 1–20
- [110] HAAB, T. C.: Estimation Using Contingent Valuation Data from a "Dichotomous Choice with Follow-Up" Questionnaire: A Comment. In: *Journal of Environmental Economics and Management* (1998), Nr. 35, S. 190 – 194
- [111] HAAB, T. C.: Nonparticipation or Misspecification? The Impacts of Nonparticipation on Dichotomous Choice Contingent Valuation. In: *Environmental and Resource Economics* (1999), Nr. 14(4), S. 443–461
- [112] HAAB, T. C. ; MCCONNELL, K. E.: Referendum Models and Economic Values: Theoretical, Intuitive, and Practical Bounds on Willingness to Pay. In: *Land Economics* 74,2 (1998), S. 216–229
- [113] HABB, T. C. ; MCCONNELL, K. E.: *Valuing Environmental and Natural Resources*. Cheltenham : Edward Elgar, 2002
- [114] HALSTEAD, J. M. ; LINDSAY, B. E. ; BROWN, C. M.: Use of the Tobit Model in Contingent Valuation: Experimental Evidence from the Pemigewasset Wilderness Area. In: *Journal of Environmental Management* (1991), S. 79–89
- [115] HALVORSEN, B. ; SAELENSMINDE, K.: Difference between Willingness-to-Pay Estimates from Open-Ended and Discrete-Choice Contingent Valuation Methods: The Effects of Heteroscedasticity. In: *Land Economics* (1998), Nr. 74(2), S. 267 – 282
- [116] HANEMANN, M. ; KANNINEN, B.: Statistical Analysis of Discrete-Response Contingent Valuation Data. In: BATEMAN, I. J. (Hrsg.) ; WILLIS, K. G. (Hrsg.): *Valuing Environmental Preferences*. Oxford University Press, 2001
- [117] HANEMANN, W. ; LOOMIS, J. ; KANNINEN, B.: Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. In: *American Journal of Agricultural Economics* (1991), Nr. 73, S. 1255–1263
- [118] HANEMANN, W. M.: Theory versus Data in the Contingent Valuation Debate. In: *Contingent Valuation of Environmental Resources*. Cheltenham : Edward Elgar, S. 38–61
- [119] HANEMANN, W. M.: Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. In: *American Journal of Agricultural Economics* (1984), Nr. 66, S. 332–341
- [120] HANEMANN, W. M.: Willingness to Pay and Willingness to Accept: How Much Can they Differ? In: *American Economic Review* (1991), Nr. 81(3), S. 635–647
- [121] HANLEY, N. ; BELL, D. ; ALVAREZ-FARIZO, B.: Valuing the Benefits of Coastal Water Quality Improvements Using Contingent and Real Behaviour. In: *Environmental and Resource Economics* (2003), Nr. 24(3), S. 273–285

- [122] HANLEY, N. ; SHOGREN, J. F. ; WHITE, B.: *Environmental economics in theory and practice*. Macmillan, Basingstoke, 1997
- [123] HERRIGES, J. A. ; SHOGREN, J. F.: Starting Point Bias in Dichotomous Choice Valuation With Follow-Up Questioning. In: *Journal of Environmental Economics and Management* (1996), Nr. 30, S. 112–131
- [124] HICKS, J. R.: Foundations of Welfare Economics. In: *Economic Journal* (1939), Nr. 49, S. 696–712
- [125] HICKS, J. R.: The Four Consumer's Surpluses. In: *Review of Economic Studies* (1943), Nr. 11, S. 31–41
- [126] HOEVENAGEL, R.: An Assessment of the Contingent Valuation Method. In: PETHIG, R. (Hrsg.): *Valuing the Environment: Methodological and Measurement Issues*. Dordrecht : Kluwer, 1994, S. 195–227
- [127] HOEVENAGEL, R. ; LINDEN, J. W. d.: Effects of Different Descriptions of the Ecological Good on Willingness to Pay Values. In: *Ecological Economics* (1993), Nr. 7, S. 223–238
- [128] HOLM-MÜLLER, K. ; MUTHKE, T.: National and International Benefit Transfer Testing With A Rigorous Test Procedure. In: *Environmental and Resource Economics* (2005)
- [129] HOROWITZ, J. *Binary Response Models*. Department of Economics, Aarhus School of Business. 2004
- [130] HOROWITZ, J. L. ; NEUMANN, G. R.: Semiparametric Estimation of Employment Duration Models. In: *Econometric Reviews* (1987), Nr. 6(1), S. 25–40
- [131] HOROWITZ, J. L. ; SAVIN, N. E.: Binary Response Models: Logits, Probits and Semiparametrics. In: *Journal of Economic Perspectives* (2001), Nr. 15/4, S. 43–56
- [132] HUETING, R.: *New Scarcity and Economic Growth*. Amsterdam : North Holland, 1974
- [133] HUETING, R.: Correcting National Income for Environmental Losses: a Practical Solution for a Theoretical Dilemma. In: CONSTANZA, R. (Hrsg.): *Ecological Economics, the Science and Management of Sustainability*. Columbia University Press, 1991, S. 94–213
- [134] HUETING, R.: The Use of the Discount Rate in Cost-Benefit Analysis for Different Uses of Humid Tropical Forest Area. In: *Ecological Economics* (1991), Nr. 43(1), S. 43–57
- [135] HULL, C.H. ; AUGUSTUS M. KELLEY (NEW YORK, 1963): *The Economic Writings of Sir William Petty. Together with the Observations upon the Bills of Mortality, more probably by Captain John Graunt*. 1899
- [136] HURWICZ, L. ; UZAWA, H.: On the Integrability of Demand Functions. In: CHIPMAN, J. S. (Hrsg.) ; HURWICZ, L. (Hrsg.) ; RICHTER, M. K. (Hrsg.) ; SONNENSCHNEIN, H. F. (Hrsg.): *Preferences, Utility and Demand*. New York : Harcourt Brace Jovanovich Inc., 1971, S. 114–148
- [137] IBRAHIM, J. G. ; CHEN, M. H.: Power Prior Distributions for Regression Models. In: *Statistical Science* (2000), Nr. 15, S. 46–60

- [138] IBRAHIM, J. G. ; CHEN, M.-H. ; MACEACHERN, S. N.: Bayesian Variable Selection of Proportional Hazards Models. In: *The Canadian Journal of Statistics* (1999), Nr. 27, S. 701–717
- [139] IBRAHIM, J. G. ; CHEN, M. H. ; SINHA, D.: *Bayesian Survival Analysis*. New York Heidelberg Berlin : Springer, 2002 (Springer Series in Statistics)
- [140] IBRAHIM, J. G. ; CHEN, M. H. ; SINHA, D.: On Optimality Properties of the Power Prior. In: *Journal of the American Statistical Association* (2003), Nr. 98, S. 204–213
- [141] ICHIMURA, H. ; LINTON, O.: Asymptotic Expansions for Some Semiparametric Program Evaluation Estimators. In: *CeMMAP working papers, number CWP04/01* (2001), S. 1–19
- [142] JAYNES, E. T.: Prior Probabilities. In: *IEEE Transactions On Systems Science and Cybernetics* (1968), S. 227–241
- [143] JORDAN, J. L. ; ELNAGHEEB, A. H.: Consequences of Using Different Question Formats in Contingent Valuation: A Monte Carlo Study. In: *Land Economics* (1994), Nr. 70(1), S. 97 – 110
- [144] JORGENSEN, G. J. S.: Protest Responses in Contingent Valuation. In: *Environmental and Resource Economics* (1999), Nr. 14, S. 131 – 150
- [145] JUSTER, T.: A Framework for the Measurement of Economic and Social Performance. In: MOSS, M. (Hrsg.): *The Measurement of Economic and Social Performance*. New York : Columbia University Press, 1973, S. 25–84
- [146] KAHNEMAN, D. ; KNETSCH, J. L.: Valuing Public Goods: The Purchase of Moral Satisfaction. In: *Journal of Environmental Economics and Management* (1992), Nr. 22, S. 57–70
- [147] KAHNEMAN, D. ; RITOV, I.: Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method. In: *Journal of Risk and Uncertainty* (1994), Nr. 1, S. 5–38
- [148] KALBFLEISCH, J. D. ; PRENTICE, R. L.: *The Statistical Analysis of Failure Time Data*. New York : Wiley, 1980
- [149] KALDOR, N.: Welfare Propositions and Interpersonal Comparisons of Utility. In: *Economic Journal* (1939), Nr. 49, S. 549–552
- [150] KANNINEN, B.: Optimal Experimental Design for Double-Bounded Choice Contingent Valuation. In: *Land Economics* (1993), Nr. 69(2), S. 138–146
- [151] KANNINEN, B. J. ; KRISTRÖM, B.: Sensitivity of Willingness-to-Pay Estimates to Bid Design in Dichotomous Choice Valuation Models: Comment. In: *Land Economics* (1993), Nr. 69(2), S. 199 – 202
- [152] KEALY, M. J. ; TURNER, R.: At Test of the Equality of Close-Ended and Open-Ended Contingent Valuation. In: *American Journal of Agricultural Economics* (1993), Nr. 75, S. 321– 331
- [153] KENDRICK, W.: The Historical Development of National Income Accounts. In: *History of political economy* (1970), S. 284 –315

- [154] KIRCHHOFF, S. ; COLBY, B. G. ; LAFRANCE, J. T.: Evaluating the Performance of Benefit Transfer: An Empirical Inquiry. In: *Journal of Environmental Economics and Management* (1997), Nr. 33, S. 75–93
- [155] KNETSCH, J. L.: Environmental Valuation: Some Problems of Wrong Questions and Misleading Answers. In: *Environmental Values* (1994), Nr. 3, S. 351–368
- [156] KOCHI, I. ; HUBBELL, B. ; KRAMER, R.: An Empirical Bayes Approach To Combining And Comparing Estimates Of The Value Of A Statistical Life. In: *2 World Congress of Environmental and Resource Economists*. Monterey, California, 2002
- [157] KOLSTAD, C.: *Environmental Economics*. New York, Oxford : Oxford University Press, 2000
- [158] KRISTROEM, B.: *Valuing Environmental Benefits Using the Contingent Valuation Method*. University of Umea, 1990
- [159] KRUTILLA, J.: Conservation reconsidered. In: *American Economic Review* (1967), Nr. 56, S. 777–786
- [160] KUO, L. ; CHEN, Z.: Parametric Discrete Choice Models Based on the Scale Mixtures of Multivariate Normal Distributions / University of Connecticut. 1999 (9942). – Forschungsbericht
- [161] KUZNETS, S.: *National Income and Capital Formation, 1919-1935*. 1941
- [162] LANCASTER, K.: Consumer Demand. A New Approach. In: *Journal of Political Economy* (1966), Nr. 74, S. 132–157
- [163] LANCASTER, K.: *A New Approach to Consumer Theory*. New York-London : Columbia University Press, 1971
- [164] LANGFORD, I. H. ; KONTOGIANNI, A. ; SKOURTOS, M. ; GEORGIU, S. ; BATEMAN, I. J.: Multivariate Mixed Models for Open-Ended Contingent Valuation Data. In: *Environmental and Resource Economics* (1998), Nr. 12, S. 44– 456
- [165] LEE, L. F.: Semiparametric Nonlinear Least-Squares Estimation of Truncated Regression Models. In: *Economic Theory* (1992), Nr. 8, S. 52 – 94
- [166] LEHR, U.: Bewertung vor Gericht - Schadensmessung mit Hilfe der "Kontingenten Evaluierungsmethode". In: *RWI-Mitteilungen* (2002), Nr. 53, S. 361 – 377
- [167] LEON, C. J. ; GONZALEZ, R. L. ; VAZQUEZ-POLO, F. J.: Elicitation of Expert Opinion in Benefit Transfer of Environmental Goods. In: *Environmental and Resource Economics* (2003), Nr. 26(2), S. 199–210
- [168] LEON, C. J. ; VASQUEZ-POLO, F. J.: A Bayesian Approach to Double Bounded Contingent Valuation. In: *Environmental and Resource Economics* (1998), Nr. 11, S. 197–215
- [169] LEON, C. J. ; VAZQUEZ-POLO, F. J. ; GUERRA, N. ; RIERAS, P.: A Bayesian Model for Benefit Transfer: Application to National Parks in Spain. In: *Applied Economics* (2002), Nr. 34, S. 749–757
- [170] LEONHARDT, D.: Adding Art to the Rigor of Science. In: *New York Times*, 28. April (2001)

- [171] LEONTIEF, W. W.: *The Structure of American Economy 1919-1939. An Empirical Application of Equilibrium Analysis*. 1941
- [172] LEWBEL, A. ; LINTON, O.: Nonparametric Censored and Truncated Regression. In: *Working Papers of the Boston College* (<http://www2.bc.edu/lewbel>) (2000)
- [173] LEWBEL, A. ; LINTON, O. ; MCFADDEN, D.: Estimating Features of a Distribution from Binomial Data. In: *Working Papers of the Boston College* (<http://www2.bc.edu/lewbel>) (2002)
- [174] LI, C. Z.: Semiparametric Estimation of the Binary Choice Model for Contingent Valuation. In: *Land Economics* (1996), Nr. 72(4), S. 462– 473
- [175] LINTOTT, J.: Environmental Accounting: Useful to Whom and for What? In: *Ecological Economics* (1995), Nr. 16, S. 179 – 190
- [176] LOOMIS, J. ; LOCKWOOD, M. ; DELACY, T.: Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection. In: *Journal of Environmental Economics and Management* (1993), Nr. 25(1), S. 45–55
- [177] LOOMIS, J. B.: Comparative Reliability of the Dichotomous Choice and Open-Ended Contingent Valuation Techniques. In: *Journal of Environmental Economics and Management* (1990), Nr. 18, S. 78 – 85
- [178] LOOMIS, J. B.: The Evolution of a More Rigorous Approach to Benefit Transfer: Benefit Function Transfer. In: *Water Resource Research* (1992), Nr. 28(3), S. 701–705
- [179] LUKEN, R. A. ; JOHNSON, F. R. ; KIBLER, V.: Benefits and Costs of Pulp and Paper Effluent Controls under the Clean Water Act. In: *Water Resource Research* (1992), Nr. 28(3), S. 665–674
- [180] LUNANDER, A.: Inducing Incentives to Understate and to Overstate Willingness to Pay within the Open-Ended and the Dichotomous Choice Elicitation Formats: An Experimental Study. In: *Journal of Environmental Economics and Management* (1998), Nr. 35(1), S. 88–102
- [181] MANSKI, C. F.: Maximum Score Estimation of the Stochastic Utility Model of Choice. In: *Journal of Econometrics*, 3 (1975), S. 205–228
- [182] MANSKI, C. F.: Semiparametric Analysis of Discrete Response. Asymptotic Properties of the Maximum Score Estimator. In: *Journal of Econometrics* (1985), Nr. 27, S. 31– 333
- [183] MANSKI, C. F.: Semiparametric Analysis of Random Effects Linear Models From Binary Panel Data. In: *Econometrica* (1987), Nr. 55(2), S. 357 – 362
- [184] MARSCHAK, J.: Binary Choice Constraints on Random Utility Indications. In: ARROW, K. (Hrsg.): *Stanford Symposium on Mathematical Methods in the Social Sciences*. Stanford, Ca., 1960, S. 312–329
- [185] MATZKIN, R. L.: Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models. In: *Econometrica* (1991), Nr. 59(5), S. 1315 – 1327
- [186] MAX-NEEF, M.: Economic Growth and Quality of Life: A Threshold Hypothesis. In: *Ecological Economics* (1995), Nr. 16, S. 115– 118

- [187] MCCONNELL, K. E.: Model Building and Judgement: Implications for Benefit Transfer with Travel Cost Models. In: *Water Resource Research* (1992), Nr. 28(3), S. 795–700
- [188] MCCONNELL, K.E. ; STRAND, I.: Measuring the cost of time in recreation demand analysis: an application to sportfishing. In: *American Journal of Agricultural Economics* (1981), Nr. 63, S. 153–156
- [189] MCFADDEN, D.: Conditional Logit Analysis of Qualitative Choice Behavior. In: ZAREMBKA, P. (Hrsg.): *Frontiers in Econometrics, Economic Theory and Mathematical Economics*. Academic Press, NY, 1974, S. 105–142
- [190] MCFADDEN, D. ; LEONARD, G. K.: Methodologies for Data Collection and Analysis. In: HAUSMAN, J. A. (Hrsg.): *Contingent Valuation - A Critical Assessment*. Amsterdam, North Holland : Elsevier Science Publishers, 1993
- [191] MESSONNIER, M. L. ; BERGSTROM, J. C.: Survey Response-Related Biases in Contingent Valuation: Concepts, Remedies, and Empirical Application to Valuing Aquatic Plant Management. In: *American Journal of Agricultural Economics* (2000), Nr. 83, S. 438–450
- [192] MILON, J. W.: Contingent Valuation Experiments for Strategic Behavior. In: *Journal of Environmental Economics and Management* (1989), Nr. 17, S. 293 – 308
- [193] MITCHELL, R. C. ; CARSON, R. T.: *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington : Resources for the Future, 1989
- [194] MÜLLER, P. ; QUINTANA, F. A.: Nonparametric Bayesian Data Analysis. In: *Statistical Science* (2004), Nr. 19(1), S. 95–110
- [195] MOGAS, J. (Hrsg.) ; RIERA, P. (Hrsg.) ; EARE (Veranst.): *A Value Inference Test for Forest Functions Using Contingent Valuation and Choice Experiment*. 2003 . – 12th Annual Conference
- [196] MORRISON, M. D. ; BLAMEY, R. K. ; BENNET, J. W.: Minimising Payment Vehicle Bias in Contingent Valuation Studies. In: *Environmental and Resource Economics* (2000), Nr. 16, S. 407–422
- [197] NAVRUD, S. ; KRISTOFERSSON, D.: Validity Tests of Benefit Transfer - Are We Performing the Wrong Tests? In: *Discussion Papers of the Agricultural University of Norway* (2001), Nr. no. D-13
- [198] NUNES, D. ; AUGUSTO, P. ; SCHOKKAERT, E.: Warm Glow and Embedding in Contingent Valuation. In: *FEEM Working Paper No. 73* (2001)
- [199] OTT, K.: Zu einer Konzeption "starker" Nachhaltigkeit. In: BOBBERT, M. (Hrsg.) ; DÜWELL, M. (Hrsg.) ; JAX, K. (Hrsg.): *Umwelt – Ethik – Recht*. 2003
- [200] PARK, T. ; LOOMIS, J. B. ; CREEL, M.: Confidence Intervals for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation Studies. In: *Land Economics* (1991), Nr. 67(1), S. 64– 73
- [201] PARSONS, G. ; KEALY, M. J.: Benefits Transfer in a Random Utility Model of Recreation. In: *Water Resource Research* (1994), Nr. 30, S. 2477–2484

- [202] PATTANAYAK, S. K. ; WING, J. M. ; DEPRO, B. M. ; HOUTVEN, G. L. V. ; CIVITA, P. D. ; STIEB, D. M. ; HUBBELL, B.: International Health Benefits Transfer Application Tool: The Use of PPP and Inflation Indices / Economic Analysis and Evaluation Division, Office of Policy Coordination and Economic Analysis, Policy and Planning Directorate, Healthy Environments and Consumer Safety Branch Health Canada. 2002. – Forschungsbericht
- [203] PEARCE, D. *Valuing Risk to Life and Health: Towards Consistent Transfer Estimates in the European Union and Accession States*. Paper prepared for the European Commission, Workshop on Valuing Mortality and Valuing Morbidity, November 13, 2000, Brussels. 2000
- [204] POULOS, C.: Improving the Reliability of the Benefit Transfer Method: A New Approach and an Application to Malaria Prevention. In: *Discussion papers of the Center for Environmental and Resource Economic Policy, North Carolina State University* (2000)
- [205] RACINE, A. ; GRIEVE, A. P. ; FLÜHLER, H.: Bayesian Methods in Practice: Experiences in the Pharmaceutical Industry. In: *Applied Statistics* (1986), S. 93–150
- [206] RADERMACHER, W.: Societies' Maneuver Towards Sustainable Development: Information and the Setting of Target Values. In: MÜLLER, F. (Hrsg.) ; LEUPOLT, M. (Hrsg.): *Eco Targets, Goal Functions, and Orientors*. 1998
- [207] RANDALL, A.: Total and Nonuse Value. In: BRADEN, J. B. (Hrsg.) ; KOLSTAD, C. D. (Hrsg.): *Measuring the Demand for Environmental Quality*. Amsterdam : Elsevier, 1991, S. 303–321
- [208] RANDALL, A. ; GRUNEWALD, O. ; JOHNSON, S. ; AUSNESS, R. ; PAGOULATOS, A.: Reclaiming Coal Surface Mines in the Central Appalachia: A Case Study of the Benefits and Costs. In: *Land Economics* (1978), Nr. 54, S. 472–489
- [209] READY, R. ; HU, D.: Statistical Approaches to the Fat Tail Problem for Dichotomous Choice Contingent Valuation. In: *Land Economics* (1995), Nr. 71(4), S. 491–499
- [210] READY, R. ; NAVRUD, S. ; DAY, B. ; DUBOURG, R. ; MACHADO, F. ; MOURATO, S. ; SPANNINKS, F. ; RODRIGUEZ, M. X. V. *Benefit Transfer in Europe: Are Values Consistent Across Countries?* Working Papers, Pennstate University. 1999
- [211] REAVES, D. W. ; KRAMER, R. A. ; HOLMES, T. P.: Does Question Format Matter? Valuing an Endangered Species. In: *Environmental and Resource Economics* (1999), Nr. 14, S. 365–383
- [212] REISER, B. ; M., Shechter: Incorporating Zero Values in the Economic Valuation of Environmental Program Benefits. In: *Environmetrics* (1999), Nr. 10(1), S. 87–101
- [213] RESEARCH, ECOTEC ; CONSULTING. *A Cost Benefit Analysis of Reduced Acid Deposition: UK Natural and Semi-Natural Aquatic Ecosystems: a Contingent Valuation Study of Aquatic Ecosystems*. 1993
- [214] RIDKER, R. G.: *Economic Costs of Air Pollution: Studies in Measurement*. New York : Springer, 1967
- [215] ROSEN, H. S.: Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. In: *Journal of Political Economy* (1974), Nr. 82, S. 34–55

- [216] ROSENBERGER, R. S. ; LOOMIS, J. B.: Benefit Transfer of Outdoor Recreation Use Values / U. S. Department of Agriculture, Forest Service. 2000. – Forschungsbericht
- [217] ROSENBERGER, R. S. ; LOOMIS, J. B.: Benefit Transfer of Outdoor Recreation Use Values: A Technical Document Supporting the Forest Service Strategy Plan / U. S. Department of Agriculture, Forest Service. 2001. – Forschungsbericht
- [218] ROWE, R.D. ; D'ARGE, R.C. ; BROOKSHIRE, D.S.: An Experiment on the Economic Value of Visibility. In: *Journal of Environmental Economics and Management* (1980), Nr. 7, S. 1–19
- [219] SAMUELSON, P. A.: Complementarity - An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory. In: *Journal of Economic Literature* (1974), Nr. 12, S. 1255–1289
- [220] SANDERS, L.B. ; WALSH, R.G. ; LOOMIS, J.B.: Toward Empirical Estimation of the Total Value of Protecting Rivers. In: *Water Resources Research* 26,7 (1990), S. 1345–1357.
- [221] SAVAGE, L. J.: *The Foundations of Statistics*. 2nd revised edition (June 1, 1972). Dover Publications, 1954
- [222] SCHÄFER, D: Interpretation und Verknüpfung von Nachhaltigkeitsindikatoren (Interpretation and interlinking of sustainability indicators). In: HARTARD, S. (Hrsg.) ; STAHRMER, C. (Hrsg.) ; HINTERBERGER, F. (. (Hrsg.): *Magische Dreiecke – Berichte für eine nachhaltige Gesellschaft, vol. 1: Stoffflussanalysen und Nachhaltigkeitsindikatoren*. Marburg : metropolis, 2000
- [223] SCHNEIDER, F. ; POMMEREHNE, W. W.: Free Riding and Collective Action: An Experiment in Public Microeconomics. In: *Quarterly Journal of Economics* (1981), Nr. 97, S. 689–702
- [224] SCHULZ, W. ; SCHULZ, E.: Zur umweltpolitischen Relevanz von Nutzen-Kosten-Analysen in der Bundesrepublik. In: *Zeitschrift für Umweltpolitik* (1991), Nr. 14(3), S. 299 – 321
- [225] SCHULZE, W. D.: Embedding and Calibration in Measuring Nonuse Values. In: *Resource and Energy Economics* (1998), Nr. 20(2), S. 163–178
- [226] SCHWARTZ, G.: Estimating the dimension of a model. In: *Annals of Statistics* (1978), Nr. 6, S. 461–464
- [227] SHECHTER, M. ; FREEMAN, S.: Nonuse Value: Reflections on the Definition and Measurement. In: PETHIG, R. (Hrsg.): *Valuing the Environment: Methodological and Measurement Issues*. Dordrecht : Kluwer, 1994, S. 171–194
- [228] SMITH, K. V. ; PATTANAYAK, S. K. ; HOUTVEN, G. L. V. ; BINGHAM, T. H.: Improving the Practice of Benefit Transfer: A Preference Calibration Approach / U. S. Environmental Protection Agency. 2000. – Forschungsbericht
- [229] SMITH, V. K.: On Separating Defensible Benefit Transfers from "Smoke and Mirrors". In: *Water Resource Research* (1992), Nr. 28(3), S. 685–694
- [230] SMITH, V. K. ; OSBORNE, L.: Do Contingent Valuation Estimates Pass a "Scope" Test? A Meta Analysis. In: *Journal of Environmental Economics and Management* (1996), Nr. 31(3), S. 287–301

- [231] SPIEGELHALTER, D. J. ; BEST, N. G. ; CARLIN, B. P. ; LINDE, A. van d.: Bayesian Measures of Model Complexity and Fit (with discussion). In: *Journal of the Royal Statistical Society B* 64 (2002), S. 583–640
- [232] SPIEGELHALTER, D. J. ; THOMAS, A. ; BEST, N. G.: WinBUGS Version 1. User Manual. (1999)
- [233] STACY, E. W.: A Generalization of the Gamma Distribution. In: *Annals of Mathematical Statistics* (1962), Nr. 33, S. 1187 – 1192
- [234] STEPHAN, G. ; AHLHEIM, M.: *Ökonomische Ökologie*. Berlin u. a. O. : Springer-Verlag, 1996
- [235] STONE, N. ; MEADE, J. E.: *National Income and Expenditure*. Oxford University Press, 1944
- [236] STRAND, J. ; NAVRUD, S.: *Norway*. Oxford University Press, 1992, S. 108–134
- [237] STUMBORG, B. E. ; BAERENKLAU, K. A. ; BISHOP, R. C.: Nonpoint Source Pollution and Present Values: A Contingent Valuation Study of Lake Mendota. In: *The Review of Agricultural Economics* (2001), Nr. 23, S. 120–132
- [238] SUTTON, A. J. ; ABRAMS, K. R. ; JONES, D. R. ; SHELDON, T. A. ; SONG, F.: *Methods for Meta-Analysis in Medical Research*. 2002
- [239] SVEDSATER, H.: Contingent Valuation of Global Environmental Resources: Test of Perfect and Regular Embedding. In: *Journal of Economic Psychology* (2000), Nr. 21(6), S. 605–623
- [240] THURSTONE, L.: A Law of Comparative Judgement. In: *Psychological Review* (1927), Nr. 34, S. 273–286
- [241] TINBERGEN, J.: *Kan hier te lande, al dan niet na Overheidsingrijpen, een verbetering van de binnenlandse conjunctuur intreden, ook zonder verbetering van onze exportpositie? Welke lering kan ten aanzien van dit vraagstuk worden getrokken uit de ervaringen van andere landen? [Is a Recovery in the Domestic Economic Situation of this Country Possible, with or without Action on the Part of the Government, even without an Improvement in our Export Position? What Can Be Learned about this Problem from the Experience of other Countries?]*. 1936
- [242] TRAIN, K.: A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit. In: *Working Paper, Department of Economics, University of California, Berkeley* (2001)
- [243] TRAIN, K.: *Discrete Choice Methods with Simulation*. Cambridge University Press, 2003
- [244] TURNER, K. ; PEARCE, D. ; BATEMAN, I.: *Environmental Economics: A elementary introduction*. New York : Harvester Wheatsheaf, 1994
- [245] UM, M.-J. ; KWAK, S.-J. ; KIM, T.-Y.: Estimating Willingness to Pay for Improved Drinking Water Quality Using Averting Behavior Method with Perception Measure. In: *Environmental and Resource Economics* (2002), Nr. 21(3), S. 287–302
- [246] VARTIA, Y. O.: Efficient Methods of Measuring Welfare Change and Compensated Income in Terms of Ordinary Demand Functions. In: *Econometrica* (1983), Nr. 51(1), S. 79 – 98

- [247] WALSH, R. G. ; JOHNSON, D. M. ; MCKEAN, J. R.: Benefit Transfer of Outdoor Recreation Demand Studies, 1968 - 1988. In: *Water Resource Research* (1992), Nr. 28(3), S. 707-713
- [248] WEIKARD, H.-P. *The Existence Value Does not Exist and Non-use Values are Useless*. Annual Conference of the European Public Choice Society, Belgirate, Italy, 4-7 April 2002. 2002
- [249] WEISBROD, B. A.: Collective-consumption services of individual-consumption goods. In: *Quarterly Journal of Economics* (1964), Nr. 78, S. 471-477
- [250] WERNER, M.: Allowing for Zeros in Dichotomous-Choice Contingent-Valuation Models. In: *Journal of Business and Economic Statistics* (1999), Nr. 17(4), S. 479-486
- [251] WHITEHEAD, J. C. ; HAAB, T. C. ; HUANG, J. C.: Partwhole Bias in Contingent Valuation: Will Scope Effects be Detected with Inexpensive Survey Methods? In: *Southern Economic Journal* (1998), Nr. 65(1), S. 160-168
- [252] WHITTINGTON, D. ; SMITH, V. K.: Giving Respondents Time to Think in Contingent Valuation Studies: A Developing Country Application. In: *Journal of Environmental Economics and Management* (1992), Nr. 22, S. 205 - 225
- [253] WILLIS, K. G. ; GARROD, G. D. ; SAUNDERS, C. M.: Benefits of Environmentally Sensitive Area Policy in England: A Contingent Valuation Assessment. In: *Journal of Environmental Management* (1995), Nr. 44, S. 105-125
- [254] WOOLDRIDGE, J. M.: *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2002
- [255] ZHOU, M.: Nonparametric Bayes Estimator of Survival Functions for Doubly/Interval Censored Data / University of Kentucky. 2002. - Forschungsbericht

HOHENHEIMER VOLKSWIRTSCHAFTLICHE SCHRIFTEN

- Band 1 Walter Deffaa: Anonymisierte Befragungen mit zufallsverschlüsselten Antworten. Die Randomized-Response-Technik (RRT). Methodische Grundlagen, Modelle und Anwendungen. 1982.
- Band 2 Thomas Michael Baum: Staatsverschuldung und Stabilisierungspolitik in der Demokratie. Zur neoinstitutionalistischen Kritik der keynesianischen Fiskalpolitik. 1982.
- Band 3 Klaus Schröter: Die wettbewerbspolitische Behandlung der leitungsgebundenen Energiewirtschaft. Dargestellt am Beispiel der Fernwärmewirtschaft der Bundesrepublik Deutschland. 1986.
- Band 4 Hugo Mann: Theorie und Politik der Steuerreform in der Demokratie. 1987.
- Band 5 Max Christoph Wewel: Intervallararithmetische Dependenzanalyse in der Ökonometrie. Ein konjekturaler Ansatz. 1987.
- Band 6 Heinrich Pascher: Die U.S.-amerikanische Deregulation Policy im Luftverkehrs- und Bankenbereich. 1987.
- Band 7 Harald Lob: Die Entwicklung der französischen Wettbewerbspolitik bis zur Verordnung Nr. 86-1243 vom 01. Dezember 1986. Eine exemplarische Untersuchung der Erfassung der Behinderungsstrategie auf der Grundlage des Konzepts eines wirksamen Wettbewerbs. 1988.
- Band 8 Ulrich Kirschner: Die Erfassung der Nachfragemacht von Handelsunternehmen. Eine Analyse der ökonomischen Beurteilungskriterien und der wettbewerbsrechtlichen Instrumente im Bereich der Verhaltenskontrolle. 1988.
- Band 9 Friedhelm Herb: Marktwirtschaftliche Innovationspolitik. 1988.
- Band 10 Claus Schnabel: Zur ökonomischen Analyse der Gewerkschaften in der Bundesrepublik Deutschland. Theoretische und empirische Untersuchungen von Mitgliederentwicklung, Verhalten und Einfluß auf wirtschaftliche Größen. 1989.
- Band 11 Jan B. Rittaler: Industrial Concentration and the Chicago School of Antitrust Analysis. A Critical Evaluation on the Basis of Effective Competition. 1989.
- Band 12 Thomas März: Interessengruppen und Gruppeninteressen in der Demokratie. Zur Theorie des Rent-Seeking. 1990.
- Band 13 Andreas Maurer: Statistische Verfahren zur Ermittlung von oligopolistischen Strukturen. 1990.
- Band 14 Peter Mandler: Zur ökonomischen und politisch-institutionellen Analyse öffentlicher Kredithilfen. 1992.
- Band 15 Heinrich J. Engelke: Die Interpretation der Rundfunkfreiheit des Grundgesetzes: Eine Analyse aus ökonomischer Sicht. 1992.
- Band 16 Thomas Fischer: Staat, Recht und Verfassung im Denken von Walter Eucken. Zu den staats- und rechtstheoretischen Grundlagen einer wirtschaftsordnungspolitischen Konzeption. 1993.
- Band 17 Stefan Eißer: Innovationswettbewerb. Determinanten und Unternehmensverhalten. 1993.
- Band 18 Reinhard Scharff: Regionalpolitik und regionale Entwicklungspotentiale. Eine kritische Analyse. 1993.
- Band 19 Karin Beckmann: Probleme der Regionalpolitik im Zuge der Vollendung des Europäischen Binnenmarktes. Eine ökonomische Analyse. 1995.

- Band 20 Bernd Nolte: Engpaßfaktoren der Innovation und Innovationsinfrastruktur. Eine theoretische und empirische Analyse für ländliche Wirtschaftsräume in Baden-Württemberg. 1996.
- Band 21 Klaus-Rainer Brintzinger: Die Nationalökonomie an den Universitäten Freiburg, Heidelberg und Tübingen 1918 - 1945. Eine institutionenhistorische, vergleichende Studie der wirtschaftswissenschaftlichen Fakultäten und Abteilungen südwestdeutscher Universitäten. 1996.
- Band 22 Steffen Binder: Die Idee der Konsumentensouveränität in der Wettbewerbstheorie. Telemekratische vs. nomokratische Auffassung. 1996.
- Band 23 Alexander Burger: Deregulierungspotentiale in der Gesetzlichen Rentenversicherung. Reformnotwendigkeiten versus Reformmöglichkeiten. 1996.
- Band 24 Burkhard Scherer: Regionale Entwicklungspolitik. Konzeption einer dezentralisierten und integrierten Regionalpolitik. 1997.
- Band 25 Frauke Wolf: Lorenzkurvendisparität. Neuere Entwicklungen, Erweiterungen und Anwendungen. 1997.
- Band 26 Hans Pittlik: Politische Ökonomie des Föderalismus. Föderative Kompetenzverteilung im Lichte der konstitutionellen Ökonomik. 1997.
- Band 27 Stephan Seiter: Der Beitrag Nicholas Kaldors zur Neuen Wachstumstheorie. Eine vergleichende Studie vor dem Hintergrund der Debatte über den Verdoorn-Zusammenhang. 1997.
- Band 28 André Schmidt: Ordnungspolitische Perspektiven der europäischen Integration im Spannungsfeld von Wettbewerbs- und Industriepolitik. 1998.
- Band 29 Bernd Blessin: Innovations- und Umweltmanagement in kleinen und mittleren Unternehmen. Eine theoretische und empirische Analyse. 1998.
- Band 30 Oliver Letzgas: Die Ökonomie internationalen Umweltschutzes. 1999.
- Band 31 Claudia Hafner: Systemwettbewerb versus Harmonisierung in Europa. Am Beispiel des Arbeitsmarktes. 1999.
- Band 32 Jürgen Kulle: Ökonomie der Musikindustrie. Eine Analyse der körperlichen und unkörperlichen Musikerverwertung mit Hilfe von Tonträgern und Netzen. 1998.
- Band 33 Michael Ganske: Intertemporale Aspekte von Staatsverschuldung und Außenhandel. 1999.
- Band 34 Margit Ströbele: Die Deregulierungswirkungen der europäischen Integration. Das Beispiel der Sondermärkte. 1999.
- Band 35 Marion Benesch: Devisenmarktinterventionen in Theorie und Praxis. Eine umfassende Analyse ihrer Zielsetzungen, Wirkungsweisen und wirtschaftspolitischen Bedeutung. 1999.
- Band 36 Torsten Gruber: Unterschiedliche geldpolitische Transmissionsmechanismen und Stabilitätskulturen als mögliche Ursachen geldpolitischer Spannungen in der Europäischen Währungsunion. 2000.
- Band 37 Bertram Melzig-Thiel: Arbeit in der Informationsgesellschaft. Chancen und Risiken neuer Informations- und Kommunikationstechnologien für die Beschäftigung. 2000.
- Band 38 Annette Fritz: Die Entsorgungswirtschaft im Spannungsfeld zwischen Abfallpolitik und Kartellrecht. Eine industrieökonomische Branchenstudie. 2001.
- Band 39 Harald Strotmann: Arbeitsplatzdynamik in der baden-württembergischen Industrie. Eine Analyse mit amtlichen Betriebspaneldaten. 2002.

- Band 40 Dietrich Benner: Qualitätsungewißheit bei Gütern mit Vertrauenseigenschaften. Entwicklung und Anwendung eines entscheidungstheoretisch fundierten Analyserahmens. 2002.
- Band 41 Jürgen M. Schechler: Sozialkapital und Netzwerkökonomik. 2002.
- Band 42 Kay-Uwe May: Haushaltskonsolidierung durch Ausgabekürzungen. Restriktionen und Strategien. 2002.
- Band 43 Peter Kühni: Der Wechselkurs als Zwischenziel der Geldpolitik im Aufholprozess. Die monetärkeynesianische Entwicklungsstrategie der Berliner Schule vor dem Hintergrund der makroökonomischen Entwicklung ausgewählter Länder Mittel- und Osteuropas. 2003.
- Band 44 Steffen Wirth: Nichtparametrische Analyse von Bildungsertragsraten. Neuere Entwicklungen und Anwendungen. 2003.
- Band 45 Bernhard Holwegler: Innovation, Diffusion und Beschäftigung. Die ökonomische Theorie der Technologiediffusion und ihr Beitrag zur Erklärung technologischer Arbeitslosigkeit. 2003.
- Band 46 Guntram R. M. Hepperle: Zukunftsorientierte Industriepolitik. Möglichkeiten und Grenzen. 2004.
- Band 47 Udo Vullhorst: Stabilisierungspolitik bei supranationaler Geldpolitik und nationaler Fiskalpolitik. Eine spieltheoretische Betrachtung. 2004.
- Band 48 Matthias Rösch: Die Bedeutung von Investivlöhnen und Gewinnbeteiligungen für Einkommensverteilung und Beschäftigung. 2004.
- Band 49 Michael Bubik: Erfolgskriterien für Unternehmenszusammenschlüsse. Eine theoretische und exemplarische Analyse. 2005.
- Band 50 Jörg Weltin: Internationale Unternehmensbesteuerung. Allokation der Besteuerungsrechte unter veränderten Rahmenbedingungen. 2005.
- Band 51 Susanne Reichart: Zum Konvergenzprozess der mittel- und osteuropäischen EU-Beitrittsländer. 2005.
- Band 52 Daniel Hartmann: Geldpolitik und Beschäftigung. Die geldpolitische Strategie der Federal Reserve: Vorbild oder Auslaufmodell? 2005.
- Band 53 Marc Peter Radke: Explaining Financial Crises. A Cyclical Approach. 2005.
- Band 54 Katja Hölsch: Umverteilungseffekte in Europa. Eine Analyse für ausgewählte Länder. 2006.
- Band 55 Ulrike Lehr: Contingent Valuation Daten und Bayes'sche Verfahren. Ein Vorschlag zur Verbesserung von Umweltbewertung und Nutzentransfer. 2006.