

# Der Überschätzungsfaktor in Zahlungsbereitschaftsanalysen – ein in-sample Test

The overestimation in willingness to pay analyses – a within-subject-test

Uta SAUER und Rainer MARGGRAF

## Zusammenfassung

Aufgrund der Vielzahl möglicher Verzerrungsgründe ist eine zentrale Frage in der Kontingenten Bewertung, ob die ermittelten Werte der hypothetischen Studien die tatsächlichen Beträge reflektieren, die Individuen für das betreffende Gut zahlen würden. Der vorliegende Beitrag bezieht sich auf eine in-sample Studie mit der Bewertung eines öffentlichen Gutes, bei der sowohl die hypothetischen Zahlungsbereitschaften als auch die tatsächlichen Zahlungen im Rahmen einer repräsentativen Befragung im Landkreis Northeim in Südniedersachsen verglichen wurden. Im Ergebnis haben nur acht von 89 zahlungsbereiten Befragten tatsächlich gezahlt. Dies entsprach einem Überschätzungsfaktor von 17,47, der deutlich höher ist als allgemein angenommen wird. In der Diskussion werden potentielle Gründe für dieses Ergebnis betrachtet.

**Schlagnote:** Agrarumweltprogramm, intra-personeller Vergleich, Überschätzungsfaktor, öffentliche Güter, reale Zahlungsbereitschaften

## Summary

Due to a great number of potential distortions, one of the main questions within contingent valuation is to which extent the stated willingness to pay (WTP) reflects the actual WTP of the respondents for the good in question. The present article refers to a study conducted in the district of Northeim, Lower Saxony, which combines a within-subject-test, a public good valuation, and the comparison of hypothetical and actual WTP in the context of a field study. The main

finding in this study is that only eight of 89 respondents stating a positive WTP did actually transfer money. As a consequence, a 17.47-times overestimation occurred. Hence potential reasons for the strong difference of hypothetical and actual WTP are discussed.

**Keywords:** Actual willingness to pay, agri-environmental scheme, overestimation, within-subject-test, public goods

## 1. Einleitung

Mit der Erhebung von Zahlungsbereitschaften wird die gesellschaftliche Nachfrage für umweltbezogene, öffentliche Güter und Leistungen ermittelt. Zahlungsbereitschaften werden in der Kontingenten Bewertung (KB) durch direkte Preisbefragungen erhoben. Dabei können eine Reihe möglicher Faktoren auftreten, die zu einer Divergenz hypothetischer und realer Zahlungsbereitschaften führen und in eine Überbewertung der betreffenden Umweltgüter münden.

Studien, die sich mit dem Ausmaß von Verzerrungen in der Kontingenten Bewertung befassen, wenden unterschiedliche Methoden an. Unterschiede bestehen z. B. in der Art des Vergleiches (Vergleich innerhalb einer Stichprobe oder zwischen verschiedenen Stichproben), des Frageformats (offen, geschlossen), der Art des Gutes (privat, öffentlich), des Untersuchungstyps (Laborexperimente, Feldstudien) und der Erhebungsmethode (KB, Discrete Choice). Sogenannte in-sample Studien führen einen Vergleich innerhalb einer Gruppe von TeilnehmerInnen durch, während split-sample Studien die Antworten zwischen verschiedenen Gruppen vergleichen, d. h. zwischen einer hypothetischen und einer realen Zahlungsbereitschaftsstichprobe. Bisher haben sich nur wenige Studien mit dem Unterschied hypothetischer und realer Zahlungsbereitschaften innerhalb einer Stichprobe von Befragten im intra-personellen Vergleich befasst. In der vorliegenden Studie wurde dieses Erhebungsformat mit der Bewertung eines öffentlichen Gutes im Rahmen einer Feldstudie kombiniert.

## 2. Vermeidung von Antwortverzerrungen in der Studie

Zur ex-ante Vermeidung bekannter Verzerrungsmöglichkeiten wurde das Befragungsdesign und das Szenario sowohl für die Zahlungsbereitschaft als auch für die Zahlung bereits im Vorfeld entsprechend entwickelt (s. Abbildung 1).

Zahlungsbereitschaft	Zahlung
<p><b>Befragungsdesign</b></p> <ul style="list-style-type: none"> <li>- Ziehung repräsentative Stichprobe</li> <li>- Bewertungsszenario plausibel und verständlich</li> <li>- zu bewertendes Gut in Menge und Qualität genau definiert</li> </ul> <p><b>Interviewerbias</b></p> <ul style="list-style-type: none"> <li>- Standardisierung des Interviews und Interviewer-Training</li> </ul> <p><b>Hypothetischer Bias</b></p> <ul style="list-style-type: none"> <li>- Vorteil eines nicht hypothetischen Marktes und in der Region durch die lokale Presse bekannt</li> </ul> <p><b>Embedding Effekt</b></p> <ul style="list-style-type: none"> <li>- Gut war nicht Teil eines Güterbündels und verlangte keine Abstraktionsfähigkeiten</li> </ul>	<p><b>Strategisches Verhalten</b></p> <ul style="list-style-type: none"> <li>- Anreizstruktur bzgl. der Bereitstellung: Budgetuntergrenze (PP) für die Anlage eines Gewässerrandstreifens in Verbindung mit Geld-zurück-Garantie (MBG)</li> </ul> <p><b>Zahlungsmodalität</b></p> <ul style="list-style-type: none"> <li>- geringe Transaktionskosten durch vorgefertigten Überweisungsträger</li> </ul> <p><b>Soziale Erwünschtheit</b></p> <ul style="list-style-type: none"> <li>- Projektmitglieder konnten keine Zahlungen zuordnen – Datenschutzbestimmung</li> </ul> <p><b>Erinnerung</b></p> <ul style="list-style-type: none"> <li>- nach der Befragung Erinnerung in der lokalen Presse für die Überweisung mit Fristsetzung</li> </ul>

Abb. 1: Instrumente zur Vermeidung von Antwortverzerrungen in der Studie

Quelle: Verändert nach SAUER, 2010, 151

Bei der empirischen Erhebung von Präferenzen können zunächst Rahmeneffekte zu potentiellen Verzerrungen der Ergebnisse führen. Auswahlfehler, Teilnahmeverweigerungen und die Bereitstellung unterschiedlicher oder mangelhafter Informationen gehören zu diesen allgemeinen Befragungsfehlern, die auch die Vergleichbarkeit der Antworten untereinander erheblich einschränken können. Zur Vermeidung von Verzerrungen durch Auswahlfehler wurde eine repräsentative Stichprobe im Landkreis Northeim in Südniedersachsen gezogen. Alle Stichproben wurden zufällig gezogen, um einen systematischen Ausschluss von Teilen der Grundgesamtheit zu vermeiden (SCHNELL et al., 2005).

Um Informationsverzerrungen durch fehlende, mangelhafte oder einseitige Informationen über das zu bewertende Gut zu verhindern, wurden die Informationen standardisiert entwickelt. Auch die Befragungsdurchführung wurde zur Vermeidung eines Interviewerbias standardisiert und die Sozialkompetenz der InterviewerInnen mithilfe eines Trainings optimiert.

Um gültige und politikrelevante Antworten mit der Kontingenten Bewertungsmethode erzielen zu können, sind insbesondere zur Gestaltung des Bewertungsszenarios einige weitere Anforderungen zu beachten. Das Szenario sollte plausibel und verständlich sein, und das Bewertungsgut muss, um ökonomische Voraussetzungen zu erfüllen, in Menge und Qualität genau beschrieben werden. Hierbei sind vor allem das mögliche Auftreten des Hypothetischen Bias und Embedding Effekte zu berücksichtigen. Das mit hypothetischen Märkten verbundene fiktive Szenario hat zur Folge, dass die Antworten nicht mit tatsächlichen finanziellen Konsequenzen verknüpft sind. Dies ist den Befragten bewusst, und sie berücksichtigen bei ihrer Antwort bspw. ihr monatliches Haushaltsbudget nicht. Die meisten der Zahlungsbereitschaftsanalysen mit öffentlichen Gütern haben kein realistisches Gut, das glaubwürdig präsentiert wird und für das potentielle Gelder auch tatsächlich eingesetzt werden könnten. Deshalb arbeiten viele Studien mit Spenden für einen Fonds, ohne ein Referenzszenario für die Umsetzung der Gelder zu besitzen. Oder die ForscherInnen greifen auf private Güter zurück, die wiederum wenig mit der Natur von öffentlichen Gütern gemein haben. In der vorliegenden Studie wurde für die Ermittlung der Zahlungsbereitschaft ein Umweltgut eingesetzt, das einem öffentlichen Gut entspricht, mit dem zusätzlichen Vorteil eines vorhandenen Marktes, auf dem es potenziell bereitgestellt werden konnte (SAUER und FISCHER, 2010, 4).

Der Untersuchungsrahmen war das Northeimer Modell, welches innerhalb eines Pilot-Projektes (2000-2009) entwickelt und im Landkreis Northeim in die Praxis umgesetzt wurde. Hierbei handelte es sich um ein regionalisiertes und ergebnisorientiertes Honorierungskonzept für ökologische Leistungen der Landwirtschaft auf der Basis marktwirtschaftlicher Prinzipien. Durch die Umsetzung der Honorierung im Landkreis Northeim seit 2004 gab es ein Referenzszenario für die Gelder. Auf diese Weise wurde der möglichen Entstehung eines hypothetischen Bias bezüglich der Zahlungsbereitschaftsäußerungen entgegengewirkt.

Im Rahmen einer Befragung (RÜFFER, 2006) zu den Präferenzen der Bevölkerung in der Projekt-Region wurde eine große Nachfrage nach Gewässerrandstreifen ermittelt, und deshalb wurden diese als Gut für das Bewertungsszenario verwendet. Da eine genaue Beschreibung des

zu bewertenden Gutes für Zahlungsbereitschaftsanalysen genauso grundlegend ist wie die Nutzungsmöglichkeiten des Gutes durch die Befragten, wurden konkret einige wenige Flüsse für das Referenzszenario ausgewählt, die nahezu alle Städte und Gemeinden des Landkreises Northeim abdeckten (SAUER, 2010, 74f).

Innerhalb der Beschreibung des Honorierungssystems im ersten Abschnitt des Fragebogens wurde gleichzeitig die Bedeutung des Gutes für die zu bewertende Politik bezüglich aktueller Agrarumweltprogramme verdeutlicht und die Qualität des Gewässerrandstreifens beschrieben. Ferner war innerhalb der Zahlungsbereitschaftsfrage das zu bewertende Gut in der Menge genau definiert: „Wären Sie bereit, die Anlage von 100 km Gewässerrandstreifen durch einen einmaligen finanziellen Beitrag zu unterstützen?“ und „Wie viel wären Sie bereit zu zahlen?“. Embedding Effekte können aufgrund einer räumlichen, zeitlichen oder inhaltlichen kontextuellen Einbettung des zu bewertenden Gutes zu einer Überbewertung führen. Da das betreffende Gut in der Studie diskret, also nicht Teil eines Güterbündels war, wurden zumindest keine Abstraktionsfähigkeiten von den Befragten verlangt.

Im Fall öffentlicher Güter wird häufig diskutiert, dass die Durchführung tatsächlicher Zahlungen in nicht geringem Maße von der Anreizstruktur abhängig ist, die durch den Zahlungskontext und durch das Ausmaß der Überwindung des *free rider* Problems bestimmt wird (FOSTER et al., 1997). Um Anreize für eine Zahlung zu geben und strategisches Verhalten durch Unter- oder Übertreibung zu verhindern, wurden speziell für die Zahlung verschiedene Techniken in der Studie integriert. Zum einen wurde ein Provision Point (PP) als Budgetuntergrenze für die Bereitstellung von Gewässerrandstreifen eingeführt, der zum anderen mit einer Money Back Garantie (MBG) verknüpft war. Mit der Übergabe des Überweisungsträgers wurden die Befragten darauf hingewiesen, dass bei Nicht-Erreichung eines Schwellenwertes keine Gewässerrandstreifen angelegt und die eingegangenen Gelder zurück überwiesen würden.

Die Zahlungsmodalitäten wurden mithilfe eines vorgefertigten Überweisungsträgers so einfach wie möglich gestaltet. Außerdem wurde eine Datenschutzbestimmung ausgehändigt, die im Rahmen der Zahlung u. a. zur Vermeidung einer Verzerrung durch soziale Erwünschtheit beitragen sollte. Durch die Anonymität der

Überweisung und die damit fehlende personenbezogene Zuordnungsmöglichkeit sollten sich die Befragten bei der Zahlung nicht genötigt fühlen, ihren vorher geäußerten Betrag tatsächlich zu überweisen. Zusätzlich wurde etwa einen Monat nach der Befragung eine Danksagung für die hohe Befragungsbeteiligung in lokalen Zeitungen veröffentlicht inkl. einer Erinnerung, die die zahlungsbereiten Befragten darauf hinwies, dass sie noch bis zu einem bestimmten Zeitpunkt die Gelegenheit für die Durchführung einer Überweisung hatten (SAUER und FISCHER, 2010, 7).

### 3. Vergleich der Zahlungsbereitschaften und Zahlungen

Die Zahlungsbereitschaftsfrage beantworteten 89 der 305 Befragten positiv, davon acht ohne einen Betrag anzugeben. Da insgesamt 11 Fälle keine gültigen Beträge aufwiesen, wurde für die Berechnungen eine Basispopulation von  $n=294$  zugrundegelegt.

Aus Tabelle 1 sind der Mittelwert, die Standardabweichung, der niedrigste und höchste genannte Betrag sowie die Gesamtsumme der hypothetischen Zahlungsbereitschaften und der tatsächlichen Zahlungen für 100 km Gewässerrandstreifen zu entnehmen. Im Vergleich zeigt der Mittelwert der hypothetischen Zahlungsbereitschaften von € 11,29 einen deutlichen Unterschied zum Mittelwert der tatsächlichen Zahlungen von € 0,65. Die Gesamtsumme der hypothetisch geäußerten Zahlungsbereitschaften lag bei € 3.320,- mit Einzelbeträgen, die bis zu € 200,- reichten. Insgesamt haben nur acht der 89 zahlungsbereiten Befragten ihre Zahlungsbereitschaft auch in die Tat umgesetzt und gezahlt. Daraus ergab sich eine zur Verfügung stehende Gesamtsumme von € 190,-.

Tab. 1: Vergleich der Zahlungsbereitschaften und Zahlungen

<b>n=294</b>	<b>Zahlungsbereitschaften in €</b>	<b>Zahlungen in €</b>
Mittelwert	11,2925	0,6463
Standardabweichung	27,58226	6,19016
Minimum-Maximum	0-200	0-100
Summe	3320	190

Quelle: Eigene Darstellung

In der Literatur herrscht Konsens darüber, dass hypothetisch geäußerte Beträge die tatsächlichen in den meisten Fällen übersteigen. Um eine potentiell signifikante Überschätzung der Zahlungsbereitschaftsergebnisse ex-post zu verhindern, kann ein Instrument eingesetzt werden, dass die erhobenen Zahlungsbereitschaften mithilfe eines Faktors kalibriert (BRÄUER und SUHR, 2005; FOX et al., 1998).

Der Überschätzungsfaktor spiegelt das Verhältnis von aufgedeckten Präferenzen zu hypothetischen Zahlungsbereitschaften wider und gibt damit an, um wie viel die geäußerten die beobachteten Zahlungsbereitschaften überschätzen. In Tabelle 2 sind neben dem Mittelwert des Überschätzungsfaktors der vorliegenden Studie (n=294) auch die Mittelwerte von allen Studien (N=131), die einen Vergleich hypothetischer und realer Zahlungsbereitschaften durch Aufdeckung individueller Präferenzen durchgeführt haben, auf Grundlage der Untersuchung von BRÄUER und SUHR (2005) aufgeführt. Diese sind um einige nach 2002 veröffentlichte Studien ergänzt. Außerdem sind Mittelwertvergleiche von Studien mit Erhebungsformaten, die die vorliegende Studie kombiniert hat, dargestellt.

In der zugrundeliegenden Studie haben insgesamt nur sehr wenige Befragte ihre Zahlungsbereitschaftsaussage in die Tat umgesetzt. Die Ergebnisse aus Tabelle 1 entsprechen einem Überschätzungsfaktor von 17,47. Damit ist der beobachtete Überschätzungsfaktor der Studie deutlich höher als die Überschätzungsfaktoren aller Einzelvergleiche bisheriger Studien.

Tab. 2: Beobachtete Überschätzungsfaktoren

	Mittelwert	n=Anzahl Befragte N=Anzahl Studien
Studie	17,47	n=294
Total	2,5	N=131
Offenes Frageformat	2,71	N=57
In-sample	2,39	N=29
Öffentliche Güter	2,68	N=48
Feldexperimente	2,61	N=57

Quelle: Eigene Darstellung

Bei der Interpretation der Ergebnisse muss allerdings die Erhebungsform berücksichtigt werden. Die in-sample-Vergleiche, die mit öffentlichen Gütern oder im Feld durchgeführten Studien weisen

nur einen Mittelwert der Überschätzungsfaktoren von 2,39 (in-sample mit wenigen Vergleichsstudien insgesamt) bis 2,68 (öffentliche Güter) auf. Aber es gibt bis dato keine Studie, die diese Erhebungskombination mit einer zusätzlichen Echtzahlung für ein öffentliches Gut, das potentiell bereit gestellt werden konnte, durchgeführt hat.

#### 4. Diskussion

Im Ergebnis haben nur sehr wenige Befragte die vorher geäußerte Zahlungsbereitschaft in die Tat umgesetzt und gezahlt. Untersuchungen zur Repräsentativität der Stichprobe zeigten, dass die Northeimer Grundgesamtheit hinreichend abgebildet wurde. Auch Interviewereffekte konnten weder bzgl. der Zahlungsbereitschaft noch der Zahlung identifiziert werden (SAUER, 2010, 94ff). Deshalb stellt sich die Frage, welche Gründe für die geringe Zahlung denkbar sind.

Einer der Gründe könnte sein, dass, obwohl die Zahlungsmodalität mit einem vorgefertigten Überweisungsträger so einfach wie möglich gestaltet wurde, die Zahlung an sich trotzdem noch mit zu hohen Transaktionskosten verbunden war. Eine alternative Zahlungsmöglichkeit mit EC-Karte bspw. hätte allerdings direkt vorgenommen werden müssen und die Befragten evtl. zu sozial wünschenswerten Handlungen gedrängt. Ein weiterer Punkt betrifft die Einbettung des zu bewertenden Gutes in das Honorierungssystem für LandwirtInnen. Denn das Anlegen eines Gewässerrandstreifens war gleichzeitig mit der Bezahlung der Landwirte für diese Leistung verknüpft. Allerdings würde man unterstellen, dass sich eine ablehnende Haltung bzgl. der Honorierung von LandwirtInnen bereits bei der Beantwortung der hypothetischen Zahlungsbereitschaftsfrage gezeigt hätte und damit eher eine Erklärung für die Anzahl nicht zahlungsbereiter Befragter wäre, als für eine geringe Umsetzung der von ihnen geäußerten Zahlungsbereitschaften. Dies wurde bei der Beantwortung der Protestantworten auch bestätigt (SAUER, 2010, 99ff).

Ein weiteres Argument für eine geringe Umsetzung der Zahlungsbereitschaften ist die Art und Weise, wie die Echtzahlung eingeleitet wurde. Die Befragten waren durch die tatsächliche Zahlungsaufforderung mit der Übergabe des Überweisungsträgers im Anschluss an die hypothetisch gestellte Frage möglicherweise

überrascht. Dennoch sollte eine Verhaltensintention in Verhalten münden, auch wenn die Vorgehensweise ungewöhnlich sein mag. Um keinen sozialen Druck zu erzeugen, wurde den Befragten, neben der Information zur Anonymität, der Überweisungsträger überlassen. Allerdings kann es im Rahmen der Zahlungsbereitschaftsfrage zu Effekten sozialer Erwünschtheit gekommen sein, die die starke Differenz zwischen hypothetischen Zahlungsbereitschaften und tatsächlichen Zahlungen zumindest in gewissem Ausmaß begründen könnten. Während die Anwesenheit von InterviewerInnen bei der Befragung die Antworten der TeilnehmerInnen möglicherweise beeinflusst hat, war die Zahlung dagegen völlig anonym (SAUER und FISCHER, 2010, 7).

Auch die Heterogenität innerhalb der individuellen Entscheidungsfindung kann eine bedeutende Rolle spielen und zu Verhalten führen, das nicht vorausgesagt werden kann. Die Duale Prozess-Theorie bspw. beschreibt auf der einen Seite einen holistisch-heuristischen Prozess, der von äußeren Signalen als affektive Faktoren geleitet wird, und auf der anderen Seite einen systematisch-analytischen Prozess, zu denen Präferenzen und die wahrgenommene Kontrolle der Befragten gehören. Während manche TeilnehmerInnen ihre Einstellungen, Überzeugungen und Werte bei der Formulierung ihrer Zahlungsbereitschaft erwägen und einbeziehen, antworten andere eher habitualisiert und in impulsiver Art und Weise (LIENHOOP und FISCHER, 2009).

Nichtsdestotrotz stellt sich die Frage, ob Einstellungen ggü. dem betreffenden Gut möglicherweise ein geeigneterer Indikator für die Messung individueller Präferenzen für öffentliche Güter ist. Allerdings werden sogenannte „ernsthafte Interessen“, bei denen Individuen bereit sind, für das betreffende Gut etwas anderes einzutauschen, damit nicht ermittelt. Aber inwieweit kann die Kontingente Bewertung dazu beitragen, „ernsthafte Interessen“ zu identifizieren (SAUER, 2010, 152f)? Um Vergleichsmöglichkeiten und weitere Erkenntnisse zu Überschätzungen geäußerter Präferenzen in der vorliegenden Kombination zu gewinnen, sind weitere Untersuchungen erforderlich. Erst dann kann eingeschätzt werden, ob die Ergebnisse der Studie möglicherweise als Ausreißerwert zu betrachten sind. Zusätzlich kann der individuelle Entscheidungsprozess als psychologische Erklärungsvariable Berücksichtigung finden.

Die tatsächlich überwiesene Summe wurde für die Anlage und Pflege eines ca. 400 Meter langen und 5 Meter breiten Gewässerrandstreifens über einen Zeitraum von fünf Jahren im Landkreis Northeim verwendet. Neben einer Veröffentlichung in regionalen Zeitungen wurde der Gewässerrandstreifen auch mit einer Informationstafel gekennzeichnet (SAUER, 2010, 154ff).

### Literatur

- BRÄUER, I. und SUHR, A. (2005): Ergebnisse von Zahlungsbereitschaftsanalysen – Interpretation und Verwendung. In: Marggraf, R., Bräuer, I., Fischer, A., Menzel, S., Stratmann, U. und Suhr, A. (Hrsg.): Ökonomische Bewertung bei umweltrelevanten Entscheidungen – Einsatzmöglichkeiten von Zahlungsbereitschaftsanalysen in Politik und Verwaltung. Marburg: metropolis, 149-184.
- FOSTER, V., BATEMAN, I. J. und HARLEY, D. (1997): Real and hypothetical willingness to pay for environmental preservation: A non-experimental comparison. *Journal of Agricultural Economics*, 48, 123-138.
- FOX, J. A., SHOGREN, J. F., HAYES, D. J. und KLIEBENSTEIN, J. B. (1998): CVM-X. Calibrating contingent values with experimental auction markets. *American Journal of Agricultural Economics*, 80, 455-465.
- LIENHOOP, N. und FISCHER, A. (2009): Can you be bothered? The role of participant motivation in the valuation of species conservation measures. *Journal of Environmental Planning and Management*, 52, 519-534.
- RÜFFER, C. (2006): Merit goods determined by society value judgements – Political implications for public participation. Dissertation an der Georg-August-Universität Göttingen.
- SAUER, U. (2010): Werte und tatsächliches Verhalten in der Kontingenten Bewertung – Eine empirische Studie im Landkreis Northeim. Dissertation an der Georg-August-Universität Göttingen.
- SAUER, U. und FISCHER, A. (2010): Willingness to pay, attitudes and fundamental values – on the cognitive context of public preferences for diversity in agricultural landscapes. *Ecological Economics*, 70, 1-9.
- SCHNELL, R., ESSER, E. und HILL, P. B. (2005): Methoden der empirischen Sozialforschung. München: Oldenbourg Verlag.

### Anschrift der VerfasserInnen

*Dr. Uta Sauer und Prof. Dr. Rainer Marggraf  
 Department für Agrarökonomie und Rurale Entwicklung  
 Georg-August-Universität Göttingen  
 Platz der Göttinger Sieben 5, 37073 Göttingen, Deutschland  
 Tel.: +49 551 39 4814  
 eMail: usauer@uni-goettingen.de*