

Joachim Weimann*

Die Rolle von Verhaltensökonomik und experimenteller Forschung in Wirtschaftswissenschaft und Politikberatung

DOI 10.1515/pwp-2015-0017

Zusammenfassung: Verhaltensökonomische Ansätze und experimentelle Methoden sind in den vergangenen dreißig Jahren zu festen Bestandteilen der ökonomischen Methodik geworden. Joachim Weimann zeichnet nach, wie es zu dieser Erweiterung gekommen ist und wie sich die neuen Konzepte und Methoden mit dem paretianischen Programm des „Rational choice“ vertragen. Auf der Grundlage dieser Analyse schlägt er eine Integration aller Methoden in ein Konzept evidenzbasierter Politikberatung vor. In diesem Konzept ordnen die Dimensionen „Kausalität“, „externe Validität“ und „Breite der Anwendbarkeit“ die Methoden. Dabei zeigt sich, dass die Verhaltensökonomik und die experimentelle Methode hervorragend geeignet sind, komplementär zur neoklassischen Theorie und zur empirischen Forschung eingesetzt zu werden.

JEL-Klassifikation: A11, A12, B10, B13, C10, C18, C90, C91, C93

Schlüsselwörter: Verhaltensökonomie, experimentelle Wirtschaftsforschung, evidenzbasierte Politikberatung, neoklassische Wirtschaftstheorie, psychologische Forschung, Dogmengeschichte.

1 Ungeklärte Verhältnisse

„[The] status of experimental economics remains controversial“, schrieb Robert Sugden vor zehn Jahren in der Einleitung zu einem Tagungsband über die Rolle von Experimenten in den Wirtschaftswissenschaften (Sugden 2005, S. 177). An dieser Einschätzung hat sich seither nicht viel geändert, obwohl die ökonomische Wissenschaft in dieser Zeit eine dynamische Entwicklung durchlaufen hat. Dass insbesondere im deutschsprachigen Raum die Stellung der experimentellen Forschung weitgehend un-

geklärt ist, zeigt sich in exemplarischer Deutlichkeit daran, dass sie in der Politikberatung – immerhin einem der wichtigsten Tätigkeitsbereiche von Ökonomen – praktisch nicht vorkommt. Man sucht experimentelle Arbeiten in den Literaturverzeichnissen der Gutachten des Sachverständigenrates vergeblich, und an den Wirtschaftsforschungsinstituten sind experimentelle Abteilungen eine große Seltenheit.

Zwar haben sich die meisten deutschsprachigen Ökonomen damit abgefunden, dass ihre Disziplin eben auch einen experimentellen Ableger hat. Doch die Mehrheit erwartet offensichtlich nicht, dass aus diesem Bereich Erkenntnisse kommen könnten, die für die praktische Politikberatung bedeutsam sind. Nach den Gründen gefragt, wird schnell auf den artifiziellen Charakter von Experimenten verwiesen und deren externe Validität in Zweifel gezogen. Das ist insofern eine erstaunliche Begründung, als die allermeisten Ökonomen keinerlei Problem damit haben, höchst abstrakte (und damit artifizielle) Modelle zur Grundlage ihrer Aussagen zu machen, deren externe Validität ebenfalls mit einem gewissen Recht bezweifelt werden kann. Beispielsweise haben Finanzwissenschaftler in aller Regel keine Scheu, wirtschaftspolitische Ratschläge auf der Grundlage der Optimalsteuertheorie zu geben, ohne dabei jemals überprüft zu haben, ob die Zensiten tatsächlich auf die Besteuerung so reagieren, wie es die Theorie voraussetzt.

Eine intensive akademische Diskussion darüber, welche Bedingungen wissenschaftliche Methoden erfüllen müssen, damit sie als Grundlage für wirtschaftspolitische Beratung dienen können, hat zumindest im deutschsprachigen Raum bis heute nicht stattgefunden. Sie wird auch in der gegenwärtig intensiv geführten Diskussion über evidenzbasierte Politikberatung¹ nur partiell nachgeholt. Die immer stärkere Hinwendung zu einer strikt evidenzbasierten Politikberatung ist ganz sicher notwendig, aber die methodische Diskussion darüber, welche Ansätze

*Kontaktperson: Joachim Weimann, Lehrstuhl für Wirtschaftspolitik, Otto-von-Guericke-Universität Magdeburg, E-Mail: joachim.weimann@ww.uni-magdeburg.de

¹ Vgl. die Beiträge von Bookmann et al. (2014) und Christoph Schmidt (2014) in dieser Zeitschrift.

über die empirische Evaluationsforschung hinaus geeignet sind, wissenschaftliche Evidenz zu liefern, bleibt bisher aus.

Das alles steht in einem gewissen Widerspruch dazu, dass die Verhaltensökonomik in den vergangenen dreißig Jahren einen enormen Bedeutungszuwachs erfahren hat. Dieser Wandel hat sich nicht im Verborgenen abgespielt, sondern findet seit wenigstens zwanzig Jahren in den ökonomischen Top-Journals statt. In diesen finden sich immer häufiger Beiträge, deren Verfasser explizit nicht neoklassische, sondern verhaltensbasierte Modelle benutzen und auf psychologischen Erkenntnissen aufbauen.² Experimente spielen in dieser Literatur eine wichtige Rolle, aber es wäre falsch, deshalb die experimentelle Wirtschaftsforschung mit der Verhaltensökonomik gleichzusetzen. Das Verhältnis ist komplexer, als es auf den ersten Blick erscheint. Die experimentelle Forschung hat in den Wirtschaftswissenschaften lange ohne jeden Bezug zur Psychologie existieren können und war in einem gewissen Sinne in die Neoklassik integriert. Gleichwohl hat sie den Weg dafür geebnet, dass sich verhaltensökonomische (psychologische) Ansätze in der Beletage der ökonomischen Forschung durchsetzen konnten.

In Deutschland hat die wissenschaftliche Aufwertung der Verhaltensökonomik nicht mit sich gebracht, dass jene Ökonomen, die sich in der Politikberatung engagieren, die Möglichkeiten dieses neuen Ansatzes ernsthaft aufgegriffen haben. Dafür hat sich freilich die Politik der Verhaltensökonomik zugewendet und dem aus ihr hervorgegangenen „Nudging“³. In der deutschen Öffentlichkeit hat die Ausschreibung von entsprechenden Mitarbeiterstellen im Kanzleramt einiges Aufsehen erregt. Die Entscheidung der Politik, die Beratung von Verhaltensökonomien in Anspruch zu nehmen, ist auch in anderen Ländern zu beobachten. Sie ist weitgehend unabhängig von den sonst üblichen ökonomischen Beratungsprozessen und den sie tragenden Institutionen.

Die Grundlagen der Verhaltensökonomik sind in der psychologischen Forschung zu suchen. Dass die Politik die Psychologie für sich entdeckt hat, ist nicht verwunderlich. Schließlich verspricht das Nudging dem Politiker eine ganze Reihe neuer Werkzeuge, mit denen sich Politik ma-

chen lässt. Von ökonomischen Beratern wurden diese Instrumente bisher mit keinem Wort erwähnt. Es wäre allerdings falsch zu glauben, dies sei eine Sache zwischen Politik und Psychologie, welche die Ökonomen nichts angehe. Es ist kein Zufall, dass das Nudging unter dem Oberbegriff „Verhaltensökonomik“ firmiert und nicht unter „Verhaltenspsychologie“. Der Einfluss psychologischer Forschung auf die Ökonomik und die Zusammenarbeit zwischen Psychologen und Ökonomen haben in den vergangenen dreißig Jahren stetig an Bedeutung gewonnen. Die experimentelle Forschung dient dabei gleichsam als Brücke zwischen den Disziplinen.

Eine Gemeinsamkeit von experimenteller Forschung und Verhaltensökonomik besteht darin, dass in beiden Fällen die Stellung innerhalb der ökonomischen Disziplin nicht abschließend geklärt ist. Wie steht es um das Verhältnis von Ökonomie und Psychologie? Wie spielen psychologische Erkenntnisse über die Funktion von „Verzerrungen und Heuristiken“ mit der Theorie rationalen Handelns zusammen, die nach wie vor die Grundlage der meisten ökonomischen Modelle ist? Wie lassen sich verhaltensökonomische Erkenntnisse in die *ökonomische* Politikberatung integrieren? Oder haben sie dort nichts zu suchen?

Ganz gleich, wie man zur experimentellen Wirtschaftsforschung oder zur Verhaltensökonomik steht, eine grundlegende Klärung der Verhältnisse dürfte als hilfreich empfunden werden. International wird diese Diskussion intensiv geführt, aber sie ist noch jung, und man kann noch nicht davon sprechen, dass sie zu einem Konsens geführt habe.⁴ In einem aktuellen Beitrag schlägt Chetty (2015) einen pragmatischen Umgang mit der Verhaltensökonomik vor, bei dem diese gleichberechtigt zur neoklassischen Standardtheorie des „Rational choice“, also der rationalen Entscheidung, in die Politikberatung integriert werden soll. Auf die Rolle der ökonomischen experimentellen Forschung und ihre Einordnung in den Methodkanon geht Chetty dabei nicht explizit ein, so dass auch in seinem Vorschlag dieses Instrument in seiner Bedeutung ungeklärt bleibt.

Hier gilt es nun einerseits den Stand der Diskussion zu erläutern sowie die experimentelle Forschung und die Verhaltensökonomik so in den Kanon der wirtschaftswissenschaftlichen Instrumente einzuordnen, dass es möglich wird, diese noch jungen Instrumente vorteilhaft in der evidenzbasierten Politikberatung einzusetzen. Zu-

² Ein Überblick findet sich in DellaVigna (2009).

³ Ein „Nudge“ ist der Versuch, eine Verhaltensänderung durch leichtes Anschubsen zu erreichen, ohne bestimmte Verhaltensweisen bindend vorzuschreiben. Ein Beispiel dafür ist in der Kantine das Platzieren gesunder Speisen auf Augenhöhe. Damit wird erreicht, dass die Kantinenbesucher häufiger zu diesen Speisen greifen als sonst. Vgl. Thaler und Sunstein (2008), Bruttel und Stolley (2014), Weimann (2014).

⁴ Rabin (1999) und DellaVigna (2009) liefern einen Überblick über die Zusammenarbeit von Psychologie und Ökonomie, Madrian (2014) analysiert die Bedeutung der Verhaltensökonomik für die Wirtschaftspolitik.

nächst sei nachgezeichnet, wie es zu dem heutigen Methodenpluralismus gekommen ist. Dabei wird die entsprechende dogmengeschichtliche Entwicklung in groben Zügen dargestellt und die Reaktionen beleuchtet, die die jeweiligen Entwicklungen provoziert haben. Anschließend versuche ich, das Methodenknäuel, das sich entwickelt hat, zu entwirren und eine sinnvolle Einordnung vorzunehmen.

2 Von Paretos Programm zur Verhaltensökonomik

Die Wirtschaftswissenschaft ist eine Verhaltenswissenschaft. Ihr Erkenntnisgegenstand sind die Wahlhandlungen, die Menschen im Zusammenhang mit der Verwendung knapper Ressourcen treffen, für die es alternative Verwendungen gibt. Diese Definition der Disziplin ist allgemein akzeptiert und sehr alt. Dennoch ist es notwendig, sie an den Anfang dieses Kapitels zu stellen: Wenn man dieses disziplinäre Verständnis voraussetzt, sollte es eigentlich naheliegen, dass ökonomische Forschung auf das Engste mit der Psychologie verbunden ist. Dass dies nicht der Fall ist, dass die Vertreter der Ökonomik es im Gegenteil verstanden haben, ein volles Jahrhundert ohne jeden Bezug zu psychologischen Erkenntnissen zu arbeiten und dabei erfolgreich zu forschen, bedarf der Erklärung.

2.1 Paretos Programm der Entpsychologisierung

Die Vertreter der neoklassischen Wirtschaftstheorie haben sich zu keinem Zeitpunkt darum bemüht, eine psychologische Begründung für ihre Verhaltenstheorie zu finden. Jedenfalls gilt dies für die Zeit nach Pareto.⁵ Das paretianische Forschungsprogramm ließ sich ohne den Rückgriff auf psychologische Erkenntnisse bearbeiten – und das war kein Zufall. Paretos Ziel war es, die Ökonomik zu einer eigenständigen Wissenschaft zu machen – und dafür war es notwendig, sie gegen die Philosophie und die Psychologie abzugrenzen.⁶ Möglich wurde dies durch die Idee,

⁵ Darauf weisen Bruni und Sugden (2007) in ihrem überaus lesenswerten Artikel hin. Dort wird der „Paretian turn“ beschrieben, der zum Ziel hatte, die Ökonomik von der Psychologie zu lösen, um ihr den Status einer eigenständigen Wissenschaft zu verleihen. Ebenda, S. 154 ff.

⁶ Ebenda.

dass der Erkenntnisgegenstand ökonomischer Forschung eben solche Entscheidungen sind, die ausschließlich auf der Grundlage rationaler Kalküle getroffen werden.

Um dabei ohne psychologische Erkenntnis auszukommen, bedurfte es zweier wichtiger Voraussetzungen. Erstens mussten die Entscheider über konsistente Präferenzen verfügen, das heißt, die Bewertung von alternativen Güterbündeln musste auf der Grundlage einer ordinalen Bewertung dieser Güterbündel erfolgen, die vollständig und transitiv und damit eine Ordnung war. Zweitens musste diese Ordnung exogen, stabil und stationär sein. Exogen musste sie in dem Sinne sein, dass die Präferenzen des Entscheiders unabhängig vom Kontext einer Entscheidung als vorgegeben gelten durften.⁷ Stabil musste sie sein in dem Sinne, dass sich die Präferenzen durch Entscheidungsvorgänge und -prozesse nicht veränderten.⁸ Und stationär musste sie sein in dem Sinne, dass sie im Zeitablauf konstant blieb. Wenn diese Voraussetzungen erfüllt waren, ließ sich jeder Auswahlvorgang als ein Optimierungskalkül abbilden, für dessen Beschreibung psychologische Kategorien vollständig entbehrlich wurden, Mathematik aber essentiell war.

Mit dem „Paretian turn“, wie ihn Bruni und Sugden (2007) nennen, eröffnete sich die Möglichkeit, die Mathematik als Verhaltenstheorie zu verwenden. Entscheidungen waren gleichbedeutend mit Optimierung; sie ließen sich deshalb immer auch mathematisch beschreiben und analysieren. Bruni und Sugden weisen aber auch darauf hin, dass Pareto diese methodische Ausrichtung mit einer wesentlichen Einschränkung versehen hatte, die seine Nachfolger nicht weiter beachteten. Diese Einschränkung bestand darin, nur solche Entscheidungen mit Hilfe der Rational-choice-Theorie abzubilden, die wiederholt auf Märkten stattfanden, und bei denen die Entscheider deshalb ausreichend Gelegenheit hatten, Erfahrungen zu sammeln. Diese durchaus gravierende, in Vergessenheit geratene⁹ Einschränkung des Anwendungsbereichs der ökonomischen Entscheidungstheorie wurde im weiteren Verlauf durch einen universellen Erklärungsanspruch ersetzt.

Vorangetrieben wurde diese Entwicklung nicht zuletzt durch die Theorie offener Präferenzen. Diese klärte, unter welchen Bedingungen man davon ausgehen darf, dass Menschen über eine Präferenzordnung verfügen, die durch eine Nutzenfunktion abgebildet werden kann, deren Maximierung die Lösung des Entscheidungsproblems lie-

⁷ Was insbesondere ausschließt, dass sie von Referenzpunkten abhängig sein können.

⁸ Was sogenannte Framing-Effekte ausschließt.

⁹ Ebenda, S. 171.

fert.¹⁰ Diese Voraussetzungen erwiesen sich als überraschend schwache Forderung an die Fähigkeit, konsistente Entscheidungen zu treffen. Damit schien der Weg frei zu sein, den Anwendungsbereich der Rational-choice-Theorie nahezu beliebig auszuweiten.

Eine weitere Entwicklung kam 1944 hinzu. Mit dem Buch „Theory of Games and Economic Behavior“ von John von Neumann und Oskar Morgenstern erblickte die Spieltheorie das Licht der Welt. Mit einem Aufsatz von John F. Nash aus dem Jahr 1950 wurde sie rasch zu einem der wichtigsten Werkzeuge der ökonomischen Theorie. Die Spieltheorie ließ sich leicht und einfach in die Rational-choice-Theorie integrieren, denn wer sie anwendet, benutzt alle Voraussetzungen, die für diese Theorie kennzeichnend sind. Die Spieler sind rational, ihre Rationalität ist „Common knowledge“, sie verhalten sich strategisch und bilden rationale Erwartungen.¹¹

Mit dem Nash-Gleichgewicht und dem Nachweis, dass jedes endliche Normalformspiel mindestens ein solches Gleichgewicht besitzt, wurde den Ökonomen ein ganzer Kosmos strategischer Interaktionen erschlossen, die nun analysierbar waren und für die gesichert war, dass jedes Rätsel, das sie enthalten, auch mindestens eine Lösung besitzt. Die Spieltheorie bot die Möglichkeit, auch komplexe und schwierige strategische Interaktionen ohne jeglichen Rückgriff auf psychologische Erkenntnisse zu analysieren. Wann ist eine Drohung glaubwürdig? Die einfache spieltheoretische Antwort lautet, dass sie es dann ist, wenn es eine beste Antwort (das heißt: eine rationale Wahl) ist, die Drohung wahr zu machen. Psychologen dürften dies vollständig anders sehen, aber darum muss sich ein Spieltheoretiker nicht kümmern. Er weiß a priori genau, wann eine Drohung glaubwürdig ist.

Mit der Theorie offenbarer Präferenzen und mit der nicht-kooperativen Spieltheorie verfügt die Wirtschaftswissenschaft über ein Instrumentarium, das zwei wichtige Eigenschaften besitzt:

1. Es lässt sich universell auf nahezu jedes beliebige Entscheidungsproblem anwenden, einschließlich solcher, bei denen strategische Interaktionen eine Rolle spielen.
2. Es unterscheidet sich durch den Einsatz formaler Methoden grundlegend von dem Instrumentarium, das in

Nachbardisziplinen bzw. in Konkurrenzdisziplinen zum Einsatz kommt.

Damit ist das „paretianische Programm“ verwirklicht – aber nicht nur das. Die Verwirklichung geschah in einer Weise, die der ökonomischen Disziplin Möglichkeiten eröffnete, die anderen Disziplinen der Verhaltenswissenschaften nicht offen stehen. Beispielsweise ist es mit mikroökonomischen Modellen möglich, kausale Zusammenhänge zweifelsfrei nachzuweisen. Zwar gelten diese Zusammenhänge zunächst nur innerhalb des jeweiligen Modells, dort sind sie aber mit mathematischer Genauigkeit intersubjektiv nachprüfbar. Diese Eigenschaft der neoklassischen Theorie wird im Zusammenhang mit der Diskussion über die Methoden, die sich für die evidenzbasierte Politikberatung eignen, noch eine wichtige Rolle spielen. Jedenfalls gibt es in der ökonomischen Forschung nun den mathematischen Beweis als Beleg dafür, dass eine Modellimplikation kausal auf Modellannahmen zurückgeführt werden kann. Es ist nur eine Spekulation, aber es spricht einiges dafür, dass dies der neoklassischen Mikrotheorie über viele Jahrzehnte hinweg wissenschaftliche Grabenkämpfe und ideologische Dispute erspart hat, wie sie in anderen Disziplinen vorkamen.

Ein letzter Vorteil der neoklassischen Wirtschaftstheorie sei noch erwähnt: Sie verfügt über ein ausgezeichnetes Immunsystem. Ihr methodischer Kern ist die Annahme, dass Menschen rationale Entscheidungen treffen. Aber Rationalität ist dabei nichts anderes als die Forderung nach Konsistenz der Entscheidung im Hinblick auf die individuellen Präferenzen. Bezüglich dieser Präferenzen werden nur sehr schwache Annahmen getroffen, die offen lassen, welcher Art die Präferenzen sind. Diese Zweiteilung des Entscheidungsmodells ist sehr praktisch. Die restriktive Forderung, dass Entscheidungen rational gefällt werden, kann geschützt werden, wenn man hinreichend flexibel mit den Annahmen über die Inhalte der Präferenzordnung verfährt.

Binmore und Shaked (2010) weisen beispielsweise darauf hin, dass die Entdeckung, dass Menschen auch soziale Präferenzen besitzen können und keineswegs ausschließlich ihren materiellen Wohlstand maximieren, die neoklassische Theorie nicht in Gefahr bringen kann. Das ist in der Tat richtig. Die Reaktionen der Theorie auf die experimentellen Nachweise sozialer Präferenzen beweisen das eindrucksvoll. Angefangen bei Rabin (1993), über Bolton und Ockenfels (2000), Fehr und Schmidt (1999) bis Charness und Rabin (2002) handelt es sich bei allen diesen Theorien um Modelle, mit denen man experimentelle Befunde erfolgreich in das Gerüst der Rational-choice-Theorie integrieren und an den Grundannahmen rationalen

¹⁰ Vgl. dazu auch die Ausführungen im Anhang von Weimann et al. 2010 und 2015.

¹¹ Damit ist insbesondere gemeint, dass sie alle verfügbaren Informationen fehlerfrei verarbeiten und ihre Erwartungen bayesianisch anpassen. Darunter versteht man, dass sie die Veränderung der Wahrscheinlichkeiten, die durch zusätzliche Information eintritt, richtig, d. h. unter Verwendung des Theorems von Bayes, berechnen.

Verhaltens ohne jede Einschränkung festhalten kann. Für psychologische Erklärungen ist auch in diesen Theorien kein Platz. Sie postulieren bestimmte Inhalte der individuellen Präferenzordnungen – beispielsweise Ungleichheitsaversion (Bolton und Ockenfels 2000) oder Fairness (Rabin 1993) – und zeigen, dass damit Verhalten rationalisierbar ist, das nicht allein auf die Maximierung materieller Wohlfahrt zielt. Warum Menschen solche Präferenzen besitzen und wie sie entstehen, ist dabei nicht von Interesse.

Ihre universelle Anwendbarkeit, ihre herausragenden Fähigkeiten in der wissenschaftlichen Königsdisziplin des Kausalitätsnachweises und ihre weitgehende Immunität gegen empirische Kritik haben dazu beigetragen, dass die ökonomische Theorie des Rational choice zu einem großen Erfolg wurde. Sie bestimmt bis heute zu einem großen Teil das Geschehen in den führenden Journalen. Aber es hat sich insbesondere in den vergangenen 35 Jahren gezeigt, dass sich die ökonomische Theorie mit der Spieltheorie eine Methodik einverleibt hat, die in sich den Kern für eine Methode trägt, die letztlich dazu beigetragen hat, das Programm der Entpsychologisierung zu konterkarieren. Ursache dafür ist, dass sich jede Beschreibung eines Spiels auch als Anleitung zu einem Experiment interpretieren lässt.

2.2 Das Experiment als neue ökonomische Methode

Ganz am Anfang der experimentellen Wirtschaftsforschung stand allerdings eine Forschungsfrage, die nicht unmittelbar mit der spieltheoretischen Analyse zusammenhängt.¹² Es ging um die Frage, wie Märkte in ein Gleichgewicht gelangen. Diese Frage war lange nicht befriedigend zu beantworten. Es bedurfte einiger Hilfskonstruktionen wie des walrasianischen Auktionators, um eine Vorstellung davon zu bekommen, warum sich auf Märkten die vorteilhaften Gleichgewichte tatsächlich einstellen. Diese offenkundige Lücke im ökonomischen Theoriegebäude versuchte als erstes Edward Chamberlin (1948) mit Hilfe eines Experiments zu schließen. Dies war eines der ersten ökonomischen Experimente, die überhaupt durchgeführt wurden und prominent veröffentlicht werden konnten. Aber erst der spätere Nobelpreisträger Vernon Smith¹³ hat mit seinen Arbeiten zeigen können, dass sich Märkte unter experimentellen Bedingungen ausgesprochen rasch und sicher in ihr Gleichgewicht bewegen. Es reichte dazu aus,

Anbieter und Nachfrage mit privater Information über Zahlungsbereitschaft (bzw. den „Wert“ eines Gutes) und Kosten zu versorgen und die Akteure miteinander handeln zu lassen. Die so entstandenen experimentellen Märkte fanden mit großer Sicherheit und Schnelligkeit in ihr Gleichgewicht.

Diese ersten Experimente waren glücklich gewählt: Ihre Ergebnisse ließen sich leicht akzeptieren und in die herrschende Theorie integrieren. Sie halfen sogar, einen wunden Punkt der Theorie zu beseitigen. Das hieß natürlich noch lange nicht, dass damit die Experimente in den Methodenkasten der Ökonomik aufgenommen waren. Aber ihr erster Auftritt war zumindest gelungen. Vielleicht hat das den Weg dafür bereitet, dass sich ein zweites Forschungsprogramm durchsetzen konnte, dessen Wurzeln in den bereits erwähnten Eigenschaften spieltheoretischer Modelle lagen. Nahezu jedes theoretisch lösbares Spiel lässt sich auch tatsächlich spielen. Man muss nur die fiktiven Spieler durch reale Personen ersetzen und einen Weg finden, wie man die im Modell angenommenen Anreize ins Labor überträgt. Die „Induced-value“-Methode beschreibt, wie man dies auf einfache und elegante Weise erreichen kann, indem man Nutzenfunktionen in monetäre Auszahlungsfunktionen überführt.

Wenn jedes Spiel, das unter der Annahme des Rational choice analysiert wird, auch gespielt werden kann – wer will dann neugierige Wissenschaftler davon abhalten, genau das auch zu tun und zu schauen, was dabei herauskommt? Zu den ersten, die ihrer Neugier nachgingen, zählten Melvin Dresher und Merrill Flood. Sie erfanden das Gefangenendilemma. Damit konnten sie experimentell der Frage nachgehen, ob ein Nash-Gleichgewicht (also ein Gleichgewicht, in dem jeder Spieler seine beste Antwort auf die Strategie der anderen Spieler spielt), noch dazu bei dominanten Strategien (wenn eine Strategie immer beste Antwort ist, gleichgültig, was die anderen Spieler tun), auch dann noch gespielt wird, wenn es haarsträubend ineffizient ist.¹⁴ Ihr Ergebnis ähnelt dem Standardresultat der vielen Zwei- oder N-Personen-Gefangenendilemma-Experimente, die bis heute stattfinden: Teilweise folgen die Teilnehmer der Theorie, teilweise begannen sie zu kooperieren.

Das Beispiel zeigt, dass die nicht-kooperative Spieltheorie von Anfang an von der Frage begleitet wurde: „Was machen reale Menschen, wenn sie in genau der Situation sind, die im Modell beschrieben ist?“ Diese Frage liegt auf der Hand, aber dennoch hat es nach Dresher und Flood noch gute 25 Jahre gedauert, bis ihre systematische

¹² Vgl. Sugden (2005), S. 178.

¹³ Vgl. beispielsweise Smith (1962).

¹⁴ Vgl. dazu Goeree und Holt (2001).

Beantwortung in Schwung kam. Seit dem Ende der siebziger Jahre aber steht endgültig die Frage im Raum, wie weit die Rational-choice-Theorie trägt. Die für viele Theoretiker ernüchternde Antwort lautete, dass sie in vielen Zusammenhängen, wenn auch längst nicht in allen, ziemlich schnell am Ende war.

Exemplarisch lassen sich dieser Vorgang und die daraus erwachsenden Konflikte an der Literatur zu „Ultimatum-Experimenten“ verdeutlichen. Dieses Experiment wurde 1982 von Güth, Schmidtberger und Schwarzer in die Literatur eingeführt. Seine Regeln sind extrem einfach und das teilspielperfekte Gleichgewicht des Spiels ist so gut intuitiv fassbar, dass man es jedem ökonomischen Laien in weniger als einer Minute erklären kann. Zwei Spieler verhandeln über die Aufteilung eines Geldbetrages X . Einer der beiden (der „Proposer“) schlägt eine Aufteilung vor, der andere (der „Responder“) kann annehmen oder ablehnen. Bei Annahme wird aufgeteilt wie vorgeschlagen, bei Ablehnung bekommen beide Spieler eine Auszahlung von 0. Für den Responder ist die Annahme jedes Angebots, das ihm mehr auszahlt als 0, eine beste Antwort. Dies antizipierend, bietet der Proposer dem Responder tatsächlich nur einen minimalen Anteil an X und der Responder nimmt diesen Vorschlag an.

Die Entdeckung von Güth et al. (1982) war, dass dieses Gleichgewicht im Experiment nicht beobachtet werden konnte. Vielmehr boten die Proposer deutlich höhere Beträge an als im Gleichgewicht, und die Responder waren bereit, Angebote zurückzuweisen, die ihnen positive Auszahlungen beschert hätten. Die Einfachheit des Spiels machte es unmöglich, die Abweichung vom Gleichgewicht darauf zurückzuführen, dass die Spieler das Spiel nicht verstanden hatten oder kognitive Beschränkungen aufwiesen. Deshalb waren die Ergebnisse von Güth et al. ein Angriff auf den Anspruch der Spieltheorie, reales Verhalten in strategischen Interaktionen voraussagen zu können. Binmore, Shaked und Sutton (1985) nahmen den Fehdehandschuh auf und versuchten ihrerseits mit Hilfe eines Experiments zu zeigen, dass die Spieltheorie doch mit Recht einen positiven theoretischen Anspruch erhebt. Die Auseinandersetzung war heftig und kurz. Sie hatte einen klaren Sieger: In den vielen Ultimatum-Experimenten, die seit 1982 durchgeführt worden sind¹⁵, hat sich das Ergebnis von Güth et al. qualitativ immer wieder bestätigen lassen.

Der Widerspruch zwischen den Ergebnissen der Ultimatum-Experimente und der Prognose der neoklassischen Theorie erwies sich als persistentes Problem. In immer neuen experimentellen Anordnungen konnte man

beobachten, dass viele Menschen auch dann nicht ihre Auszahlung maximieren, wenn das dazu notwendige Optimierungsproblem sehr einfach und überschaubar ist. Im „Trust game“, dem Vertrauensspiel¹⁶, beobachtete man ebenso wenig das teilspielperfekte Gleichgewicht wie im „Gift exchange game“, dem Geschenkaustausch-Spiel¹⁷. In zahlreichen Experimenten wurde außerdem deutlich, dass Versuchspersonen im Labor in der Lage sind, soziale Dilemmata, wie sie in „Öffentliches-Gut-Experimenten“ und bei „Allmende-Problemen“ auftreten¹⁸, zumindest teilweise zu überwinden. Schließlich zeigte sich in „Diktator-Experimenten“, in denen Versuchspersonen die Möglichkeit hatten, einen Geldbetrag in beliebiger Weise zwischen sich und einem anderen Spieler aufzuteilen, dass ein großer Anteil der Versuchspersonen darauf verzichtete, den gesamten Betrag für sich zu behalten. Es mag sein, dass viele Spieltheoretiker diese Ergebnisse als Niederlage empfunden haben, aber so muss man es nicht sehen.

2.3 Eingeschränkte Rationalität kann nützlich sein

Die Ergebnisse der Ultimatum-Experimente und der anderen Experimente, die das Rational-choice-Modell zu widerlegen scheinen, bieten Ansatzpunkte für eine durchaus positive Bewertung. Diese entsteht, wenn man das Modell des strikt rationalen Menschen hinsichtlich seiner Bedeutung kritisch hinterfragt. Damit ist nicht die Frage gemeint, ob Menschen sich tatsächlich rational verhalten, sondern ob es für Menschen eigentlich immer vorteilhaft und gut ist, sich strikt rational zu verhalten. Auf den ersten Blick scheint es so zu sein, denn nichtrationales Verhalten bedeutet schließlich, dass man sich im Widerspruch zu seinen wahren Interessen befindet. Aber es gibt zwei Stellen, an denen sich der Homo oeconomicus gewissermaßen selbst im Wege steht.

¹⁶ Dabei erhalten zwei Spieler einen Betrag X . Spieler A kann an Spieler B dann einen beliebigen Anteil seiner Anfangsausstattung an B übergeben. Was übergeben wird, wird vom Experimentator verdreifacht, d. h. bei B kommt der dreifache Betrag an, den A abgegeben hat. Danach kann B entscheiden, ob er etwas an A zurückgibt. Im Gleichgewicht hat B keinen Anreiz etwas an A zurückzugeben. Das antizipiert A und gibt nichts an B ab. Der mögliche Effizienzgewinn ist verloren. Berg et al. (1995).

¹⁷ Ein Experiment, das dem Vertrauensspiel sehr ähnlich ist, aber häufig als Spiel zwischen Firma und Beschäftigtem gespielt wird. Fehr et al. (1997).

¹⁸ Das dabei auftretende Problem entspricht dem des Gefangenendilemmas.

¹⁵ Einen frühen Überblick liefert Roth (1995).

Die erste hängt unmittelbar mit dem Ultimatum-Spiel zusammen. Strikt rationales Verhalten hat in Situationen, in denen die Spieler unter vollständiger und perfekter Information agieren wie im Ultimatum-Spiel, zur Folge, dass die Entscheidungen eines Homo oeconomicus exakt berechenbar sind. Das macht ihn zu einem leichten Opfer. Im Ultimatum-Spiel kann der Proposer sich vollständig gegen den Responder durchsetzen. Im Stackelberg-Modell des Duopols kann der „Mengenführer“ den „Mengenfolger“ an jede beliebige Stelle seiner Reaktionsfunktion bugsieren – die Rationalität der Spieler macht sie berechenbar und damit ausbeutbar.

Ganz anders sieht das experimentelle Ergebnis aus. Im Ultimatum-Experiment ist die Auszahlungsdifferenz zwischen Proposer und Responder sehr viel kleiner als im Gleichgewicht. Ähnliche Ergebnisse lassen sich für weitere Spiele zeigen, in denen sich Spieler in einer schlechten strategischen Position befinden, weil die einzigen Drohungen, die sie aussprechen können, im Sinne der Spieltheorie unglaubwürdig sind.¹⁹ Man beachte dabei, dass auch solche Spieler die Vorteile genießen können, die der Erstziehende in diesen Spielen haben kann, die in der Responder-Rolle bereit sind, positive Angebote abzulehnen.

Die zweite Stelle, an der Rationalität eher hinderlich ist, findet sich dort, wo auf den Homo oeconomicus die berühmten Rationalitätsfallen lauern, die überall dort aufgestellt sind, wo sich Gefangenen-Dilemmata in ihren verschiedenen Formen auftun. Ob es sich um externe Effekte handelt, um öffentliche Güter oder um Allmenden, rationale Akteure sind grundsätzlich nicht in der Lage, das Wohlfahrtsoptimum zu realisieren. Sie bleiben in hoffnungslos ineffizienten Lösungen stecken. Das ist für das ökonomische Rational-choice-Modell ein sehr ernsthaftes Problem – jedenfalls dann, wenn man mit diesem Modell einen universellen Erklärungsanspruch verbindet und eine positive theoretische Gültigkeit reklamiert. In der anthropologischen Forschung ist man sich ziemlich sicher, dass der Aufstieg des Homo Sapiens eng damit zusammenhängt, dass Menschen in der Lage sind zu kooperieren. Aber selbst wenn man von einer solchen entwicklungs geschichtlichen Interpretation absieht, fällt es schwer, sich vorzustellen, wie die hochkomplexen Gesellschaften unserer Zeit existieren könnten, wenn die Menschen unfähig sein sollten, kooperative Lösungen zu finden.

Angesichts dieser beiden Stellen, an denen Rationalität erkennbar nicht vorteilhaft ist, ist es eine gute Nachricht, dass Menschen genau in solchen Situationen, in denen sie entweder in der Gefahr sind, ausgebeutet zu werden

oder in Rationalitätsfallen zu laufen, *nicht* vollständig rational handeln. Dies gilt ungeachtet der Tatsache, dass soziale Dilemmata durchaus ein gravierendes Problem darstellen und beispielsweise Kooperation in großen Gruppen kein Selbstläufer ist. Wie sich noch zeigen wird, wecken gerade die Erkenntnisse der experimentellen Forschung und der Verhaltensökonomik die Hoffnung, dass solche Probleme grundsätzlich lösbar sind. Diese Hoffnung könnte man kaum haben, wenn das Rational-choice-Modell das menschliche Verhalten korrekt beschrieb.

2.4 Seltens methodischer Dualismus

Zwischen den Befunden der experimentellen Forschung und dem neoklassischen Modell rationaler Entscheidungen ließe sich relativ leicht eine methodische Versöhnung durch eine strikte Arbeitsteilung herbeiführen. Die Rational-choice-Modelle hätten dabei die Rolle, eine normative Theorie zu entwickeln; die experimentelle Forschung (zusammen mit der Verhaltensökonomik) wäre für die positive theoretische Forschung zuständig. Wie eine solche Arbeitsteilung begründet werden könnte, sei an dieser Stelle mit Hilfe einer Anekdote verdeutlicht, die sich auf einer Jahrestagung des Vereins für Socialpolitik zugetragen hat²⁰.

Auf dieser Tagung hielt Reinhard Selten die Thünen-Vorlesung. Der Vortrag bestand im Wesentlichen aus verhaltensökonomischen Erkenntnissen und Vorschlägen dazu, wie diese mit einer Theorie eingeschränkt rationalen Verhaltens erklärt werden könnten. Im Anschluss an den Vortrag fragte der damalige Vorsitzende des Vereins, Hans-Werner Sinn, der die Moderation inne hatte, Professor Selten scherzhaft, ob er nicht vielleicht ein bisschen schizophren sei. Schließlich habe er den Nobelpreis für eine Arbeit bekommen, in der er davon ausgehe, dass sich Menschen hyperrational verhielten, gleichzeitig aber forsche er intensiv an Theorien, in denen er annehme, dass Menschen nur eingeschränkt rational seien. Wenig überraschend bestritt Selten, schizophren zu sein, und erklärte, er sei ein „methodologischer Dualist“. Auf der einen Seite strebten wir alle danach, rational zu sein, also keine Fehler zu begehen, wenn wir unsere Ziele verfolgten. Deshalb sei es wichtig zu überlegen, wie denn eine vollständig rationale Lösung aussähe. Auf der anderen Seite sei klar, dass es uns nicht gelinge, immer vollständig rationale Entscheidungen zu treffen. Deshalb ergebe es genauso viel Sinn, danach zu forschen, wie denn tatsächliche, nur einge-

¹⁹ Vgl. beispielsweise Brosig et al. (2004), Mitropoulos et al. (2007).

²⁰ Nach Erinnerung des Autors war es die Tagung im Jahr 1999.

schränkt rational gefällte Entscheidungen zustande kämen.

Mit Seltens methodischem Dualismus wäre eine perfekte Arbeitsteilung möglich. Eine solche schlösse keineswegs aus, dass die positive Theorie auch Situationen identifiziert, in denen Menschen sehr wohl in der Lage sind, rationale Entscheidungen zu treffen, und in denen deshalb die Rational-choice-Theorie prognostischen Wert besitzt. Die Laborexperimente der vergangenen drei Dekaden haben keineswegs ausschließlich gezeigt, dass die Menschen vom Rationalmodell abweichen. Im Gegenteil hat sich in vielen Experimenten beispielsweise gezeigt, dass die Versuchspersonen sehr wohl auf Anreize reagieren – und das durchaus auch so, wie es das Rationalmodell beschreibt.²¹ Ein in dieser Weise organisiertes Nebeneinander von Rational choice und Verhaltensökonomik entspräche im Übrigen durchaus Paretos ursprünglichen Programm, denn er wollte Rationalität nur in bestimmten Situationen als Erklärungsansatz heranziehen.²² Es entspräche auch der pragmatischen modernen Sicht, wie sie Chetty (2015) formuliert hat. Nach seiner Auffassung ist die Verhaltensökonomik fallweise anzuwenden, wenn sich abzeichnet, dass sie bessere Erklärungen liefern kann als das neoklassische Modell.

Allerdings setzt eine solche Arbeitsteilung voraus, dass die Vertreter der neoklassischen Theorie ihren Universalanspruch aufgeben und eingestehen, dass ihr Ansatz nur in einer Teilmenge der Menge aller Entscheidungen anwendbar ist. Wenn man das Publikationsgeschehen in den ökonomischen Top-Journalen betrachtet, ergibt sich ein differenziertes Bild einer sehr dynamischen Entwicklung. In den vergangenen dreißig Jahren hat es eine deutliche Hinwendung zu empirischen Arbeiten gegeben²³, und die Zahl von Publikationen, deren Verfasser explizite verhaltensökonomische Alternativen zur Neoklassik behandeln, ist stetig gestiegen. Aber zugleich hat man den Eindruck, dass das Immunsystem, mit dem das Rational-choice-Paradigma ausgestattet ist, in den vergangenen zwanzig Jahren gute Arbeit geleistet hat.

²¹ Um ein aktuelles Beispiel zu nennen: Brosig-Koch et al. (2015, im Erscheinen) zeigen, dass Ärzte auf unterschiedliche Entlohnungssysteme exakt so reagieren, wie es das neoklassische Rationalmodell prognostiziert.

²² Vgl. Bruni und Sugden (2007).

²³ Hamermesh (2013) untersucht das Publikationsverhalten zwischen 1980 und 2010 und kommt zu dem Ergebnis, dass der Anteil der empirischen Arbeiten, die in „General-interest“-Journals veröffentlicht wurden, von 38 auf 72 Prozent gestiegen ist.

2.5 Der neoklassische Reparaturbetrieb

Die vielen Befunde, die mit einfachen ökonomischen Experimenten diagnostiziert wurden und die zeigten, dass reale Menschen vom Rationalmodell abweichen, bildeten den Nährboden für eine Entwicklung, an deren Ende der Begriff der „sozialen Präferenzen“ zu einem festen Bestandteil des ökonomischen Jargons geworden ist.²⁴ Die Beobachtungen im Ultimatum-Experiment wiesen darauf hin, dass die Menschen einen gewissen Sinn für Fairness besitzen. Die Ergebnisse des Gift-exchange-Experiments wurden als Beleg dafür gelesen, dass sich die Versuchspersonen *reziprok* verhielten²⁵. Das bedeutet, dass sie bereit waren, auf eigene Auszahlungen zu verzichten, um jemanden, der zuvor „nett“ zu ihnen war, zu belohnen, bzw. um jemanden, der nicht nett oder „unnett“ war, zu bestrafen. Ökonomen begannen sich dafür zu interessieren, wie man diese Befunde und Ideen in das neoklassische Theoriegerüst einbauen kann. Den Anfang machte Rabin (1993) in einer Arbeit, in der er Fairness als Konzept in die Standard-Spieltheorie einführte. Falk und Fischbacher (2006) lieferten ein theoretisches Konzept, das erlaubte, mit dem Phänomen reziproken Verhaltens umzugehen. Mit den Arbeiten von Bolton und Ockenfels (2000), Fehr und Schmidt (1999) sowie schließlich Charness und Rabin (2002) wurden soziale Präferenzen endgültig in der neoklassischen Theorie heimisch.

Der Erfolg der neuen theoretischen Ansätze ist leicht zu erklären. Die Idee, dass Menschen so etwas wie soziale Präferenzen besitzen, kann durch eine relativ geringfügige Änderung der Annahmen über die Inhalte der Präferenzordnung in die Standardtheorie intergriert werden. Beispielsweise benutzten Bolton und Ockenfels für ihr ERC-Modell (Equity, reciprocity, competition) die Annahme, dass Menschen eine Präferenz sowohl für die eigene absolute Auszahlung hegen als auch für die eigene relative Auszahlung. Mit dieser vergleichsweise kleinen Änderung der Annahmen über die individuellen Präferenzen ist man in der Lage, eine große Zahl von experimentellen Beobachtungen in Einklang mit der theoretischen Prognose zu bringen, was nicht möglich ist, wenn allein die absolute eigene Auszahlung als Argument in der Nutzenfunktion steht.

Die neoklassische Theorie ist wie gesagt nicht grundsätzlich auf die Annahme angewiesen, dass Menschen allein an ihrem eigenen materiellen Wohlstand interessiert

²⁴ Einen Überblick bieten Fehr und Schmidt (2003) sowie Camerer (2003).

²⁵ Vgl. Falk und Fischbacher (2006), Nowak und Roch (2006), Nowak und Siegmund (2005) und Levine (1998).

sind (Binmore und Shaked 2010). Altruistische Präferenzen oder, allgemeiner, „*Other-regarding preferences*“ sind mit dem neoklassischen Modell deshalb im Prinzip leicht vereinbar, weil man in der reinen Theorie darauf verzichtet, Angaben zum Inhalt der Präferenzordnungen zu machen. In der Praxis der Modellbildung ist diese Offenheit aber nicht wirklich durchzuhalten. Damit man zu gehaltvollen Prognosen gelangt, ist eine Präzisierung der Präferenzinhalte unumgänglich. Der Regelfall ist dabei die Annahme, dass sich Menschen strikt eigennützig verhalten.²⁶ Das Immunsystem der Neoklassik funktioniert vor allem deshalb so gut, weil sie auf der einen Seite die grundsätzliche Offenheit besitzt, Veränderungen der Annahmen über die Präferenzinhalte zuzulassen, andererseits aber über eine einfache Standardannahme verfügt, mit der die Theorie sehr weit gekommen ist: Mehr materieller Wohlstand ist besser als weniger.

Bei dieser Standardannahme kann es auch nach der Entdeckung sozialer Präferenzen bleiben. Die sozialen Präferenzen bleiben dann den wenigen Fällen vorbehalten, in denen das Modell strikt eigennütziger Akteure versagt. Larry Samuelson hat auf den Punkt gebracht, wie die Vertreter der neoklassischen Theorie mit der Entdeckung sozialer Präferenzen fertig wurden: „Whatever motivates the subjects, whether money or fairness or something else, it is something that we can model with the familiar optimization tools of economics, without abandoning rational behavior as a unifying principle.“ (Samuelson 2005, S. 70)

Bis hierhin kommt die experimentelle Forschung innerhalb der Wirtschaftswissenschaft vollständig ohne Bezüge zur psychologischen Forschung aus. Solange der Gegenstand der experimentellen Forschung allein der Test von ökonomischer Theorie ist, bleibt der Abstand zur Psychologie gewahrt und die ökonomische Forschung als Ganzes behält ihre Eigenständigkeit. Dabei hilft auch, dass sich die ökonomischen Experimente methodisch sehr deutlich von psychologischen abgrenzen lassen. Es sind vor allem zwei wichtige Eigenschaften ökonomischer Versuche, die den Unterschied ausmachen. Erstens laufen wirtschaftswissenschaftliche Experimente praktisch immer unter Verwendung monetärer Anreize, zweitens verzichten Ökonomen grundsätzlich darauf, die Versuchspersonen zu manipulieren. Ein Teilnehmer an einem ökonomischen Experiment kann sicher sein, dass das, was der Experimentator beschreibt, auch tatsächlich das ist, was im Experiment abläuft. In vielen psychologischen

²⁶ Wenn im Folgenden von „eigennützigem Verhalten“ die Rede ist, dann ist damit Eigennutz im materiellen Sinne gemeint. Diese Klärung ist notwendig, weil sich auch ein optimierender Altruist streng genommen „eigennützig“ verhält.

Experimenten kann er das nicht. Mit dieser Abgrenzung ist es möglich, nicht nur die Resultate, sondern die gesamte experimentelle Forschung in die Neoklassik zu integrieren, ohne dass das paretianische Programm der Entpsychologisierung in Gefahr gerät.

Auf den ersten Blick ist damit eine relativ reibungslose Assimilation der experimentellen Ökonomik gelungen. Es bleiben allerdings einige „Restbefunde“ der Forschung zu sozialen Präferenzen, die davon ausgenommen sind. Einer betrifft die stabile Beobachtung, die sich in verschiedenen Kontexten sicher reproduzieren lässt, zur Wirkung von „Cheap talk“²⁷. Darunter versteht man eine Kommunikation zwischen den Spielern, die aus spieltheoretischer Sicht für die strategische Situation und damit für das Gleichgewicht des Spiels keine Rolle spielen dürfte.²⁸ Tatsächlich wirkt aber die Kommunikation auch dann sehr stark, wenn Reputationseffekte ausgeschlossen werden können.

Brosig et al. (2003) zeigten darüber hinaus, dass die Kommunikationswirkung stark vom Kommunikationskanal abhängt. Die stärkste Wirkung entfaltet sich, wenn die Spieler von Angesicht zu Angesicht kommunizieren können. In diesem Experiment waren die Gruppen, die so kommunizieren konnten, in der Lage, ein öffentliches Gut effizient bereitzustellen.²⁹ Auch das Koordinationsproblem im „Minimum effort coordination game“³⁰ verschwindet vollständig, wenn man den Versuchspersonen die Gelegenheit gibt, vor dem Experiment miteinander zu sprechen.³¹ Dieses Beispiel zeigt, dass die experimentelle Forschung durchaus auch Ergebnisse hervorgebracht hat, die sich nicht ohne weiteres in die neoklassische Theorie integrieren

²⁷ Brosig et al. (2003) zeigen die Wirkung von Kommunikation in Öffentliches-Gut-Spielen, Riechmann und Weimann (2008) im „Minimum effort coordination game“, Brosig et al. (2004) in einem einfachen Verhandlungsspiel und Hoffmann et al. (2015) in einem Verhandlungsspiel mit asymmetrischer Information.

²⁸ Sie spielt beispielsweise dann eine Rolle, wenn Kommunikation Reputationseffekte verursachen kann.

²⁹ Dass Reputationseffekte dabei bestenfalls eine untergeordnete Rolle spielen, konnte dadurch gezeigt werden, dass eine reine Identifikation der Spieler (ohne Kommunikationsmöglichkeit) ohne jeglichen Effekt blieb.

³⁰ In diesem Spiel steht eine Gruppe von N Spielern vor der Aufgabe, zu entscheiden, wie viel Aufwand jeder einzelne zur Erstellung eines Gutes zu leisten bereit ist. Die Auszahlung an die Gruppe hängt davon ab, wie groß die minimale Anstrengung in der Gruppe gewählt wird. Aufwand verursacht Kosten, aber die Auszahlung ist so gestaltet, dass es die Pareto-dominante Lösung ist, wenn alle die höchste Anstrengung wählen. Risiko-dominant ist jedoch das Gleichgewicht, in dem alle die niedrigste Anstrengung wählen. Von Hyck et al. (1990) zeigten erstmals, dass bereits bei Gruppen mit mehr als fünf Mitgliedern ein massives Koordinationsversagen auftritt.

³¹ Riechmann und Weimann (2008).

ren lassen. Solche Stellen sind potentiell „anfällig“ für psychologische Erklärungsversuche.

2.6 Interne und externe Validität von Experimenten

Trotz der zuletzt genannten Einschränkung stellt sich die Frage, warum die experimentelle Forschung in der Politikberatung eine derart untergeordnete Rolle spielt, obwohl sie erfolgreich in die Theorie integriert worden ist. Der Grund dürfte im asymmetrischen Verhältnis von externer und interner Validität von Laborexperimenten zu finden sein. Die interne Validität betrifft die Frage, wie gut es mit Hilfe von Laborexperimenten gelingt, theoretische Modelle zu testen. Dafür ist es notwendig, einerseits die Anreize, die im Modell als wirksam angenommen werden, möglichst 1:1 ins Labor zu übertragen und andererseits eine randomisierte Zuordnung der Versuchspersonen zu den verschiedenen Anordnungen (einschließlich der Kontrollgruppen) zu leisten. Beides kann bei Laborexperimenten relativ einfach sichergestellt werden, so dass es gerechtfertigt ist, von einer hohen internen Validität auszugehen.³² Diese gesicherte interne Validität macht es für die Vertreter der reinen Theorie schwierig bis unmöglich, die experimentellen Resultate (insbesondere wenn sie für die Theorie negativ ausfallen) zu ignorieren.

Vollkommen anders stellt sich die Situation im Hinblick auf die externe Validität von Laborexperimenten dar. Solange Experimente ausschließlich dafür benutzt werden, Theorien zu testen, ist die Übertragbarkeit auf die Realität nebensächlich. Beim Theorientest kommt es ausschließlich auf die interne Validität an. Manche experimentellen Ökonomen nutzen dies, um der Kritik zu entgehen, dass die Experimente zu artifiziell seien, um auf die Realität übertragbar zu sein (Croson und Gächter 2010). Wenn aber das Experiment zu artifiziell ist, um etwas über Realität zu sagen, dann gilt das auch für die Theorie, die mit dem Experiment getestet werden soll. Dann werden Modelltheorie und Experiment zu einem selbstreferenziellen Gebilde, für dessen Rechtfertigung man eigentlich kein Experiment braucht.

Im Gegensatz dazu ist die Politikberatung darauf angewiesen, dass die Kausalitäten, die Ökonomen entdeckt haben, auch in der Realität eine Rolle spielen. Deshalb muss jede Methode, auf deren Grundlage Politikberatung

stattfinden soll, über eine entsprechend ausgeprägte externe Validität verfügen. Das gilt – und das ist ein wichtiger Punkt – sowohl für das Experiment als auch für das formale Modell.

Levitt und List (2007, 2008) haben in zwei viel beachteten und prominent publizierten Arbeiten eine Reihe von Vorbehalten gegen die Übertragbarkeit von Laborexperimenten auf die Realität formuliert. In einem *Science*-Artikel kamen sie zu folgendem Schluss: „Perhaps the greatest challenge facing behavioral economics is demonstrating its applicability in the real world.“ (Levitt und List 2008, S. 909) Nur am Rande sei bemerkt, dass Levitt und List offensichtlich die experimentelle Forschung mit der Verhaltensökonomik gleich setzten, was gerade an dieser Stelle zu einem Missverständnis führt. Die externe Validität von *verhaltensökonomischen* Befunden ist weit aus weniger problematisch als jene von ökonomischen Laborexperimenten, auf die sich ihre Kritik ausschließlich bezieht. Levitt und List boten eine Liste von Gründen an, die ihrer Meinung nach dagegen sprachen, dass sich experimentelle Ergebnisse zu sozialen Präferenzen auf die Realität übertragen lassen. Die meisten dieser Punkte kann man als eine allgemeine Kritik an Laborversuchen verstehen, die nicht auf die spezifische Untersuchung von sozialen Präferenzen eingeschränkt werden muss. Die wichtigsten Kritikpunkte von Levitt und List seien hier aufgeführt:

- 1) *Versuchspersonen stehen unter Beobachtung.* Die typische Situation in einem Laborexperiment ist aus der Sicht einer Versuchsperson dadurch gekennzeichnet, dass ihr Verhalten genau beobachtet, aufgezeichnet und später ausgewertet wird. Diese spezielle Situation unterscheidet die Versuchsperson im Labor von Personen „im Feld“ und sie hat, so die These von Levitt und List, Einfluss auf das Verhalten.
- 2) *Die Auszahlungen in Experimenten sind zu gering.* Diese Kritik hat die experimentelle Forschung von Anbeginn an begleitet. Die dahinter stehende Hypothese ist, dass bei höheren Auszahlungen die Anreize, eine rationale Entscheidung zu treffen, massiv steigen. Folglich müsste man bei hinreichend hohen Auszahlungen auch ein entsprechend rationales Verhalten beobachten. Auch die Existenz sozialer Präferenzen (die ja nicht im Widerspruch zu rationalem Verhalten steht) könnte eine Frage des Preises sein. Wenn es zu teuer wird, „sozial“ zu sein, kann das die Rückkehr zum Eigennutz bedeuten.
- 3) *Die Auswahlmengen sind künstlich beschränkt.* Die Spieler können nur aus einer vorgegebenen Menge von möglichen Aktionen auswählen, und diese ist unter Umständen zu eng gewählt. Als Beispiel führ-

³² Zugleich bedeutet das auch, dass Laborexperimente bei der Identifikation von Kausalitäten ähnlich leistungsfähig sind wie das theoretische Modell.

ten Levitt und List das Diktator-Experiment an. Der Strategieraum des Diktators ist insofern beschränkt, als er nur positive Abgaben an den Rezipienten leisten kann. List (2007) und Bardsley (2008) zeigten, dass sich das Abgabeverhalten im Diktator-Experiment dramatisch ändert, wenn man zulässt, dass die Diktatoren den Rezipienten auch Geld wegnehmen dürfen.

- 4) *Die Versuchspersonen sind nicht repräsentativ; es gibt einen „Selektionseffekt“.* Üblicherweise werden ökonomische Experimente mit studentischen Versuchspersonen durchgeführt. Aus dem Verhalten von Studenten jedoch kann nicht unbedingt darauf geschlossen werden, wie sich nichtstudentische Versuchspersonen verhalten hätten bzw. wie sich reale Menschen (Manager, Angestellte, Menschen ohne akademischen Abschluss) verhalten. Außerdem nehmen Versuchspersonen an den Experimenten freiwillig teil. Bei der Rekrutierung der Versuchspersonen kann es folglich zu einem Selektionseffekt kommen. Es kann sein, dass bestimmte Typen oder Gruppen von Studenten eher bereit sind, an einem Experiment teilzunehmen, als andere.
- 5) *Die Versuchspersonen haben keine Gelegenheit zu lernen.* In ökonomischen Experimenten wird in der Regel sehr viel Wert darauf gelegt sicherzustellen, dass die Versuchspersonen das Entscheidungsproblem, mit dem sie konfrontiert werden, genau verstanden haben. Beispielsweise ist es üblich, Proberunden zu veranstalten und den Versuchspersonen Testfragen zu stellen, um zu überprüfen, ob sie das Spiel tatsächlich verstanden haben. Außerdem werden die meisten Experimente wiederholt, das heißt, die Versuchspersonen müssen ein und dieselbe Entscheidung mehrfach innerhalb einer experimentellen Sitzung treffen. Insofern haben sie durchaus Gelegenheit zu lernen. Dennoch ist der Versuch als solcher neuartig für sie, ebenso wie die experimentelle Situation, in der sie sich befinden. Und die Erfahrung dieser Situation als Ganzes wiederholt sich üblicherweise nicht. Das ist in der Realität anders; Menschen geraten immer wieder in Situationen, in denen sie sehr ähnliche Entscheidungen treffen müssen. Dabei kann zwischen den einzelnen Ereignissen mehr oder weniger Zeit vergehen. Die Effekte, die durch solche wiederholten Entscheidungen auftreten, werden in singulären Experimenten nicht erfasst. Beispielsweise ist über die Dynamik sozialer Präferenzen relativ wenig bekannt. Implizit wird in der Literatur für diese Präferenzen genauso Stabilität unterstellt wie für die Standardpräferenzen der neoklassischen Theorie.

Alle diese Einwände lassen sich ihrerseits zum Gegenstand experimenteller Untersuchung machen, und inzwischen gibt es eine ganze Reihe von Experimenten, in denen genau das geschieht.³³ Die dabei erzielten Ergebnisse lassen sich vorsichtig wie folgt zusammenfassen: Keiner der von Levitt und List aufgeführten Punkte hat sich als derart schwerwiegend herausgestellt, dass eine Übertragung experimenteller Ergebnisse auf die Realität grundsätzlich unmöglich erscheint. Andererseits gibt es systematische Einschränkungen, derer man sich bewusst sein muss, wenn man aus experimentellen Ergebnissen Gestaltungsvorschläge für die Realität ableiten will.

Dass Versuchspersonen im Labor unter strikter Beobachtung stehen, scheint beispielsweise wenig Einfluss auf ihr Verhalten zu haben (Barmettler et al. 2012, Bolton et al. 1998). Die Anonymität, mit der Versuchspersonen ausgestattet werden können, ist im Experiment zudem gestaltbar. In sogenannten Doppelblindversuchen kann nicht nur ein Proband den anderen nicht beobachten, es sieht außerdem auch der Experimentator nicht, was die einzelne Versuchsperson tut. Die Beobachtbarkeit ist dann maximal eingeschränkt. Hoffman et al. (1994) haben beispielsweise gezeigt, dass unter einer solchen Doppelblind-Anordnung die Zahl der Personen, die im Diktator-Experiment Geld an eine andere Person abgeben, dramatisch zurückgeht.³⁴

Die experimentellen Ergebnisse zu der Frage, wie sich höhere Auszahlungen auf das Laborverhalten auswirken, sind uneindeutig. Carpenter et al. (2005) haben weder im Ultimatum-Experiment noch im Diktator-Experiment einen Unterschied zwischen Versuchsanordnungen festgestellt, bei denen es um 10 oder um 100 Dollar ging. Allerdings konnten sie im Ultimatum-Experiment nur Aussagen über das Verhalten der Proposer machen, weil alle Angebote (bis auf eines) angenommen wurden und deshalb über die Annahmegränze der Responder nichts gesagt werden kann. Andersen et al. (2011) haben gezeigt, dass diese Grenze bei höheren Auszahlungen fällt, und behaupten, dass sie für hinreichend große Auszahlungen gegen Null gehen würde. Slonim und Roth (1998) haben beobachtet, dass die Varianz im Verhalten mit steigenden Auszahlungen abnimmt. Camerer und Hogarth (1999) finden dagegen in einer Meta-Analyse keine starken Effekte der Auszahlungshöhe, ebenso wenig wie Fehr et al. (2002).

³³ Überblicke bieten beispielsweise Falk und Heckman (2009) sowie Barmettler et al. (2012).

³⁴ Vgl. auch Cherry et al. (2002), die ebenfalls beobachten, dass bei einer Doppelblind-Anordnung die Abgaben im Diktatorspiel gegen Null gehen.

Auch im Hinblick auf die Frage, ob vom Verhalten studentischer Versuchspersonen auf die Entscheidungen „normaler“ Menschen geschlossen werden kann, sind die Ergebnisse nicht eindeutig. Neben Befunden, die zeigen, dass Experten sich anders verhalten als studentische Versuchspersonen (beispielsweise Feltovich 2012), gibt es eine Reihe von Experimenten, die keinen Unterschied ergeben (Dohmen et al. 2008). Es scheint dabei eine Rolle zu spielen, ob die Experten in der Realität spezielle Fertigkeiten oder Daumenregeln entwickeln können, die ein bestimmtes Entscheidungsproblem zu bewältigen helfen. Es liegt nahe, dass Studenten, wenn sie ein einziges Mal vor die Aufgabe gestellt werden, ein solches Entscheidungsproblem zu lösen, über diese speziellen Fertigkeiten nicht verfügen und sich deshalb anders verhalten als Experten. Weiterhin hat sich gezeigt, dass die befürchtete Selbstselektion der Versuchsteilnehmer nicht vorzuliegen scheint (Falk et al. 2013).

Sehr wenig lässt sich bisher zu der Frage sagen, ob das im Labor beobachtete Verhalten über die Zeit stabil ist, denn dazu gibt es vergleichsweise wenige Experimente. Insbesondere von Interesse ist, ob soziale Präferenzen in diesem Sinne stabil sind. Sass und Weimann (2015a, 2015b) und Sass et al. (2015) zeigen, dass in verschiedenen Standardspielen eine bestimmte Dynamik des Verhaltens zu beobachten ist, wenn die Experimente mit einem gewissen zeitlichen Abstand wiederholt werden. In aller Regel lässt sich beobachten, dass der Anteil der nichteigennützigen Entscheidungen bei identischer Wiederholung des Experimentes zurückgeht. Dies gilt auch dann, wenn Lerneffekte ausgeschlossen werden können.³⁵ Im Hinblick auf soziale Präferenzen scheint dabei ein Phänomen wichtig zu sein, das aus der Sozialpsychologie bekannt ist: „Moral self-licensing“. Mit diesem Begriff ist beschrieben, dass Menschen, nachdem sie sich uneigennützig verhalten haben, daraus die Berechtigung ableiten, das nächste Mal eher an sich denken zu dürfen. Sass et al. (2015) zeigen, dass diese Art der Verhaltensänderung sozial akzeptiert ist.³⁶

Die Frage der externen Validität von Experimenten kann damit nicht allgemein beantwortet werden. Vielmehr kommt es darauf an, wofür ein Experiment benutzt werden soll und welches Entscheidungsproblem auf dem Programm steht. Wenn man mit dem Experiment eine Theorie

³⁵ Dies wird dadurch erreicht, dass nach den einzelnen Durchgängen des Experiments die Versuchsteilnehmer keinerlei Feedback über den Ausgang des Experiments bzw. das Verhalten der anderen Teilnehmer bekommen.

³⁶ Dazu wird das Norm-Aufdeckungsverfahren benutzt, das Krupka und Weber (2010) in die Literatur eingeführt haben.

überprüfen will, die rationale Entscheidungen in einem allgemeinen, d. h. kontextfreien Sinne beschreiben soll, ist das nur recht und billig, auch wenn eine Widerlegung der Theorie dabei herauskommen kann. Der Verweis auf den Laborkontext, für den die Theorie nicht gemacht sei, tut dann wenig zur Sache. Der Experimentator braucht sich um die externe Validität seines Experiments keine Gedanken zu machen.³⁷

Wenn es allerdings nicht um die Überprüfung eines ökonomischen Modells geht, sondern um die Erforschung eines ökonomisch relevanten Phänomens, für das entweder keine Theorie zur Verfügung steht oder die neoklassische Theorie als positive Theorie versagt, dann *muss* man sich zur Übertragbarkeit auf die Realität Gedanken machen. Damit wird die Nähe zur neoklassischen Theorie zu einem entscheidenden Aspekt der experimentellen Forschung.

Sugden (2005) benutzt ein sehr instruktives Beispiel, an dem sich verdeutlichen lässt, wie die Beziehung zwischen Theorie und Experiment in diesem Zusammenhang zu bewerten ist: Die neoklassische Theorie legt nahe, dass Steuerhinterziehung eine Tat ist, die sich als Ergebnis einer Wahl zwischen einem sicheren Verlust (der Steuerzahlung) und einer Lotterie (der Hinterziehung mit gegebener Entdeckungswahrscheinlichkeit) abbilden lässt. Wenn man sich in einem Experiment mit Steuerhinterziehung befassen möchte, muss man die Versuchspersonen deshalb mit einer entsprechenden Wahl konfrontieren. Im Blick auf beide, auf die Theorie wie auch auf das Experiment, stellt sich die Frage der externen Validität genau dann, wenn man den Anspruch erhebt, nicht nur die Wahl zwischen einem sicheren Verlust und einer Lotterie abzubilden, sondern etwas darüber zu sagen, wie sehr Menschen dazu neigen, Steuern zu hinterziehen.

So kommen wir zum Ausgangspunkt unserer Überlegungen zurück, denn damit läuft alles auf die Frage hinaus, wie extern valide das Modell rationaler Entscheidungen ist. Wenn es valide ist, spielt der Kontext einer Entscheidung genauso wenig eine Rolle wie deren moralische Bewertung oder die Existenz sozialer Normen. Wenn sich aber Diskrepanzen zwischen der Modellaussage und dem zeigen, was in der Realität beobachtet werden kann, schlägt die Stunde der wissenschaftlichen Konkurrenz. Diese kann dann – beispielsweise in Gestalt der Psychologie – darauf verweisen, dass soziale Normen bei der Entscheidung über die eigene Steuerehrlichkeit eine wichtige Rolle spielen. Ein großer Vorteil der experimentellen Methode besteht darin, dass sie damit nicht aus dem

³⁷ Sugden (2005), S. 182.

Feld ist. Die experimentellen Ökonomen brauchen ihre Labore nicht zu schließen, wenn klar wird, dass die Modelle, die sie im Rahmen der reinen ökonomischen Forschung bisher untersucht haben, ihrem universellen Erkenntnisanspruch nicht gerecht werden können. Im Gegenteil: statt die Labore zu schließen, öffnen sie diese für die Kollegen aus der Psychologie.

2.7 Verhaltensökonomik: Die Wiederentdeckung der Psychologie

1979 erschien in *Econometrica* ein Artikel, der die ökonomische Wissenschaft in einem Maße beeinflusst hat wie kaum ein anderer. Tatsächlich wurde seit 1970 nur ein anderer Artikel häufiger zitiert³⁸ als „Prospect Theory: An analysis of decision under risk“ von Daniel Kahneman und Amos Tversky. Das ist schon deshalb erstaunlich, weil man es hier nicht mit Ökonomen, sondern mit Psychologen zu tun hat. Wäre Amos Tversky nicht bereits 1996 im Alter von nur 59 Jahren verstorben, hätte sehr wahrscheinlich auch er im Jahr 2002 den Nobelpreis für Ökonomie erhalten, den sich Daniel Kahneman mit Vernon Smith teilte.

Mit einem gewissen Recht kann man die Veröffentlichung der Prospect theory als das Erweckungserlebnis der Verhaltensökonomik bezeichnen. In ihrer ursprünglichen Fassung erklärt diese Theorie die Wahl zwischen Lotterien – und damit die Wahl zwischen mit Risiko behafteten Alternativen – mit Hilfe von vier speziellen Annahmen:³⁹

- 1) Die Bewertung einer Lotterie erfolgt nicht in einem absoluten Sinne, sondern relativ zu einem Referenzpunkt.
- 2) Die Bewertungsfunktion $v(x)$, die der Lotteriauszahlung x eine Bewertung zuordnet, hat einen Knick an der Stelle des Referenzpunktes und ist für Werte, die kleiner sind als der Referenzpunkt, steiler als für Werte, die größer sind. Dieser Verlauf von $v(x)$ ist Ausdruck der sogenannten Verlustaversion.
- 3) $v(x)$ ist für positive Werte konkav und für negative Werte konvex, das heißt die Gewichtung der Lotteriauszahlungen wird umso geringer, je weiter entfernt sie vom Referenzpunkt liegen.
- 4) Entscheider transformieren die Wahrscheinlichkeiten, mit denen die Lotteriergebnisse eintreten, durch eine Gewichtungsfunktion. Infolge dessen werden kleine

Wahrscheinlichkeiten überschätzt und große unterschätzt.

Die Prospect theory erklärt Verhalten unter Verwendung von Annahmen, die nicht a priori gesetzt werden (wie in der Neoklassik die Annahme rationalen Verhaltens), sondern die mit Ergebnissen psychologischer – insbesondere experimenteller – Forschung empirisch begründet sind. Dabei handelt es sich einerseits um Annahmen hinsichtlich der Präferenzen (zum Beispiel die Verlustaversion), andererseits um Annahmen über systematische Verzerrungen bei der Entscheidungsfindung (beispielsweise die verzerrte Wahrnehmung von Wahrscheinlichkeiten). Das entspricht dem Muster verhaltensökonomischer Theorien. Sie setzen an den Stellen an, wo menschliche Entscheidungsprozesse systematisch von einer rationalen Auswahl abweichen.

Menschen benutzen Heuristiken, um komplizierte Entscheidungen zu vereinfachen, und sie unterliegen Verzerrungen. Sie treffen mithin Entscheidungen, die nicht in ihrem wahren Interesse liegen. Die Ursachen und Ausprägungen solcher Verzerrungen sind vielgestaltig. Chetty (2015) und DellaVigna (2009) schlagen eine formale Charakterisierung vor, die im Wesentlichen drei Bereiche identifiziert, in denen es zu Abweichungen vom Rationalmodell kommen kann:

- 1) Die Präferenzordnung weist nicht die Standardeigenschaften auf, die im neoklassischen Modell vorausgesetzt werden. Beispielsweise verstößt die Abhängigkeit der Präferenzen von einem Referenzpunkt gegen die Standardannahmen der Neoklassik. DellaVigna (2009) zählt auch die Existenz sozialer Präferenzen zu den Verletzungen der Standardannahmen, obwohl diese, wie wir gesehen haben, durchaus mit dem neoklassischen Modell in Einklang gebracht werden können.
- 2) Menschen bilden keine rationalen Erwartungen. Beispielsweise überschätzen sie systematisch ihre eigenen Fähigkeiten, was zu zeitinkonsistenten Entscheidungen führt. Die Unfähigkeit, bayesianische Erwartungen zu bilden, hat Phänomene wie die „Gamblers fallacy“ zur Folge. Darunter versteht man, dass jemand irrtümlich annimmt, dass eine Kugel zufällig aus einer Urne gezogen wird, in der die zuvor gezogene Kugel fehlt („ohne Zurücklegen“). Deshalb glaubt auch ein Roulettespieler, der der Gamblers fallacy unterliegt, dass die Wahrscheinlichkeit für „Rot“ höher ist als für „Schwarz“, wenn zuvor mehrfach hintereinander „Schwarz“ fiel.
- 3) Bei gegebenen Erwartungen und Präferenzen kann die Entscheidungsfindung selbst nicht rational erfolgen. Beispielsweise lassen sich Menschen von den Randbe-

³⁸ Vgl. Kim et al. (2006). Häufiger zitiert wurde nur ein sehr technischer *Econometrica*-Artikel von White (1980), in dem es um einen Test auf Heteroskedastizität geht.

³⁹ Vgl. DellaVigna (2009), S. 325.

dingungen, unter denen eine Entscheidung zu treffen ist, beeinflussen: Man spricht dabei von „Framing-Effekten“. Oder sie beachten wichtige Entscheidungsfaktoren nicht oder zu wenig, weil sie unachtsam sind oder sich von sozialem Druck beeinflussen lassen.

Für die politischen Implikationen, die sich aus entscheidungstheoretischen Ansätzen ableiten lassen, ist ein Aspekt sehr wichtig, anhand dessen die Unterschiede zur neoklassischen Modellierung sehr deutlich werden. Während im neoklassischen Modell die Präferenzordnung des Entscheiders durch eine Nutzenfunktion wahrheitsgemäß abgebildet wird, deren Maximierung dann letztlich die rationale Entscheidung liefert, unterscheidet man in der Verhaltensökonomik zwischen dem Erfahrungsnutzen $u(c)$ und dem Entscheidungsnutzen $v(c)$.⁴⁰ Ersterer ist der Nutzen, den der Entscheider tatsächlich erfährt, wenn seine Entscheidung wirksam wird. Letzterer ist der Nutzen, von dem der Entscheider zum Zeitpunkt der Entscheidung glaubt, es sei sein wahrer Nutzen. Während in der Neoklassik davon ausgegangen wird, dass $u(c) = v(c)$, lässt die Verhaltensökonomik zu, dass beide voneinander abweichen. In Analogie zu einem externen Effekt lässt sich $e(c) = u(c) - v(c)$ als *Internalität* interpretieren, als einen Keil, den der Entscheider selbst zwischen seinen wahren Nutzen und jenen Nutzen treibt, den er der Entscheidung zugrunde legt.⁴¹ Eine Politikimplikation, die aus einer Internalität folgt, ist die gleiche wie die im Falle eines externen Effekts: Durch staatliche Intervention kann es möglich sein, eine Wohlfahrtsverbesserung zu erreichen, indem man quasi den Keil zwischen den beiden Nutzen entfernt.

Ein Beispiel für eine Verzerrung, die zu einer Differenz zwischen Erfahrungsnutzen und Entscheidungsnutzen führt, ist mangelnde Selbstkontrolle, die zeitinkonsistente Präferenzen zur Folge haben kann. O'Donoghue und Rabin (1999) haben eine relativ einfache Theorie entwickelt, mit der sie in der Lage sind, dieses Phänomen zu beschreiben. Im Kern besteht die Erklärung darin, dass die Diskontierung zwischen der Gegenwart und einem künftigen Zeitpunkt immer höher ist als zwischen zwei künftigen Zeitpunkten. Das führt zu einer Überbewertung der Gegenwart, die für Investitionsentscheidungen aller Art weitreichende Folgen haben kann. Bei solchen Entscheidungen kommt es in der Gegenwart zu einer negativen Auszahlung. Beispielsweise muss man auf Konsum verzichten, unter Opfern das Rauchen aufgeben oder sich im Fitnessclub für die eigene

Gesundheit abstrampeln. Die Erträge fallen dagegen erst später an. Die Überbewertung der in der Gegenwart zu tragenden Kosten führt dazu, dass man dazu neigt, die Investition aufzuschieben, weil der im Augenblick notwendige Nutzenverlust zu schwer wiegt. Die Folge ist, dass bereits geringe Opportunitätskosten in der Gegenwart ausreichen können, um eine Entscheidung immer wieder zu vertagen, die zu fällen ohne die verzerrende Bewertung des Gegenwartskonsums rational wäre. Dies erklärt einen Effekt, der aus der Sicht des neoklassischen Rationalmodells nicht auftreten dürfte, der sich aber als empirisch äußerst stabil erwiesen hat.

Madrian und Shea (2001) haben erstmalig darauf hingewiesen, dass die Ausgestaltung der „Voreinstellung“ (Default-Lösung) bei der Entscheidung über die Teilnahme an Maßnahmen der betrieblichen Altersvorsorge in den Vereinigten Staaten einen gravierenden Einfluss besitzt. Wenn die Default-Lösung darin besteht, dass ein Beschäftigter keinen Rentenvertrag abschließt⁴², liegt die Partizipationsrate bei 20 Prozent. Wenn aber die Default-Lösung umgekehrt wird und ein Beschäftigter aktiv entscheiden muss, wenn er nicht teilnehmen will, steigt sie auf 80 Prozent. Ein ähnlicher Effekt tritt ein, wenn die Mitarbeiter gezwungen werden, aktiv eine Entscheidung zu treffen. Dies gilt jedenfalls dann, wenn diese Entscheidung von einem persönlichen Beratungsgespräch begleitet wird, wie Benarzi und Thaler (2004) festgestellt haben. Das Modell von O'Donoghue und Rabin (1999) liefert eine Erklärung für diesen Default-Effekt. Es reichen bereits die geringen Opportunitätskosten der Beitrittserklärung⁴³, um zu verhindern, dass die Beschäftigten die eigentlich vorteilhafte und durchaus auch ihren Präferenzen entsprechende Entscheidung für den Sparvertrag fällen.

Das Beispiel macht deutlich, dass verhaltensökonomische Erklärungen den Weg für neuartige Politikinstrumente öffnen können. Chetty et al. (2014) konnten anhand dänischer Daten zeigen, dass der Default-Effekt nicht dazu führt, dass diejenigen, die sich nach der Umstellung des Defaults für den Sparvertrag entscheiden, an anderer Stelle ihre Ersparnisse zurückfahren. Tatsächlich kann durch einfache Umstellung der Default-Regel die Nettoersparnis deutlich angehoben werden. Das Beispiel macht zugleich deutlich, wie Nudging im Idealfall wirken kann. Menschen, die eine Internalität aufweisen, die sie daran hin-

⁴⁰ Dabei darf $v(c)$ nicht mit der Bewertungsfunktion der Prospect theory verwechselt werden. Die Variable c kann für Konsum stehen oder für einen anderen Entscheidungsgegenstand.

⁴¹ Vgl. Chetty (2015), S. 9.

⁴² Dabei handelt es sich um sogenannte 401(k)-Verträge, die steuerlich gefördert werden und bei denen der Arbeitgeber in der Regel einen Zuschuss zu den Sparbeträgen zahlt.

⁴³ Üblicherweise werden 3 Prozent des Einkommens einbehalten und in den Sparplan eingezahlt, zuzüglich eines Zuschusses vom Arbeitgeber.

dert, den eigentlich vorteilhaften Sparvertrag abzuschließen, werden durch die geänderte Default-Regel zu einer Entscheidung „geschubst“, die in ihrem eigenen wohlverstandenen Interesse liegt. Wer frei von einer solchen Internalität ist und sich gegen den Sparvertrag entscheidet, wird nicht daran gehindert, an dieser Entscheidung festzuhalten. Der Nudge verzerrt seine Entscheidung nicht, doch er korrigiert die Fehlentscheidung jener Personen, deren Präferenzen verzerrt sind. Mit dem Nudge ist die Politik damit auf der sicheren Seite. Sie muss sich nicht dem Vorwurf aussetzen, paternalistisch zu handeln, sorgt aber dennoch dafür, dass wohlfahrtsreduzierende Verzerrungen korrigiert werden.

Das Beispiel zeigt, dass die Verhaltensökonomik unmittelbare Implikationen auch für die deutsche Politik haben kann. Die Einführung der gesetzlichen Rentenversicherung und des damit einhergehenden Versicherungszwangs kann man zwar nicht als Nudge bezeichnen, aber die Rentenreformen der vergangenen Jahre bieten durchaus Ansatzpunkte für Nudging. Beispielsweise könnte die Entscheidung, ob eine private kapitalgedeckte Rentenversorgung neben der gesetzlichen aufgebaut werden soll (Stichwort: Riester-Rente), mit einem Entscheidungszwang versehen werden, der jeden Beschäftigten verpflichtet, sich aktiv für oder gegen eine solche Versicherung zu entscheiden. Man kann mit hoher Wahrscheinlichkeit davon ausgehen, dass sich ähnliche Effekte zeigen würden wie in den Vereinigten Staaten – vor allem wenn das von Benarzi und Thaler (2004) vorgeschlagene SMT-Programm aufgegriffen würde. SMT steht für „Save more tomorrow“. Dabei wird ein Sparvertrag abgeschlossen, bei dem Anteile zukünftiger Lohnerhöhungen eingezahlt werden. Das hat zur Folge, dass bei Abschluss dieses Vertrages keine Verluste in Form eines Konsumverzichts entstehen. Auch dieser „Trick“ ist verhaltensökonomisch begründet. Er dient dazu, die Verlustaversion zu überwinden, die Menschen davon abhält, heute auf etwas zu verzichten, um in der Zukunft mehr konsumieren zu können.

Nicht immer sind die politischen und wohlfahrtsökonomischen Implikationen verhaltensökonomischer Erkenntnisse indes so klar und unproblematisch wie im Fall des Default-Effekts. Beispielsweise zeigen Blaufuß et al. (2013) und Fochmann und Weimann (2013) in einem Laborexperiment, dass die Besteuerung von Arbeitseinkommen verzerrt wahrgenommen wird. Insbesondere bei einfachen linearen Steuertarifen verwechseln die Besteuerten Brutto- und Nettoeinkommen, was dazu führt, dass die Besteuerung bei konstantem Nettolohn zu steigenden Arbeitsanstrengungen führt. Welche Implikationen ergäben sich, wenn sich dergleichen nicht nur im Labor, sondern auch im Feld beobachten ließe? Auf der einen Seite könnte

man argumentieren, dass der richtige Nudge darin bestünde, die Besteuerten besser darüber aufzuklären, wie hoch ihre Nettoeinkünfte tatsächlich sind, um die Internalität zu korrigieren, die sich in der höheren Arbeitsleistung ausdrückt. Auf der anderen Seite hat diese Internalität einen durchaus positiven Effekt, denn sie reduziert die Zusatzlast der Besteuerung, wirkt also wohlfahrtssteigernd. Sollte man mithin die Menschen ihrem Fehler überlassen, weil sich damit Zusatzlasten sparen lassen?

In dem zuletzt genannten Beispiel lässt sich der verhaltensökonomische Befund einer verzerrten Wahrnehmung von Steuern bisher lediglich im Labor erheben. Seine externe Validität ist damit nicht vollständig gesichert. Allerdings wurden die Experimente mit Beschäftigten durchgeführt (nicht mit Studenten), also mit Menschen, die es gewohnt sind, Einkommen zu erzielen, das der Besteuerung unterliegt. Es spricht deshalb einiges dafür, dass sich im Feld zumindest ähnliche Effekte zeigen dürften. Ganz allgemein gilt für die Verhaltensökonomik jedoch, dass sie der Vorwurf mangelnder externer Validität nicht trifft. Denn die Sammlung von Feldbeobachtungen, die geeignet sind, Laborbefunde zu bestätigen, ist mittlerweile beeindruckend groß.

So konnten DellaVigna und Malmendier (2006) zeigen, dass mangelnde Selbstkontrolle beziehungsweise falsche Erwartungen über das eigene künftige Verhalten die Ursache dafür sind, dass Fitnessclubs überwiegend Verträge mit festen Laufzeiten und einer Flatrate verkaufen, obwohl für die meisten Mitglieder gilt, dass sie wesentlich preisgünstiger trainieren würden, wenn sie jeden Besuch einzeln bezahlten. Die Erwartung, wie oft man in der Zukunft trainieren wird, ist häufig falsch. Fast immer wird die künftige Bereitschaft, die Kosten des Trainings zu tragen, überschätzt.⁴⁴ Eine falsche Einschätzung der eigenen Fähigkeiten scheint überhaupt ein verbreitetes Phänomen zu sein. Zumindest bei Aufgaben, die den Probanden einfach erscheinen, überschätzen sich sowohl Studenten als auch Firmenchefs (Camerer und Lavallo 1999). Wenn sie die Aufgaben dagegen als schwierig wahrnehmen, neigen die Menschen zur Unterschätzung der eigenen Fähigkeiten (Kruger 1999).

Read und van Leeuwen (1998) haben in einem Feldversuch gezeigt, dass Menschen dazu neigen, ihre Gegenwartspräferenz in die Zukunft fortzuschreiben, ohne dabei zu beachten, dass sich die Umstände verändern können.

⁴⁴ Auch dieses Phänomen ist Bestandteil der Theorie von O'Donoghue und Rabin (1999). „Naive“ Menschen gehen davon aus, dass die Diskontierung zwischen zwei Zeitpunkten in der Zukunft konstant ist, sich also auch dann nicht ändert, wenn der erste der beiden Zeitpunkte zur Gegenwart geworden ist.

Dazu haben sie Beschäftigte aufgefordert, eine Bestellung für ein Essen aufzugeben, das eine Woche später stattfinden soll. Diejenigen, die dies am späten Nachmittag taten (und dabei eher hungrig waren), bestellten signifikant mehr Speisen als jene Kollegen, welche die gleiche Aufgabe direkt nach dem Mittagessen erfüllten. Colin et al. (2007) haben diesen „Projection bias“ nachweisen können, als sie die Onlinebestellungen von warmer Kleidung untersuchten. Für die Bestellungen, die an kalten Tagen aufgegeben wurden, zeigte sich eine signifikant höhere Retourenquote als für Bestellungen, die an eher warmen Tagen aufgegeben wurden.

Besonders reichhaltig ist inzwischen die Literatur zur „Inattention“, also dazu, dass wichtige Entscheidungsparameter keine Beachtung finden. Chetty et al. (2009) stellten fest, dass Supermarkt-Kunden Steuern, die nicht auf den Preisschildern ausgewiesen wurden, systematisch ignorierten, obwohl ihnen grundsätzlich bekannt war, dass diese an der Kasse auf den Preis aufgeschlagen werden. Die Autoren schätzen den „Unaufmerksamkeitsparameter“ auf 0,94, wobei der Wert 1 einer vollständigen Nichtbeachtung entspricht.

Dass sozialer Druck Entscheidungen beeinflusst, konnten Garicano et al. (2005) in einer Untersuchung des Verhaltens von Fußballschiedsrichtern feststellen. Sie beobachteten die Entscheidung darüber, wie lange nachzuspielen ist. Wenn die Heimmannschaft ein Tor zurück lag, fiel die Nachspielzeit doppelt so lang aus wie bei einem Vorsprung von einem Tor. Mas und Moretti (2009) zeigten, dass sozialer Druck auch am Arbeitsplatz wirksam ist. So arbeiteten Kassiererinnen in einem Supermarkt deutlich effizienter, wenn sie wussten, dass hinter ihnen eine Kollegin saß, von der bekannt war, dass sie sehr schnell arbeitet. Dass auch Emotionen bei wichtigen Entscheidungen eine Rolle spielen, zeigt Simonsohn (2010). Er wies nach, dass Studenten, die eine Elite-Universität besichtigt hatten, sich dort mit höherer Wahrscheinlichkeit einschrieben, wenn am Besichtigungstag schlechtes Wetter herrschte. Denn dann galt ihre Aufmerksamkeit mehr den akademischen Angeboten als den Möglichkeiten der Freizeitgestaltung an der Universität.

Die Liste ließe sich fortsetzen.⁴⁵ Nur eine Beobachtung, die schon früh im Labor gemacht wurde und die besonders eindrucksvoll die Prospect theory zu bestätigen scheint, sei noch genannt: der „Ausstattungseffekt“ (Endowment effect). Kahnemann et al. (1990) verteilten an die Hälfte ihrer Versuchspersonen Kaffeebecher mit einem

Universitätslogo. Danach wurde anreizkompatibel erhoben, wie hoch die Zahlungsbereitschaft (Willingness to pay, WTP) für einen solchen Becher bei denen war, die keinen erhalten hatten, und wie hoch der Mindestpreis war, von dem an die Becherbesitzer bereit waren, ihren Becher wieder zu verkaufen (Willingness to accept, WTA). Es zeigte sich, dass die WTA doppelt so hoch ausfiel wie die WTP, obwohl sie aufgrund der zufälligen Zuweisung der Versuchspersonen in die beiden Gruppen eigentlich gleich hoch hätte sein müssen. Die Prospect theory kann die massive Diskrepanz zwischen WTA und WTP erklären, weil sich durch die Zuteilung des Bechers der Referenzpunkt ändert und die Verlustaversion die Bewertung erhöht.

List (2003, 2004) sowie Plott und Zeiler (2005) sind allerdings in Feldversuchen darauf gestoßen, dass der Ausstattungseffekt nicht stabil ist. Das zentrale Ergebnis der Arbeiten von List ist, dass der Ausstattungseffekt verschwindet, wenn die Verkäufer Erfahrungen sammeln. Bei professionellen Händlern war dieser Effekt nicht mehr nachweisbar. Das hat eine wichtige Implikation: Die in der Verhaltensökonomik thematisierten Verzerrungen können unter Bedingungen wiederholter Interaktionen verschwinden. Das schränkt die Anwendbarkeit der Verhaltensökonomik für die Beschreibung von Märkten in einem gewissen Sinne ein, denn viele Märkte sind Orte, auf denen solche wiederholten Interaktionen stattfinden.

Damit schließt sich der Kreis, denn Paretos ursprüngliches Programm bestand ja gerade in der Analyse von Markttransaktionen, die den Akteuren genug Zeit und Raum geben, rationales Verhalten zu lernen. Die moderne Verhaltensökonomik verweist implizit wieder auf diese alte Arbeitsteilung. Verzerrungen und Heuristiken dürften vor allem dort eine wichtige Rolle spielen, wo kein oder nur eingeschränktes Lernen in der für rationales Verhalten notwendigen Form stattfinden kann. Wo es stattfindet, kommt die Theorie der rationalen Entscheidung wieder zu ihrem Recht. Der zweite Bereich, in dem das um die Annahme eigennützigem Verhaltens erweiterte Rationalmodell in Schwierigkeiten gerät, besteht aus Situationen, in denen der Eigennutz mit sozialen Normen in Konflikt gerät. Aber hier kann man eine rationale Lösung herbeiführen, in dem man neue Elemente in die Nutzenfunktion integriert – sei es Pflichtgefühl, Altruismus oder etwas Ähnliches.

Zusammenfassend sei festgehalten, dass die Existenz von systematischen Verzerrungen individueller Entscheidungen und die Verwendung von Heuristiken als empirisch im Feld nachgewiesene Tatbestände gelten können. Welche Bedeutung sie im Hinblick auf ökonomisch relevante Phänomene haben und welche Wohlfahrtsimplika-

⁴⁵ Der interessierte Leser sei auf die umfangreiche Fallsammlung in DellaVigna (2009) verwiesen.

tionen sich aus ihnen ergeben, muss jeweils im Einzelfall geklärt werden. Angesichts des schon jetzt gelungenen Nachweises, dass verhaltensökonomisch begründete politische Interventionen erhebliche Wirkungen haben können, wäre es kaum zu rechtfertigen, der Verhaltensökonomik in der Politikberatung (zu) wenig Beachtung zu schenken. Das Gleiche gilt für die experimentelle Forschung. Die Relevanz von Feldversuchen und natürlichen Experimenten ist offensichtlich, doch auch für das Laborexperiment kann man diese Relevanz beanspruchen – nicht zuletzt aus methodischen Gründen.

3 Neusortierung des ökonomischen Instrumentenkastens

Zu Beginn dieses Abschnitts sei der Versuch unternommen, drei Begriffe noch einmal gegeneinander abzugrenzen, bei denen die Gefahr besteht, dass sie fälschlicherweise als sehr ähnlich angesehen werden. Es geht dabei um den Bezug der experimentellen Methode zur Verhaltensökonomik und um die Frage, welche Rolle soziale Präferenzen spielen. In einem Gedankenexperiment sei einmal angenommen, dass Laborexperimente eine vollständige interne Validität aufweisen, dass sich alle Anreize, die in der Theorie angenommen werden, 1:1 ins Labor übertragen lassen und dass die Menschen tatsächlich rationale Entscheidungen treffen. In diesem Fall muss jedes korrekt durchgeführte Experiment jede neoklassisch konzipierte Theorie, die es testen soll, vollständig bestätigen.

Wenn dies nicht der Fall ist, wenn es also zu Abweichungen des Laborverhaltens von der theoretischen Prognose kommt (bei vollständiger interner Validität), dann bleiben zwei mögliche Erklärungen dafür: (1) Die Versuchspersonen verhalten sich rational, verfolgen aber noch andere Ziele, als man in der Modellbildung angenommen hat. Das heißt, die tatsächlichen Nutzenfunktionen der Versuchspersonen enthalten Argumente, die nicht in der Auszahlungsfunktion des Experiments enthalten sind. Soziale Präferenzen sind eine mögliche Abweichung dieser Art. Oder (2) die Versuchspersonen verhalten sich nicht rational, weil sie Verzerrungen unterliegen, also $u(c) \neq v(c)$, oder weil sie Heuristiken benutzen, die nichtrationale Entscheidungen hervorrufen.

Die experimentelle Methode ist zunächst einmal nur dazu da, Abweichungen per se festzustellen. Laborexperimente sind dafür besonders gut geeignet, weil sie die Möglichkeit eröffnen, Theorien exakt zu testen. Diese Möglichkeit ist ihnen gegeben, weil durch den Vergleich von verschiedenen Anordnungen („Treatments“) (bei randomi-

erter Zuweisung der Versuchspersonen zu den Treatments) kausale Aussagen möglich sind, die denen eines formalen Modells nahe kommen. Die Frage, ob eine Abweichung der theoretischen Prognose von dem experimentellen Befund auf (1) oder (2) zurückzuführen ist, ist nachrangig zu klären. Läuft es auf (1) hinaus, können Wissenschaftler für die weitere Forschung das Standardmodell rationaler Entscheidung ohne Einschränkung benutzen, eine Konsultation psychologischer Expertise ist nicht notwendig. Liefert hingegen (2) die bessere Erklärung der Daten, führt die experimentelle Analyse zur Verhaltensökonomik.

Sehr ähnlich verhält es sich mit Experimenten, mit denen nicht explizit ein ökonomisches Modell getestet wird – beispielsweise weil bereits bewiesen ist, dass das entsprechende Modell die experimentellen Daten nicht abbilden kann. Explorative Experimente können in diesem Fall dazu beitragen, eine Entscheidung darüber herbeizuführen, ob ein deskriptiv erfolgreicherer Modell eher im Bereich der Verhaltensökonomik zu suchen ist oder ob es unter Verwendung alternativer Präferenzen konstruiert werden kann. Auch diese Fähigkeit besitzen Laborexperimente aus dem Grund, dass sie kausale Zusammenhänge aufdecken können. Die Frage der externen Validität des Laborexperiments ist nachgeordnet. Sie kann nur fallweise geklärt werden, im Ideal durch ein entsprechend konstruiertes Feldexperiment.

Bei Gültigkeit dieser Abgrenzung stellt sich die Frage, welche Rolle eine so verstandene experimentelle Forschung in der Wirtschaftswissenschaft im Allgemeinen und in der Politikberatung im Besonderen spielen kann. Im Hinblick auf die Grundlagenforschung dürfte sie leicht zu beantworten sein; die Antwort steckt in der soeben formulierten Abgrenzung zur Verhaltensökonomik. Experimente sind unverzichtbare Werkzeuge für die Überprüfung von Theorien. In dieser Funktion sind sie unabhängig von ihrer externen Validität. Spannender ist dagegen die Frage, welche Rolle sie in der Politikberatung spielen können.

Damit sich eine wissenschaftliche Methode dafür eignet, in der evidenzbasierten Politikberatung eingesetzt zu werden, muss sie über drei wichtige Fähigkeiten verfügen: (1) die Fähigkeit, kausale Zusammenhänge nachzuweisen, (2) die Fähigkeit zur externen Validität und (3) die möglichst breite und universelle Anwendbarkeit. Zwischen Kausalität und externer Validität besteht dabei ein ausgeprägter Trade-off.⁴⁶ Je eher eine Methode geeignet ist, Kausalität nachzuweisen, umso weniger können ihre Er-

⁴⁶ Vgl. dazu auch Weimann (2015).

gebnisse externe Validität reklamieren. Der Grund ist sehr einfach: Kausalität lässt sich am besten nachweisen, wenn die für einen ökonomischen Zusammenhang wichtigen Elemente isoliert werden können. Dazu ist eine Abstraktion notwendig, die mitunter sehr weit geht. Je abstrakter aber ein Modell ist, umso größer ist auch die Wahrscheinlichkeit, dass es in der Realität Einflüsse gibt, die in diesem Modell nicht abgebildet sind. Wie sich gleich zeigen wird, weisen die Methoden, die am ehesten geeignet sind, den Trade-off zwischen Kausalität und externer Validität aufzuheben, Schwächen in der dritten Dimension – der breiten Anwendbarkeit – auf, weil sie nur in ausgewählten Fällen herangezogen werden können.

Wenn man die verschiedenen Methoden, die Ökonomen zur Verfügung stehen, um wissenschaftliche Evidenz zu erzeugen, gedanklich in einem Raum einsortiert, den die Dimensionen Kausalität, externe Validität und Anwendbarkeit aufspannen, dann ergibt sich folgendes Bild:

- 1) Die *neoklassische Theorie* ist perfekt geeignet, kausale Zusammenhänge zu bestimmen. Sie ist praktisch unbeschränkt anwendbar, weist aber eine stark eingeschränkte externe Validität auf.
- 2) Die *empirische Wirtschaftsforschung*, die ebenfalls über ein breites Anwendungsspektrum verfügt, besteht über weite Strecken im Nachweis von Korrelationen. Der Nachweis von Kausalitäten scheitert hier häufig, weil die auftretenden Identifikationsprobleme aufgrund eingeschränkter Datenverfügbarkeit nicht lösbar sind.
- 3) Die Vertreter der *empirischen Evaluationsforschung* wiederum bemühen sich um kausale Analysen und haben in den vergangenen Jahren enorme methodische Fortschritte erzielt. Aber sie sind darauf angewiesen, dass sich natürliche Experimente finden lassen, die es erlauben, eine Kontrollgruppe zu konstruieren – also eine Gruppe von Menschen, die sich von einer anderen nur darin unterscheidet, dass ihr keine spezifische „Behandlung“ zugekommen ist. Das schränkt die Anwendbarkeit dieser Methode erheblich ein.
- 4) Der *randomisierte Feldversuch* ist als Königsweg ebenfalls geeignet, den Trade-off zwischen Kausalität und externer Validität aufzuheben. Doch er ist nur in relativ seltenen Fällen zu realisieren und besitzt deshalb einen noch weiter eingeschränkten Anwendungsbereich.
- 5) Wo es um die Evaluation von institutionellen Regelungen geht, die erst in der Zukunft eingerichtet werden sollen, für die es also per se keine Daten geben kann, ist empirische Forschung gänzlich aussichtslos.
- 6) Die *Laborökonomik* bewegt sich zwischen empirischer und theoretischer Forschung. Mit ihr ist es möglich, Kausalitäten mit fast derselben Präzision nachzuweisen wie die Theorie – das gilt zumindest für Treatment-Effekte, deren Robustheit in reproduzierten Experimenten nachgewiesen worden ist. Auch hinsichtlich des Anwendungsbereiches nimmt es das Laborexperiment sowohl mit der theoretischen als auch mit der empirischen Forschung auf. Hinsichtlich der externen Validität ist es der empirischen Forschung und erst recht dem Feldexperiment unterlegen, es weist aber eine größere externe Validität auf als das formale Modell.

Dieser systematische Überblick über die Methodenlandschaft der Wirtschaftswissenschaft macht deutlich, dass es die eine ideale Methodik der ökonomischen Forschung nicht gibt. Chetty (2015) schlägt vor, bei der Entscheidung, ob man empirische Daten eher mit einem neoklassischen Modell oder mit einem verhaltensökonomischen Ansatz erklärt, pragmatisch vorzugehen. Sein Argument ist, dass a priori nicht klar sei, an welchen Stellen welcher Ansatz tatsächlich mehr verspreche. Die alte Idee, der viele Theoretiker anhängen, man könne letztlich alles mit dem Rationalmodell erklären, ist von der wissenschaftlichen Evidenz längst überrollt worden. Daher ist der pragmatische Weg, den Chetty vorschlägt, letztlich der einzig mögliche.

Doch dieser Pragmatismus sollte sich nicht nur auf die Entscheidung zwischen verhaltensökonomischer oder neoklassischer Modellierung beziehen, sondern auch auf die Wahl der Methode, mit der man ein Erkenntnisproblem angeht. Ob man einen modelltheoretischen Zugang wählt, einen empirischen, einen experimentellen oder einen quasi-experimentellen, sollte man von den konkreten Bedingungen und Anliegen abhängig machen, beispielsweise davon, ob es bereits vorhandene oder erst neu zu schaffende Institutionen zu evaluieren gilt. Auch die Verfügbarkeit von Daten muss eine Rolle spielen, und es ist zu klären, ob die Möglichkeit besteht, ein Feldexperiment durchzuführen oder ein natürliches Experiment zu nutzen. Bei allen diesen Abwägungen ist die Wahl einer Methode keine Entscheidung gegen eine andere. Vielmehr kann man komplementäre Methoden gewinnbringend kombinieren.

Wie dies gehen kann, soll ein abschließendes Beispiel verdeutlichen. Eine in der klimapolitischen Debatte wichtige Frage betrifft die Wirkungen einer Vorreiterrolle eines Landes oder einer Gruppe von Ländern bei internationalen Klimaverhandlungen. Hilft es, wenn Deutschland mit gutem Beispiel vorangeht, oder ist ein solcher Alleingang eher schädlich? Aus nahe liegenden Gründen ist es nicht möglich, diese Frage empirisch zu beantworten. Was

bleibt, sind die mathematische Modellierung einer Verhandlungssituation und das Experiment. Konrad und Thum (2014) haben dazu ein Zwei-Länder-Modell vorgeschlagen, in dem die Länder asymmetrische Informationen besitzen. Jedes Land kennt nur die eigenen Kosten, die entstehen, wenn es Klimaschutz betreibt. Das Modellergebnis ist, dass Verhandlungen, die unter der Voraussetzung geführt werden, dass eines der beiden Länder bereits eine Vorleistung erbracht hat, mit geringerer Wahrscheinlichkeit zu einer effizienten Verhandlungslösung führen als Verhandlungen, die ohne Vorleistung erfolgen.

Das Modell, das Konrad und Thum benutzen, ist ein neoklassisches Modell, in dem beide Akteure vollständig rational handeln, strategisch denken und ausschließlich die Maximierung der eigenen Auszahlung anstreben. Verhaltensökonomische Erkenntnisse spielen dabei ebenso wenig eine Rolle wie soziale Präferenzen. Hoffmann et al. (2015) haben das Modell experimentell getestet. Dabei wurde den Versuchspersonen Schritt für Schritt immer mehr Freiraum gegeben, der es ihnen erlaubte, von den Modellprognosen abzuweichen. Beispielsweise konnten sie im letzten Schritt (anonym) miteinander kommunizieren. Der experimentelle Befund bestand zum einen darin, dass die Versuchspersonen zahlreiche Abweichungen von rationalem Verhalten zeigten. Es kam zu falschen Erwartungen über das Verhalten des anderen Spielers; es wurden Ungleichheitsaversion und Reziprozität beobachtet; die Kommunikation, die vollständig den Charakter von „Cheap talk“ hatte, beeinflusste das Verhalten. Andererseits aber zeigte sich, dass das zentrale Modellergebnis von Konrad und Thum in allen Anordnungen bestätigt werden konnte. Stets war die Anzahl der erfolgreichen Verhandlungen größer, wenn zuvor *keine* Vorleistung erbracht wurde.⁴⁷

Das Beispiel zeigt, wie Theorie und Experiment zusammenwirken können, um gemeinsam Evidenz zu erzeugen. Die Theorie zeigt zunächst ein überraschendes und kontraintuitives Resultat. Sie benutzt dabei einschränkende Annahmen indem sie nur Rationalität der Akteure zulässt. Das Experiment öffnet den Akteuren den Freiraum, davon abzuweichen, und bestätigt dennoch das Modellergebnis. Auf diese Weise stärkt ein Experiment, obwohl es ein neoklassisches Modell mit verhaltensökonomischen „Anomalien“ konfrontiert, genau dieses Modell, und es stützt die Intuition, die das Modell erzeugt hat.

⁴⁷ Und bis auf eine Ausnahme waren die Unterschiede jeweils signifikant.

4 Fazit: Die Pluralität der Methoden ist eine Stärke der ökonomischen Disziplin

Die ökonomische Methodenlandschaft hat in den vergangenen dreißig Jahren erheblich an Diversität gewonnen. Ein Methodenpluralismus, wie er heute herrscht und sich in den führenden Journalen der Wissenschaft zeigt, wäre vor 1970 und auch einige Zeit danach nicht möglich gewesen. Dabei haben die Psychologie und der Einzug experimenteller Methoden eine wesentliche Rolle gespielt. Das damit verbundene Aufkommen der Verhaltensökonomik sollte nicht als Versuch einer wissenschaftlichen Revolution im Sinne von Thomas S. Kuhn verstanden werden. Kuhnscher Fortschritt bestünde darin, dass ein herrschendes Paradigma (in diesem Falle die Neoklassik) durch ein anderes ersetzt wird. Genau das aber sollte nicht geschehen.

Deshalb ist es notwendig, die Pluralität der Methoden und der theoretischen Ansätze nicht als eine Schwäche, sondern als eine Stärke der Disziplin zu begreifen. So wie Experimente sich als ein komplementäres Element methodisch längst bewährt haben, so kann sich die Verhaltensökonomik als komplementärer Erklärungsansatz für menschliches Verhalten etablieren und lässt sich fruchtbar mit der Neoklassik kombinieren. Wenn es gelingt, dieses Bewusstsein in der Profession zu verankern, dann ist auch die Integration der experimentellen und verhaltensökonomischen Forschung in die Politikberatung ein notwendiger und konsequenter Schritt, der früher oder später auch in Deutschland vollzogen wird.

Mitunter hört man die Klage, dass es die Wirtschaftswissenschaft an Pluralität vermissen lasse. Wer diese Klage erhebt, sollte sich bewusst machen, wie vielfältig, vieltalig, schillernd und faszinierend ökonomische Forschung inzwischen geworden ist. Dass sie so werden konnte, spricht uneingeschränkt für diese Disziplin.

Danksagung: Friedrich Breyer, Sönke Hoffmann, Markus Sass und Florian Timme sei für wertvolle Hinweise gedankt. Außerdem geht mein Dank an Karen Horn für ihre wertvolle Unterstützung.

Literatur

- Andersen, S., S. Ertac, U. Gneezy, M. Hoffman und J.A. List (2011), Stakes matter in ultimatum games, *American Economic Review* 101(7), S. 3427–39.
- Bardsley, N. (2008), Dictator game giving: altruism or artefact?, *Experimental Economics* 11(2), S. 122–33.

- Barmettler, F., E. Fehr und Ch. Zehnder (2012), Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory, *Games and Economic Behavior* 75(1), S. 17–34.
- Benartzi, S. und R.H. Thaler (2004), Save more tomorrow: Using behavioral economics to increase employee saving, *Journal of Political Economy* 112(1), S. 164–87.
- Berg, J., J. Dickhaut und K. McCabe (1995), Trust, reciprocity, and social history, *Games and Economic Behavior* 10(1), S. 122–42.
- Binmore, K., A. Shaked und J. Sutton (1985), Testing noncooperative bargaining theory: A preliminary study, *American Economic Review* 75(5), S. 1178–80.
- Binmore, K. und A. Shaked (2010), Experimental economics: Where next?, *Journal of Economic Behavior & Organization* 73(1), S. 87–100.
- Blaufuß, K., M. Fochmann, J. Hundsdörfer, D. Kiesewetter und J. Weimann (2013), Net wage illusion in a real effort experiment, *Scandinavian Journal of Economics* 115(2), S. 476–84.
- Bolton, G.E., R. Zwick und E. Katok (1998), Dictator game giving: Rules of fairness versus acts of kindness, *International Journal of Game Theory* 27(2), S. 269–99.
- Bolton, G.E. und A. Ockenfels (2000), ERC: A theory of equity, reciprocity and competition, *American Economic Review* 90(1), S. 166–93.
- Boockmann, B., C. Buch und M. Schnitzer (2014), Evidenzbasierte Wirtschaftspolitik in Deutschland: Defizite und Potentiale, *Perspektiven der Wirtschaftspolitik* 15(4), S. 307–23.
- Brosig-Koch, J., H. Hennig-Schmidt, N. Kairies-Schwarz und D. Wiesen (2015): Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision, erscheint in: *Journal of Economic Behavior & Organization*.
- Brosig, J., J. Weimann und C.L. Yang (2004), Communication, reputation and punishment in sequential bargaining experiments, *Journal of Institutional and Theoretical Economics* 160(4), S. 576–606.
- Brosig, J., J. Weimann und A. Ockenfels (2003), The effect of communication media on cooperation, *German Economic Review* 4(2), S. 217–41.
- Bruttel, L.V. und F. Stolley (2014), Ist es im Interesse der Bürger, wenn ihre Regierung Nudges implementiert?, *Wirtschaftsdienst* 94(11), S. 767–69.
- Bruni, L. und R. Sugden (2007), The road not taken: How psychology was removed from economics, and how it might be brought back, *The Economic Journal* 117(1), S. 146–73.
- Camerer, C.F. und R.M. Hogarth (1999), The effects of financial incentives in experiments: A review and capital-labor-production framework, *Journal of Risk and Uncertainty* 19(1–3), S. 7–42.
- Camerer, C.F. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, Princeton.
- Camerer, C.F. und D. Lovallo (1999), Overconfidence and excess entry: An experimental approach, *American Economic Review*, 89(1), S. 306–18.
- Carpenter J., E. Verhoogen und S. Burks (2005), The effect of stakes in distribution experiments, *Economics Letters* 86(3), S. 393–98.
- Chamberlin, E. (1948), An experimental imperfect market, *Journal of Political Economy* 56(2), S. 95–108.
- Charness, G. und M. Rabin (2002), Understanding social preferences with simple tests, *Quarterly Journal of Economics* 117(3), S. 817–69.
- Cherry, T.L., P. Frykblom und J.F. Shogren (2002), Hardnose the dictator, *American Economic Review* 92(4), S. 1218–21.
- Chetty, R. (2015), Behavioral economics and public policy: A pragmatic perspective, Discussion Paper Harvard University and NBER.
- Chetty, R., A. Looney und K. Kroft (2009), Salience and taxation: Theory and evidence, *American Economic Review* 99(4), S. 1145–77.
- Chetty, R., J.N. Friedman, S. Leth-Peterson, T. H. Nielsen und T. Olsen (2014), Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark, *Quarterly Journal of Economics* 129(3), S. 1141–219.
- Conlin, M., T. O'Donoghue und T.J. Vogelsang (2007), Projection bias in catalog orders, *American Economic Review* 97(4), S. 1217–49.
- Croson, R. und S. Gächter (2010), The science of experimental economics, *Journal of Economic Behavior & Organization* 73(1), S. 122–31.
- DellaVigna, S. und U. Malmendier (2006), Paying not to go to the gym, *American Economic Review* 96(3), S. 694–719.
- DellaVigna, S. (2009), Psychology and Economics: Evidence from the Field, *Journal of Economic Literature* 47(2), S. 315–372.
- Dohmen, T., A. Falk, D. Huffman und U. Sunde (2008), Representative trust and reciprocity: Prevalence and determinants, *Economic Inquiry* 46(1), S. 84–90.
- Falk, A., S. Meier und C. Zehnder (2013), Do lab experiments misrepresent social preferences? The case of self-selected student samples, *Journal of the European Economic Association* 11(4), S. 839–52.
- Falk, A. und U. Fischbacher (2006), A theory of reciprocity, *Games and Economic Behavior* 54(2), S. 293–315.
- Falk, A. und J. Heckman (2009), Lab experiments are a major source of knowledge in the social sciences, *Science* 326, S. 535–38.
- Fehr, E., U. Fischbacher und E. Tougareva (2002), Do high stakes and competition undermine fairness? Evidence from Russia, IEW Working Papers 120, Universität Zürich.
- Fehr, E., S. Gächter und G. Kirchsteiger (1997), Reciprocity as a contract enforcement device: Experimental evidence, *Econometrica* 65(4), S. 833–60.
- Fehr, E. und K.M. Schmidt (2003), Theories of fairness and reciprocity, Evidence and economic application, in: Dewatripont, M., L.P. Hansen und S.J. Turnovsky (Hrsg.), *Advances in Economics and Econometrics, 8th World Congress, Econometric Society Monographs*, Cambridge University Press, Cambridge, S. 208–57.
- Fehr, E. und K.M. Schmidt (1999), A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114(3), S. 817–68.
- Feltovich, N. (2012), What's to know about laboratory experimentation in economics? *Journal of Economic Surveys* 25(2), S. 371–79.
- Fochmann, M. und J. Weimann (2013), The effects of tax salience and tax experience on individual work efforts in a framed field experiment, *Finanzarchiv/Public Finance Analysis* 69(4), S. 1–32.
- Garicano, L., I. Palacios-Huerta und C. Prendergast (2005), Favoritism under social pressure, *Review of Economics and Statistics* 87(2), S. 208–16.
- Goeree, J. und C.A. Holt (2001), Ten little treasures of game theory and ten intuitive contradictions, *American Economic Review* 91(5), S. 1402–22.
- Güth, W., R. Schmittberger und B. Schwarze (1982), An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior & Organization* 3(4), S. 367–88.
- Hamermesh, D.S. (2013), Six decades of top economics publishing: Who and how?, *Journal of Economic Literature*, 51(1), S. 162–72.

- Hoffman, E., K. McCabe, K. Shachat und V.L. Smith (1994), Preferences, property rights, and anonymity in bargaining games, *Games and Economic Behavior* 7(3), 346–80.
- Hoffmann, S., B. Mißm und J. Weimann (2015), To commit or not to commit? An experimental investigation of pre-commitments in bargaining situations with asymmetric information, *Journal of Public Economics* 121(1), S. 95–105.
- Kahneman, D. und A. Tversky (1979), Prospect Theory: An analysis of decision under risk, *Econometrica* 47(2), S. 263–91.
- Kahneman, D., J.L. Knetsch und R.H. Thaler (1990), Experimental tests of the endowment effect and the Coase Theorem, *Journal of Political Economy* 98(6), S. 1325–48.
- Kim, E.H., A. Morse und L. Zingales (2006), What has mattered to economics since 1970, *Journal of Economic Perspectives* 20(4), S. 189–202.
- Konrad, K. und M. Thum (2014), Climate policy negotiations with incomplete information, *Economica* 81(3), S. 244–56.
- Kruger, J. (1999), Lake Wobegon be gone! The ‘below-average effect’ and the egocentric nature of comparative ability judgments, *Journal of Personality and Social Psychology* 77(2), S. 221–32.
- Krupka, E. und R. Weber (2013), Identifying social norms using coordination games: Why does dictator game sharing vary?, *Journal of the European Economic Association* 11(3), S. 495–524.
- Levine, D.K. (1998), Modeling altruism and spitefulness in experiment, *Review of Economic Dynamics* 1(3), S. 593–622.
- Levitt, S.D. und J.A. List (2008), Homo economicus evolves, *Science* 31 (Feb.), S. 909.
- Levitt, S.D. und J.A. List (2007), What do laboratory experiments measuring social preferences reveal about the real world?, *Journal of Economic Perspectives* 21(2), S. 153–74.
- List, J.A. (2003), Does market experience eliminate market anomalies?, *Quarterly Journal of Economics* 118(1), S. 41–71.
- List, J.A. (2004), Neoclassical theory versus Prospect Theory: Evidence from the marketplace, *Econometrica* 72(2), S. 615–25.
- List, J.A. (2007), On the interpretation of giving in dictator games, *Journal of Political Economy* 115(3), S. 482–93.
- Madrian, B.C. und D.F. Shea (2001), The power of suggestion: Inertia in 401(k) participation and savings behavior, *Quarterly Journal of Economics* 116(4), S. 1149–87.
- Madrian, B.C. (2014), Applying insights from behavioral economics to policy design, *Annual Review of Economics* 6(1), S. 663–88.
- Mas, A. und E. Moretti (2009), Peers at Work, *American Economic Review* 99(1), S. 112–45.
- Mitropoulos, A., J. Weimann und C.L. Yang (2007), An alternative approach to explaining bargaining behavior in simple sequential games, *Pacific Economic Review* 11(2), S. 201–22.
- Nash, J.F. (1950), Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences of the United States of America* 36(1), S. 48–9.
- Neumann, J. von und O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.
- Nowak, M.A. und S. Roch (2006), Upstream reciprocity and the evolution of gratitude. *Proceedings of the Royal Society* 274 (1610), S. 605–09.
- Nowak, M.A., und K. Sigmund (2005), Evolution of indirect reciprocity, *Nature* 10, S. 1291–98.
- O’Donoghue, T. und M. Rabin (1999), Doing it now or later, *American Economic Review* 89(1), S. 103–24.
- Plott, C.R. und K. Zeiler (2005), The willingness to pay-willingness to accept gap, the ‘endowment effect,’ subject misconceptions, and experimental procedures for eliciting valuations, *American Economic Review* 95(3), S. 530–45.
- Rabin, M. (1993), Incorporating fairness into game theory and economics, *American Economic Review* 83(5), S. 1281–302.
- Read, D. und B. van Leeuwen (1998), Predicting hunger: The effects of appetite and delay on choice, *Organizational Behavior and Human Decision Processes* 76(2), S. 89–205.
- Riechmann, T. und J. Weimann (2008), Competition as a coordination device: Experimental evidence from a minimum effort coordination game, *European Journal of Political Economy* 24(2), S. 437–54.
- Roth, A.E. (1995), Bargaining Experiments, in: J. Kagel und A.E. Roth (Hrsg.), *Handbook of Experimental Economics*, Princeton University Press, Princeton, S. 253–348.
- Samuelson, L. (2005), Economic Theory and Experimental Economics, *Journal of Economic Literature* 43(1), S. 65–107.
- Sass, M. und J. Weimann (2015a), Moral self-licensing and the direct touch effect, *Cesifo Working Paper* 5174.
- Sass, M. und J. Weimann (2015b), The peculiar power of pairs, *Cesifo Working Paper* 5246.
- Sass, M., F. Timme und J. Weimann (2015), The dynamics of dictator behavior, *CESifo Working Paper* 5348.
- Schmidt, C.M. (2014), Wirkungstreffer erzielen – Die Rolle der evidenzbasierten Politikberatung in einer aufgeklärten Gesellschaft, *Perspektiven der Wirtschaftspolitik* 15(3), S. 219–33.
- Simonsohn, U. (2010), Weather to go to college, *Economic Journal* 120(543), S. 270–80.
- Slonim, R. und A.E. Roth (1998), Learning in high stakes ultimatum games: An experiment in the Slovak Republic, *Econometrica* 66(3), S. 569–96.
- Smith, V.L. (1962), An experimental study of competitive market behavior, *Journal of Political Economy* 70(2), S. 111–37.
- Sugden, R. (2005), Introduction: Experiment, theory, world: A symposium on the role of experiments in economics, *Journal of Economic Methodology* 12(2), S. 177–184.
- Thaler, R.H. und C.R. Sunstein (2008), *Nudge: Improving decisions about health, wealth and happiness*, Penguin Books, London.
- Thaler, R.H. und S. Benartzi (2004), Save more tomorrow: Using behavioral economics to increase employee saving, *Journal of Political Economy* 112(1), S. 164–87.
- Van Huyck, J., R.C. Battalio, R.O. Beil (1990), Tacit coordination games, strategic uncertainty, and coordination failure, *American Economic Review* 80(1), S. 234–48.
- Weimann, J., A. Knabe und R. Schöb (2012), *Geld macht doch glücklich. Wo die ökonomische Glücksforschung irrt*, Schäffer-Poeschel, Stuttgart.
- Weimann, J., A. Knabe und R. Schöb (2015), *Measuring Happiness, The Economics of Well-being*, MIT Press, Cambridge.
- Weimann, J. (2015), Verhaltensökonomik und Politikberatung, in: Held, M., G. Kubon-Gilke und R. Storn (Hrsg.), *Reformen und ihre politisch-ökonomischen Fallstricke, Jahrbuch normative und institutionelle Grundfragen der Ökonomik*, Bd. 14.
- Weimann, J. (2014), Nudging: Missbrauch der Verhaltensökonomie oder sinnvolle Politikberatung?, *Wirtschaftsdienst* 94(11), S. 779–781.
- White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48(4), S. 817–838

Autoreninformation



Joachim Weimann
Lehrstuhl für Wirtschaftspolitik,
Otto-von-Guericke-Universität Magdeburg
joachim.weimann@ww.uni-magdeburg.de

Joachim Weimann ist Inhaber des Lehrstuhls für allgemeine Wirtschaftspolitik an der Fakultät für Wirtschaftswissenschaft der Otto-von-Guericke-Universität Magdeburg. Er ist Vorsitzender der Gesellschaft für experimentelle Wirtschaftsforschung und wissenschaftlicher Direktor des MaXLab (Magdeburger Labor für experimentelle Wirtschaftsforschung). Er ist Mitglied der Akademie der Technikwissenschaften (acatec). Seine Forschungsschwerpunkte liegen in der Verhaltensökonomik, der experimentellen Wirtschaftsforschung, der Lebenszufriedenheitsforschung und der Umweltökonomik. Er ist Autor mehrerer Monographien, zweier Lehrbücher und zahlreicher Beiträge in wissenschaftlichen Journalen.