

Utilitarismus als Methode der Ethik

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

eingereicht am 06.07.2010

an der Philosophischen Fakultät I
der Humboldt-Universität zu Berlin

von Dipl.-Volksw. Malte C. Daniels M.A.

Datum der Promotion: 19.10.2010

Der Dekan der Philosophischen Fakultät I: Prof. Michael Seadle, PhD

1. Gutachter: Prof. Dr. Olaf Müller (Humboldt-Universität zu Berlin)
2. Gutachter: Prof. Dr. Thomas Schmidt (Humboldt-Universität zu Berlin)

Abstract

In diesem Buch schlage ich eine radikal neuartige Sicht auf den Utilitarismus vor. Meine Hauptthese ist, dass der Utilitarismus selbst keine vollständige normative Theorie ist, aus der sich per se Handlungsbewertungen ableiten ließen, sondern normativ untersättigt und neutral ist. Sein normativer Gehalt ist vollständig abhängig von angenommenen Nutzenfunktionen. Jede konsistente Menge von moralischen Regeln (Moralsystem) kann, wie ich im Anhang beweise, als ein Spezialfall des Utilitarismus interpretiert werden. Um dies zu explizieren, stelle ich utilitaristische Interpretationen verschiedener Moralpositionen vor: Zwei alltagsmoralische Regeln (das Gebot, Versprechen zu halten und das Tötungsverbot), egalitäre Positionen zu Verteilungsfragen sowie die Rawls'sche Theorie der Gerechtigkeit. Hierbei gebe ich stets korrespondierende Nutzenfunktionen an, deren Vorliegen Utilitaristen auf die interpretierten moralischen Positionen verpflichtet. Die Frage nach dem normativen Gehalt des Utilitarismus *schlechthin* ist somit falsch gestellt. Ein Großteil der Kritik des Utilitarismus als moralische Theorie läuft ins Leere, denn sie kritisiert einen solchen nur scheinbar eigenständigen normativen Gehalt des Utilitarismus, den sie selbst durch unterstellte Nutzenfunktionen erst erzeugt. Die Verteidigung des Utilitarismus verfehlt ebenso oft das Ziel, wenn sie zur Rettung des Utilitarismus *an sich* konkurrierende Nutzenfunktionen ins Feld führt, ohne die normative Neutralität des Utilitarismus generell zu thematisieren. Diese normative Neutralität macht den Utilitarismus gleichsam zur Normalform der Ethik, denn jeder moralische Disput lässt sich in einen Disput über korrespondierende Nutzenfunktionen überführen. Der Utilitarismus ist also nicht moralische Theorie, sondern vielmehr die *Methode der Ethik*.

Schlagwörter: Philosophie, Ethik, Moralphilosophie, Moraltheorie, Utilitarismus, Nutzentheorie, Nutzenfunktion, Normalform, Egalitarismus, Alltagsmoral, Rawls, Sen

Abstract

In this book I suggest to look at utilitarianism in a radical new way. My key starting point is that utilitarianism in itself is not a complete normative theory from which judgements of actions could be drawn but instead that utilitarianism is normatively undersaturated and neutral. Its normative content is entirely dependent on assumed utility functions. Every consistent set of moral rules (moral system) can, as I prove in the appendix, be interpreted as a special case of utilitarianism. To explicate this, I give utilitarian interpretations of a number of different moral positions: two everyday moral rules (the rule to keep promises and the prohibition of killing), egalitarian positions on distribution, and the Rawlsian theory of justice. In each case I propose corresponding utility functions that commit utilitarians to the interpreted moral position. Thus, asking about the normative content of utilitarianism *in itself* is meaningless. Much of the critique of utilitarianism as a moral theory misses the point as it criticizes such an assumed normative content of utilitarianism in itself while imputing the existence of certain utility functions. Many defendants of utilitarianism fall for the same mistake, as they propose rival utility functions to rescue their assumed normative content of utilitarianism without addressing the general normative neutrality of utilitarianism. For this normative neutrality, utilitarianism can be viewed as the normal form of ethics, because every moral dispute can be translated into a dispute over corresponding utility functions. Therefore, Utilitarianism is not a moral theory, but *rather the method of ethics*.

Keywords: philosophy, ethics, moral philosophy, moral theory, utilitarianism, utility theory, utility function, normal form, egalitarianism, everyday morality, Rawls, Sen

*Für Eta und Ulrich,
die mir das Denken ermöglichten.*

Inhaltsverzeichnis

1	Einführung und Übersicht	7
2	Utilitarismus	9
3	Die utilitaristische Form zweier alltagsmoralischer Regeln .	16
3.1	Die Institution des Versprechens	16
3.1.1	Der vermeintlich nicht vorhandene Wert utilitaristischer Versprechen	19
3.1.2	Eine simple Konzeption utilitaristischer Versprechen	21
3.1.3	Möglichkeit und Wert utilitaristischer Versprechen	30
3.2	„Du sollst nicht töten!“	38
4	Drei Einwände gegen den Utilitarismus	42
4.1	Freundschaft und Parteilichkeit	42
4.2	Die Unterscheidung zwischen Tun und Unterlassen	47
4.3	Exkurs: der Einwand der Überforderung.....	50
5	Egalitärer Utilitarismus	54
5.1	Egalitarismus	54
5.2	Egalitäre Kritik am Utilitarismus	57
5.3	Zwei egalitäre Varianten des Utilitarismus	58
5.3.1	Utilitaristische Gleichverteilung aus abnehmendem Grenznutzen	58
5.3.2	Utilitaristische Gleichverteilung aus interdependenten Nutzenfunktionen	64
5.4	Exkurs zur utilitaristischen Form des extremen Egalitarismus.....	74
5.5	Egalitarismus aus konkaven Funktionen der sozialen Wohlfahrt	76
5.6	Zwischenfazit	79
6	Der Rawls'sche Utilitarismus	82
6.1	Rawls' Prinzipien der Gerechtigkeit	82

6.2	Der Rawls'sche Utilitarismus.....	85
6.3	Die utilitaristische Form der Kritik an der Rawls'schen Theorie der Gerechtigkeit.....	89
7	Utilitarismus als Methode	93
7.1	Zur Annahme sachverhaltsabhängiger Präferenzen	94
7.2	Zur Annahme interdependenter Nutzenfunktionen	97
7.3	Korrespondierende Nutzenfunktionen als utilitaristische Normalform moralischer Regeln.....	100
8	Zusammenfassung	106
9	Anhang: Beweis der Existenz korrespondierender Nutzenfunktionen	109
	Literaturverzeichnis.....	112
	Selbständigkeitserklärung	118

„Jeder Werth ist eine Vergleichungsgröße, und sogar steht er nothwendig in doppelter Relation: denn erstlich ist er relativ, indem er für Jemanden ist, und zweitens ist er komparativ, indem er im Vergleich mit etwas Anderm, wonach er geschätzt wird, ist. Aus diesen zwei Relationen hinausgesetzt, verliert der Begriff Werth allen Sinn und Bedeutung. Dies ist zu klar, als daß es noch einer weitem Auseinandersetzung bedürfte.“

(Schopenhauer 1840/1977: 201f.)

1 Einführung und Übersicht

Mirrlees eröffnet seinen Beitrag „The economic uses of utilitarianism“ mit den Worten: „There have been so many papers presenting versions of utilitarianism, or defending it against criticism [...], that it is hard to defend writing another.“¹ Man muss ihm hier doppelt Recht geben. Der Utilitarismus als moralische Theorie, aber auch als Startpunkt der Wohlfahrtstheorie hat seit seiner Formulierung im England des 18. und 19. Jahrhunderts durch Bentham, Mill und Sidgwick eine überaus große wissenschaftliche Beachtung und Diskussion erfahren. Mit Mirrlees müsste man demnach eine weitere Beschäftigung mit dem Thema ablehnen, wenn diese nicht einen neuen Blick auf den Utilitarismus eröffnen könnte. In dieser Arbeit werde ich mich genau dieser Herausforderung stellen und einen radikal neuartigen Umgang mit dem Utilitarismus vorschlagen.

Zentral für meine Abhandlung ist die These, dass jegliches moralische Verhalten als utilitaristisches, also gesamtutzenmaximierendes, Verhalten *ganz bestimmter Akteure* verstanden werden kann. Anders formuliert lautet diese These: Jede konsistente Menge von moralischen Regeln (Moralsystem) kann als eine spezielle Form des Utilitarismus interpretiert werden. Diese These gilt es anhand von einigen Beispielen zu plausibilisieren, zu erläutern, zu qualifizieren und zu beweisen.

Der weitere Aufbau der Arbeit gliedert sich wie folgt: In Kapitel 2 werde ich den Utilitarismus in gebotener Kürze vorstellen, da es auf diesem Gebiet schon viele hervorragende Publikationen gibt.² Um den Rahmen dieser Ausarbeitung nicht zu sprengen, wird es außerdem notwendig sein, einige einschränkende Annahmen zu treffen und zu begründen.

¹ Mirrlees (1982: 63).

² Zum Beispiel Shaw (1998) oder Mulgan (2007).

Im 3. Kapitel stelle ich utilitaristische Interpretationen zweier alltagsmoralischer Regeln vor: das Gebot, Versprechen zu halten, und das Verbot zu töten. Es soll dabei jeweils gezeigt werden, dass sich die Einhaltung einzelner alltagsmoralischer Gebote und Verbote auch als utilitaristisches Verhalten von Akteuren mit passend zu wählenden Nutzen- oder auch Glücksfunktionen verstehen ließe.

Anschließend werde ich in Kapitel 4 drei Standardeinwände gegen den Utilitarismus vorstellen und darlegen, dass diese jeweils nur bestimmte Formen des Utilitarismus treffen, nicht aber die utilitaristische Form der hinter dem Einwand jeweils stehenden moralischen Position.

Im 5. Kapitel werde ich verschiedene egalitäre Positionen zu Verteilungsfragen vorstellen und mittels Sens Maßzahlen der Ungleichverteilung charakterisieren. Mithilfe von interdependenten Nutzen- oder auch Glücksfunktionen wird es auch hier gelingen, die scheinbar anti-utilitaristischen Intuitionen über Verteilungsfragen als Spezialfälle des Utilitarismus darzustellen.

In Kapitel 6 werde ich nach einer kurzen Vorstellung der Rawls'schen Theorie der Gerechtigkeit Nutzenfunktionen angeben, deren Vorliegen Utilitaristen darauf verpflichtet, den Rawls'schen Prinzipien gemäß zu entscheiden. In einem zweiten Schritt wird die Hauptkritik an der Rawls'schen Position vorgestellt und ebenfalls utilitaristisch interpretiert.

Im 7. Kapitel schließlich wird die von mir vorgebrachte These philosophisch qualifiziert. Dabei wird es insbesondere um die Frage gehen, was in der Ethik durch eine Betrachtung von korrespondierenden Nutzenfunktionen – also solchen Nutzenfunktionen, die einen Utilitaristen zu denselben Handlungen verpflichten, auf die ein abzubildendes Moralsystem verpflichtet – vorausgesetzt wird und worin ihr Gewinn besteht. Im Anhang unter 9. wird die eingangs erwähnte These bewiesen.

Starten wir also mit einer kurzen Vorstellung des Utilitarismus.

2 Utilitarismus

In diesem Kapitel werde ich in gebotener Kürze den Utilitarismus vorstellen und auf die gängigsten Kritiken eingehen, um auch den mit dem Utilitarismus nicht vertrauten Lesern einen Einstieg in das Thema zu ermöglichen.³ Da in diesem Abschnitt wichtige Annahmen und Einschränkungen getroffen werden, bitte ich alle anderen, dieses Kapitel nicht zu überspringen.

Man kann den Utilitarismus auf die einfache Maxime reduzieren: „Handle so, dass das größtmögliche Maß an Glück entsteht!“. Diese Maxime des Utilitarismus ergibt sich aus den Grundprinzipien des Utilitarismus: Eudämonismus oder Hedonismus und Konsequentialismus.

Der Eudämonismus legt den Utilitarismus dabei auf die Zielgröße fest: Glück oder auch Nutzen. Andere an dieser Stelle vorausgesetzte und diskutierte Zielgrößen sind Wohlergehen, Zufriedenheit, die Menge der Lustgefühle, die die Menge der Unlustgefühle übersteigt, Freude, der Grad der Präferenz Erfüllung, Vorteil, Gewinn oder auch individuelles Wohl. Für die theoretische Struktur des Utilitarismus ist es dabei unerheblich, was genau die Zielgröße ist. Ich werde auf diese Varianten des Utilitarismus deshalb nicht näher eingehen und im Folgenden nur noch von Nutzen sprechen, ohne diesen hier bereits näher zu charakterisieren.⁴ Repräsentiert werden kann dieser Nutzen standardmäßig durch Nutzenfunktionen, die jeweils einer Menge von Argumenten, von denen der Nutzen annahmegemäß abhängt, einen Wert zuweisen.⁵

³ Der näher interessierte Leser sei auf eine der zahlreichen einführenden Publikationen verwiesen, zum Beispiel auf die bereits weiter oben genannten Bücher Shaw (1998) und Mulgan (2007).

⁴ Ähnlich Sen (1997: 16).

⁵ Vgl. Debreu (1954) und Broome (1991: 175).

Der Konsequentialismus als zweiter wesentlicher Bestandteil des Utilitarismus besagt, dass der moralische Wert einer Handlung allein vom Wert ihrer Konsequenzen abhängt.⁶ Folglich ergibt sich für den Utilitaristen der moralische Wert einer Handlung allein daraus, wie diese Handlung die Zielgröße des Utilitarismus, also Glück oder auch Nutzen, beeinflussen wird. Weiter fordert der Utilitarismus als teleologische Ethik, dass man genau die Handlung auswählen muss, die das Glück oder auch den Nutzen maximiert.

Im Einpersonenfall sind die vom Utilitarismus vorgeschriebenen Handlungen übrigens genau diejenigen, die ein eigennutzorientierter Akteur oder auch Egoist wählen würde. Im Mehrpersonenfall muss die utilitaristische Theorie hingegen Aussagen darüber treffen, wie die „Individualglücke“ oder auch Individualnutzen der betrachteten Individuen⁷ zu einem für den Utilitaristen allein entscheidungsrelevanten Maß aggregiert werden. Mögliche Kandidaten für dieses Maß sind beispielsweise der Summen- und der Durchschnittsnutzenutilitarismus. Um Summen- und Durchschnittsnutzenutilitarismus zu erläutern, gehe ich von einer Gesellschaft von n Individuen ($I = \{1, 2, 3, \dots, n\}$) aus und bezeichne den Nutzen jedes Individuums i mit U_i . Entweder man betrachtet nun den Gesamtnutzen U , also die Summe aller individuellen Nutzen U_i :

⁶ Eine konsequentialistische Ethik steht also im Gegensatz zur deontologischen Ethik, die einigen Handlungen einen intrinsischen moralischen Wert zuschreibt. In den Abschnitten 3.1.2 und 3.2 werde ich jedoch dafür argumentieren, dass es in extensionaler Hinsicht keinen *notwendigen* Unterschied zwischen deontologischen und utilitaristischen und somit konsequentialistischen Positionen gibt.

⁷ Ich werde mich aus Gründen der Vereinfachung im Folgenden auf den Nutzen von Menschen beschränken. Prinzipiell spricht einiges dafür, den Nutzen aller empfindenden Lebewesen zum Gegenstand der utilitaristischen Maximierung zu machen. Einige Fragen, wie die nach der praktischen Möglichkeit, etwas über Nutzenniveaus auszusagen, werden dann jedoch noch schwieriger, als sie es selbst bei einer Beschränkung der Betrachtung auf Menschen schon sind.

$$(1) \quad U = \sum_{i=1}^n U_i,$$

oder man betrachtet den Durchschnittsnutzen U_D :

$$(2) \quad U_D = \frac{\sum_{i=1}^n U_i}{n}.$$

Man erkennt leicht, dass bei konstantem n die Verteilungen des Nutzens, die (1) maximieren, auch jeweils die Verteilungen des Nutzens sind, die (2) maximieren. Aus Vereinfachungsgründen gehe ich im Folgenden von einer konstanten Anzahl n von Menschen aus, so dass zwischen (1) und (2) für das weitere Vorgehen kein relevanter Unterschied besteht.⁸

Neben diesen beiden klassischen Spielarten des Utilitarismus ließen sich – durchaus abweichend von der Literatur – zusätzlich auch beliebige andere Funktionen⁹ aus der Menge der individuellen Nutzen als *utilitaristische* Funktionen der sozialen Wohlfahrt charakterisieren.¹⁰

⁸ Für einen Einstieg in die Probleme des Utilitarismus bei nicht konstanter Anzahl Menschen siehe Parfit (1986: Kapitel 17 und 19) sowie für eine formale Behandlung Roemer (1996: 153-159).

⁹ Üblich sind jedoch plausible Qualifizierungen wie zum Beispiel ein positiver Einfluss der Individualnutzen. Vgl. beispielsweise Sen (1997: 8) und Harsanyi (1955: 268).

¹⁰ Insofern könnte hier der Begriff des Utilitarismus ausgeweitet werden. Beliebige andere Aggregationsfunktionen oder Funktionen der sozialen Wohlfahrt könnten jeweils mit der Begründung *als besondere Form des Utilitarismus* verstanden werden, dass die Argumente dieser Funktionen der sozialen Wohlfahrt *ausschließlich die Indi-*

Es spricht jedoch wenig dafür, von der einfachen und überzeugenden Form von (1) oder (2) abzuweichen.¹¹ Insbesondere – das sei bereits an dieser Stelle vorweggenommen – ist es zur Modellierung egalitärer Positionen nicht notwendig, eine von (1) oder (2) abweichende Funktion der sozialen Wohlfahrt anzunehmen.¹²

Zu unterscheiden wäre jedoch zwischen der idealen utilitaristischen Aggregationsfunktion (1) und einer praktisch von einem Utilitaristen unter nicht-idealen Bedingungen zu wählenden Aggregationsfunktion.¹³ Während (1) für den *idealen* Entscheider die rechte Aggregationsfunktion ist, hat der *Utilitarist unter nicht-idealen Bedingungen* gegebenenfalls unvollständige Informationen über Wahrscheinlichkeiten, Konsequenzen und Nutzeneffekte. Wenn der Utilitarist beispielsweise darum weiß, dass er sicherere Aussagen über Veränderungen seines Nutzens und über die Nutzen guter Freunde treffen kann, könnte dies begründen, diesen sichereren Nutzenveränderungen ein höheres Gewicht in seiner

vidualnutzen sind. Der so ausgeweitete Begriff „Utilitarismus“ ist dann keiner, der selbst schon eine bestimmte Art von Verteilung normativ beinhalten würde. Im Gegenteil: Prinzipiell ließen sich alle Verteilungen von Nutzen oder Gütern als Maximierungslösung einer ganz bestimmten utilitaristischen Funktion der sozialen Wohlfahrt darstellen.

¹¹ An dieser Stelle sei auf die Argumentation von Harsanyi verwiesen, der aus relativ simplen Annahmen über die Struktur rationaler Entscheidungen für (1) als Funktion der sozialen Wohlfahrt argumentiert. Insbesondere macht er dabei Annahmen über rationale Entscheidungen hinter dem Schleier des Nichtwissens, die er als Entscheidungen unter Unsicherheit versteht. Für Einzelheiten dieses viel beachteten Arguments siehe Harsanyi (1955) und zuvor Vickrey (1945). Für eine sehr gute Aufarbeitung und eine treffende Kritik siehe Roemer (1996: 138-150).

¹² Siehe hierzu Kapitel 5, insbesondere die Diskussion im Abschnitt 5.5, und Roemer (1996: 128ff.).

¹³ Vgl. hierzu den Ansatz von Bailey (1997), der ausführlich untersucht, wie sich Utilitaristen in nicht-idealen Welten wie der unsrigen verhalten würden, und zu dem Schluss kommt, dass unser Verhalten und unsere Institutionen utilitaristischer sind, als gemeinhin angenommen wird.

Aggregationsfunktion zu geben.¹⁴ Ebenso kann es aus Gründen einer kognitiven oder rationalen Beschränktheit der Individuen sinnvoll sein, zwischen einem Handlungs- und Regelutilitarismus zu unterscheiden.¹⁵ Auch diesen Unterschied werde ich nicht weiter diskutieren, da ich, abgesehen von Einzelfällen, im weiteren Verlauf dieser Arbeit von idealen utilitaristischen Bedingungen ausgehen werde und deswegen den Handlungsutilitarismus und die einfache Summenfunktion (1) als utilitaristische Aggregationsfunktion voraussetze.

Ein prinzipieller Einwand gegen den Utilitarismus besteht darin, die interpersonale Vergleichbarkeit von Nutzen zu bezweifeln.¹⁶ Diesen Zweifeln an der Möglichkeit interpersonaler Nutzenvergleiche steht jedoch unsere alltägliche Praxis entgegen, Annahmen über Nutzen verschiedener Personen zu treffen. Deutlich wird dies auch an der üblichen Denkfigur, sich in andere hineinzusetzen. So regt sich selten Widerspruch, wenn wir Sätze äußern wie: „Ich wäre lieber die Person A in Si-

¹⁴ Eine solche Annahme nicht-idealer Bedingungen ist ein Weg, dem so genannten Nearest-and-dearest-Einwand oder auch Freundschaftseinwand am Utilitarismus zu begegnen. Im Abschnitt 4.1 hingegen interpretiere ich die hinter diesem Einwand stehende Position als Spezialfall des idealen Utilitarismus. Für einen Einstieg in die Debatte um den Nearest-and-dearest-Einwand siehe Jackson (1991) und die anschauliche Darstellung bei Mulgan (2007: 93-114).

¹⁵ Hare (1981) argumentiert beispielsweise dafür, dass prinzipiell oder aktuell beschränkte Individuen Regeln folgen sollten, die selbst nach utilitaristischen Grundsätzen aufgestellt sind. Eine umfassende Verteidigung des Regelutilitarismus findet sich in Hooker (2000). Allerdings spricht unter idealen utilitaristischen Bedingungen für unbeschränkte Individuen alles für eine extensionale Gleichheit von Handlungs- und Regelutilitarismus. Vgl. dazu Smart (1993: 11f.), Lyons (1965) und auch Hare (1963: 130-136).

¹⁶ Aus diesem Grund verzichtet die Wohlfahrtstheorie weitgehend auf die Annahme kardinaler Nutzenfunktionen. Für einen Einstieg in die Debatte um die interpersonale Vergleichbarkeit von Nutzen eignet sich der hervorragende Sammelband von Elster und Roemer (1991).

tuation X als Person B in Situation Y“.¹⁷ Auch diese Debatte werde ich in dieser Arbeit nicht weiter betrachten und im Folgenden die interpersonale Vergleichbarkeit von Nutzen voraussetzen, ohne die sich das utilitaristische Programm überhaupt nicht sinnvoll formulieren lässt.¹⁸

Häufig wurde der Utilitarismus kritisiert, indem ins Feld geführt wurde, dass er in seinen Forderungen unseren moralischen Intuitionen widerspricht.¹⁹ Eine Variante der Verteidigung des Utilitarismus gegen eine solche Kritik wäre, dafür zu argumentieren, dass hier eben unsere moralische Intuition trügt und der Utilitarismus recht darin tut, unseren Intuitionen zu widersprechen.²⁰ Eine andere Variante der Verteidigung des Utilitarismus besteht darin, dass verneint wird, dass der Utilitarismus unter nicht-idealen Bedingungen diese unseren moralischen Intuitionen widersprechenden Forderungen überhaupt stellen würde.²¹ Diese Debatten hier im Einzelnen nachzuvollziehen, soll jedoch nicht Ziel dieser Abhandlung sein. Stattdessen werde ich eine andere Sicht auf den Utilitarismus vorstellen, die zwar *auch* als Verteidigung des Utilitarismus verstanden werden kann, letztlich aber die Notwendigkeit einer Verteidigung des Utilitarismus hinterfragt.

Die genannten klassischen Verteidigungen des Utilitarismus gehen jeweils davon aus, dass der Utilitarismus selbst schon eine vollständige

¹⁷ Vgl. Sen (1997: 46) und Harsanyi (1996: 50).

¹⁸ Vgl. Roemer (1996: 128f.). Generell ist jedoch davon auszugehen, dass eine Lockerung dieser Voraussetzung, die zum Beispiel an zentraler Stelle von Bailey (1997) vorgenommen wird, sinnvoll und erkenntnisbringend sein kann. Im Rahmen dieser Arbeit werde ich diesem Ansatz jedoch nicht weiter folgen.

¹⁹ Vgl. hierzu die inzwischen klassisch zu nennende Kritik von Williams (1979) und die Aufarbeitung bei Mulgan (2007: 93-114).

²⁰ Vgl. beispielsweise Smart (1993: 68), Kagan (1998) und Mirrlees (1982: 75).

²¹ Vgl. Bailey (1997), der von einer eingeschränkten Sicherheit in interpersonalem Nutzenvergleichen ausgeht, um Alltagsmoral und Utilitarismus zu versöhnen, und Smart (1993), der dafür argumentiert, dass ein Regelutilitarismus viele unserer moralischen Intuitionen abbilden muss.

normative Theorie ist, aus der sich Handlungsbewertungen ableiten lassen. Sie verkennen dabei die normative Untersättigung des Utilitarismus, der ohne die Annahme oder Voraussetzung konkreter Nutzenfunktionen keinen eigenen normativen Gehalt hat. Der Utilitarismus ist vielmehr *Methode*, mit der Handlungen oder Umweltzustände – aber immer nur relativ zu den angenommenen Nutzenfunktionen – bewertet werden. Dies ist der Ansatzpunkt für die von mir in dieser Ausarbeitung vorgestellte These, dass jedes konsistente Moralsystem als Spezialfall des Utilitarismus interpretiert werden kann, der durch entsprechend zu wählende, so genannte korrespondierende Nutzenfunktionen charakterisiert ist. Auch wenn ich in den folgenden Kapiteln diese These zunächst anhand von Beispielen der Vereinbarkeit von Alltagsmoral und Utilitarismus erläutere und plausibilisiere, ist diese Vereinbarkeit selbst *nicht* die Hauptaussage meiner Arbeit. Stattdessen möchte ich dafür werben, den Status des Utilitarismus als eigenständige normative Theorie zu hinterfragen und den Blick auf die Zusammenhänge zwischen Moralsystemen und ihren korrespondierenden Nutzenfunktionen zu werfen.

Nachdem wir nun die notwendigen Grundlagen gelegt und einige Konventionen getroffen haben, wollen wir im nächsten Kapitel versuchen, zwei alltagsmoralische Regeln utilitaristisch zu interpretieren.

3 Die utilitaristische Form zweier alltagsmoralischer Regeln

3.1 Die Institution des Versprechens

Wenden wir uns nun der (alltags-)moralischen Regel zu, dass gegebene Versprechen einzuhalten sind.²² Eine Extremposition in Bezug auf die Geltung dieser Regel ist die deontologische Auslegung von Versprechen, wie sie bei Kant zu finden ist. Nach deontologischen Maßstäben ist die Befolgung moralischer Prinzipien schon für sich genommen wertvoll und auch das einzige Kriterium für die moralische Bewertung von Handlungen.²³ In der deontologischen Bewertung von Versprechen kommt es somit nicht auf die Folgen des Einhaltens eines Versprechens oder auf die Folgen des Brechens eines Versprechens an, sondern allein darauf, ob das Versprechen gehalten wurde.²⁴

²² Praxis und (ideale) Regeln, wie Versprechen ausgesprochen werden können, wie sie empfangen werden können und wie sie Verhalten formen sollten, lassen sich dabei als *Institution des Versprechens* charakterisieren. Vgl. Bailey (1997: 68f.), ausgehend von Knight (1992): „Institutions are to be conceived as a set of rules, which (1) provide information about how agents are expected to act in certain situations, (2) can be recognized by those who are members of a relevant group as the rules to which other agents conform in these situations, and (3) structure the strategic choices of actors in such a way as to produce equilibria“.

²³ Vgl. beispielsweise Patzig (1983: 135).

²⁴ Diese deontologische und, wie ich zeigen möchte, nur scheinbar anti-utilitaristische Position kann dabei durchaus zu derselben Bewertung der Folgen kommen, wie eine utilitaristische. Darauf weist zum Beispiel Smart hin: „He [the non-utilitarian] might agree with us in the evaluation of total consequences but disagree with us in the evaluation of possible actions. He might say: ‘The total consequences of A are better than the total consequences of B, but it would be *unjust* to do A, for you *promised* to do B’.“ Vgl. Smart (1993: 14). Ich werde dafür argumentieren, dass es – anders als hier

In vielen Fällen ist unsere alltagsmoralische Position, die wir bei der Bewertung von Versprechenseinhaltungen oder -brüchen einnehmen, recht nah an der deontologischen Position. Auf den ersten Blick ist Hare zuzustimmen, wenn er fragt: „Do we not think that some promises ought to be kept just because they are promises, although satisfactions are not maximized thereby?“²⁵

Jedoch ist die Geltung dieser alltagsmoralischen Regel nicht ganz ohne Ausnahmen, die in der Literatur auch heftig diskutiert wurden. Charakteristisch ist hierbei, dass eine zweite alltagsmoralische Intuition, nämlich die, dass Handlungen eben *auch* nach ihren Folgen zu bewerten sind, der Prima-facie-Verpflichtung entgegensteht, gegebene Versprechen einzuhalten. Wenn wir etwas versprechen, wollen wir selten den Satz äußern „unter egal welchen Umständen werde ich dies oder jenes tun oder lassen“, sondern eher den Satz „unter erwarteten und hinreichend wenig außergewöhnlichen Umständen werde ich dies oder jenes tun oder lassen“.²⁶

Patzig schreibt dazu: „Soviel ist wahr: ein Versprechensbruch [...] [macht] jede Handlung sittlich verwerflich, falls nicht noch andere sittlich relevante Züge an ihr auftreten.“²⁷ Nach dieser Sichtweise ist es also isoliert gesehen moralisch falsch, beispielsweise eine versprochene Schachverabredung zu versäumen. Wenn wir aber die versprochene Schachverabredung nur deswegen versäumen, weil wir einen unschul-

unterstellt – dem Utilitarismus generell nicht verwehrt bleiben sollte, das „*promised*“ mit in die Bewertung einfließen zu lassen. Hierzu mehr unter 3.1.3.

²⁵ Hare (1963: 132).

²⁶ Patzig (1983: 59). Ähnlich Harsanyi (1996: 58f.): „As compared with act utilitarianism, rule utilitarianism will be much closer to traditional morality in maintaining that promises should be kept, subject only to rather rare exceptions.“ Weiter spricht unsere Praxis, gebrochene Versprechen zu entschuldigen, wenn derjenige, der das Versprechen gebrochen hat, dafür nur gute Gründe angeben kann, für eine eher moderate alltagsmoralische Auslegung der Pflicht, Versprechen einzuhalten.

²⁷ Patzig (1983: 59f.).

dig Ertrinkenden retten (müssen), wäre dieser Versprechensbruch auch nach Patzig utilitaristisch erlaubt, wenn nicht gar gefordert.²⁸ Diese abwägende Position scheint mir nicht allzu weit von unserer alltagsmoralischen Bewertung dieser Situation entfernt zu sein. Letztlich wird die moderate alltagsmoralische Bewertung der Verpflichtung, Versprechen einzuhalten, somit in jedem Einzelfall von den begleitenden Umständen abhängen.

Patzig geht an dieser Stelle noch einen Schritt weiter und stellt die den Utilitarismus verteidigende These auf, dass sich unsere Entscheidungspraxis bezüglich Versprechen über weite Strecken im Einklang mit der utilitaristischen Doktrin bewegt.²⁹ Doch diesen Schritt, auch wenn er nicht unplausibel erscheint, muss ich an dieser Stelle und auf diese Weise nicht mitgehen. Weder treffe ich Aussagen über den konkreten Gehalt unserer alltagsmoralischen Praxis noch vergleiche ich allein genau diese mit „der utilitaristischen Doktrin“.

Denn meine These ist neutraler und allgemeiner: dass jede konsistente Menge von moralischen Regeln (Moralsystem) als Spezialfall des Utilitarismus interpretiert werden kann, was insbesondere eine wie auch immer geartete alltagsmoralische Versprechenspraxis, aber auch die extreme deontologische Auslegung der Pflicht, Versprechen einzuhalten, mit einschließt.

Halten wir zunächst fest: Es gibt absolute und nicht absolute oder moderate Setzungen der moralischen Regel, dass Versprechen einzuhalten

²⁸ In einem solchen Fall liefe eine deontologische Auslegung der moralischen Pflicht zur Nothilfe und der Pflicht, Versprechen einzuhalten, auf ein moralisches Dilemma hinaus. Eine Einordnung solcher inkonsistenter Mengen moralischer Regeln in der von mir vorgeschlagenen Weise ist nicht möglich, weswegen ich solche Mengen in der eingangs formulierten These ausdrücklich ausnehme. Vergleiche dazu die Diskussion unter Fußnote 39 und in Abschnitt 7.3.

²⁹ Patzig (1983: 60).

sind. Bevor wir unter 3.1.2 diese zwei Varianten der moralischen Regel, Versprechen einzuhalten, jeweils in einem simplen Modell als Spezialfall des Utilitarismus interpretieren, wollen wir unter 3.1.1 die zu widerlegende These formulieren, dass es echte Versprechen von Utilitaristen gar nicht geben könnte.

3.1.1 Der vermeintlich nicht vorhandene Wert utilitaristischer Versprechen

Es ließe sich wie folgt dafür argumentieren, dass utilitaristische Versprechen keinen Wert haben oder dass es überhaupt keinen Sinn macht, von utilitaristischen Versprechen zu sprechen.

Zunächst müssen wir uns verdeutlichen, dass Versprechensakte Sprechakte sind, für die die Konvention einer gemeinsamen Sprache von Versprechensgeber und Versprechensnehmer eine notwendige Voraussetzung ist. Hierzu schreibt Wärneryd: „Languages are mappings that associate conventional meanings to symbolic actions. A symbolic action sigma can be said to be a statement of the intention to do some action x in the future only if it is the case that people can reliably be expected to do x upon having done sigma.“³⁰ Es kann somit nur sinnvoll sein, von echten Versprechen zu sprechen, wenn auch zuverlässig erwartet werden kann, dass sich der Versprechensgeber daran hält.

Wärneryd weiter: „Consider, however, a particular expression sigma in natural language. For sigma to mean s, the speaker must be expected to follow the action of uttering sigma by the action s. For such an expectation to arise generally, it must usually be the case that it is rational to actually do s when expected to do s. This leads to an intriguing a priori observation about the kinds of situations in reality to which statements of intention in language are applicable. [...] There can be no statements

³⁰ Wärneryd (1990: 19f.).

of intention attached to actions which no one would ever perform.“³¹ Bei der Frage, ob es in einer Gesellschaft Versprechen geben kann, geht es somit im Kern um die Frage, ob die versprochenen Handlungen überhaupt glaubhaft zugesichert und eingehalten werden können. Schauen wir uns nun zunächst zwei unmögliche Sprachgemeinschaften von Versprechensgebern an, in denen nicht erwartet werden kann, dass Versprechen eingehalten werden:

- (A) *Sprachgemeinschaft von Versprechensignorierern*: Ihre Äußerung „Ich verspreche, dass ich H tun werde“ muss wie folgt uminterpretiert werden: „Ich verspreche nicht, dass ich H tun werde“. Es kann somit nicht erwartet werden, dass die Handlung H folgt. In dieser Sprachgemeinschaft gibt es somit keine echten Versprechen von Wert.
- (B) In einer *aleatorischen Sprachgemeinschaft* lösen alle immerzu aus, ob sie ihre Versprechen halten. Ihre Äußerung „Ich verspreche, dass ich H tun werde“ muss wie folgt uminterpretiert werden: „Ich werde auslösen, ob ich H tun werde“. Es kann somit nicht erwartet werden, dass die Handlung H folgt. Auch in dieser Sprachgemeinschaft gibt es somit keine echten Versprechen von Wert.

Analog ließe sich eine utilitaristische Sprachgemeinschaft untersuchen und dafür argumentieren, dass es auch in ihr keine echten Versprechen von Wert geben könne:

- (C) In der *utilitaristischen Sprachgemeinschaft* machen es alle immer von den glücksmaximierenden Folgen ihre Handelns abhängig, ob sie ihr Versprechen halten. Ihre Äußerung „Ich verspreche, dass ich H tun werde“ muss wie folgt uminterpretiert werden: „Ich werde H genau dann tun, wenn H im Vergleich zu allen Alternativen die besten Folgen nach sich zieht.“ Da dies, so die Leugner uti-

³¹ Wärneryd (1990: 34).

litaristischer Versprechen, genau das ist, was der Utilitarist *sowie so immer sagen würde und zu befolgen hat*, können wir *nicht mehr als vorher* erwarten, dass Handlung H folgt. Folglich, so die Leugner utilitaristischer Versprechen, gibt es auch in der utilitaristischen Sprachgemeinschaft keine echten Versprechen von Wert.

Auf diese Weise könnten die Leugner utilitaristischer Versprechen zu dem voreiligen Schluss kommen, dass Versprechen von Utilitaristen ohne Wert sind, da diese ihr Versprechen brechen müssen und werden, sobald dies aggregiert nutzenmaximierend ist. Ich denke, dieser Schluss ist falsch, denn er unterschlägt die auch für einen Utilitaristen möglichen Bindungswirkungen von Versprechen. Um dies zu erläutern, versuche ich im Folgenden zu modellieren, wie sich auch ein Utilitarist verpflichtet fühlen kann, Versprechen einzuhalten.³²

3.1.2 Eine simple Konzeption utilitaristischer Versprechen

Stellen wir uns eine Gemeinschaft vor, in der das Glück der einzelnen Menschen neben den üblichen Dingen auch davon abhängt, dass die ihnen gegebenen Versprechen eingehalten werden.³³ Generell gelte also, dass

³² Auch Smart (1993: 71) sieht die Möglichkeit, dass Utilitaristen sogar besonders gute Versprechenshalter sein könnten. Er führt diese Möglichkeit jedoch nicht weiter aus.

³³ Prinzipiell sind drei Varianten denkbar, Nutzeneffekte von Versprechenseinhaltungen oder -brüchen zu verorten, die sich nicht gegenseitig ausschließen: beim Versprechensgeber, beim Versprechensnehmer und bei allen anderen. Anwander (2008: 115-184) unterscheidet dieselben drei Varianten, jedoch hinsichtlich der hier nicht zu diskutierenden Frage, um wen es bei der moralischen Pflicht, Versprechen einzuhalten, geht. Ich betrachte hier zunächst den einfachen und vielleicht intuitivsten Fall, dass Nutzeneffekte nur beim Versprechensnehmer vorliegen, die eine Einhaltung von Versprechen durch Utilitaristen zur Folge haben.

$$(3) \quad \frac{\partial U_{V,i}((\dots), V_1, V_2 \dots V_n)}{\partial V_j} > 0 \text{ für alle } j.$$

Dabei soll $U_{V,i}$ den Nutzen bezeichnen, den die Person i in Abhängigkeit von den Graden der Einhaltung V_1, V_2, \dots, V_n der ihm gegebenen n Versprechen hat. Beispielhaft sei die additive und vielleicht einfachste Form einer solchen von Versprechenseinhaltungen abhängigen Nutzenfunktion vorgestellt:

$$(4) \quad U_{V,i}((\dots), V_1, V_2 \dots V_n) = U_i(\dots) + \sum_{j=1}^n a_{ji} V_j .$$

Als Argumente (...) der angeführten Funktionen $U_{V,i}$ und U_i kommen die üblichen Argumente von Nutzenfunktionen in Betracht, zum Beispiel Güterbündel oder auch Ausstattungen mit Rechten. V_j steht für den Grad der Einhaltung des gegenüber der Person i ausgesprochenen Versprechens j , mit $V_j > 0$, falls das Versprechen eingehalten wird, und $V_j < 0$, falls das Versprechen gebrochen wird. a_{ji} ist ein Gewichtungsfaktor, der den Einfluss oder anzunehmenderweise auch die Wichtigkeit eines Versprechens j für die Person i widerspiegelt.

Den Wert eines Versprechens oder eben auch seines Bruches könnte man natürlich noch um einige Feinheiten erweitern. Beispielsweise vertrauen wir im Allgemeinen den Versprechen von engen Freunden mehr oder sind in größerem Maße über ein Nichteinhalten enttäuscht. Der Bruch eines Versprechens, das uns ein Freund gegeben hat, dürfte somit negativere Folgen für die eigene Glücksbilanz haben als der Bruch eines Versprechens, das uns ein Fremder gegeben hat, weil wir mehr darauf vertraut haben, dass es eingehalten wird. Ein weiterer Punkt: Je nachdem, wie fest uns etwas versprochen wurde, vertrauen wir mehr oder weniger darauf, dass es auch eingehalten wird. Wird ein besonders festes Versprechen gebrochen, so hat dies negativere Folgen für die eigene Nutzenbilanz als losere Versprechen oder auch nur Ankündigungen.

Gebrochene Versprechen von Personen, die uns gegenüber schon häufig Versprechen gebrochen haben, werden weniger negative Folgen für unsere Glücksbilanz haben, da wir ihre Versprechen von vornherein nicht sonderlich ernst genommen haben. All dies ließe sich durch jeweils geeignete Wahl des a_{ji} abbilden.³⁴

Ein Utilitarist, der in dieser Gesellschaft nun verschiedene Handlungen gegeneinander abwägt, muss folglich die Effekte seiner möglichen Versprechensbrüche mit bedenken.³⁵ Aggregiert er über Nutzenfunktionen vom oben angenommenen Typ, hat *ceteris paribus* das Halten eines Versprechens einen Wert und er muss also jeweils gute Gründe haben, um ein Versprechen zu brechen.³⁶ Diese Konzeption utilitaristischer Versprechen kann somit unsere moderate alltagsmoralische Praxis, mit Versprechen umzugehen, gut abbilden.

Aber auch die deontologische Extremposition, dass Versprechen unter allen Umständen und ohne weitere Betrachtung der anderweitigen Folgen einzuhalten sind, lässt sich als Spezialfall des Utilitarismus interpretieren.³⁷ Denn Utilitaristen, die die folgende Nutzenfunktion haben,

³⁴ Natürlich könnte man auch modellieren, dass der Versprechensbrecher *ceteris paribus* einen negativen Nutzeneinfluss aus dem Versprechensbruch hat. Für einen negativen Nutzeneinfluss bei einem Versprechensbrecher könnte zum Beispiel sein Gewissen sorgen. Vgl. hierzu die Diskussion in Fußnote 42.

³⁵ Den kritischen Leser, dem eine Auseinandersetzung damit fehlt, dass hier solche Folgen wie „ein Versprechen gebrochen haben“ oder unter 3.2 solche Folgen wie „jemanden getötet haben“ bewertet werden, verweise ich auf die Diskussion im Abschnitt 7.1.

³⁶ Vgl. Patzig (1983: 59).

³⁷ Die These, dass es keinen notwendigen Widerspruch zwischen einer deontologischen Moral und dem Utilitarismus gibt, wird zum Beispiel von Bailey (1997: 172) vertreten, der sich dabei auf Hare bezieht. Eine ähnliche Position ist in Cumminsky (1990) zu finden. In diesem Abschnitt wie auch unter 3.2 versuche ich Spezialfälle des Utilitarismus explizit vorzustellen, die mit einer deontologischen Moral kompatibel sind.

sollten unter allen Umständen davon Abstand nehmen, ein Versprechen zu brechen.

$$(5) \quad U_{V,i}((\dots), V_1, V_2, \dots, V_n) = \begin{cases} -\infty & \text{falls } \exists j \text{ mit } V_j < \alpha \\ U_i(\dots) & \text{andernfalls} \end{cases}$$

Hierbei sollen V_1, V_2, \dots, V_n für die Grade der Einhaltung der n gegenüber Person i ausgesprochenen Versprechen stehen, und α gerade für den Schwellenwert, ab dem ein Versprechen als nicht eingehalten gilt. Hier hätten wir jedoch den unerwünschten Effekt, dass nach dem ersten gebrochenen Versprechen die folgenden gebrochenen Versprechen keinen weiteren negativen Effekt auf die Glücksbilanz mehr hätten. Dies ließe sich leicht dadurch beheben, dass die Anzahl der gebrochenen Versprechen gezählt wird und für jeden weiteren Versprechensbruch ein noch negativerer Wert anzusetzen ist. Nehmen wir dabei jedoch einen zu kleinen negativen Wert für jeden Versprechensbruch an, so wäre es vorstellbar, dass er durch (etwa in einem Szenario konstruierte) anderweitige positive Effekte aufgewogen wird. Eine Lösung zur Modellierung der deontologischen Position wäre hier beispielsweise, eine anzunehmende maximale Glücksmenge im Leben eines Menschen mit der angenommenen maximalen Anzahl der betrachteten Menschen zu multiplizieren ($= \lambda$) und den Wert eines jeden Versprechensbruches mit dem Negativen ($= -\lambda$) hiervon in der Nutzenfunktion anzusetzen.³⁸ Wir erhalten:

³⁸ Der kritische Leser fragt zu Recht, wie die maximale Glücksmenge im Leben eines Menschen und die maximale Anzahl der betrachteten Menschen zu bestimmen sind. Zum Glück ist es an dieser Stelle nicht notwendig, belastbare Größen anzugeben, da wir hier lediglich auf der Suche nach einer plausiblen Untergrenze für eine hinreichend große Zahl sind. Trotz der Schwierigkeiten der interpersonalen Nutzenverglei-

$$(6) \quad U_{V,i}((\dots), V_1, V_2 \dots V_n) = U_i(\dots) - \lambda \sum_{j=1}^n V_j$$

mit $V_j = 1$, falls das j -te gegenüber der Person i ausgesprochene Versprechen gebrochen wird. In diesem Fall hat also annahmegemäß jeder Versprechensbruch einen negativen Nutzeneffekt auf den Versprechensnehmer, der in keinem Fall durch etwaig damit verbundene andere Nutzeneffekte aufgewogen werden kann. Versprechen zu brechen, kann unter diesen Voraussetzungen folglich nie die glücksmaximierende und somit nie die richtige Handlung sein.

Durch die Setzung der negativen Folgen eines Versprechensbruches für das Individuum i auf $-\infty$ in Formel (5), aber auch durch die prohibitiv hohe Setzung des λ in Formel (6) wird quasi eine lexikographische Priorität der Variablen V über die Argumente von $U_i(\dots)$ ausgedrückt. Dies hat die beabsichtigte Folge, dass ein Abweichen von der als absolut verstandenen abgebildeten moralischen Regel für einen Utilitaristen unter keinen Umständen in Betracht kommt.³⁹

Alternativ wäre es *in gleicher Weise*⁴⁰ möglich, negative Nutzeneffekte eines Versprechensbruches beim Versprechensbrechenden anzuneh-

che, auf die bereits unter 2 hingewiesen wurde, und auch trotz der Schwierigkeiten der gegenwärtigen Zählung zukünftiger Menschen wird es für jeden modellierten Fall leicht sein, eine solche hinreichend große Zahl anzugeben.

³⁹ Schwierigkeiten könnte es hierbei geben, wenn *mehrere als absolut angenommene* moralische Regeln auf diese Weise in den Nutzenfunktionen utilitaristisch modelliert werden sollen. Wie ich jedoch im Abschnitt 9 beweise, können wir stets geeignete Zahlenverhältnisse für die in dieser Weise aufzustellenden Nutzenfunktionen finden, falls das abzubildende Moralsystem konsistent ist.

⁴⁰ Hieran sieht man, dass die von mir vorgeschlagene Methode prinzipiell *neutral* ist bezüglich des Ortes der Modellierung von Nutzeneffekten. Weiter unten werde ich einen Sonderfall eines moralisch zu bewertenden Versprechens vorstellen, bei dem dies anders ist: das berühmte Versprechen am Sterbebett.

men.⁴¹ Hierzu müssten wir nur Formel (5) unter der Maßgabe interpretieren, dass V_1, V_2, \dots, V_n für die Grade der Einhaltung der *n* von Person *i* *ausgesprochen* Versprechen stehen, und analog Formel (6) unter der Maßgabe interpretieren, dass $V_j = 1$, falls Person *i* das *j*-te von *ihr* *ausgesprochene* Verspreche bricht. Als Quelle dieser Nutzeneinflüsse ließe sich beispielsweise das schlechte Gewissen des Versprechensbrechenden ausmachen.⁴² Darüber hinaus ist leicht ersichtlich, dass eine Kombination der beiden Interpretationen von Formel (6) auch eine gleichzeitige Modellierung von Nutzeneinflüssen sowohl auf den Versprechensnehmer als auch auf den Versprechensgeber zulässt.⁴³

⁴¹ Der interessierte Leser, der sich fragt, worin diese negativen Nutzeneffekte des Versprechensbrechers bestehen könnten, sei auf Anwander (2008: 115-134) verwiesen, der untersucht, um wen es bei der moralischen Pflicht, Versprechen einzuhalten, geht. Anders als Anwander suche ich hier jedoch nicht nach Begründungen für oder gegen eine moralische Regel, Versprechen einzuhalten, sondern zeige lediglich verschiedene Möglichkeiten der *Annahme* solcher Nutzeneffekte – einmal beim Versprechensgeber und einmal beim Versprechensnehmer – auf, die gleichsam eine Befolgung dieser Regel durch Utilitaristen zur Folge hätten.

⁴² Häufig aufgeworfen wird diesbezüglich die Frage, ob ein Utilitarist, der mit einem Versprechensbruch das Gesamtwohl maximiert, überhaupt ein schlechtes Gewissen haben *sollte*. Ein Grund dagegen wäre, dass er ja das Beste getan hat und sich somit nichts vorzuwerfen hat. Ein regelutilitaristischer Grund dafür könnte sein, dass gerade das Vorliegen eines schlechten Gewissens mit dazu führt, dass die Regel, nur Versprechen zu brechen, wenn dies klar besser ist, stabiler wird. Diese Frage will ich hier nicht weiter diskutieren oder entscheiden. Unabhängig davon, ob ein Utilitarist ein schlechtes Gewissen haben sollte, kann hier die Position vertreten werden, dass es mit in das Nutzenkalkül einbezogen werden muss, *wenn er eines hat*.

⁴³ An dieser Stelle geht es mir *nicht* darum, eine Aussage darüber zu treffen, bei wem Nutzeneffekte von eingehaltenen oder gebrochenen Versprechen *wirklich* anfallen. Wenn mich meine alltagspsychologischen Kenntnisse nicht täuschen, scheint es aber nicht von der Hand zu weisen, dass der, dem etwas Versprochenes nicht eingehalten wird, einen nicht unerheblichen Teil der negativen Folgen des Versprechensbruches trägt. Dies ist jedoch für die von mir vorgeschlagene Methode nicht erheblich.

Nun haben wir das Rüstzeug beieinander, um zwei beliebte moralische Gedankenexperimente in der vorgestellten Weise nachzuvollziehen. Diese beiden Gedankenexperimente sind an dieser Stelle deswegen für uns relevant, weil sie häufig genannt werden, um für eine Unvereinbarkeit von deontologischen und utilitaristischen Positionen zu argumentieren. Um die These von der Interpretierbarkeit jeder Menge von moralischen Regeln als Spezialfall des Utilitarismus beibehalten zu können, gilt es, die deontologische Position jeweils utilitaristisch zu modellieren.

Das erste Gedankenexperiment hat ein einem Sterbenden⁴⁴ gegebenes Versprechen zum Inhalt, von dem kein Dritter etwas wissen kann⁴⁵ und dessen Einhaltung große und zum Zeitpunkt des Versprechens nicht vorhergesehene negative Folgen nach sich ziehen würde.⁴⁶ Es lässt sich dafür argumentieren und es wird auch allgemein angenommen, dass hier die deontologische Bewertung eine andere sein *muss* als die utilitaristische.⁴⁷ Denn der Utilitarist müsse angesichts der großen und zum Zeitpunkt des Versprechens unvorhersehbaren Folgen eindeutig für einen Bruch des Versprechens optieren, durch den zudem der Versprechensnehmer durch zwischenzeitliches Ableben nicht negativ affiziert

⁴⁴ Die besondere Situation des Versprechens am Sterbebett schließt, zumindest unter der nicht abwegigen Annahme, dass es kein oder zumindest kein beeinflusstes Leben nach dem Tod gibt, einen negativen Nutzeneffekt auf den Versprechensnehmer für den Fall des Bruches des Versprechens aus.

⁴⁵ Dass Dritte von diesem Versprechen nichts mitbekommen haben, schließt dabei für diesen Fall die übliche Argumentation aus, dass es einen Vorbildeffekt eines Einhaltens oder eines Brechens des Versprechens geben kann, der selbst unter Nutzenaspekten zu bewerten wäre. Der interessierte Leser sei auf Williams (1979: 63) verwiesen, der solche Vorbildeffekte auf Dritte zugesteht. Anwander (2008: 135-153) bezeichnet die soziale Konvention, Versprechen einzuhalten, als „wertvolle Praxis“ und nennt „Interessen, die wir alle an einer Institution des Versprechens haben“, ohne dabei jedoch eine explizit utilitaristische Begründung der Institution des Versprechens im Sinn zu haben.

⁴⁶ Siehe Smart (1993: 63) und Hare (1963: 133-134).

⁴⁷ Siehe Hare (1963: 133-134) und Smart (1993: 5, 13).

werden könne. Wie wir jedoch weiter oben bereits gesehen haben, würde sich der Utilitarist der deontologischen Forderung, auch in diesem Fall das Versprechen einzuhalten, unter der Voraussetzung anschließen, dass die negativen Nutzeneffekte des Versprechensbruches die anderweitigen großen Nutzeneffekte aufwiegen würden.⁴⁸

Im zweiten Gedankenexperiment kann jemand sein Versprechen nur einhalten, wenn er es damit fünf anderen verunmöglicht, ihre Versprechen einzuhalten.⁴⁹ Die deontologische als explizit nicht-konsequentialistische Position fordert auch in diesem Fall, das *erste* Versprechen einzuhalten, egal wie vielen *anderen* Versprechern man damit ihr Versprechenseinhalten unmöglich macht.⁵⁰ Es findet also keine Abwägung statt, ob ein Versprechensbruch oder gegebenenfalls fünf besser sind.⁵¹ Nun wird aber auch die deontologische wie die utilitaristische Position nichts von einer Person fordern können, das unmöglich ist. Verändern wir also die Nutzenfunktionen (5) und (6) so, dass nur dann

⁴⁸ Eine Stufe weiter in Richtung eines argumentatorischen Overkills treiben wir dieses Gedankenexperiment, wenn wir *zusätzlich* annehmen, dass der Versprechende zwischenzeitlich sein Versprechen *vergessen* hat und es nur *zufällig* bricht oder einhält. Die deontologische Position würde wohl auch in diesem Fall das Einhalten des Versprechens als besser oder als geboten bewerten, obwohl dies annahmegemäß überhaupt keinen Unterschied in der Welt bewirken könnte. Sinnvoll fordern vom unwissend Handelnden könnte die deontologische Position das Einhalten des Versprechens in einem solchen Fall natürlich nicht. Jedenfalls nicht mehr, als ein Utilitarist gefordert wäre, unbekannte Nutzeneffekte in sein Kalkül mit einzubeziehen.

⁴⁹ Das Beispiel stammt von Williams (1972: 52).

⁵⁰ Williams (1979: 52), aber auch Broome (1991: 5) sehen dies als Beleg dafür, dass sich Konsequentialismus und die deontologische Position nicht miteinander vereinbaren lassen. Ich denke, dass sie sich in diesem Punkt irren.

⁵¹ Genau an diesem Punkt kann man erkennen und kritisieren, dass es der deontologischen Position gar nicht um Versprechen oder die Einhaltung von Versprechen generell geht, genauso wenig, wie es um die beteiligten Menschen oder ihr Glück geht. Es zählt allein das Befolgen oder das Bemühen um das Befolgen einer allermeist nicht kontextsensitiven Regel.

$V_j < \alpha$ im Fall von (5) oder nur dann $V_j = 1$ im Fall von (6), wenn es dem Versprechensgeber *überhaupt möglich* ist, sein Versprechen einzuhalten. Qualifizieren wir die Nutzenfunktionen in dieser Art um, so wird auch die utilitaristische Abwägung zu dem Schluss kommen, dass das *erste* Versprechen stets einzuhalten ist, egal wie viele andere Versprechenseinhaltungen in der Folge damit unmöglich gemacht werden. Denn die fünf folgenden Versprechensbrüche hätten, da sie nach dem ersten Versprechensbruch annahmegemäß unvermeidlich sind, bei wie oben umqualifizierten Nutzenfunktionen (5) und (6) keinen prohibitiv negativen Effekt.⁵²

Sollten uns unsere moralischen Überlegungen von der deontologischen Position abweichend zu dem Schluss kommen lassen, dass es bei Vorliegen geeigneter Umstände geboten ist, einzelne Versprechen zu brechen, um eine gewisse Anzahl anderer Versprechenseinhaltungen möglich werden zu lassen, so ließe sich dies durch geeignete Wahl des λ in (6) modellieren.

In diesem Abschnitt wurde anhand von beispielhaften Nutzenfunktionen dafür argumentiert, dass ein Utilitarist verpflichtet sein kann, Versprechen zu halten. Hierbei wurden verschiedene moralische Intuitionen bezüglich der Verpflichtung, Versprechen einzuhalten, modelliert.⁵³ Bevor wir unter 3.2 zu einer analogen Betrachtung des Verbotes zu töten kommen, werde ich zunächst untersuchen, ob Utilitaristen überhaupt Versprechen geben sollten.

⁵² Es ist zugegebenermaßen so, dass die so umqualifizierten Nutzenfunktionen überraschende Eigenschaften haben. Dies ist jedoch den Eigenschaften der utilitaristisch abzubildenden deontologischen Position geschuldet.

⁵³ Aus Platzgründen verzichte ich auf eine Behandlung des Verbotes zu lügen. Die Argumente scheinen mir ganz analog dem Fall des Versprechensbruches zu liegen.

3.1.3 Möglichkeit und Wert utilitaristischer Versprechen

Aus der Argumentation im letzten Abschnitt können wir folgern, dass bei passenden Präferenzen auch Utilitaristen Versprechen einhalten werden, *wenn sie welche gegeben haben*. Diese Analyse ist so lange unvollständig, wie wir nicht zeigen können, dass sie auch echte Versprechen *geben können und würden*.

In diesem Abschnitt will ich also dafür argumentieren, dass es auch in einer Gesellschaft von Utilitaristen Versprechen geben könnte.⁵⁴ Die Gegenposition liegt ja bereits auf dem Tisch: In Abschnitt 3.1.1 sahen wir, dass ein Utilitarist, egal, ob er ein Versprechen gegeben hat oder nicht, genau dieselbe Maxime befolgt: „Ich werde H genau dann tun, wenn H im Vergleich zu allen Alternativen die besten Folgen nach sich zieht.“ Die Frage ist nun, ob die beste Handlung vor und nach Abgabe des Versprechens *dieselbe* Handlung ist. Dies hängt, wie wir im letzten Kapitel gesehen haben, von den angenommenen Präferenzen, aber auch davon ab, ob wirklich echte Versprechen zu Stande gekommen sind.⁵⁵

⁵⁴ Ich versuche, das hier nur für den leichtesten Fall zu zeigen, in dem die gesamte Gesellschaft aus Utilitaristen besteht. Kompliziertere Fälle, in denen ein gewisser Anteil der Gesellschaft nach anderen, zum Beispiel egoistischen, Kriterien handelt, werden sich nur im Grad der Nützlichkeit von Versprechen und der Sicherheit ihrer Einhaltung unterscheiden.

⁵⁵ Ob also ein Sprechakt in obigen Nutzenfunktionen (3) bis (6) als V_j zu zählen ist oder nicht.

Hier gibt es, wie ich zeigen werde, zwei Möglichkeiten⁵⁶, die jeweils Gleichgewichte in Erwartungen sind.⁵⁷

Fall 1: Der „Versprechensnehmer“ erwartet vom „Versprechensgeber“ *nicht mehr als vor Abgabe des so verstandenen „Versprechens“* ein Verhalten, das eine „Einhaltung“ des „Versprechens“ wäre. Genauer: Er erwartet keinen Einfluss auf das Verhalten des utilitaristischen „Versprechensgebers“. Ein solcher Sprechakt wird also auf Seiten des „Versprechensnehmers nicht die Erwartungen an zukünftige Handlungen und insbesondere auch nicht die Bewertung von zukünftigen Handlungen verändern. Somit wird der Einfluss von Handlungen des „Versprechensgebers“ auf seinen Nutzen der gleiche sein wie vor Abgabe des „Versprechens“. Der „Versprechensgeber“ wiederum erwartet genau dies, dass also die Abgabe eines so verstandenen „Versprechens“ die Nutzensauswirkungen seiner zukünftigen Handlungen nicht verändert. Er wird sich daher nach Abgabe eines so verstandenen „Versprechens“ genau derselben Entscheidungssituation gegenübergestellt sehen, in der er ohne die Abgabe eines solchen „Versprechens“ wäre, und wird sich als Utilitarist folglich für dieselbe Handlung entscheiden. Hier liegt ein Gleichgewicht vor, denn die Gesamtmenge von Erwartungen und Handlungen ist in sich stimmig. In diesem Fall liegt offensichtlich kein echtes, d.h. kein Erwartungen veränderndes oder entscheidungsrelevantes, Versprechen vor.⁵⁸

⁵⁶ Ich gehe hier und im Folgenden davon aus, dass die beteiligten Personen Sicherheit darüber haben, wie der jeweils andere den Satz „ich verspreche, dass p“ interpretieren wird und ob und wie die Äußerung des Satzes die Erwartungen und Nutzen von zukünftigen Handlungen verändert.

⁵⁷ Im Folgenden nehme ich aus Gründen der Vereinfachung Nutzenfunktionen wie zum Beispiel (3) und (4) an, die die Effekte eines Einhaltens oder eines Bruches des Versprechens allein bei den Versprechensnehmern lokalisieren.

⁵⁸ Weswegen obige Begriffe in Anführungszeichen zu setzen waren.

Fall 2: Der Versprechensnehmer erwartet vom Versprechensgeber *mehr als vor Abgabe des Versprechens* ein Verhalten, das eine Einhaltung des Versprechens wäre. Somit verändert sich der Einfluss von Handlungen des Versprechensgebers auf seinen Nutzen durch den Sprechakt der Abgabe des Versprechens.⁵⁹ Denn der Versprechensgeber wiederum bezweckt genau dadurch eine *Selbstbindung*, dass die Abgabe eines Versprechens die Nutzensauswirkungen seiner zukünftigen Handlungen verändert. Obwohl er als Utilitarist weiterhin nach derselben Maxime der Nutzenmaximierung zu handeln hat wie vor Abgabe des Versprechens, sieht er sich *nach Abgabe eines Versprechens* einer veränderten Entscheidungssituation gegenübergestellt. Denn durch das Abgeben seines Versprechens hat er die Folgen seiner zukünftigen Handlungen in einer Weise beeinflusst, die ein Einhalten seines Versprechens nun wertvoller machen. Handlungen, die ein Einhalten des Versprechens konstituieren, bewirken nun relativ mehr Nutzen⁶⁰ als vor Abgabe des Versprechens.⁶¹ Er wird sich als Utilitarist daher *ceteris paribus* in mehr

⁵⁹ Und zwar annahmegemäß so, dass ein Einhalten des Versprechens positiv bewertet wird, oder genauer: dass eine Handlung, die ein Einhalten des Versprechens konstituiert, für ihn von höherem Nutzen ist, als wenn das Versprechen nicht gegeben worden wäre.

⁶⁰ Zum Beispiel allein schon dadurch, dass der Versprechensnehmer im Vertrauen auf ein wahrscheinliches Einhalten des Versprechens entsprechend disponiert hat. In der Unsicherheit darüber, wie stark der Versprechensnehmer auf ein Einhalten des Versprechens baut, läge ein guter utilitaristischer Grund dafür, eine utilitaristische Regel einzuführen, die Versprechenseinhaltungen fordert, solange nur nicht sehr viel dagegen spricht.

⁶¹ Siehe Smith (1766: 12, 87ff.) und auch Mill (1859: 174): „When a person, either by express promise or by conduct, has encouraged another to rely upon his continuing to act in a certain way — to build expectations and calculations, and stake any part of his plan of life upon that supposition, a new series of moral obligations arises on his part towards that person, which may possibly be overruled, but can not be ignored.“ Die hierzu heiß diskutierte Frage, woher genau die moralische Verpflichtung, Versprechen einzuhalten, kommt, ist somit für den Utilitaristen kurz und einfach zu beantworten: Sie folgt in den betrachteten Situationen bereits aus der allgemeinen Pflicht, den Nut-

Situationen für Handlungen entscheiden müssen, die ein Einhalten seines Versprechens konstituieren, als vor Abgabe seines Versprechens.⁶² Hier liegt das zweite Gleichgewicht vor, denn die Gesamtmenge von Erwartungen und Handlungen ist in sich stimmig.⁶³ Im Fall solcher hier angenommenen Präferenzen und Erwartungen liegen somit echte, d.h. entscheidungsrelevante, utilitaristische Versprechen vor.⁶⁴

Wie wir sehen, liegen somit zwei Gleichgewichte vor. Natürlich könnten wir einfach *annehmen*, dass die Utilitaristen diese im Fall 2 vorausgesetzten Präferenzen und Erwartungen kontingenterweise haben, um die moralische Regel, Versprechen einzuhalten, als Spezialfall des Utilitarismus interpretieren zu können. Damit jedoch der Nachweis von Möglichkeit und Wert utilitaristischer Versprechen nicht unvollständig bleibt, müssen wir über den gebrachten Nachweis, dass ein solches Gesamtsystem aus Erwartungen und Handlungen stabil ist, hinaus noch dafür argumentieren, dass die Institution des Versprechens in einer Gesellschaft von Utilitaristen sinnvoll sein kann.

zen zu maximieren. Vgl. Hume (1739/40: 524) und die Diskussion bei Anwander (2008: 10).

⁶² Deutlich werden hier eine Pfadabhängigkeit und Selbstverstärkung der Geltung moralischer Regeln: In diesem Fall werden dadurch, dass Versprechen eingehalten werden, die Erwartungen an Einhaltungen von Versprechen gestärkt, was wiederum über die plausible Annahme steigender Differenzen in Nutzeneffekten bei stärkeren Erwartungen an Einhaltungen von Versprechen dazu führt, dass utilitaristische Versprechen eher eingehalten werden.

⁶³ Vgl. hierzu Fußnote 22. In dieser hier beschriebenen utilitaristischen Gesellschaft besteht also eine Institution des Versprechens, *für die jedoch keine speziellen Regeln bestehen müssen*, sondern lediglich die oben spezifizierten Nutzenfunktionen und die allgemeine Pflicht, nach der utilitaristischen Maxime zu handeln.

⁶⁴ Der Unterschied zwischen Versprechen und bloßen Absichtserklärungen ist in dieser Konzeption lediglich ein gradueller. Im hier vorgestellten Fall binden ausgesprochene Versprechen stärker als Absichtserklärungen, weil es gängige und gewollte Praxis ist, dass neben einer erklärten Absicht stets das Angebot mitschwingt, *sich darauf verlassen zu können*.

Dabei will ich die Frage undiskutiert lassen, ob und wie sich Erwartungen und Präferenzen bezüglich Versprechenseinhaltungen und -brüchen verändern können, zum Beispiel durch Gruppenevolution⁶⁵ oder durch Wahl durch die Personen⁶⁶. Ich werde also nicht den Weg beschreiben, auf dem sich utilitaristische Gesellschaften ohne Institution des Versprechens zu utilitaristischen Gesellschaften mit der Institution des Versprechens wandeln könnten, sondern ich werde lediglich diese beiden denkbaren Gesellschaften miteinander vergleichen und dafür argumentieren, dass auch und gerade für Utilitaristen die Institution des Versprechens eine sinnvolle Institution sein könnte.⁶⁷ Weiter ist zu erklären, warum in einer Gesellschaft von Utilitaristen, in der es die Institution des Versprechens gibt, überhaupt jemals ein Utilitarist ein Versprechen geben sollte.⁶⁸

Nach obigen Ausführungen ist klar, dass der Utilitarist eine gute Rechtfertigung dafür braucht, ein Versprechen abzugeben. Denn zumindest in einigen Fällen wird sein zukünftig optimales Handeln nach Abgabe eines Versprechens ein anderes sein, als es ohne Abgabe dieses Verspre-

⁶⁵ Die These hierbei wäre die folgende: Im Prozess der Evolution und Selektion von Gruppen sind diejenigen Gruppen erfolgreicher, deren Gruppenmitglieder sich explizit oder implizit utilitaristisch bezüglich ihrer Gruppe verhalten. Für eine Übersicht über spieltheoretische Erklärungen von evolutionsstabilem Verhalten siehe Güth (1999: 34, 173-177) und die weiterführende Literatur bei Hammerstein/Selten (1994).

⁶⁶ Die These wäre hier, dass die Mitglieder der utilitaristischen Gesellschaft aus utilitaristischen Gründen ihre Erwartungen und Präferenzen bezüglich Versprechenseinhaltungen und -brüchen so wählen sollten, dass eine erfolgreiche Praxis des Versprechens möglich wäre. Ohne diese beiden Thesen weiter diskutieren zu wollen, hoffe ich, dass sich aus dem Geschriebenen zumindest ein Startpunkt für eine Begründung der in dieser und der vorherigen Fußnote erwähnten Thesen ergibt.

⁶⁷ Genau diese gleichgewichtsauswählende Funktion von Institutionen untersucht Bailey (1997: 68-89). Vgl. Fußnote 22.

⁶⁸ Gründe wie Nutzeneffekte aus aufzubauendem Vertrauen und etwaigem Prozessnutzen brauchen wir dabei nicht gesondert betrachten: Sie sind in der utilitaristischen Bewertung bereits enthalten.

chens gewesen wäre. Das ist Sinn und Zweck von Versprechen. Letztlich sind Versprechen für Utilitaristen aber immer nur dann relevant verhaltensändernd, wenn ein Versprechensbruch beispielsweise beim Versprechensnehmer zu so negativen Folgen führen würde, dass *deswegen* die vor Abgabe des Versprechens optimale Handlung nicht mehr ausgeführt werden darf. Durch Abgabe des Versprechens verpflichtet sich der utilitaristische Akteur somit darauf, auf die ohne Abgabe des Versprechens optimale Handlung unter Umständen zu verzichten. Die Frage ist aber: Warum sollte ein Utilitarist dies tun?

Die Antwort ist, dass die unter der Voraussetzung, dass kein Versprechen gegeben wurde, erwartete optimale zukünftige Handlung mitunter schlechter sein kann als die abweichend hiervon erwartete zukünftige Handlung *nach Abgabe eines Versprechens*. Dies wird regelmäßig dann der Fall sein, wenn der Gesamtnutzen einer Handlung, die ein Einhalten des Versprechens konstituiert, deutlich größer ist als der Gesamtnutzen *derselben, jedoch nicht versprochenen* Handlung. Der Grund hierfür kann in der Koordinierungswirkung eines Versprechens liegen, die es dem utilitaristischen Versprechensnehmer überhaupt erst ermöglicht, seinerseits Handlungen in Betracht zu ziehen, die nur dann überaus nutzenbringend sind, wenn der Versprechensgeber sein Versprechen hält.⁶⁹ Es ist die Existenz solcher positiven Nutzeneffekte koordinierten Handelns, die dem utilitaristischen Versprechensgeber nicht nur die glaubhafte Abgabe von Versprechen ermöglicht, sondern ihn qua utilitaristischer Logik *unter bestimmten Umständen* sogar zur Abgabe von Versprechen verpflichtet.

Im hinreichend bekannten Gedankenexperiment des Gefangenendilemmas beispielsweise wären Utilitaristen verpflichtet, sich kurz vor Gefangennahme gegenseitig der Kooperation zu versichern. Eine bloße

⁶⁹ Ähnlich Scanlon (1990, 1998: 304), der den Wert der durch eine Zusicherung gewonnenen Gewissheit betont, ohne explizit utilitaristisch zu argumentieren.

Erörterung der Nutzeneffekte der Handlungen „Gestehen“ und „Leugnen“ allein würde hier noch keine Verpflichtung zu koordiniertem Handeln zur Folge haben, falls Unsicherheit über das jeweilige Handeln des anderen bestünde. Allerdings sind die zwei Utilitaristen nicht zwingend zur Abgabe eines Versprechens verpflichtet, denn eine *einfache Absprache* genügt bereits häufig zur Hebung der Vorteile koordinierten Handelns im Fall von Situationen, die als einfache Koordinationsspiele beschrieben werden können.⁷⁰ Die Abgabe eines Versprechens wird hingegen erforderlich, wenn der Versprechensgeber nicht genau über die Vorteile koordinierten Handelns informiert werden kann⁷¹ oder darf⁷² oder wenn dies zu kostspielig⁷³ wäre.⁷⁴

⁷⁰ Vgl. Güth (1999: 215-251).

⁷¹ Wir weichen an dieser Stelle also die implizit angenommenen idealen Bedingungen etwas auf und gleichen die Annahmen über Informationsstände etwas der realen Welt an, zumindest die der utilitaristischen Versprechensgeber. Denkbar ist beispielsweise, dass der Versprechensgeber aus Zeitgründen nicht im Detail darüber informiert werden kann, warum sein Versprechen wichtig ist.

⁷² Wenn er von den potentiellen großen positiven Nutzeneffekten wissen dürfte, dann könnte er einfach über diese informiert werden und würde als Utilitarist so handeln, dass sie realisiert würden. Es gibt aber solche Situationen, in denen es essentiell ist, dass der Versprechende nicht weiß, warum er etwas versprechen soll. Ein Beispiel wäre das Gunnar abgerungene Versprechen gegenüber Julia, ihr an seinem Geburtstag beim Umzug zu helfen, der aber nur ein Vorwand für eine Überraschungsparty ist.

⁷³ Es ließe sich leicht eine Situation modellieren, in der in einer Gesellschaft von Utilitaristen mit Informations- oder Überzeugungskosten die Institution des Versprechens kostensenkend wirken würde. Gegenüber dem Abschluss notarieller Verträge beispielsweise zeichnen sich bindende Versprechen durch deutlich geringere Transaktionskosten aus.

⁷⁴ Natürlich besteht hier immer noch ein Koordinationsproblem. Der utilitaristische Versprechensgeber, der nicht genau über die Höhe der positiven Nutzeneffekte des Versprechens informiert ist, muss sich auf die diesbezüglichen Signale des potentiellen Versprechensnehmers verlassen können. Da wir an dieser Stelle von utilitaristischen Versprechensnehmern und vollständiger Information auf ihrer Seite ausgehen, ist dieses Problem hier nicht virulent.

Es zeigt sich also, dass die Institution des Versprechens in utilitaristischen Gesellschaften das Gesamtnutzenniveau steigern kann, weil sie es den utilitaristischen Akteuren erlaubt, sich glaubhaft selbst zu binden⁷⁵ und so in neuen Gleichgewichten utilitaristische Koordinationsvorteile realisieren können.⁷⁶

Somit ist aber auch klar, dass in utilitaristischen Gesellschaften *manche* Versprechen nicht gegeben werden können, einfach weil sich kein utilitaristischer Akteur glaubhaft daran binden kann, etwas sehr Schlechtes zu tun. Je nach Annahme der Nutzenfunktionen ist dabei die Menge des glaubhaft Versprechbaren eine andere.⁷⁷

Neben diesen unmöglichen Versprechen gibt es jedoch, so hoffe ich in diesem Abschnitt gezeigt zu haben, bei geeigneten Annahmen über Nutzeneffekte von Versprechenseinhaltungen und -brüchen jeweils eine Menge von möglichen, sinnvollen und einzuhaltenden Versprechen in utilitaristischen Gesellschaften.

⁷⁵ Hier sehen wir, dass das aus der Spieltheorie bekannte strategische Verhalten der Selbstbindung nicht nur bei Annahme von eigennutzinteressierten Spielern auftaucht, sondern auch bei im utilitaristischen Sinne altruistischen Spielern. Zur Anwendung der strategischen Selbstbindung in Gefangenendilemmasituationen siehe Schelling (1980). Zur Funktion der Moral als Selbstbindung zur Erreichung besserer Gleichgewichte siehe Baurmann/Kliemmt (1995: 19-23).

⁷⁶ Um diese Koordinationsvorteile in nicht vollständig utilitaristischen Gesellschaften zu sichern, kann es im Einzelfall utilitaristisch geboten sein, die Nichteinhaltung von Versprechen zu sanktionieren und die Geltung der moralischen Regel, dass Versprechen einzuhalten sind, in der Erziehung zu übermitteln.

⁷⁷ In einer utilitaristischen Gesellschaft, die die moralische Verpflichtung, Versprechen einzuhalten, deontologisch auslegt, wie etwa durch die Nutzenfunktionen (5) und (6) beschrieben, gibt es keine Grenze des glaubhaft Versprechbaren. So könnte zum Beispiel glaubhaft versprochen werden, die gesamte Menschheit auszulöschen. Dass dies versprochen werden *kann*, heißt für einen Utilitaristen in dieser gedachten Gesellschaft jedoch noch nicht, dass er dieses Versprechen geben *sollte*.

3.2 „Du sollst nicht töten!“

Nachdem wir unter 3.1 das moralische Gebot, sein Versprechen zu halten, utilitaristisch interpretiert haben, werde ich im Folgenden versuchen, das moralische Verbot „Du sollst nicht töten!“ als allgemeine Handlungsregel für Utilitaristen mit speziellen Nutzenfunktionen darzustellen.

Stellen wir uns dazu eine Gemeinschaft vor, in der das Glück der einzelnen Menschen neben den üblichen Dingen auch negativ davon abhängt, ob sie jemanden töten.⁷⁸ Generell gelte also, dass

$$(7) \quad \frac{\partial U_{T,i}((...), T_1, T_2 \dots T_n)}{\partial T_j} < 0 \text{ für alle } j.$$

Dabei soll $U_{T,i}$ den Nutzen bezeichnen, den die Person i auch in Abhängigkeit der durch sie begangenen Tötungen T_1, T_2, \dots, T_n hat. Beispielhaft sei die additive und vielleicht einfachste Form einer solchen von Tötungen abhängigen Nutzenfunktion vorgestellt:

$$(8) \quad U_{T,i}((...), T_1, T_2 \dots T_n) = U_i(...) - \sum_{j=1}^n a_{ji} T_j .$$

Als Argumente (...) der angeführten Funktionen $U_{T,i}$ und U_i kommen die üblichen Argumente von Nutzenfunktionen in Betracht, zum Beispiel Güterbündel oder auch Ausstattungen mit Rechten. T_j nehme im Fall einer Tötung von j durch i den Wert 1, andernfalls den Wert 0 an. a_{ji} ist ein Gewichtungsfaktor, der den Einfluss der Tötung von j durch und für die Person i widerspiegelt.

⁷⁸ Ähnlich auch Portmore (2009: 329). Der Grund für solche Nutzeneffekte könnte beispielsweise ein schlechtes Gewissen sein.

Verstehen wir das Tötungsverbot als nicht verhandelbar, so können wir uns also keine Situation vorstellen, in der Töten erlaubt oder geboten wäre. Soll dies der Utilitarist genauso sehen, muss der Effekt eines Tötens demnach so negativ sein, dass dieser durch keine anderen Effekte aufgewogen werden kann. Wir könnten die folgende Nutzenfunktion des Individuums i aufstellen:

$$(9) \quad U_{T,i}((...),T) = \begin{cases} -\infty & \text{falls } T = 1 \\ U_i(...) & \text{falls } T = 0 \end{cases}$$

mit $T = 1$, falls Individuum i tötet, andernfalls $T = 0$. Der utilitaristische Akteur muss nun bei der Entscheidung, ob er tötet, die sonstigen, gegebenenfalls auch positiv zu bewertenden, Folgen auf alle zu aggregierenden Nutzen dem Einfluss seines Tötens auf seine Nutzenbilanz gegenüberstellen und wird zu dem Schluss kommen, dass es utilitaristisch verboten ist zu töten. Durch die Setzung der negativen Folgen eines Tötens für das Individuum i auf $-\infty$, aber auch durch die prohibitiv hohe Setzung des λ im nachfolgend betrachteten Fall wird quasi eine lexikographische Priorität der Variablen T über die Argumente von $U_i(...)$ ausgedrückt. Dies hat die beabsichtigte Folge, dass ein Abweichen von der als absolut verstandenen abgebildeten moralischen Regel für einen Utilitaristen unter keinen Umständen in Betracht kommt.⁷⁹

⁷⁹ Schwierigkeiten könnte es hingegen geben, wenn mehrere als absolut angenommene moralische Regeln auf diese Weise in den Nutzenfunktionen utilitaristisch modelliert werden sollen. Ich bin jedoch zuversichtlich, dass wir immer dann, wenn die Menge moralischer Regeln konsistent ist, auch geeignete Zahlenverhältnisse für die in dieser Weise aufzustellenden Nutzenfunktionen finden können. Das heißt aber auch, dass wir nicht erwarten können, echte moralische Dilemmata mithilfe der Modellierung äquivalenter utilitaristisch zu maximierender Nutzenfunktionen auflösen zu können.

Hier hätten wir aber den unerwünschten Effekt, dass nach dem ersten Mord die folgenden Morde keinen negativen Effekt auf die Glücksbilanz mehr hätten.⁸⁰ Dies ließe sich leicht dadurch beheben, dass die Anzahl der Tötungen gezählt wird und für jedes weitere Töten ein noch negativerer Wert anzusetzen ist. Nehmen wir jedoch einen zu kleinen negativen Wert für je eine Tötung an, so wäre es vorstellbar, dass er durch (etwa in einem Szenario konstruierte) anderweitige positive Effekte aufgewogen wird. Eine Lösung zur Modellierung der deontologischen Position des absoluten Tötungsverbot es wäre hier beispielsweise, eine anzunehmende maximale Glücksmenge im Leben eines Menschen mit der angenommenen maximalen Anzahl der betrachteten Menschen zu multiplizieren ($= \lambda$) und den Wert eines jeden Tötens mit dem Negativen hiervon ($= -\lambda$) in der Nutzenfunktion anzusetzen.⁸¹ Wir erhalten:

$$(10) \quad U_{T,i}((\dots), T_1, T_2 \dots T_n) = U_i(\dots) - \lambda \sum_{j=1}^n T_j$$

mit $T_j = 1$, für das j-te Töten durch Individuum i. Hier wird also für jeden Fall des Tötens bei dem Tötenden mehr Glück abgezogen, als alle Menschen jemals zusammen haben könnten. Das Töten kann unter diesen Voraussetzungen folglich nie glücksmaximierende Handlung sein. Sollten uns unsere moralischen Überlegungen zum Schluss kommen lassen, dass es geboten ist, einen Einzelnen zu töten, sobald man nur dafür mindestens n Menschen retten kann, so ließe sich dies durch geeignete Wahl des λ in Formel (10) oder a_{ji} in Formel (8) und in Abhän-

⁸⁰ Zugegeben ein eher theoretisches Problem. Ein Utilitarist mit einer solchen Nutzenfunktion würde natürlich niemals über eine zweite Tötung nachdenken müssen, da er die erste bereits verworfen hätte. Eine Tötung einer Vielzahl von Menschen würde zudem ebenso ausgeschlossen wie die Tötung nur eines Menschen.

⁸¹ Vgl. Fußnote 38.

gigkeit der $U_j(\dots)$ der zu rettenden Menschen ebenso modellieren. Alternativ, und vielleicht nicht weniger plausibel, wäre es möglich, die negativen Nutzeneffekte einer Tötung nicht nur dem Tötenden zuzuschreiben, sondern auch beim Getöteten zu verorten.⁸²

Ebenso ließe sich dafür argumentieren, dass bei allen Menschen ein gewisser negativer Nutzeneinfluss wirksam wird, weil sie sich weniger sicher fühlen können, nicht selbst getötet zu werden, oder weil sie nicht gerne in einer Gesellschaft leben möchten, deren Mitglieder sich gegenseitig töten. In einer anderen Ausgestaltung könnte man annehmen, dass jeder Mensch Freunde oder Verwandte hat, die mehr als andere unter einer Tötung dieses Menschen leiden würden.

Egal, wie genau man die prohibitiv negativen Nutzeneinflüsse von Tötungen modelliert – und die in diesem Unterabschnitt vorgeschlagenen sollen nur beispielhaft angeführt sein –, es ist deutlich geworden, dass wir bei geeigneter Wahl der Nutzenfunktionen die alltagsmoralische Regel „Du sollst nicht töten!“ als Spezialfall des Utilitarismus interpretieren können.

⁸² Dies ist möglich, obwohl der Getötete nicht weiterlebt. Der negative Nutzeneffekt besteht hierbei jedoch, anders als beim Tötenden, gerade nicht darin, dass er den Rest seines Lebens mit dem Makel dieser Tat leben muss. Hierzu bleibt dem Getöteten keine Zeit. Denkbar wäre eine starke negative Nutzenbeeinflussung im Moment der Kenntnisnahme der eigenen Tötung, beispielsweise durch Schmerz, Unverständnis, Wut oder Enttäuschung. Aber auch eine überraschende, schmerzlose und unbemerkte Tötung hat einen negativen Nutzeneinfluss beim Getöteten zur Folge, der allerdings nicht gesondert modelliert werden muss, da er darin besteht, dass jeglicher sonst im Fall des Weiterlebens entstehende Nutzen für ihn entfällt.

4 Drei Einwände gegen den Utilitarismus

Wie bereits im Kapitel 2 angemerkt, besteht Kritik am Utilitarismus häufig darin, eine Differenz zwischen unseren alltagsmoralischen Überzeugungen und dem vermeintlichen normativen Gehalt des Utilitarismus zu beschreiben. Im Folgenden werde ich drei solcher Einwände gegen den Utilitarismus vorstellen und jeweils dafür argumentieren, dass keine *notwendige* Unvereinbarkeit von Alltagsmoral und Utilitarismus vorliegt. Mein Ziel ist dabei nicht, den Utilitarismus als bestimmte normative Moraltheorie zu verteidigen. Ich werde deshalb anders als viele Autoren weder dafür argumentieren, dass unsere alltagsmoralischen Einstellungen andere sein sollten, noch dafür argumentieren, dass der Utilitarismus bestimmte normative Urteile enthalten müsse. Stattdessen werde ich im Folgenden zu jeweils angenommenen und für die Zwecke des Arguments akzeptierten alltagsmoralischen Regeln passende Nutzenfunktion diskutieren, die auch einen Utilitaristen auf die Einhaltung dieser Regeln verpflichten würden.

4.1 Freundschaft und Parteilichkeit

Häufig wurde die Unvereinbarkeit von Alltagsmoral und Utilitarismus damit begründet, dass der Utilitarismus von uns eine starke Unparteilichkeit fordert, der wir im alltagsmoralisch durchaus erlaubten bevorzugenden Umgang mit Freunden und Familie nicht entsprechen.⁸³ Hintergrund dieser Position ist, dass der Utilitarismus vom Handelnden

⁸³ Dies ist im Kern der im zweiten Kapitel angesprochene Nearest-and-dearest-Einwand gegen den Utilitarismus, der in der Literatur häufig im Zusammenhang mit dem so genannten Argument der Overdemandingness diskutiert wird. Siehe die klassische Quelle Williams (1979) sowie für eine weitergehende Diskussion Jackson (1991) und Mulgan (2007: 124). Die so genannte Kritik der Overdemandingness des Utilitarismus wird im übernächsten Unterkapitel betrachtet.

bekanntlich fordert, unter den möglichen Handlungsoptionen die jeweils beste auszuwählen. Dies legt die Einschätzung nahe, dass im Zuge dieses Abwägungsprozesses alltagsmoralisch wohlfundierte Freundschaftsbande oder Parteilichkeit keine Rolle spielen dürfen. Wir können dieser Kritik den Wind aus den Segeln nehmen, indem wir zwischen geforderter Unparteilichkeit auf verschiedenen Ebenen unterscheiden. Richtig ist, dass *bei gegebenen Nutzenfunktionen* der utilitaristisch Handelnde keine anderen Entscheidungsgrößen wie etwa Parteilichkeit beachten darf. Dies bedeutet jedoch nicht, dass sein Handeln in einem starken Sinn unparteilich sein muss, denn das utilitaristisch Gebotene hängt von den zugrunde liegenden Nutzenfunktionen ab, die durchaus eine Präferenz für parteiliches Handeln ausdrücken können. Hierzu müssen wir nur annehmen, dass die Existenz spezieller Beziehungen zwischen Menschen einen Niederschlag in den jeweiligen Nutzenfunktionen findet.⁸⁴

Im Folgenden werde ich am Beispiel des Schenkens illustrieren, wie sich durch geeignete Wahl der Nutzenfunktionen der unparteilich maximierende Utilitarist einem alltagsmoralischen Gebot der Parteilichkeit entsprechend verhält.⁸⁵ Nehmen wir dazu an, dass ein Utilitarist vor der

⁸⁴ Stellen wir uns den gegenteiligen Fall vor: eine Freundschaft zwischen Menschen, bei der das Wohlbefinden der Befreundeten durch alle möglichen Handlungen auf exakt die gleiche Weise beeinflusst wird, wie es in dem Fall wäre, wenn zwischen ihnen keine Freundschaft bestünde. Es ist fraglich, ob dies überhaupt eine Freundschaft genannt werden kann, denn im Allgemeinen beeinflussen uns Handlungen unserer Freunde anders, als es die Handlungen von Fremden tun.

⁸⁵ Natürlich gibt es verschiedene Arten von Freundschaften und speziellen Relationen zwischen Menschen. Das Schenken soll an dieser Stelle repräsentativ für Handlungen stehen, denen innerhalb von Freundschaften und anderen speziellen Relationen eine besondere Bedeutung zugesprochen wird. Andere Praktiken, die als konstitutiv für Freundschaften und andere spezielle Relationen anzusehen sind, lassen sich analog der Vorgehensweise unter 3.1.2 utilitaristisch interpretieren, denn sie zeigen große Ähnlichkeiten zur Praxis des Versprechens.

Wahl steht, einem von drei Menschen ein Geschenk zu machen, bestehend aus materiellen Gütern mit dem Wert $G > 0$. Die Ausgangssituation der Ausstattung mit materiellen Gütern (MG) sei die folgende:

A: 100

B: 100

C: 80

Nehmen wir weiterhin an, dass $U_i = U(MG_i)$, dass MG einen fallenden Grenznutzen haben und dass die utilitaristische Gesamtfunktion einfach die Summe der U_i ist. Unter diesen Annahmen ist es offenkundig glücksmaximierend, C das Geschenk zukommen zu lassen. Wenn das Geschenk teilbar ist und $G > 20$, so wäre es glücksmaximierend das Geschenk so zu verteilen, dass $MG_1 = MG_2 = MG_3$. Das Gleiche gilt für den Fall, dass A, B und C mit derselben Ausstattung an MG starten. Es wäre also nicht glücksmaximierend das Geschenk lediglich einem zu geben, auch nicht, wenn er ein Freund ist.

Dies liegt genau daran, dass wir die Eigenschaften oder speziellen Umstände von Freundschaft im obigen Modell nicht berücksichtigt haben. Um die Effekte von Freundschaft auf das Glück der Individuen im Fall eines Geschenks zu modellieren, nehmen wir an, dass Geschenke an Freunde und von Freunden anders bewertet werden als Geschenke von Fremden. Betrachten wir beispielsweise die Nutzenfunktionen $U_i = U(MG_i) + W_i(g_i)$, wobei $W_i(g)$ der Nutzen eines Geschenks g aus der Menge der gelungenen, gut gemeinten und wohlverstandenen Geschenke zwischen Freunden ist, der den Nutzeneffekt aus der Erhöhung der MG_i in $U(MG_i)$ übersteigt. Somit ist die Bewertung des Geschenks nicht für alle potentiell Beschenkten gleich, sondern im obigen Beispiel für A größer, weil anzunehmenderweise A ein Freund des utilitaristischen Schenkenden ist. Wir sehen, dass der Utilitarist bei geeigneter Wahl der Nutzenfunktion hinsichtlich der Größe des abnehmenden Grenznutzens

der MG_i und hinsichtlich der $W_i(g_i)$ zum Ergebnis kommen wird, seinen Freund und keinen anderen zu beschenken.⁸⁶

Eine andere Weise, Freundschaften oder andere spezielle Relationen in Nutzenfunktionen zu modellieren, sind interdependente Nutzenfunktionen. Die hierbei betrachteten Nutzenfunktionen haben den Nutzen anderer Individuen als Argument.⁸⁷ Für den Utilitaristen ergibt sich dadurch je nach Ausgestaltung der Nutzenfunktionen unter Umständen das Problem rekursiver Nutzenfunktionen. Betrachten wir beispielsweise folgende interdependente Nutzenfunktionen

$$(11) \quad U_i^I(U_1, U_2, \dots, U_n) = U_i + \sum_{j=1, j \neq i}^n \lambda_{ij} U_j^I,$$

wobei λ_{ij} die Gewichtung des Gesamtnutzens von j für den Gesamtnutzen von i festlegt. Falls nun $\lambda_{ij} \neq 0$ und $\lambda_{ji} \neq 0$, so sind die Gesamtnutzen von j und i zwar nur rekursiv definiert, werden aber bei geeignet gewählten Werten für die λ nach endlich vielen Iterationsschritten stabile Werte annehmen.⁸⁸ Wenn der Utilitarist nun den Nutzen über alle i aggregiert, wird er die indirekten Nutzeneffekte von Nutzenveränderungen zu berücksichtigen haben. Prinzipiell können wir also auf diese Weise mit interdependenten Nutzenfunktionen die Wertschätzungen modellieren, die einzelne Individuen dem Wohlergehen anderer entgegen-

⁸⁶ Analog zu Geschenken kann hier von Dienstleistungen, Freundschaftsdiensten, Aufmerksamkeit oder auch einfach von gemeinsam verbrachter Zeit gesprochen werden. Die Argumentationsstruktur bleibt dabei jeweils dieselbe.

⁸⁷ So ließe sich modellieren, dass uns Freunde und Familie besonders am Herzen liegen und ihr Wohlergehen einen Einfluss auf unser Wohlergehen hat. Falls dies der Fall ist, muss der Utilitarist dies beachten.

⁸⁸ Ein Erläuterungsbeispiel: Wenn der Nutzen von Max positiv durch den Nutzen von Moritz beeinflusst wird, dieser wiederum aber auch vom Nutzen von Max abhängt, so müssen einige Denkschritte durchlaufen werden, um etwas über den Nutzen der beiden sagen zu können. Broome (1991: 182) gibt ein Beispiel für einen einfachen Zweipersonenfall.

genbringen, und somit parteiliches Verhalten in speziellen Beziehungen als Spezialfall des Utilitarismus interpretieren.⁸⁹

So könnte es bei Vorliegen solcher Nutzenfunktionen beispielsweise utilitaristisch geboten sein, einer Person bei einer Entscheidung zur exklusiven Rettung den Vorzug zu geben, deren Wohlbefinden für viele andere Personen sehr wichtig ist, obwohl der Einfluss auf den direkten Nutzen dieser Person *geringer* sein möge als bei den anderen, nicht geretteten Personen.⁹⁰

Der aufmerksame Leser wird die große Ähnlichkeit zur Argumentation für utilitaristische Versprechen unter 3.1 bemerkt haben. Während dort das Geben eines Versprechens annahmegemäß die Bewertung zukünftiger Handlungen durch die beteiligten Personen veränderte, ist es hier das Eingehen von Freundschaften, die anzunehmenderweise den Wert zukünftiger Handlungen verändern.

Es bleibt die Frage, warum ein Utilitarist überhaupt Freundschaften schließen sollte. Ein Grund gegen Freundschaften könnte für den Utilitaristen mithin sein, dass ihn eingegangene Freundschaften in Zukunft dazu zwingen würden, Handlungen zu begehen, die *unter gleichen Voraussetzungen* weniger gut sind als die Alternativen, die ihm zukünftig offenstünden, wäre er diese Freundschaften nicht eingegangen. Wie im Fall des Versprechens verändert aber auch der Utilitarist diese Voraussetzungen, indem er Freundschaften eingeht. Er wird deshalb Freundschaften und andere spezielle Beziehungen zu begründen suchen, *weil* solche Beziehungen ein besonders geeigneter Ort sind, Gutes zu tun. Wenn wir, um bei den Nutzenfunktionen zu bleiben, im ersten Beispiel

⁸⁹ Vgl. hierzu auch die Betrachtung unter 5.3.2. Die Annahme interdependenter Nutzenfunktionen selbst wird unter 7.2 diskutiert und gerechtfertigt.

⁹⁰ In einem durch entsprechende Nutzenfunktionen konkretisierten Beispiel müsste der Utilitarist den von den Massen geliebten Star retten, nicht aber mehrere Kinder mit jeweils angenommen potentiell höheren *direkten* Nutzeneffekten.

nur genügend große Werte für $W_i(g)$ annehmen, wird auch der unparteilichste Utilitarist ein glühender Verfechter der Freundschaft sein.⁹¹

Die eingangs in diesem Abschnitt vorgestellte Kritik am Utilitarismus trifft also nur insofern zu, als es dem Utilitaristen nicht erlaubt ist, in den Fällen parteilich zu sein, in denen keine aggregierte Präferenz für parteiliches Handeln besteht. Wie dieser Abschnitt zeigen sollte, ist der Utilitarismus durchaus zusammen mit aggregierten Präferenzen für Parteilichkeit vorstellbar. Es besteht also kein notwendiger Unterschied zwischen unparteilich-utilitaristischem Maximieren und parteilichem Handeln. Die eingangs vorgestellte Kritik am Utilitarismus und an der Vereinbarkeit von Alltagsmoral und Utilitarismus vermag somit nicht zu überzeugen, denn sie setzt einen einseitigen und verengten Begriff des Utilitarismus voraus.

4.2 Die Unterscheidung zwischen Tun und Unterlassen

Häufig wird dem Utilitarismus vorgeworfen, dass er die moralisch relevante Unterscheidung zwischen Tun und Unterlassen missachtet.⁹² Auf den Punkt gebracht lautet der Vorwurf wie folgt: Indem der Utilitarismus allein auf die Konsequenzen des Handelns rekurriert, das in einer Tat oder in dem Unterlassen einer Tat bestehen kann, ist er blind für die moralische Relevanz dessen, ob eine Veränderung in der Welt aktiv herbeigeführt wurde oder ob diese lediglich nicht verhindert wurde.⁹³

⁹¹ Ebenso wird an dieser Stelle mit Spezialisierungsvorteilen argumentiert. Als Freund kennt man seine Freunde besser, als man es als Fremder je könnte, und kann umso besser wissen, wie man ihnen Gutes tun kann. Vgl. Bailey (1997: 147ff.) oder auch Jackson (1991; 473ff.), der diesbezüglich von spezifischem Vertrauenskapital spricht.

⁹² Siehe beispielsweise Williams (1979: 56ff, 73) und Mulgan (2007: 138). Einführend zur Unterscheidung von Tun und Unterlassen sei Birnbacher (1995) empfohlen.

⁹³ Als Beispiel sei die unterschiedliche moralische und rechtliche Bewertung einer Tötung und einer unterlassenen Hilfeleistung mit Todesfolge genannt.

Verdeutlichen wir uns dieses Argument und präzisieren wir es anhand eines Gedankenbeispiels: Stellen wir uns zwei Welten vor, die sich in ihrem Ablauf *ausschließlich* dadurch unterscheiden, dass eine bestimmte Folge in der einen Welt durch eine Tat T und *genau diese Folge* in der anderen Welt durch die Unterlassung U dieser Tat bewirkt werden könnte.⁹⁴ Der Utilitarist muss nun, so die Kritik, die Tat T in der einen Welt und die Unterlassung U in der anderen Welt absolut gleich bewerten, da diese annahmegemäß zu gleichen Konsequenzen führen. Anders jedoch die oben erwähnte Kritik am Utilitarismus, die auch in diesem Szenario einen moralisch relevanten Unterschied zwischen Tun und Unterlassen sieht.

Undiskutiert lassen möchte ich an dieser Stelle die Frage, ob eine solche Unterscheidung zwischen Tun und Unterlassen *wirklich* moralisch relevant ist.⁹⁵ Weiter werde ich nicht diskutieren, ob überhaupt sinnvoll von einem *Unterschied* zwischen T in der einen und U in der anderen Welt gesprochen werden kann, obwohl zwischen diesen wie angenommen eine vollständige kausale Äquivalenz besteht. Stattdessen werde ich in diesem Abschnitt dafür argumentieren, dass auch Utilitaristen T und U unterschiedlich bewerten können, ja müssen, wenn für sie diese Unterscheidung relevant ist.

Erinnern wir uns dazu an die Ausführung zum utilitaristischen Tötungsverbot unter 3.2, in der sachverhaltsabhängige Präferenzen

⁹⁴ Diese Präzisierung ist notwendig: Niemals wird in *einer* Welt eine Tat *dieselben* Konsequenzen haben wie eine Unterlassung. Um im Beispiel der vorherigen Fußnote zu bleiben: Im Allgemeinen unterscheiden sich Tötungen von unterlassenen Hilfeleistungen in einer Weise, die den Utilitaristen bei geeigneten Nutzenfunktionen darauf verpflichtet, die beiden *Handlungen* unterschiedlich zu bewerten. Der Vorwurf der Ignoranz gegenüber der moralischen Relevanz der Unterscheidung zwischen Tun und Unterlassen könnte dem Utilitarismus in diesem Fall somit nicht in jedem Fall gemacht werden.

⁹⁵ Der interessierte Leser sei hierzu auf Bennett (1966) und die daran anschließende Diskussion verwiesen.

zugrunde gelegt wurden. Dabei ergaben sich annahmegemäß Nutzeneffekte aus dem Sachverhalt „dass der Tötende tötete“.⁹⁶ Aus der Struktur der angenommenen Nutzenfunktionen (8), (9) und (10) folgt direkt, dass andere Handlungen – zum Beispiel Unterlassungen –, die zwar *ansonsten* dieselben Folgen haben, aber angenommenerweise keine Tötungen darstellen, grundlegend anders zu bewerten sind als Handlungen, die Tötungen darstellen.⁹⁷

Allgemein ließe sich der Unterschied zwischen Tun und Unterlassen in folgender Formel fassen:

$$(12) \quad U_{T,i}((\dots), T_1, T_2 \dots T_n) = U_i(\dots) - \sum_{j=1}^n a_{ij} T_j .$$

Als Argumente (...) der angeführten Funktionen $U_{T,i}$ kommen die üblichen Argumente von Nutzenfunktionen in Betracht, zum Beispiel Güterbündel oder auch Ausstattungen mit Rechten. T_j nehme im Fall einer Nichtunterlassung der Handlung j durch i den Wert 1, andernfalls den Wert 0 an. a_{ij} ist ein Gewichtungsfaktor, der den Einfluss der Nichtunterlassung der Handlung j durch und für die Person i widerspiegelt.⁹⁸

⁹⁶ Die Annahme von solchen sachverhaltsabhängigen Präferenzen wird unter 7.1 diskutiert und plausibilisiert.

⁹⁷ Natürlich liegt hier der Vorwurf der Spitzfindigkeit auf der Hand, denn fraglich könnte ja sein, ob Unterlassungen mit Todesfolge nicht auch als Tötungen im Sinne der Formeln (8), (9) und (10) zu zählen wären. Man muss aber an dieser Stelle dem Utilitarismus dieselben Differenzierungsmöglichkeiten zubilligen wie seinen Kritikern.

⁹⁸ Ein solcher Einfluss könnte in den psychischen Kosten bestehen, die mit manchen Taten verbunden sind. Vgl. hierzu Williams (1979: 65f.), der allerdings hiermit nicht den Utilitarismus verteidigt.

Falls Nutzenfunktionen dieses Typs vorliegen, wird auch bei sonst gleichen Folgen ein Unterschied in den Bewertungen von Unterlassungen und Nichtunterlassungen vorliegen.

Wir sehen daran, dass der Utilitarismus nicht notwendig ignorant ist bezüglich der Unterscheidung zwischen Tun und Unterlassen. Die eingangs in diesem Abschnitt erwähnte Kritik kann deshalb nicht überzeugen, wenn sie den Utilitarismus *an sich* zum Gegenstand hat. Sie vermag nur den Utilitarismus zu treffen, der sich keiner unterschiedlichen aggregierten Präferenz bezüglich Unterlassungen und Nichtunterlassungen gegenübergestellt sieht.

4.3 Exkurs: der Einwand der Überforderung

An dieser Stelle werde ich abseits der Hauptargumentation in gebotener Kürze einen weiteren Einwand gegen die Vereinbarkeit von Utilitarismus und Alltagsmoral vorstellen. Das Argument ist hierbei das folgende⁹⁹: Der Utilitarismus unterscheidet sich wesentlich von der Alltagsmoral, weil er anders als unsere Alltagsmoral eine strikte Gebots- und Verbotsethik ist. Betrachten wir noch einmal die Handlungsregeln des Utilitarismus, um diesen Einwand besser zu verstehen:

(U_g) Eine Handlung H* eines Akteurs ist genau dann moralisch geboten, wenn sie bessere Folgen nach sich zieht als *jede* ihrer Handlungsalternativen H₁, H₂, H₃, ... , die dem Akteur zum selben Zeitpunkt offenstehen.

(U_v) Eine Handlung H- eines Akteurs ist genau dann moralisch verboten, wenn es unter ihren Handlungsalternativen H₁, H₂, H₃ eine Handlung H_i gibt, die bessere Folgen nach sich zieht als H-.

⁹⁹ Vgl. zum Beispiel Williams (1979: 45-56) oder Mulgan (2007: 131-146).

(U_e) Eine Handlung H₀ eines Akteurs ist genau dann moralisch erlaubt, wenn es unter ihren Handlungsalternativen H₁, H₂, H₃ keine Handlung H_i gibt, die bessere Folgen nach sich zieht als H₀.

Für den Utilitaristen gibt es diesem Schema zufolge in den allermeisten Fällen nur eine erlaubte und somit gebotene Handlung $H = H^*$. Im Rest der Fälle gibt es eine Klasse von erlaubten Handlungen H₀, von denen jedoch eine zwingend zu wählen ist. In unserer Alltagsmoral hingegen gibt es neben klaren Ge- und Verboten eine große Menge an erlaubten und somit keineswegs zwingend erforderlichen Handlungen. Dieser Unterschied sei es, der für eine Unvereinbarkeit von Utilitarismus und Alltagsmoral spreche.

Allerdings lässt sich dieser scheinbare Unterschied zwischen Alltagsmoral und Utilitarismus leicht erklären. Er resultiert aus der Voraussetzung, dass dem utilitaristischen Entscheider alle Folgen und somit auch die Nutzeneffekte aller möglichen Handlungen in einer gegebenen Entscheidungssituation bekannt sind. Abstrahiert man von dieser Idealisierung und denkt man sich etwa Unsicherheit, Informationsbeschaffungskosten, Entscheidungsfindungskosten (beispielsweise durch Zeitablauf) hinzu, so verändern sich die Grenzen zwischen den Mengen der in einer Entscheidungssituation jeweils gebotenen, verbotenen und erlaubten Handlungen.¹⁰⁰

Nehmen wir beispielsweise an, ein Utilitarist habe vier Handlungen zur Verfügung und erwartet begründet folgende unsicheren Gesamtnutzeneffekte dieser Handlungen:

H₁: U ∈ (10, 20)

H₂: U ∈ (5, 25)

¹⁰⁰ Erneut sei der interessierte Leser auf Bailey (1997) verwiesen, der dieses Argument ausführlich entwickelt.

H₃: U ∈ (15)

H₄: U ∈ (0, -10)

Diese Konstellation könnte sich zum Beispiel ergeben, wenn sich ein Utilitarist drei Möglichkeiten H₁ bis H₃ gegenübergestellt sieht, eine Menschenmenge vor einer Bombenexplosion zu bewahren, deren genauer Erfolg im Fall von H₁ und H₂ unsicher ist. H₄ stehe hier für das genaue Evaluieren der Handlungsmöglichkeiten H₁ bis H₃, das allerdings aufgrund des damit einhergehenden Zeitverlustes angesichts der tickenden Zeitbombe klar die schlechteste Alternative sei. In dieser Situation sind im utilitaristischen Kalkül H₁ bis H₃ erlaubte Handlungen, H₄ jedoch eine verbotene Handlung. Aus diesem Beispiel wird deutlich, dass sich ein nicht-idealer Utilitarist in nicht-idealen Entscheidungssituationen einer Vielzahl von erlaubten Handlungen gegenübergestellt sehen kann.¹⁰¹ Ebenso ist es denkbar, dass ein Utilitarist in weniger dramatischen Situationen aus einer Vielzahl von erlaubten Handlungen wählen darf, von denen auch eine darin bestehen könnte, zunächst in aller Ruhe nachzudenken oder auf eine bessere Gelegenheit für eine gute Tat zu warten.

Somit zeigt sich, dass der Einwand der Überforderung nur einen idealen Utilitarismus trifft, nicht jedoch *den Utilitarismus an sich*. Denn wie ich in dieser Arbeit versuche zu zeigen, ist der Utilitarismus selbst noch keine normative Theorie, sondern vielmehr eine Methode. Einen normativen Gehalt bekommt ein Utilitarismus erst durch Hinzudenken von Nutzenfunktionen und, wie wir in diesem Unterabschnitt ergänzen müssen, durch Hinzudenken oder Vorliegen von Kontextbedingungen. Zu den Kontextbedingungen zählt neben weiteren Determinanten der Entscheidungssituation auch die epistemische Verfassung des utilitaris-

¹⁰¹ Zudem ist stets eine utilitaristische Gleichwertigkeit von Handlungsalternativen modellierbar, was ebenso jeweils eine Vielzahl von gleichermaßen erlaubten Handlungen zur Folge hat. Siehe den Beweis im Abschnitt 9.

tischen Entscheiders. Gehen wir, wie in diesem Abschnitt, von nicht-idealen Entscheidern aus, so wird der normative Gehalt des Utilitarismus regelmäßig ein anderer sein, als wenn wir ideale Entscheider annehmen.

Der Einwand der Überforderung als Kritik an der Vereinbarkeit von Utilitarismus und Alltagsmoral oder als Kritik *am Utilitarismus an sich* kann schlicht deshalb nicht überzeugen, weil er irreführenderweise eine bestimmte ideale Form des Utilitarismus einer bestimmten Menge an moralischen Regeln bei speziellen Kontextbedingungen gegenüberstellt.

Abweichend vom Argumentationsgang der restlichen Kapitel besteht das verteidigende Argument hier neben der normativen Untersättigung des Utilitarismus hinsichtlich der Nutzenfunktionen darüber hinaus auch noch in der normativen Untersättigung hinsichtlich der Kontextbedingungen der Entscheidungssituation und der epistemischen Verfassung des Entscheiders. Weil ich diesen abweichenden Ansatz im weiteren Verlauf nicht weiter verfolgen möchte, bleibt dieses Unterkapitel, was es ist: ein Exkurs. Gehen wir deshalb im Folgenden weiter wie im Kapitel 2 festgelegt von idealen Entscheidern aus.

5 Egalitärer Utilitarismus

5.1 Egalitarismus

Der Egalitarismus vertritt die Position, dass Gleichheit zwischen Individuen etwas Gutes ist.¹⁰² Dabei ist erstens zu klären, *für wen* die Gleichheit zwischen Individuen etwas Gutes ist, und zweitens, *welche Gleichheit genau* zwischen Individuen als etwas Gutes eingeschätzt wird.

Die erste Frage wird häufig erst gar nicht gestellt und wenn doch, dann wird sie so beantwortet, dass Gleichheit zwischen Individuen etwas *an sich Gutes* sei oder dass es *gut für die Gesellschaft* sei, in der sie leben. Gleichheit wird hier also als ein suprapersonales Gut gesehen. Indem ich annehme, dass Individuen Präferenzen bezüglich Gleichheit haben können, werde ich in diesem Kapitel versuchen, egalitäre Intuitionen utilitaristisch zu interpretieren.

Vorher ist jedoch die zweite Frage zu klären, welche Gleichheit genau zwischen Individuen als etwas Gutes eingeschätzt wird oder bezüglich welcher Eigenschaften eine Gleichheit von Individuen als etwas Gutes beurteilt wird. Plausible und in der Literatur diskutierte Kandidaten sind beispielsweise die Gleichheit von Einkommen¹⁰³, von Vermögen, von Wohlergehen, von Glück, von Chancen auf Wohlstand¹⁰⁴ oder auch Gleichheit von Nutzen.¹⁰⁵ Da es hier jeweils strukturelle Parallelitäten in den Argumentationen und in den Beispielen gibt, schlage ich vor, für die weitere Betrachtung der egalitären Position genau zwei Klassen von gewünschten Gleichheiten zu unterscheiden: die Gleichheit bezüglich der

¹⁰² Für eine Übersicht siehe Broome (1991: 174).

¹⁰³ Siehe hier insbesondere Sen (1997).

¹⁰⁴ Sehr überzeugend argumentiert diesbezüglich Roemer (1996: 263ff.).

¹⁰⁵ Vgl. Broome (1991: 174). Für eine Übersicht über weitere Ansätze in der Literatur siehe Bailey (1997: 126).

Ausstattung mit materiellen Gütern (MG) und die Gleichheit von Wohlergehen oder auch von Nutzen¹⁰⁶ (U).

Um überhaupt sinnvoll von mehr oder weniger Gleichheit sprechen zu können, müssen wir dabei die Ausstattungen mit materiellen Gütern, aber auch die Nutzen der Individuen vergleichen können. Im Fall der materiellen Güter ist das einfach. Natürlich wird von der egalitären Position nicht die Gleichheit der Güterbündel gefordert, sondern die Gleichheit *des Wertes* dieser Güterbündel. In einer funktionierenden Tauschökonomie ist daher der Vergleich der Ausstattung mit MG kein besonderes Problem, leicht ließen sich ein Nullpunkt und eine Einheit MG definieren.

Der Vergleich von Nutzen bereitet dagegen bedeutend größere theoretische Probleme. Grundlegend ist hier die Kritik von Robbins zu nennen, der keine wissenschaftliche Basis für interpersonale Vergleiche sieht.¹⁰⁷ Ähnlich schreibt Johansson: „Unfortunately, there is no universally accepted principle for making [...] interpersonal comparisons. In fact some economists argue that such interpersonal comparisons are impossible; there is no ‘scientific’ way of comparing utility or ‘happiness’ across households.“¹⁰⁸ Dies ist natürlich nicht nur ein Problem des Egalitarismus allein, sondern sämtlicher Positionen, die in irgendeiner Weise Nutzen von verschiedenen Individuen gemeinsam betrachten.¹⁰⁹

¹⁰⁶ Der Leser sei an die Ausführungen des Kapitels 2 und die Festlegung erinnert, die Zielgröße des Utilitarismus mit dem Terminus „Nutzen“ auszudrücken. Erneut sei daran erinnert, dass es für die theoretische Struktur des Utilitarismus unerheblich ist, was genau diese Zielgröße ist.

¹⁰⁷ Robbins (1938).

¹⁰⁸ Johansson (1995: 38).

¹⁰⁹ Zum Beispiel sind interpersonale Nutzenvergleiche für jede utilitaristische Abwägung vorauszusetzen. Siehe dazu die Anmerkungen und die Festlegung in Kapitel 2. Ebenso vgl. Bailey (1997: 5) und Roemer (1996: 128f., 160).

Allerdings ist Sen Recht zu geben, der darauf hinweist, dass wir interpersonale Vergleiche praktisch und systematisch anstellen, beispielsweise, wenn wir sagen, dass wir in einer bestimmten Situation lieber die Person A als die Person B wären.¹¹⁰ Ich möchte das Problem der interpersonalen Vergleichbarkeit an dieser Stelle deswegen nicht weiter diskutieren.¹¹¹ Der Egalitarismus muss, um überhaupt Aussagen über bessere oder schlechtere Verteilungen treffen zu können, die Erkennbarkeit dieser Verteilungen voraussetzen. Dass sich der Utilitarismus, wie wir in Kapitel 2 gesehen haben, in derselben Position befindet, ist somit kein prinzipielles Argument gegen eine utilitaristische Modellierbarkeit egalitärer Bewertungen von Verteilungen oder gegen weiter unten vorzustellende egalitäre Präferenzen. Gehen wir also für die Zwecke der Argumentation im Folgenden von einer interpersonalen Vergleichbarkeit von Nutzen aus.¹¹²

¹¹⁰ Sen (1997: 14).

¹¹¹ Ein weiteres Problem, das für die hier vorgestellten Formen des Egalitarismus nicht virulent ist, wäre das Problem der *intertemporalen* Vergleichbarkeit von Nutzen. Die Frage ist hier, wie weit man Vergleichen vertrauen kann, die man zwischen Nutzenniveaus einer Person *zu verschiedenen Zeiten* anstellt. Es ergeben sich ähnliche Probleme, wie sie bereits Wittgenstein im wohlbekannten Privatsprachenargument erkannt hat.

¹¹² Einen anderen, aber sehr fruchtbaren Weg geht Bailey an dieser Stelle. Er untersucht verschiedene Welten, in denen interpersonale Nutzenvergleiche mit jeweils unterschiedlicher Sicherheit möglich sind. Er kommt zu dem Ergebnis, dass der Utilitarist umso mehr Umverteilung wollen kann, je genauer die interpersonalen Nutzenvergleiche angestellt werden können. In einem weiteren Schritt differenziert er zwischen solchen Gütern, deren Nutzeneffekte auf andere leichter erkennbar sind, und solchen, bei denen dies nicht der Fall ist, und argumentiert aus der utilitaristischen Position kraftvoll und überzeugend für eine Absicherung des sozialen Minimums durch Umverteilung, da in diesen Fällen die ebenso utilitaristisch begründbaren prozeduralen Gerechtigkeitsauffassungen von Locke, Nozick und Gauthier ausgestochen werden. Vgl. Bailey (1997: 100-102, 120-144).

5.2 Egalitäre Kritik am Utilitarismus

Die egalitäre Kritik am Utilitarismus besteht im Wesentlichen in dem Vorwurf, der Utilitarismus wäre indifferent gegenüber Fragen der Verteilung, der Gleichheit oder der Gerechtigkeit.¹¹³ Vielleicht der prominenteste Vertreter dieser Kritik ist Sen, der dem Utilitarismus einen perversen Umgang mit Verteilungsungerechtigkeiten vorwirft¹¹⁴ und feststellt: „It seems fairly clear that fundamentally utilitarianism is very far from an egalitarian approach.“¹¹⁵

Bereits an dieser Stelle ist anzumerken, dass ein prinzipieller Unterschied zwischen Ungleichheit in Einkommen oder materiellen Gütern und Ungleichheit in Nutzen besteht. Diesen Punkt betont auch Johansson: „Note, however, that we are talking about inequality in terms of units of *utility*. This does not indicate indifference to inequality of *incomes*, because the marginal utility derived from an additional dollar may vary between individuals.“¹¹⁶ Die auf diesen Umstand bauende utilitaristische Standardantwort auf egalitäre Kritik an Ungleichheiten in Einkommen oder materiellen Gütern wird unter 5.3.1 vorgestellt. Sen widerspricht dieser Standardantwort des egalitären Utilitarismus, indem er in einem Gegenbeispiel die scheinbar notwendig perversen und anti-egalitären Forderungen des Utilitarismus herausstellt.¹¹⁷ Ich werde gegen die Kritik von Sen argumentieren, indem ich die Annahmen von Sen, speziell die von ihm zugrunde gelegten Nutzenfunktionen, in Zweifel ziehe. Auch für Gleichheiten, die über die utilitaristisch durch abnehmende Grenznutzen wünschbare Gleichheit hinausgehen, lässt sich utilitaristisch argumentieren. Zu diesem Zweck werden in 5.3.2 Nutzen-

¹¹³ Vgl. beispielsweise Johansson (1995: 33), Williams (1979: 102), Mulgan (2007: 163), Sen (1997: 16), Rawls (1971: 26), Roemer (1996: 130) und Bailey (1997: 5).

¹¹⁴ Siehe Sen (1997: 111).

¹¹⁵ Sen (1997: 18).

¹¹⁶ Johansson (1995: 34), Hervorhebungen im Original.

¹¹⁷ Sen (1997: 16).

interdependenzen betrachtet. Eine weitere zu beleuchtende egalitäre Position, dass Ungleichheiten *in Nutzen* etwas Schlechtes sind, werde ich unter 5.3.2 utilitaristisch interpretieren. Dabei wird für den Zweck der Modellierung erneut davon ausgegangen, dass etwas einen Wert *für irgendjemanden* hat, wenn es überhaupt einen Wert hat.¹¹⁸

5.3 Zwei egalitäre Varianten des Utilitarismus

5.3.1 Utilitaristische Gleichverteilung aus abnehmendem Grenznutzen

In diesem Abschnitt werde ich ein einfaches und inzwischen klassisch¹¹⁹ zu nennendes Argument explizieren, das dafür spricht, dass auch die utilitaristische Bewertung egalitäre Züge tragen kann. Unter relativ einfachen und plausiblen Annahmen wird das utilitaristische Kalkül dabei gleichere Verteilungen¹²⁰ von materiellen Gütern als besser bewerten als ungleichere Verteilungen.¹²¹

Nehmen wir an, dass alle betrachteten n Individuen die gleiche Nutzenfunktion mit positivem, aber abnehmendem Grenznutzen in MG_i haben, also zum Beispiel:

¹¹⁸ Vgl. das an den Anfang dieser Arbeit gestellte Zitat von Schopenhauer sowie Broome (1991: 187).

¹¹⁹ Vgl. Marshall (1920: 471), Pigou (1932: 89) sowie zum Beispiel Sen (1997: 16, 83f.), Johansson (1995: 34), Bailey (1997: 102) und Broome (1991: 175ff.).

¹²⁰ Im Folgenden wird mit Gleichverteilung keine Wahrscheinlichkeitsverteilung beschrieben, sondern eine gleiche Verteilung materieller Güter.

¹²¹ Im Unterschied zu der Vorgehensweise in den Kapiteln 3 und 4 geht es hier also nicht mehr um Bewertungen von Handlungen, sondern um Bewertungen von Verteilungen von MG. Es ist allerdings leicht ersichtlich, wie diese Bewertungen in Bewertungen von Handlungen überführt werden könnten, die die Verteilung von MG beeinflussen.

$$(13) \quad U_i = U(MG_i) \text{ mit } \frac{\partial U}{\partial MG_i} > 0 \text{ und } \frac{\partial^2 U}{\partial^2 MG_i} < 0 .$$

Nehmen wir weiter an, dass die Menge der zu verteilenden Güter $MG = MG_1 + MG_2 + \dots + MG_n$ konstant ist, egal wie man diese Menge auf die n Individuen aufteilt, so lässt sich leicht folgern, dass die utilitaristisch gebotene Verteilung eine Gleichverteilung ist. Denn nehmen wir einmal das Gegenteil an: Es gäbe mindestens ein i und j , so dass $MG_i > MG_j$. Aus den identischen Nutzenfunktionen mit positivem und abnehmendem Grenznutzen folgte dann, dass

$$\frac{\partial U}{\partial MG_i} < \frac{\partial U}{\partial MG_j} .$$

Da in die utilitaristische Zielfunktion alle U_i gleichgewichtet eingehen, ist es somit utilitaristisch geboten, MG von Individuum i zu Individuum j umzuverteilen, da der Verlust an Nutzen bei Individuum i kleiner ist als der Gewinn an Nutzen bei Individuum j , solange $MG_i > MG_j$. Bei diesem simplen Modell sorgen also die angenommenen Eigenschaften der Nutzenfunktionen für eine utilitaristische Bevorzugung der Gleichverteilung. Die Annahmen des positiven, aber abnehmenden Grenznutzens von MG sind hierbei Standardannahmen für Nutzenfunktionen, die ich an dieser Stelle nicht weiter zu begründen brauche.¹²² Die Annahmen identischer Nutzenfunktionen und der Konstanz der Gesamtmenge der materiellen Güter MG hingegen sind echte Einschränkungen der Allgemeinheit dieser Argumentation. Gegen die Konstanz der Gesamtmenge

¹²² Jedes Kind kennt jedoch die anekdotische Evidenz, dass die fünfte Kugel Eis weniger mundet als die erste.

der MG würden zum Beispiel Umverteilungskosten erster und zweiter Art sprechen. Umverteilungskosten erster Art wären Kosten, die direkt bei der Umverteilung der MG von einem Individuum zu einem anderen Individuum anfallen könnten. Je nachdem, wie hoch diese sind, wird ein gewisses Maß an Ungleichverteilung¹²³ utilitaristisch geboten sein, jedoch wird auch unter Annahme dieser Umverteilungskosten der Utilitarist in diesem Szenario niemals für eine Umverteilung sein, die die Ungleichverteilung erhöht.¹²⁴ Umverteilungskosten zweiter Art wären Kosten, die aus der Anpassung des Verhaltens der Individuen aufgrund der Umverteilung resultieren könnten. Hängt beispielsweise die Menge der zur Verfügung stehenden MG vom Arbeitseinsatz der Individuen ab, der selbst wieder von ihren Anreizen abhängt, die selbst durch Umverteilung abnehmen können, so liegen Umverteilungskosten zweiter Art vor.¹²⁵ Je nachdem, wie hoch diese Nutzenverluste durch Anreizverluste sind, wird ein gewisses Maß an Ungleichverteilung utilitaristisch geboten sein.

Insgesamt bleibt also auch bei Existenz gewisser Umverteilungskosten die generelle Folgerung bestehen, dass der Utilitarismus bei gleichen Nutzenfunktionen mit positivem, aber abnehmendem Grenznutzen gewisse egalitäre Züge aufweist. Genau dies stellt auch Bailey fest: „In general, however, the principle is a fairly sound normative generalization [gleicher abnehmender Grenznutzen von MG], which provides a good reason for a fair amount of egalitarianism in distribution.“¹²⁶

¹²³ Der Utilitarist würde jede Umverteilung gutheißen, bei der die Gewinne aus der Umverteilung zwischen Individuen mit unterschiedlichen Grenznutzen die dabei anfallenden Nutzenverluste durch Verteilungskosten übersteigen.

¹²⁴ Denn zu den Nutzenverlusten durch Verteilungskosten kämen hier nur negative Effekte aus der Umverteilung von Individuen mit größerem Grenznutzen zu Individuen mit kleinerem Grenznutzen.

¹²⁵ Vgl. Johansson (1995: 35f.).

¹²⁶ Bailey (1997: 102).

Die zweite oben angesprochene Einschränkung der Allgemeinheit dieses Arguments verdient eine etwas intensivere Betrachtung. Sind die Nutzenfunktionen nicht gleich und hat beispielsweise ein Individuum stets einen höheren Grenznutzen von MG als ein anderes Individuum, so wird eine sehr ungleiche Verteilung utilitaristisch geboten sein. Sen kritisiert genau dies als eine perverse und anti-egalitäre Forderung des Utilitarismus. Er konkretisiert seine Kritik anhand eines Beispiels, in dem ein Gesunder jeweils einen doppelt so hohen Nutzen von Einkommen haben soll wie ein körperlich Behinderter.¹²⁷ Der Utilitarist, so Sen, müsste in diesem Szenario ausgehend von einer Gleichverteilung für eine Umverteilung von MG von dem körperlich Behinderten zum Gesunden argumentieren.¹²⁸ Sen schließt daraus: „Consequently, the utilitarian ordering can be quite perverse in dealing with distributional inequality when distinct persons have different utility functions.“¹²⁹

Sicher ist Sen Recht zu geben, wenn er diese Forderung nach Umverteilung von MG von dem körperlich Behinderten zum Gesunden als pervers brandmarkt. Und natürlich ist hier sein Schluss richtig, dass die utilitaristische Bewertung eine perverse sein *kann* und sein muss, *wenn man entsprechende Nutzenfunktionen annimmt*. Es stellt sich jedoch die Frage, ob diese Perversität in der utilitaristischen Methode ihre Ursache hat oder ob sie in den Eigenschaften der angenommenen Nutzenfunktion begründet ist. Denn es ließe sich vortrefflich dafür argumentieren, dass der Grenznutzen von MG für einen körperlich Behinderten eben *nicht kleiner* als der des Gesunden ist, sondern im Gegenteil er-

¹²⁷ Siehe Sen (1997: 16).

¹²⁸ Siehe Sen (1997: 17). Ein Beispiel, das ähnlich motiviert, aber nicht so kontrovers ist, findet sich bei Bailey (1997: 5). Er stellt sich zwei Personen, Lucky und Unlucky, vor, deren Nutzenfunktionen ähnliche Eigenschaften wie im Beispiel von Sen haben, und kommt zum selben Schluss, dass aus utilitaristischer Sicht hier eine Ungleichverteilung gefordert werden müsste.

¹²⁹ Sen (1997: 111).

heblich höher.¹³⁰ Wie dem auch sei, mein Ansatz steht und fällt nicht mit der empirischen Frage, wie es um die Nutzenfunktionen in diesem Beispiel genau bestellt ist. Hieraus wird deutlich, dass es Sens Vorwurf der Perversität des Utilitarismus selbst ist, der von ganz speziell gewählten Nutzenfunktionen abhängt. Da die von Sen vorgestellte Konstellation von Nutzenfunktionen nicht denkunmöglich ist, zeigt sein Argument zwar, dass der Utilitarismus *nicht notwendigerweise* egalitäre Verteilungen bevorzugen muss. Viel weiter kommt Sen mit seinem Beispiel jedoch nicht. Denn ebenso denkbar ist auch der gegenteilige Fall, dass Nutzenfunktionen relativ gleich und mit positivem, aber abnehmendem Grenznutzen sind, sowie dass nicht allzu hohe Umverteilungskosten bestehen. Daraus folgt, wie wir gesehen haben, dass der Utilitarismus nicht notwendigerweise anti-egalitäre Verteilungen bevorzugen muss. Je stärker abnehmend die Grenznutzen, je geringer die Umverteilungskosten und je gleicher die Nutzenfunktionen, desto mehr wird der Utilitarismus egalitäre Züge aufweisen.

Es zeigt sich hieran, dass der Utilitarismus an sich keine *eigene* Präferenz für eine mehr oder weniger egalitäre Verteilung hat. Die egalitäre Kritik am Utilitarismus trifft somit nur in dem Maße zu, wie er keine genuin egalitäre Theorie ist. Denn dem Utilitarismus kommt es allein auf die Summe der Nutzen an und so ergibt sich eine utilitaristische Besser- oder Schlechterbewertung gleicherer Verteilungen allein aus den Nutzenfunktionen der betrachteten Individuen.

Wenn wir also die Standardannahmen der Nutzentheorie treffen, also einen positiven, aber abnehmenden Grenznutzen von materiellen Gütern, so werden gleichere Verteilungen der materiellen Güter zwischen

¹³⁰ Das würdevolle Leben eines körperlich Behinderten wird beispielsweise eine speziell auf seine Bedürfnisse zugeschnittene Wohnungseinrichtung oder ständige Betreuung erfordern. Im Umkehrschluss heißt dies: Das Fehlen von MG, die dies ermöglichen, hat große negative Nutzensauswirkungen – aller Wahrscheinlichkeit nach größere als beim Gesunden.

den Individuen als utilitaristisch besser bewertet, zumindest wenn wir davon ausgehen, dass die betrachteten Individuen nicht allzu unterschiedlich sind.¹³¹ Diejenigen Egalitaristen, denen ein solcher Utilitarismus bereits ausreichend egalitär ist, könnten wir als schwache Egalitaristen bezeichnen.

Der moderate Egalitarist bezüglich der Verteilung materieller Güter wäre durch diese Argumentation hingegen noch nicht zufrieden gestellt, weil er darauf bestehen würde, Ungleichverteilungen aus noch *weiteren* Gründen zu kritisieren als dem der abnehmenden Grenznutzen. Denn wir haben in diesem Abschnitt gesehen, dass, wenn die oben diskutierten restriktiven Bedingungen (gleiche Nutzenfunktionen, positive, aber abnehmende Grenznutzen, keine Transferkosten) gelockert werden, der Utilitarist zwar tendenziell für gleichere Verteilungen optieren würde, jedoch nicht notwendigerweise für das vom moderaten Egalitaristen gewünschte Maß¹³² an Gleichverteilung. Um auch diese egalitäre Position als Spezialfall des Utilitarismus interpretieren zu können, werden wir im nächsten Abschnitt interdependente Nutzenfunktionen als ein weiteres Instrument der utilitaristischen Modellierung betrachten.¹³³

¹³¹ Diese Annahmen sind, so denke ich, nicht allzu unplausibel. Aber bereits ein solcher anspruchsloser Utilitarismus reicht meines Erachtens aus, die große Verteilungsungerechtigkeit auf der Welt zu kritisieren.

¹³² Den aufmerksamen Leser, der zu Recht bemängelt, dass wir ein Maß der Gleichverteilung oder Ungleichverteilung noch nicht definiert haben, bitte ich noch um zwei Seiten Geduld.

¹³³ Natürlich reichen die in diesem Abschnitt angenommenen Eigenschaften der Nutzenfunktionen bereits aus, um für eine utilitaristische Gleichverteilung zu argumentieren. Theoretisch sind wir hier also schon am Ziel, indem wir gezeigt haben, dass es eine Welt geben kann, in der utilitaristisch die Gleichverteilung optimal wäre. Auf diese Weise die egalitäre Kritik am Utilitarismus beiseitezuschieben, hieße aber, es sich zu leicht zu machen. Das Argument wird im nächsten Abschnitt entscheidend dadurch an Wert gewinnen, dass wir weniger restriktive Annahmen über Nutzenfunktionen treffen.

5.3.2 Utilitaristische Gleichverteilung aus interdependenten Nutzenfunktionen

Wir haben im letzten Abschnitt gesehen, dass der Utilitarist für gleichere Verteilungen argumentieren kann, wenn abnehmende Grenznutzen vorliegen. Dies gilt eingeschränkt auch bei Existenz von Verteilungskosten. Der Utilitarist ist also in diesem Fall bereit, für eine gleichere Verteilung eine Reduktion der Summe der zu verteilenden MG hinzunehmen. Dies deswegen, weil annahmegemäß die kleinere, aber gleich verteilt Menge MG zu aggregiert höheren Nutzen führt. Die Argumentation in diesem Kapitel wird darauf hinauslaufen, dass es Umstände geben kann, unter denen der kluge Utilitarist sogar eine Reduktion der Gesamtmenge *der direkten Nutzen aus materiellen Gütern gutheißen* wird. Natürlich ist der Utilitarist dabei weiterhin daran gebunden, diejenige Verteilung zu bevorzugen, die den höchsten Gesamtnutzen verspricht. Was ich also vorschlage, ist, zwischen direkten und indirekten Nutzen aus materiellen Gütern zu unterscheiden. Als direkten Nutzen materieller Güter werde ich den Nutzen bezeichnen, der bei dem Individuum anfällt, dem diese Güter zukommen. Mit indirektem Nutzen materieller Güter möchte ich diejenigen Nutzeneffekte bezeichnen, die bei *allen anderen* Individuen daraus entstehen, dass genau diesem Individuum diese materiellen Güter zukommen.¹³⁴ Bevor ich dazu übergehe, diese Unterscheidung durch interdependente Nutzenfunktionen zu verdeutlichen, werde ich aber noch einige Maße der Ungleichverteilung sowie verschiedene Typen von Egalitaristen vorstellen.

¹³⁴ Eine ähnliche sprachliche Einteilung findet sich bei Broome (1991: 180), allerdings für den weniger allgemeinen Zweipersonenfall und bezogen auf indirekte Nutzeneffekte von Nutzen und nicht wie hier auf indirekte Nutzeneffekte aus MG.

Das Standardwerk über ökonomische Ungleichheiten¹³⁵ ist Sens „On economic inequality“.¹³⁶ Im zweiten Kapitel in Sens Buch werden verschiedene Maße von Ungleichverteilungen vorgestellt¹³⁷, von denen ich an dieser Stelle nur die beiden einfachsten Varianten aufführen möchte: die relative Spannbreite und die mittlere relative Abweichung.¹³⁸

Betrachten wir die Verteilung der materiellen Güter MG auf n Personen und bezeichnen die Ausstattung der Person i mit materiellen Gütern mit MG_i , bezeichnen wir weiter das größte MG_i mit MG_H und das kleinste MG_i mit MG_L und den arithmetischen Durchschnitt der MG_i mit MG_{av} , dann ist die relative Spannbreite

$$(14) \quad S = \frac{MG_H - MG_L}{MG_{av}},$$

also der Unterschied zwischen den extremen Werten von MG_i , gemessen in Einheiten des arithmetischen Durchschnitts der MG_i . Die relative Spannbreite S wird dabei Werte zwischen 0, wenn alle MG_i gleich sind,

¹³⁵ Sen spricht über Einkommen. Ich betrachte materielle Güter (MG), um auch Vermögenseffekte einzuschließen.

¹³⁶ Sen (1997).

¹³⁷ Sen (1997: 24-46).

¹³⁸ Weiter unten soll gezeigt werden, wie durch Annahme von interdependenten Nutzenfunktionen, die Präferenzen für Gleichverteilung ausdrücken, die utilitaristisch optimale Verteilung ein egalitär gewünschtes Maß an Gleichverteilung einhalten wird. Ich beschränke mich hier auf die beiden einfachsten Maße von Ungleichverteilungen. Generell bin ich zuversichtlich, dass auch die komplizierteren Maße auf diese Weise behandelt werden können, weil alle Werte, die Sen zur Berechnung seiner Maße einsetzt – die materiellen Güter eines jeden Individuums – prinzipiell auch den interdependenten Nutzenfunktionen als Argumente zur Verfügung stehen und somit einer Abbildbarkeit der egalitären Wertschätzung einer gewissen Gleichverteilung in den Nutzenfunktionen nichts entgegensteht.

also $MG_H = MG_L = MG_{av}$, und n annehmen, wenn MG_H maximal und alle anderen $MG_i = MG_L = 0$.

Die mittlere relative Abweichung vom Mittelwert ergibt sich aus

$$(15) \quad A = \sum_i \frac{|MG_i - MG_{av}|}{nMG_{av}}$$

und nimmt Werte zwischen 0 und $2(n - 1)/n$ an.¹³⁹

Somit haben wir nun zwei Maßzahlen von Ungleichverteilungen definiert. Bei der Modellierung der egalitären Positionen als Spezialformen des Utilitarismus werde ich weiter unten in den anzunehmenden Nutzenfunktionen eine *aggregierte utilitaristische Abneigung* gegen die in diesem Sinne zu messenden Ungleichheiten in der Verteilung materieller Güter ausdrücken. Bevor wir dazu kommen, müssen wir allerdings die *Höhe der Abneigung* gegen diese so gemessenen Ungleichheiten betrachten. Dies bringt uns zu der Unterscheidung von verschiedenen egalitaristischen Positionen.

Dem schwachen Egalitaristen, den wir bereits im letzten Abschnitt untersucht haben, sind Ungleichverteilungen nur insofern etwas Schlechtes, wie sie die Summe der direkten Nutzen der Individuen aus ihren materiellen Gütern verringern.¹⁴⁰

¹³⁹ Sen diskutiert und kritisiert die hier vorgestellten und andere Maßzahlen der Ungleichverteilung. Diese Diskussion soll an dieser Stelle nicht weiter kommentiert oder wiederholt werden. Der interessierte Leser sei auf Sen (1997: 24-46) verwiesen.

¹⁴⁰ Dies ist exakt die Position des Utilitaristen, wenn keine interdependenten Nutzenfunktionen angenommen werden. Je nach Nutzenfunktionen kann der optimale Grad an Ungleichheit also auch ein von null verschiedener Wert sein, wenn die im letzten Abschnitt diskutierten Bedingungen nicht gegeben sind. Letztlich ließe sich auch die These vertreten, dass hier keine genuin egalitäre Forderung durch den Utilitarismus

Die Position des moderaten Egalitarismus¹⁴¹ bezüglich der Verteilung materieller Güter geht darüber hinaus: Sie bewertet gleichere Verteilungen als besser, obwohl diese die Summe der *direkten Nutzen* der Individuen aus ihren materiellen Gütern nicht maximieren. Diese Position wird deswegen moderat genannt, weil sie nicht *nur* auf die Gleichverteilung achtet, sondern *auch* auf die Summe der direkten Nutzen. Der moderate Egalitarist sieht sich somit einem Trade-off zwischen der Gleichverteilung materieller Güter und der Summe der direkten Nutzen der Individuen aus ihren materiellen Gütern gegenübergestellt.¹⁴²

Die extreme Position des Egalitarismus bezüglich der Verteilung materieller Güter hingegen ist die, dass ungleichere Verteilungen immer schlechter sind als gleichere, *egal wie groß* der Unterschied in den Summen der direkten Nutzen der Individuen aus ihren materiellen Gütern dabei ist.¹⁴³

Nun endlich haben wir alles beisammen, um zum Kernstück dieses Abschnitts voranschreiten zu können: der Vorstellung interdependenter Nutzenfunktionen und der Argumentation dafür, dass eine Annahme solcher Nutzenfunktionen der utilitaristischen Position die nachzubil-

erhoben wird. Man könnte dieses utilitaristische Kalkül, das von nicht-interdependenten Nutzenfunktionen ausgeht, somit auch als *egalitaristisch neutral* bezeichnen, denn es werden weder interdependente Nutzenfunktionen angenommen, die eine *Gleichverteilung* über den Ausgleich von Grenznutzen hinaus utilitaristisch vorteilhafter erscheinen ließen, noch werden interdependente Nutzenfunktionen angenommen, die eine solche *Ungleichverteilung* als utilitaristisch vorteilhafter erscheinen ließen.

¹⁴¹ Vgl. die Einordnungen bei Bailey (1997: 11f.).

¹⁴² Vgl. hierzu Blackorby/Donaldson (1977), die mehrere moderate Versionen des Egalitarismus untersuchen.

¹⁴³ Vgl. Broome (1991: 184f.). Nicht in dieses Schema einordnen lässt sich die viel beachtete Theorie von Rawls (1971), die mit ihrer Fokussierung auf den Schlechtestgestellten in einer Gesellschaft eher als lexikalischer Prioritarismus zu bezeichnen wäre. Die Rawls'sche Position ist Gegenstand des nächsten Kapitels.

denden egalitären Züge verpasst.¹⁴⁴ Dabei werde ich jeweils eine utilitaristische Version der moralischen Position angeben, die durch Kombination der moderaten oder extremen egalitären Position mit einem der vorgestellten Maße der Ungleichverteilung entsteht.

Betrachten wir zunächst einen Typ von Nutzenfunktionen, durch die auch Nutzeneffekte ausgedrückt werden, die ein Individuum aufgrund von Ungleichverteilungen hat, die durch S und A charakterisiert sind:

$$(16) \quad U_{i,S,A}(MG_i, S, A) \text{ mit } \frac{\partial U_{i,S,A}}{\partial MG_i} > 0 \text{ und } \frac{\partial^2 U_{i,S,A}}{\partial^2 MG_i} < 0 .$$

Im Folgenden werde ich ein Beispiel für eine solche Nutzenfunktion vorstellen¹⁴⁵ und betrachten:

$$(17) \quad U_{i,S}(MG_i, S) = U_i(MG_i) - \alpha_i S .$$

Individuen mit diesen Nutzenfunktionen haben also einen Nutzeneinfluss in Höhe von $\alpha_i S$, der eine gewisse Präferenz bezüglich der relativen Spannbreite der Verteilung der materiellen Güter ausdrückt.¹⁴⁶ An die-

¹⁴⁴ Die prinzipielle Möglichkeit einer Vereinbarkeit von Egalitarismus und Utilitarismus scheint auch Roemer (1996: 130) zu sehen: „The theorems of utilitarianism, it turns out, do not challenge, the usual kind of aversion to inequality.“ Er konkretisiert diesen Gedanken jedoch nicht.

¹⁴⁵ Für einen ähnlichen Ansatz, jedoch rudimentärer und nur für zwei Individuen, siehe Broome (1991: 180).

¹⁴⁶ Eine Klarstellung: Ziel dieser Arbeit ist es nicht, die Sinnhaftigkeit dieses speziellen Maßes und der anderen Maße von Ungleichverteilungen zu diskutieren. Sen selbst zieht diese überzeugend in Zweifel, vgl. Sen (1997: 24-46). Es soll hier anhand von Beispielen gezeigt werden, wie der Egalitarismus als Spezialfall des Utilitarismus interpretiert werden kann, wenn man interdependente Nutzenfunktionen und die damit

sem Beispiel lässt sich die oben getroffene Unterscheidung zwischen direktem und indirektem Nutzen aus materiellen Gütern illustrieren. Direkter Nutzen aus materiellen Gütern ist der Teil des Nutzens, der dem Individuum i aus seinen MG_i erwächst. Indirekter Nutzen aus materiellen Gütern sind die Nutzeneffekte aus der Verteilung der materiellen Güter auf die Individuen, hier durch $\alpha_i S$ repräsentiert.

Die individuelle Einschätzung α_i der relativen Spannbreite könnte dabei zwischen den Individuen differieren. Denkbar wäre beispielsweise, dass für Individuen mit größerer Ausstattung von materiellen Gütern das α_i kleiner ist als bei den Individuen mit geringerer Ausstattung von materiellen Gütern. Dies entspräche einer höheren Aversion gegen gesamtgesellschaftliche Ungleichverteilungen bei den materiell Schlechtergestellten als bei den materiell Bessergestellten einer Gesellschaft. In einem solchen Fall litte annahmegemäß der materiell Schlechtestgestellte also *mehr* unter einer Vergrößerung der Spannbreite als der Bestgestellte. Denkbar wäre natürlich auch jede andere Kombination der α_i , insbesondere also auch diejenige, dass alle $\alpha_i = 0$. In diesem Fall gäbe es *keine* individuellen Präferenzen bezüglich der relativen Spannbreite der Verteilung der materiellen Güter und es bestünde keine Differenz zwischen $U_{i,S}(MG_i, S)$ und dem bisher betrachteten einfachen $U_i(MG_i)$.

Entscheidend ist aber in jedem Fall die aggregierte Präferenz bezüglich der relativen Spannbreite der Verteilung der materiellen Güter. Denn der Utilitarist, der die Summe über die einzelnen $U_{i,S}$ bildet, erhält die durch die Verteilung der materiellen Güter zu maximierende Gesamtsumme

verbundene Unterscheidung von direktem und indirektem Nutzen aus materiellen Gütern betrachtet. Weiter unten in diesem Abschnitt werde ich generell für die Sinnhaftigkeit von interdependenten Nutzenfunktionen argumentieren.

$$(18) \quad \sum_i U_{i,S}(MG_i, S) = \sum_i U_i(MG_i) - \sum_i \alpha_i S .$$

Er wird also bei der Wahl der Verteilung der materiellen Güter neben der Summe der direkten Nutzen der materiellen Güter auch auf die Summe der indirekten Nutzeneffekte aus der Verteilung der materiellen Güter auf die Individuen zu achten haben. Je nach dem Wert der $\sum \alpha_i$ wird der Utilitarist bereit sein, von der größten Summe der direkten Nutzen aus materiellen Gütern abzuweichen¹⁴⁷, soweit die indirekten Nutzeneffekte (aus einer aggregierten Bewertung der relativen Spannweite der Verteilung von MG) überwiegen. Dies ist exakt die moderate egalitäre Position.

In genau gleicher Weise lässt sich die egalitäre Position der Negativbewertung einer mittleren relativen Abweichung der MG_i als Spezialfall des Utilitarismus verstehen. Nehmen wir dazu an, dass die Individuen einer Gesellschaft Nutzenfunktionen des folgenden Typs haben:

$$(19) \quad U_{i,A}(MG_i, A) = U_i(MG_i) - \alpha_i |MG_i - MG_{av}| .$$

Der Utilitarist, der nun die Summe über die einzelnen $U_{i,A}$ bildet, erhält die durch die Verteilung der materiellen Güter zu maximierende Gesamtsumme

¹⁴⁷ Der Term $\sum \alpha_i$ bestimmt also die relative Wichtigkeit der mit S gemessenen Ungleichverteilung im egalitären utilitaristischen Kalkül.

$$(20) \quad \sum_i U_{i,A}(MG_i, A) = \sum_i U_i(MG_i) - \sum_i \alpha_i |MG_i - MG_{av}|.$$

Es lässt sich dabei ein β so finden, dass

$$(21) \quad \sum_i \alpha_i |MG_i - MG_{av}| = \beta \sum_i |MG_i - MG_{av}|,$$

wobei β ein Maß ist für die aggregierte Präferenz bezüglich der Abweichung der Individualausstattungen mit materiellen Gütern vom Mittelwert. Setzen wir nun $\lambda = \beta n MG_{av}$, können wir folglich (20) umschreiben als

$$(22) \quad \begin{aligned} \sum_i U_{i,A}(MG_i, A) &= \sum_i U_i(MG_i) - \lambda \sum_i \frac{|MG_i - MG_{av}|}{n MG_{AV}} \\ &= \sum_i U_i(MG_i) - \lambda A, \end{aligned}$$

was gezeigt werden sollte. Wieder muss der Utilitarist bei der Wahl der Verteilung der materiellen Güter nicht nur die Summe der direkten Nutzen der materiellen Güter maximieren, sondern auch auf den Einfluss der indirekten Nutzeneffekte aus der Verteilung der materiellen Güter auf die Individuen achten. Genau wie es der moderate Egalitarist fordert, wird er also bereit sein, von der utilitaristisch optimalen Summe der *direkten* Nutzen aus materiellen Gütern abzuweichen, da er die indirekten Nutzen, die sich aus der mittleren relativen Abweichung der MG_i ergeben, in seinem Kalkül mit zu berücksichtigen hat.

Nun lassen sich über die Wahl der α_i in (17) und (19) verschieden strenge egalitäre Positionen als Utilitarismus modellieren. Nehmen wir an, dass durch eine Ungleichverteilung die Summe der direkten Nutzen aus materiellen Gütern erhöht werden kann – beispielsweise trotz abnehmender Grenznutzen durch Anreizeffekte –, so wird der Utilitarist bei angenommenen $\sum \alpha_i \neq 0$ im Fall von (17) trotzdem nicht die Summe des direkten Nutzens aus MG maximieren, da er natürlich den Gesamtnutzen aus Formel (18) maximieren muss. Generell gilt: Je größer die Summe¹⁴⁸ der α_i , desto weniger Ungleichverteilung wird der Utilitarist bereit sein zu akzeptieren. Die durch einen solchen moderat-egalitären Utilitaristen ausgewählte Menge an Ungleichverteilung ist davon gekennzeichnet, dass sich die gegengerichteten marginalen Effekte aus direkten und indirekten Nutzenveränderungen durch Umverteilung gerade aufheben.

Es lassen sich natürlich prohibitiv große α_i annehmen, so dass der Utilitarist zu einem extremen Egalitaristen wird und prinzipiell keine Ungleichverteilungen akzeptieren wird, egal wie viel durch diese an direktem Nutzen aus materiellen Gütern gewonnen werden könnte.¹⁴⁹ Somit ist gezeigt, wie sich moderate und extreme egalitaristische Positionen bezüglich der Verteilung materieller Güter als utilitaristischer Spezialfall interpretieren lassen.

Bevor ich im nächsten Abschnitt utilitaristische Interpretationen eines extremen Egalitarismus vorstelle, werde ich noch eine weitere egalitäre Position untersuchen, die nicht Ungleichverteilungen materieller Güter, sondern Ungleichheiten im Nutzen der Personen als negativ bewertet.¹⁵⁰ Auch diese werde ich als Spezialfall des Utilitarismus interpretieren,

¹⁴⁸ Die gewichtete Summe der α_i im Fall von (19).

¹⁴⁹ Zur Unterscheidung von moderaten und extremen egalitären Positionen und Kritik an der extremen Position siehe Bailey (1997: 11) und Broome (1991: 184).

¹⁵⁰ Vgl. Broome (1991: 174).

indem ich auf interdependente Nutzenfunktionen zurückgreife. Die hierbei betrachteten Nutzenfunktionen haben als Argument jedoch *direkt den Nutzen anderer Individuen*.¹⁵¹ Hierin unterscheiden sie sich von den weiter oben in diesem Abschnitt vorgestellten Nutzenfunktionen, die als interdependente Elemente *die materiellen Güter anderer Individuen* als Argument hatten. Für den Utilitaristen ergibt sich dadurch je nach Ausgestaltung der Nutzenfunktionen unter Umständen das Problem rekursiver Nutzenfunktionen.¹⁵²

Ein weiterer Unterschied zu den weiter oben in diesem Abschnitt betrachteten Nutzenfunktionen ergibt sich daraus, dass Nutzen *an sich* wohl weniger einfach umzuverteilen ist als materielle Güter. Eine Nutzenumverteilung hingegen, die sich allein aus einer Umverteilung von MG oder Ähnlichem ergibt, steht hier gerade nicht zur Debatte, da sie bereits weiter oben behandelt wurde. Allerdings ist hier auch nicht die Frage zu klären, ob ein solcher Egalitarismus überhaupt handlungsrelevant werden kann, der Ungleichheiten im Nutzen der Personen als negativ bewertet, die sich nicht allein aus Ungleichverteilungen von leichter umzuverteilenden Dingen wie MG oder Ähnlichem ergeben. Denn wir sollten dem Utilitaristen dieselben Handlungsoptionen der Umverteilung von Nutzen zubilligen, die wir auch dem Egalitaristen zugestehen. Gehen wir also im Folgenden zur Modellierung eines Egalitarismus, der Ungleichheiten im Nutzen der Personen als negativ bewertet, davon aus, dass eine solche Bewertung handlungsrelevant sein kann und Nutzen im gewünschten Maß durch geeignete Handlungen umverteilbar sind.

Betrachten wir beispielsweise folgende interdependente Nutzenfunktionen

¹⁵¹ Vgl. Broome (1991: 182).

¹⁵² Vgl. Fußnote 88.

$$(23) \quad U_i^I(MG_i, U_1, U_2, U_{i-1}, U_{i+1} \dots U_n) = U_i(MG_i) + \sum_{j=1, j \neq i}^n \lambda_{ij} U_j^I,$$

wobei λ_{ij} die Gewichtung des Gesamtnutzens von j für den Gesamtnutzen von i festlegt. Falls hierbei $\lambda_{ij} \neq 0$ und $\lambda_{ji} \neq 0$, so sind die Gesamtnutzen von j und i zwar nur rekursiv definiert, werden aber bei geeignet gewählten Werten für die λ nach endlich vielen Iterationsschritten stabile Werte annehmen.

Wenn der Utilitarist nun den Nutzen über alle i aggregiert, wird er die indirekten Nutzeneffekte von möglichen Nutzenumverteilungen zu berücksichtigen haben. Ganz analog zum obigen Vorgehen könnten wir durch Wahl geeigneter λ_{ij} nun Formeln ähnlich wie (17) oder (19) aufstellen und die Argumentation wie oben führen.

Prinzipiell können wir also auf diese Weise mit interdependenten Nutzenfunktionen die Wertschätzung von Nutzengleichverteilungen in den Individuen modellieren und somit egalitäre Positionen bezüglich der wünschenswerten Gleichheit von Nutzen als Spezialfall des Utilitarismus interpretieren.

5.4 Exkurs zur utilitaristischen Form des extremen Egalitarismus

Betrachten wir den oben vorgestellten extremen Egalitarismus, der dadurch gekennzeichnet ist, dass prinzipiell keine Ungleichverteilungen akzeptiert werden, egal wie viel durch diese an direktem Nutzen aus materiellen Gütern gewonnen werden könnte. Für den extremen Egalitaristen wird also jede Menge ungleich verteilter MG_i schlechter sein als irgendeine Menge gleichverteilter MG_i . Es gilt also:

$$(24) \quad \sum_i U_i(MG_1, MG_2, \dots, MG_n) < \sum_i U_i(x, x, \dots, x) \text{ für alle } x \text{ und } MG_i,$$

wobei x die Art oder den Wert der materieller Güter bezeichnet, die jedem Individuum zukommen. Insbesondere gilt diese Schlechterbewertung jeder Ungleichverteilung auch für den Vergleichsfall der vollkommenen Abwesenheit materieller Güter, so dass für alle ungleich verteilten MG_i gilt:

$$(25) \quad \sum_i U_i(MG_1, MG_2, \dots, MG_n) < \sum_i U_i(0, 0, \dots, 0) .$$

Spätestens bei dieser Konsequenz daraus, dass der Wert von Gleichverteilung als absolut gesetzt wird, fällt nun die Kritik am extremen Egalitarismus leicht. Die Argumentation könnte dabei wie folgt lauten:

Die extreme Form des Egalitarismus ist zu verwerfen, weil Menschen, egal wie wichtig ihnen Gleichverteilung auch sein möge, ein gewisses Maß an Ausstattung mit materiellen Gütern bei Ungleichverteilung einzeln und aggregiert lieber ist als eine gleiche, aber für alle völlig unzureichende Ausstattung mit materiellen Gütern.

Letztlich ist dies nicht mehr als die begründete Aussage, dass die Nutzenfunktionen von Menschen nicht vom Typ (17) oder (19) mit prohibitiv großen α_i sind und dass sie im Speziellen nicht die in (24) ausgedrückten Eigenschaften aufweisen. Diese Kritik könnte nun dergestalt präzisiert werden, dass – obiger Begründung folgend – ein Schwellenwert, beispielsweise das Existenzminimum M , angegeben wird, ab dem erst eine Gleichverteilung stets einer Ungleichverteilung von MG_i vorzuziehen ist:

$$(26) \quad \sum_i U_i(MG_1, MG_2, \dots, MG_n) < \sum_i U_i(M, M, \dots, M) .$$

Natürlich wären Gegenargumente des extremen Egalitaristen möglich. Beispielsweise könnte er darauf beharren, dass marginale Unterschreitungen des Existenzminimums für alle gerade noch zu tolerieren sind, wenn man dadurch andernfalls sehr große Ungleichverteilungen verhindern könnte.

Verlassen wir jedoch an dieser Stelle diesen Disput zwischen moderateren und extremeren Egalitaristen und verdeutlichen uns, dass ihre Diskussion, die durchaus als moralische zu charakterisieren ist, als Streit über korrespondierende Nutzenfunktionen geführt wurde. Das Beispiel zeigt somit, wie sich ein moralischer Diskurs über verschiedene Spielarten des Egalitarismus oder auch eine Abwägung zwischen Gesamtnutzen und Gleichheit in einen Disput über Nutzenfunktionen übersetzen lässt. Eine These dieser Arbeit ist ja, dass sich jede konsistente Menge moralischer Regeln – sprich jede konsistente moralische Position – als spezieller Utilitarismus interpretieren lässt. Eine weitere These schließt sich hier nahtlos an: dass sich jeder moralische Disput in eine Argumentation über korrespondierende Nutzenfunktionen überführen lassen kann. Das ist es, was ich als utilitaristische Methode in der Ethik bezeichnen möchte.¹⁵³

5.5 Egalitarismus aus konkaven Funktionen der sozialen Wohlfahrt

Widmen wir uns nun der Frage, ob es einen sinnvollen Egalitarismus geben kann, der *strukturell* anders ist als die in den letzten drei Abschnitten vorgestellten und als Sonderfälle des Utilitarismus interpretierten egalitären Positionen. Vorgegangen waren wir in den letzten drei Kapiteln so, dass wir interdependente Nutzenfunktionen betrachtet ha-

¹⁵³ Diese Diskussion wird im Abschnitt 7.3 fortgesetzt.

ben, durch die eine aggregierte Präferenz für gleichere Verteilungen zum Ausdruck kam, die somit der Utilitarist zu beachten hat.

Fraglich ist, ob es möglich ist, sich einen sinnvollen Egalitarismus vorzustellen, der einen Wert in der Gleichheit von Nutzen sieht, der prinzipiell nicht durch eine solche utilitaristische Interpretation abgebildet werden kann, weil sich dieser Wert prinzipiell nicht in den Nutzenfunktionen der Individuen widerspiegeln kann. Broome nennt einen solchen Egalitarismus „Gemeinschafts-Egalitarismus“ und definiert ihn so: „It is the view that general good can be increased without increasing the total of individual good, by distributing individual good more equally.“¹⁵⁴ Ähnlich ist wohl auch die Position von Scanlon: „Fairness and equality do not represent ways in which individuals may be made *better off*. They are, rather, special morally desirable features of states of affairs and institutions.“¹⁵⁵

Die einzig mögliche utilitaristische Interpretation für einen solchen Gemeinschafts-Egalitarismus besteht darin, eine andere als die in Kapitel 2 gewählte Aggregationsart zu betrachten. Die konkrete Wahl einer solchen Funktion der sozialen Wohlfahrt kann dabei als eine Modellierung eines bestimmten Verteilungsprinzips verstanden werden.

Einen Vorschlag macht Roemer: „As we have seen, if utility is ratio-scale fully comparable, then social choice functions which are highly averse to inequality – as utilitarianism is not – are admissible, namely, those that order states accordingly to $\sum U_i^{1/r}$, where r is a large positive number.“¹⁵⁶ Diese Funktion der sozialen Wohlfahrt unterscheidet sich

¹⁵⁴ Broome (1991: 178).

¹⁵⁵ Scanlon (1977: 81).

¹⁵⁶ Roemer (1996: 160), Notation der Formel angepasst. Um ein mögliches Missverständnis an dieser Stelle auszuräumen: Roemer trifft hier nicht die Aussage, dass der Utilitarismus per se anti-egalitär sein müsse. Er muss hier vielmehr so verstanden

von der normalen Form der utilitaristischen Aggregationsfunktion vor allem dadurch, dass der Einfluss zusätzlicher Individualnutzen mit steigenden Individualnutzen sehr stark abnimmt. Dadurch ist in dieser Aggregationsfunktion gleichsam ein hohes Maß an Aversion gegen Ungleichverteilung ausgedrückt, so dass auch *unabhängig* von einer Schlechterbewertung von Ungleichverteilungen durch die Individuen selbst eine Gleichverteilung zu fordern ist.

Eine solche Wahl einer von der einfachen Summenfunktion der Individualnutzen abweichenden Aggregationsfunktion ist meines Erachtens jedoch aus mehreren Gründen nicht überzeugend. Denn erstens bewerten solche von der einfachen Summenfunktion der Individualnutzen abweichenden Aggregationsfunktionen zusätzliche Nutzen unterschiedlich in Abhängigkeit davon, bei wem sie entstehen, ohne dass hierfür auf den ersten Blick eine überzeugende Begründung vorliegt.

Zweitens bleibt unklar, *wer es ist*, der diese Aversion gegen Ungleichverteilungen hat, die durch solche Funktionen der sozialen Wohlfahrten ausgedrückt werden. Die naheliegende Antwort wäre, dass dies eine Aversion *der Gesellschaft selbst* gegen Ungleichverteilungen ist. Allerdings hätte die Gesellschaft selbst dann eine Wertschätzung für eine Gleichheit oder Gleichverteilung, für die die Gesamtheit der betroffenen Mitglieder der Gesellschaft keine oder gar eine negative Wertschätzung hätte. Es bleibt unklar, in genau welchem Sinn und warum die über das von den Individuen gewünschte Maß hinaus bestehende Gleichheit oder Gleichverteilung *moralisch wünschbar* wäre. Diesen Punkt betont auch Broome, der keine Notwendigkeit für einen solchen Gemeinschafts-Egalitarismus sieht: „If equality is valuable, individualistic egalitaria-

werden, dass die utilitaristische Aggregationsfunktion allein noch keine Festlegung bezüglich der Ungleichheitsaversion trifft.

nism captures its value properly, and communal egalitarianism does not.“¹⁵⁷

Schließlich kann ein solcher Gemeinschafts-Egalitarismus auch nicht damit begründet werden, dass er prinzipiell eine andere oder größere Gleichverteilung materieller Güter fordern könnte als die weiter oben in diesem Kapitel vorgestellten und auf der einfachen Summenfunktion der Individualnutzen basierenden Varianten des egalitären Utilitarismus.¹⁵⁸

Fassen wir zusammen: Zwar lassen sich egalitäre Intuitionen *auch* durch die Wahl von der einfachen Summenfunktion der Individualnutzen abweichender Aggregationsfunktionen ausdrücken, doch hat eine solche Betrachtung einen geringeren explanatorischen Gehalt. Denn eine in der Aggregationsfunktion selbst ausgedrückte Aversion gegen ungleiche Verteilungen verunmöglicht die weitere Diskussion anhand von Nutzenfunktionen darüber, *warum überhaupt* und *warum genau diese* Aversion gegen Ungleichverteilungen angenommen werden soll.

5.6 Zwischenfazit

In diesem Kapitel habe ich Nutzenfunktionen vorgestellt, die dem Utilitarismus mehr oder weniger egalitäre Züge geben können. Im Abschnitt 5.3.1 habe ich die Standardannahme der abnehmenden Grenznutzen materieller Güter als Grund für eine gewisse utilitaristische Wertschätzung von gleicheren Verteilungen materieller Güter vorgestellt.¹⁵⁹ Der moderate Egalitarist fordert jedoch mehr Gleichheit in der Verteilung

¹⁵⁷ Broome (1991: 199).

¹⁵⁸ Dies folgt unmittelbar aus dem in Abschnitt 9 geführten Beweis: Jede Funktion der sozialen Wohlfahrt lässt sich als konsistentes Moralsystem interpretieren, das eine Rangordnung von Verteilungen erzeugt.

¹⁵⁹ Gewisse oben diskutierte und wenig restriktive Annahmen vorausgesetzt.

materieller Güter, als der Utilitarist schon allein aus abnehmenden Grenznutzen fordern würde. Um auch diese Position als eine utilitaristische interpretieren zu können, habe ich im Abschnitt 5.3.2, ähnlich wie schon unter 3.1.2, die Menge der Argumente der Nutzenfunktionen ausgeweitet, so dass sich in den diskutierten Nutzenfunktionen Präferenzen über die Verteilung materieller Güter oder auch direkt von Nutzen widerspiegeln. Dabei konnten wir sehen, dass die extreme egalitäre Position als Sonderfall des moderaten Egalitarismus verstanden werden kann und auf gleiche Weise der utilitaristischen Modellierung zugänglich war.

Im Kapitel 5.4 wurde deutlich, wie eine Diskussion zwischen verschiedenen Spielarten des Egalitarismus in einen Diskurs über die Nutzenfunktionen überführt wurde. Im Kapitel 5.5 schließlich wurde eine weitere denkbare Modellierung egalitärer Bewertungen in konkaven Funktionen der sozialen Wohlfahrt vorgestellt, kritisiert und als explanatorisch minderwertig verworfen.

Betrachten wir nach diesen Ausführungen die in Abschnitt 5.2 vorgestellte Kritik von Sen erneut, die wir hier für unsere Zwecke durch folgendes Zitat aufgreifen können: „It seems to me clear that a concern for equity must militate against the use of utilitarianism.“¹⁶⁰ Ich hoffe dieses Kapitel und insbesondere die Argumentation unter 5.3 haben gezeigt, dass Sen hier falsch liegt und dass kein notwendiger Widerspruch zwischen Egalitarismus und Utilitarismus besteht. Der Utilitarist kann und muss Gleichheit oder Gleichverteilung von materiellen Gütern einen Wert zuordnen, *wenn* die Individuen dies tun, d.h., wenn sie Nutzenfunktionen haben, die solche Präferenzen für Gleichheit oder Gleichverteilung von materiellen Gütern ausdrücken.¹⁶¹ Wir sehen somit, dass der Utilitarismus weder notwendigerweise eine egalitäre noch notwendigerweise eine anti-egalitäre Bewertung nach sich zieht. Vielmehr

¹⁶⁰ Sen (1997: 45).

¹⁶¹ Vgl. Broome (1991: 240).

kommt es auf die jeweils angenommenen Nutzenfunktionen der Individuen an.

Natürlich kann man sich Situationen, sprich: Konstellationen von Nutzenfunktionen, vorstellen, in denen der Utilitarismus anti-egalitär ist. Die egalitäre Kritik am Utilitarismus greift jedoch zu kurz, wenn sie in dem Vorwurf besteht, dass der Utilitarismus nicht in *allen denkbaren* Situationen egalitäre Forderungen erhebt.

6 Der Rawls'sche Utilitarismus

6.1 Rawls' Prinzipien der Gerechtigkeit

Nachdem wir im letzten Kapitel gesehen haben, dass der Utilitarismus nicht notwendig anti-egalitär ist, werde ich in diesem Kapitel die Rawls'sche Position als einen Spezialfall des Utilitarismus modellieren. Diese Gegenüberstellung der Rawls'schen „Theorie der Gerechtigkeit“ und des Utilitarismus ist auch deswegen spannend, weil man die Rawls'sche quasi-kontraktualistische Theorie gerade als Gegenentwurf zum Utilitarismus verstehen kann. Wie der Egalitarist ist Rawls der Meinung, dass die Kriterien der Pareto-Optimalität und Effizienz nicht zur zufrieden stellenden Lösung von Güterverteilungsproblemen ausreichen: „The principle of efficiency does not by itself select one particular distribution of commodities as the efficient one. To select among the efficient distributions some other principle of justice, [...] is necessary.“¹⁶²

Im vorherigen Kapitel haben wir gesehen, wie der Utilitarist zwischen verschiedenen Verteilungsalternativen durchaus in einem egalitaristischen Sinn entscheiden kann. Rawls' Ansatz unterscheidet sich darüber hinaus scheinbar in zwei weiteren und für ihn wesentlichen Punkten vom Utilitarismus: Erstens argumentiert Rawls für eine Maximierung von miteinander zu vereinbarenden Freiheits- und Grundrechten und nicht direkt von Nutzen und zweitens argumentiert er dafür, nach Maximierung dieser Freiheitsrechte nur den Nutzen des Schlechtestgestellten zu betrachten.¹⁶³ Genau diese scheinbaren Unterschiede gilt es weiter unten bei der utilitaristischen Modellierung der Rawls'schen Position zu berücksichtigen. Zunächst ist es jedoch notwendig, die Ar-

¹⁶² Rawls (1999: 59).

¹⁶³ Vgl. hierzu Roemer (1996: 128).

gumentation von Rawls – für unsere Zwecke jedoch nur skizzenhaft und ohne ins Detail zu gehen – nachzuzeichnen.¹⁶⁴

Rawls' Startpunkt ist ein gedachter vertragstheoretischer Urzustand, der anders als der Hobbes'sche oder Locke'sche Naturzustand keinen historischen Ort hat. In diesem abstrakten und hypothetischen Urzustand – Rawls nennt ihn die „original position“ – sollen nun die Individuen über die allgemeinsten Regeln der Gesellschaft übereinkommen. Die original position ist durch einen „Schleier des Nichtwissens“ gekennzeichnet, der bewirkt, dass die Individuen kein Wissen von ihren konkreten Zielen oder konkreten individuellen Lebensumständen haben: „First of all, no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength and the like. Nor, again, does anyone know his conception of the good, the particulars of his rational plan of life, or even the special features of his psychology, such as *aversion to risk* or liability to optimism and pessimism. [...] they do not know its economic or political situation, or the level of civilization and culture it has been able to achieve.“¹⁶⁵ Diese original position bezeichnet Rawls als fair, da alle Akteure hinter diesem Schleier von absolut gleichen Voraussetzungen ausgehen. Die Individuen der original position verschmelzen gleichsam zu einem einzigen moralischen Subjekt, das über die allgemeinsten Regeln der Gesellschaft entscheidet, die nach Rawls mindestens den folgenden zwei Prinzipien der Gerechtigkeit genügen müssen:

„*First Principle*: Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.

¹⁶⁴ Der näher interessierte Leser sei auf die Einführung von Kersting (2008) verwiesen.

¹⁶⁵ Rawls (1999: 118), eigene Hervorhebung.

Second Principle: Social and economic inequalities are to be arranged so that they are [...] to the greatest benefit of the least advantaged [...].¹⁶⁶

Während das erste Prinzip hier also eine absolute Gleichverteilung maximaler und miteinander zu vereinbarender Freiheits- und Grundrechte fordert, besagt das zweite, auch Differenzprinzip genannte, dass eine sonstige Ungleichverteilung nur dann gerechtfertigt ist, wenn sie dem Schlechtestgestellten zum Vorteil gereicht.¹⁶⁷

Zur Ableitung des Differenzprinzips aus der original position nimmt Rawls an, dass die Individuen die Maximin-Entscheidungsregel¹⁶⁸ anwenden. Er begründet dies damit, dass den Individuen die üblicherweise angenommenen Informationen fehlen, die eine Strategie der Maximierung ihres Erwartungswertes sinnvoll erscheinen lassen.¹⁶⁹ Die entscheidende Rolle spielt hierbei die Rawls'sche Annahme, dass die Individuen hinter dem Schleier des Nichtwissens über *keine Einstellung zu Risiko* verfügen. Es fehlt ihnen also insbesondere an der Möglichkeit,

¹⁶⁶ Rawls (1999: 266). Ich unterschlage hier für die Gradlinigkeit des Arguments Rawls' zweiten Teil, der wie folgt lautet: „b) attached to offices and positions open to all under conditions of fair equality of opportunity“. Dies ist deshalb statthaft, weil diese Einschränkung zulässiger Ungleichverteilungen im Allgemeinen keine bindende Einschränkung ist und sich zudem bereits aus dem ersten Prinzip ergeben kann.

¹⁶⁷ Ich beschränke mich im Folgenden darauf, die Rawls'sche Argumentation für das zweite Prinzip nachzuzeichnen, denn dieses wurde im Gegensatz zum eher weniger umstrittenen ersten Prinzip der maximal miteinander zu vereinbarenden Freiheits- und Grundrechte stärker zum Gegenstand der philosophischen Debatte.

¹⁶⁸ Die Maximin-Entscheidungsregel wählt diejenige Entscheidung aus, die das beste minimale Ergebnis erreicht. Die Maximin-Strategie unterscheidet sich grundlegend von der Strategie der Maximierung des Erwartungswertes und wird regelmäßig dann die gewählte Strategie sein, wenn nichts über Wahrscheinlichkeitsverteilungen bekannt ist oder die Bewertung von Auszahlungen unterhalb eines bestimmten Minimums sehr niedrig ist. Jeder Akteur kann bei Befolgung der Maximin-Strategie sicher gehen, dass er mindestens ein solches Minimum erhält.

¹⁶⁹ Vgl. Rawls (1999: 148ff.).

Aussagen darüber zu treffen, auf wie viel materielle Güter MG sie im schlechtesten Fall verzichten würden, um sich eine wie auch immer wahrscheinliche Chance auf eine Mehrausstattung mit MG zu eröffnen.¹⁷⁰ Die rationale Entscheidung der Individuen hinter dem Schleier des Nichtwissens, so Rawls, ist deswegen, in Befolgung der Maximin-Entscheidungsregel nur solche Ungleichverteilungen zuzulassen, die dem jeweils Schlechtestgestellten zum Vorteil gereichen.

6.2 Der Rawls'sche Utilitarismus

Um eine Modellierung der Rawls'schen Theorie vorzustellen, nehme ich im Folgenden gleiche Nutzenfunktionen bei allen Individuen an, sowie dass ihr Nutzen von zwei Güterarten abhängt. Aus Rawls' Prinzipien ist ersichtlich, dass wir hier prinzipiell zwei Argumente der von uns zu betrachtenden Nutzenfunktionen zu unterscheiden haben: zum einen Freiheits- und Grundrechte, im Folgenden GG genannt, und zum anderen sonstige materielle Güter, im Folgenden MG genannt. Eine Nutzenfunktion, die obigen Festlegungen entspricht, ist:

$$(27) \quad U_i(GG_i, MG_i) = \begin{cases} -\infty & \text{falls } GG_i < GG^* \\ f(MG_i) & \text{andernfalls} \end{cases},$$

wobei MG_i und GG_i die dem Individuum i jeweils zukommenden Mengen an GG und MG seien und GG^* die von Rawls intendierte maximale Menge miteinander vereinbarter gleicher Freiheits- und Grundrechte

¹⁷⁰ Zum Zusammenhang von Risikoaversion, der Maximin-Strategie und Grenznutzen siehe Hammond (1975: 465-467).

sei. Die utilitaristische Aggregationsfunktion einer Verteilung der Güter GG und MG auf n Individuen ist dann die Summe der U_i

$$(28) \quad \sum_i U_i(GG_i, MG_i) = \begin{cases} -\infty & \text{falls } \exists i(GG_i < GG^*) \\ \sum_i f(MG_i) & \text{andernfalls} \end{cases} .$$

Die Fallunterscheidung in Formel (27) und (28) drückt die von Rawls geforderte lexikographische Priorität von Freiheits- und Grundrechten über sonstige Güter aus. Falls diese Freiheits- und Grundrechte nicht für alle den maximal miteinander zu vereinbarenden Umfang erreichen, wird hier ein prohibitiv negativer Nutzeneffekt angenommen, so dass ein Utilitarist nicht für eine solche Verteilung der GG_i optieren könnte.

Somit lässt sich eine von Rawls' gegen den Utilitarismus vorgebrachte und weithin akzeptierte Kritik entkräften, dass ihm zufolge Einzelne, also auch die Schlechtestgestellten, zur Erhöhung des Gesamtwohls der Gesellschaft ihrer Rechte beraubt werden dürften. Rawls ist bekanntlich anderer Meinung: „Each member of society is thought to have an inviolability founded on justice or, as some say, on natural right, which even the welfare of every one else cannot override. Justice denies that the loss of freedom for some is made right by a greater good shared by others.“¹⁷¹ Diese Rawls'sche Sicht findet sich jedoch auch in (27) wieder, denn die Rate der Substitution von GG_i gegen MG_i ist bei $GG_i = GG^*$ unendlich hoch ist. Im Aggregat hat somit ein Unterschreiten von GG^* auch nur für ein Individuum so negative Auswirkungen, dass in der Maximierungslösung für alle i gelten muss: $GG_i = GG^*$, d.h., GG sind somit nicht gegen MG austauschbar. Diese Rawls'sche Kritik am Utilitarismus läuft deswegen ins Leere, da sie kein *Wesensmerkmal* des Utilitarismus

¹⁷¹ Rawls (1999: 24f.).

kritisiert, sondern lediglich diejenigen Formen des Utilitarismus, in denen Freiheits- und Grundrechte als verhandelbare oder aufwiegbare Nutzenkomponenten verstanden werden. Der Utilitarist jedoch ist, wie beispielsweise im Fall von Nutzenfunktionen des Typs (27), nicht darauf festgelegt, stets eine Aufwiegbarkeit von Freiheits- und Grundrechten vorzusetzen.

Eine andere und von mir an dieser Stelle nicht zu entscheidende Frage ist, ob die genaue Form der lexikographischen Priorität von Freiheits- und Grundrechten, die in der Nutzenfunktion (27) festgelegt ist, gänzlich überzeugt. Der Rawls'sche Ansatz, Freiheits- und Grundrechte in Höhe von GG^* als prinzipiell nicht aufwiegbare anzusehen, ließe sich kritisieren, indem eine andere Gewichtung GG zu MG oder eine andere Höhe von GG^* vorgeschlagen wird.¹⁷² Es zeigt sich daran erneut, wie eine moralische Diskussion in eine Diskussion über korrespondierende Nutzenfunktion überführt werden kann.

Die bisherige Betrachtung hat also gezeigt, dass eine vorgestellte Nutzenfunktion vom Typ (27) dem ersten Rawls'schen Gerechtigkeitsprinzip entspricht, oder genauer, dass eine Gesellschaft von Utilitaristen mit den Nutzenfunktionen vom Typ (27) so verfasst ist, dass sie das erste Rawls'sche Prinzip der Gerechtigkeit erfüllt. Ob sie auch dem zweiten Rawls'schen Prinzip der Gerechtigkeit – dem Differenzprinzip – entspricht, hängt maßgeblich von den Eigenschaften der Funktion $f(MG_i)$ ab. Bis hier haben wir keine Aussagen über die Eigenschaften von f ge-

¹⁷² Beispielsweise mit der Begründung, dass eine Vergrößerung von GG_i , die aber immer noch unter GG^* zurückbleibt, eine schlechte Situation immerhin verbessern sollte. Im Fall von Nutzenfunktionen des Typs (27) hingegen würde eine solche Veränderung keinen Unterschied in der Bewertung ausmachen. Natürlich ließen sich die Nutzenfunktionen des Typs (27) so verfeinern, dass auch eine unter GG^* zurückbleibende Vergrößerung von GG_i einen positiven Nutzeneinfluss bewirkt, ohne dass es ein von Rawls unerwünschtes Austauschverhältnis zwischen GG_i und MG_i geben muss. Daran ist zu sehen, dass es eine ganze Klasse von Nutzenfunktionen gibt, deren utilitaristische Gesellschaft das erste Rawls'sche Prinzip der Gerechtigkeit erfüllt.

treffen. In den Wirtschaftswissenschaften wird davon ausgegangen, dass $f' > 0$ und $f'' < 0$ sind, was besagt, dass der Grenznutzen von MG positiv, aber abnehmend ist. Welche Funktion auch immer f ist, sie legt die Risikoaversion der Individuen hinsichtlich der Verteilung der MG fest. Betrachten wir zum Beispiel

$$(29) \quad f(MG_i) = -MG_i^{-\alpha},$$

wobei α ein Maß für die Risikoaversität ist, so ergibt sich als utilitaristische Summenfunktion

$$(30) \quad \sum_i U_i(GG_i, MG_i) = \begin{cases} -\infty & \text{falls } \exists i(GG_i < GG^*) \\ \sum_i -MG_i^{-\alpha} & \text{andernfalls} \end{cases}.$$

Unterstellen wir nun eine maximale Risikoaversion und lassen α gegen unendlich gehen, so erhalten wir als Spezialfall die utilitaristische Summenfunktion

$$(31) \quad U^R = \sum_i U_i(GG_i, MG_i) = \begin{cases} -\infty & \text{falls } \exists i(GG_i < GG^*) \\ \sum_i g(\min(MG_i)) & \text{andernfalls} \end{cases},$$

deren Wert im Fall von $GG_i = GG^*$ nur noch in einer von g ausgedrückten Weise vom minimalen MG_i abhängt.¹⁷³ Eine utilitaristische Gesellschaft, in der die Nutzenfunktionen der Individuen vom Typ U^R sind, wird somit nicht nur fordern, dass allen die maximalen miteinander zu vereinbarenden Freiheits- und Grundrechte zukommen, sondern darüber hinaus auch nur solche ungleichere Verteilungen zulassen, die dem jeweils Schlechtestgestellten zum Vorteil gereichen.

Der Utilitarist, der sich Nutzenfunktionen vom Typ U^R gegenübergestellt sieht, wird deswegen genau für solche Verteilungen von GG und MG eintreten, die für eine nach Rawls'schen Prinzipien verfasste Gesellschaft konstitutiv sind. Die Rawls'schen Prinzipien der Gerechtigkeit sind somit nicht unvereinbar mit einem Utilitarismus.

6.3 Die utilitaristische Form der Kritik an der Rawls'schen Theorie der Gerechtigkeit

Die Rawls'sche Denkfigur der idealen Abstraktion zur Begründung von Moralprinzipien hat sich als überaus wirkungsmächtig erwiesen. Die Kritik an der Rawls'schen Theorie der Gerechtigkeit hat so auch weniger diese Denkfigur oder sein erstes Prinzip zum Gegenstand, sondern konzentriert sich auf das Differenzprinzip. Dem Rawls'schen Ansatz ist es inhärent, dass die angenommene Ausgestaltung der *Eigenschaften* der Individuen in der original position entscheidend dafür ist, welche allgemeinsten Regeln diese für sich festlegen. Wie wir weiter oben gesehen haben, folgt das Differenzprinzip aus der Rawls'schen Annahme, dass die Individuen der original position *keinerlei Kenntnisse* über ihre Risikoaversionen haben. Die Kritik am Rawls'schen Differenzprinzip war

¹⁷³ Vgl. Arrow (1973: 256f.), Sen (1979: 472) und Hammond (1975). Ähnlich auch die formale Behandlung des Zusammenhangs zwischen Utilitarismus und Maximin bei Roemer (1992: 373-350).

deswegen im Kern immer eine Kritik an der Ausgestaltung der original position und insbesondere an den angenommenen Eigenschaften der idealen Subjekte hinter dem Schleier des Nichtwissens.¹⁷⁴

Verkompliziert wurde diese Debatte dadurch, dass Rawls zwar den Individuen in der original position wie oben beschrieben keine Einstellung zu Risiko zugestehen möchte, an anderer Stelle aber recht konkrete Aussagen über Nutzenniveaus macht und hierbei verkennt, dass diese implizit eine Einstellung zum Risiko enthalten. So begründet er die Motivation der Individuen in der original position, nach der das Differenzprinzip begründenden Maximin-Strategie zu handeln, mit folgendem Gedankengang: „[...] the person that is choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the maximin stipend [...]“¹⁷⁵ Im Kern ist dies die Aussage, dass ab einem gewissen erreichbaren und zu sichernden Minimum von MG der Grenznutzen dieser Güter stark zurückgeht oder sogar null wird, was gleichbedeutend ist mit einer starken Risikoaversion *an dieser Stelle*. Implizit relativiert hiermit Rawls also seine Annahme über die Unkenntnis der Risikoaversionen durch die idealen Entscheider an dieser Stelle mit dem Ziel, die Entscheidung für das Differenzprinzip zu plausibilisieren. Diese Unklarheit bei Rawls bezüglich der Einstellungen der Individuen zu Risiko fußt auf dem Rawls'schen Abgrenzungsbedürfnis gegen den Utilitarismus, dem er eine Begründung des Differenzprinzips nicht zutraut. Auch Arrow sieht diesen Fehler Rawls': „Rawls therefore errs when he argues that average utilitarianism assumes risk neutrality [...] on the contrary, the degree of risk aversion of the individuals is already incorporated in the utility function.“¹⁷⁶ Die Rawls'sche Ablehnung des Utilitarismus kann vor diesem Hintergrund als bloßes Missver-

¹⁷⁴ Siehe beispielsweise Arrow (1973), Bailey (1997: 11-12, 43), Broome (1991: 51-58), Buchanan (1976), Lyons (1972: 544f.), Roemer (1996: 172ff.) und Zamagni (1981: 14ff.).

¹⁷⁵ Rawls (1999: 134).

¹⁷⁶ Arrow (1973: 256).

ständnis gedeutet werden. Auch Utilitaristen müssten, wie wir im letzten Abschnitt gesehen haben, in der original position für das Differenzprinzip optimieren, wenn sie die von Rawls vorausgesetzte maximale Risikoaversion hinsichtlich der Verteilung von materiellen Gütern hätten.

Vielfach wurde allerdings eine solche maximale Risikoaversion als nicht sinnvolle Annahme kritisiert. Konstruktiv wird diese Kritik, wenn sie eine andere Einstellung der Individuen zur Verteilung vom MG als Alternative vorschlägt, was sich als Vorschlag einer alternativen Nutzenfunktion verstehen lässt. Stimmt man beispielsweise mit Rawls in der Annahme einer extremen Risikoaversion hinsichtlich der Verteilung von Freiheits- und Grundrechten überein, sieht jedoch für die Sphäre der MG eine Maximierung des Gesamtnutzens aus MG als sinnvollere Annahme an, so schlägt man abweichend von U^R vor, für den einfachsten Fall $U(MG_i) = MG_i$ die folgende utilitaristische Summenfunktion zu betrachten:

$$(32) \quad \sum_i U_i(GG_i, MG_i) = \begin{cases} -\infty & \text{falls } \exists i(GG_i < GG^*) \\ \sum_i MG_i & \text{andernfalls} \end{cases} .$$

Ein anderes Beispiel der Kritik an der in U^R ausgedrückten maximalen Risikoaversion *entlang aller MG* findet sich bei Roemer, der zwar auch eine festzulegende minimale Ausstattung von MG fordert, allerdings abweichend von Rawls dafür plädiert, oberhalb von dieser Schwelle MG^{\min} einen von null unterschiedenen Grenznutzen vorauszusetzen: „Once this level [a certain minimal level of functioning or opportunity] is reached by all, the commandment to distribute remaining resources

equally is much less compelling.¹⁷⁷ Eine utilitaristische Summenfunktion, die Roemers Vorschlag aufgreift, wäre beispielsweise

$$(33) \quad U^* = \sum_i U_i(GG_i, MG_i) = \begin{cases} -\infty & \text{falls } \exists i(GG_i < GG^*) \\ -\infty & \text{falls } \exists i(MG_i < MG^{\min}) \\ \sum_i f(MG_i) & \text{andernfalls} \end{cases},$$

mit $f' > 0$ und $f'' < 0$. Wie wir in Kapitel 5 gesehen haben, lassen sich weitere egalitaristische Positionen durch geeignete Wahl des $f(MG_i)$ utilitaristisch modellieren, indem abnehmende Grenznutzen oder interdependente Nutzenfunktionen angenommen wurden. U^* umgeht ebenso wie U^R oder Formel (32) die Rawls'sche Kritik an einer Abwägbarkeit von Grundrechten und materiellen Gütern, indem GG eine indexikalische Priorität gegenüber MG eingeräumt wird, bedient sich jedoch auf Seiten der MG nicht des umstrittenen Maximin-Prinzips.

In die Alltagssprache zurückübersetzt würde U^* ungefähr Folgendes bedeuten: „Materielle Güter erhöhen mein Wohlbefinden, aber ohne Freiheits- und Grundrechte und ein Mindestmaß an materiellem Wohlstand geht es mir unendlich schlecht“, oder mit anderen Worten: Ein menschenwürdiges Dasein sowie Freiheits- und Grundrechte sind das höchste Gut.

¹⁷⁷ Roemer (1996: 161).

7 Utilitarismus als Methode

In den vorangegangenen Kapiteln habe ich meine These, dass sich jedes konsistente Moralsystem als Spezialfall des Utilitarismus interpretieren lässt, anhand von Beispielen aufgestellt und erläutert. Die Beispiele, die es utilitaristisch zu interpretieren galt, waren dabei leicht gefunden, denn sie spiegeln die Standardeinwände gegen den Utilitarismus wider. Auch wenn ich bisher einen expliziten Beweis für meine These schuldig geblieben bin, ist bereits die Verschiedenartigkeit der betrachteten und utilitaristisch interpretierten Moralsysteme ein starkes Indiz für eine allgemeine utilitaristische Interpretierbarkeit.¹⁷⁸ Andere Autoren haben in einigen der gewählten Beispiele ganz ähnlich für eine Vereinbarkeit von Utilitarismus und Alltagsmoral argumentiert, um so *den oder wie wir besser sagen sollten: einen* Utilitarismus gegen Kritik zu verteidigen. Hierbei wurden zumeist implizit, selten jedoch explizit alternative zugrunde liegende Nutzenfunktionen vorgeschlagen. Durchweg übersehen wurde jedoch, dass das Fehlen vorgegebener Nutzenfunktionen eine normative Untersättigung des Utilitarismus zur Folge hat, die einen neuen Blick auf den Utilitarismus ermöglicht.

Eine positive Ausnahme hiervon ist Portmore, der ein dem Ansatz dieser Arbeit nicht unähnliches Vorgehen wählt. In seinem Aufsatz „Consequentializing“ argumentiert er überzeugend für eine prinzipielle Übersetzbarkeit von nicht-konsequentialistischen in konsequentialistische Moralsysteme, jedoch ohne diese zu beweisen oder anhand von Beispielen ausführlich zu diskutieren.¹⁷⁹ Allerdings verwirft auch er an gleicher Stelle den Utilitarismus aufgrund seiner scheinbaren Unvereinbarkeit mit unseren alltagsmoralischen Überzeugungen. Das von

¹⁷⁸ Für einen formalen Beweis der utilitaristischen Interpretierbarkeit siehe Kapitel 8.

¹⁷⁹ Portmore (2009: 329ff.).

ihm vorgeschlagene bloße Konsequentialisieren kommt deswegen über allgemeine Aussagen zu Präferenzen nicht hinaus und bleibt abstrakt.

Bevor ich den Zusammenhang der normativen Untersättigung des Utilitarismus und seiner von mir vorgeschlagenen Rolle als Methode der Ethik unter 7.3 beleuchte, möchte ich zunächst erneut den Blick auf die Nutzenfunktionen lenken. Anders als in den vorherigen Kapiteln sind wir hier allerdings nicht auf der Suche nach passenden Nutzenfunktionen zur utilitaristischen Interpretation eines bestimmten Moralsystems. Stattdessen sollen zwei Typen von Präferenzen näher diskutiert werden, von denen in den vorherigen Kapiteln undiskutiert Gebrauch gemacht wurde: sachverhaltsabhängige Präferenzen und interdependente Nutzenfunktionen.

7.1 Zur Annahme sachverhaltsabhängiger Präferenzen

In den Kapiteln 3.1, 3.2 und 4.2 habe ich bestimmte Moralsysteme als Spezialfall des Utilitarismus interpretiert. Hierzu habe ich Typen von Nutzenfunktionen vorgestellt, die über die üblichen Argumente (wie zum Beispiel Güterbündel oder Rechte) hinaus noch weitere Argumente haben. So haben wir beispielsweise unter 3.1 die Sachverhalte „dass das Versprechen V_j gebrochen wurde“ und „dass das Versprechen V_j gehalten wurde“ und unter 3.2 den Sachverhalt „dass das Individuum nicht getötet hat“ als Argumente der Nutzenfunktionen betrachtet. Wie sinnvoll die Annahme solcher sachverhaltsabhängigen Präferenzen innerhalb eines utilitaristischen Kalküls ist, werde ich in diesem Abschnitt diskutieren.

Kritisiert wurden sachverhaltsabhängige Präferenzen in der Literatur hauptsächlich aus Gründen der theoretischen Hygiene oder aus Gründen der gewünschten theoretischen Abgrenzbarkeit. So möchte Nelson

beispielsweise den psychologischen Egoismus¹⁸⁰ ausschließen, um ähnlich wie Gauthier¹⁸¹ einen möglichst großen Teil unserer Moralprinzipien aus individueller Rationalität abzuleiten. Wir müssen ihm zustimmen, wenn er schreibt: „If ‘avoiding immoral actions’ is an argument in everyone’s function lexicographically prior to commodities like ‘beer’ and ‘cars’, then no further linkage between morality and rationality as maximization is required.“¹⁸² Die Relevanz dieses Zitats für unsere Sache müssen wir dabei abschwächen: Nelson schließt solche sachverhaltsabhängigen Präferenzen und somit den psychologischen Egoismus aus, um dafür zu argumentieren, dass Moral in stärkerem Maße einer rationalen Begründung bedarf. Anders als Nelson und Gauthier geht es uns hier jedoch nicht um eine *Begründung* von Moral im Spannungsfeld von Egoismus und Rationalität, sondern um eine *Interpretation* eines Moralsystems als Spezialfall des Utilitarismus. Mit der Sorge um eine allzu leichte und verkürzende Begründung von Moral kann ein Ausschluss von sachverhaltsabhängigen Präferenzen somit nicht begründet werden.

Weitaus einschlägiger für das in dieser Arbeit vorgeschlagene Vorgehen sind diese Ausführungen von Smart:

„What Bentham, Mill and Moore are all agreed on is that the rightness of an action is to be judged solely by consequences, states of affairs brought about by the action. Of course we shall have to be careful here not to construe ‘state of affairs’ so widely that any ethical doctrine becomes utilitarian. *For if we did so we*

¹⁸⁰ Der psychologische Egoismus geht davon aus, dass alles Handeln des Menschen final darauf zielt, seinen Nutzen zu steigern. Beispielsweise wertet der psychologische Egoist eine anscheinend altruistische milde Gabe als egoistische Tat, da sie durch psychische Effekte – beispielsweise durch das Gewissen des Gebenden – in ausschlaggebender Weise den Nutzen des Gebenden befördert.

¹⁸¹ Gauthier (1986).

¹⁸² Nelson (1988: 149).

*would not be saying anything at all in advocating utilitarianism. If, for example, we allowed 'the state of having just kept a promise' then a deontologist who said we should keep promises simply because they are promises would be a utilitarian. And we do not wish to allow this.*¹⁸³

Interessant ist an diesem Zitat zunächst, dass Smart wohl meiner These zustimmen würde, dass sich jedes konsistente Moralsystem als Sonderfall des Utilitarismus interpretieren lässt, falls man wie ich einen weiten Begriff von Nutzenfunktionen zugrunde legt. Er nimmt diese Erkenntnis der Abhängigkeit des normativen Gehalts des Utilitarismus von den angenommenen Nutzenfunktionen jedoch nicht zum Anlass, den Status des Utilitarismus innerhalb der Ethik zu thematisieren. Stattdessen setzt er einen eigenständigen normativen Gehalt des Utilitarismus schlichtweg voraus. Folgerichtig ergibt sich für Smart die Notwendigkeit, sachverhaltsabhängige Präferenzen auszuschließen, denn ebendieser eigenständige normative Gehalt wäre bei Zulassung von sachverhaltsabhängigen Präferenzen in Gefahr.

Wir müssen Smart jedoch nicht folgen in diesem Bestreben, den Utilitarismus gegen andere Moralsysteme abzugrenzen. Teilen wir seine Intuitionen hinsichtlich der normativen Eigenständigkeit des Utilitarismus nicht, so besteht auch nicht die Notwendigkeit, sachverhaltsabhängige Präferenzen vom oben diskutierten Typ auszuschließen. Um sachverhaltsabhängige Präferenzen auszuschließen, bedarf es also schlicht einer besseren Begründung als einer gewünschten Abgrenzbarkeit des Utilitarismus gegen andere Moralsysteme, die hier gerade infrage steht.

Die Beweislast in dieser Frage trägt die Position, die sich prinzipiell dagegen ausspricht, solche Sachverhalte wie „dass Handlung H ausgeführt wurde“ als zu utilitaristisch zu bewertende Folgen der Handlung H zu-

¹⁸³ Smart (1993: 13), eigene Hervorhebungen.

zulassen.¹⁸⁴ Denn *denkbar* ist es allemal, dass solche Sachverhalte für das Wohlergehen von Menschen von Bedeutung sein könnten.¹⁸⁵

7.2 Zur Annahme interdependenter Nutzenfunktionen

In Kapitel 4.1 und 5.3.2 habe ich interdependente Nutzenfunktionen angenommen, um bestimmte Moralsysteme als Spezialfall des Utilitarismus zu interpretieren. Dieser Typ von Nutzenfunktionen soll in diesem Abschnitt diskutiert werden.

Die Annahme interdependenter Nutzenfunktionen und die Zulässigkeit ihrer utilitaristischen Berücksichtigung werden in der Literatur unterschiedlich bewertet. Buchanan beispielsweise argumentiert dafür, dass Wohlwollen, Missgunst, Neid, Altruismus durchaus in den Präferenzen enthalten sein können, was sich als Plädoyer für interdependente Nutzenfunktionen verstehen lassen muss.¹⁸⁶ Ähnlich argumentiert Mirrlees in Bezug auf Ungleichheit: „It is the case that many people are affected by inequality, and have tastes about it. Therefore, inequality in the society affects their utility.“¹⁸⁷

¹⁸⁴ Sachverhalte als utilitaristisch zu bewertende Folgen von Handlungen nehmen auch Broome (1991: 4) und Scheffler (1995: 1f.) an. Siehe dazu auch Williams (1979: 51) und Buchanan (1994: 125), der es für *möglich* hält, dass Individuen eine Präferenz dafür haben, (Moral-)Regeln einzuhalten. Für Ansatzpunkte dieser Position bei Schopenhauer siehe Weise (1995: 261f.).

¹⁸⁵ Ich setze an dieser Stelle weder voraus noch muss ich dafür argumentieren, dass Menschen *tatsächlich* sachverhaltsabhängige Präferenzen haben. Allein schon die logische Möglichkeit sachverhaltsabhängiger Präferenzen spricht *prima facie* dagegen, sie *prinzipiell* auszuschließen.

¹⁸⁶ Koslowski/Buchanan (1996: 66).

¹⁸⁷ Mirrlees (1982: 76). Eine empirische Studie zu interdependenten Nutzenfunktionen findet sich in Zizzo (2003).

Harsanyi sieht hingegen die Notwendigkeit, gewisse Präferenzen auszuschließen: „All clearly anti-social preferences, such as sadism, envy, resentment and malice should be excluded from the social utility function.“¹⁸⁸ Diese Präferenzen sollen also bei der Auswahl der besten Handlung unberücksichtigt bleiben. Der Grund dafür scheint auf der Hand zu liegen: Eine Welt zum Beispiel ohne Sadismus – so möchte man meinen – wäre eine bessere Welt. Unklar ist hierbei jedoch, ob die Menge der eindeutig anti-sozialen Präferenzen klar zu umreißen ist.¹⁸⁹ Bei den von Harsanyi als Beispiel genannten Präferenzen handelt es sich um Instanzen *antagonistischer Präferenzen*, deren utilitaristische Berücksichtigung im Allgemeinen nicht, auch nicht von Harsanyi, infrage gestellt wird. Aber selbst wenn hier eine eindeutige Grenzziehung möglich sein sollte, ist dadurch der Ausschluss solcher scheinbar anti-sozialen Präferenzen noch nicht ausreichend begründet. Immerhin wäre es einem Utilitaristen bei Ausschluss solcher scheinbar anti-sozialen Präferenzen nicht möglich, bei der Auswahl der besten Handlung etwaige Nutzeneffekte beispielsweise aus Neid zu berücksichtigen. Zudem möchte Harsanyi mit dem Ausschluss solcher scheinbar anti-sozialen Präferenzen Gefahren für das allgemeine Wohl begegnen, die der Utilitarist natürlich auch zu berücksichtigen hat. Beispielsweise müssen wir im Allgemeinen keine Angst davor haben, dass der Utilitarist eine Folterung durch Sadisten gutheißen würde, selbst wenn er – anders als Harsanyi es hier fordert – die Präferenzen des Sadisten dafür, andere zu quälen, akzeptiert und berücksichtigt. Denn der Lustgewinn des Sadis-

¹⁸⁸ Harsanyi (1996: 56).

¹⁸⁹ Arrow (1973: 255) beispielsweise spricht sich dagegen aus, Neid und Missgunst prinzipiell auszuschließen, möchte aber trotzdem anti-soziale Präferenzen ausgeschlossen sehen, soweit sie sich unzulässig auf ein bestimmtes Individuum oder auf eine bestimmte Gruppe von Individuen beziehen.

ten wird in den allermeisten Fällen nicht ausreichen, um die negativen Effekte des *Gequältwerdens* aufzuwiegen.¹⁹⁰

Ebenso, wie sich widerspruchsfrei ein Moralsystem vorstellen lässt, das Folterungen erlaubt, ist es vorstellbar, dass ein Utilitarist zu Folterungen verpflichtet ist, *wenn entsprechend radikale Nutzenfunktionen angenommen werden*. Um gegen ein solches Moralsystem zu argumentieren, ist es jedoch nicht notwendig, die ganze Klasse interdependenter oder auch nur anti-sozialer Nutzenfunktionen auszuschließen, sondern es genügt diese *radikalen* Nutzenfunktionen zu kritisieren. Genau dies hieße, einen moralischen Disput in einen Disput über Nutzenfunktionen zu überführen.

Broome schließlich diskutiert interdependente Nutzenfunktionen mit einem dem Ansatz dieser Arbeit ähnelnden Ziel der Versöhnung des Utilitarismus mit dem Egalitarismus.¹⁹¹ Ähnlich versucht Gesang (2003) einen „humanen Utilitarismus“ dadurch zu plausibilisieren, dass er interdependente Präferenzen zulässt, die er „externe Präferenzen“ nennt. Auch diese Autoren versuchen sich somit an einer Verteidigung des Utilitarismus an sich und verkennen die Kraft oder die Implikationen des von ihnen gewählten argumentativen Schachzugs, die es im nächsten Abschnitt zu beleuchten gilt.

¹⁹⁰ Bailey weist ergänzend darauf hin, dass wir uns in einer nicht-idealen Welt, in der der Utilitarist ähnlich wie in unserer Welt keine genauen Informationen über Nutzeneffekte haben kann, der negativen Nutzeneffekte des Gequältwerdens sehr viel sicherer sein können als der Lustgewinne für den Sadisten. Dies spreche gegen die utilitaristische Begründbarkeit von Folterungen in solchen Welten. Vgl. Bailey (1997: 96f., 144f.).

¹⁹¹ Broome (1991: 182).

7.3 Korrespondierende Nutzenfunktionen als utilitaristische Normalform moralischer Regeln

Wenn man, wofür ich in den letzten beiden Abschnitten argumentiert habe, sachverhaltsabhängige und interdependente Nutzenfunktionen nicht ausschließt, so können zu jedem konsistenten Moralsystem *korrespondierende* Nutzenfunktionen angegeben werden, deren Vorliegen einen Utilitaristen auf genau diejenigen Handlungen verpflichtet, auf die auch das Moralsystem verpflichtet. Beispiele für solche korrespondierenden Nutzenfunktionen wurden in den Kapiteln 3 bis 6 gegeben. Ein formaler Beweis dafür, dass sich für jedes konsistente Moralsystem korrespondierende Nutzenfunktionen angeben lassen, findet sich im Anhang. In diesem Abschnitt werde ich thematisieren, welchen philosophischen Wert in der Betrachtung von korrespondierenden Nutzenfunktionen liegt und welchen Status korrespondierende Nutzenfunktionen in der Ethik haben.

Zunächst gilt es, einen naheliegenden Einwand zu entkräften, der die angegebenen korrespondierenden Nutzenfunktionen als *beliebige Ad-hoc-Setzungen* verwirft. Wie wir im Abschnitt 3.1.2 gesehen haben, konnte das Einhalten von Versprechen *auf mehrere Weisen* als utilitaristisches Verhalten interpretiert werden: zum einen indem wir Nutzeffekte versprechenseinhaltender oder versprechensbrechender Handlungen nur beim Versprechensnehmer und zum anderen indem wir diese Effekte nur beim Versprechensnehmer verortet haben. Diese multiple utilitaristische Modellierbarkeit moralischer Regeln sollte jedoch nicht verwundern. Denn die moralische Regel, dass Versprechen einzuhalten sind, erzeugt allein noch keine fein ausdifferenzierte Ordnung in der Menge der möglichen Handlungen. Es stellt sich somit auch nicht die Frage, welche der betrachteten Nutzenfunktionen denn nun *die richtige* korrespondierende Nutzenfunktion zur moralischen Regel ist, dass Versprechen eingehalten werden sollen.

Allgemein gilt: je weniger vollständig das zu interpretierende Moralsystem, desto größer die Menge der möglichen korrespondierenden Nutzenfunktionen. Dies ist jedoch kein Manko der korrespondierenden Nutzenfunktionen, sondern ist darin begründet, dass sie zum einen *mindestens* den normativen Gehalt des zu interpretierenden Moralsystems abbilden und zum anderen selbst eine vollständige Ordnung in der Menge der möglichen Handlungen erzeugen, die das zu interpretierende Moralsystem nicht notwendigerweise erzeugt. Im Umkehrschluss gilt auch: je feiner ausdifferenziert die moralische Ordnung, die ein Moralsystem in der Menge der möglichen Handlungen erzeugt, desto kleiner auch die Menge der möglichen korrespondierenden Nutzenfunktionen.

Der philosophische Wert der Betrachtung korrespondierender Nutzenfunktionen ergibt sich daraus, dass sie utilitaristische Welten charakterisieren, die hinsichtlich des Geforderten dem Ausgangsmoralsystem gleichen.¹⁹² Die Angabe korrespondierender Nutzenfunktionen leistet somit eine utilitaristische Abbildung oder auch Interpretation eines Moralsystems. Der normative Gehalt des Moralsystems wird dabei analysiert, denn er ergibt sich äquivalent aus dem Zusammenspiel des utilitaristischen Maximierens und der korrespondierenden Nutzenfunktionen.

Die Normativität selbst ergibt sich im korrespondierenden Utilitarismus anders als im Ausgangsmoralsystem nicht aus der Geltung der verschiedenen moralischen Regeln, sondern aus der *einen und einfachen* utilitaristischen Pflicht zur Maximierung. Weil sie in utilitaristischen Welten den normativen Gehalt des abzubildenden Moralsystems erzeu-

¹⁹² Mit Portmore (2009: 340) ließe sich an dieser Stelle zwischen normativ extensionaler und intensionaler Gleichheit moralischer Systeme unterscheiden. Auch wenn sich ein moralisches System und sein korrespondierender Utilitarismus hinsichtlich des Geforderten nicht unterscheiden – also normativ extensional gleich sind –, ließe sich doch ein Unterschied ausmachen in der intensionalen Begründung, *warum* etwas gefordert ist. Portmore weist jedoch darauf hin, dass es sich hierbei *um unterschiedliche Auffassungen derselben Begründung* handeln könnte.

gen, sind korrespondierende Nutzenfunktionen nicht willkürlich oder ad hoc gewählt. Die Tatsache, dass für alle konsistenten Moralsysteme korrespondierende Nutzenfunktionen existieren, erlaubt es, diese in Anlehnung an die Terminologie der Logik als utilitaristische Normalformen der Moralsysteme zu bezeichnen.¹⁹³

Die Vorteile einer Betrachtung von korrespondierenden Nutzenfunktionen gegenüber einer Betrachtung des Moralsystems werden weniger deutlich, wenn dieses Moralsystem isoliert diskutiert wird. Hier kann die Betrachtung der korrespondierenden Nutzenfunktionen lediglich das Verständnis für Grenzen und Lücken dieses Moralsystems erleichtern. Anders verhält es sich hingegen, wenn verschiedene konkurrierende Moralsysteme thematisiert werden. Diskutiert man die konkurrierenden moralischen Regeln direkt, so muss diese Diskussion anhand von Beispielen geführt werden. Hierbei wird typischerweise eine fragliche moralische Regel dahingehend kritisiert, dass sie im betrachteten Beispiel weniger überzeugt als die alternativ diskutierte moralische Regel. Hierbei bleibt häufig unklar, *was genau es ist*, das an einer moralischen Regel in einem bestimmten Beispiel zu überzeugen vermag. Patzig nimmt diesbezüglich an, dass das Prinzip einer rationalen Abwägung zwischen solchen verschiedenen Alternativen der Utilitarismus sein müsse.¹⁹⁴ Dann aber wäre der Streit zwischen verschiedenen moralischen Regeln im Kern ein Streit um angenommene Nutzenfunktionen zur Entscheidung darüber, welche moralische Regel im betrachteten Einzelbeispiel zum überzeugenderen Ergebnis kommt.

¹⁹³ Eine interessante Frage, die ich hier nur am Rande erwähnen möchte, ist, ob sich auch andere moralische Theorien als Normalform eignen. Ließen sich beispielsweise alle denkbaren konsistenten Moralsysteme als Varianten des Kant'schen Rigorismus interpretieren? Siehe hierzu Portmore (2007: 59f.), der in einem ähnlichen Kontext eine Art Kant'sche Normalform diskutiert und verwirft. Meine *Vermutung* ist, dass die Kombination aus der Pflicht zur Maximierung und unbestimmtem normativen Gehalt eine Sonderstellung des Utilitarismus begründet.

¹⁹⁴ Patzig (1983: 136).

Anstatt konkurrierende moralische Regeln anhand eines Beispiels zu diskutieren, liegt es somit nahe, die unterschiedlichen korrespondierenden Nutzenfunktionen zu betrachten, so dass der moralische Disput in einen Disput über korrespondierende Nutzenfunktionen überführt wird. Die Vorteile dieses Vorgehens liegen auf der Hand: Erstens bilden die korrespondierenden Nutzenfunktionen die zu diskutierenden moralischen Regeln in *allen Situationen* und nicht nur im betrachteten Beispiel ab, so dass in der Diskussion von den Eigenschaften oder Unzulänglichkeiten des Beispiels abstrahiert werden kann.

Zweitens tritt der Unterschied zwischen den konkurrierenden moralischen Regeln viel deutlicher zutage, wenn die formalen Unterschiede in den korrespondierenden Nutzenfunktionen betrachtet werden. Dies wurde zum Beispiel an den gegeneinander gestellten Nutzenfunktionen in den Abschnitten 5.4 und 6.3 deutlich. Der einheitliche utilitaristische Rahmen erlaubt es dabei, unterschiedliche moralische Regeln systematisch einheitlich zu betrachten.

Drittens haben wir Intuitionen über Nutzeneffekte oder Nutzenfunktionen. Wir gelangen zu diesen Intuitionen durch moralische Erziehung, durch Beobachten anderer Menschen sowie durch Introspektion. Vor diesem Hintergrund können wir vorgeschlagene korrespondierende Nutzenfunktionen bewerten. Besagt eine Intuition beispielsweise, dass Nutzenfunktionen keine extremen Eigenschaften aufweisen, so wären auf Grundlage dieser Intuition die zwei zu deontischen Moralverständnissen korrespondierenden Nutzenfunktionen (5) und (9) zurückzuweisen. Besagt eine andere Intuition beispielsweise, dass Nutzen zumindest *auch* von der individuellen Ausstattung mit MG abhängt und nicht *nur* von der Ausstattung Dritter mit MG, so wären auf Grundlage dieser Intuition die zwei zur extremen egalitären Position und zur Rawls'schen Position korrespondierenden Nutzenfunktionen (25) und (31) zurückzuweisen. Je sicherer wir uns einer Intuition über Nutzeneffekte sind, desto weniger werden wir bereit sein, eine moralische Regel als gültig zu akzeptieren, deren korrespondierende Nutzenfunktionen zu dieser Intu-

ition im Widerspruch stehen. Ein Gewinn der Betrachtung von korrespondierenden Nutzenfunktionen besteht somit darin, dass wir genauer erkennen können, warum Intuitionen über Nutzeneffekte ein Argument gegen moralische Regeln darstellen können.

Dabei setzt eine Betrachtung korrespondierender Nutzenfunktionen *keineswegs* voraus, dass die betrachteten Individuen *tatsächlich* auch diese Nutzenfunktionen haben. Denn dass das Vorliegen korrespondierender Nutzenfunktionen den Utilitaristen auf dieselbe Handlung verpflichtet, auf die eine moralische Regel regelbefolgende Moralsubjekte verpflichtet, sagt ebenso wenig darüber aus, dass diese Pflicht besteht, wie darüber, dass diese Nutzenfunktionen vorliegen. Bei einer Übersetzung oder Interpretation – und darum handelt es sich bei der Angabe korrespondierender Nutzenfunktionen – wird weder das eine noch das andere präsupponiert. Aus demselben Grund müssen wir zur Annahme korrespondierender Nutzenfunktionen auch nicht in revisionistischer Weise voraussetzen, dass die betrachteten Individuen *in Wirklichkeit* Utilitaristen sind.

Natürlich kann eine solche Überführung eines ethischen Streites zwischen einzelnen konkurrierenden moralischen Prinzipien in eine Diskussion über verschiedene korrespondierende Nutzenfunktionen den eigentlichen Streit selbst noch nicht entschieden. Im Streit ist der Utilitarismus keine Partei, weil er *allein* noch keine vollständige normative Theorie ist und sein normativer Gehalt vollständig davon abhängt, welche Nutzenfunktionen betrachtet werden. Eine Entscheidung über einen Disput zwischen verschiedenen Moralregeln erfordert, dass zwischen diesen Moralregeln abgewogen oder vermittelt wird. Diese Abwägungen erfolgen zumeist anhand von Beispielen und häufig implizit oder explizit unter Bezugnahme auf Nutzeneffekte. Ein solcher Gang der Argumentation kann mithilfe von korrespondierenden Nutzenfunktionen formalisiert werden. Wie beispielhaft in den Abschnitten 5.4 und 6.3 deutlich wurde, verpflichten die Auf- und Gegeneinanderstellung von konkurrierenden korrespondierenden Nutzenfunktionen den Mo-

ralphilosophen zu einem erhöhten Maß an Präzision in der Formulierung moralischer Positionen. Dies sowie die Tatsache, dass auch voneinander verschiedenste Moralsysteme auf einheitliche Weise durch verschiedene korrespondierende Nutzenfunktionen charakterisiert werden können, lassen den Utilitarismus als Methode der Ethik und natürlichen Ort der moralischen Diskussion erscheinen.

8 Zusammenfassung

Ziel dieser Arbeit war es, die im Anhang bewiesene These vorzustellen und zu diskutieren, dass jedes konsistente Moralsystem als Spezialfall des Utilitarismus interpretiert werden kann. Zunächst habe ich in Kapitel 2 in der gebotenen Kürze den Utilitarismus vorgestellt, einige einschränkende Annahmen getroffen und auf die normative Untersättigung des Utilitarismus hingewiesen.

Im 3. Kapitel wurden zwei alltagsmoralische Regeln utilitaristisch interpretiert: das Gebot, Versprechen zu halten, und das Verbot zu töten. Hierbei habe ich unter 3.1 zunächst die Möglichkeit und den Wert von Versprechen in utilitaristischen Gesellschaften untersucht. Darauf folgend wurden Typen von Nutzenfunktionen vorgestellt, bei deren Vorliegen der Utilitarist die alltagsmoralische Regel befolgt, dass Versprechen einzuhalten sind. Unter 3.2 haben wir Typen von Nutzenfunktionen betrachtet, bei deren Vorliegen der Utilitarist das Verbot zu töten einhält.

Anschließend wurden in Kapitel 4 drei Standardeinwände gegen den Utilitarismus dargelegt. Es konnte gezeigt werden, dass diese Einwände jeweils nur bestimmte Formen des Utilitarismus treffen, nicht aber die utilitaristische Form der hinter dem Einwand jeweils stehenden moralischen Position.

In Kapitel 5 wurden verschiedene egalitäre Positionen zu Verteilungsfragen präsentiert und mittels Sens Maßzahlen der Ungleichverteilung charakterisiert. Auch hier habe ich Nutzenfunktionen vorgestellt, bei deren Vorliegen der Utilitarist gleichsam eine egalitäre Position vertreten muss. Zunächst betrachteten wir die egalitären Implikationen der standardmäßig anzunehmenden abnehmenden Grenznutzen von materiellen Gütern. Darüber hinausgehende egalitäre Positionen bezüglich der Verteilung von materiellen Gütern ließen sich durch Annahme von interdependenten Nutzen als Spezialfälle des Utilitarismus interpretieren.

Im 6. Kapitel habe ich nach einer kurzen Vorstellung der Rawls'schen Theorie der Gerechtigkeit korrespondierende Nutzenfunktionen angegeben, deren Vorliegen Utilitaristen darauf verpflichtet, den Rawls'schen Prinzipien gemäß zu entscheiden. In einem zweiten Schritt wurden ganz ähnliche Nutzenfunktionen beschrieben, die die utilitaristische Form der Hauptkritik an der Rawls'schen Position abbilden, ohne die weithin akzeptierten Aussagen der Theorie der Gerechtigkeit infrage zu stellen.

In Kapitel 7 folgten die notwendige philosophische Einordnung und Flurbereinigung. Zunächst wurde das Betrachten von sachverhaltsabhängigen und interdependenten Nutzenfunktionen problematisiert und verteidigt. Im Abschnitt 7.3 schließlich wurden der Status der von mir vorgeschlagenen korrespondierenden Nutzenfunktionen in der Ethik thematisiert. Die Überführung eines Moralsystems in seinen korrespondierenden Utilitarismus wurde dabei als nicht revisionistisch charakterisiert. Als Übersetzung oder Interpretation kann uns der korrespondierende Utilitarismus jedoch hilfreich dabei sein, Moralsysteme und ihre Differenzen auf einheitliche und systematische Weise zu betrachten.

Zentral für den in dieser Ausarbeitung vorgestellten Ansatz ist die Erkenntnis, dass der Utilitarismus selbst noch keine vollständige normative Theorie ist, aus der sich per se Handlungsbewertungen ableiten lassen. Der Utilitarismus ist vielmehr *Methode*, mit der Handlungen oder Umweltzustände – aber immer nur relativ zu den angenommenen Nutzenfunktionen – bewertet werden. Die Frage nach dem handlungsrelevanten Gehalt des Utilitarismus *schlechthin* führt somit in die Irre, da sie nicht vor dem Hintergrund angenommener Nutzenfunktionen diskutiert wird. Stattdessen muss von einer normativen Neutralität des Utilitarismus ausgegangen werden, die es nahelegt, ihn als Normalform der Ethik zu bezeichnen. Denn jedes konsistente Moralsystem kann als Spezialfall des Utilitarismus interpretiert werden, der durch korrespondierende Nutzenfunktionen charakterisiert ist. Ebenso lässt sich jeder

moralische Disput in einen Disput über korrespondierende Nutzenfunktionen überführen. Hierdurch werden, neben einem Gewinn an Präzision und Systematik, der Mensch als moralisches Wesen und seine Eigenschaften in den Fokus gestellt. Nutzenfunktionen – gleichsam als Ausdruck des zugrunde liegenden Menschenbildes – zu diskutieren, ist es, was ich als vordringlichste Aufgabe der Moralphilosophie verstehe.

9 Anhang: Beweis der Existenz korrespondierender Nutzenfunktionen

Sei $H = \{h_1, h_2, \dots, h_n\}$ die Menge der möglichen Handlungen.¹⁹⁵

\geq_{mor} sei eine totale Quasi-Ordnung auf H , die durch eine konsistente Menge moralischer Regeln M auf H erzeugt wird. \geq_{mor} ist dabei also eine zweistellige Relation, die folgende Eigenschaften hat:

- a) für beliebige Elemente h_i und h_j aus H gilt $h_i \geq_{\text{mor}} h_j$ und/oder $h_j \geq_{\text{mor}} h_i$ sowie
- b) für beliebige Elemente $h_i \geq_{\text{mor}} h_j$ und $h_j \geq_{\text{mor}} h_k$ aus H gilt $h_i \geq_{\text{mor}} h_k$.

Aus dieser Relation kann man zwei weitere Relationen ableiten:

$h_i \sim h_j$ (h_i und h_j sind moralisch gleichwertig), gdw. $h_i \geq_{\text{mor}} h_j$ und $h_j \geq_{\text{mor}} h_i$

sowie

$h_i > h_j$ (h_i ist moralisch besser als h_j), gdw. $h_i \geq_{\text{mor}} h_j$ und nicht $h_i \geq_{\text{mor}} h_j$.

Wir bezeichnen die Menge der Äquivalenzklassen H/\sim , also die Menge aller Mengen moralisch gleichwertiger Elemente, als A und schließen

¹⁹⁵ Der gesamte erste Teil des Beweises bis zur Fallunterscheidung ist eine teilweise Wiedergabe der Argumentation von Debreu (1954) mit kleinen Adaptionen.

die trivialen Fälle aus, in denen A nur ein einziges Element hat,¹⁹⁶ sowie den Fall der überabzählbar vielen Handlungen.

Eine reellwertige Funktion $\Phi(h)$ nennen wir ordnungserhaltend, wenn für beliebige Elemente h_i und h_j aus H gilt:

$$\Phi(h_i) \geq \Phi(h_j) \text{ gdw } h_i \geq_{\text{mor}} h_j.$$

Wenn wir nun die Elemente von A der moralischen Güte nach anordnen, ist es offensichtlich möglich, durch Induktion über die so entstandene Rangfolge eine ordnungserhaltende Funktion $\theta(h)$ zu konstruieren.¹⁹⁷

1. Spezialfall: gleiche korrespondierende Nutzenfunktionen

Nehmen wir an, dass die Individualnutzen $U_1(\cdot) = U_2(\cdot) = \dots = U_m(\cdot)$ und $\theta(\cdot)$ eine ordnungserhaltende Funktion ist, so dass gilt: $h_p \geq_{\text{mor}} h_q$ gdw. $\theta(h_p) \geq \theta(h_q)$. Wir hätten mit $\theta(\cdot)$ eine utilitaristische Interpretation von \geq_{mor} gefunden, wenn wir zeigen können, dass $\theta(\cdot) = U_1(\cdot) + \dots + U_m(\cdot)$. Betrachten wir hierzu die Individualnutzen $U_1(\cdot) = \theta(h)/m$, so ergibt sich als zu maximierende utilitaristische Summenfunktion $U(\cdot) = U_1(\cdot) + \dots + U_m(\cdot) = m \cdot U_1(\cdot) = \theta(\cdot)$, was zu zeigen war. Diese zu maximierende Summenfunktion $\theta(\cdot)$ ist somit eine der zahlreichen utilitaristischen Interpretationen der konsistenten Menge moralischer Regeln, die die Ordnung \geq_{mor} erzeugt.

¹⁹⁶ Zum Beispiel weil nur eine mögliche Handlung betrachtet wird oder alle möglichen Handlungen moralisch gleichwertig sind.

¹⁹⁷ Vgl. das analoge Vorgehen von Debreu (1954: 161). Hierzu eine Anmerkung: Je weniger umfangreich die zu ordnende Menge der möglichen Handlungen ist und je weniger ausdifferenziert die konsistente Menge moralischer Regeln ist, desto mehr verschiedene ordnungserhaltende Funktionen lassen sich aufstellen.

2. Allgemeinfall: verschiedenartige korrespondierende Nutzenfunktionen

Nehmen wir an, dass $\theta(\cdot)$ eine ordnungserhaltende Funktion ist. Es gilt somit: $h_p \geq_{\text{mor}} h_q$ gdw. $\theta(h_p) \geq \theta(h_q)$. Betrachten wir nun beliebige Individualnutzenfunktionen $U_1(\cdot), U_2(\cdot), \dots, U_{i-1}(\cdot), U_{i+1}(\cdot), \dots, U_m(\cdot)$ und bezeichnen deren Summe mit $\gamma(\cdot)$. Es lässt sich nun stets ein $U_i(\cdot)$ finden, so dass $\gamma(\cdot) + U_i(\cdot) = \theta(\cdot) = U_1(\cdot) + \dots + U_m(\cdot)$. Somit ist $\theta(\cdot)$ eine utilitaristische Interpretation von \geq_{mor} . Diese zu maximierende Summenfunktion $\theta(\cdot)$ ist somit eine der zahlreichen utilitaristischen Interpretationen der konsistenten Menge moralischer Regeln, die die Ordnung \geq_{mor} erzeugt.

Literaturverzeichnis

- Anwander, N. (2008): „Versprechen und Verpflichten“, Paderborn: Mentis.
- Arrow, K. (1973): „Some Ordinalist-Utilitarian Notes on Rawls’s Theory of Justice by John Rawls“, in: *The Journal of Philosophy*, Vol. 70, No. 9, S. 245-263.
- Bailey, J. W. (1997): „Utilitarianism, institutions, and justice“, New York: Oxford University Press.
- Baurmann, M. / Kliemmt, H. (1995): „Zur Ökonomie der Tugend“, in: Weise, P. u.a. (Hrsg.): „Markt, Norm und Moral“, Frankfurt am Main u.a.: Campus Verlag, S. 13-44.
- Birnbacher, D. (1995): „Tun und Unterlassen“, Stuttgart: Reclam.
- Blackorby, C. / Donaldson, D. (1977): „Utility vs. Equity: Some Plausible Quasi-orderings“, in: *Journal of Public Economics* 7, S. 365-381.
- Broome, J. (1991): „Weighing Goods: Uncertainty, Equality, and Time“, Cambridge u.a.: Basil Blackwell.
- Buchanan, J. M. (1994): „Choosing What to Choose“, in: *Journal of Institutional and Theoretical Economics (Zeitschrift für die gesamte Staatswissenschaft)* 150(1), S. 123-135.
- Cumminsky, D. (1990): „Kantian Consequentialism“, in: *Ethics* 100, S. 586-630.

- Debreu, G. (1954): „Representation of a Preference Ordering by a Numerical Function“, in: Thrall, R. M. u.a. (Hrsg.): „Decision Processes“, New York: John Wiley, S. 159-165.
- Elster, J. / Roemer, J. E. (1991): „Interpersonal Comparisons of Well-being“, New York: Cambridge University Press.
- Fitzgerald, P. J. (1967): „Acting and Refraining“, in: *Analysis*, Vol. 27, No. 4, S. 133-139.
- Gauthier, D. P. (1986): „Morals by Agreement“, Oxford: Clarendon Press.
- Gesang, B. (2003): „Eine Verteidigung des Utilitarismus“, Ditzingen: Reclam.
- Güth, W. (21999): „Spieltheorie und ökonomische (Bei)Spiele“, Berlin u.a.: Springer.
- Hammerstein, P. / Selten, R. (1994): „Game Theory and Evolutionary Biology“, in: Aumann, R. J. / Hart, S. (Hrsg.): „Handbook of Game Theory“, Volume II, Amsterdam u.a.: Elsevier, S. 931-993.
- Hammond, P. J. (1975): „A Note on Extreme Inequality Aversion“, in: *Journal of Economic Theory*, Volume 11, Issue 3, S. 465-467.
- Hare, R. M. (121963): „Freedom and reason“, Oxford: Clarendon Press.
- Hare, R. M. (91981): „Moral thinking: its levels, methods, and point“, Oxford: Clarendon Press.

- Harsanyi, J. C. (1955): „Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility“, in: *Journal of Political Economy* 63, S. 309-323. Die Seitenzahlen im Text beziehen sich auf den Wiederabdruck in Phelps, E. S. (Hrsg.) (1973): „Economic Justice“, Harmondsworth: Penguin Books, S. 266-285.
- Harsanyi, J. C. (1996): „Morality and the theory of rational behaviour“, in: Sen, A. / Williams, B.: „Utilitarianism and beyond“, Cambridge u.a.: Cambridge University Press, S. 39-62.
- Hooker, B. (2000): „Ideal Code, Real World“, Oxford: Oxford University Press.
- Hume, D. (1739/40): „A Treatise of Human Nature“, hrsg. von Selby-Bigge, I. A. / Nidditch, P. H. (2¹⁹⁷⁸), Oxford: Clarendon Press.
- Johansson, P. (1995): „An introduction to modern welfare economics“, Cambridge u.a.: Cambridge University Press.
- Kagan, S. (1989): „The limits of morality“, Oxford: Clarendon.
- Kersting, W. (2008): „John Rawls zur Einführung“, Hamburg: Junius.
- Knight, J. (1992): „Institutions and Social Conflict“, New York: Cambridge University Press.
- Koslowski, P. / Buchanan, J. M. (1996): „Ethics of Capitalism and Critique of Sociobiology: Two Essays with a Comment by James M. Buchanan“, Berlin u.a.: Springer.

- Lyons, D. (1965): „The forms and limits of Utilitarianism“, London: Oxford University Press.
- Marshall, A. (⁸1920): „Principles of Economics“, London: Macmillan.
- Mill, J. S. (1859): „On Liberty“, hrsg. von Himmelfarb, G. (1974), Harmondsworth: Penguin.
- Mirrlees, J. A. (1982): „The economic uses of utilitarianism“, in: Sen, A. / Williams, B. (1996): „Utilitarianism and beyond“, Cambridge u.a.: Cambridge University Press, S. 63-84.
- Mulgan, T. (2007): „Understanding Utilitarianism“, Stocksfield: Acumen.
- Nelson, A. (1988): „Review: Economic Rationality and Morality“, in: Philosophy and Public Affairs, Vol. 17 (2), S. 149-166.
- Parfit, D. (1986): „Reasons and Persons“, Oxford: Oxford University Press.
- Patzig, G. (²1983): „Ethik ohne Metaphysik“, Göttingen: Vandenhoeck & Ruprecht.
- Pigou, A. C. (⁴1932): „The economics of Welfare“, London: Macmillan.
- Portmore, D. (2007): „Consequentializing Moral Theories“, in: Pacific Philosophical Quarterly 88, S. 39-73.
- Portmore, D. (2009): „Consequentializing“, in: Philosophy Compass 4, S. 329-347.

- Rawls, J. (1999): „A Theory of Justice: Revised Edition“, Oxford: Oxford University Press.
- Robbins, L. (1938): „Interpersonal Comparisons of Utility: A comment“, in: *The Economic Journal* 48 (192), S. 635-641.
- Roemer, J. E. (1996): „Theories of distributive justice“, Cambridge u.a.: Harvard University Press.
- Scanlon, T. (1977): „Rights, Goals and Fairness“, in: *Erkenntnis* 11(1), S. 81-95.
- Scanlon, T. (1990): „Promises and Practices“, in: *Philosophy and Public Affairs* 19 (3), S. 199-226.
- Scanlon, T. (1998): „What we owe to each other“, Cambridge, Mass.: Harvard University Press.
- Scheffler, S. (1995): „The rejection of consequentialism: a philosophical investigation of the considerations underlying rival moral conceptions“, Oxford: Clarendon Press.
- Schelling, T. (1980): „Strategy of conflict“, Cambridge, Mass.: Harvard University Press.
- Schopenhauer, A. (1840/1977): „Über die Grundlage der Moral“, Zürich: Diogenes.
- Sen, A. / Williams, B. (1996): „Utilitarianism and beyond“, Cambridge u.a.: Cambridge University Press.

- Sen, A. (1997): „On economic inequality“, Oxford: Clarendon Press.
- Shaw, W. H. (1999): „Contemporary Ethics: Taking Account of Utilitarianism“, Oxford: Blackwell Publishers.
- Smart, J. J. C. (1993): „An Outline of a system of utilitarian ethics“, in: Smart, J. J. C. / Williams, B. (1993): „Utilitarianism for and against“, Cambridge u.a.: Cambridge University Press.
- Smith, A. (1766): „Lectures of Jurisprudence“, hrsg. von Meek, R. L. u.a. (1978), Oxford: Oxford University Press.
- Vickrey, W. (1945): „Measuring marginal utility by reactions to risk“, in: *Econometrica* 13, S. 319-333.
- Wärneryd, K. (1990): „Economic Conventions: Essays in Institutional Evolution“, Stockholm: Stockholm School of Economics.
- Weise, P. (1995): „Moral zwischen Markt und Norm: Die Moraltheorie Arthur Schopenhauers aus ökonomischer Sicht“, in: Weise, P. u.a. (Hrsg.): „Markt, Norm und Moral“, Frankfurt am Main u.a.: Campus Verlag, S. 240-264.
- Williams, B. (1979): „Kritik des Utilitarismus“, übersetzt von Köhler, W. R., Frankfurt am Main: Klostermann.
- Zizzo, D. J. (2003): „Empirical evidence on interdependent preferences: Nature or nurture?“, in: *Cambridge Journal of Economics* 27, S. 867-880.

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation mit dem Titel „Utilitarismus als Methode der Ethik“ selbständig verfasst und dabei keine anderen als die angegebenen Hilfsmittel benutzt habe. Die vorliegende Arbeit wurde bisher noch nicht anderweitig als Dissertation eingereicht und veröffentlicht

Malte C. Daniels

Juist, 11. 9.2011