

Felix Schläpfer, Peter Zweifel*

Nutzenmessung bei öffentlichen Gütern: Konzeptionelle und empirische Probleme in der Praxis

Die Abwägung von Kosten und Nutzen spielt bei der Bereitstellung öffentlicher Güter eine zentrale Rolle. Mithilfe welcher Methode lässt sich der Nutzen öffentlicher Güter am besten schätzen? Welche typischen konzeptionellen und empirischen Probleme treten bei den einzelnen Schätzmethoden häufig auf?

Aus ökonomischer Sicht ist die Kosten-Nutzen-Analyse das zentrale Paradigma im öffentlichen Sektor. Öffentliche Güter sollen dann bereitgestellt werden, wenn die über alle betroffenen Individuen aggregierten Nutzen die Kosten überwiegen.¹ Der Nutzen (englisch: benefit) ist dabei definiert als derjenige Geldbetrag, den eine Person maximal zu zahlen bereit wäre, um das öffentliche Gut zu erhalten. Falls das öffentliche Gut für die Person einen negativen Nutzen aufweist, so entspricht der Nutzen dem Geldbetrag, mit dem eine betroffene Person entschädigt werden müsste, um bei Bereitstellung des Gutes auf dem ursprünglichen Nutzenniveau (englisch: utility) zu bleiben.

Während die Kosten der Bereitstellung öffentlicher Güter in vielen Fällen bekannt sind, liegt der Nutzen der Bereitstellung oft nicht auf der Hand. In jüngster Zeit wird jedoch zunehmend versucht, auch den Nutzen von staatlichen Maßnahmen in Geldeinheiten zu schätzen und explizit in die Kosten-Nutzen-Analyse politischer Entscheidungen einzubeziehen. Einerseits wird die Bedeutung von öffentlichen Gütern, etwa im Zusammenhang mit verkehrs- und energiepolitischen Entscheidungen in der Gesellschaft stärker wahrgenommen als früher. Andererseits ist bei Ämtern und Interessenvertretern auch eine zunehmende Tendenz festzustellen, ihre politischen Vorschläge mit Zahlen zu legitimieren. Vor diesem Hintergrund hat sich die Entwicklung von Methoden der Präferenzfassung zu einem der aktivsten Gebiete in der akademischen umwelt- und gesundheitsökonomischen Forschung entwickelt.

Den Nutzen öffentlicher Güter zu schätzen, ist aus zwei Gründen schwierig. Erstens sind Individuen nicht geübt darin, öffentliche Güter gegen Geld abzuwägen. Der Alltag gibt kaum Gelegenheit, Erfahrungen dabei zu sammeln und sich über die eigene Zahlungsbereitschaft für diese Güter Rechenschaft zu geben. Zweitens haben Bürger und Konsumenten – wenn sie sich über ihre individuelle Bewertung eines öffentlichen Gut im Klaren sind – wenig Anreiz dazu, ihre wahre Zahlungsbereitschaft im Rahmen von Befragungen offen zu legen. Mit diesen beiden grundlegenden Problemen der Präferenzfassung müssen Bewertungsmethoden fertig werden. Über die grundlegenden Probleme der Präferenzfassung hinaus sind die einschlägigen Studien jedoch auch oft mit vermeidbaren konzeptionellen und methodischen Problemen behaftet.

Für den Laien – auch den nicht spezialisierten Ökonomen – ist die Beurteilung der Qualität von angewandten Nutzenschätzungen sehr schwierig. Als Folge davon werden die Ergebnisse solcher Studien in den Tageszeitungen oft unkritisch aufgenommen und verbreitet. Natürlich ist manchem Leser bewusst, dass die Zahlen möglicherweise auch durch die Optik eines Auftraggebers geprägt sein können. Dies reicht aber nicht aus, um zwischen verlässlichen und weniger verlässlichen Zahlen zu unterscheiden. Abhilfe leisten kann hier nur eine verbessertes Verständnis der konkreten Möglichkeiten und Grenzen der empirischen Nutzenmessung in Forschungsinstituten, Ämtern und Redaktionen. Dieses Verständnis ist die Voraussetzung dafür, dass die Studien im politischen Betrieb

Prof. Dr. Peter Zweifel, 61, ist ordentlicher Professor für Volkswirtschaftslehre und Direktor am Sozialökonomischen Institut der Universität Zürich; Dr. Felix Schläpfer, 38, ist dort wissenschaftlicher Mitarbeiter.

* Die Autoren danken S. Banfi und M. Filippini (ETH Zürich) für wertvolle Kommentare.

¹ Der Begriff „öffentliches Gut“ wird hier in einem weiten Sinn verwendet. Er schließt auch Güter ein, für die die Eigenschaften der „Nicht-ausschließbarkeit“ und der „Nicht-rivalität“ nur unvollkommen zutreffen. Wichtige Beispiele von öffentlichen Gütern sind etwa die Bereitstellung der Infrastrukturen des öffentlichen und privaten Verkehrs oder auch staatliche Maßnahmen zur Verminderung externer Effekte.

ihrer Qualität entsprechend berücksichtigt werden. Außerdem kann aus einer kritischeren Rezeption ein Anreiz für eine hohe Qualität der Studien erwachsen.

Der vorliegende Text will einen Beitrag zu diesen Zielen leisten, indem er die wichtigsten Probleme im Zusammenhang mit der angewandten Nutzenmessung bei öffentlichen Gütern anhand von Beispielen darlegt. Zunächst werden grundlegende konzeptionelle Fallstricke erläutert. Dann werden häufige Probleme im Zusammenhang mit den einzelnen herkömmlichen Bewertungsansätzen beleuchtet. Schließlich werden alternative Ansätze vorgestellt, die auf der Analyse politischer Entscheidungen beruhen.

Konzeptionelle Fehler

Grundlegende konzeptionelle Fehler, die auch in großangelegten Auftragsstudien nicht selten vorkommen, betreffen

- (1) die Erhebung von unklar definierten Totalwerten statt der Bewertung realer Politik-Alternativen,
- (2) die Verwechslung von Nutzen (benefits) und Wertschöpfung und
- (3) die Überschätzung der Genauigkeit von Schätzungen, wenn die zugrundeliegenden Variablen mit systematischem Messfehler erhoben wurden.

Erhebung von irrelevanten Totalwerten

Eine Studie des Beratungsbüros ECONCEPT schätzte im Auftrag der Forstdirektion des (vormaligen) Bundesamtes für Umwelt, Wald und Landschaft (BUWAL) den „monetären Wert des Schweizer Waldes für die Erholung“.² Die Autoren wollten die Zahlungsbereitschaft für die Erhaltung des Schweizer Waldes als Erholungsquelle messen. Dabei stellt sich unmittelbar die Frage nach der Alternative. Was wäre die konkrete Alternative zur Erhaltung des Schweizer Waldes? Abholzen des ganzen Waldes und Ersetzen mit Kuhweiden oder Parkplätzen? Einzäunen und Abriegeln für Erholungsuchende? Einzäunen als Naturreservat oder als privates Jagdrevier? Die Alternative bleibt in der Studie unklar.

Dies ist typisch für Versuche, Totalwerte von Gütern zu erfassen, deren „Abschaffung“ gar nicht zur Disposition steht. Die relevante Handlungsalternative bildet fast ohne Ausnahme eine bestimmte – unter Umstän-

den geringfügige – Veränderungen dieser Güter. Dann also macht es aus ökonomischer Sicht keinen Sinn, Totalwerte von öffentlichen Gütern erfassen zu wollen;³ der Gegenstand der Bewertung ist letztlich nicht definiert. Und auch wenn der Alternativzustand klar definiert wäre, etwa als „Abholzen und Parkplätze bauen“, wäre die Bewertung nicht entscheidungsrelevant, denn die Ersetzung des gesamten Schweizer Waldes durch Parkplätze steht schlicht nicht zur Debatte. Studien dieser Art sind somit konzeptionell nicht sinnvoll angelegt und als Grundlage für politische Entscheidungen nicht hilfreich. Ein sinnvoller Bewertungsgegenstand im Zusammenhang mit dem Erholungswert des Schweizer Waldes wäre z.B. eine räumlich und zeitlich genau definierte Erhöhung des Anteils von naturnahen Wäldern oder eine bestimmte alternative Erschließung der Wälder für Erholungsuchende.

Verwechslung von Nutzen und Wertschöpfung

Aus Sicht der Kosten-Nutzen-Analyse sollen Infrastrukturen und Dienstleistungen, die vom Markt nicht angeboten werden, dann vom Staat bereitgestellt (oder mitfinanziert) werden, wenn deren volkswirtschaftlicher Nutzen höher ist als die Kosten. Die Höhe des Nutzens ist daher von wesentlichem Interesse für die Frage der Subventionswürdigkeit – von unrentablen lokalen Verkehrsverbindungen bis hin zu Flughäfen. Der Nutzen setzt sich in diesen Fällen zusammen aus der Zahlungsbereitschaft der Kunden für die Dienstleistung (die sich im Beispiel der Bahnlinien anhand der Nachfrage für Bahnbillette schätzen lässt) und der (politischen) Zahlungsbereitschaft der Steuerzahler für die Erhaltung des Bahnbetriebs, etwa aus regionalpolitischen Gründen. Statt diese beiden relevanten Nutzenkomponenten zu erfassen, schätzen einschlägige Studien aber oft die Wertschöpfung im Zusammenhang mit der fraglichen Dienstleistung.

Ein Beispiel dafür liefert eine Studie von INFRAS über den „Nutzen des öffentlichen Verkehrs“.⁴ Nur ist die Wertschöpfung zur Beurteilung des volkswirtschaftlichen Nutzens (und damit der Subventionswür-

³ Es ist auch nicht möglich, aus der Zahlungsbereitschaft für die Erhaltung des ganzen Waldes die Zahlungsbereitschaft für die Erhaltung von z.B. 99% des Waldes herzuleiten. Aus Totalwerten kann nicht auf Zahlungsbereitschaften für marginale Veränderungen geschlossen werden – bei einem (räumlich) heterogenen Gut wie dem Wald erst recht nicht, da nicht definiert ist, um welches Prozent es sich handelt.

⁴ M. Peter, M. Maibach, D. Sutter (INFRAS): Volkswirtschaftliche Bedeutung des öffentlichen Verkehrs in der Schweiz. Verband öffentlicher Verkehr, Bern 2004. Abgesehen von der Interpretation der Wertschöpfung als „Nutzen“ sind auch die „induzierten“, „multiplikativen“ und „katalytischen“ Effekte, die in der Studie veranschlagt werden, konzeptionell problematisch; vgl. R. Föllmi, U. Meister: Konsum hängt nicht von Bahn und Bus ab. Kritik an Studie zum Nutzen des öffentlichen Verkehrs, in: Neue Zürcher Zeitung vom 23.10.2004.

² Vgl. W. Ott, M. Baur: Der Monetäre Wert des Schweizer Waldes für die Erholung, Schlussbericht ECONCEPT, Forstdirektion BUWAL, Bern 2005. Die Autoren betonen sogar in der Zusammenfassung explizit, dass ihre Studie zum ersten Mal die Wertschätzung der Erholungsleistungen des gesamten Schweizer Waldes ermittelt (S. 6).

digkeit) einer öffentlichen Leistung schlichtweg nicht brauchbar. Dies ist leicht ersichtlich, wenn man sich überlegt, dass der öffentliche Verkehr auch dann eine hohe Wertschöpfung hätte, wenn er von keinem einzigen Kunden nachgefragt würde, für ihn keinerlei politische Zahlungsbereitschaft bestünde, und während er gewaltige externe Kosten durch Lärm und Abgase verursachte. Einen volkswirtschaftlichen Nettonutzen hätte ein solcher öffentlicher Verkehr offensichtlich nicht, ja man müsste ihn aus volkswirtschaftlicher Sicht sofort einstellen. Die gleiche Überlegung belegt auch die Fragwürdigkeit von Studien, die den volkswirtschaftlichen Nutzen von Flughäfen anhand von deren Wertschöpfung zu belegen versuchen, wie z.B. INFRAS.⁵ In beiden Beispielen – demjenigen des öffentlichen Verkehrs und der Flughäfen – besteht überdies das oben thematisierte Problem der Bewertung irrelevanter Gesamtnutzen anstelle von konkreten politischen Alternativen.

Überschätzung der Genauigkeit von Schätzungen

Oft werden monetäre Bewertungen anhand von Stichproben geschätzt, die aus einer Grundgesamtheit entnommen werden. Im Rahmen der Schätzung werden dann aufgrund der Stichprobe Vertrauensintervalle berechnet, die angeben, in welchem Bereich sich die geschätzten Populations-Parameter mit einer bestimmten Wahrscheinlichkeit befinden. Dabei besteht allerdings die Gefahr, die Genauigkeit der gewonnenen Aussagen nach wie vor zu überschätzen. Denn diese Vertrauensintervalle gelten nur, wenn es gelang, den Populationsparameter ohne Verzerrung zu schätzen. Während Verzerrungen durch nicht repräsentative Stichprobenziehung anhand der üblicherweise erhobenen sozioökonomischen Merkmale der Befragten leicht erkennbar sind, sind Verzerrungen durch systematische Messfehler in Standardstudien oft nicht nachweisbar.

Ein Beispiel möge dies illustrieren. Angenommen, Zahlungsbereitschaften für ein öffentliches Gut werden anhand einer Stichprobe mittels einer Befragung geschätzt. Es gibt gute Gründe dafür, dass die Befragten systematisch überhöhte Zahlungsbereitschaften angeben, wie etwa in der Studie von Schmitt et al.⁶ Statistische Vertrauensintervalle für den Mittelwert der

⁵ INFRAS: Volkswirtschaftliche Bedeutung des Flughafens Zürich – Auswirkungen verschiedener Szenarien, im Auftrag des Amtes für Verkehr Kanton Zürich (AFV), Zürich 2005.

⁶ M. Schmitt, F. Schläpfer, A. Roschewitz: Bewertung von Landschaftsveränderungen im Schweizer Mittelland aus Sicht der Bevölkerung. Eine Anwendung der Choice-Experiment-Methode; Forschungsanstalt für Wald, Schnee und Landschaft WSL, Birmensdorf 2005.

in der Befragung geäußerten Zahlungsbereitschaften – wie in der genannten Studie berechnet – sagen dann nur aus, mit welcher Wahrscheinlichkeit die geäußerte Zahlungsbereitschaft in einem bestimmten Bereich liegt, nicht aber die tatsächliche Zahlungsbereitschaft. Um beurteilen zu können, inwiefern es gelang, die gemessenen Größen mit dem Populations-Parameter in Deckung zu bringen, sind spezielle Validitätstests erforderlich. Sie geben Auskunft über die Einflüsse, die zu systematischen Abweichungen zwischen gemessenen und wahren Werten führen, sowie allfällige Gegenmaßnahmen oder Wertberichtigungen der Autoren.

Traditionelle Methoden der Präferenzfassung

Im Folgenden werden Probleme im Zusammenhang mit herkömmlichen Bewertungsansätzen für öffentliche Güter erläutert, die auf beobachteten Entscheidungen (im Gegensatz zu hypothetischen Entscheidungen in Befragungen) beruhen. Diese Methoden versuchen, aus privaten Konsumententscheidungen oder aus Marktpreisen auf die Wertschätzung für öffentliche Güter zu schließen. Dies ist dann möglich, wenn öffentliche Güter messbare Auswirkungen auf Konsumententscheidungen oder Marktpreise haben, oder wenn – im Fall des Reparatur-/Schadenskosten-Ansatzes – feststeht, dass der Wert eines beschädigten Gutes mindestens so hoch ist, dass die Wiederherstellung die beste Option darstellt.

Ein generelles Problem aller dieser Ansätze besteht darin, dass sie jeweils nur denjenigen Teil des Wertes öffentlicher Güter erfassen, der sich in Konsumententscheidungen oder Marktpreisen niederschlägt. So wird sich höchstens ein Teil der Zahlungsbereitschaft für einen Naturpark im Preis der umliegenden Immobilien (Hedonic-Pricing-Ansatz) oder im Besucherverhalten (Reisekosten-Ansatz) niederschlagen. Ebenso schlägt sich der durch Luftverschmutzung oder Lärmbelastung verursachte gesamtwirtschaftliche Schaden nur zum Teil in den Arztrechnungen der Betroffenen (Schadens-/Reparaturkosten-Ansatz) nieder. Den politischen Entscheidungsträgern fehlten damit immer noch wichtige Informationen über weitere Bestandteile des gesellschaftlichen (Grenz-)Nutzens von Maßnahmen, die für eine umfassende Kosten-Nutzen-Analyse relevant sind (vgl. die Erhebung irrelevanter Totalwerte).

Bewertung anhand von Marktpreisen

Wenn öffentliche Güter wie Umweltveränderungen einen direkten Einfluss auf kommerzielle Aktivitäten, wie etwa die Fischerei oder die Forstwirtschaft haben,

lässt sich der Nutzen oder Verlust anhand der Marktpreise der produzierten Güter abschätzen.⁷ Soweit solche Einflüsse auf (nicht wesentlich verzerrte) Marktpreise messbar sind, ist dieser Ansatz konzeptionell unproblematisch. Seine Anwendungsmöglichkeiten sind allerdings dadurch begrenzt, dass sehr viele öffentliche Güter keine direkt messbaren Auswirkungen auf die Preise kommerziell produzierter und gehandelter Güter haben.

Hedonic-Pricing-Methode

Der Hedonic-Pricing-Ansatz beruht darauf, dass die Nachfrage z.B. für das private Gut „Wohnen“ auch die Nachfrage für lokale öffentliche Güter widerspiegelt. Für eine ruhige Wohnlage wird auf dem Immobilienmarkt in barer Münze bezahlt. Der Wert der Ruhe kann anhand einer Stichprobe von Liegenschaften geschätzt werden, indem man den Einfluss lokaler Lärmimmissionen auf den Immobilienpreis mit dem statistischen Verfahren der multiplen Regression isoliert.

Bei geringfügigen lokalen Unterschieden des öffentlichen Gutes ist die Interpretation der Preisunterschiede als Zahlungsbereitschaften konzeptionell unproblematisch. Probleme ergeben sich allerdings daraus, dass nicht jede Beziehung, die in einer multiplen Regressionsgleichung gefunden wird, kausal interpretiert werden kann. Ein interessantes Beispiel liefert eine neuere Studie der Eidgenössischen Technischen Hochschule (ETH) Zürich, die ergab, dass Mobilfunkantennen einen signifikant negativen Effekt auf den Preis umliegender Immobilien haben.⁸ Für eine negative Beziehung zwischen Antennen-Nähe und Immobilienpreise ist aber auch eine alternative oder ergänzende Erklärung denkbar, bei der der Immobilienpreis die Ursache und der Standort der Antenne den Effekt bilden (umgekehrte Kausalität). Es genügt, anzunehmen, dass die Mobilfunk-Betreiber die Standorte für Antennen bevorzugt in Quartieren mit niedrigen Wohnungspreisen wählen, weil sie dort weniger Widerstand aus der Bevölkerung erwarten. Die beobachteten Korrelationen wären dann Ausdruck davon, dass die Immobilienpreise die Standorte der Antennen beeinflussen und nicht umgekehrt die Antennen die Bewertung der Immobilien.

⁷ Bei kleinen Veränderungen ist der Marktpreis eine gute Annäherung an die Bewertung einer zusätzlichen Einheit des Gutes, bei größeren Veränderungen kann der Wert des öffentlichen Gutes anhand der Konsumentenrente und der Produzentenrente geschätzt werden.

⁸ S. Banfi, M. Filippini, A. Horehájová: Hedonic price functions for Zürich and Lugano with special focus on electrosmog. Paper presented at the Monte Verità Conference on Sustainable Resource Use and Economic Dynamics (SURED), Mai 2006.

Reisekosten-Ansatz

Beim Reisekosten-Ansatz wird der Zusammenhang zwischen Reisekosten und Häufigkeit des Besuchs von räumlich begrenzten öffentlichen Gütern wie Naherholungsgebieten analysiert. Die Preisvariation, die für die Identifikation der Nachfrage benötigt wird, ergibt sich aus den individuellen Reisedistanzen und den damit verbundenen Transport- und Zeitkosten. Der Wert des öffentlichen Gutes wird mit Hilfe der Konsumentenrente erfasst. Die Konsumentenrente summiert die Differenzen zwischen der maximalen Zahlungsbereitschaft und den Reisekosten, die ein Besucher für jeden einzelnen Besuch auf sich nimmt.

Ein generelles Problem dieses Ansatzes besteht darin, die Substitutions- bzw. Ausweichmöglichkeiten angemessen zu erfassen und zu berücksichtigen. Schelbert et al. schätzten mit dem Reisekosten-Ansatz die Zahlungsbereitschaft für das Naherholungsgebiet A (Zürichberg-/Adlisbergwald).⁹ Angenommen, es bestehe außer A ein alternatives Erholungsgebiet B, und eine Besucherin sei indifferent zwischen A und B. Sie wird dann bei einem Verlust von A ohne Nutzeneinbuße das Erholungsgebiet B aufsuchen. Die Nachfragekurve wird jedoch eine Nutzeneinbuße anzeigen. Dieser Fehler lässt sich nur vermeiden, wenn die Reisekosten derselben Person für das Aufsuchen sowohl von A und B (und gegebenenfalls weiterer Alternativen C, D,...) bekannt sind. Diese Bedingung ist kaum je erfüllt.

Zuweilen werden die Kosten des Besuches fälschlicherweise direkt als untere Grenze für den Wert des öffentlichen Gutes aufgefasst. Dieser Ansatz, der auch in der bereits erwähnten ECONCEPT-Studie verwendet wurde, verwechselt das private Gut „Erholung“ mit dem öffentlichen Gut „Erholungsgebiet“. Die Kosten der Anreise selbst können wohl als Mindestmaß für den Wert des (privaten) Gutes „Erholungsausflug“ interpretiert werden, aber nicht als Mindestmaß für die Bewertung des Erholungsgebietes als öffentliches Gut. Die Zahlungsbereitschaft für das Naherholungsgebiet als öffentliches Gut (für die Naherholung) ergibt sich als Differenz zwischen der Zahlungsbereitschaft für die Anreise und den Kosten der Anreise.

Schadens-/Reparaturkosten-Ansatz

Der Schadens- oder Reparaturkosten-Ansatz besteht darin, jene Kosten zu erfassen, die auf dem

⁹ H. Schelbert, T. Lang, I. Buse, J. Henzmann, R. Maggi, R. Iten, C. Nielsen: Wertvolle Umwelt. Ein wirtschaftswissenschaftlicher Beitrag zur Umwelteinschätzung in Stadt und Agglomeration Zürich, Zürcher Kantonalbank, Zürich 1988.

Markt gezahlt werden müssen, um Schäden zu beheben, welche die Veränderung des betrachteten öffentlichen Gutes verursacht. Beispiele sind etwa zusätzliche Kosten der Trinkwasseraufbereitung wegen einer Erhöhung der Gewässerbelastung, oder Gesundheitskosten durch erhöhte Luftverschmutzung oder Lärmbelastung.

Die Schätzung von Wiederherstellungskosten stellt keine Bewertungsmethode im eigentlichen Sinne dar. Denn Sie erfasst nicht den ökonomischen Wert des Gutes, sondern nur die Reparaturkosten, welche der Eigentümer bereit ist, aufzuwenden. Reparaturaufwendungen, aber auch Aufwendungen zur Vermeidung von Schäden, bleiben jedoch in aller Regel hinter dem Wert des Gutes zurück. Sie können als Mindestwert des öffentlichen Gutes aufgefasst werden – als Mindestwert deshalb, weil z.B. eine erhöhte Luftverschmutzung nicht nur gesundheitliche Schäden verursacht und weil sich die gesundheitlichen Schäden nur unvollständig in „Reparaturkosten“ (Aufwendungen für ärztliche Leistungen) manifestieren.

Zahlungsbereitschafts-Befragungen

Die traditionellen Bewertungsansätze können – wie bereits erwähnt – nur Teile des Wertes von gewissen Gütern erfassen. Diese Einschränkungen fallen bei Befragungen grundsätzlich weg. Man kann problemlos auch nach der gesamten Zahlungsbereitschaft für die Bereitstellung jeglicher öffentlicher Güter fragen, vom Wert einer um 10 Dezibel verminderten Lärmbelastung an der Wohnstraße bis zur Zahlungsbereitschaft für einen besseren Schutz der Wale. Aus diesem Grund hat das Interesse an Befragungsmethoden in den letzten zwei Jahrzehnten enorm zugenommen. Weitaus die meisten Studien zur Bewertung öffentlicher Güter verwenden heute Methoden, die auf Befragungen beruhen.

Die Bereitstellung öffentlicher Güter setzt typischerweise einen politischen Entscheid über die Finanzierung voraus. Konsistent mit diesem politischen Entscheidungsrahmen namentlich einer direkten Demokratie sind Erhebungen zur Zahlungsbereitschaft dann, wenn sie die Form einer Abstimmungsfrage betreffend die Finanzierung des öffentlichen Gutes haben. Das zu finanzierende Gut wird dabei detailliert und insbesondere auch hinsichtlich der unterstellten eigentumsrechtlichen Ausgangslage realistisch beschrieben und die Verteilung der Kosten klar definiert.¹⁰ Die Abstimmungsfrage kann dabei auf mindestens drei Arten gestellt werden:

mungsfrage kann dabei auf mindestens drei Arten gestellt werden:

- (1) Als direkte Frage nach der maximalen sich ergebenden Erhöhung der Steuerbelastung, bei welcher der Befragte der Vorlage gerade noch zustimmen würde (Open-Ended Contingent Valuation).
- (2) Als Entscheidungsfrage (Ja/Nein) bei einer vorgegebenen Erhöhung der Steuerlast, die variiert wird, um aus der Verteilung der Ja/Nein-Antworten bei verschiedenen Werten die Verteilung der Zahlungsbereitschaften zu bestimmen (Dichotomous-Choice Contingent Valuation).
- (3) Als Entscheidung zwischen zwei oder mehreren alternativen Politiken, die sich hinsichtlich mehrerer Merkmale unterscheiden, wobei eines der Merkmale eine mit der Politik verbundene Veränderung der Steuerlast ist.¹¹ Die Merkmals-Kombinationen der zur Auswahl gestellten Alternativen werden bei der Befragung variiert. Die Zahlungsbereitschaften für die Merkmale des öffentlichen Gutes werden ermittelt, indem statistisch analysiert wird, wie stark diese Merkmale – relativ zur Steuerlast – in den Entscheidungen berücksichtigt wurden (Conjoint Analysis, Discrete Choice Experiment).

Während die universelle Anwendbarkeit eine eindeutige Stärke dieser Befragungsansätze ist, haben sie zwei wichtige Schwachpunkte: (A) Wie kann dem Problem strategischer Antworten begegnet werden? (B) Werden die Befragten möglicherweise überfordert?

Strategisches Antworten

Das erste der oben genannten Frageformate lädt zu strategischen Antworten geradezu ein. Aber auch das zweite Format lässt strategische Antworten zu. Da die Fragen jeweils hypothetische Steuerkosten nennen, die nichts mit den effektiven Steuerlasten zu tun haben, können strategisch denkende Befragte zu einer vorgeschlagenen Politik auch dann ja sagen, wenn die Steuerkosten für sie unakzeptabel hoch sind. Die Contingent Valuation kann deshalb zu einer krassen Überschätzung der Zahlungsbereitschaften führen. Beim dritten Ansatz besteht die Gefahr strategischer Antworten zwar grundsätzlich auch, doch die strategischen Möglichkeiten sind für die Befragten weniger leicht erkennbar.

¹¹ Sowohl beim Format (2) als auch beim Format (3) wird den Befragten eine Steuererhöhung in rein hypothetischem Ausmaß vorgelegt, eine Steuererhöhung also, die nichts zu tun hat mit den effektiv zu erwartenden individuellen Kosten der Bereitstellung des Gutes (vgl. die Diskussion zum Problem des strategischen Antwortens).

¹⁰ Die Interpretation der Ergebnisse von Befragungen, die von diesen Vorgaben abweichen, sind mit Unsicherheiten befrachtet, die noch weit über die im Folgenden diskutierten Probleme hinausgehen.

Dem Problem strategischer Antworten kann begegnet werden, indem als (hypothetische) Kosten der Vorlagen realistische prozentuale Steuererhöhungen in die Abstimmungsfragen eingesetzt werden. Im Gegensatz zum üblichen Frageformat mit seinen gemäß einem experimentellen Design zufällig eingesetzten Geldbeträgen, die strategisches Antworten bei Ja/Nein-Fragen erst ermöglichen, ordnet man den Befragten individuell realistische und damit auch glaubwürdige Steuerkosten zu. Am einfachsten geschieht dies durch eine Formulierung der Kosten in Prozenten der aktuellen Steuerlast. Die Zahlungsbereitschaften lassen sich bei diesem modifizierten Frageformat anhand der Verteilung der Zahlungsbereitschaften in Steuerprozenten und der Verteilung der Einkommen berechnen.¹²

Kognitive Begrenztheit und vereinfachende Heuristiken

Bei Entscheidungen über private Güter helfen die Marktpreise vergleichbarer Produkte, den Kaufentscheid zu treffen: So behelfen sich wohl die meisten Leute bei der Auswahl einer Flasche Wein gelegentlich mit der Annahme, dass der Marktpreis etwas über die Qualität des Produkts aussagt. Diese Annahme ist sogar berechtigt, denn Marktpreise werden durch die marginalen Nachfrager bestimmt, deren Zahlungsbereitschaft gerade noch den Preis deckt. Solche Käufer dürften zu den besser informierten gehören. Da andere Leute besser informiert sind, ist der Preis der Flasche als Ausdruck der marginalen Zahlungsbereitschaft für ahnungslose Käufer durchaus informativ. Auch nur verbal geäußerte Zahlungsbereitschaften für vertraute private Güter wie Wohnungen sagen deshalb die Zahlungsbereitschaften sehr verlässlich voraus, da sich die Abwägung von Kosten und Nutzen auf vereinfachende Heuristiken, wie den Vergleich mit anderen Wohnungen ähnlicher Art oder Preislage, stützen kann.

Bei Entscheidungen über öffentliche Güter in Abstimmungen oder Umfragen sind keine Gleichgewichtspreise als Qualitätsindikatoren vorhanden. Für die Bewertung öffentlicher Güter anhand von Befragungen hat dieser Befund die Konsequenz, dass die Antworten die persönlichen Werte und Interessen nur dann abbilden, wenn die Befragten verlässliche „Shortcuts“ zur Verfügung haben, auf welche sich vereinfachende Heuristiken stützen können. Diese Voraussetzung ist erfüllt bei Befragungen zu Themen,

die in derselben Form bereits Gegenstand medialer politischer Debatten in den Medien waren. Nicht erfüllt ist die Voraussetzung hingegen häufig bei konventionellen Zahlungsbereitschafts-Befragungen über öffentliche Güter; jene haben denn auch die üblichen Validitäts-Tests kaum je überstanden.¹³

Ein möglicher Ausweg aus dem Dilemma besteht darin, das öffentliche Gut „als privates Gut zu verkaufen“. So schätzen Zweifel und Schneider den Wert des an sich öffentlichen Guts „Reaktorsicherheit“, indem sie den Teilnehmern des Discrete Choice Experiments klar machten, dass eine ausgebaute Haftpflichtversicherung zu höheren Strompreisen führen müsste und den Strompreis als Attribut ins Experiment aufnehmen.¹⁴ Die individuell zu treffende Entscheidung wird damit einfacher; dafür werden ähnlich wie bei den traditionellen Ansätzen nur Teilwerte des öffentlichen Gutes, in diesem Fall die versicherten materiellen Verluste, bewertet.¹⁵

Eine zweite Lösung besteht darin, die Stellungnahmen von Meinungsführern, politischen Parteien und Verbänden in die Beschreibung der Entscheidungssituation aufzunehmen. Diese Stellungnahmen informieren den Stimmbürger gleichsam darüber, ob die ihm nahestehende Partei „kaufen“ oder „nicht kaufen“ empfiehlt. Wenn die Stimmbürger die Übereinstimmung ihrer eigenen Interessen mit denen der Partei- und Verbandsinteressen abschätzen können, sind sie in der Lage, Entscheidungen zu treffen, die in ihrem Interesse liegen, auch wenn sie die Einzelheiten der Vorlagen nicht kennen.¹⁶ Fehlen hingegen solche Signale oder „Shortcuts“, wie die Politikwissenschaftler sagen, so sind sie bei Fragen über öffentliche Güter sehr orientierungslos, fallen auf „Mogelpackungen“ herein und nennen Zahlen, die nicht als konsistente

¹² M. Schmitt et al., a.a.O.; N. E. Flores, A. Strong: Cost credibility and the stated preference analysis of public goods, in: *Resource and Energy Economics*, Vol. 29 (2007), Nr. 3, S. 195-205.

¹³ Eine Befragung zur Zahlungsbereitschaft für den Schutz der Landschaft im Zürcher Weinland ergab 1996 einen Medianwert von 240 Franken pro Jahr und Person. Drei Monate später wurde eine Vorlage über die Erhöhung des kantonalen Budgets für Natur und Heimatschutz im Weinland von fast allen Testgemeinden der Befragung abgelehnt. Die Kosten der Vorlage betragen rund 20 Franken pro Person und Jahr, kantonsweit wurde die Vorlage angenommen; vgl. F. Schläpfer, N. D. Hanley: Contingent valuation and collective choice, in: *Kyklos*, 59, 2006, S. 115-135.

¹⁴ P. Zweifel, Y. Schneider: Marginale Zahlungsbereitschaft für eine erhöhte Internalisierung des Risikos von Kernkraftwerken, Bundesamt für Energie, Bern 2002.

¹⁵ Siehe auch Y. Schneider, P. Zweifel: How much internalization of nuclear risk through liability insurance?, in: *Journal of Risk and Uncertainty*, Vol. 29 (2004), S. 219-240.

¹⁶ A. Lupia: Shortcuts versus encyclopedias – information and voting behavior in California insurance reform elections, in: *American Political Science Review*, Vol. 88 (1994), Nr. 1, S. 63-76.

Zahlungsbereitschaften interpretiert werden können.¹⁷ Diese Alternative soll im nächsten Abschnitt näher erläutert werden.

Befragungen mit Angabe von Parteiempfehlungen

Für die Teilnehmer von Zahlungsbereitschafts-Befragungen über öffentliche Güter kann die Aufgabe dadurch vereinfacht werden, dass man ihnen vorgängig eingeholte Parteiempfehlungen zur Verfügung stellt. Damit können sich die Entscheidungen wie in einer Volksabstimmung auf die Empfehlungen vertrauter Organisationen mit bekannten politischen Interessen abstützen. Eine erste Studie, die diesen Ansatz verfolgte, ergab Zahlungsbereitschaften, die gerade noch 10% bis 50% der Werte ausmachten, die aus der Befragung ohne solche Empfehlungen resultierten.¹⁸ Angesichts der bekannten Tendenz zur Überschätzung von Zahlungsbereitschaften in Befragungen über öffentliche Güter deutet dieses Resultat darauf hin, dass man mit diesem Ansatz näher an jene Werte herankommt, welche die realen Abstimmungsentscheide steuern als Antworten „aus dem hohlen Bauch“.

Analyse von Volksabstimmungen über die Finanzierung öffentlicher Güter

Reale Abstimmungsentscheide betreffend Finanzierungsvorlagen für öffentliche Güter liefern einerseits einen Validitätstest.¹⁹ Andererseits lassen sie, wenn Ergebnisse für verschiedene Einkommensgruppen vorliegen, auch Rückschlüsse auf die Verteilung der Zahlungsbereitschaften zu, die für eine akzeptable Verteilung der Kosten von öffentlichen Gütern von Interesse ist. Der Rückschluss auf die Präferenzen wird vom Umstand begünstigt, dass der gesellschaftlich-politische Entscheidungsrahmen klar definiert und die Kostenfolgen real sind. Überdies findet eine freie, öffentliche Debatte statt, welche die Wähler mit einer Vielfalt von relevanten Informationen konfrontiert. Und schließlich können die Wähler anhand von „Parolen“ auch dann relativ zuverlässig in ihrem Interesse abstimmen, wenn sie den Inhalt der Vorlage nur rudimentär kennen. Bis-

herige ökonomische Analysen des individuellen oder aggregierten Abstimmungsverhaltens haben konsistente Muster im Abstimmungsverhalten verschiedener Wählergruppen ergeben. Daten aus Volksabstimmungen über spezifische Finanzierungsvorlagen stellen deshalb eine wertvolle, wenn auch beschränkte, Quelle von Informationen über Präferenzen für öffentliche Güter dar.

Schlussfolgerungen

Häufig auftretende Fehler in der Bewertung öffentlicher Güter sind die Erfassung von nicht entscheidungsrelevanten totalen statt entscheidungsrelevanten marginalen Werten, die Verwechslung von Nutzen (Benefits) und Wertschöpfung und die Überschätzung der Genauigkeit geschätzter Parameter.

Traditionelle Bewertungsansätze, die auf beobachteten Präferenzen beruhen, sind in der Anwendung begrenzt, da die meisten öffentlichen Güter Wertkomponenten aufweisen, die mit diesen Ansätzen nicht erfasst werden können. Teilwerte sind jedoch oft wenig hilfreich für politische Entscheidungen, welche eine umfassende Kosten-Nutzenanalyse voraussetzen.

Universell anwendbar sind nur Befragungen, die aber ihrerseits zwei wesentliche Nachteile aufweisen. Der erste sind strategische Antworten. Zu ihrer Vermeidung kann die Bewertungsfrage als Entscheidungsfrage (Ja/Nein) gestellt werden, wobei die genannten Kosten des Gutes dann wie bei Volksabstimmungen – und anders als bei herkömmlichen Umfragen – als realistische Steuererhöhung des Teilnehmers ausgedrückt werden. Das zweite Problem ist die mangelnde Fähigkeit vieler Befragter, konsistente Entscheidungen ohne Rückgriff auf vereinfachende Heuristiken – anhand von Preisen, Debatten in den Medien oder Partei-Empfehlungen – zu treffen. Eine mögliche Lösung besteht darin, den Befragten Parteiempfehlungen zur Kenntnis zu bringen und so eine Orientierungshilfe zu leisten.

Was die Marktentscheidungen bei privaten Gütern sind, sind Volksabstimmungen über Finanzierungsvorlagen bei öffentlichen Gütern. Aufgrund der begrenzten strategischen Antwortmöglichkeiten und des offenen, demokratisch legitimierten Wettbewerbs der Argumente ist die bindende Volksabstimmung, welche die Bewertung mit der Entscheidung verbindet, eine gute Methode zur Erfassung der Präferenzen für öffentliche Güter.

¹⁷ M. Schmitt et al., a.a.O.; F. Schläpfer, R. Soliva: The market for public lemons – a field experiment on benefit uncertainty and the political market mechanism. Paper presented at the 15th Annual Conference of the European Association of Environmental and Resource Economists, Thessaloniki, Griechenland, Juni 2007.

¹⁸ F. Schläpfer, M. Schmitt, A. Roschewitz: Competitive politics, simplified heuristics, and preferences for public goods, in: Ecological Economics, 2008 in Druck.

¹⁹ Vgl. Fußnote 13.