

# Präferenzen, Rationales Entscheiden und Moralische Urteile

Vortrag von DDr. Johannes Wallacher,

Institut für Gesellschaftspolitik an der Hochschule für Philosophie München,  
beim zweiten Workshop Ethik, vom 5. - 7. März 2003 in Arnoldshain (Taunus), zum Thema  
*Wie wollen wir leben, was sollen wir tun? Zum Verhältnis zwischen Gutem und Gerechtem*  
veranstaltet von der Evangelische Akademie Arnoldshain, dem Forschungsinstitut für  
Philosophie Hannover und der Katholische Akademie Rabanus Maurus

---

Im Zentrum des Verhaltensmodells der modernen Ökonomie steht der „Homo Oeconomicus“, der seine Entscheidungen emotionslos auf der Basis eines bestimmten Rationalität-Konzeptes trifft. In diesem Zusammenhang kontrovers diskutiert wird vor allem die Frage, inwieweit sich in diesen Ansatz ganz allgemein moralische Urteile integrieren lassen. Die Unterscheidung zwischen Gutem und Gerechtem spielt dabei eine wichtige Rolle, was im Weiteren aufgezeigt werden soll.

## 1 Das Verhaltensmodell der neoklassischen Ökonomie

Die ökonomische Logik der wirtschaftswissenschaftlichen Standardlehre orientiert sich an einem vergleichsweisen engen Konzept der *Zweck-Mittel-Rationalität*. Ausgangspunkt dafür ist das einzelne Individuum, welches seine von ihm selbst vorgegebenen Zwecke mit Hilfe der ihm zur Verfügung stehenden Mitteln bestmöglich erreichen will. Grundlage dafür ist eine permanente Kosten-Nutzen-Abwägung. Die individuellen Zwecke bzw. Ziele werden mit Hilfe von Präferenzen abgebildet, die ihre Nutzenvorstellungen repräsentieren und an denen die Akteure ihre Handlungen ausrichten.

Dabei wird allerdings ein erheblich *ingeschränktes Nutzenkonzept* zugrunde gelegt. Während die Utilitaristen des 18./19. Jahrhunderts noch von einem breiten Verständnis von Nutzen ausgingen – so unterschied beispielsweise Jeremy Bentham nicht weniger als 14 Komponenten von Nutzen (greifbare Güter wie Reichtum, Erfahrung, Macht, aber auch nicht greifbare Güter wie Reputation, Wohltätigkeit, Ehrfurcht) –, vollzog sich mit Beginn des 20. Jahrhunderts eine radikale Hinwendung zu einem rein materiell basierten Nutzenkonzept, das jegliche psychologische Grundlagen und Erklärungsmuster ausgeblendet hat. Dabei wird Abschied genommen von einer direkten Nutzenmessung durch Befragung oder Kommunikation. Statt dessen werden indirekt Rückschlüsse über die individuellen Nutzenvorstellungen aus den empirisch beobachtbaren Wahlhandlungen (Behaviorismus) abgeleitet. Im Zentrum der Untersuchungen stehen die individuelle Nutzeinschätzung der Akteure, externe Bewertungsmaßstäbe oder Nutzenvergleiche zwischen verschiedenen Individuen (interpersoneller Nutzenvergleich) werden systematisch ausgeschlossen.

Im Zentrum des ökonomischen Verhaltensmodells steht die Figur des *Homo Oeconomicus*, der im Rahmen der neoklassischen Gleichgewichtstheorie entwickelt wurde, inzwischen aber auch auf viele menschliche Handlungsbereiche jenseits von Märkten angewendet wird. Beispiele dafür sind die ökonomische Theorie der Demokratie, des Rechts, aber auch der Familie oder des Selbstmords (Becker 1993). Menschliches Verhalten wird dabei als rationale Wahl aus verschiedenen Alternativen (Rational-Choice) verstanden. Dieses Modell rationaler Wahlhandlungen beruht auf folgenden grundlegenden Prinzipien (vgl. Frey/Benz 2002; Erlei u.a. 1999; Kirchgässner 2000, 12-64):

- Handlungseinheit ist der Einzelne: Dies bedeutet, dass die individuellen Präferenzen der betroffenen Akteure Ausgangspunkt der Überlegungen sind. Dies entspricht der Annahme des methodischen Individualismus, nach der alle gesellschaftlichen Phänomene auf individuelle Entscheidungsakte zurückzuführen sind.
- Das Akteursverhalten ist anreizgeleitet: Die Individuen orientieren ihre Handlungen an Anreizen, die ihnen eine Abwägung der Vor- und Nachteile der Handlungsalternativen erlauben. Da sie stets auf ihren eigenen Vorteil bedacht sind, entscheiden sich zweckrationale Akteure für die Alternative mit dem günstigsten Kosten-Nutzen-Verhältnis.
- Anreize sind durch Präferenzen und Restriktionen bestimmt, zwischen denen strikt getrennt wird: Die Knappheitsbedingung wird dadurch modelliert, dass die Bedürfnisse prinzipiell unersättlich sind und der individuelle Handlungsspielraum durch monetäre wie zeitliche Restriktionen eingeschränkt ist. Die Präferenzen werden als konstant betrachtet, so dass Verhaltensänderungen auf leichter beobachtbare und messbare Kosten- bzw. Restriktionsänderungen zurückgeführt werden.
- Auf dieser Grundlage wird ein allgemeines ökonomisches Nachfragegesetz formuliert, das menschliches Verhalten folgendermaßen erklärt: Wenn sich die relativen Preise bzw. Kosten eines Gutes im Vergleich zu den Alternativen ändern, werden die davon betroffenen Akteure in vorhersehbarer Weise reagieren. Bei relativer Teuerung sinkt und bei relativer Kostensenkung steigt tendenziell die Nachfrage nach den entsprechenden Gütern.
- Dem Homo-Oeconomicus-Modell geht es um eine aggregierte Betrachtung und nicht darum, das Verhalten bestimmter einzelner Individuen zu prognostizieren. Es unterstellt lediglich, dass sich das Verhalten der Mehrzahl der Individuen in dieser Weise beschreiben und vorhersagen lässt.

Die Annahme des ökonomischen Verhaltensmodells wurden im Rahmen der Theorie der *gezeigten oder offengelegten Präferenzen* („*revealed preferences*“) formalisiert. Grundlage dafür ist ein ordinales Nutzenkonzept, in dem der Nutzen nur indirekt über die tatsächlichen Wahlhandlungen von Individuen erfasst wird. Dies macht weitere Annahme über das Entscheidungsverhalten der Individuen notwendig. So müssen sie Informationen über die ihnen zur Verfügung stehenden Alternativen besitzen und diese vollständig und in logisch konsistenter Weise in eine Präferenzordnung bringen können. Nach dem Präferenzkonzept wird die Wertschätzung für ei-

ne bestimmte Wahl von Alternativen  $x_1, x_2, \dots$  folgendermaßen ausgedrückt:

- $x_1 \checkmark x_2$  (Präferenz von  $x_1$  gegenüber  $x_2$ )
- $x_1 \sim x_2$  (Indifferenz von  $x_1$  gegenüber  $x_2$ )

Eine Person verhält sich dann rational, wenn seine Präferenzen folgende Bedingungen erfüllen:

- für alle  $x_i$  gilt:  $x_i \checkmark x_i$  (Reflexivität, d.h. jede Alternative wird mindestens so gut eingeschätzt wie sie selbst)
- für alle  $x_i, x_j$  gilt: entweder  $x_i \checkmark x_j$  oder  $x_j \checkmark x_i$  (Vollständigkeit)
- für alle  $x_i, x_j, x_k$  gilt: wenn  $x_i \checkmark x_j$  und  $x_j \checkmark x_k$ , dann  $x_i \checkmark x_k$  (Transitivität)

Sind die drei Bedingungen der Reflexivität, Vollständigkeit und Transitivität erfüllt, so spricht man von einer wohldefinierten Präferenzordnung, gilt zusätzlich noch die vierte Bedingung der Stetigkeit:

- für alle  $x_i, x_j, x_k$  gilt: wenn  $x_i \overset{TM}{\sim} x_j \overset{TM}{\sim} x_k$ , dann muss es eine Kombination von  $x_i$  und  $x_k$  geben, etwa  $y$  mit  $y \sim x_j$  (Stetigkeit der Präferenzordnung),

so kann die Präferenzordnung als Nutzenfunktion definiert werden.

## **2 Grenzen des behavioristischen Verhaltensmodells**

Die Bedeutung des ökonomischen Verhaltensmodells und der ökonomischen Methode liegt darin, dass sie unter den gegebenen Annahmen für bestimmte Problemanalysen einheitliche und klare Voraussagen erlaubt. Dafür können auch Abstraktionen und Reduktionen wie in der Verhaltensannahme, die zu diesem Zweck bewusst vorgenommenen werden, gerechtfertigt sein. Dieses Vorgehen ermöglicht es, theoretische Hypothesen zu entwickeln und sie empirisch zu testen. Eine solch positive Erklärung bestimmter Verhaltensmuster ist auch für eine wirtschaftsethische Argumentation wichtig. Dabei geht es nicht darum, moralisch erwünschte Verhaltensweisen zu formulieren, sondern darum, welche Konsequenzen sich aus bestimmten Verhaltensmustern für die politische Gestaltung von Strukturen und Institutionen ergeben.

Allerdings wachsen die Zweifel darüber, inwieweit das ökonomische Verhaltensmodell wirklich zur Erklärung ökonomischen Handelns von Individuen beiträgt, und ob die daraus abgeleiteten Politikempfehlungen hilfreich sind. Eine der Kritikpunkte bezieht sich darauf, ob und inwieweit Werturteile sinnvoll in dieses Konzept integriert werden können. Interessanterweise werden diese Anfragen nicht in erster Linie von „Moralisten“, sondern vermehrt aus der eigenen Zunft der Ökonomen vorgetragen, die ihre Kritik sowohl mit grundsätzlich theoretischen Erwägungen als auch mit Hilfe von empirischen Untersuchungen begründen.

### **2.1 Grundsätzliche Anfragen und Erweiterung des Präferenzmodells**

Ein wesentlicher grundsätzlicher Einwand gegen das ökonomische Verhaltensmodell bezieht sich auf die zu *einfache Struktur des Präferenz-Konzepts* (vgl. Sen 1984). Verhalten soll näm-

lich erklärt werden in der Terminologie von Präferenzen, die ihrerseits wieder durch das Verhalten definiert sind. „Präferiert“ bedeutet in diesem Sinne nichts anderes als „gewählt“. Es dürfte jedoch Informationsquellen über Präferenzen in ihrem alltäglichen Verwendungssinn geben, die nicht auf das Entscheiden bzw. Auswählen von Alternativen zurückzuführen sind. Ein Beispiel dafür ist etwa das in der Philosophiegeschichte immer thematisierte Phänomen der Willensschwäche (*acrasia*). Die Präferenz für  $x_1$  gegenüber  $x_2$  ( $x_1 \checkmark x_2$ ) bedeutet nach dem ökonomischen Verhaltensmodell nämlich dreierlei:

- Damit wird der Wahlakt beschrieben und ausgedrückt, dass die Person P  $x_1$  gewählt hat.
- Gleichzeitig bedeutet dies jedoch, dass für P der persönliche Nutzen von  $x_1$  höher ist als für  $x_2$ , da der persönliche Nutzen als numerische Repräsentation dieser Präferenz definiert wird.
- Aufgrund dieser beiden Annahmen kommt Person P kaum umhin, durch diese Wahl ihren eigenen Nutzen zu steigern, d.h. es handelt sich um einen Egoismus per Definitionem.

Die indirekte Nutzenerfassung durch die tatsächliche Wahlhandlung berücksichtigt außerdem nicht, dass die Auswahl in der Regel einen Kompromiss zwischen einer Vielzahl von Überlegungen wider gibt, die nicht allein auf die persönliche Wohlfahrtsteigerung reduziert werden können. Dabei stellt sich vor allem die Frage der möglichen *Berücksichtigung moralischer Präferenzen*. In diesem Zusammenhang scheint die moralphilosophische Unterscheidung von Gutem und Gerechten von Bedeutung zu sein. Wenn man gute Handlungen als solche definiert, durch die sich die Handelnden als diejenigen verwirklichen, die sie sein möchten, so lassen sich diese vergleichsweise einfach in das Präferenz-Konzept integrieren, wenn man den Vorteilsbegriff sehr weit definiert und argumentiert, dass alle dadurch motivierten Handlungen den eigenen Nutzen steigern. Wenn man jedoch jegliche Motivation als persönliche Nutzensteigerung definiert, stellt sich allerdings die Frage, ob ein solches Konzept noch einen ersichtlichen Erkenntnisgewinn bringt, oder ob ein solch weites Nutzenverständnis nicht tautologisch wird. Ein offensichtliches Problem hat das Homo-Oeconomicus-Modell mit Handlungen, die motiviert sind durch „verantwortliche Verpflichtung“ bzw. im Einklang mit allgemein verbindlichen Regeln stehen. Hier werden bewusst auch persönliche Nachteile in Kauf genommen. In diesem Fall wird ein „Keil zwischen persönlicher Auswahl und persönlicher Wohlfahrt“ (Sen 1984, 211) getrieben, was das behavioristische Verhaltensmodell nicht mehr abbilden kann.

Aus diesem Grund stellt sich die Frage einer *Erweiterung des Präferenzmodells* mit einer differenzierteren Struktur, die auch moralische Urteile erlaubt. An dieser Stelle kommt nun ein entscheidender Unterscheid zwischen Moralität und ökonomischer Rationalität zum Tragen. Während Letztere sich auf das Verhältnis von Entscheidungen und Präferenzen konzentriert, verlangt Moral ein Urteil über die Präferenzen. Eine mögliche Differenzierung des Präferenzmodells bietet der Vorschlag von John Harsanyi (1955 in Sen 1984, 217), „ethische“ von „subjektiven“ Präferenzen zu trennen. Während erstere die bevorzugte Entscheidung des Individuums auf der Basis moralischer Abwägungen ausdrückt, geben „subjektive“ Präferenzen die tatsächliche Wahl unter Berücksichtigung persönlicher Interessen wider.

Über diesen Dualismus von Präferenzen aus einem moralischen und einem persönlichen Standpunkt hinaus führt die von Amartya Sen (1985) vorgeschlagene *moralische Ordnung von Präferenzordnungen*. Sie beruht auf wesentlich mehr Informationen und gestattet ein Spektrum von moralischer Artikulation, was für gesellschaftliche Verständigungsprozesse von großer Bedeutung ist. Diese moralische Rangliste der Ranglisten von Handlungen hat folgende formale Gestalt: Sei  $X$  die Menge alternativer und sich ausschließender Handlungskombinationen und  $Y$  die Menge aller möglichen Ordnungen der Elemente von  $X$ , so kann mit  $Q$  eine Quasi-Ordnung der Elemente von  $Y$  erstellt werden, die reflexiv und transitiv ist, aber nicht notwendigerweise vollständig, was eine moralische Auffassung auch nicht zu sein braucht.

Dies sei an einem Beispiel erläutert. Eine Person  $P$  ist vor die Entscheidung gestellt, dass er aus beruflichen Gründen zweimal im Monat von seinem Wohnort München nach Berlin reisen muss. Für  $P$  stellt sich nun die Frage, welches Verkehrsmittel er benutzt, wobei in seine Entscheidungen verschiedene Aspekte (Zeitfrage, ökologisches Gewissen, etc.) einfließen. Nach reiflicher Überlegung kommt  $P$  zu folgenden Präferenzordnungen ( $A, B, C, M \succ Y$ ):

- $A$ : „Fliege zweimal im Monat“, was die persönliche Wohlfahrtsordnung von  $P$  unter Berücksichtigung von Interessen einiger Betroffener darstellt, etwa weil  $P$  seiner Familie nicht zumuten möchte, vier volle Tage abwesend zu sein, was die Fahrt mit der Bahn notwendig machen würde;
- $B$ : „Fliege unter allen Umständen“ repräsentiert „isolierte“ persönliche Interessen;
- $C$ : „Fliege maximal 10 mal im Jahr, wenn ich besonders günstigen Tarif bekomme; setze mich aber für eine Kerosinbesteuerung für innerdeutsche Flüge ein“, tatsächlich von  $P$  getroffene Entscheidung („moralischer Kompromiss“);
- $M$ : „Fliege nie, weil ich innerdeutsche Flüge aus ökologischen Gründen kategorisch ablehne“ repräsentiert „moralisch beste“ Lösung im Sinne eines kategorischen Imperativs.

Diese Präferenzordnungen ordnet  $P$  nun unter moralischen Gesichtspunkten und kommt zur moralischen Meta-Ordnung  $M, C, A, B$ . Dies erlaubt ein variables Ausmaß moralischer Artikulation, nicht nur im Sinne einer Unterscheidung zwischen moralischen Präferenzen und tatsächlicher Wahl (was etwa auf Willensschwäche zurückzuführen wäre), sondern vermag auch moralisches Abwägen oder moralische Kompromisse auszudrücken, was vor allem für gesellschaftliche Abstimmungsprozesse von Nutzen ist.

Die detailliertere Struktur der moralischen Ordnung von Präferenzordnungen erweist sich insbesondere bei der *Analyse von sozialen Dilemma-Situationen* als nützlich, die durch Spielmodelle wie das Gefangenendilemma charakterisiert werden können (vgl. Sen 1985, 188-193). Damit kann nämlich Moral statt zwischen Handlungsweisen in Form einer Entscheidung zwischen Präferenzmustern ausgedrückt werden. Das bekannte Gefangendilemma kann auch in einer allgemeinen Notation dargestellt, wobei für  $A$   $a_1$ : Gestehen bzw. Defektieren und  $a_0$ : Leugnen bzw. Kooperieren bedeutet und dementsprechend für  $B$ :  $b_1$ : Gestehen und  $b_0$ : Leugnen.

A/B	leugnen	gestehen
leugnen	(2,2)	(20,0)
gestehen	(0,20)	(10,10)

bzw.

A/B	leugnen	gestehen
leugnen	$(a_0, b_0)$	$(a_0, b_1)$
gestehen	$(a_1, b_0)$	$(a_1, b_1)$

Die Homo-Oeconomicus-Präferenzordnungen (HO-Präferenzen) zweckrationaler Spieler A und B sind (in abnehmender Priorität):

- A:  $(a_1, b_0)$ ,  $(a_0, b_0)$ ,  $(a_1, b_1)$ ,  $(a_0, b_1)$
- B:  $(a_0, b_1)$ ,  $(a_0, b_0)$ ,  $(a_1, b_1)$ ,  $(a_1, b_0)$

Eine kooperative Lösung kann bei diesen Präferenzen nur erreicht werden, wenn die Einhaltung des Vertrags mit Zwang durchgesetzt werden kann, da ein permanenter Anreiz zum Vertragsbruch besteht. Folgende Präferenzordnungen der beiden Gefangenen signalisieren, dass beide zu einer bedingten Kooperation bereit sind (BK),

- A:  $(a_0, b_0)$ ,  $(a_1, b_0)$ ,  $(a_1, b_1)$ ,  $(a_0, b_1)$
- B:  $(a_0, b_0)$ ,  $(a_0, b_1)$ ,  $(a_1, b_1)$ ,  $(a_1, b_0)$

Beide wollen grundsätzlich kooperieren, allerdings nur dann, wenn man davon überzeugt ist, dass auch der andere sich kooperativ verhält. Wenn einer erwartet, der andere werde gestehen, um ihm zu schaden, wird auch er aus Absicherung oder Fairnessgesichtspunkten vorziehen zu gestehen. Bei diesen Präferenzen gibt es nun zwei Gleichgewichtspunkte  $(a_0, b_0)$  und  $(a_1, b_1)$ , d.h. das Ergebnis hängt vom Verhalten ab, das jeder vom anderen erwartet. Im Gegensatz zu den HO-Präferenzen braucht es hier keinen Zwang zur Einhaltung einer Vereinbarung, da aufgrund der Präferenzordnungen keine dauernde Versuchung besteht, den Vertrag zu brechen.

Bei nachfolgender Präferenzordnung (AK) geben beide Spieler der Kooperation absoluten Vorrang, etwa weil sie davon überzeugt sind, dass Kooperation an sich geboten ist.

- A:  $(a_0, b_0)$ ,  $(a_0, b_1)$ ,  $(a_1, b_0)$ ,  $(a_1, b_1)$
- B:  $(a_0, b_0)$ ,  $(a_1, b_0)$ ,  $(a_0, b_1)$ ,  $(a_1, b_1)$

Bei BK- und AK-Präferenzen ist  $(a_0, b_0)$  unzweifelhaft das beste Ergebnis für alle und bei HO-Präferenzen das einzige Ergebnis, das beide besser stellt als die nicht-kooperative Lösung des Gefangenendilemmas. BK-Präferenzen garantieren, dass dieses „kollektive“ Optimum bei einem Vertrauen auf Kooperation erreicht wird, AK-Präferenzen sichern dieses Optimum bedingungslos, während bei HO-Präferenzen dieses Ergebnis nie erreicht wird, es sei denn mit Hilfe eines durchsetzbaren Vertrages, der allerdings aufgrund der Präferenzstruktur permanent gefährdet ist. Eine Klassifizierung der drei Präferenzordnungen nach moralischen

Gesichtspunkten könnte dann folgendermaßen aussehen: AK-Präferenzen, VS-Präferenzen, GD-Präferenzen.

## 2.2 Ergebnisse der experimentellen Ökonomie

Ein zweiter Einwand gegen das ökonomische Verhaltensmodell richtet sich gegen die *unzureichende empirische Bestätigung* der Homo-Oeconomicus-Annahmen. Die kritischen Anfragen kommen wiederum aus den Wirtschaftswissenschaften selbst, wo sich in den letzten Jahren angeregt durch sozialpsychologische Forschungen mit der experimentellen Ökonomie eine Richtung entwickelt hat, die mit Hilfe von Laborexperimenten empirische Belege für individuelles Verhalten in ökonomischen Entscheidungssituationen geliefert hat. Für Arbeiten auf diesem Gebiet wurden etwa Vernon L. Smith (USA) und Daniel Kahnemann (Israel) 2002 mit dem Nobelpreis für Ökonomie ausgezeichnet, auch der letzte deutsche Nobelpreisträger für Ökonomie aus dem Jahre 1994, Reinhard Selten, beschäftigt sich inzwischen intensiv mit diesen Fragen. Diese Laborexperimente werden unter realen Bedingungen durchgeführt. Dies bedeutet, dass die Wahlhandlungen der untersuchten Teilnehmer für diese unmittelbare monetäre Auswirkungen haben und sie für ihre jeweiligen Entscheidungen entsprechend zahlen müssen oder Geld bekommen.

Die Untersuchungen legen nahe, dass die Annahme strikt rational und eigennützig agierender Individuen zu kurz greift, und zwar in dreierlei Hinsicht. Wirtschaftliches Handeln ist nicht nur geprägt von beschränkter Rationalität („bounded rationality“), sondern auch von eingeschränkter Willenskraft und einem begrenzten Egoismus. Die empirischen Untersuchungen weisen vor allem auf *reziproke Verhaltensmuster* hin, wobei Reziprozität im Sinne einer bedingten Kooperationsbereitschaft verstanden wird (vgl. Fehr/Schmidt 2000). Dies bedeutet, dass Akteure auch in strategischen ökonomischen Entscheidungssituationen auf freundliches Verhalten freundlich reagieren und unfreundliches Verhalten entsprechend bestrafen, und zwar auch dann, wenn dies mit persönlichen Einbußen verbunden ist.

Dies lässt sich an dem so genannten Ultimatum Spiel zeigen. Bei diesem sehr einfachen Spiel müssen sich zwei Personen gemäß folgender Regel über die Aufteilung einer bestimmte Geldsumme, etwa 100 € verständigen. Person A macht einen Vorschlag, wie das Geld aufzuteilen ist. Person B muss auf diesen Vorschlag antworten und kann den Vorschlag akzeptieren oder zurückweisen. Wenn B zurück weist, bekommt keiner der beiden etwas, wenn B annimmt, wird das Geld gemäß dem Vorschlag von A aufgeteilt. Unter der üblichen spieltheoretischen Annahme der Zweckrationalität beider Spieler (Common Knowledge of Rationality) ist die Lösung des Spiels, dass A dem Mitspieler B einen minimalen Betrag von  $X > 0$ , z.B. 10 Cent, zugesteht. Da nämlich A weiß, dass dies für B besser ist als gar nichts, wird er ihm nur einen minimalen Teil überlassen, um selbst einen möglichen großen Anteil einzustreichen. Die Auswertung einer Vielzahl solcher Spiele in Laborexperimenten zeigt nun allerdings andere Lösungen. Wenn A

dem Antworter B weniger als 20 Prozent der aufzuteilenden Summe anbietet, wird dieses Angebot mit einer Wahrscheinlichkeit von 0,4-0,6 abgelehnt. Als Motiv für die Ablehnung wird in der Regel angegeben, dass ein niedriges Angebot als „unfair“ angesehen und daher zurückgewiesen wird, obwohl damit ein persönlicher Verlust verbunden ist.

Die empirische Evidenz dieser auf Fairness ausgerichteten Verhaltensmuster – bzw. genauer gesagt als Ausschluss von als unfair angesehenen Praktiken –, auch in ökonomischen Entscheidungssituationen, konnte inzwischen übrigens auch unter verschiedensten experimentellen Bedingungen in verschiedenen Kulturen nachgewiesen werden, so dass zu vermuten ist, dass *Reziprozität ein allgemein kulturübergreifendes Verhaltensprinzip* ist. Armin Falk spricht daher bereits vom Homo Reciprocans, den er dem Modell des Homo Oeconomicus gegenüberstellt (Falk 2003). Das Phänomen der bedingten Kooperationsbereitschaft differenziert Falk in fünf Verhaltensprinzipien, die durch zahlreiche empirische Experimente bestätigt wurden:

- Prinzip 1: Die Mehrheit der Individuen belohnt faires und bestraft unfaires Verhalten, selbst wenn dies mit Kosten verbunden ist.
- Prinzip 2: Eine Handlung bzw. ein Ergebnis wird jedoch nur dann als fair bzw. unfair eingeschätzt, wenn die handelnde Person hierfür tatsächlich verantwortlich ist, d.h. wenn die betreffende Person sich auch anders hätte entscheiden können. Dies bedeutet, dass nicht nur die Konsequenzen von Handlungen und Ergebnissen für eine Fairness-Beurteilung relevant sind, sondern auch Intentionen bzw. Verfahren.
- Prinzip 3: In sozialen Dilemma-Situationen zeigt die Mehrheit der Individuen ein bedingt kooperatives Verhalten: „Kooperieren die anderen, kooperiere ich auch; defektieren die anderen, defektiere ich auch.“ Hieraus ergibt sich u.U. die Existenz multipler Gleichgewichte, sogar in Situationen, in denen gemäß den Annahmen des Homo Oeconomicus Modells nur ein einziges Gleichgewicht existiert.
- Prinzip 4: Reziprozität spielt auch für das aggregierte Ergebnis auf Märkten eine Rolle. Entscheidend ist jedoch, ob es einen Spielraum für reziproke Reaktionen gibt: Während Reziprozität auf unvollständigen Vertragsmärkten das Marktergebnis beeinflussen kann, gilt dies für vollständige Vertragsmärkte kaum.
- Prinzip 5: In Gruppen, in denen (wie es typischerweise der Fall ist) eigennützig und reziproke Individuen miteinander interagieren, hängt es vom institutionellen Umfeld ab, ob sich reziprokes oder eigennütziges Verhalten „durchsetzt“: In sozialen Dilemma-Situationen ohne individuelle Bestrafungsmöglichkeit wird sich häufig eigennütziges Verhalten durchsetzen. Existiert hingegen die Möglichkeit zu (informellen) Sanktionen, können Egoisten durch reziproke Individuen diszipliniert werden, so dass Kooperation erreicht wird.

### **3 Folgerungen für Wirtschaftsethik und Wirtschaftspolitik**

Die theoretischen und empirischen Einwände gegen das ökonomische Verhaltensmodell haben *weitreichende Implikationen für Wirtschaftsethik und Wirtschaftspolitik*. Die experimentellen Ergebnisse zeigen, dass auch ökonomische Akteure entgegen der Homo-Oeconomicus-Annahme bedingt kooperationsbereit sind, was nicht bedeutet, dass die Akteure per se altruistisch oder solidarisch handeln, sondern dass sie grundsätzlich bereit sind, sich an gewissen Fairnesskriterien zu orientieren und sich unter bestimmten Voraussetzungen kooperativ verhalten. Eine entscheidende Rolle spielen dabei die Erwartungen über das Verhalten anderer und die Einschätzung des institutionellen Umfeldes. Das hat eine wichtige Bedeutung etwa im Hinblick auf die Bereitstellung öffentlicher Güter. Bedingt kooperative Akteure beteiligen sich an der Finanzierung öffentlicher Güter, wenn sie davon ausgehen, dass andere dies auch tun, und wenn die Verteilung der damit verbundenen Lasten als fair bzw. nicht als unfair empfunden wird. Die empirischen Ergebnissen stützen also in gewisser Weise die BK-Präferenzen von oben, was in der Analyse des Gefangenendilemmas zu zwei Gleichgewichtspunkten führt, einem „kooperativen“ und einem „unkooperativen Erwartungsgleichgewicht“. Das heißt aber auch, dass solche Gesichtspunkte wie Fairness oder Gerechtigkeit in gewisser Weise handlungsleitend sein können.

Die Existenz von zwei Erwartungsgleichgewichten weist daraufhin, dass soziale Interaktionen nicht nur auf der Basis zwangsbewehrter rechtlicher Vereinbarungen gesteuert werden, sondern dass regelkonformes Verhalten wesentlich auch auf informellen Mechanismen beruht. Solche Normen bedingter Kooperation, die ein wichtiges Element des *Sozialkapitals* einer Gesellschaft (Wallacher 2002) darstellen, können durch das ausschließlich anreizorientierte Homo-Oeconomicus-Modell nicht erklärt werden. Dies zeigt sich etwa an der Steuermoral, die als eine Form bedingt kooperativen Verhaltens aufgefasst werden kann (Falk 2003, 155f). Empirische Untersuchungen zeigen, dass die Bereitschaft, Steuern zu entrichten, auch davon abhängt, ob das Steuersystem als fair angesehen wird und die Erwartung vorherrscht, dass ein Großteil der Bevölkerung seine Steuern entrichtet. Wenn dies nicht der Fall ist, wird auch der „Homo Reciprocans“ in der Regel nur eine geringe Steuermoral aufweisen.

Das hat wesentliche *Konsequenzen für die Politik*, welche die Verhaltensreziprozität durch gezielte Maßnahmen nutzen und somit das Sozialkapital einer Gesellschaft fördern kann. Dabei gilt es zu beachten, dass die Rechtsordnung und andere formale Rahmenbedingungen wechselseitig mit den informellen sozialen Normen verbunden sind. Wenn das formale ordnungspolitische Institutionengefüge (Rechtssystem, Steuersystem, Sozialstaat) als fair eingeschätzt wird, stützt dies das „kooperative Erwartungsgleichgewicht“ und fördert entsprechend auch regelkonformes Verhalten. Umgekehrt können institutionelle Rahmenbedingungen, die als unfair oder ungerecht wahrgenommen werden, die freiwillige Kooperationsbereitschaft untergraben.

Empirische Untersuchungen weisen schließlich darauf hin, dass externe Anreizmechanismen auch zur *Verdrängung intrinsischer Motivation* führen können, was Rückwirkungen auf grund-

legende Formen gesellschaftlicher Kooperation hat. So zeigen Untersuchungen, dass die Bezahlung von bestimmten ehrenamtlichen Tätigkeiten die Bereitschaft zu freiwilligen Leistungen untergraben kann. Dies spielt vor allem auf unvollständigen Vertragsmärkten eine Rolle, wo Reziprozität erwiesenermaßen das Marktergebnis beeinflussen kann. Das Paradebeispiel dafür sind die Arbeitsmärkte, die in besonderem Maße unvollständig sind, da die Arbeitsanforderungen in den entsprechenden Verträgen nur sehr unvollständig spezifiziert werden können und somit ein erheblicher Handlungsspielraum verbleibt. Zahlreiche Untersuchungen legen nahe, dass die Leistungsbereitschaft nicht nur von monetären Anreizen bestimmt wird, sondern wesentlich auch intrinsisch motiviert ist (Frey 1997). Unter bestimmten Voraussetzungen können externe Anreize die Arbeitsmotivation verstärken, unter anderen Bedingungen wird durch zusätzliche Geldzahlungen oder Arbeitsvorschriften intrinsische Motivation und damit auch die Leistungsbereitschaft untergraben. Diese „verborgenen Kosten der Belohnung“ werden von Sozialpsychologen damit erklärt (vgl. ebd., 23f), dass dadurch die individuelle Selbstbestimmung eingeschränkt und die Selbsteinschätzung vermindert wird, was schließlich die Möglichkeit reduziert, intrinsische Motivation durch entsprechende Möglichkeiten explizit auszudrücken.

### **Literatur:**

**Becker**, Gary S. (<sup>2</sup>1993/1976), Der ökonomische Ansatz zur Erklärung menschlichen Verhaltens, Tübingen (orig.: The Economic Approach to Human Behavior, Chicago).

**Erlei**, Mathias/**Leschke**, Martin/**Sauerland**, Dirk (1999), Neue Institutionenökonomik, Stuttgart.

**Falk**, Armin (2003), Homo Oeconomicus versus Homo Reciprocans, in: Perspektiven der Wirtschaftspolitik 4, Nr. 1, 141-172.

**Fehr**, Ernst/**Schmidt** Klaus M. (2000), Theories of Fairness and Reciprocity – Evidence and Economic Applications, CESifo Working Paper No. 403, München.

**Frey**, Bruno (1997), Markt und Motivation. Wie ökonomische Anreize die (Arbeits)-Moral verdrängen, München.

**Frey**, Bruno/**Benz**, Matthias (2002), Ökonomie und Psychologie: eine Übersicht, in: D. Frey/Rosenstiel v., Lutz (Hrsg.), Enzyklopädie der Wirtschaftspsychologie, Göttingen.

**Kirchgässner**, Gebhard (<sup>2</sup>2000), Homo oeconomicus. Das ökonomische Modell individuellen Verhaltens und seine Anwendung in den Wirtschafts- und Sozialwissenschaften, Tübingen.

**Sen**, Amartya (1984/1977), Rationalclowns: eine Kritik der behavioristischen Grundlagen der Wirtschaftstheorie, in: K.-P. Markl (Hrsg.), Analytische Politikphilosophie und ökonomische Rationalität, Band 2, Opladen, 200-229 (orig.: Rational Fools: A Critique of the Behavioral Foundation of Economic Theory, in: Philosophy and Public Affairs 6, 317-344).

**Sen**, Amartya (1985/1974), Entscheidungen, Präferenzen und Moral, in: K.-P. Markl (Hrsg.), Analytische Politikphilosophie und ökonomische Rationalität, Band 1, Opladen, 186-199 (orig.:

Choices, Orderings and Morality, in: S. Körner (ed.), Practical Reason, Oxford).

**Sen**, Amartya (2002), Rationality and Freedom, Cambridge.

**Wallacher**, Johannes (2002), Das soziale Kapital, in: Stimmen der Zeit 219, Nr. 5, 306-318.