

# **Gesundheitsökonomie**

Vorlesung und Übung, Wintersemester 2006/07

Silke Übelmesser\*\*/Martin Werding\*

## **– Gliederung –**

### **I. Einführung**

1. Problemstellung und stilisierte Fakten
2. Individuelle Gesundheit als Gut und als Produkt

### **II. Leistungsanbieter**

3. Ärzte
4. Krankenhäuser

### **III. Rahmenbedingungen und Anreizmechanismen**

5. Honorierung
6. Versicherungen

### **IV. Gesundheitssystem und Gesundheitspolitik**

7. Übergeordnete Organisationen
8. Politische Ökonomie
9. Herausforderungen

\* ifo Institut für Wirtschaftsforschung und CESifo

Kontakt: [werding@ifo.de](mailto:werding@ifo.de)

\*\* CES, Universität München, und CESifo

Kontakt: uebelmesser@lmu.de

## Literatur

- Arrow, K. J. (1985), „The Economics of Agency“, in: J. W. Pratt und R. W. Zeckhauser (Hrsg.), *Principals and Agents. The Structure of Business*, Harvard Business School Press: Boston.
- Barr, N. (1993), *The Economics of the Welfare State*, 2<sup>nd</sup> edition, Oxford University Press: Oxford u. a.
- Becker, G. S. and K. M. Murphy (1988), „A Theory of Rational Addiction“, *Journal of Political Economy* 96, 675–700.
- Breyer, F. (1984a), „Anbieterinduzierte Nachfrage nach ärztlichen Leistungen und die Zieleinkommens-Hypothese“, *Jahrbücher für Nationalökonomie und Statistik* 199, 415–432.
- Breyer, F. (1984b), *Die Nachfrage nach medizinischen Leistungen: Eine empirische Analyse von Daten aus der Gesetzlichen Krankenversicherung*, Springer: Berlin.
- Breyer, F. und V. Ulrich (2000): „Gesundheitsausgaben, Alter und medizinischer Fortschritt: Eine Regressionsanalyse“, *Jahrbücher für Nationalökonomie und Statistik* 220, 1-17.
- Breyer, F. und P. Zweifel (1997), *Gesundheitsökonomie*, 2. Aufl., Springer: Berlin u. a.
- Breyer, F., P. Zweifel und M. Kifmann (2005), *Gesundheitsökonomie*, 5. Aufl., Springer: Berlin u. a.
- Breyer, F. et al. (1988): *Die Krankenhaus-Kostenfunktion. Der Einfluß von Diagnosespektrum und Bettenauslastung auf die Kosten im Krankenhaus*, AOK-Verlag, Bonn.
- Cutler, D. M. (2002), „Health Care and the Public Sector“, *Handbook of Public Economics* 4, 2143-2243.
- Darby, M. R. und E. Karnli (1973), „Free Competition and the Optimal Amount of Fraud“, *Journal of Law and Economics* 16, 67–88.
- Dranove D. und M. Satterthwaite (2000): „The Industrial Organization of Health Care Markets“, in: A. Culyer and J. Newhouse (Hrsg.): *Handbook of Health Economics*, Bd. 1B, Elsevier, Amsterdam, 1093-1139.
- Eckstein, Z. et al. (1985), „Uncertain Lifetimes and the Welfare Enhancing Properties of Annuity Markets and Social Security“, *Journal of Public Economics* 26, 303–320.
- Ehrlich, I. and G. S. Becker (1972), „Market Insurance, Self-Insurance, and Self-Protection“, *Journal of Political Economy* 80, 623–648.
- Fenge, R. (2003): „Generationengerechtigkeit im Gesundheitswesen“, ifo Schnelldienst 14, 8-13.
- Fuchs, V. (1978), „The Supply of Surgeons and the Demand for Operations“, *Journal of Human Resources* 13 (Suppl.), 35–56.
- Grossman, M. (1972), „On the Concept of Health Capital and the Demand for Health“, *Journal of Political Economy* 80, 223–255.
- Bundesministerium für Gesundheit und Soziale Sicherung (2003): *Bericht der Kommission, Kommission für die Nachhaltigkeit in der Finanzierung der sozialen Sicherungssysteme* (Rürup Kommission), Berlin.
- Laffont, J.-J und J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge, Mass.
- Lauterbach, K. W. und M. Lungen (2000), „Auswirkungen von DRGs auf die Krankenhausfinanzierung“, *Das Krankenhaus* 3/2000, 168-175.
- Neubauer, G. (1994), „Probleme der Leistungsdefinition und der Preisfindung für Fallpauschalen und Sonderentgelte“, in: H. R. Vogel, K. Hässner und M. Gerharz (Hrsg.), *Sonderentgelte und Fallpauschalen im Krankenhaus*, Gustav Fischer Verlag: Stuttgart, 29-46.
- Neubauer, G. und R. Nowy (2000), *Analyse der DRG-Fallkostenkalkulation der Vergütungsfindung und der Zu- und Abschläge in Australien*, Gutachten im Auftrag der Deutschen Krankenhausgesellschaft. Institut für Gesundheitsökonomik: München.
- OECD (1993), *OECD Health Systems: Facts and Trends 1960–1991*, Health Policy Studies No. 3, OECD: Paris.
- OECD (1996), *Health Care Reform: The Will to Change*, Health Policy Studies No. 8, OECD: Paris.

- OECD (2001), *OECD Health Data 2001: A Comparative Analysis of 30 Countries*, OECD: Paris.
- OECD (2002), *Measuring Up: Improving Health System Performance in OECD Countries*, OECD: Paris.
- Österreichische Ärztekammer (2000), *Comparison of the Health Care Systems of the EU Member Countries*, ÖÄK: Wien, <http://www.aek.or.at>.
- Rothschild, M. und J. E. Stiglitz (1976), „Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Incomplete Information“, *Quarterly Journal of Economics* 90, 624–649.
- Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen (2001), *Bedarfsgerechtigkeit und Wirtschaftlichkeit*, Gutachten 2000/2001.
- Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2003): *Staatsfinanzen konsolidieren, Steuersystem reformieren – Jahresgutachten 2003/04*, Stuttgart.
- Schneider, M. et al. (1995), *Gesundheitssysteme im internationalen Vergleich*, Basys: Augsburg, <http://www.basys.de>.
- Schulenburg, Graf von der, J.-M., Kielhorn, A., Greiner, W. und T. Volmer (1998), *Praktisches Lexikon der Gesundheitsökonomie*, Asgard-Verlag: Sankt Augustin.
- Steinmann, L. und P. Zweifel (2002): „On the (In)Efficiency of Swiss Hospitals“, *Applied Economics* (erscheint demnächst)
- Van de Ven, W. und R. Ellis (2000): „Risk Adjustment in Competitive Health Plan Markets“, in: A. Culyer and J. Newhouse (Hrsg.): *Handbook of Health Economics*, Bd. 1A, Elsevier, Amsterdam, 755-845.
- Wagner, G. (2003): „Pauschalprämien setzen das Konzept der Bürgerversicherung am besten um“, *ifo Schnelldienst* 17, 3-6.
- Wagstaff, A. (1986), „The Demand for Health: Some New Empirical Evidence“, *Journal of Health Economics* 5, 195–233.
- WHO (1998), *World Health Report 1998*, WHO: New York.

## 1. Einführung: Problemstellung und stilisierte Fakten

Literatur: Breyer, Zweifel und Kifmann (2005, Kap. 1)  
Barr (1993)  
Cutler (2002)

### 1.1 Motivation

Maßnahmen zur Verbesserung des Gesundheitszustandes, zur Vermeidung und zur Behandlung von Krankheiten sowie zur Verlängerung des Lebens haben – unter anderem – eine ganz wichtige ökonomische Dimension. Sie beruhen stets auch auf abwägenden Entscheidungen der betroffenen Individuen, die unter bestimmten (Budget-)Restriktionen gefällt werden, so sehr die Individuen auch an ihrer Gesundheit und ihrem Leben interessiert sein mögen. Beeinflusst und teilweise ersetzt werden diese individuellen Entscheidungen in vielen Ländern ferner durch kollektive Entscheidungen, und die Verwendung von Ressourcen im Gesundheitssektor wird in nennenswertem Ausmaß öffentlich reguliert, finanziert oder sogar organisiert. In jedem Fall ergeben sich zahlreiche Fragestellungen, die einer ökonomischen Analyse zugänglich sind, und in den Bereich klassischer Erkenntnisgegenstände der Volkswirtschaftslehre und der Finanzwissenschaft fallen.

Ein einfaches Beispiel (Barr 1993, 293) mag genügen, um zu illustrieren, dass selbst der Wert des menschlichen Lebens – so unschätzbar er aus der Sicht einer individualistischen Ethik erscheinen mag – einem ökonomischen Kalkül zugänglich ist. Da Sofortmaßnahmen am Unfallort medizinisch gesehen ganz entscheidend sind für Überlebens- und Heilungschancen im Fall schwerer Unfälle, ließe sich beides deutlich steigern, wenn an jeder Straßenkreuzung eine Ambulanz bereit stünde oder Notarztwagen mit medizinisch geschulter Besatzung ständig auf den Straßen patrouillieren würden. Kaum ein Individuum und keine Gesellschaft ist jedoch bereit, die erforderlichen Mittel aufzubringen, um ein solches Netz zur Notfallversorgung flächendeckend vorzuhalten. Offenbar werden hier zumindest implizit die entstehenden Kosten gegen die Erträge einer Minderung von Gesundheits- und Lebensrisiken abgewogen und damit eine ökonomische Entscheidung getroffen, aufgrund derer der Wert dieser Dinge zumindest indirekt auch in finanziellen Größen messbar wird.<sup>1</sup>

---

<sup>1</sup> Dabei kann der rationale Charakter dieser Entscheidung hier eindeutiger unterstellt werden als bei anderen gängigen Beispiele, z. B. beim individuellen Zigarettenkonsum angesichts der damit verbundenen Gefahren (von den zusätzlichen Externalitätenproblemen, die in diesem Bereich auftreten, sei hier abgesehen). Vgl. allerdings Becker und Murphy (1988) für eine ökonomische Analyse „rationalen Suchtverhaltens“.

Eine ökonomische Theorie des Gesundheitswesens befasst sich vor diesem Hintergrund mit einer ganzen Reihe mikro-ökonomischer Aspekte. Im Mittelpunkt der öffentlichen und politischen Diskussion stehen demgegenüber allerdings zumeist Probleme, die ihre Ursachen zwar möglicherweise im Bereich individueller Entscheidungen haben, aber in erster Linie an volkswirtschaftlichen Aggregaten sichtbar werden und stark mit staatlichen Maßnahmen im Bereich der Wirtschafts- und Sozialpolitik verknüpft sind. Kostenexplosion im Gesundheitswesen, Lohnfortzahlungen im Krankheitsfall, Risikostrukturausgleich zwischen gesetzlichen Krankenkassen, Fallpauschalen für Ärzte und Selbstbeteiligungen der Patienten bei der Finanzierung von Behandlungskosten sowie die Regulierung des Pharmamarktes sind nur einige einschlägige Stichworte aus den Diskussionen der vergangenen Jahre. Ein Schlüsselthema all dieser Debatten bildet dabei die Rolle des Staates – im Vergleich zur Rolle von Märkten – im Gesundheitswesen (Cutler 2002). Entscheidend für das Weitere ist, dass eine Antwort auf die Frage nach dieser Rolle auch aus ökonomischer Sicht nicht leicht zu geben ist.

## **1.2 Allokations- und Gerechtigkeitsprobleme im Gesundheitswesen: Gliederung der Veranstaltung**

Die grundlegenden Probleme, die einer rein wettbewerblichen Allokation im Bereich des Gesundheitswesens entgegen stehen, beginnen bereits damit, dass das eigentliche Ziel „Gesundheit“ – zu maximierendes Argument in individuellen Nutzenfunktionen und einschlägigen Ansätzen zur Formulierung sozialer Wohlfahrtsfunktionen – nur schwer definierbar und messbar ist (es gibt Krankheiten ohne Symptome; die Grenze zwischen Hypochondrie und psychischer Krankheit ist kaum zu ziehen; etc.). Hinzu kommt, dass sich die meisten Entscheidungen in diesem Bereich auf „Inputs“ (Behandlungen, Medikamente) beziehen, deren kausale Beziehung zur Gesundheit – neben Selbstheilungskräften, Überlebenswillen, sonstigen Lebensumständen und individuellem Lebensstil – auch für Experten reichlich unklar ist, nicht zuletzt weil sie einer rapiden, insgesamt unvorhersehbaren technologischen Entwicklung unterliegen.

Aus dieser Tatsache ergeben sich zum einen massive Probleme wegen möglicherweise fehlender Konsumentensouveränität (die notwendigen Informationen fehlen; sie sind, soweit vorhanden, äußerst komplex; sie werden häufig nur von Ärzten verfügbar gemacht, die zugleich Anbieter möglicher Leistungen zur Wiederherstellung oder Verbesserung der Gesundheit sind; in vielen Fällen fehlt die Zeit, weitere Informationen einzuholen oder Preise und Leistungen zu vergleichen; vgl. Kapitel 2). Zum anderen erhalten Anbieter von Gesundheitsinputs eine prekäre Schlüsselrolle für die effektive Inanspruchnahme ihrer Leistungen, die eine effiziente Allokation stark von ihrer – intrinsisch nicht-profitorientierten ? – Motivation abhängig

macht (vgl. Kapitel 3 und 4). Es besteht die Möglichkeit, dass sich Märkte für Gesundheitsinputs ganz grundsätzlich als dysfunktional erweisen, wie folgendes Zitat pointiert hervorhebt:

*„That any sane nation, having observed that you could provide for the supply of bread by giving bakers a pecuniary interest in baking for you, should go on to give a surgeon a pecuniary interest in cutting off your leg, is enough to make one despair of political humanity.“* (George Bernard Shaw 1911).

Zu suchen ist daher nach vernünftigen Rahmenbedingungen und Anreizmechanismen namentlich für die Entscheidungen von (potentiellen) Patienten und Ärzten, die allerdings alles andere als leicht zu konstruieren sind.<sup>2</sup> Die Gestaltung der Honorierungssysteme (vgl. Kapitel 5) und von Versicherungsverträgen (vgl. Kapitel 6) nimmt hier eine Schlüsselrolle ein.

Im Zusammenhang mit Versicherungen gegen die finanziellen Folgen (Einkommensausfall und Behandlungskosten) von Beeinträchtigungen des Gesundheitszustands existieren gewichtige „pekuniäre“ Externalitäten, die auf asymmetrischen Informationen basieren und wahrscheinlich besser unter den Einzelbezeichnungen „adverse Selektion“ („*hidden information*“) und „*Moral hazard*“ („*hidden action*“) bekannt sind. Diese Externalitäten haben im Wesentlichen ihr *fundamentum in re*.

Neben diesen Aspekten gibt es eine Reihe weiterer Allokationsprobleme, die auch anderweitig bekannt sind. Im Bereich der Gesundheitsvorsorge und -pflege ergeben sich gewisse reale (oder „technologische“) Externalitäten (Hygiene, Impfschutz). Davon konzeptionell zu unterscheiden sind weitere Typen pekuniärer, nämlich „fiskalischer“ Externalitäten, die sich erst aus einer bestimmten Organisation des Gesundheitswesens, nämlich mit einer starken Rolle des Staates und seiner fiskalischen Aktivitäten ergeben und verbunden sind mit „*Third-party payments*“ über das Staatsbudget oder para-fiskalisch organisierte (d. h. separat geführte) öffentliche Gesundheitsbudgets.

---

<sup>2</sup> Eine Wanderlegende der zeitgenössischen Gesundheitsökonomie besagt, dass einige Indianerstämme – erwähnt werden gelegentlich die Navajos – ihre Medizinmänner („Ärzte“) nach einem völlig anderen Schema vergütet haben: Letztere wurden von den Mitgliedern ihres Stammes entlohnt, solange diese gesund waren; im Falle von Krankheit würden die Leistungen ausgesetzt und erst wieder aufgenommen, wenn der Medizinmann eine „Behandlung“ erfolgreich abgeschlossen hatte. Nicht ganz untypisch für die Überlieferung dieser Legende ist, dass Breyer und Zweifel (1997, 260) diese Form der Vergütung von Ärzten im alten China ansiedeln. Die Frage ist, ob und gegebenenfalls wie dieses kombinierte Versicherungs- und Krankenversorgungsarrangement unter den Gegebenheiten einer anonymen Massengesellschaft nachgebildet werden könnte.

Diese Externalitäten, die teilweise durch die Besonderheiten des Gesundheitsbereichs bedingt sind und teilweise mit der Rolle des Staates zusammenhängen, machen es notwendig, Hinweise auf Markversagen sorgfältig zu analysieren und das Engagement des Staates vor diesem Hintergrund zu evaluieren.

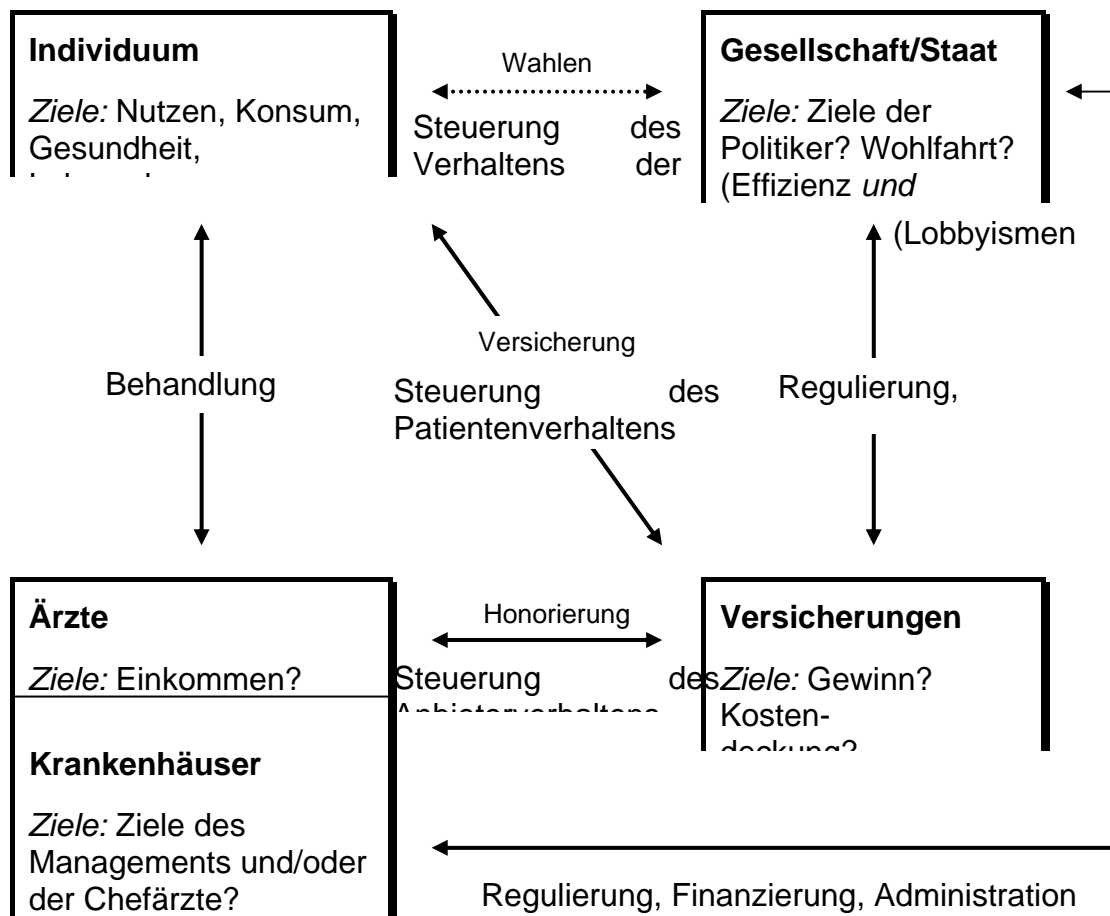
Dafür bietet es sich an, die Form der Organisation des Gesundheitswesens auch im internationalen Vergleich zu betrachten (vgl. Kapitel 7). Wird Handlungsbedarf deutlich, dann müssen die Reformen jedoch unter Berücksichtigung politökonomischer Überlegungen beurteilt werden (vgl. Kapitel 8). In einer Demokratie müssen Mehrheiten überzeugt und das Verhalten von Lobbygruppen berücksichtigt werden.

Im Hinblick vor allem auf aggregierte Entwicklungen ergeben sich bestimmte bereits heute absehbare Herausforderungen (vgl. Kapitel 9) – zu nennen sind dabei insbesondere die für die Dynamik der Gesundheitsausgaben sehr ungünstige demographische Entwicklung der meisten entwickelten Volkswirtschaften und die zukünftigen, in ihren Auswirkungen auf diese Ausgaben nur schwer einzuordnenden technologischen Entwicklungen.

Bei jeder weitgehend privatwirtschaftlichen oder vorrangig staatlichen Organisation des Gesundheitswesens kommen schließlich auch noch aus ökonomischer Sicht schwer zu lösende Gerechtigkeitsprobleme ins Spiel. Für das Gesundheitswesen wird häufig eine Art gutsspezifischer Egalitarismus vertreten, wobei aber zugleich oft nur unzureichend zwischen (teilweise sinnlosen, teilweise unerfüllbaren) Forderungen nach einer Gleichheit der Versorgung mit einschlägigen Leistungen, einer Chancengleichheit im Hinblick auf Gesundheitsversorgung oder die Gesundheit oder einer Gleichheit im Hinblick auf den effektiven Gesundheitszustand unterschieden wird. All dies steht jedoch vor dem Hintergrund einer erkennbaren Ungleichheit von Gesundheitszuständen und -risiken und von Mitteln zur Gesundheitsvorsorge und Krankheitsbekämpfung (Bildungsstand, Haushaltseinkommen, Pro-Kopf-Einkommen im Falle von Haushalten mit zahlreichen ökonomisch abhängigen Mitgliedern), die hier durchaus beachtenswert sein kann.

Die Vorlesung wird alle hier angesprochenen Themen, im Wesentlichen gegliedert nach bestimmten Gruppen von Akteuren im Gesundheitswesen, der Reihe nach behandeln und am Ende die sich im Einzelnen wie im Hinblick auf ein schlüssig konstruiertes Gesamtsystem der Gesundheitssicherung ergebenden wirtschaftspolitischen Implikationen zusammenfassen. Abbildung 1.1 gibt noch einmal einen Überblick über die Zusammenhänge der einzelnen Untersysteme und Akteure im Gesundheitswesen.

Abb. 1.1: Systemzusammenhänge im Gesundheitswesen

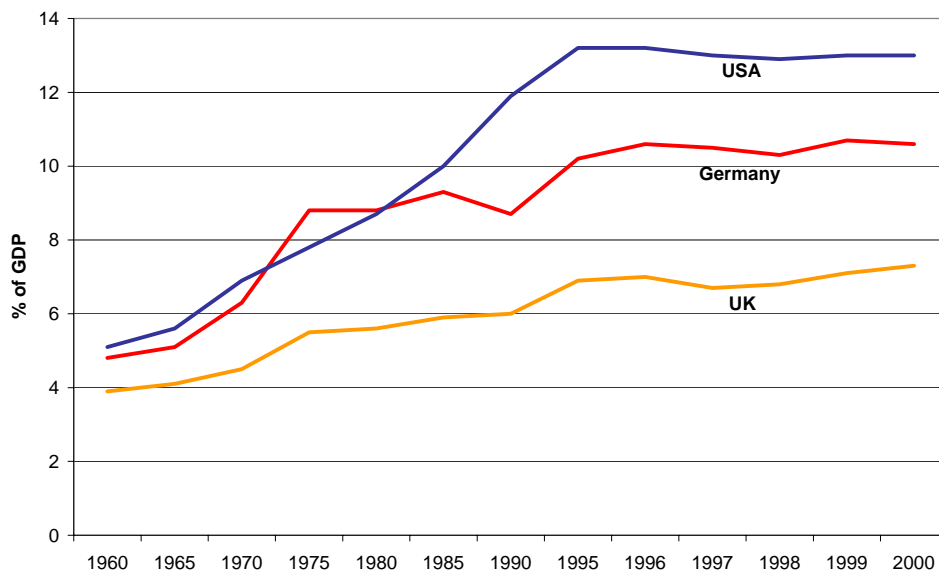


### 1.3 Stilisierte Fakten: Gesundheitsausgaben und die Rolle des Staates

Vergleicht man die (privaten und öffentlichen)<sup>3</sup> Ausgaben in allen Teilsektoren eines typischen Gesundheitswesens auf internationaler Ebene und konzentriert sich dabei zunächst auf drei Länder – die USA, Deutschland und Großbritannien – so fällt auf (vgl. Abb. 1.2): In allen drei Volkswirtschaften hat das Gesundheitswesen eine nennenswerte Größe; es umfasst einen nicht unwichtigen Teil der gesamtwirtschaftlichen Nachfrage, Produktion und Beschäftigung. In allen drei Ländern lässt sich im Zeitablauf ein deutlicher Anstieg der gesamten Gesundheitsausgaben beobachten, allerdings bei ebenfalls beachtlichen und im Ländervergleich recht stabilen Unterschieden hinsichtlich Niveau und Dynamik.

<sup>3</sup> Eine Zusammenfassung ist schon wegen der massiven Systemunterschiede bei internationalen Vergleichen zwingend. Vorausgesetzt werden muss dabei allerdings, dass die – hier von der OECD getroffene – Abgrenzung von „Gesundheitsausgaben“ einigermaßen klar, vollständig und für alle erfassten Länder gleich vorgenommen werden kann.

Abb. 1.2: Gesundheitsausgaben (1960–00) in % des BIP

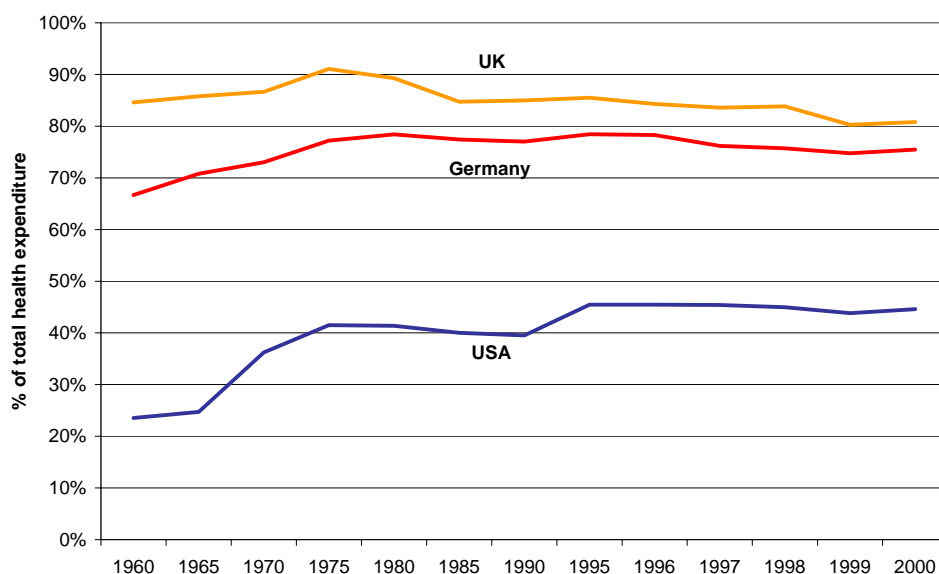


Quellen: OECD (2001)

Angaben für Deutschland bis 1990: früheres Bundesgebiet; ab 1995: Deutschland.

Interessant ist die vergleichsweise eindeutige Reihung, weil sie exakt umgekehrt zur Rolle des Staates in den Gesundheitssystemen dieser Länder ausfällt (vgl. Abb. 1.3): Das höchste Ausgangsniveau und der stärkste längerfristige Anstieg ergeben sich in den USA, wo der Anteil öffentlicher Ausgaben an den Gesamtausgaben im Gesundheitswesen durchgängig am niedrigsten ist; das exakte Gegenteil gilt in beiderlei Hinsicht für Großbritannien, während Deutschland – mit einer stärker „gemischten“ Finanzierung der Gesundheitsausgaben – insgesamt eine mittlere Position einnimmt.

Abb. 1.3: Öffentliche Gesundheitsausgaben (1960–00) in % aller Gesundheitsausgaben



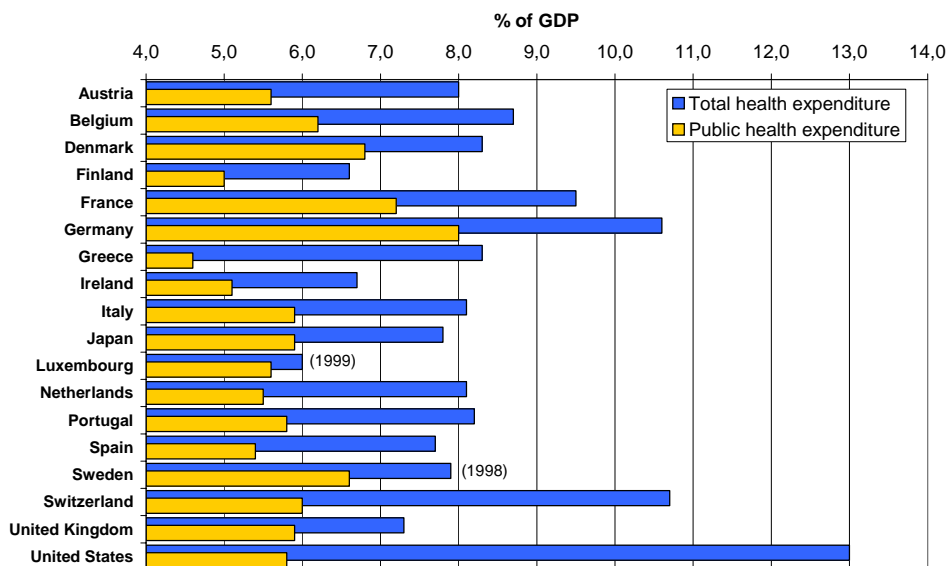
Quellen: OECD (1993; 1996; 2001).

Angaben für Deutschland bis 1990: früheres Bundesgebiet; ab 1995: Deutschland.

Während man staatliche Finanzierung und Administration aus ökonomischer Sicht häufig mit Ressourcenverschwendung und polit-ökonomisch oder bürokratiethoretisch schlüssig erklärbarem Ausgabenwachstum assoziiert, scheint im Gesundheitswesen demnach eher das Gegenteil zu gelten.

Im Hinblick auf eine bei aller Grobheit der Maßstäbe zumindest konzeptionell sinnvolle Bewertung ist dieser Vergleich allerdings unvollständig. Mit den Gesundheitsausgaben wird hier lediglich ein Input gemessen, während für eine Beurteilung der Effizienz der Systeme der jeweilige Output – Gesundheit, Schnelligkeit und Häufigkeit von Heilungsprozessen, Überlebens- und Sterblichkeitsraten, etc. – entscheidend sein müsste. Anders ausgedrückt: die nationalen Gesundheitssysteme können sich auch hinsichtlich ihrer hier nicht beobachteten Qualität stark unterscheiden. Wenn höhere Kosten jedoch mit höherer Qualität einhergehen, ist an den Entwicklungen in Abb. 1.2 unter Umständen nichts weiter besonders. Bezogen auf die drei betrachteten Länder würde eine übereinstimmende Reihung von Ausgabenniveaus und „Qualität“ landläufigen Vorstellungen und fallweiser Evidenz möglicherweise weitgehend entsprechen, und Deutschland nähme seine mittlere Position – beim Kostenniveau und beim Grad staatlicher Einflussnahme – in gewisser zu Recht ein, wenn auf dieser Basis eine akzeptable Qualität von Gesundheitsdienstleistungen („effizient“) produziert wird, die zwischen der der beiden anderen Länder liegt.

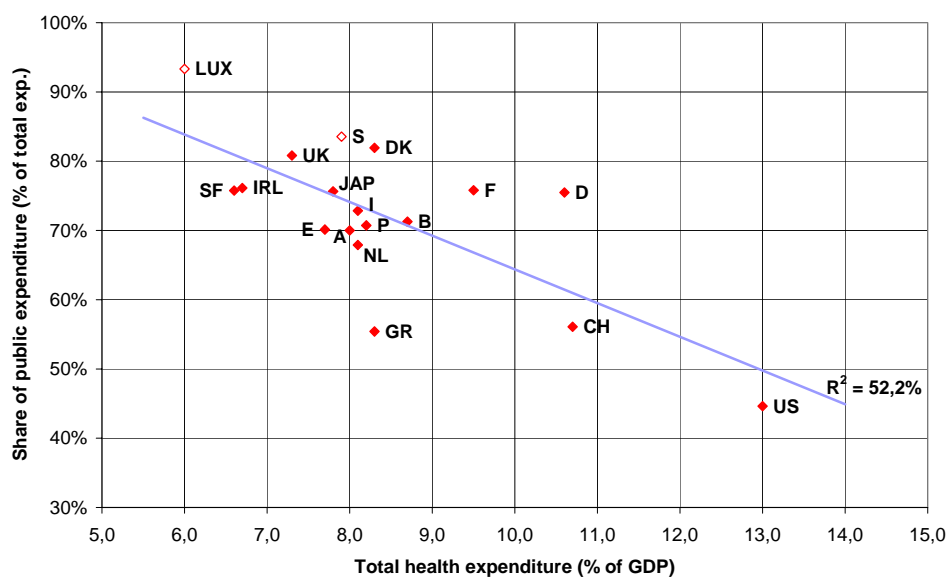
Abb. 1.4: Gesamte und öffentliche Gesundheitsausgaben (2000) in % des BIP



Quelle: OECD (2001)

Ein breiter angelegter Vergleich zwischen allen EU-Staaten und ausgewählten weiteren OECD-Ländern (USA, Japan, Schweiz) lässt allerdings zu Tage treten, dass Deutschland im Hinblick auf das aktuelle Niveau der gesamten Gesundheitsausgaben nicht wirklich eine mittlere Position einnimmt (vgl. Abb. 1.4). Vielmehr belegt es unter den Industrieländern, trotz eines großen Abstandes zum Spitzenreiter USA, sowohl bei den Gesamtausgaben als auch bei den öffentlichen Ausgaben für Gesundheit selbst einen der vordersten Plätze. Dies gilt, obwohl Organisationsform als auch Qualität des Gesundheitswesens bei zahlreichen kontinentaleuropäischen Ländern mit Deutschland vergleichbar sein dürften oder kaum so sehr streuen wie die Ausgabenniveaus. Insgesamt findet man im Mehr-Länder-Vergleich einen negativen Zusammenhang (vgl. Abb. 1.5) zwischen staatlichem Engagement im Gesundheitssektor und den gesamten Gesundheitsausgaben, der allerdings stark vom Outlier USA bestimmt ist.

Abb. 1.5: Zusammenhang zwischen gesamten und öffentlichen Gesundheitsausgaben



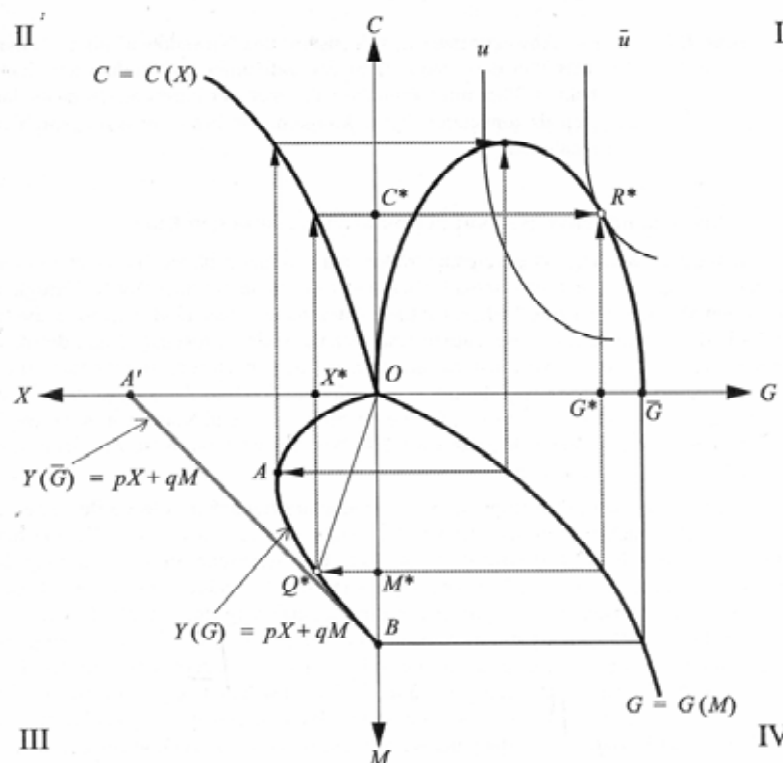
Wichtige Gründe sowohl für das interessante Puzzle, das sich beim Blick auf USA, Deutschland und Großbritannien ergibt, als auch für das weit uneinheitlichere Bild bei einer erweiterten internationalen Perspektive ergeben sich daraus, dass das Gesundheitswesen – nach den Maßstäben der ökonomischen Standardtheorie – durch eine ganze Reihe von Problemen gekennzeichnet ist, die das Funktionieren marktlicher wie nicht-marktlicher Allokationsprozesse in diesem Bereich behindern. Mit einfachen Aussagen über Rahmenbedingungen, Inputs und Outputs des Gesundheitssystems bewegt man sich daher zumeist ohne die gebotene Vorsicht in vermintem Gelände.

Wichtig ist auch, sich (noch einmal) klar zu machen, dass ein internationaler Vergleich des Anteils der gesamten Gesundheitsausgaben am BIP oder der Bedeutung öffentlicher und privater Gesundheitsausgaben auch aus folgenden Gründen problematisch sein kann: einerseits sagen diese Größen nichts über die mit den Inputs erreichte Gesundheitsqualität und andererseits bleiben die möglicherweise von Land zu Land unterschiedlichen Präferenzen der Bürger unberücksichtigt.

### 1.4. Optimale Ausgaben für Gesundheit

Unter diesen Gesichtspunkten ist auch eine staatliche Steuerung der Gesundheitsausgaben bzw. der Gesundheitsquote als Anteil der Gesundheitsausgaben am BIP problematisch. Das kann mit Hilfe eines einfachen Modells graphisch (vgl. Abb. 1.6) veranschaulicht werden, in dem zwei Eigenschaften von Gesundheit berücksichtigt werden: je besser der Gesundheitszustand, desto höher ist das auf dem Arbeitsmarkt erzielbare Einkommen und desto höher ist der konsumtive Nutzen aus diesem Einkommen.

Abb. 1.6: Optimale Ausgaben für Gesundheit



Quelle: Breyer, Zweifel, Kifmann (2005, Abb. 1.1).

Die Individuen wählen die Menge an medizinischen Leistungen  $M$  und Konsumgütern  $X$  so, dass gegeben ihr Budget  $Y(G)$  der Nutzen aus Gesundheit  $G = G(M)$  und

Konsum  $C = C(X)$  maximal ist. Die erste Eigenschaft von Gesundheit lässt sich dabei im 1.Quadranten ablesen und die zweite Eigenschaft im 3.Quadranten.<sup>4</sup> In der Ausgangssituation resultiert aus diesem Maximierungsproblem als Tangentialpunkt der höchsten erreichbaren Indifferenzkurve an die Kurve der Wahlmöglichkeiten das individuelle Nutzenmaximum  $(C^*, G^*)$  im Punkt  $R^*$ , das durch Wahl von  $M^*$  und  $X^*$  erreicht werden kann. Die für das Individuum optimale Aufteilung des Budgets gibt Punkt  $Q^*$  an.

Diese mit einem einzelwirtschaftlichen Optimierungsproblem gefundenen Ergebnisse lassen sich leicht auf die gesamtwirtschaftliche Ebene übertragen. Dafür ist es nur nötig die Bezeichnungen anzupassen. Man spricht dann von gesamtwirtschaftlichem Gesundheitsausgaben bzw. Konsum, betrachtet statt Gesundheit die Zahl der gesund verbrachten Personenjahre und meint mit  $Q^*$  nicht mehr die individuelle Budgetaufteilung, sondern die Aufteilung des BIP auf beide Güter.

Was passiert jetzt, wenn es zu technologischem Fortschritt kommt wie in Abb. 1.7 durch die Verschiebung im 4.Quadranten dargestellt? Die Budgetgerade (3.Quadrant) und die Wahlmöglichkeitenkurve (1.Quadrant) verschieben sich nach außen. Ohne staatlichen Eingriff kann das Folgende beobachtet werden, wenn der besseren Übersichtlichkeit halber angenommen wird, dass die Bürger die Ausweitung der Wahlmöglichkeiten ausschließlich für eine Verbesserung des Gesundheitszustandes verwenden: Die Nachfrage nach medizinischen Leistungen steigt von  $M^*$  auf  $M^{**}$ , während die Nachfrage nach Konsumgütern konstant bleibt,  $X^* = X^{**}$ . Das bewirkt, dass die Gesundheitsquote auf  $Q^{**}$  wächst und sich das Nutzenniveau auf  $R^{**}$  erhöht.

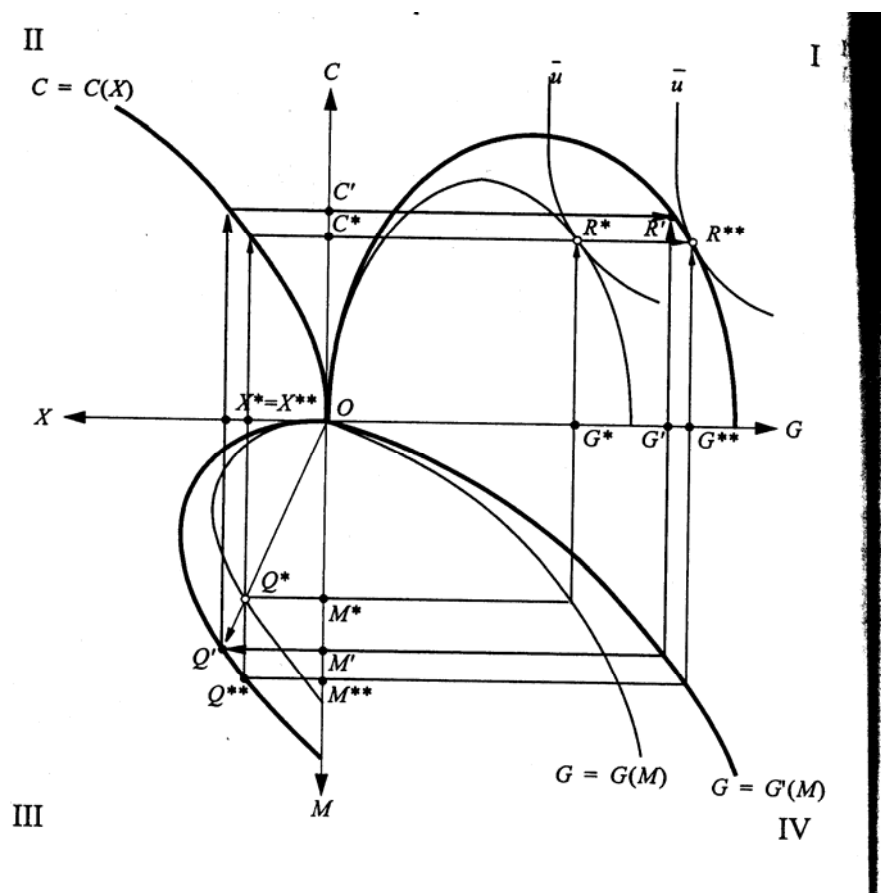
Wird jetzt angenommen, dass der Staat die Gesundheitsquote stabilisieren möchte und deshalb auf  $Q' < Q^{**}$  fixiert, so führt dies dazu, dass die Individuen zu viele konsumtive Leistungen und zu wenige medizinische Leistungen konsumieren und deshalb ein niedrigeres Nutzenniveau erreichen als ohne staatlichen Eingriff,  $R^{**} < R^*$

Dieses einfache Modell zeigt wiederum, dass die Gesundheitsquote kein sehr geeignetes Maß ist, da zum einen die Gesundheitsqualität nicht in die Analyse eingeht und zum anderen die Präferenzen der Bürger keine Berücksichtigung finden. Es gibt somit mehrere Gründe, die gegen die Vorgabe einer bestimmten Gesundheitsquote sprechen, wobei besonders die potentiellen Effizienzverluste anzumerken sind.

---

<sup>4</sup> Für die genaueren Spezifikationen, vgl. Breyer, Zweifel, Kifmann (2003) Kap. 1.

Abb. 1.7: Optimale Gesundheitsausgaben und technologischer Wandel



Quelle: Breyer, Zweifel, Kifmann (2005, Abb. 1.2).

## 2. Individuelle Gesundheit als Gut und als Produkt

Literatur: Breyer und Zweifel (1997, Kap. 2 )  
Breyer, Zweifel und Kifmann (2005, Kap. 2 und 3)  
Grossman (1972)  
Wagstaff (1986)

### 2.1 Motivation

Individuen, die innerhalb des Gesundheitssektors keine spezielle Funktion haben (wie Ärzte oder Versicherungsmanager), treten im Kontext gesundheitsökonomischer Überlegungen in drei verschiedenen Rollen auf: Als Patienten (siehe dazu Kapitel 3, 4 und 5), als Versicherte (siehe Kapitel 6) sowie – ganz fundamental – als Konsumenten und Produzenten ihrer eigenen Gesundheit. Überlegungen zum letzten dieser Aspekte setzen voraus, dass die Individuen eine gewisse Konsumentensouveränität besitzen und gewisse grundlegende Kenntnisse von der Wirkung gesundheitserhaltender und –fördernder Maßnahmen haben. Diese Annahme wird in den folgenden Kapiteln zwar teilweise relativiert. Aus ökonomischer Sicht wäre die Annahme völliger Inkompetenz und Ignoranz der Individuen in dieser Hinsicht allerdings mehr als unbefriedigend.

Ziel dieses Kapitels ist es, mit den in diesem Kapitel behandelten Aspekten zu illustrieren, dass eine Analyse von individuellen Entscheidungen im Zusammenhang mit der eigenen Gesundheit möglich ist und durchaus interessante, auch empirisch relevante Erkenntnisse liefert.

### 2.2 Bewertung von Gesundheitsmaßnahmen

Bevor wir das Verhalten von Patienten thematisieren, müssen wir fragen, ob dieses Verhalten überhaupt der ökonomischen Analyse zugänglich ist. Damit etwas als Marktverhalten, also als mehr oder weniger kompetitives Zusammenspiel von Nachfrage und Angebot (hier: von Gesundheitsleistungen), beschrieben werden kann, muss ein Markt für das entsprechende Gut, d. h. insbesondere ein Marktpreis, existieren. Erst in Bezug auf einen gegebenen, endlichen Marktpreis lassen sich Angebot und Nachfrage definieren.

### 2.2.1 Grundsätzliche Überlegungen

In Bezug auf „Leben“ und „Gesundheit“ als Güter müssen wir vor jeder weiteren ökonomischen Analyse daher zwei Fragen beantworten:

- 1) Kann Leben überhaupt in Geld gemessen werden?
- 2) Wenn ja, ist der Wert endlich?

Zur Beantwortung der Frage 1 hilft es, sich klarzumachen, dass in vielen Fällen, in denen es um die Bewertung des Nutzens von Gesundheitsmaßnahmen geht, nicht identifiziertes (individuelles) Leben auf dem Spiel steht, sondern so genanntes statistisches Leben. Eine beliebige Person aus einer Gruppe wird also mit einer bestimmten Wahrscheinlichkeit ihr Leben verlieren. Außerdem geht es bei der Bewertung oft nicht um ein Leben schlechthin, sondern um die Verlängerung oder Verkürzung der Dauer eines Lebens bzw. um (mögliche) Beeinträchtigungen der Lebensbedingungen.

Wichtig ist bei der Frage, ob man Leben in Geldeinheiten bewerten kann und sollte, der Umstand, dass bei der Befürwortung oder Ablehnung von gesundheitsdienlichen Maßnahmen im öffentlichen Bereich implizit solche Bewertungen sowieso ständig stattfinden. Ob sich die Kosten für eine solche Maßnahme lohnen, kann man nur dann beantworten, wenn man den Nutzen der Maßnahme durch einen mit anderen Optionen vergleichbaren Maßstab, also z. B. in Geldeinheiten, misst. Solche impliziten Bewertungen finden aber verdeckt, unbewusst und unsystematisch statt, so dass erst durch eine offene, nachvollziehbare Bewertung Fehlentscheidungen verhindert werden könnten.

Wird auf eine monetäre Bewertung des Nutzens einer Maßnahme verzichtet, besteht die Gefahr, dass kostenaufwändige Maßnahmen zur Lebensrettung ergriffen und kostengünstige unterlassen werden, da die Nutzenseite der Maßnahme nicht kalkuliert wird. Der Verzicht auf eine solche Bewertung kann also auf Kosten einer längeren durchschnittlichen Lebenserwartung und eines höheren Konsums in der Volkswirtschaft gehen. Kosten bzw. Nutzen, die nicht in Geld bewertet werden können oder sollen, nennt man „intangibel“. Gehen intangible Kosten nicht ins Kalkül der Nettovorteile eines Gesundheitsprojekts ein, ergibt sich folgendes Problem: Sicherheitsgefährdende Projekte werden zu leicht bewilligt und solche, die die Sicherheit erhöhen, zu leicht abgelehnt.

Zur Beantwortung der Frage 2 lassen sich zunächst zwei Definitionen dafür aufstellen, welcher Geldbetrag einem Individuum sein eigenes Leben wert ist.

- A) Geldbetrag , den es zu zahlen bereit ist, um den sicheren Tod zu vermeiden;
- B) Geldbetrag, den man ihm als Entschädigung dafür bietet; seinen eigenen Tod freiwillig in Kauf zu nehmen.

Das Problem der Definition A ist, dass bei einer Befragung regelmäßig herauskommen würde, der Befragte sei bereit, sein gesamtes Vermögen inklusive des verpfändbaren zukünftigen Einkommens zu zahlen. Eine solche Aussage gibt aber nichts über die Präferenzen des Individuums preis, sondern eher über sein Vermögen und seine Verschuldungskapazität. Damit verfehlt diese Definition das, was man erreichen möchte: eine subjektive Einschätzung darüber, wieviel einem das eigene Leben wert ist. Es sollen schließlich die Präferenzen des Individuums und nicht seine Budgetrestriktionen abgefragt werden. Bei Definition B) wird als Antwort hingegen in der Regel kein endlicher Betrag genannt. Beide Definitionen unterscheiden sich gleichsam in der Verteilung der Eigentumsrechte. Bejaht man ein absolutes Recht auf Leben, dann ist der Wert eines identifizierten Lebens unendlich.

Der Unterschied zwischen identifiziertem und statistischem Leben kann auch so gefasst werden: Bei einem statistischen Leben liegt lediglich der Schleier der Ungewissheit über der Tatsache, welches Leben von einer gesundheitsgefährdenden Situation betroffen ist. Typischerweise sind die Risiken so, dass die Anzahl der Opfer vorher nicht mit Sicherheit bekannt sind. Es ist sogar unklar, ob es überhaupt Opfer gibt. Das heißt, mit einer gewissen positiven Wahrscheinlichkeit kommt niemand, kommen einige, kommen alle betroffenen Personen ums Leben. Ist dieser Risikotyp zu analysieren, dann sind statistische Leben das relevante Konzept.

Dann lässt sich aber feststellen, dass die meisten Menschen bereit sein dürften, zur Kompensation kleiner Risiken, die ihr Leben oder ihre Gesundheit betreffen, endliche Geldbeträge entgegenzunehmen. Wenn Individuen ihrem Leben aber einen endlichen Wert zumessen, dann sollte dies auch der Staat bei der Bewertung von gesundheitsfördernden und Sicherheitsmaßnahmen tun.

### **2.2.2 Theoretische Konzepte**

Zwei Konzepte zur Messung des Wertes von Leben und Gesundheit sind der Humankapital- und der Zahlungsbereitschaftsansatz.

#### **a) Humankapitalansatz**

Wird ein Leben über das ihm zugeordnete Humankapital bewertet, so ist der Wert des Lebens gleich der diskontierten Summe aller zukünftigen marginalen Beiträge

zum Sozialprodukt oder, bei Entlohnung nach dem Wertgrenzprodukt, aller zukünftigen Arbeitseinkommen. Hierbei wird unterstellt, dass das Bruttosozialprodukt das geeignete Kriterium für das Wohlergehen der Gesellschaft ist.

Ein Problem dieses Ansatzes ist, dass dieses Konzept nicht auf individuellen Wertungen beruht und damit der in der Ökonomie herrschenden subjektiven Wertlehre widerspricht. Ein weiteres Problem ist, dass der Wert des Lebens von Menschen, die keinen Beitrag zum BSP mehr erbringen (Rentner, Nicht-Arbeitsfähige), Null wäre. So wird zum Beispiel die Freude am Leben nicht als ein Zustand verstanden, der dem Leben einen Nutzen gibt.

## b) Zahlungsbereitschaft

Bei diesem Ansatz wird gefragt:

- A) Wie viel ist einem Individuum eine kleine Änderung der Wahrscheinlichkeit, in einem bestimmten Zeitraum zu sterben, wert? Oder:
- B) Wie viel ist einem Individuum eine kleine Änderung der erwarteten Lebensdauer wert?

Mögliche Antworten auf Frage A sind, den Geldbetrag  $Z$  anzugeben, den das Individuum bereit ist zu zahlen, wenn dafür das Risiko zu sterben um einen kleinen Wert  $\Delta\pi$  gesenkt wird. Dieser Geldbetrag wird **äquivalente Variation** genannt. Eine andere mögliche Antwort wäre der Geldbetrag  $K$ , der erforderlich ist, um das Individuum für eine Erhöhung des Sterberisikos zu kompensieren. Dieser Geldbetrag wird entsprechend als **kompensierende Variation** bezeichnet.

Die Antwort auf Frage B ist ein Geldbetrag, den das Individuum für eine Verlängerung des Lebens um ein Jahr zu opfern bereit wäre. Dieser Geldbetrag kann als Zahlungsbereitschaft bzw. Wert eines gewonnenen Lebensjahres gedeutet werden.

## Modell

Die Zahlungsbereitschaft hänge von der Höhe des Ausgangsniveaus des Sterberisikos  $\pi$  und von der Höhe des Vermögens  $y$  ab. Der Nutzen, der zur Definition herangezogen wird, hängt vom Vermögen ab und davon, ob das Individuum die nächste Periode überlebt oder sofort stirbt: die binäre Variable  $L$  kann die Werte 1 („Leben“) oder 0 („Tod“) annehmen.

Die Nutzenfunktion  $u$  erfüllt folgende Annahmen:

- 1)  $u(L, y)$  sei in  $y$  zweimal stetig differenzierbar, strikt zunehmend und strikt konkav, d. h. es gilt

$$u_y = \frac{\partial u}{\partial y}[L, y] > 0, u_{yy} = \frac{\partial^2 u}{\partial y^2}[L, y] < 0 \quad \text{für alle } (L, y)\text{-Kombinationen.}$$

- 2) Für alle  $y$  gelte:  $u[1, y] > u[0, y]$ .
- 3) Für alle  $y$  gelte:  $u_y[1, y] > u_y[0, y]$

Die Erwartungsnutzenfunktion lautet:

$$EU(\pi, y) \equiv (1 - \pi) \cdot u_1(y) + \pi \cdot u_0(y) \quad \text{wobei } u(1, y) \equiv u_1(y) \text{ und } u(0, y) \equiv u_0(y).$$

Wird das totale Differential gleich Null gesetzt und umgeformt, erhält man die Grenzrate der Substitution zwischen dem Vermögen und der Sterbewahrscheinlichkeit:

$$m(\pi, y) = \left. \frac{dy}{d\pi} \right|_{dEU=0} = \frac{u_1[y] - u_0[y]}{(1 - \pi)u'_1[y] + \pi u'_0[y]} > 0$$

Diese Grenzrate gibt an, wie viel das Vermögen des Individuums zunehmen muss, damit es bereit ist, eine infinitesimal kleine Erhöhung des Sterberisikos, ausgehend vom Niveau  $\pi$ , in Kauf zu nehmen. Das nennen wir die marginale Zahlungsbereitschaft.

Wie ändert sich diese Grenzrate der Substitution  $m$ , wenn sich das Ausgangsniveau des Sterberisikos  $\pi$  erhöht?

$$\frac{\partial m}{\partial \pi} = \left. \frac{d^2 y}{d\pi^2} \right|_{dEU=0} = \frac{-\{u_1[y] - u_0[y]\} \cdot \{u'_0[y] - u'_1[y]\}}{\{(1 - \pi)u'_1[y] + \pi u'_0[y]\}^2} > 0$$

Intuitiv heißt das, je höher die Sterbewahrscheinlichkeit schon ist, umso höher muss die Kompensation für eine zusätzliche Erhöhung der Sterbewahrscheinlichkeit sein. Denn die Wahrscheinlichkeit, dass diese Kompensation nicht das Vermögen im Erlebensfall sondern den Nachlass im Todesfall erhöht, steigt, und daher muss die Kompensation mit dem geringeren Grenznutzen des Geldes bewertet werden.

Wie ändert sich die Grenzrate der Substitution mit dem Ausgangsniveau des Vermögens?

$$\frac{\partial m(\pi, y)}{\partial \pi} = \frac{\partial EU / \partial y \cdot \{u'_1[y] - u'_0[y]\} - \{u_1[y] - u_0[y]\} \cdot \partial^2 EU / \partial y^2}{\{(1 - \pi)u'_1[y] + \pi u'_0[y]\}^2} > 0$$

Die marginale Zahlungsbereitschaft steigt mit dem Vermögen, da mit zunehmendem Vermögen der Grenznutzen des Geldes sinkt und sich somit die Grenzrate der Substitution zwischen Geld und Überlebenswahrscheinlichkeit zugunsten der letzteren verändert.

### Der Wert eines statistischen Lebens

Wie wird ein Gesundheitsprojekt bewertet? Es seien  $n$  Personen betroffen, und das Sterberisiko des  $i$ -ten Individuums werde um  $\delta_i$  gesenkt. Sei nun die absolute Zahlungsbereitschaft des  $i$ -ten Individuums für eine Senkung der Sterbewahrscheinlichkeit  $\pi$  um  $\delta$  gegeben durch  $Z_i$ . Dann lässt sich die Zahlungsbereitschaft für eine Senkung um  $\delta_i$  errechnen als:  $Z_i \cdot \delta_i / \delta$ . Die Summe dieser individuellen Zahlungsbereitschaften ergibt dann den Nutzen oder den Wert eines „Gesundheitsprojektes“:

$$B = \sum_{i=1}^n \delta_i \cdot \frac{Z_i}{\delta} \approx \sum_{i=1}^n \delta_i \cdot m_i$$

Der Beweis für den Zusammenhang zwischen absoluter und marginaler Zahlungsbereitschaft:  $\lim_{\delta \rightarrow 0, Z \rightarrow 0} (Z_i / \delta) = m_i$  findet sich in Breyer/Zweifel (1997, Satz 2.1, S. 37). Um diesen Wert eines Projektes ausrechnen zu können, muss der Planer von jedem einzelnen Betroffenen wissen, um welchen Betrag  $\delta_i$  sein individuelles Sterberisiko gesenkt wird und wie hoch seine Grenzrate der Substitution  $m_i$  ist. Da diese individuellen Daten kaum zu beschaffen sein werden, kann er aber auch folgende Daten abschätzen und verwenden:

- die Anzahl der geretteten statistischen Leben, also die Summe der  $\delta_i$
- die durchschnittliche marginale Zahlungsbereitschaft  $\bar{m} = n^{-1} \sum_{i=1}^n m_i$

Zwischen den individuellen und den hier genannten Daten besteht folgender Zusammenhang:

$$B = \sum_{i=1}^n [\delta_i \cdot m_i] = \bar{m} \cdot \sum_{i=1}^n \delta_i + n \cdot \text{Cov}(\delta_i, m_i)$$

Das heißt, sind die individuellen Größen Risikoreduktion und Zahlungsbereitschaft unkorreliert, so fällt die Kovarianz weg, und der Nutzen des Projekts ergibt sich als Produkt aus den Abschätzungen für die durchschnittlichen marginalen Zahlungsbereitschaft  $\bar{m}$  und die Anzahl der geretteten statistischen Leben  $\sum \delta_i$ .

Um zu Beurteilen, ob das Gesundheitsprojekt durchgeführt werden sollte, müssen der Nutzen  $B$  und die Kosten des Projektes  $C$  verglichen werden. Übersteigt der Nutzen die Kosten, dann bedeutet die Realisierung eine zumindest potentielle Pareto-Verbesserung.

Die Analyse zeigt, dass der Nutzen eines Projektes nicht nur von der Summe der geretteten statistischen Leben abhängt, sondern auch von den individuellen Zahlungsbereitschaften. Diese hängen u.a. – wie wir gesehen haben – positiv von der Sterbewahrscheinlichkeit und dem Vermögen ab. Es ist wichtig, sich zu verdeutlichen, was die Folgen sein würden, wenn man Projekte höher gewichtet, die besonders das Risiko höher exponierter und reicherer Personen reduzieren. In beiden Fällen – besonders aber im zweiten – ist es wichtig, sich die Folgen vor Augen zu halten und ggf. zu überprüfen, ob die Folgen so zu akzeptieren sind.

### 2.3 Gesundheit als Produkt

Gesundheitsproduktion lässt sich als Investition auffassen, die den Kapitalverzehr infolge von Erkrankungen wieder wettmacht und sogar zu einem Nettoanstieg des Kapitalbestandes an „Gesundheit“ führen kann. Die Investition kommt durch den Einsatz medizinischer Leistungen zustande. Der Bestand an Gesundheitskapital ist einer Abschreibung unterworfen, wobei deren Rate nicht nur von der Zeit, sondern auch von anderen Einflüssen wie etwa dem Verzehr bestimmter Konsumgüter abhängt.

**Das Grossman Modell** (Grossman 1972; Wagstaff 1986):

$T$  sei der Lebenshorizont eines Individuums. In jeder Periode  $\tau$  verbringe es eine Zeitspanne  $t^k$  in krankem Zustand. Dabei gelte, dass die Zeitspanne um so kleiner ist, je größer der Bestand an Gesundheitskapital ist. Der Nutzen des Individuums hänge positiv vom Konsum  $X$  und negativ von der krank verbrachten Zeitspanne  $t^k$  ab. Die Zeitpräferenzrate, mit der der Nutzenstrom diskontiert wird, sei  $\rho$ . Das Individuum maximiert dann:

$$W = \int_0^T U\{t^k(\tau); X(\tau)\} \cdot e^{-\rho\tau} d\tau \quad \text{mit}$$

$$\partial U / \partial t^k(\tau) \leq 0 \quad ; \quad \partial U / \partial X(\tau) \geq 0 \quad ; \quad \partial t^k(\tau) / \partial H(\tau) < 0 \quad \forall \tau$$

Die Netto-Investition  $\dot{H}$  in den Kapitalbestand an Gesundheit ergibt sich aus der Brutto-Investition  $I$ , bei der medizinische Leistungen  $M$  gekauft werden und präventive Anstrengungen  $t^l$  getätigt werden. Das Gesundheitskapital schreibt sich mit der Rate  $\delta$  ab, so dass der Bestand  $H$  ständig abnimmt und ohne weitere Investitionen das Individuum immer kränker würde. Die Abschreibungsrate hängt je nach Lebensstil positiv oder negativ vom Konsum ab und nimmt mit dem Lebensalter  $\tau$  zu:

$$\dot{H}(\tau) = I[M(\tau), t^l(\tau) - \delta[\tau, X(\tau)] \cdot H(\tau), \quad \text{mit}$$

$$\frac{\partial I}{\partial M} > 0, \quad \frac{\partial I}{\partial t^l} > 0, \quad \frac{\partial \delta(\tau)}{\partial \tau} > 0, \quad \frac{\partial \delta(\tau)}{\partial X(\tau)} \begin{matrix} > \\ < \end{matrix} 0$$

Der Preis der medizinischen Leistungen  $q$  setzt sich zusammen aus dem Preis  $P$  für  $M$  und dem durch  $t^l$  entgangenen Arbeitseinkommen. Der Preis des Konsumgutes sei  $D$ . Das Arbeitseinkommen  $Y$  ist umso höher, je weniger Krankheitszeit anfällt und je weniger Zeit für präventive Anstrengungen verwendet wird. Das Vermögenseinkommen sind die Zinseinnahmen aus dem Vermögen  $A$ , wobei der Zinssatz  $r$  ist. Der Vermögenszuwachs  $\dot{A}$  ergibt sich dann als:

$$\dot{A}(\tau) = rA(\tau) + Y[t^k(\tau) + t^l(\tau)] - P(\tau) \cdot M(\tau) - P(\tau) \cdot X(\tau) \quad \text{mit}$$

$$\frac{\partial Y(\tau)}{\partial t^k(\tau)} = \frac{\partial Y(\tau)}{\partial t^l(\tau)} < 0$$

Die notwendige Bedingung für ein Maximum dieses dynamischen Optimierungsproblems ist:

$$\left[ \frac{\partial U / \partial t^k(\tau)}{\lambda(0)} \cdot e^{-(\rho-r)\tau} + \frac{\partial Y(\tau)}{\partial t^k(\tau)} \right] \frac{\partial t^k(\tau)}{\partial H(\tau)} = \left[ r + \delta(\tau) - \frac{\dot{q}(\tau)}{q(\tau)} \right] q(\tau)$$

Der Grenzertrag einer Investition muss den Grenzkosten entsprechen.

Zur Interpretation dieser Gleichung:

- Der zweite Multiplikator auf der linken Seite ist negativ, d. h. eine Erhöhung des Kapitalstocks an Gesundheit verringert die krank verbrachte Zeit in einer Periode. Der Klammerausdruck auf der linken Seite muss dann auch negativ sein, damit der Wert der linken Seite insgesamt einen positiven Grenznutzen einer Gesundheitsinvestition ergibt.

- Der erste Term in der linken Klammer sagt zum Einen, dass die Verringerung der krank verbrachten Zeit den Nutzen direkt erhöht, der auf den Entscheidungszeitpunkt abdiskontiert wird. Daneben wird der Gewinn an gesund verbrachter Zeit dazu verwendet, die Vermögensrestriktion zu lockern. Der erste Summand in der eckigen Klammer der linken Seite gibt als Komponente des Grenznutzens einer Gesundheitsinvestition nur den Aspekt der Gesundheit als Konsumgut wieder.
- Der zweite Summand bewertet Gesundheit als Investition. Die Verringerung der krank verbrachten Zeit erhöht das Arbeitseinkommen. Daher hätte eine Investition in die Gesundheit auch dann einen Ertrag, wenn krank verbrachte Zeit als solche nicht direkt Nutzen mindernd wäre. Solche indirekten Effekte auf den Gesamtnutzen  $W$ , die nur über den Effekt auf das Vermögen wirken, werden in Investitionsmodellen betrachtet.
- Auf der rechten Seite stehen schließlich die Kosten einer Gesundheitsinvestition. Diese setzen sich zusammen aus den entgangenen Zinseinnahmen für die Ausgaben für medizinische Leistungen, die sonst auf dem Kapitalmarkt hätten angelegt werden können. Außerdem schreibt sich das erworbene Gesundheitskapital sofort mit der Rate  $\delta$  ab. Schließlich gibt es einen dynamischen Kosteneffekt. Nimmt der Marktpreis in der Zeit zu:  $\dot{q}(\tau) > 0$ , dann reduzieren sich die Kosten der Investition, je früher sie getätigt wird und umgekehrt.

Insgesamt lässt sich also festhalten, dass das Grossman-Modell Gesundheit und Vermögen als zwei miteinander verbundene Aktiva darstellt, deren Werte das Individuum im Zeitablauf optimal steuern kann. Für die Gesundheit gilt, dass der Grenznutzen des Haltens einer Einheit seines Bestandes eine konsumtive und eine investive Komponente enthält, die aus Zins, Abschreibung und einer zeitlichen Wertänderung des Kapitalgutes "Gesundheit" besteht.

Empirische Überprüfungen haben die vom Modell postulierten Zusammenhänge weitgehend bestätigt. Abweichungen einzelner Ergebnisse werden damit erklärt, dass die verwendeten empirischen Methoden nicht optimal gewählt oder weitere Erklärerfaktoren notwendig wären. Außerdem lässt sich kritisieren, dass Unsicherheit im Modell nicht berücksichtigt ist. Erweiterungen zielen deshalb u.a. darauf ab, den Gesundheitszustand als stochastischen Prozess zu modellieren.

### 3. Ärzte

Literatur: Breyer, Zweifel und Kifmann (2005, Kap. 8)  
Breyer (1984a; 1984b)  
Darby und Karni (1973)  
Fuchs (1978)

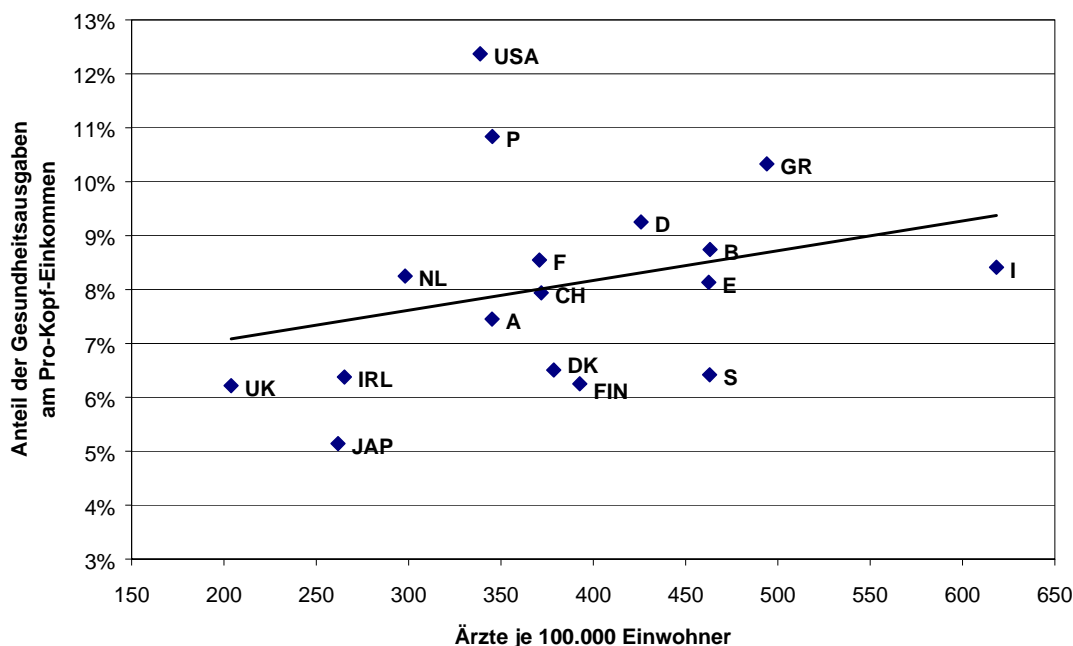
Ärzte, insbesondere niedergelassene Ärzte, nehmen innerhalb der meisten Gesundheitssysteme eine Schlüsselstellung ein: Als erste Anlaufstelle im Falle einer möglichen Erkrankung stellen sie Diagnosen und entscheiden über Therapien, auch wenn sie die Patienten zu weiteren Untersuchungen, ambulanten oder stationären Behandlungen u. U. an spezialisierte Fachärzte oder an Krankenhäuser weiter verweisen. Das Problem dabei ist, dass Ärzte im Rahmen der meisten Honorierungssysteme ein finanzielles Interesse an Art und Umfang der Behandlung haben, das sie v. a. angesichts ihres Informationsvorsprungs ohne wirksame Kontrolle und unabhängig von den Interessen der Patienten und Versicherten an Gesundheit und effizienter Versorgung verfolgen können.

#### 3.1 Anbieterinduzierte Nachfrage

Die These, dass Ärzte in ihrer Doppelrolle als Anbieter medizinischer Leistungen und Berater ihrer Patienten die für sie relevante Nachfrage in ihrem Interesse steuern oder sogar erzeugen – man spricht in diesem Zusammenhang auch von der „doppelten Dynamik der Ärzteneinkommen“ – nährt sich unter anderem aus der Beobachtung einer positiven Korrelation von Ärztedichte und Behandlungsvolumen. Die Tatsache, dass pro Kopf der Bevölkerung um so mehr Behandlungen vorgenommen werden, je mehr Ärzte es gibt, erscheint dabei allerdings zunächst nur als verwunderlich, wenn man „Gesundheit“ als klar definierten Zustand ansieht, der sich jeweils mit einem im Durchschnitt gegebenen Aufwand an Zeit, Dienstleistungen, Medikamenten etc. herstellen lässt (und wenn man davon ausgeht, dass sie im Normalfall mit einem kleineren Aufwand erreicht wird als sich in der Realität beobachten lässt).

Ein grobes Bild, das die These einer anbieterinduzierten Nachfrage motiviert, geben die in Abbildung 3.1 enthaltenen Daten zum Zusammenhang zwischen Ärztedichte und (mit den unterschiedlichen Pro-Kopf-Einkommensniveaus gewichteten) Pro-Kopf-Ausgaben für Gesundheit auf der Basis eines bewusst einfach gehaltenen Querschnittsvergleichs für 17 OECD-Länder.

Abb. 3.1: Ärztedichte und Gesundheitsausgaben in 17 OECD-Ländern (1998)



Quelle: WHO (1998).

Betrachtet man Gesundheit eher als beliebig skalierbares Phänomen, das von Individuen – je nach Preis und Budget – verschieden stark nachgefragt wird, erscheint der rein quantitative Zusammenhang zwischen Ärztedichte und Behandlungsvolumen noch nicht als verwunderlich. Vielmehr kann es sich auch um eine normale Marktreaktion auf eine Ausweitung der Angebotskapazität handeln, infolge derer – bei sinkenden Preisen! – mehr nachgefragt wird. Wenn die Preiselastizität der Nachfrage größer als Eins ist, werden dabei die gesamten Behandlungskosten (aggregiert und pro Kopf der Bevölkerung) steigen. Dieser Sicht der Dinge widerspricht jedoch die Beobachtung, dass die Zunahme des Behandlungsvolumens häufig auch bei konstanten oder sogar zunehmenden, administrierten Kosten je Behandlungseinheit erfolgt.

Ein realistisches Bild des Marktes für ärztliche Dienstleistungen dürfte implizieren, dass es – unabhängig von steuernden Maßnahmen der Ärzte – eine wohldefinierte (preisunelastische) Nachfrage nach Gesundheitsleistungen nur im Sinne einer zumindest konzeptionell bestimmbaren „Primärnachfrage“ gibt, die bei gegebenen Krankheitsbildern und gegebenem medizinischen Kenntnisstand einer ordnungsgemäßen Behandlung entspricht, welche die Ärzte als perfekte „Sachwalter“ der Patienteninteressen vornehmen. Jenseits dessen gibt es jedoch tendenziell Spielräume für zusätzliche, anbieterinduzierte „Nachfrage“, hinsichtlich derer eine Unterscheidung von Angebots- und Nachfragekurve konzeptionell unmöglich wird.

### 3.2 Ein einfaches Modell ärztlichen Verhaltens

Ein Versuch, das – über ein medizinisch erforderliches oder sogar sinnvolles Maß hinaus gehende – nachfrageschaffende Verhalten von Ärzten modellhaft zu fassen, unterstellt den Ärzten auf der Basis einer allgemeinen Nutzenfunktion  $u = u(Y, t, s)$  grundsätzlich ein Interesse an hohem Einkommen ( $Y$ , mit  $u_Y > 0$ ), kurzer Arbeitszeit ( $t$ , mit  $u_t < 0$ ) und angemessener medizinischer Behandlung (ohne nachfragesteuernde Aktivitäten  $s$ ,  $u_s < 0$ ).

Der Ansatz führt allerdings erst dann zu klaren Ergebnissen, wenn man die Nutzenmotive auf der Basis der sogenannten „Zieleinkommens-Hypothese“ (Breyer 1984a) nur für verschiedene Definitionsbereiche von  $Y$  wirksam werden lässt. In ihrer zugespitztesten Form besagt diese Hypothese, dass Ärzte ein bestimmtes Zieleinkommen  $Y^*$  realisieren wollen, wobei ihr Einkommensmotiv nur dann operabel ist, wenn  $Y \leq Y^*$ , während die beiden anderen Nutzendeterminanten nur spürbar werden, wenn  $Y > Y^*$ :

$$u_Y \begin{cases} > 0 & \text{falls } Y \leq Y^* \\ = 0 & \text{falls } Y > Y^* \end{cases}$$

$$u_t, u_s \begin{cases} = 0 & \text{falls } Y \leq Y^* \\ < 0 & \text{falls } Y > Y^* \end{cases}$$

Denkbar sind allerdings auch unter diesen Annahmen noch mehrere Konstellationen, die teilweise ebenfalls einen positiven Zusammenhang von Ärztedichte und Behandlungsaufwand pro Kopf der Bevölkerung nahe legen, ohne dabei die problematischen Konnotationen der These anbieterinduzierter Nachfrage aufzuweisen:

- Rationierte Nachfrage: Wenn die Ärztedichte so gering ist, dass sie – selbst bei maximalem Arbeitseinsatz aller Ärzte – nicht ausreicht, die Primärnachfrage (bei  $s = 0$ ) zu befriedigen, nimmt das Volumen effektiver Behandlungen mit steigender Ärztedichte zu, weil der Nachfrageüberhang abgebaut werden kann. Dieses Phänomen deutet also eher auf anderweitige Probleme der Gesundheitsversorgung als auf problematisches Arztverhalten. Ob die Ärzte in dieser Situation tatsächlich ihre maximale Kapazität mobilisieren, hängt von ihrem Einkommen (im Vergleich zum Zieleinkommen) ab.
- Befriedigung der Primärnachfrage ohne Nachfrageschaffung: Reicht die maximale bzw. die realisierte Kapazität der Ärzte aus, genau die Primärnachfrage der Patienten zu befriedigen und erzielen die Ärzte dabei (mindestens) ihr Zieleinkommen, unterlassen sie nachfrageschaffendes Verhalten ( $s = 0$ ).
- Nachfrageschaffendes Verhalten: Wenn die Ärztedichte so groß wird, dass der einzelne Arzt durch den auf ihn entfallenden Teil der Primärnachfrage nicht mehr

sein Zieleinkommen realisieren kann, wird er bei den Patienten zusätzliche Nachfrage erzeugen ( $s > 0$ ) und seinen Zeiteinsatz dadurch wieder so ausdehnen, dass er exakt sein Zieleinkommen erreicht. Dies ist der gemeinhin für problematisch erachtete Fall anbieterinduzierter, medizinisch aber sinnloser oder sogar kontraproduktiver Nachfrage.

Für die zugespitzte Variante der Zieleinkommenshypothese, bei der überhaupt keine Abwägung zwischen Einkommensziel, Arbeitsleid und medizinischem Ethos stattfindet, ist die zweite dieser Lösungen nur punktuell, d. h. für eine bestimmte Ärztedichte denkbar. Größerer Spielraum für ein Szenario mit ausreichender Versorgung der Patienten ohne künstliche Nachfrageschaffung entsteht, wenn angenommen wird, dass  $u_Y$  durchgängig positiv ist, dass das Zieleinkommen nicht ein fester Wert ist, sondern durch einen gewissen Einkommensbereich beschrieben wird, oder dass die Ärzte auch ein „intrinsisches Motiv“ für die Behandlung von Patienten haben, aufgrund dessen sie ihr Zieleinkommen auch überschreiten, ohne sofort ihre Arbeitszeit herabzusetzen. Unter diesen Bedingungen reagieren sie innerhalb eines gewissen Wertebereichs zunehmender Ärztedichte mit Anpassungen der Arbeitszeit, bis ihr Einkommen unter  $Y^*$  bzw. unter die Untergrenze des Zieleinkommensbereichs fällt. (Weitere denkbare Konstellationen, etwa dass der Arzt bei gegebenen Preisen für seine Leistungen selbst bei maximalem Zeiteinsatz nicht sein Zieleinkommen erreicht, werden hier vernachlässigt.)

### 3.3 Alternative Erklärungen

Für einen positiven Zusammenhang zwischen Ärztedichte und Behandlungsvolumen bzw. Behandlungskosten pro Kopf der Bevölkerung gibt es neben der Erklärung durch manipulatives Verhalten der Ärzte auch alternative Erklärungen, die teilweise bereits angesprochen wurden:

- 1) Normale Marktreaktion: Die erhöhte Ärztedichte bewirkt bei gegebener, preiselastischer Nachfragefunktion eine Mengenausweitung bei gleichzeitiger Preissenkung für ärztliche Leistungen. In diesem Fall steigt das Behandlungsvolumen, aber nicht zwingend die Behandlungskosten pro Kopf der Bevölkerung (je Mengeneinheit müssen sie sogar sinken).
- 2) Beseitigung eines bestehenden Nachfrageüberhangs: Die (gemessen an der Primärnachfrage oder bei gegebenen, zumeist administrierten Preisen) zuvor ungenügende Versorgung der Patienten mit ärztlichen Leistungen kann bei steigender Ärztedichte sukzessiver verbessert werden. Behandlungsvolumen und Behandlungskosten pro Kopf der Bevölkerung steigen.
- 3) Rückgang indirekter Kosten: Die steigende Ärztedichte mindert für die Nachfrager indirekte Kosten des Arztbesuchs, insbesondere Wege- und Wartezeiten. Diese

in der Gegenüberstellung von Ärztedichte und direkten Behandlungskosten nicht berücksichtigte Senkung indirekter Kosten ruft eine höhere Nachfrage wach.

- 4) Umgekehrter Kausalzusammenhang: Wenn Ärzte effiziente Standortentscheidungen treffen, werden sie eine Ausbildung zum Arzt wählen bzw. dorthin gehen (und die Ärztedichte damit im Zeitablauf oder regional erhöhen), wenn bzw. wo die Nachfrage (bei gegebenen Krankheitsbildern, Altersstrukturen etc.) am größten ist. Hier ist die (verschieden große) Nachfrage exogen, die Ärztedichte dagegen endogen bestimmt.

## 4. Krankenhäuser: Leistungserbringung und Abrechnungssysteme

Literatur: Breyer, Zweifel und Kifmann (2005, Kap. 9)  
Breyer et al. (1988)  
Steinmann und Zweifel (2002)  
Lauterbach und Lungen (2000)  
Neubauer (1994)  
Neubauer und Nowy (2000)  
Sachverständigenrat f. d. Konz. Aktion i. Gesundheitswesen (2001)  
Schulenburg *et al.* (1998)

### 4.1 Motivation

Der Anteil an Gesundheitsausgaben für Krankenhäuser ist in den meisten entwickelten OECD Ländern beträchtlich gestiegen. Die Höhe der Ausgaben sagt aber nicht unbedingt darüber etwas aus, ob die verwendeten Mittel effizient und den Präferenzen der Bürger eingesetzt wurden. Wichtiger als die schiere Höhe der Ausgaben ist also eine Analyse der ökonomischen Probleme auf diesem Markt. Dazu zählen die Struktur des Marktes, die Struktur der Krankenhäuser und die Heterogenität des Outputs. Wirtschaftlichkeit erfordert, dass die richtige Menge von Gesundheitsgütern dem richtigen Konsumenten in der richtigen Zusammensetzung angeboten und zu geringst möglichen volkswirtschaftlichen Kosten hergestellt wird.

### 4.2 Strukturprobleme

Im Krankenhaussektor tauchen zwei ökonomische Probleme auf:

#### a) *Struktur des Marktes*

Weil Gesundheitskosten in der Regel über eine soziale Zwangsversicherung abgedeckt werden (abgesehen von einem minimalen Selbstbehalt), tragen die Konsumenten nicht die vollen Kosten der Behandlung. Die Versicherung deckt das Risiko der Krankheitskosten ab, das im Krankenhausbereich für den Einzelnen besonders hoch sein kann.

Da der Verbraucher den Preis nicht tragen muss, drückt seine Nachfragemenge nicht seine relative Wertschätzung des Gesundheitsgutes im Verhältnis zur Knappheit aus. Der Preismechanismus kann damit nicht mehr als Allokationsmechanismus für die den Verbraucherwünschen entsprechende Zuteilung von Ressourcen an die Wirtschaftszweige sorgen. Konsumenten fragen dann viel von allen Gesundheitsgütern nach, auch im Verhältnis zu anderen Konsumgütern, so dass ein

Anbieter nicht mehr durch den Preis in die Produktionssparte gelenkt wird, die am dringendsten nachgefragt wird.

Aufgrund dieses Institutionenversagens wird zwischen Krankenhäusern und Krankenkassen ausgehandelt, welche Mittel in den Sektor fließen. Die Gefahr ist, dass der Krankenhaussektor je nach Verhandlungsstärke zu viele oder zu wenige Leistungen erbringt.

#### b) Struktur der Firmen

Krankenhäuser arbeiten zumeist nicht gewinnorientiert. Dadurch haben sie auch keinen Anreiz an einer kostenminimalen Produktion. Die Zielfunktion eines Krankenhauses ist deshalb nicht einfach zu bestimmen. Sie setzt sich aus den einzelnen Interessen der Akteure zusammen. Dazu gehören:

- a) leitende und sonstige Krankenhausärzte
- b) Krankenschwestern und technisches Personal
- c) der Verwaltungsleiter
- d) der Krankenhausträger

Da ein Krankenhaus ein soziales Gebilde aus verschiedenen Gruppen von Akteuren mit jeweils eigenen Interessen darstellt, stellt die Zielfunktion eines Krankenhauses einen Kompromiss aus den Zielen der einzelnen Gruppen dar:

$$u = u(Q, \Pi, L^S, K^S) \quad \text{mit} \quad \frac{\partial u}{\partial Q} > 0, \quad \frac{\partial u}{\partial \Pi} > 0, \quad \frac{\partial u}{\partial K^S} > 0, \quad \frac{\partial u}{\partial L^S} > 0$$

Positiv bewertet werden vermutlich die Behandlungsqualität  $Q$ , der Einnahmenüberschuss  $\Pi$ , sowie das Ausmaß der brachliegenden Produktionsfaktoren  $L^S$  und  $K^S$ . Ziel eines Honorierungssystems muss es dann sein, zu technischer Effizienz zu führen, d.h. den Anreiz, Produktionsfaktoren brachliegen zu lassen, zu vermindern, und gleichzeitig maximale Behandlungsqualität zu gewährleisten.

### 4.3 Heterogenität des Krankenhaus-Outputs

#### a) Allgemein

Die Anwendung des Prinzips der Wirtschaftlichkeit auf diesen Markt erfordert zunächst eine Definition des Ertrages oder des Outputs eines Krankenhauses, was nicht so einfach ist. Die eigentliche Leistung ist *nicht* eine Operation, Bestrahlung, Medikation, Wundversorgung, Unterbringung oder Verköstigung, auch nicht die

medizinische Leistung, Pflegeleistung, Hotelleistung etc. Sondern diese Leistungen sind lediglich Mittel zu einem anderen Zweck. Dieser Zweck wird deutlich, wenn man sich klarmacht, was der Patient nachfragt: Positive Beeinflussung des Gesundheitszustandes, d.h. Heilung oder Eindämmung der Krankheit, Linderung der Schmerzen. Aber diese allgemeine Definition zieht weitere Probleme der Messung und Zurechnung der Leistung nach sich.

*Problem der Messung:* Soll der Gesundheitszustand bei Beginn und Ende des Krankenhausaufenthaltes gemessen und verglichen werden? Problem: Gesundheit ist mehrdimensional und die subjektive Erwartung der Patienten unterschiedlich.

*Problem der Zurechnung:* Der relevante Vergleich für die Beurteilung einer erfolgreichen Leistung des Krankenhauses wäre der Vergleich des Gesundheitszustand nach Krankenhausaufenthalt mit dem fiktivem Zustand ohne Krankenhausbehandlung, der sich am Ende des gleichen Zeitraumes eingestellt hätte. Dieser Vergleich ist allerdings hypothetisch. So entzieht sich das Ausmaß der durch das Krankenhaus verhinderten Verschlechterung des Gesundheitszustandes der Messung. Auch während des Aufenthaltes im Krankenhaus werden Leistungen erbracht, wenn z. B. Leiden verringert werden sollen bei unheilbar Kranken. Problem: Das subjektive Wohlbefinden ist genauso wenig verlässlich und objektiv messbar wie der Einfluss des Krankenhausaufenthaltes auf den Gesundheitszustand (Selbsteilungskräfte). Zudem gibt die Existenz des Krankenhauses den Einwohnern des Einzugsbereiches Sicherheit, im Falle einer Erkrankung, dort behandelt werden zu können. Diese Optionsnachfrage wird durch das Vorhalten von Betten einschließlich des zugehörigen Personals und der Geräte befriedigt.

*Fazit:* Der Output des Krankenhauses ist in der Verbesserung des Gesundheitszustandes der Patienten und der Bereithaltung von Kapazitäten zur Befriedigung der Optionsnachfrage zu sehen. Ersteres ist aber schwer messbar und dem Krankenhaus kaum adäquat zurechenbar.

#### *b) Mögliche Indikatoren des Outputs*

- Mengen der eingesetzten Produktionsfaktoren: Arbeitszeit der Ärzte und des Pflegepersonals, medizinischer Bedarf, Strom, Brennstoffe etc.
- Mengen der medizinischen und pflegerischen Einzelleistungen: Untersuchungen, Operationen, Medikation, Injektionen, Krankengymnastik, Mahlzeiten etc.
- Anzahl der Patienten bzw. Behandlungsfälle, differenziert nach Krankheitsart.
- Anzahl der Pflagetage, differenziert nach Intensität.

Die Produktionsfaktoren erstellen Einzelleistungen (sekundäre Inputs). Die Minimierung des Faktoreinsatzes kann als technische Effizienz bezeichnet werden.

Zwei Sichtweisen des Produktionsprozesses im Krankenhaus sind denkbar:

b1) Behandlungsfälle und Pflgetage als gleichrangige Zwischenprodukte (Behandlungsfälle: medizinische Komponente; Pflgetage: pflegerische Komponente). Unterschiedliche Einzelleistungen können ihnen als Inputs zugeordnet werden, so dass zwei Formen der Effizienz jeweils auftreten können: Pflegerische und interne medizinische Effizienz. Ersteres bedeutet den Einsatz möglichst geringer pflegerischer Leistungen je Tag und zweiteres möglichst geringe medizinische Leistungen je Fall- jeweils bei gegebener medizinischer Qualität.

Als externe medizinische Effizienz wird die Aufgabe des Gesundheitssystems bezeichnet, ein gegebenes Maß an Gesundheit mit möglichst geringen Kosten für Krankenhausaufenthalte zu erreichen.

b2) Pflgetage als Inputs in die Gesamtbehandlung eines Patienten. Dann wird Minimierung der Verweildauer bei gegebenem Krankheitsspektrum die Verweil-Effizienz genannt. Dies setzt voraus, dass ein Gesundheitszustand als Norm bei Entlassung des Patienten angesehen wird, ansonsten könnte die Verweil-Effizienz beliebig gesteigert werden, indem Patienten immer früher und kränker entlassen werden.

Das Bett als Input ist ein limitationaler Faktor für den Output Behandlungsfälle. Bei gegebener Anzahl gleichzeitig behandelter Fälle liegt der Bedarf an Betten eindeutig fest. Zusätzliche Betten haben eine Grenzproduktivität von Null, es sei denn man berücksichtigt die Optionsnachfrage, bei der auch ein leeres Bett eine Leistung des Krankenhauses darstellt. Dann wird der Teil der Faktorausstattung, der unabhängig von der Patientenzahl vorhanden sein muss, als Erstellung einer Vorhalteleistung begriffen und die Bettenzahl zu einem eigenständigen Indikator für den Output des Krankenhauses.

### c) *Klassifizierung des Outputs*

Die Darstellung des Krankenhauses als Produktionsprozess mit mehreren Inputs und einem Output „Gesundung“ klammert jedoch das Problem aus, dass gerade deshalb entsteht, weil der Output eines Krankenhauses durch große Heterogenität gekennzeichnet ist. 1.000 Fälle im Krankenhaus A sind nicht gleich 1.000 Fällen im Krankenhaus B. Zu Klassifizierung von ähnlichen und unterschiedlichen Behandlungsfällen können folgende Merkmale herangezogen werden:

- Art der Krankheit (Hauptdiagnose)
- Schwere der Krankheit bzw. Komplikationen
- Stadium der Erkrankung

- Etwaige weitere Krankheiten (Nebendiagnosen)
- Eigenschaften des Patienten (Alter, Geschlecht, etc.,)

Es kann nicht jeder Patient als besondere Produktart verstanden werden, da sonst die Leistungen der Krankenhäuser nicht miteinander verglichen werden könnten. Ein Kompromiss ist die Patientenklassifikation. Ziel dabei ist, möglichst homogene, aber überschaubare Gruppen von Behandlungsfällen zu bilden. Die gebräuchlichsten Patienten-Klassifikationen sind:

- *International Classification of Diseases (ICD)*  
Basis für Todesursachenstatistik; bezieht sich ausschließlich auf Hauptdiagnosen; über 900 Gruppen; 110 Obergruppen bereits sehr grob.
- *Diagnosis Related Groups (DRG)*  
Neben Hauptdiagnose als Einteilungsmerkmal auch Nebenerkrankungen und Komplikationen, das Alter des Patienten sowie die Behandlungsart (operativ oder konservativ), dennoch nur 467 Gruppen.
- *Patient Management Categories (PMC)*  
Stärkeres Gewicht auf Begleiterkrankungen sowie die gewählte Behandlungsstrategie, 840 Gruppen.

Im Folgenden soll die DRG-Klassifikation genauer betrachtet werden, da diese Klassifikation Grundlage der Krankenhausreform in Deutschland ist.

### *c1) Diagnosis-Related-Groups (DRG) als Grundlage der pauschalierten Vergütung*

Die DRG dienen als Grundlage für Fallpauschalen. Unter Fallpauschalen wird eine Honorierungsform verstanden, bei der alle anfallenden Leistungen (Personal und Sachmittel) je Behandlungsfall mit einem Pauschalbetrag abgegolten werden (Schulenburg *et al.* 1998, 39). Die Vergütung ist dabei unabhängig von der Verweildauer und vom individuellen Ressourcenverbrauch. Dafür muss jeder Behandlungsfall einer Fallgruppe zugeordnet werden (z. B. Kreuzbandruptur). Als wesentliches Ziel der Fallgruppenbildung ist die Homogenität des Ressourcenaufwands in den einzelnen Gruppen zu nennen. Die Übereinstimmung innerhalb einer Gruppe soll mit einer möglichst deutlichen Unterschiedlichkeit zwischen verschiedenen Fallgruppen einhergehen. Ferner ist zu fordern, dass Fallgruppen eine Zuordnung von Kosten erlauben und als Bezugsbasis für Preise geeignet sind. Außerdem sollte die Fallgruppenbildung medizinisch plausibel sein und die Überarbeitung aufgrund des medizinischen Fortschritts aber auch durch Änderungen im Krankheitsspektrum erlauben.

Die DRG wurden in den USA in einem *Top-down*-Ansatz auf breiter empirischer Basis definiert und von vornherein am Ressourcenverbrauch und auf

Kostenhomogenität ausgerichtet (Lauterbach und Lungen 2000, 168). Dabei sollten medizinisch vergleichbare Krankenhausfälle mit ähnlichen Kosten in Gruppen zusammengefasst werden. Die DRG werden seit 1983 als Basis für die Vergütung von Krankenhausleistungen durch Fallpauschalen eingesetzt. Die Grundprinzipien der DRG sind:

- Jedem Patienten wird pro Krankenhausaufenthalt genau eine DRG zugeordnet (eindimensionales, hierarchisiertes System). Die Abrechnung mehrerer Fallpauschalen oder zusätzlicher Sonderentgelte ist nicht möglich.
- Es existieren statistisch ermittelte Grenzverweildauern. Das heißt, Krankenhäuser können Zuschläge zu den Fallpauschalen erhalten, wenn für einzelne Patienten die Grenzverweildauer überschritten wird (*day outlier*).
- Es muss genau eine Hauptdiagnose bestimmt werden.
- Multimorbidität wird durch Nebendiagnosen und Schweregrad berücksichtigt. Ferner wird das Alter des Patienten einbezogen.
- Die Zuordnung zu den DRG erfolgt technisch über eine spezielle "Grouper-Software".

Das ursprünglich für die Behandlung von älteren Patienten konzipierte System wurde später zu den *All Patient Diagnosis Related Groups* (AP-DRG) weiter entwickelt, mit denen alle Behandlungsfälle im Krankenhaus erfasst werden sollen. Die nun ausgewählten „verfeinerten“ AR-DRG enthalten 661 abrechenbare Patientengruppen. Die Zuordnung des Einzelfalls erfolgt dabei zunächst in eine von 23 Hauptkategorien („*Main Diagnostic Category*“). Diese Hauptkategorien orientieren sich vor allen an Organen und Organsystemen oder – falls möglich – an Krankheitsursachen (z. B. Herzinfarkt). In der nächsten Stufe wird zwischen operativen und nichtoperativen Fallgruppen unterschieden. Dann werden Schweregrad, Nebendiagnosen und Komplikationen berücksichtigt. Für die Erfassung einiger schwer klassifizierbarer Fälle ist es nötig, auf Sammel-DRG („Fehler-DRG“) zurückzugreifen.

Als Grundlage des pauschalierten Vergütungssystems in Deutschland wurden die *Diagnosis Related Groups* (DRG) festgelegt. Dieses ursprünglich in den USA entwickelte Patientenklassifikationssystem wird inzwischen weltweit eingesetzt und wurde international unterschiedlich ausgestaltet. Mittlerweile wurde die Einführung der Australischen *Refined Diagnosis-Related-Groups* (AR-DRG) als Basis für das deutsche System beschlossen.

Letztendlich werden in den meisten Ländern, in denen die Vergütung über Fallpauschalen erfolgt, schätzungsweise lediglich 60 % der Krankenhausleistungen über DRG abgerechnet. Von vornherein ausgeschlossen sind Krankenhausleistungen u. a. in den Bereichen Psychiatrie, Onkologie und Pädiatrie,

da bei diesen Diagnosen lange Behandlungszeiten bzw. eine hohe Variabilität des Ressourcenverbrauchs gegeben sind. Insbesondere bilden die DRG keine Leistungen zur Krankheitsfolgenbewältigung ab und sind in dieser Hinsicht nicht für die medizinische Rehabilitation geeignet.

Ein weiterer wesentlicher Schritt bei der Entwicklung von Fallpauschalen ist die Festlegung der Entgelte. Beim amerikanischen DRG-System setzt sich die Vergütung aus mehreren Komponenten zusammen, deren Höhe sich nicht nur nach der zugeordneten DRG richtet. In der Regel ist jeder DRG ein relatives Gewicht zugeordnet, das aus den durchschnittlichen Kosten je Fall der jeweiligen Gruppe ermittelt wurde. In Deutschland sollen diese Relativgewichte den hiesigen spezifischen Verhältnisse angepasst werden. Das Relativgewicht wird mit einem Basispunktwert multipliziert, der das durchschnittliche Entgelt für alle Behandlungsfälle wiedergibt. Das so ermittelte Entgelt wird oftmals durch zusätzliche Faktoren multiplikativ verändert. Damit werden strukturelle Besonderheiten von Kliniken ausgeglichen, wie z. B. Kosten für Aus- und Weiterbildung, Teilnahme an der Notfallversorgung oder auch eine exponierte Lage. Solche Zu- und Abschläge sollen auch in Deutschland definiert werden.

### *c2) Reformen in Deutschland*

Durch die Novelle des Krankenhausfinanzierungsgesetzes vom 22.12.1999 wurde eine grundlegende Reform der Vergütung von Krankenhausleistungen eingeleitet. Ab dem Jahr 2003 sollen die Leistungen der Krankenhäuser über Fallpauschalen vergütet werden. Die Einführung der DRG soll im 2003 zunächst „budgetneutral“ erfolgen. Das heißt, die für die einzelnen Krankenhäuser ausgehandelten Vergütungsrahmen werden durch die DRG zunächst nicht verändert. Die DRG dienen damit vorerst vor allem der Dokumentation der Behandlungsfälle. Bis zum Jahr 2007 soll dann aber der Übergang zu einem pauschalierten Vergütungssystem geschaffen sein. Zum jetzigen Zeitpunkt sind allerdings wesentliche Aspekte des ordnungspolitischen Rahmens für die Einführung der Vergütung mit Fallpauschalen noch offen. U. a. geht es darum, ob die Fallpauschalen regional differenziert werden sollen, um beispielsweise höhere Kosten in Ballungsräumen ausgleichen zu können. Außerdem werden zurzeit Fragen der Mengensteuerung und des Ausgleiches von Mehr- oder Mindererlösen diskutiert. Das heißt, die Frage des finanziellen Ausgleichs, wenn ein Krankenhaus mehr oder weniger Patienten behandelt hat als zunächst geplant. Ferner stehen die Regelungen für Zuschläge zum Beispiel für die Ausbildung und Notfallversorgung oder die Behandlung von seltenen Erkrankungen oder Verletzungen (u. a. Brandverletzungen) noch nicht fest.

Die Reform der Vergütung verfolgt vor allem das Ziel, die Transparenz und die Wirtschaftlichkeit der Krankenhausleistungen zu erhöhen. Eine effizientere Leistungserbringung soll insbesondere über eine Verkürzung der Verweildauer und die Stärkung des Wettbewerbs zwischen den Krankenhäusern erreicht werden. Diese Zielsetzung verdeutlicht, dass sich durch die Fallpauschalen in den Krankenhäusern mehr ändern wird als lediglich die Abrechnung der Leistungen. In diesem Beitrag soll beschrieben werden, was unter Fallpauschalen zu verstehen ist und welche tief greifenden Veränderungen durch die Vergütung mit Fallpauschalen zu erwarten sind.

Fallpauschalen sollen Anreize für eine effiziente Leistungserbringung schaffen. Bei Fallpauschalen sind die Leistungserbringer bestrebt, den Ressourcenbedarf für die Behandlung der Patienten möglichst gering zu halten, da das Entgelt anders als beispielsweise bei der Vergütung über Kostenerstattung unabhängig vom Behandlungsaufwand ist (Lauterbach und Lungen 2000, 169). Eine Strategie, dieses Ziel zu erreichen, liegt in der Verkürzung der Verweildauer (Neubauer 1994, 39). Folglich wurde auch in anderen Ländern nach Einführung von Fallpauschalen generell eine Abnahme der Verweildauer beobachtet. Die im internationalen Vergleich relativ langen Verweildauern in Deutschland waren letztendlich ein wichtiger Grund für die nun erfolgte Reform. Während die Patienten im Jahr 1998 noch durchschnittlich 10,7 Tage im Krankenhaus verbrachten, wird damit gerechnet, dass durch die DRG-Einführung eine Verkürzung der Verweildauer, ähnlich wie in Australien, auf bis zu 6 Tage möglich ist (Neubauer und Nowy 2000, 21).

Als ein wesentlicher *Vorteil* von Fallpauschalen wird die Transparenz der Vergütung angesehen. Da die Höhe der Fallpauschalen für alle Leistungserbringer gleich ist, können sich alle Akteure unmittelbar über die Vergütung informieren. Damit ist eine Einschätzung der eigenen Leistungsfähigkeit im Medizinmarkt für die Leistungsanbieter möglich. Über diesen Vergleich wird ein Anreiz für die Erschließung von Wirtschaftlichkeitsreserven geschaffen. Durch Fallpauschalen kommt es außerdem zu einer Dezentralisierung von Entscheidungskompetenzen und Verantwortung in Richtung des Leistungserbringers, der eigenständig über den Ressourceneinsatz entscheiden kann.

Den hier beschriebenen Vorteilen von Fallpauschalen stehen aber auch *Fehlanreize* gegenüber, die u. a. zu einer Verschlechterung der Behandlungsqualität führen können (vgl. hierzu auch die Pressemitteilung der Sachverständigenrates für die Konzertierte Aktion im Gesundheitswesen vom 20.3.2001).

Die wichtigsten Fehlanreize von Fallpauschalen seien kurz genannt. Der von Fallpauschalen ausgehende Anreiz zur Minimierung des Ressourceneinsatzes kann

zu einer Reduktion der erbrachten Leistungen pro Behandlungsfall führen. Dies kann die Behandlungsqualität verbessern, wenn auf medizinisch fragwürdige oder eindeutig unnötige Therapieelemente verzichtet wird und die Belastung des Patienten durch die geringere Therapiedichte sinkt bzw. bestimmte unerwünschte Nebenwirkungen nicht auftreten. Zumindest sollte die Qualität dadurch nicht beeinträchtigt sein. Falls jedoch medizinisch oder pflegerisch erforderliche Therapien nicht gegeben werden, ist eine schlechtere Behandlungsqualität die Folge. Eine weitere Möglichkeit der Kosteneinsparung besteht bei der Entscheidung zwischen zwei oder mehreren Therapieverfahren, die jeweils medizinisch in Frage kommen, sich aber in den Kosten unterscheiden (Lauterbach und Lungen 2000, 169). Dabei kann ein Interessenkonflikt zwischen betriebswirtschaftlichen und medizinischen Überlegungen entstehen, wenn beispielsweise ein kostengünstigeres Verfahren einem aus medizinischer Sicht besseren Verfahren gegenübersteht.

Der Sachverständigenrat (Sachverständigenrat 2001, 90) weist ferner darauf hin, dass einige der Einsparungen in der Krankenhausversorgung durch die nachgehenden Versorgungsbereiche getragen werden müssen. Dies betrifft vor allem sowohl den Sektor der niedergelassenen Ärzte als auch die medizinische Rehabilitation, die häufig im Anschluss an Krankenhausbehandlungen durchgeführt wird. Nach einer Schätzung von Neubauer beträgt die zu erwartende Verlagerung aufgrund der verkürzten Verweildauer im Krankenhaus ca. 60 Mio. Leistungstage, die von anderen Systemen aufgefangen werden müssen, wenn es nicht zu einer Unterversorgung der Patienten in Teilbereichen kommen soll (Neubauer 2000, 2).

Ein weiterer, oft genannter Fehlanreiz bei Fallpauschalen ist das so genannte „Rosinenpicken“ (Neubauer 1994, 40). Damit wird die Tendenz bezeichnet, möglichst unkomplizierte Fälle zu behandeln und Patienten, bei denen die Behandlung aufwendiger erscheint, nicht aufzunehmen bzw. an andere Behandler zu verweisen. Der Anreiz zum Rosinenpicken wird umso geringer, je stärker die Vergütung der Fallpauschale nach Schwere der Erkrankung differenziert wird.

Nach Ansicht des Sachverständigenrates bedarf deshalb die Einführung von Fallpauschalen von Beginn an sowohl flankierende Maßnahmen der Qualitätssicherung als auch vorausschauende Konzepte zur Sicherstellung einer angemessenen nachgehenden Versorgung der Patienten (Nachsorge, Rehabilitation, Pflege; Sachverständigenrat 2001, 86). Die Einführung der Fallpauschalen sollte ferner durch die Entwicklung bzw. Anwendung entsprechender Behandlungsleitlinien begleitet werden. Eine vergleichbare Strategie wurde in den USA nach Einführung der DRG eingeleitet.

Durch die Einführung der Fallpauschalen wird es zu elementaren Umwälzungen sowohl in den betriebswirtschaftlichen Rahmenbedingungen und der Struktur der Krankenhäuser als auch in den Behandlungsabläufen kommen, die sich unmittelbar auf die Patientenversorgung auswirken werden. Wie tief greifend diese Veränderungen sein werden, hängt auch davon ab, wie der ordnungspolitische Rahmen für die Fallpauschalen letztendlich gestaltet wird. Beispielsweise wurde von Seiten der Krankenkassen die Einführung von Höchstpreisen für die Fallpauschalen in die Diskussion eingebracht. Auf diese Weise hätten die Krankenkassen die Möglichkeit, mit einzelnen Krankenhäusern niedrigere Preise zu verhandeln und diese Kliniken dann stärker zu belegen. Es ist offensichtlich, dass eine Höchstpreisregelung, deren Einführung zurzeit eher unwahrscheinlich scheint, den Wettbewerb zwischen den Kliniken deutlich verschärfen würde. Dabei sehen die Krankenkassen selbst die Preiskomponente als entscheidenden Faktor an, der so gestaltet sein sollte, dass ein ruinöser Preiswettbewerb unter den Leistungsanbietern zu Lasten des Versorgungsangebotes und der Versorgungsqualität vermieden werden kann.

Fallpauschalen können aber auch in ein Vergütungssystem eingebettet sein, das auf Krankenhausbudgets beruht, wie es für eine Übergangsphase ab 2003 vorgesehen ist. Dann wären die Auswirkungen voraussichtlich geringer. Die Ausgestaltung des Fallpauschalensystems wird sicher auch von den Erfahrungen in der Einführungsphase abhängig sein. Als sehr wahrscheinlich gilt mittlerweile, dass der ehrgeizige Zeitplan nicht einzuhalten sein wird.

Was die Folgen der Einführung von Fallpauschalen betrifft, so sind unmittelbare von mittelfristigen Effekten zu unterscheiden. Unmittelbar führt die Einführung der Fallpauschalen nicht zu Einsparungen im Gesundheitswesen. Dies liegt auch daran, dass die Kalkulation der Vergütung empirisch über die derzeitigen durchschnittlichen Behandlungskosten in den einzelnen Fallgruppen erfolgt. Mittelfristig wird die Reduktion der Verweildauer jedoch auch zu einer Abnahme der Vergütung pro Behandlungsfall führen und damit zu geringeren Aufwendungen im Krankenhausbereich beitragen. Ob sich die Gesundheitsausgaben insgesamt reduzieren, hängt davon ab, inwieweit es zu einer Kostenverlagerung in vorgeschaltete oder nachgehende Versorgungsbereiche kommt. Die zu erwartende deutliche Reduktion der Verweildauer in den Krankenhäusern wird voraussichtlich nicht durch eine Zunahme der Fallzahlen ausgeglichen werden. Deshalb wird die DRG-Einführung in den nächsten Jahren zwangsläufig zu freien Kapazitäten in den Krankenhäusern führen. Der Ideenwettbewerb um die Nutzung dieser Vakanzen hat bereits begonnen.

Die Einführung der DRG erfordert umfangreiche Vorbereitungen. Die Entwicklung und stetige Pflege des neuen Krankenhausfinanzierungssystems erfordert nicht unerheblichen personellen und finanziellen Ressourceneinsatz. Im Einzelnen geht es zunächst um die notwendige Anpassung der australischen Klassifikation an die Versorgungsstrukturen und Besonderheiten in Deutschland, die Kalkulation der Fallpauschalen und Bewertungsrelationen sowie die Entwicklung von Kodierrichtlinien. Nach Einführung des Systems ist eine ständige Pflege notwendig, die seine Anpassung an medizinische und medizintechnische Entwicklungen und Kostenentwicklungen gewährleistet. Dazu wurde zunächst eine DRG-Projektstelle eingerichtet, die die methodischen Voraussetzungen zur Kalkulation der Fallpauschalen erarbeiten soll. Zudem befindet sich ein DRG-Institut im Aufbau. Für die Finanzierung der beschriebenen Aufgaben wurde nun die Erhebung eines DRG-Systemzuschlags auf jeden Behandlungsfall in Höhe von 60 Pfennig beschlossen. Der Sachverständigenrat geht davon aus, dass nach Einführung des DRG-Systems die administrativen Kosten der Krankenhäuser dauerhaft oberhalb des Niveaus bei dem jetzigen Vergütungssystem liegen werden (Sachverständigenrat 2001, 88).

In den Krankenhäusern haben mittlerweile die Vorbereitungen für die Vergütung mit DRG begonnen. Dabei geht es zum einen um die Kodierung der behandelten Fälle. Es müssen u. a. die Hauptdiagnose, die Nebendiagnosen und die angewandten medizinischen Prozeduren dokumentiert werden. Dabei sind mit Prozeduren nicht nur chirurgische Eingriffe, sondern auch diagnostische und nichtoperative Maßnahmen gemeint. Die Qualität der Kodierungen ist letztendlich mitentscheidend für die Höhe der Vergütung. Als weitere Vorbereitungsmaßnahme auf die DRG wird in vielen Krankenhäusern die Kostenrechnung weiterentwickelt. Da sich die Vergütung nun auf den einzelnen Fall bezieht, ist es für das Krankenhaus wichtig zu wissen, wie sich die Gesamtkosten auf die Behandlung der Patienten verteilen.

#### *d) Die Kostenfunktion des Krankenhauses*

Die Kostenfunktion ist im Gegensatz zur Produktionsfunktion auch für Mehrprodukt-Unternehmen definiert: Jedem Outputbündel werden die minimalen Kosten der Erstellung dieser Ausbringungsmengen zugeordnet. Die empirische Ermittlung der Kostenfunktion gibt Aufschluss über:

- a) Die Art der Skalenerträge, die sich durch den Verlauf der Kostenfunktion ergeben. Daraus lassen sich Aussagen über die optimale Betriebsgröße ableiten. Das hat eine wirtschaftspolitische Bedeutung für den staatlich regulierten Krankenhausesektor wie in Deutschland und für Krankenhäuser, die ihre Bettenzahl nicht selbst festlegen können.
- b) Die Ableitung der Kostenfunktion nach Anzahl der Patienten eines bestimmten Falltyps ergibt die Grenzkosten der Behandlung dieser Patiententypen. Dadurch

erhält man Informationen zur Berechnung von Preisen im Rahmen leistungsorientierter Vergütungsformen.

- c) Die Höhe der Residuen einer Kostenschätzung beschreibt die Differenz zwischen tatsächlichen und geschätzten Kosten für die verschiedenen Krankenhäuser. Dadurch lassen sich Aussagen über die relative Wirtschaftlichkeit treffen. Das erleichtert die Kontrolle der Wirtschaftlichkeit von Krankenhäusern in einem System der Kostenerstattung.

Die mikroökonomische Kostenfunktion hängt ab von der Outputmenge, den Faktorpreisen und im kurzfristigen Fall von der Menge fixer Produktionsfaktoren. Eine Schätzung setzt voraus, dass alle einbezogenen Unternehmen sich kostenminimierend verhalten. Aufgrund der überwiegenden Existenz von gemeinnützigen und öffentlichen Trägern von Krankenhäusern ist aber von einem Gewinnstreben und damit auch von einer Kostenminimierung nicht auszugehen.

#### d1) Parametrische Kostenfunktion

Evans (1971) *Behavioural cost function* soll die bei tatsächlich beobachtetem Verhalten resultierenden Kosten erklären. Hier gehen dann auch Determinanten ein wie die Art der Trägerschaft; die Organisationsform des Krankenhauses; die Form der Vergütung, die sämtlich keinen Einfluss auf die Minimalkosten haben dürften, mit denen aber systematische Abweichungen vom Kostenminimum erklärt werden können.

Verhaltens-Kostenfunktion (nach Evans):

$$C = C(X_1, X_2, \dots, X_m; Y; Z; W; T; D_1, D_2, \dots, D_n)$$

$C$	Gesamtkosten des Krankenhauses
$X_1, \dots, X_m$	Vektor der Fallzahlen in $m$ verschiedenen Patientengruppen
$Y$	Zahl der Pflgetage
$Z$	Bettenzahl
$W$	Faktorpreisniveau
$T$	Trägerschaft
$D_1, \dots, D_n$	weitere Merkmale des Krankenhauses.

Quadratische Variablen sagen etwas über die Skaleneffekte aus. Bei Division der Gleichung durch die Fallzahl  $X = \sum X_i$  werden die Kosten je Fall als unabhängige Variable geschätzt.

Als Beispiel sei hier auf eine Studie von Breyer *et al.* (1988, siehe Breyer, Zweifel, Kifmann, 2005) verwiesen auf der Basis von 614 bundesdeutschen Krankenhäusern aus dem Jahr 1983.

Parametrische Kostenfunktionen haben jedoch auch Nachteile. Zum einen handelt es sich um mittlere Kostenfunktionen, d.h. die Regression geht durch die „Mitte der Punktwolke“. Auch ineffiziente Krankenhäuser beeinflussen die Grenzkosten und die optimale Betriebsgröße. Die Verfügbarkeit von Informationen zu den Faktorpreisen ist außerdem nicht immer gegeben. Ferner hängen die Resultate von der gewählten funktionalen Form ab. Nicht-parametrische Kostenfunktionen vermeiden diese Probleme.

#### *d2) Nicht-parametrische Kostenfunktion*

Ein Beispiel für eine nicht-parametrische Kostenfunktion stellt die Data Envelopment Analysis (DEA) von Charnes, Coopers and Rhodes (1978) dar. Grundlage für das Verfahren sind Input- und Outputmengen. Es ist also weder ein kostenminimierendes Verhalten der Krankenhäuser notwendig noch Kenntnisse der Faktorpreise.

Das Verfahren der DEA ermittelt eine empirische Grenze der Produktionsmöglichkeiten, an der die ineffizienten Beobachtungen gemessen werden können. Mit Hilfe eines Algorithmus werden dafür zunächst die Krankenhäuser bestimmt, die die effiziente Grenze aufspannen. Das Ausmaß der technischen Ineffizienz ist dann eine Funktion der Distanz zur effizienten Grenze. Der Vorteil im Vergleich zu parametrischen Kostenfunktionen liegt also darin, dass ineffiziente Krankenhäuser hier das Ergebnis nicht beeinflussen.

Der Nachteil dieses Verfahrens ist, dass eine konvexe Technologie definiert ist, die nicht-zunehmende Skalanelastizitäten impliziert. Diese Annahme ist jedoch problematisch, wenn die Größe des Krankenhauses keine Entscheidungsvariable des Managements ist. Eine geeignete Anpassung des Algorithmus ist also notwendig, um so variable Skalenerträge abzubilden und Skaleneffizienz, die z.B. der Träger zu verantworten hat, von technischer Ineffizienz, die auf schlechtes Management zurückzuführen ist, zu trennen. Auch dann bleibt aber noch das Problem, dass Ausreißer einen Einfluss auf die effiziente Grenze und somit auf die Resultate haben können.

Ein Beispiel stellt eine Studie von Steinmann und Zweifel dar (2002, siehe Breyer, Zweifel, Kifmann, 2005), in der die Effizienz von 310 schweizerischen Krankenhäuser für die Jahre 1993-1996 untersucht wird.

Für beide Arten von Kostenfunktionen gilt, dass ein Betriebsvergleich wegen der Heterogenität und der Komplexität sehr anspruchsvoll ist. Es ist somit nur möglich, relative Effizienzmaße zu konstruieren. Die Bewertung der Krankenhäuser erfolgt in Relation zu den besten Beobachtungen, aber es sind keine Aussagen zur absoluten Effizienz möglich.

## 5. Honorierung

Literatur: Breyer, Zweifel und Kifmann (2005, Kap. 10)  
 Breyer und Zweifel (1997, Abschnitt 7.2)  
 Laffont und Tirole (1993)  
 Dranove and Satterthwaite (2000)

Die Art der Honorierung medizinischer Leistungen ist ganz offensichtlich von zentraler Bedeutung für die Anreize, tatsächlich Kranke bzw. eine potentiell gesundheitlich beeinträchtigte Personengruppe unter Berücksichtigung von Kosten, Qualität und Verfügbarkeit der jeweils erforderlichen Behandlung effizient zu versorgen. Angesichts zahlreicher Marktunvollkommenheiten, insbesondere asymmetrischer Informationen, zu denen in der Praxis häufig auch noch Unvollkommenheiten der relevanten Institutionen treten, ist es sicherlich ein heroisches Ziel, ein „optimales Vergütungssystem“ für die Erbringer medizinischer Dienstleistungen – insbesondere Ärzte und Krankenhäuser – herzuleiten, das praktisch relevant sein soll. Aus der Analyse eines einfachen Grundmodells (Abschnitt 5.1) und einiger Modellvarianten mit um mehr Realismus bemühten Erweiterungen und Vertiefungen (Abschnitte 5.2 und 5.3) lassen sich gleichwohl Einsichten in grundlegende Eigenschaften verschiedener Vergütungsformen gewinnen und einige Aspekte, die für die Gestaltung realer Honorierungssysteme von Bedeutung sind, herausarbeiten (Abschnitt 5.4).

### 5.1 Das Grundmodell: Grundvergütung und volle Kostenverantwortung

Als „optimal“ werden Entscheidungen der Leistungserbringer und ihre Ergebnisse im Folgenden im Sinne der Zielfunktion eines objektiven Sachwalters der Patienteninteressen angesehen, dessen Rolle effektiv v.a. durch Krankenversicherungen oder staatliche Instanzen ausgefüllt werden kann. Gewisse Einschränkungen der Konsumentensouveränität von (potentiellen) Patienten werden im Modell damit anerkannt, spielen jedoch für sich genommen keine Rolle.

Der Sachwalter maximiere generell eine „Wohlfahrtsfunktion“  $W$ , in die der Behandlungsnutzen aller Patienten  $B$  (gemessen in monetären Größen) positiv, und die an den Leistungserbringer zu zahlende Vergütung  $P$  negativ eingehen. Der Nutzen des Leistungserbringers hänge ab von der Vergütung, abzüglich der ihm entstehenden Kosten der Behandlung, soweit sie durch die Vergütung nicht teilweise oder ganz kompensiert werden. Diese Kosten bestehen zum einen aus den tatsächlichen monetären Behandlungskosten  $K(e)$  mit stochastischer Natur, deren Erwartungswert  $C(e) \equiv E(K(e))$  (mit  $C_e < 0$ ,  $C_{ee} \geq 0$ ) ist, zum anderen aus Anstrengungskosten  $V(e)$  (mit  $V_e > 0$ ,  $V_{ee} > 0$ ). Die Anstrengung  $e$ , die der

Leistungserbringer im Modell unternimmt, richtet sich also darauf, die erwarteten Behandlungskosten zu senken, bereitet ihm jedoch zugleich einen Nutzenverlust. Im Grundmodell wird angenommen, dass sowohl Sachwalter als auch Leistungserbringer risikoneutral sind und bei ihren Entscheidungen daher einfach nur die Erwartungswerte für alle unsicheren Größen berücksichtigen. Falls der Leistungserbringer den vom Sachwalter angebotenen Behandlungs- und Honorarvertrag ausschlägt, realisiert er – auf der Basis irgendeiner Art von Alternativeinkommen – einen Reservationsnutzen in Höhe von  $\bar{u}$ .

Das optimale Anstrengungsniveau  $e^*$  ist jenes, bei dem der marginale Nutzenverlust  $V_e(e^*)$  exakt der marginalen erwarteten Kostensenkung  $-C_e(e^*)$  entspricht. Für den Sachwalter ist dieses Niveau von Bedeutung, weil er den Leistungserbringer für die ihm entstehenden Kosten kompensieren muss, um seine Partizipation bei ausreichendem Erwartungsnutzen  $E(u(P, e)) \geq \bar{u}$  sicher zu stellen. (Im Optimum bindet diese Nebenbedingung.) Wäre  $e$  beobachtbar, könnte der Sachwalter das erreichen, indem er neben dem Reservationsnutzen auch alle tatsächlichen Behandlungs- und Anstrengungskosten übernimmt, Abweichungen von  $e^*$  jedoch (prohibitiv) sanktioniert. Nimmt man realistischerweise an, dass  $e$  unbeobachtbar ist, lässt sich die *First-best*-Lösung trotzdem realisieren, wenn der Sachwalter eine fixe Grundvergütung anbietet, die  $\bar{u}$  zuzüglich der bei  $e^*$  erwarteten Kosten deckt und den Leistungserbringer alle Abweichungen von  $K$  gegenüber  $C$  tragen lässt. Die Höhe der Grundvergütung stellt dabei sicher, dass der Leistungserbringer den Vertrag akzeptiert; die volle Kostenverantwortung bewirkt, dass er das optimale Anstrengungsniveau  $e^*$  wählt.

Neben einer pauschalierten Grundvergütung erweisen sich im Rahmen mehrgliedriger Vergütungssysteme jedoch auch andere Komponenten als optimal, mindest im Sinne von *Second-best*-Lösungen, wenn man einige der zugegebenermaßen recht restriktiven Annahmen des Grundmodells reihum variiert bzw. zusätzliche Aspekte der Entscheidungssituation des Leistungserbringers beleuchtet und dabei weitere Annahmen über die Informations- und Kontrollmöglichkeiten des Sachwalters trifft. Wir beschränken uns im Folgenden auf Aspekte, die im Hinblick auf reale Steuerungsprobleme bei der Erbringung und Vergütung medizinischer Dienstleistungen als die vordringlichsten erscheinen.

## 5.2 Modellvarianten mit Kostenbeteiligung

Im Grundmodell bietet der Sachwalter dem Leistungserbringer keine direkte Beteiligung an den tatsächlichen Behandlungskosten – bis hin zur vollen Kostenübernahme – an, weil dies zwangsläufig einen negativen Effekt für die (als solche nicht beobachtbaren) Anstrengungen zur Kostensenkung hat.

Kostenbeteiligungen sind in der Praxis jedoch weit verbreitet, und dies kann in der Tat optimal sein, wenn den damit verbundenen Nachteilen hinsichtlich der Wirtschaftlichkeitsanreize aus der Sicht des Sachwalters andere Vorteile gegenüber stehen. Von besonderer Bedeutung ist dabei der Aspekt der Sicherung einer hinreichenden Behandlungsqualität, die im Grundmodell (bei gegebenem  $B$ ) als invariant angesehen wurde.

### 5.2.1 Risikoaverse Leistungserbringer

Eine einfache Modellvariante, in der die Grundvergütung durch eine Beteiligung an den tatsächlichen Behandlungskosten  $K$  ergänzt werden sollte, ergibt sich, wenn unterstellt wird, dass der Leistungserbringer nicht risikoneutral, sondern risikoavers ist. Während Ersteres als akzeptable Vereinfachung für das Verhalten institutioneller Leistungserbringer (Krankenhäuser) gelten kann, kann Letzteres vor allem bei niedergelassenen Ärzten eine Rolle spielen, die es jeweils mit vergleichsweise geringen Fallzahlen zu tun haben, bei denen individuelle Abweichungen von Erwartungswerten wie  $C$  nicht durch das Gesetz großer Zahlen ausgeglichen werden. Für den Sachwalter wird weiterhin angenommen, dass er risikoneutral sei.

Der Nutzen eines risikoaversen Leistungserbringers hängt neben dem Erwartungswert des Einkommens (bzw. der negativ darin eingehenden, unsicheren Kosten) z. B. auch von der Varianz der relevanten Größen ab. Bei Übertragung der vollen Kostenverantwortung müsste der Sachwalter dem Leistungserbringer daher auch eine Risikoprämie anbieten, die den daraus resultierenden Nutzenverlust kompensiert, damit dieser den Vertrag akzeptiert. Diese in der Grundvergütung enthaltene Prämie kann jedoch gesenkt werden, wenn der Sachwalter im Gegenzug einen Teil der tatsächlichen Behandlungskosten übernimmt. Die Höhe der optimalen Kostenbeteiligung  $\gamma K$  (mit  $0 < \gamma < 1$ ) hängt dabei ab von der Varianz von  $K$  und vom Grad der Risikoaversion des Leistungserbringers (sowie von den Kosteneffekten der induzierten Schwächung der Wirtschaftlichkeitsanreize).

### 5.2.2 Qualitätssicherung

In der Realität dürfte der Behandlungsnutzen oder -erfolg  $B$  nicht einfach gegeben, sondern seinerseits Gegenstand von Entscheidungen, v. a. des Leistungserbringers, sein. In einer entsprechenden Modellerweiterung kann dieser Nutzen als Funktion  $B(q)$  (mit  $B_q > 0$ ) modelliert werden, wobei  $q$  ein eindimensionales Maß für die Qualität (d. h. Intensität, Sorgfalt) der Behandlung durch den Leistungserbringer darstellt. Letzterer erbringt im erweiterten Modell somit, neben den Maßnahmen zur Kostensenkung, noch eine zweite (eigentlich „die“ typisch ärztliche) Anstrengung,

welche die erwarteten Behandlungskosten steigert ( $C_q > 0$ ), während ihr unmittelbarer Effekt für den Nutzen des Leistungserbringers ( $V_q$ ) à priori offen bleibt.

Wären die Inputgröße  $q$  beobachtbar oder zumindest der Behandlungserfolg  $B(q)$  verifizierbar (und dabei deterministisch durch  $q$  bestimmt), ließe sich auch in diesem Szenario eine *First-best*-Lösung realisieren, indem der Sachwalter  $(e^*, q^*)$  bzw.  $(e^*, B(q^*))$  vorschreibt, durch eine Grundvergütung die anfallenden Kosten und den Reservationsnutzen kompensiert und Abweichungen sanktioniert oder indem er  $q^*$  vorschreibt, die Grundvergütung auf die bei  $e^*$  erwarteten Kosten abstimmt und den Leistungserbringer bei voller Kostenverantwortung selbst die optimale Anstrengung  $e$  wählen lässt. Alternativ zum letzten Fall könnte der Sachwalter bei verifizierbarem  $B(q)$  neben der Grundvergütung auch eine vom Behandlungserfolg abhängige Komponente gewähren, die Grundvergütung selbst jedoch um den Betrag für  $B(q^*)$  kürzen, so dass der Leistungserbringer auch  $q^*$  freiwillig wählt.

Für praktische Zwecke ist jedoch eher der Fall relevant, dass weder  $q$  noch  $B(q)$  hinreichend gut beobachtbar sind bzw. dass beobachtete Realisationen von  $B(q)$  sich nicht hinreichend klar den Anstrengungen  $q$  zurechnen lassen. Ein Problem ergibt sich unter diesen Umständen, wenn (in der Umgebung von  $(e^*, q^*)$ )  $C_q + V_q > 0$ , weil der Leistungserbringer bei voller Kostenverantwortung dann nicht freiwillig die optimale Qualität  $q^*$  wählt. Eine solche Konstellation kann sich einstellen, wenn annahmegemäß  $V_e > 0$  und

- a)  $V_q > 0$  (für den Leistungserbringer bedeutet  $q$  stets einen zusätzlichen Nutzenverlust): Da der Leistungserbringer dann ohnedies stets nur die Mindestqualität anbietet, sollte er die volle Kostenverantwortung tragen, damit er wenigstens kosteneffizient arbeitet;
- b)  $V_q = 0$  (der Leistungserbringer ist indifferent bezüglich  $q$ ): Neben einer Lösung mit voller Kostenverantwortung und minimaler Qualität ist hier auch eine mit voller Kostenübernahme ( $\gamma = 1$ ) erwägenswert, weil der Leistungserbringer dann zwar keine Anstrengungen zur Kostensenkung trifft, aber zumindest  $q^*$  wählen könnte, weil ihm dies keinerlei Anstrengung abverlangt; eine Mischlösung ( $0 < \gamma < 1$ ) ist in jedem Fall suboptimal;
- c)  $V_q < 0$  (der Leistungserbringer zieht einen eigenen Nutzen aus qualitativ angemessener Behandlung; „ärztlicher Arbeitsethos“): In diesem Fall wägt der Leistungserbringer selbst ab zwischen Kostenminderung und Qualitätssteigerung und wählt eine innere Lösung; dabei kann er vom Sachwalter u. U. durch eine Grundvergütung, die einen Teil der erwarteten Kosten und den Reservationsnutzen deckt, kombiniert mit einer direkten Beteiligung an den verbleibenden Kosten so beeinflusst werden, dass sich – ganz im Sinne einer

*Second-best*-Lösung – die marginalen Wohlfahrtseffekte erhöhter Qualität und sinkender Kosteneffizienz genau die Waage halten.

### 5.3 Modellvarianten mit Fallpauschalen: Selektion von Patienten

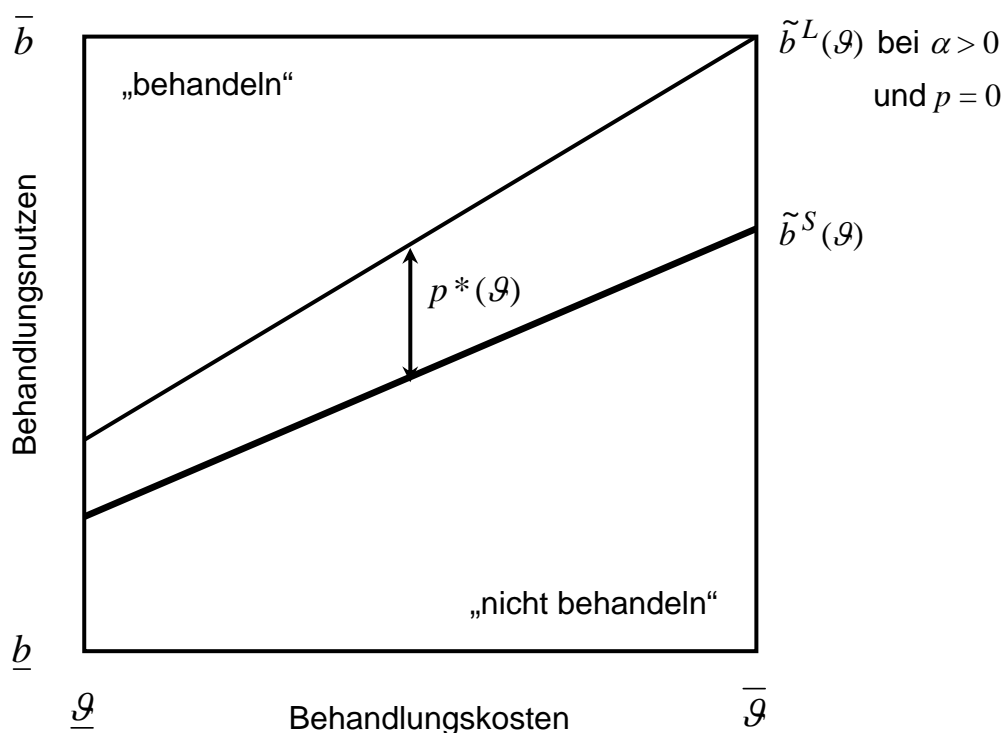
Eine weitere wichtige Entscheidung, die vom Leistungserbringer in der Regel ohne hinreichende Kontrollmöglichkeit durch die Patienten oder Dritte getroffen wird, bezieht sich auf die Auswahl von Patienten bzw. Krankheiten, denen er sich unter einer größeren Zahl potentieller Fälle widmet. Mit anderen Worten, der Behandlungsnutzen, der sich bisher stets auf einen definierten Kreis von (potentiellen) Behandlungsfällen bezog, kann individuell variieren, und das daraus resultierende Aggregat B ergibt sich auch aus Selektionsentscheidungen des Leistungserbringers, die nicht unbedingt im Sinne des Sachwalters sind.

Ein geeignetes Modell zur Analyse dieses Problems muss nun in Größen pro (potentiellem) Patient formuliert werden. Für ein Szenario mit individuellen Behandlungsnutzen  $b \in [\underline{b}, \bar{b}]$  und erwarteten individuellen Behandlungskosten  $c(\vartheta, e)$  mit fallspezifischen Maximalbeträgen  $\vartheta \in [\underline{\vartheta}, \bar{\vartheta}]$  – wobei die Verteilungen von  $b$  und  $\vartheta$  unabhängig voneinander seien – ergeben sich dabei die unterschiedlichsten denkbaren Kombinationen: Die Behandlung kann dringlich ( $b$  „hoch“), aber vergleichsweise billig sein ( $\vartheta$  „niedrig“); sie kann weniger dringlich ( $b$  „niedrig“) und dabei sehr aufwendig ( $\vartheta$  „hoch“) sein, etc. Im ersten dieser Fälle wird der Sachwalter die Behandlung selbst bei begrenzten Ressourcen für sinnvoll, im zweiten Fall möglicherweise nicht für sinnvoll halten, während die Schlussfolgerungen des Leistungserbringers jeweils von seinen Motiven und nicht zuletzt vom Vergütungssystem abhängen werden. Dabei dürfte der Leistungserbringer mit einem Informationsvorsprung mindestens hinsichtlich der effektiven Dringlichkeit der Behandlung ( $b$ ) ausgestattet sein.

Nimmt man an, dass der Sachwalter aufgrund der Basis einer Diagnose des Leistungserbringers zumindest den „Kostentyp“ ( $\vartheta$ ) des jeweiligen Patienten erkennen kann, so lässt sich eine *First-best*-Lösung realisieren, indem der Sachwalter – neben einer Grundvergütung – für jeden Behandlungsfall  $n$  eine diagnosespezifische Fallpauschale  $p(\vartheta)$  gewährt, welche die bei optimaler Anstrengung  $e^*$  erwarteten Behandlungskosten jeweils bis zu einer Obergrenze deckt, die dem fallspezifischen Behandlungsnutzen (abzüglich des möglichen internen Nutzens des Leistungserbringers aus der Behandlung verschieden dringlicher Fälle) entspricht. Die Grundvergütung hat dann die Funktion, den Reservationsnutzen des Leistungserbringers aufzuwiegen und somit seine Partizipation sicher zu stellen. Für Kosteneffizienz sorgt die Kostenverantwortung des

Leistungserbringers bei insgesamt pauschalierter und auf bei  $e^*$  erwartete Kosten abgestimmter Vergütung.

Abbildung 5.1 veranschaulicht diese Lösung für den Fall, dass der Leistungserbringer – wie in Abschnitt 5.2.2 – einen gewissen direkten Nutzen  $\alpha b$  (mit  $0 \leq \alpha < 1$ ) aus der angemessenen medizinischen Versorgung von Fällen mit unterschiedlicher Dringlichkeit zieht, der aber weniger stark ist als der vom Sachwalter berücksichtigte Behandlungsnutzen  $b$ . Für Sachwalter wie für Leistungserbringer gibt es eine von  $\vartheta$  abhängige Untergrenze  $\tilde{b}^S$  bzw.  $\tilde{b}^L$  für die Dringlichkeit eines Behandlungsfalls, ab der eine Behandlung erst sinnvoll wird, wobei diese Grenze bei „billigen“ Fällen jeweils niedriger liegt als bei „teuren“. Ohne Fallpauschalen ( $p = 0$ ), liegt die Grenze für den Leistungserbringer (wegen  $\alpha < 1$ ) höher als für den Sachwalter – sofern nur die Partizipationsbedingung erfüllt ist, kümmert jener sich um weniger Patienten als dieser für sinnvoll hält. Durch geeignet abgestufte Fallpauschalen  $p(\vartheta)$  mit  $p_\vartheta > 0$  können die beiden Untergrenzen jedoch exakt zur Deckung gebracht werden.



Für den Fall, dass der Sachwalter den Kostentyp  $\vartheta$  nicht beobachten kann, so dass er statt diagnosespezifischer Fallpauschalen nur noch ungezielte Kopfpauschalen gewähren kann, ist eine *First-best-Lösung* nicht mehr erreichbar. Bei hoher Kopfpauschale (z. B. bemessen nach  $\bar{\vartheta}$ ) behandelt der Leistungserbringer im Bereich niedriger  $\vartheta$  auch „zu wenig dringliche“ Fälle; bei niedriger Kopfpauschale

( $\vartheta$ ) vernachlässigt er im Bereich hoher  $\vartheta$  selbst „hinreichend dringliche“ Fälle; bei mittleren Werten der Pauschale trifft er Fehlentscheidungen beider Typs. (In Abbildung 5.1 lassen sich durch Kopfpauschalen nur Parallelverschiebungen von  $\tilde{b}^L$  bewerkstelligen, nicht auch die in der Regel erforderliche Drehung durch Variation von  $p$  mit  $\vartheta$ .) Hier ist erneut eine Kombination von Grundvergütung, evtl. ergänzt um Kopfpauschalen zur Beeinflussung der Zahl behandelter Patienten, mit einer direkten Kostenbeteiligung optimal. Die Kostenbeteiligung (mit  $0 < \gamma < 1$ ) hat auf die kritische Grenze  $\tilde{b}^L$  einen doppelten Effekt: einen (an sich unerwünschten) Niveaueffekt, da durch die sinkende Kosteneffizienz die erwarteten Kosten steigen und  $\tilde{b}$  nach oben verschoben wird, sowie einen (erwünschten) Struktureffekt, weil die Grenze durch die höhere (niedrigere) Beteiligung bei höheren (niedrigeren) Kosten im Sinne der Ziele des Sachwalters in die „richtige“ Richtung gedreht wird. Im *Second-best*-Optimum halten sich die marginalen Wohlfahrtswirkungen beider Effekte wiederum exakt die Waage.

#### 5.4 Zusammenfassung und Schlussfolgerungen

Eine wichtige Beobachtung aus der theoretischen Analyse ist, dass volle Kostenverantwortung des Leistungserbringers bei einer reinen Grundvergütung zu einer wirtschaftlichen Behandlung gegebener Fallzahlen und Krankheitsbilder führt. Die Grundvergütung ist dabei zwar der Höhe nach durch einige Modellparameter bestimmt, wird jedoch als reine Pauschale ausgezahlt. Eine solche Lösung ist wegen der stochastischen Natur der direkten Behandlungskosten für risikoaverse Leistungserbringer jedoch unattraktiv. Außerdem erweisen sich die damit gesetzten Anreize zur Kostendämpfung – ein sehr wichtiges Problem in den meisten realen Gesundheitssystemen – als einseitig, wenn die Leistungserbringer zugleich die Qualität der Behandlung und die Auswahl der behandelten Patienten steuern, ohne dass ihre Entscheidungen durch die Patienten oder neutrale Dritte („Sachwalter“ der Patienteninteressen) kontrolliert werden können. Lösungen („zweitbesten Art“) für diese Probleme bieten direkte Kostenbeteiligungen und im Idealfall nach der jeweiligen Diagnose differenzierte Fallpauschalen. In jedem Fall dürfte in der Regel ein gemischtes Vergütungssystem mit mehreren Honorarkomponenten optimal sein, wobei die genaue Struktur für verschiedene Bereiche medizinischer Dienstleistungen (ambulante vs. stationäre Behandlung, allgemeinmedizinische vs. fachärztliche Behandlung etc.) – je nach Art und Schwere der dort auftretenden Informationsasymmetrien – unterschiedlich ausfallen dürfte.

Bislang wurde hier drei grobe Typen von Vergütungskomponenten – Grundvergütungen ( $G$ ), Kostenbeteiligungen ( $\gamma K$ ) und Fall- bzw. Kopfpauschalen ( $pn$ ) – unterschieden. Im Hinblick auf reale Vergütungstypen in den Gesundheitssystemen entwickelter Industrieländer erweist sich diese Kategorisierung

jedoch noch als etwas grob und unvollständig (vgl. Tabelle 5.1). Zu berücksichtigen ist dabei allerdings auch, dass existierende Vergütungssysteme keinesfalls optimal sein müssen, sondern häufig historisch gewachsen und nur evolutorisch an die sichtbar werdenden Effizienzprobleme angepasst worden sind. Insbesondere ist in der Praxis häufig nicht sicher gestellt, dass alle vorhandenen oder zu geringen Kosten verfügbar zu machenden Informationen relevanter Art in die effektive Honorierung von Erbringern medizinischer Dienstleistungen eingehen, so dass nur wirklich unlösbare Informationsmängel und -asymmetrien verbleiben.

Wo in der theoretischen Analyse von „Kostenbeteiligungen“ die Rede war, implizieren die verwendeten Formalisierungen eine prozentuale Beteiligung an den tatsächlichen Behandlungskosten, die generell die Kosteneffizienz mindern. In der Praxis sind daneben im Detail abweichende Varianten zu beobachten, bei denen die Vergütung anknüpft an

- eingesetzten Produktionsfaktoren (Arbeitszeit des Arztes und seiner Mitarbeiter, Medikamente, ökonomische Abnutzung medizinischer Apparate, sonstige Kapitalkosten des Praxisbetriebes): Wenn die relativen Honorarsätze nicht exakt der marginalen Substitutionsrate der eingesetzten Faktoren bei der (Wieder-)Herstellung von „Gesundheit“ entsprechen, bietet eine solche Lösung keinen Anreiz zur effizienten Faktorkombination bei erbrachten Einzelleistungen und ihrer Mischung im Rahmen komplexer Behandlungen;
- erbrachten „Einzelleistungen“ (d. h. einzelne Beratungen und Behandlungen, die sich jeweils aus einem spezifischen Faktorbündel ergeben): In diesem Fall ist immerhin Produktionseffizienz bei der Erstellung der jeweils einzeln honorierten Leistungen gewährleistet; wenn die relativen Sätze nicht exakt darauf abgestimmt sind, führt die Lösung aber nicht zu einem optimalen Behandlungsmix im Sinne einer effizienten Leistungskombination.

Pauschale Vergütungsformen, die mit der Zahl tatsächlicher oder potentieller Patienten variieren, können anknüpfen an

- diagnostizierten Krankheitsbildern („Fallpauschalen“): Dies kann, vor allem im Hinblick auf die Selektion von Patienten mit unterschiedlich kostenträchtigen Erkrankungen, Teil einer *First-best*-Lösung sein;
- behandelte Patienten („Kopfpauschalen“ für jeden Behandlungsfall): Dies setzt zumindest Anreize, sich überhaupt einer aus der Sicht der Leistungserbringer nicht als gegeben anzusehenden Fallzahl anzunehmen; die Lösung erweist sich – zumindest für spezielle (v. a. chronische) Krankheiten, für die dann evtl. Sonderprogramme mit eigenem Budget aufgelegt werden – jedoch häufig als zu undifferenziert;
- registrierte Patienten („Kopfpauschalen“ für potentielle Behandlungsfälle): Diese Variante schafft für Leistungserbringer – v. a. „Hausärzte“ mit

Grundversorgungsfunktion – einen Anreiz, Patienten auch in qualitativer Hinsicht befriedigend zu versorgen; vorausgesetzt ist dabei, dass die Patienten die Behandlungsqualität zumindest grundsätzlich beurteilen können, ihre Informationen darüber austauschen und dass sie ihre Chancen auf eine gute Behandlung unter den verfügbaren Alternativen nötigenfalls durch einen Anbieterwechsel steigern.

Pauschalierte Grundvergütungen können gewährt werden als

- (periodenbezogenes *etc.*) Fixum (für als eigenständige Unternehmer agierende Ärzte, einschließlich eines Beitrags zur Deckung der Praxiskosten): Diese Lösung setzt v. a. die im Grundmodell erkennbar werdenden Anreize zu einer kosteneffizienten Behandlung, die durch andere Honorarkomponenten – teils jedoch mit gutem Grund – außer Kraft gesetzt werden;
- fixes Gehalt (für abhängig beschäftigte Ärzte): Eine solche Lösung bietet per se keinerlei Anreize zur Kosteneffizienz, erscheint jedoch dann als sinnvoll, wenn die zuvor diskutierten Informationsasymmetrien sich substantiell mindern lassen, wenn die Leistungserbringer als Beschäftigte des Sachwalters viel stärker unter seiner Kontrolle sind als in allen anderen Varianten; im Rahmen sogenannter *Health maintenance organisations* wird ein solches Modell (v. a. in den USA und in der Schweiz) auch in ansonsten stark privatwirtschaftlich geprägten Gesundheitssystemen praktiziert.

In der theoretischen Analyse angesprochen wurde schließlich auch die mögliche Gewährung von Erfolgsprämien (d. h. eine Vergütung oder Vergütungselemente, die den erreichten Behandlungserfolg – z. B. auf der Basis fallspezifischer Genesungsdauern, Rückfallraten, Komplikationsraten, Sterberaten o. ä. – einbeziehen). Soweit der Behandlungserfolg auf diese Weise objektiv messbar und hinreichend klar einschlägigen Anstrengungen des Leistungserbringers zuzurechnen ist, setzen erfolgsabhängige Honorarkomponenten direkte Anreize zur Qualitätssicherung.

In der Realität lässt sich in den letzten Jahren weltweit ein massiver Trend hin zu pauschalen Entlohnungsformen – Grundvergütungen, Fall- oder Kopfpauschalen – registrieren (vgl. erneut Tabelle 5.2), während bis vor kurzem noch Verfahren dominierten, die in erster Linie an Einzelleistungen anknüpften.

Tabelle 5.2: Honorierungssysteme in ausgewählten OECD-Ländern

Belgien	Einzelleistungsvergütung
Deutschland	Einzelleistungsvergütung
Japan	Einzelleistungsvergütung
Schweiz	Einzelleistungsvergütung
Dänemark	Pro-Kopf-Pauschale für registrierte Patienten, Praxiskostenunterstützung, Einzelleistungsvergütung für Sonderleistungen
Großbritannien	Pro-Kopf-Pauschale für registrierte Patienten
Irland	Pro-Kopf-Pauschale für registrierte Patienten mit Medical Card (33%), Einzelleistungsvergütung für andere Patienten
Italien	Pro-Kopf-Pauschale für registrierte Patienten plus Praxisunkostenhilfe
Niederlande	Pro-Kopf-Pauschale für registrierte Patienten
Österreich	Pro-Kopf-Pauschale für registrierte Patienten und Einzelleistungsvergütung von Sonderleistungen
Spanien	Pro-Kopf-Pauschale für registrierte Patienten für Ärzte in Gesundheitszentren; Einzelleistungsvergütung für Ärzte in Privatpraxen
USA	Pro-Kopf-Pauschale für registrierte Patienten oder Einzelleistungsvergütung
Finnland	Mischform aus Gehalt, Pro-Kopf-Pauschale für behandelte Patienten und Einzelleistungsvergütung
Frankreich	Gehalt für Ärzte in Gesundheitszentren; Einzelleistungsvergütung in Privatpraxen
Griechenland	Gehalt für Ärzte in ländlichen Gegenden; Einzelleistungsvergütung für Ärzte in Städten
Schweden	Gehalt in Gesundheitszentren (>80% der Ärzte), Einzelleistungsvergütung in Privatpraxen

Quellen: Österreichische Ärztekammer (2000), Schneider *et al.* (1995).

## 6. Krankenversicherungen

Literatur: Breyer, Zweifel und Kifmann (2005, Anhang 5.A, Kap. 5 und 6)  
Arrow (1985)  
Barr (1993, Kap. 5)  
Eckstein *et al.* (1985)  
Ehrlich und Becker (1972)  
Rothschild und Stiglitz (1976)

### 6.1 Motivation: Marktversagen auf Krankenversicherungsmärkten

Risikoaverse Individuen sehen angesichts der Möglichkeit, im Falle einer Krankheit Einkommensverluste (sowie darüber hinaus gehende Nutzeneinbußen) zu erleiden und Ressourcen für die Wiederherstellung ihrer Gesundheit aufwenden zu müssen, generell einen Vorteil darin, Maßnahmen zur Vorbeugung (Selbstversicherung) zu treffen und eine Versicherung (durch Dritte) abzuschließen. Allerdings ergeben sich auf dem Markt für Krankenversicherungen einige – auch aus anderem Kontext bekannte – Anlässe für Marktversagen, die einem effizienten Umgang mit Krankheitsrisiken, insbesondere einem aus gesellschaftlicher Sicht optimalen Versicherungsschutz, entgegen stehen. Diese Probleme – adverse Selektion und *Moral hazard* – haben eine gemeinsame Wurzel in Informationsasymmetrien: In beiden Fällen fehlen dem jeweiligen Anbieter einer Krankenversicherung Informationen, die für die Vertragsgestaltung wesentlich wären, und die die Nachfrager, gegeben die Anreizsituation vor oder nach Vertragsabschluss, letztlich nicht glaubwürdig übermitteln (signalisieren) können. Der erwünschte Versicherungsschutz kommt daher nicht oder nicht in vollem Umfang zustande.

„Chronologisch“ betrachtet tritt das Problem möglicher adverser Selektion bereits vor bzw. beim Abschluss von Krankenversicherungsverträgen auf, während *Moral-hazard*-Verhalten erst hinterher einsetzt, nämlich entweder bei der Beeinflussung des Krankheitsrisikos („*Ex-ante Moral hazard*“) oder, nachdem sich dieses Risiko materialisiert hat, bei der Beeinflussung der „Schadenshöhe“ im Krankheitsfall („*Ex-post Moral hazard*“). Da die problematischen Anreize für rationale Individuen, die asymmetrische Informationslage auszunutzen, aber jeweils vorhersehbar sind, ergeben sich in allen Fällen (Rück-)Wirkungen auf die Ausgestaltung der Verträge selbst.

Dabei unterscheiden sich die beiden, hier angesprochenen Problemkreise ihrer Natur nach sehr deutlich:

- Probleme adverser Selektion resultieren aus der unvollkommenen (im Extremfall: Un-)Beobachtbarkeit eines „*Datums*“, nämlich des individuellen Krankheitsrisikos

im Falle *heterogener* Individuen. Krankenversicherer müssen zwar die generell Verteilung dieses Risikos, etwa die Existenz und die Besetzung verschiedener Risikoklassen, kennen, um die von ihnen angebotenen Verträge überhaupt kalkulieren zu können. Sie können einzelne Nachfrager in der Regel aber nicht (oder nur eingeschränkt) in diese Verteilung einordnen. Ein voll auf das individuelle Risiko abgestellter Versicherungsschutz ist daher unmöglich.

- *Moral hazard* ergibt sich aus der unvollkommenen Beobachtbarkeit individuellen *Verhaltens* nach Abschluss einer Krankenversicherung. Krankenversicherer überschauen zwar die Anreizsituation ihrer (in dieser Hinsicht im Prinzip völlig *homogenen*) Versicherten, eigene Vorbeugung oder Maßnahmen zur Eindämmung des jeweils entstehenden Schadens zu unterlassen, wenn sie die Folgen der Krankheit nicht (in vollem Umfang) selbst zu tragen haben. Sie können aber kaum verifizieren, inwieweit die Individuen diesen Anreizen effektiv folgen. Ein Versicherungsschutz, bei dem das Auftreten solcher individuellen Verhaltensreaktionen berücksichtigt wird, ist daher ebenfalls ausgeschlossen.

Klar zum Ausdruck kommen diese Unterschiede etwa in der von Arrow (1985) vorgeschlagenen Terminologie, nach der die Ursache adverser Selektion auch als „*hidden information*“, die Ursache von *Moral-hazard*-Verhalten auch als „*hidden action*“ bezeichnet wird.

Ein weiterer Unterschied zwischen beiden Typen von Problemen sei hier bereits vorweg angesprochen: Während sich die Folgen adverser Selektion durch staatliche Eingriffe zumindest teilweise beheben lassen, ist das Phänomen des *Moral hazard* auch mit den Mitteln hoheitlicher Zwangsgewalt nicht lösbar. Unter Umständen verschärfen sich die damit verbundenen Probleme sogar, wenn die wechselseitige Verhaltenskontrolle kleiner Versichertenkollektive außer Kraft gesetzt wird und *Third-party payments* im Rahmen noch anonymere, staatlich kontrollierter oder administrierter Institutionen anfallen, bei denen überdies Mittel eingesetzt werden, deren Erhebung bereits bestimmte Effizienzverluste („*Marginal costs of public funds*“) erzeugt.

## 6.2 Adverse Selektion

Um die Probleme adverser Selektion zu analysieren, ist ein Modell mit heterogenen Individuen – mindestens zwei Typen *A* und *B* – erforderlich, die sich hinsichtlich der Höhe ihres (hier als nicht-beeinflussbar unterstellten) Krankheitsrisikos  $\pi_i, i \in \{A, B\}$  unterscheiden, wobei  $\pi_A < \pi_B$ . Vereinfachend wird im Folgenden unterstellt, dass der Nutzen der Individuen (auf der Basis einer einheitlichen *Ex-post*-Nutzenfunktion für alle Risikotypen) allein von ihrem verfügbaren Einkommen abhängt ( $u = u(y)$ ), dass der Gesundheitszustand binär ist – Individuen sind gesund

(g) oder krank ( $k$ ) – und dass eine etwaige Krankheit allein mit finanziellen Folgen (Behandlungskosten, Einkommenseinbußen) verbunden ist, die sich als Verlust  $L$  messen lassen.

Die (*Ex-ante*-)Erwartungsnutzenfunktion jedes Individuums lautet dann:

$$EU_i = \pi_i u(y_i^k) + (1 - \pi_i) u(y_i^g), \quad i \in \{A, B\}.$$

Das verfügbare Einkommen eines Individuums ergibt sich ohne Versicherung im Falle einer Krankheit aus dem „Bruttoeinkommen“  $Y$  und dem Verlust  $L$  als  $y_i^k = Y - L$ , anderenfalls ist es  $y_i^g = Y$ . Bei Abschluss einer Versicherung zahlt das Individuum in jedem Fall eine Prämie  $P$ , die von der Höhe des (frei wählbaren) Versicherungsschutzes  $I$  abhängt, im Falle der Krankheit erhält es dann allerdings die Versicherungssumme  $I$ , die es teilweise oder ganz für den eintretenden Verlust kompensiert. Sieht man von Transaktionskosten des Vertragsabschlusses und der Schadensabwicklung ab, ergibt sich eine Prämie, die linear in der Deckungssumme ist,  $P = \sigma I$ . Dabei ist  $\sigma$  der Preis jeder Geldeinheit von  $I$ .

### 6.2.1 Individuelle Optima und Soziales Optimum

Wenn das Individuum sein Einkommen in den alternativen Gesundheitszuständen, z. B. durch Abschluss einer Versicherung, frei wählen kann, wird es in diesem einfachen Modell eine Vollversicherung mit  $u'(y_i^k) = u'(y_i^g)$  und – bei einer strikt konkaven Nutzenfunktion – folglich  $y_i^k = y_i^g$  wählen. In jedem anderen Fall kann es seinen Erwartungsnutzen steigern, wenn es Einkommen vom Zustand mit dem kleineren Grenznutzen (höheren verfügbaren Einkommen) in den Zustand mit größerem Grenznutzen (kleinerem verfügbaren Einkommen) verschiebt, etwa indem es eine höhere Versicherungsprämie zahlt, um im Krankheitsfall eine höhere Versicherungssumme zu erhalten.

Da diese Überlegung unabhängig vom Risikotyp gilt, führt individuelle Nutzenmaximierung zu einer Situation, wo die (*Ex-post*-)Grenzraten der Substitution zwischen beiden Gesundheitszuständen für beide Risikotypen gleich Eins und damit identisch sind. Die *Ex-ante*-Grenzraten der Substitution, die sich unmittelbar aus der Erwartungsnutzenfunktion ( $\partial EU_i / \partial y_i^j$ ) ergeben, müssen sich dann aber zwischen den beiden Risikotypen unterscheiden, da die (im Optimum identischen) *Ex-post*-Grenzraten der Substitution jeweils mit  $\pi_i / (1 - \pi_i)$  gewichtet werden und sich die Krankheitswahrscheinlichkeiten ja systematisch unterscheiden.

Da die Überlegung zur Differenz zwischen den individuellen *Ex-post*-Grenzraten der Substitution beider Risikotypen außerdem unabhängig von den Einkommensniveaus in jedem Gesundheitszustand und vom Preis der Versicherung ( $P$  bzw.  $\sigma$ ) gilt,

impliziert sie ferner, dass die Indifferenzkurven der Erwartungsnutzenfunktionen beider Risikotypen nicht deckungsgleich sind: Vielmehr ist die Grenzrate der Substitution zwischen den Zuständen Krankheit vs. Gesundheit für Individuen des Typs  $A$  bei jeder möglichen  $\{y_i^k, y_i^g\}$ -Kombination größer als bei Typ  $B$ . Das bedeutet, dass die Indifferenzkurven von Typ  $A$  steiler im  $\{y_i^k, y_i^g\}$  Raum verlaufen als die Indifferenzkurven von Typ  $B$ .

Eine aus gesellschaftlicher Sicht optimale Allokation für den Fall heterogener Individuen unter Unsicherheit über die Entwicklung ihres Gesundheitszustandes lässt sich etwa auf der Basis einer „sozialen Wohlfahrtsfunktion“ herleiten, die die (irgendwie gewichtete) Summe aller Individualnutzen repräsentiert:

$$W = EU_A(y_A^k, y_A^g) + \delta EU_B(y_B^k, y_B^g).$$

(Wenn  $\delta$  dem Verhältnis der Bevölkerungsanteile von Individuen der Typen  $B$  und  $A$  entspricht, handelt es sich beispielsweise um eine rein „utilitaristische“ Nutzenfunktion.) Maximiert man diese Wohlfahrtsfunktion unter Beachtung der gesellschaftlichen Ressourcenbeschränkung – die mit den Bevölkerungsanteilen gewichtete Summe der erwarteten Einkommen ist gleich dem durchschnittlichen „Bruttoeinkommen“ abzüglich des bei gegebener Bevölkerungsstruktur zu erwartenden Verlustes  $L$  – gelangt man zu den selben Resultaten wie mit den Überlegungen zum individuellen Optimum: Im sozialen Optimum ist  $u'(y_i^k)/u'(y_i^g) = 1$ , d.h. beide Typen von Individuen sollten eine Vollversicherung erhalten, und die *Ex-ante*-Substitutionsraten variieren mit den unterschiedlichen Krankheitsrisiken.

### 6.2.2 Gleichgewicht im Versicherungsmarkt

Ein Gleichgewicht im Markt für Krankenversicherungen ist dadurch gekennzeichnet, dass alle Individuen den Vertrag wählen, der ihren Erwartungsnutzen maximiert, dass jeder dieser Verträge den Versicherern einen nicht-negativen Gewinn garantiert (mit Null-Gewinn als Grenzfall bei vollkommenem Wettbewerb) und dass kein Versicherer anders gestaltete Verträge anbieten kann (welche auch zumindest von einigen Individuen gekauft würden), die ihm einen höheren Gewinn versprechen. Wie üblich darf also keiner der Beteiligten einen Anreiz haben, von der Gleichgewichtssituation abzuweichen – sonst handelt es sich eben nicht um ein Gleichgewicht.

Entscheidend für die Frage, ob es auf dem Krankenversicherungsmarkt zu einer Risikoselektion, einschließlich der damit unter Umständen verbundenen Probleme, kommt, ist im Folgenden, ob es ein sogenanntes „*Pooling equilibrium*“ gibt, bei dem alle Individuen – unabhängig von ihrem individuellen Krankheitsrisiko – denselben

Vertrag abschließen, oder ein „*Separating equilibrium*“, mit unterschiedlichen Verträgen für die Risikotypen *A* und *B*.

### 6.2.3 *Common Knowledge*

Ist die Zuordnung jedes Individuums zu einer bestimmten Risikoklasse allseits (also auch dem Versicherer vor Vertragsabschluss) bekannt, existiert kein *Pooling equilibrium*. Angenommen, es würden zunächst nur einheitliche Verträge angeboten, mit einer Prämie, die dann auf das durchschnittliche Risiko aller Individuen abgestellt sein muss, könnte jeder Versicherer einen (höheren) Gewinn erzielen, wenn er Individuen mit niedrigem Krankheitsrisiko spezielle Verträge anbietet, für die er eine etwas niedrigere Prämie verlangt. Dieser Anreiz besteht zumindest solange, wie diese Prämie noch höher ist als das individuelle Risiko dieser Individuen im Hinblick auf die erforderliche Kostendeckung erzwingt. Die Vertragskonditionen für Individuen mit hohem Krankheitsrisiko müssen sich parallel dazu verschlechtern, damit die Versicherer bei diesem Typ keinen Verlust erleiden. Wettbewerb der Krankenversicherer und freie Vertragswahl der Individuen führen somit letztlich zu einer vollständigen Ausdifferenzierung individueller Verträge für Individuen vom Typ *A* bzw. *B*, wobei ein Gleichgewicht erst dann erreicht ist, wenn für beide Typen Verträge angeboten werden, deren Prämie exakt auf das individuelle Risiko abgestellt ist ( $\sigma_i = \pi_i$ ). Die Krankenversicherer machen dann unter diesen Bedingungen wieder keinerlei Gewinn, haben aber auch keinen Anreiz zur Abweichung mehr.

Bei öffentlicher Information über individuelle Krankheitsrisiken kommt es demnach zu einem *Separating equilibrium*, d. h. zur Risikoselektion, die in diesem Fall unter Effizienzgesichtspunkten jedoch kein Problem darstellt. (Die Bedingungen individueller und sozialer Optimierung werden ja erfüllt.) Lediglich unter Verteilungsgesichtspunkten kann man diese Lösung für problematisch halten, wenn man beachtet, dass die unterschiedlichen Risiken der Individuen angeboren oder in jedem Fall nicht beeinflussbar sind, und auf den Aspekt der Chancengleichheit oder der erwarteten Ungleichheit der (*Ex-post*-)Verteilung des verfügbaren Einkommens abstellt.

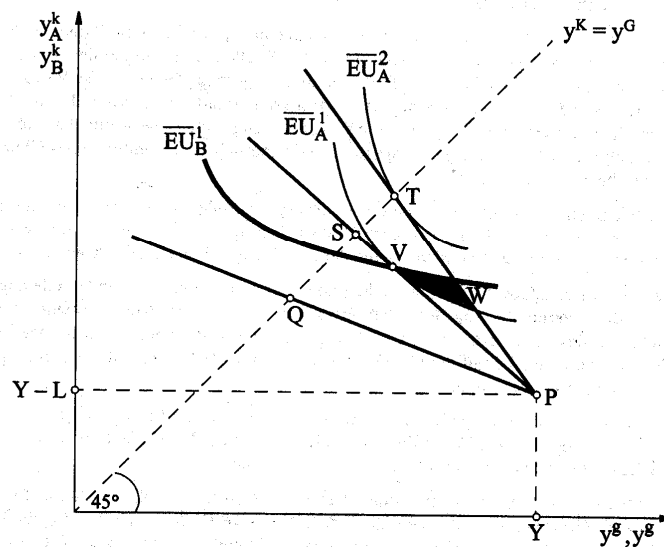
### 6.2.4 *Asymmetrische Informationen*

Ganz anders stellt sich die Situation dar, wenn die Informationen über die individuellen Krankheitsrisiken asymmetrisch sind, wobei die Krankenversicherer zwar wissen, dass es verschiedene Risikotypen gibt, aber kein Individuum einer bestimmten Klasse zuordnen können. Von den Individuen selbst wird hingegen angenommen, dass sie sehr wohl eine Vorstellung davon haben, welcher

Risikoklasse sie angehören. In diesem Fall kann der Versicherer Individuen vom Typ  $A$  bzw.  $B$  nicht gezielt unterschiedliche Verträge anbieten, die auf sie zugeschnitten sind. Er muss vielmehr einheitliche Verträge anbieten *oder* auf eine Selbstselektion angesichts seiner Vertragsangebote setzen. Ein Vertrag, der nur für Individuen mit niedrigem Krankheitsrisiko kalkuliert ist, muss also so ausgestaltet sein, dass er auch nur für solche Individuen attraktiv ist. Z. B. muss er mit einer reduzierten Deckung einher gehen, die für Individuen mit hohem Risiko – trotz der günstigeren Prämie – einer teureren, aber umfangreicheren Versicherung unterlegen ist. Somit spielen hier die unterschiedlichen Präferenzen der Individuen vom Typ  $A$  und  $B$  und entsprechend die unterschiedlichen Formen ihrer Indifferenzkurven eine entscheidende Rolle.

Ein *Pooling equilibrium* existiert auch im Fall asymmetrischer Informationen über individuelle Krankheitsrisiken nicht. Dies lässt sich graphisch leicht demonstrieren. In einem  $(y_i^k, y_i^g)$ -Diagramm repräsentiert der Punkt  $P = \{Y - L, Y\}$  die durch die Ausgangssituation jedes Individuums. Ausgehend von diesem Punkt kann jedes Individuum andere gesundheitsabhängige Einkommenskombinationen realisieren, indem es eine Versicherung mit zunehmender Deckung  $I$  und zunehmender Prämie  $P$  kauft. Das Ergebnis ist in dem einfachen Fall, der hier zugrunde gelegt ist, eine „Budgetgerade“ mit variierendem Versicherungsumfang, deren Steigung vom Preis jeder Versicherungseinheit abhängt. Individuell, d. h. für jeden Risikotyp „faire“ Prämien  $\sigma_i$  entsprechen jeweils der Krankheitswahrscheinlichkeit  $\pi_i$  und führen zu den Budgetgeraden  $\overline{PQ}$  und  $\overline{PT}$ . Die Kalkulation für eine Einheitsversicherung (also für ein *Pooling equilibrium*) basiert hingegen auf dem durchschnittlichen Krankheitsrisiko aller Individuen, mit  $\pi_A < \sigma_P = \pi_P < \pi_B$ , und ist durch die Budgetgerade  $\overline{PS}$  gegeben. Abbildung 6.1 veranschaulicht diese Zusammenhänge.

Abb. 6.1: Pooling Equilibrium bei asymmetrischer Information



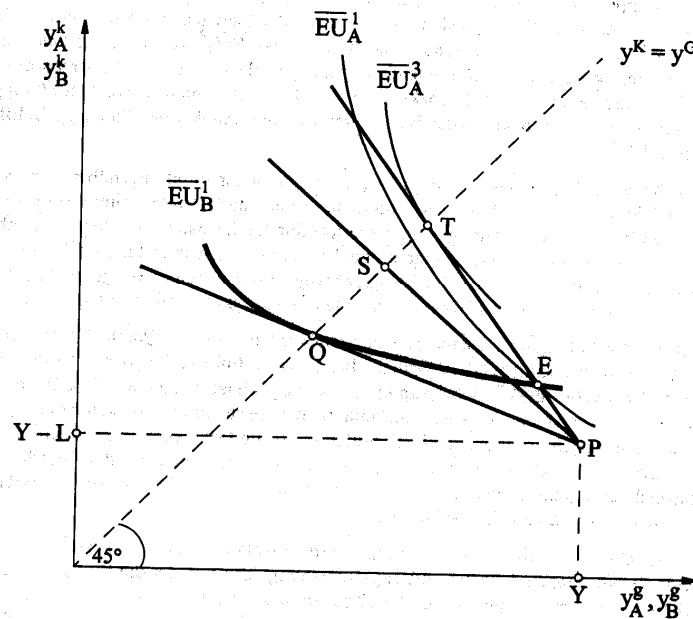
BZK, 2003, S.199

Ganz egal, wo auf der  $\overline{PS}$ -Gerade das *Pooling equilibrium* liegen soll, gilt aufgrund unserer Überlegungen in Abschnitt 6.2.1: Da Indifferenzkurven der Individuen vom Typ A im durch die Einheitsversicherung realisierten Punkt immer steiler verlaufen als die der Individuen vom Typ B, gibt es stets Verträge (mit geringerer Deckung und niedrigerer Prämie), die die Typ-A-Individuen der Einheitsversicherung vorziehen, die Typ-B-Individuen hingegen nicht. In Abbildung 6.1. liegen diese Verträge in der schwarz markierten Fläche. In diesem Fall kann eine Einheitsversicherung, die auf der  $\overline{PS}$ -Gerade liegt, also *kein* Gleichgewicht sein.

Wiederum tritt Risikoselektion ein, die nun – bei asymmetrischer Information über individuelle Krankheitsrisiken – allerdings problematisch ist: Die Lösung mit differenzierten Verträgen, falls sie überhaupt existiert, ist ineffizient, weil Individuen mit niedrigem Risiko nicht mehr der von ihnen gewünschte (Voll-)Versicherungsschutz angeboten werden kann. Vielmehr muss die Deckung von Verträgen mit niedrigen Prämien, die auf diese Individuen zugeschnitten sind, limitiert werden, damit sie nicht auch von Individuen mit hohem Risiko gekauft werden (was die Kalkulation des Versicherers untergraben würde).

Im Falle asymmetrischer Informationen ist allerdings noch nicht einmal klar, ob ein *Separating equilibrium* überhaupt existiert. In Abbildung 6.2 ist dies der Fall: Es gibt ein Gleichgewicht mit getrennten Verträgen – mit einem teuren Vollversicherungsschutz ( $\sigma_B = \pi_B$ ) für Individuen mit hohem Risiko und einem günstigeren Versicherungsschutz ( $\sigma_A = \pi_A$ ) für Individuen mit niedrigem Risiko, dessen Deckung aber so bestimmt sein muss, dass dieser Vertrag für Individuen vom Typ B im Vergleich zu „ihren“ Verträgen unattraktiv ist.

Abb. 6.2: Separating Equilibrium bei asymmetrischer Information



BZK, 2003, S.201

Ein Gleichgewicht ist dieser Punkt jedoch nur dann, wenn er für Individuen vom Typ A besser ist als ein Einheitsversicherungsschutz, d.h. wenn die Indifferenzkurve  $\overline{EU}_A^1$  die  $\overline{PS}$ -Gerade nicht tangiert oder schneidet. Wäre diese zusätzliche Bedingung verletzt (was bei einem etwas anderen Verlauf des Indifferenzkurvensystems  $\overline{EU}_A$  jederzeit eintreten kann), hätte der Wettbewerb der Versicherer um gute Risiken *kein* Gleichgewicht, und der Markt für Verträge mit Individuen vom Typ A bräche ganz zusammen.

### 6.3 Moral Hazard

Um die Verhaltensreaktionen zu untersuchen, die als *Moral hazard* bezeichnet werden, ist ein Modell mit homogenen Individuen ausreichend. Allerdings muss, abweichend vom vorangegangenen Abschnitt, nun unterstellt werden, dass entweder das individuelle Krankheitsrisiko oder die Höhe der Behandlungskosten (und anderer finanzieller Verluste) im Krankheitsfall durch das jeweilige Individuum beeinflussbar ist, und zwar durch Aktivitäten, die die Krankenversicherung nicht beobachten kann. Man spricht in diesem Zusammenhang auch von *Ex-ante-* bzw. *Ex-post-Moral hazard*, jeweils bezogen auf den möglichen Eintritt des Versicherungsfalles.

#### 6.3.1 Krankheitsrisiko und Vorbeugemaßnahmen (Ex-ante-Moral hazard)

Wir wenden uns zunächst dem Fall eines variablen Krankheitsrisikos zu (*Ex-ante-Moral hazard*), bei dem  $\pi$  eine abnehmende Funktion des Umfangs individueller

Vorbeugungsaktivitäten  $V$  ist, die allerdings gleichzeitig zu direkten Nutzeneinbußen führen, da sie finanzielle Aufwendungen mit sich bringen.

Die Erwartungsnutzenfunktion des durchschnittlichen Individuums lautet in diesem Fall generell:

$$EU = \pi(V) u(y^k) + (1 - \pi(V)) u(y^g).$$

Auf dieser Basis werden wir nun zunächst Szenarien optimaler Vorbeugung *ohne* Versicherung sowie optimaler Vorbeugung und optimaler Versicherung bei Beobachtbarkeit der individuellen Vorbeugungsmaßnahmen betrachten, um schrittweise das eigentliche *Moral hazard*-Problem zu entwickeln, das sich erst ergibt, wenn die Vorbeugungsmaßnahmen unbeobachtbar sind.

#### a) Optimale Vorbeugung ohne Versicherung

Ohne Versicherung (durch Dritte) kann das Individuum sein Erkrankungsrisiko durch Vorbeugungsmaßnahmen  $V$  steuern, die Ressourcen binden, die Wahrscheinlichkeit, tatsächlich zu erkranken, jedoch – u. U. mit abnehmenden Grenzzraten o. ä. – reduzieren. Die verfügbaren Einkommen im Falle von Krankheit und Gesundheit lauten somit  $y^k = Y - V - L$  bzw.  $y^g = Y - V$ . Als Bedingung erster Ordnung für ein optimales Maß an Vorbeugung ergibt sich (aus  $dEU / dV$ ):

$$\pi'(u(y^k) - u(y^g)) = \pi u'(y^k) + (1 - \pi) u'(y^g) \equiv E(u'(y))$$

Die linke Seite dieser Gleichung repräsentiert den marginalen Nutzenszuwachs, der mit einer zusätzlichen (Geld-)Einheit Vorbeugung verbunden ist; er ergibt sich aus der Differenz der direkten Nutzen in den alternativen Gesundheitszuständen, gewichtet mit dem marginalen Effekt von  $V$  auf  $\pi$ . Die rechte Seite bezeichnet hingegen die marginale Nutzeneinbuße durch die Aufwendungen für eine Einheit  $V$ , gemessen durch den Erwartungswert des Grenznutzens. Im Optimum betreibt das Individuum demnach gerade so viel Vorbeugung, dass diese beiden Grenzzraten identisch sind. Bei normalen Verläufen der relevanten Kurven wird das Individuum daher ein gewisses Maß an Vorbeugung betreiben, um das mit der Krankheit verbundene Verlustrisiko zu vermindern, und zwar um so mehr, je größer der Verlust im Krankheitsfall ist und je effektiver die Vorbeugung (pro eingesetzter Geldeinheit) das Risiko senkt. Allerdings hat Vorbeugung aus einzel- und volkswirtschaftlicher Sicht ihre Grenzen, wenn sie – bezogen auf den erwarteten Ertrag – zu aufwendig wird.

#### b) Optimale Versicherung bei beobachtbarer Vorbeugung

Nimmt man die Möglichkeit hinzu, sich gegen die finanziellen Folgen des Krankheitsrisikos zu versichern, so werden Vorbeugungsmaßnahmen aus der Sicht des

Individuums im Prinzip austauschbar gegen einen Versicherungsschutz durch Dritte, für den es ebenfalls Aufwendungen – die Versicherungsprämie – hat. Solange die Vorbeugungsmaßnahmen für den Versicherer beobachtbar sind, kann er den Preis der angebotenen Versicherung jedoch davon abhängig machen, ob und in welchem Umfang das Individuum nach wie vor Vorbeugung betreibt. Die Verhaltensreaktion durch die jeweils abgeschlossene Versicherung kann dadurch so beeinflusst werden, dass kein Anreiz zu *Moral hazard*-Verhalten entsteht, das hier im Unterlassen volkswirtschaftlich sinnvoller Vorbeugemaßnahmen bestünde.

Gegenüber a) ändern sich die Definitionen des verfügbaren Einkommens, das sich nun aus  $y^k = Y - V - P(I, V) - L + I$  bzw.  $y^s = Y - V - P(I, V)$  ergibt. Die Deckungssumme der Versicherung, die im Krankheitsfall ausbezahlt wird, wird dabei erneut als  $I$  bezeichnet;  $P$  ist die dafür zu entrichtende Prämie, die hier vom Umfang der Versicherung ( $I$ ) und der individuellen Vorbeugemaßnahmen ( $V$ ) abhängig gemacht werden kann. Bei einer „fairen“ (d. h. risikoorientierten und transaktionskostenfreien) Versicherung gilt z. B., dass  $P = \pi(V)I$ .

Der Erwartungsnutzen des Individuums ist somit eine Funktion sowohl von  $V$  als auch von  $I$ . Aus der Bedingung erster Ordnung für einen optimalen Versicherungsumfang ( $\partial EU / \partial I$ ) folgt, dass das Individuum im Falle einer fairen Versicherung eine Vollversicherung wählt (so dass  $u'(y^k) = u'(y^s)$ ,  $u(y^k) = u(y^s)$ ) und, bei strikt konkaver Nutzenfunktion,  $y^k = y^s$ ). Die Bedingung erster Ordnung für optimale Vorbeugung ( $\partial EU / \partial V$ ) führt zu

$$\pi'(u(y^k) - u(y^s)) = \left(1 + \frac{\partial P}{\partial V}\right)(\pi u'(y^k) + (1 - \pi) u'(y^s)) \equiv \left(1 + \frac{\partial P}{\partial V}\right)E(u'(y)).$$

Die linke Seite dieser Bedingung ist im Falle einer Vollversicherung gleich Null, während der Erwartungswert des Grenznutzens strikt größer als Null ist. Im Optimum muss daher  $1 = -\partial P / \partial V$  gelten, was bei einer fairen Vollversicherung mit  $\partial P / \partial V = \pi' I = \pi' L$  stets erfüllt ist. Diese letzte Bedingung lässt sich auch so interpretieren, dass das Individuum trotz der Vollversicherung, die es unzweideutig wählt, so viel Vorbeugung betreibt, dass die Grenzkosten (eine Einheit  $V$ ) genau dem Grenzertrag (Prämienminderung, d. h. mit dem Verlust gewichtete Risikominderung durch die zusätzliche Vorbeugungseinheit) entsprechen. Je nach den Verläufen von  $\pi(V)$  und  $u(y)$  kann dies mehr oder weniger Vorbeugung sein als im Falle ohne Krankenversicherung.

Da sich das Individuum mit vollständiger Deckung für  $L$  durch eine Krankenversicherung keinem Einkommensverlustrisiko (und den davon ausgehenden Effekten für  $u(y)$ ) mehr gegenüber sieht, konzentriert es sich nun darauf, den Erwartungswert seines Einkommens zu maximieren – die Bedingung

dafür lautet ebenfalls  $1 = -\pi' L$ . Die Gleichverteilung dieses maximalen Einkommenserwartungswerts wird dann über den Abschluss einer Krankenversicherung ohne Selbstbeteiligung bewerkstelligt. Der erwartete Nutzen ist dabei mindestens so groß wie im Fall ohne Versicherung (in der Regel jedoch höher), da ansonsten freiwillig keine Versicherung abgeschlossen würde.

### c) Optimale Versicherung bei nicht-beobachtbarer Vorbeugung

Das Problem reduzierter Vorbeugung angesichts der Möglichkeit, sich gegen die Verluste durch Krankheit extern zu versichern, tritt dann voll hervor, wenn der Krankenversicherer diese Art von Verhaltensanpassungen nicht beobachten und daher durch die Anreizwirkungen der Vertragsgestaltung auch nicht voll kontrollieren kann. Er weiß jeweils nur, wie ein rationales Individuum auf den Versicherungsschutz reagieren *würde* (nämlich in der Regel mit verringerter Vorbeugung) und muss sich bei Kalkulation und Vertragsgestaltung auf dieses Reaktionsmuster einstellen.

Formal führt die Unbeobachtbarkeit von  $V$  dazu, dass die Prämien  $P$  für die Versicherung nur noch vom gewählten Versicherungsumfang  $I$  abhängig gemacht werden können. Der Erwartungsnutzen des Individuums ist daher letztlich auch nur eine Funktion des jeweils gewählten Versicherungsschutzes  $I$ . In Abhängigkeit vom Versicherungsschutz, der durch die vom Versicherer festgelegte Prämie  $P(I)$  mitbestimmt wird, wird das Individuum jedoch – in einer logisch nachgelagerten Entscheidung – das aus seiner Sicht optimale Maß an Vorbeugeaktivitäten  $V(I)$  bestimmen. Dabei ist abzusehen, dass es desto weniger Vorbeugung betreiben wird, je umfassender sein Versicherungsschutz ist. Um eine angemessene Prämie zu kalkulieren, muss der Versicherer das Entscheidungsproblem des Individuums daher durch *Backward induction* lösen.

Maximiert man den Erwartungsnutzen des Individuums – gedanklich: bei gegebenem  $I$  und  $P(I)$  –, so erhält man aus der Bedingung erster Ordnung ( $dEU/dV$ ) nun eine „Reaktionsfunktion“  $V(I)$ , von der sich allgemein Folgendes sagen lässt: Für  $I=0$  entspricht die optimale Vorbeugung dem Ergebnis unter 5.3.1, d. h. sie ist in der Regel positiv. Für  $I=L$  (Vollversicherung) ist die optimale Vorbeugung gleich Null, da sie hier nur mit Kosten verbunden ist, jedoch nicht zu Erträgen in Form einer adäquaten, individuell spürbaren Abstufung der Versicherungsprämie führt. Im Bereich  $0 < I < L$  nimmt die optimale Vorbeugung somit tendenziell (allerdings nicht zwingend streng monoton) ab.

Die Prämie einer „fairen“ Versicherung, bei der der Preis jeder Einheit Versicherungsschutz grundsätzlich nach wie vor dem Risiko des (durchschnittlichen) Versicherten entspricht, muss diese Reaktion berücksichtigen. Allgemein gilt, dass

$P(I) = \pi(V(I)) I$ . Die zweite Ableitung dieser Prämienfunktion legt nahe, dass die Prämie progressiv mit dem gewählten Versicherungsschutz steigen muss, um das (aufgrund sinkender Vorbeugung) steigende Risiko jeweils angemessen zu reflektieren.

Gegeben diese Prämienfunktion und die (auch vom Versicherer berücksichtigte) Verhaltensreaktion bei der Vorbeugung wird das Individuum seinen optimalen Versicherungsschutz (aus  $dEU/dI$ ) schließlich gemäß der Bedingung

$$P'(I) E(u'(y)) = \pi u'(y^k)$$

wählen: Die linke Seite dieser Gleichung misst die marginale Nutzeneinbuße durch einen erhöhten Versicherungsschutz wegen des damit verbundenen Prämienanstiegs; die rechte Seite den marginalen Nutzenzuwachs durch das im Krankheitsfall erhöhte Einkommen. Im Optimum gleichen sich somit einmal mehr die Grenzkosten und Grenzerträge einer zusätzlichen Einheit der Deckungssumme aus.

Für das Individuum ist angesichts dieser Situation auch bei einer fairen Versicherung kein Vollversicherungsschutz mehr erstrebenswert – oder jedenfalls ein geringerer als unter b). (NB: Es sei daran erinnert, dass der optimale Vollversicherungsschutz zuvor an einer Reihe zusätzlicher Annahmen hing: dass die angebotene Versicherung „fair“ ist, dass die Krankheit rein finanzielle Verluste mit sich bringt, etc. Für diesen Fall steht jedoch aufgrund der formalen Optimalitätsbedingungen fest, dass eine Vollversicherung unter den Rahmenbedingungen von Fall c) als optimale Lösung ausscheidet. Die Kombination von Vollversicherung ( $I = L$ ), völlig unterlassener Vorbeugung ( $V(L) = 0$  wegen der Reaktionsfunktion  $V(I)$ ) und entsprechend hohen Prämien ( $P = \pi(0)L$ ) hätte im Prinzip zuvor schon freiwillig gewählt werden können. Sie erwies sich dort jedoch der Kombination  $\{L, V > 0, \pi(V)\}$  als unterlegen.) Um stark überproportional steigende Prämien zu vermeiden, begnügt sich das Individuum mit einer geringeren Deckungssumme und reduziert das Verlustrisiko (individuell und gesamtwirtschaftlich) auch durch eigene Vorbeugungsaktivitäten. Sein Nutzen ist dabei allerdings kleiner als im Falle beobachtbarer Vorbeugung, die durch die Konditionen des Versicherungsvertrages honoriert wird.

Die Anforderung einer (im Versicherungsschutz) nicht-linearen Prämiengestaltung setzt voraus, dass der Versicherer Kontrolle über den gesamten Versicherungsschutz des Individuums hat – jeder Versicherungsschaden darf nur mit einer Versicherung abgerechnet werden – und dass er den Verlauf der Reaktionsfunktion  $V(I)$  sowie die Effektivität der Vorbeugungsmaßnahmen  $\pi(V)$  kennt. Lassen sich diese Anforderungen, z. B. durch diverse Ergänzungsversicherungen oder aufgrund noch begrenzterer Informationen, nicht

realisieren, kann der Versicherer nur eine lineare Prämie verlangen. Wenn er dann nicht zusätzlich  $I$  limitiert, wird er davon ausgehen, dass das Individuum insgesamt einen Vollversicherungsschutz wählt (u. U. verteilt auf mehrere Versicherungen) und keinerlei Vorbeugung treibt. Die geforderte Prämie entspricht dann jeweils der maximalen Krankheitswahrscheinlichkeit.

### 6.3.2 Behandlungskosten (Ex-post-Moral hazard)

Eine etwas andere Art von (*Ex-post*-)Moral hazard-Verhalten kann eintreten, wenn das Individuum nach Eintritt einer Krankheit aufgrund des Versicherungsschutzes dem Anreiz folgt, Behandlungskosten ohne Rücksicht auf ihre Höhe entstehen zu lassen. (Je nach Honorierungssystem wird der Arzt es dabei, über das Maß medizinisch angemessener bzw. sinnvoller Behandlungen hinaus, unterstützen und u. U. sogar aktiv dazu anhalten, vgl. Kapitel 5. Die im Folgenden dargestellten Überlegungen beleuchten dasselbe Problem von einer anderen Seite, ergeben jedoch nur einen Sinn, wenn das Individuum zumindest im Grundsatz eigenständige Nachfrageentscheidungen fällt, die sich durch die Gestaltung seines Versicherungsschutzes beeinflussen lassen.) Diese im Prinzip vorhersehbare Verhaltensreaktion hat Rückwirkungen auf den Preis der Versicherung, kann bei der Gestaltung von Versicherungsverträgen jedoch auch von vornherein berücksichtigt und eingedämmt werden.

In einem einfachen Modell, mit dessen Hilfe das Problem und die Struktur möglicher Lösungen erfasst werden können, maximiert das Individuum wiederum seinen Erwartungsnutzen

$$EU = \sum_s \pi_s u(f(H_s) - M_s - P + I_s) = \sum_s \pi_s u(f(\theta_s + M_s) - M_s - P + I_s).$$

auf der Basis einer Nutzenfunktion in Abhängigkeit vom jeweils verfügbaren Einkommen über mehrere mögliche Gesundheitszustände  $H_s$ ,  $s = \{1, \dots, S\}$ , die sich mit der Wahrscheinlichkeit  $\pi_s$  realisieren. (Die Gesundheitszustände werden dabei aufsteigend von  $s = 1$  „schlecht“ bis  $S$  „gut“ gereiht.) Der verwirklichte Gesundheitszustand  $H_s$  hat dabei eine stochastische Komponente  $\theta_s$ , die sich durch Behandlungskosten  $M_s$  modifizieren („verbessern“) lässt, so dass  $H_s = \theta_s + M_s$ . Der Gesundheitszustand nach Behandlung beeinflusst zugleich das Bruttoeinkommen des Individuums,  $Y = f(H_s) = f(\theta_s + M_s)$ , aus dem sich das verfügbare Einkommen durch Abzug der Versicherungsprämie  $P$  sowie – im Krankheitsfall – durch Subtraktion der Kosten  $M_s$  und Addition der vereinbarten Versicherungsleistung  $I_s$  ergibt.

Das hier verwendete Grundmodell ist so angelegt, dass jedes Individuum *ex post* ein gewisses Maß an Behandlung – von Bagatell-Arzneimitteln bis hin zu

kostenträchtigen Operationen – in Anspruch nimmt. Dies bewirkt die Annahme  $f'(\theta_s) \geq 1$ , die impliziert, dass sich eine Behandlung  $M_s$  (mit auf Eins normiertem Preis jeder Einheit von  $M$ ) selbst beim bestmöglichen Gesundheitszustand noch lohnt. Das ist eine Stilisierung, die lediglich die Analyse vereinfacht.

*a) Optimale Versicherung bei beobachtbarem Gesundheitszustand ( $\theta_s$ )*

Wäre der spontan eintretende Gesundheitszustand  $\theta_s$  für die Versicherung beobachtbar, so dass auch die „passende“ Behandlung mit Kosten  $M_s$  feststeht, könnte die Versicherung dem Individuum freie Wahl hinsichtlich der gewünschten Deckung  $I_s$  lassen.

Ist die Versicherung „fair“, mit  $P = p_1 I_1 + \dots + p_s I_s$ , so ergibt sich für diesen Fall erstens, dass das Individuum im Optimum eine durch die Realisation von  $\theta_s$  bestimmte Behandlung nachfragt, die für jedes  $s$  zum selben End-Gesundheitszustand  $H^*$  führt, bei dem die Bedingung  $f'(H^*) = 1$  erfüllt ist. (Hier entspricht der Grenzertrag der Behandlung genau den Grenzkosten. Der Aufwand dafür wird mit steigendem  $s$  immer kleiner.) Zweitens zeigt sich, dass das Individuum in diesem Fall einen Versicherungsschutz wählt, der bewirkt, dass der Grenznutzen im Optimum unabhängig von  $s$  wird. Dies kann daher nur ein Vollversicherungsschutz, mit  $I_s = M_s$ , (bzw. eine gleich hohe Sachleistung) sein oder ein Schutz, bei dem  $I_s$  von  $M_s$  durch eine Konstante (nicht einen proportionalen Selbstbehalt o. ä.) verschieden ist.

*b) Optimale Versicherung bei nicht-beobachtbarem Gesundheitszustand ( $\theta_s$ )*

Kann die Realisation von  $\theta_s$  von der Versicherung nicht beobachtet werden, sondern nur vom Individuum und/oder seinem Arzt, haben letztere einen Anreiz, die Behandlung über das wirtschaftlich sinnvolle Maß hinaus zu treiben, solange die Versicherung für die damit verbundenen Kosten (voll) aufkommt. Da  $\theta_s$  und die bekannten Wahrscheinlichkeiten  $\pi_s$  keine Anhaltspunkte für die bei optimalen Nachfrageentscheidungen erwarteten Behandlungskosten mehr liefern, wird der unter a) beschriebene Versicherungsvertrag unkalkulierbar.

Die Versicherung kann unter diesen Umständen nur noch eine Erstattungsfunktion  $I(M)$  für die jeweils geltend gemachten Behandlungskosten  $M$  anbieten, von der vereinfachend angenommen wird, dass sie einen proportionalen Selbstbehalt  $cM$  für das Individuum vorsieht. Es lässt sich zeigen, dass die Nachfrage nach ärztlicher Behandlung nach Abschluss eines solchen Vertrages höher ausfällt als bei beobachtbarem  $\theta_s$ , solange  $c < 100\%$ . Andererseits geht sie mit zunehmendem Selbstbehalt jedoch eindeutig zurück. Entscheidend ist hier, dass der Erwartungsnutzen des Individuums in diesem Fall eindeutig zunimmt, wenn  $c$  – ausgehend von einem Vertrag ohne proportionalen Selbstbehalt ( $c = 0$ ) – erhöht wird.

Der prämiensenkende Effekt des Selbstbehalts, der aufgrund der erhöhten Kostendisziplin überproportional ausfällt, überwiegt den erwarteten Verlust durch die Übernahme eines eigenen Kostenanteils zunächst eindeutig. Im Optimum müssten sich diese Einzeleffekte exakt ausgleichen. In jedem Fall wählt das Individuum in diesem *Second-best*-Szenario jedoch freiwillig einen Vertrag mit positivem Selbstbehalt.

#### 6.4 Lösungsmöglichkeiten

Wege, die in diesem Kapitel erörterten Anreizprobleme aufgrund von Informationsasymmetrien zu mildern oder – im Sinne eines *Second-best*-Optimums – zu lösen, wurden in der Analyse jeweils bereits angesprochen. Die Aussagen dazu sollen abschließend noch einmal zusammengefasst und durch Hinweise bezüglich der praktischen Umsetzung dieser Lösungsmöglichkeiten ergänzt werden.

Risikoselektion, die auf wettbewerblich organisierten Versicherungsmärkten eintritt, lässt sich – mit oder ohne asymmetrische Informationen – durch staatlichen Zwang reduzieren oder sogar ganz ausschließen. Entsprechende Maßnahmen bestehen in der Regulierung von Versicherungsverträgen (Ausschluss bestimmter Differenzierungen, Überwachung der Tarif- und Prämiengestaltung), der Verhängung eines Kontrahierungszwanges für Versicherer und u. U. auch die Verhängung einer Versicherungspflicht für die Bevölkerung. Staatlich organisierte Pflichtversicherungen enthalten in der Regel alle diese Elemente (in extremer Variante) und können sich im Hinblick auf Transaktionskosten (kurzfristig) als überlegen gegenüber einem noch so strikt regulierten Wettbewerb unter Krankenversicherern erweisen. Dafür leiden sie (langfristig) aber häufig unter mangelnder dynamischer Effizienz.

Staatlicher Zwang, der Risikoselektion begrenzt, hat unterschiedliche Wirkungen, je nachdem ob er vor dem Hintergrund öffentlicher Informationen oder asymmetrischer Informationen über individuelle Krankheitsrisiken ausgeübt wird.

- Wenn individuelle Krankheitsrisiken bekannt sind, ist Risikoselektion eine effiziente Lösung, die lediglich dazu führt, dass jedem Individuum die von ihm gewünschte Versicherung zu einem nach dem jeweiligen Risiko abgestuften Preis angeboten wird (vgl. 6.2.3). Eine staatliche Begrenzung differenzierter Verträge führt demnach zu Effizienzeinbußen (die Marginalbedingungen für individuelle Optima werden verletzt) und verteilt durch die Vereinheitlichung der Versicherungsbedingungen von Individuen mit niedrigem Risiko zu solchen mit hohem Risiko um. Je nach verteilungspolitischer Zielsetzung kann darin ein Gewinn an Verteilungsgerechtigkeit gesehen werden, der wie üblich seinen Preis *in puncto* Effizienz hat.

- Im Falle asymmetrischer Informationen wird die Risikoselektion auch unter Effizienzgesichtspunkten problematisch (vgl. 6.2.4). Im Extremfall kommen akzeptable Versicherungen für Typ-A-Individuen unter Wettbewerb überhaupt nicht zustande. Im diesem Fall kann staatlicher Zwang zumindest eine Einheitsversicherung (*Pooling equilibrium*) herstellen, die unter Wettbewerb kein Gleichgewicht darstellt und für die Individuen einer Lösung ohne Versicherung in der Regel immer noch vorzuziehen ist. Da eine solche Lösung für Individuen mit hohem Risiko unzweideutig besser, weil billiger ist als eine risikoadäquat kalkulierte Versicherung, bringt der staatliche Eingriff demnach eine echte Pareto-Verbesserung. Sonst kommt es einem *Separating equilibrium*, das nur durch Limitierung des Versicherungsschutzes für (nicht als solche erkennbare) Individuen mit geringem Krankheitsrisiko bewerkstelligt werden kann und somit zu einer suboptimalen Deckung von deren Krankheitsrisiko führt. In diesem Fall, kann eine Einheitsversicherung mit limitierter Deckung (Basissicherung) und freien, privat organisierten Zusatzversicherungsmöglichkeiten ebenfalls Pareto-verbessernd wirken: Die einheitliche Basissicherung begünstigt Individuen mit hohem Risiko durch günstigere Konditionen für einen Teil ihres Versicherungsschutzes; für Individuen mit niedrigem Risiko schafft sie einen neuen Startpunkt, ab dem Risikoselektion betrieben wird, so dass sie u. U. insgesamt einen umfassenderen Versicherungsschutz erlangen können, als wenn ihre gesamte Versicherung dem Wettbewerb überlassen bliebe.

Zu beachten ist in diesem Zusammenhang, dass das Szenario mit asymmetrischen Informationen nicht generell ein realistisches Szenario darstellt. Abgesehen davon, dass sowohl Versicherer als auch die Individuen selbst sich über ihre exakte Risikosituation im Unklaren sein können, werden private Versicherer unter Wettbewerb alle verfügbaren Informationen nutzen, die ihnen eine direkte Risikoselektion erlauben, ohne auf die Selbstselektion der Versicherten auszuweichen. Regulierungen und Begrenzungen der Risikoselektion hinsichtlich allseits verfügbarer Informationen sind in der Praxis daher durchaus verbreitet.

Anders als beim Problem adverser Selektion bleibt staatlicher Zwang gegenüber *Moral hazard*-Problemen weitgehend wirkungslos, wenn wir nicht unterstellen, dass das, was für private Versicherer unbeobachtbar ist, für staatliche Instanzen beobachtbar wird. Verpflichtungen zu beobachtbarer Vorbeugung können durch private Versicherer ebenso durchgesetzt werden wie mit Hilfe staatlichen Zwangs. Bei wirklich unbeobachtbaren Verhaltensreaktionen auf eine Versicherung können sich die Probleme im Kontext staatlich administrierter Systeme, wie bereits in der Einleitung angesprochen, sogar noch verschärfen und risikoadäquate Prämien weiter verteuern. Auch die Möglichkeiten zur Eindämmung solcher Reaktionen sind in privaten wie in öffentlichen Krankenversicherungen dieselben: Eine Limitierung des

Versicherungsschutzes erhält zumindest einen Teil des Anreizes, eigenständige Vorbeugung gegen Krankheit und ihre finanziellen Folgen zu treffen bzw. die Folgekosten einer Erkrankung in Grenzen zu halten. (Man beachte, dass nicht-finanzielle Krankheitsfolgen, soweit sie nicht durch zusätzliche Versicherungsleistungen im Sinne einer „Entschädigung“ kompensiert werden, ohnedies eine Art Selbstbehalt der Individuen mit sich bringen.)

Eine Begrenzung des Versicherungsschutzes, wie sie zur Abwehr von *Moral hazard* oder im Falle adverser Selektion zur Abschreckung hoher Risiken von Verträgen, die für niedrige Risiken kalkuliert sind, erforderlich ist, lässt sich auf verschiedene Weise bewerkstelligen. In der Praxis üblich sind

- Beschränkungen hinsichtlich der Art der in Anspruch genommenen *Leistungen* (global nach Leistungskategorien – ambulant, stationär, zahnärztlich *etc.* – oder durch selektive Ausschlüsse bestimmter Behandlungsarten, -methoden, Arzneimittel usw.),
- Beschränkungen hinsichtlich der Art und Person der *Leistungserbringer* (zugelassene Ärzte, im Einzelnen bestimmte Vertragsärzte; Ausschluss bestimmter Berufsgruppen im Bereich medizinischer Berufe *etc.*)
- sowie Beschränkungen hinsichtlich der Höhe der Erstattung der *Kosten* im Krankheitsfall, wobei diese Beschränkungen
  - an der *Menge* bestimmter Leistungen pro Periode (Sehhilfen, psycho- oder physiotherapeutische Behandlungen),
  - am *Preis* (Gebührenordnung, Festpreise, absolut definierte „Rezeptgebühren“ pro Behandlungsfall, absolut oder proportional definierte Zuzahlungen zu einzelnen Leistungen)
  - oder – im allgemeinsten Fall – an den *Ausgaben* in Folge einer Erkrankung bzw. Behandlung

ansetzen können; Beschränkungen der zuletzt genannten Art können die Form einer proportionalen Selbstbeteiligung, einer festen Obergrenze oder – als Schutz gegen extreme Risiken – umgekehrt eines absolut definierten, festen Selbstbehalts pro Periode annehmen; *grosso modo* den selben Effekt wie ein fester Selbstbehalt haben Beitragserstattungen für leistungsfreie Perioden; schließlich gibt es, als Grenzfall einer Beschränkung der Kostenerstattung, noch den selten genutzten Fall eines „Indemnitätstarifs“, bei dem der Versicherte eine diagnosebezogene Pauschalleistung erhält, über deren Verwendung er (selbst ohne Inanspruchnahme einer Behandlung) selbst frei entscheiden kann, so dass die marginalen Anreize zur sparsamen Behandlung schon für den ersten € Behandlungskosten gelten.

In der Mehrzahl der Industrieländer spielen staatlich organisierte Versicherungssysteme eine wichtige Rolle für die Absicherung der Bürger gegen

Krankheitsrisiken – zumindest im Sinne einer verpflichtenden Absicherung auf einem bestimmten Basisniveau (vgl. Tabelle 6.1). Die wichtigste Ausnahme bilden die USA, wo staatliche Systeme nur für Bedürftige und Rentner bestehen, während die Mehrzahl der Bürger durch freiwillige, private Versicherungen abgesichert ist, die in der Regel vom Arbeitgeber bezahlt werden.

Finanziert werden die staatlichen Systeme in ihrer Mehrzahl vorrangig durch Beiträge, die nicht risikoorientiert kalkuliert werden, sondern proportional zum Erwerbseinkommen sind. Hinzu treten häufig (allerdings z. B. nicht in Deutschland) steuerfinanzierte Zuschüsse; in einigen Staaten (Dänemark, Finnland, Großbritannien und Irland) ist die staatliche Krankenversicherung sogar weitgehend oder ganz steuerfinanziert, basiert also vorrangig auf mehr oder weniger progressiven, einkommensbezogenen Abgaben. Neben einem *Pooling* alters- und anamneseabhängiger Risikounterschiede bewirken diese Formen der Finanzierung daher auch noch eine (lohn-) einkommensbezogene Umverteilung, die durch die (Mit-)Versicherung von Individuen ohne eigenes (Erwerbs-)Einkommen auch noch auf weitere Dimensionen ausgeweitet wird.

Systeme, in denen freiwillig private (selten: staatliche) Zusatzversicherungen geschlossen werden können und Systeme, in denen private Versicherungen vorrangig alternativ – auf der Grundlage bestimmter Ausnahmen von der Versicherungspflicht – geschlossen werden können, während die staatlichen Systeme im Prinzip als Vollversicherungen angelegt sind, halten sich unter den größeren OECD-Ländern annähernd die Waage. In dieser Hinsicht ist der Umgang mit möglichen (adversen) Selektionsproblemen demnach recht unterschiedlich. Allerdings gibt es in einer ganzen Reihe von Ländern auch in den staatlichen Systemen Zuzahlungen von Patienten, die *Moral hazard*-Verhalten der Versicherten und – aufgrund der dadurch beeinflussten Patienteninteressen – auch der Leistungserbringer kontrollieren sollen (vgl. Tabelle 6.2).

Tabelle 6.1: Krankenversicherungssysteme in ausgewählten OECD-Ländern

Belgien	Staatliche Pflichtversicherung <i>plus</i> freiwillige (staatliche oder private) Zusatzversicherung
Dänemark	Steuerfinanzierte, staatliche Volksversicherung <i>plus</i> freiwillige Zusatzversicherung
Deutschland	Staatliche Pflichtversicherung ( <i>plus</i> freiwillige Zusatzversicherung) <i>oder</i> private Versicherung

Finland	Steuerfinanzierter, staatlicher Gesundheitsdienst; staatliche Pflichtversicherung <i>oder</i> private Versicherung
Frankreich	Staatliche Pflichtversicherung <i>und</i> freiwillige Zusatzversicherung
Griechenland	Staatliche Pflichtversicherung <i>oder</i> private Versicherung
Großbritannien	Steuerfinanzierter, staatlicher Gesundheitsdienst <i>und</i> freiwillige Zusatzversicherung
Irland	Steuerfinanzierter, staatlicher Gesundheitsdienst <i>und</i> freiwillige (staatliche) Zusatzversicherung
Italien	Staatliche Pflichtversicherung (Gesundheitsdienst) <i>oder</i> private Versicherung
Niederlande	Staatliche Pflichtversicherung <i>oder</i> private Versicherung
Österreich	Staatliche Pflichtversicherung <i>und</i> freiwillige Zusatzversicherung
Schweden	Staatliche Pflichtversicherung (Gesundheitsdienst)
Spanien	Staatliche Pflichtversicherung <i>oder</i> private Versicherung
USA	Freiwillige, private Versicherung; staatliche Hilfsprogramme für Bedürftige und Rentner
Japan	Staatliche Pflichtversicherung

Quellen: WHO Europe (2002).

Tabelle 6.2: Patienten-Zuzahlungen in ausgewählten OECD-Ländern

Belgien	Zuzahlungen von 10%–30% der Kosten
Dänemark	—
Deutschland	— (2004: 10 € pro Quartal für Arztbesuche)
Finnland	8 € für drei Arztbesuche bzw. 16 € Jahresgebühr
Frankreich	Zuzahlung von 25% außer bei schweren Krankheiten
Griechenland	—
Großbritannien	—
Irland	—
Italien	—
Niederlande	Zuzahlung von 20%, maximal 92 € pro Jahr; Befreiung für niedrige Einkommen
Österreich	Zuzahlung von 10%–20% für 20% der Bevölkerung; 20% Zuzahlung für Bauern und Selbständige
Schweden	7–10 € für Allgemeinarztbesuch, 17–31 € für Spezialistenbesuch, 11–34 € für Notarzt
Schweiz	Zuzahlung von 10%
USA	Zuzahlung von 20%–50% für bestimmte Leistungen
Japan	Zuzahlung von 30%, maximal 1.154 ¥

Quellen: WHO Europe (2002).

## 7. Übergeordnete Organisationen

Literatur: Breyer, Zweifel und Kifmann (2005, Kap. 11 und 7)  
Dranove und Satterthwaite (2000)  
Van de Ven und Ellis (2000)

Bisher haben wir uns vor allem mit einzelnen Agenten im Gesundheitssystem und ihren paarweisen Interaktionen – etwa in den Beziehungen Arzt–Patient oder Versicherter–Versicherung – befasst. Nun wollen wir uns übergreifenden systemischen Aspekten und Fragen der Organisation des Gesundheitssystems als Ganzem zuwenden. In Abschnitt 7.1 werden verschiedene Grundformen der Organisation dieses Systems behandelt, in denen die einzelnen internen Beziehungen jeweils in charakteristischer Weise geordnet werden. In Abschnitt 7.2 werden einige Aspekte der Organisation des Krankenversicherungswesens vertieft diskutiert, die für das gesamte Gesundheitssystem zumindest potentiell von großer Bedeutung sind und über die eher mikroökonomischen Probleme, die in Kapitel 6 behandelt wurden, hinaus gehen.

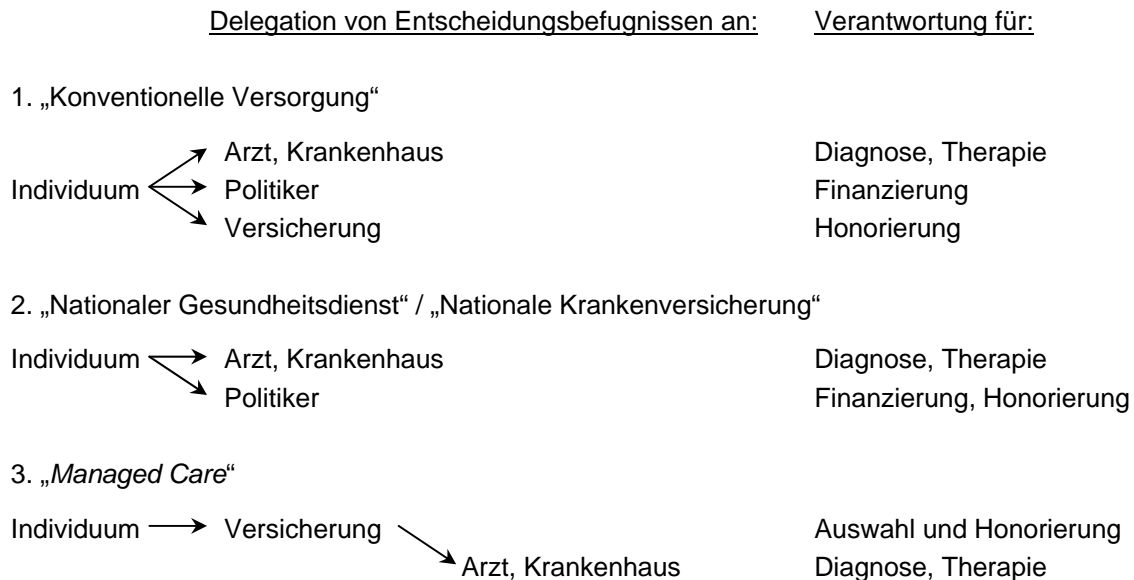
### 7.1 Sachwalterbeziehungen im Gesundheitssystem

Im Mittelpunkt aller Gesundheitssysteme steht die Beziehung des Patienten zu seinem Arzt bzw. allgemein zu Erbringern medizinischer Dienstleistungen. Da ein Individuum, das solche Dienstleistungen nachfragt, in der Regel zu geringe Kenntnisse besitzt, um die Notwendigkeit und Güte einzelner Dienstleistungen beurteilen zu können, hat diese Beziehung den Charakter einer Sachwalterbeziehung: Das Individuum überträgt dem Arzt letztlich weitgehend die Verantwortung für Diagnose und (Auswahl der) Therapie. Da es innerhalb dieser Beziehung jedoch auch Interessenkonflikte geben kann – insbesondere durch das Einkommensinteresse des Arztes (vgl. Kapitel 3) – treten im Rahmen des Gesundheitssystems neben diese zentrale Beziehung andere ergänzende Sachwalterbeziehungen. Welche ergänzenden Sachwalter dabei auftreten und welche Verantwortung sie im Einzelnen übernehmen erklärt einen Großteil der Unterschiede zwischen den Formen der Organisation des Gesundheitssystems, die sich im internationalen Vergleich beobachten lassen.

### 7.1.1 Grundformen nationaler Gesundheitssysteme

Bei grober Typisierung lassen sich anhand der jeweils etablierten Sachwalterbeziehungen drei Grundformen der Organisation des Gesundheitssystems unterscheiden (vgl. auch Abb. 7.1).

Abb. 7.1: Sachwalterbeziehungen als Organisationsmerkmal



Bei der „konventionellen Versorgung“, zu beobachten etwa in Deutschland, der Schweiz, den Niederlanden, Schweden und den USA, stehen alle anderen Agenten des Gesundheitssystems in einer Sachwalterbeziehung zum Patienten: Den Ärzten und Krankenhäusern überträgt das Individuum als Patient die Verantwortung für die Auswahl und zumeist auch für die Erbringung der Gesundheitsdienstleistungen; als Bürger wählt es die Politiker, die für die Finanzierung der Krankenhäuser (und für zahlreiche Regulierungen des Systems) zuständig sind; als Versicherter schließt es einen Vertrag mit einer Krankenversicherung, die wiederum die Honorierungsformen und Tarife mit Ärzten und Krankenhäusern aushandelt. Während die Politiker ex definitione öffentliche Agenten sind, können Leistungserbringer und Versicherungen vorrangig in öffentlicher Trägerschaft stehen oder privatwirtschaftlich verfasst sein, was jeweils gewisse Unterschiede in ihrem Verhalten bedingt und daher innerhalb des Spektrums „konventioneller Versorgungssysteme“ für eine weitere Differenzierung sorgt.

In Großbritannien wird die Bereitstellung von Gesundheitsdienstleistungen über einen „nationalen Gesundheitsdienst“ (*National Health Service*, NHS) organisiert, in Kanada über eine staatliche „nationale Krankenversicherung“. Im ersten Fall gibt es formell keine Versicherungen, sondern die medizinische Versorgung wird über

Steuern finanziert, der NHS übernimmt die Versicherungsfunktion implizit mit und die Ärzte sind im Wesentlichen öffentliche Bedienstete. Im zweiten Fall gibt es eine unmittelbar staatlich administrierte Einheitsversicherung, während Ärzte und andere Leistungserbringer etwas größere Freiheit genießen. In beiden Fällen gibt es jedenfalls neben der Beziehung des Individuums zu den Leistungserbringern nur Politiker als ergänzende Sachwalter. Da die Ärzte in Kanada frei sind, private Rechnungen zu stellen, während die nationale Krankenversicherung den Patienten jeweils nur vergleichsweise niedrige Festbeträge erstattet, hat sich hier allerdings effektiv ein System der Kostenbeteiligung bei freier Arztwahl etabliert, das einige der Probleme des NHS hinsichtlich Umfang und Qualität der Versorgung vermeidet.

In den USA findet sich schließlich zumindest als Sub-Sektor des Gesundheitssystems mit nennenswertem „Marktanteil“ noch ein grundlegend anderes System, das als „*Managed care*“ bezeichnet wird. In diesem Fall nehmen (in der Regel private) Krankenversicherer als vorrangige Sachwalter des Individuums eine Schlüsselrolle für die Organisation der gesamten Versorgung ein, indem sie ihre Mitglieder an von ihnen ausgewählte Ärzte vermitteln, welche sie gegebenenfalls an ein in der Regel ebenfalls vertraglich an die Versicherung gebundenes Krankenhaus weiterleiten. Ansätze zu *Managed care* für einzelne Krankheitsbilder oder Leistungsarten sind mittlerweile auch im Rahmen anderer Gesundheitssysteme entstanden. *Managed care* in Reinform betreiben die sogenannten „*Health Maintenance Organisations*“, auf deren Eigenarten hier abschließend noch kurz eingegangen wird.

### 7.1.2 Eigenarten der wichtigsten Sachwalterbeziehungen

Zentral ist in Gesundheitssystemen aller Art letztlich die Sachwalterbeziehung zwischen Arzt und Patient. Ihr gilt daher auch hier zuerst die Aufmerksamkeit.

**Ärzte:** Über die Arzt-Patient-Beziehung wurde schon in vorangegangenen Kapiteln (v. a. Kapitel 3, 4 und 5) Einiges gesagt. An dieser Stelle mag es genügen zu wiederholen, dass ein Patient seinem Arzt in der Regel vertrauen muss, weil dieser vom Patienten nicht zu überprüfende Fachkenntnisse hinsichtlich der Diagnose und Therapie von Krankheiten hat. Es gibt sogar Anzeichen dafür (beispielsweise die Wirksamkeit von Placebos), dass das Vertrauen als solches wichtige Voraussetzung des Heilerfolgs ist. Trotzdem gibt es mindestens zwei Komplexe von Gründen dafür, die Sachwalterbeziehung zum Arzt durch weitere Sachwalterbeziehungen zu ergänzen. Erstens sind die Qualifikation des Arztes und die Qualität der von ihm angebotenen Leistungen einer Überprüfung durch Dritte, die selbst Fachleute sind, zugänglich. Zweitens berücksichtigt der Arzt bei der Behandlung neben den gesundheitlichen Belangen des Patienten nicht ohne Weiteres auch seine

finanziellen Interessen. Vielmehr kann hier ein Interessenkonflikt auftreten, bei dessen Lösung der Patient mangels medizinischer Fachkenntnisse generell, oft aber auch wegen der Dringlichkeit seines Gesundheitsproblems, in einer schwachen Verhandlungsposition ist. Qualitätssicherung und Kostenkontrolle sind daher Aspekte, im Hinblick auf die das Individuum die Hilfe ergänzender Sachwalter in Anspruch nehmen kann.

Im Folgenden wird eine Reihe möglicher weiterer Sachwalter angesprochen, deren Rolle in real existierenden Gesundheitssystemen mehr oder weniger stark ausgeprägt sein kann.

**Arbeitgeber:** Arbeitgeber haben in der Regel große Erfahrungen bei der Gestaltung von Sachwalterbeziehungen, treten sie doch selber als Auftraggeber ihrer Arbeitnehmer auf, die Sachwalter ihrer Interessen sind. Hier haben sie ähnliche Probleme zu lösen, wie die Patienten in Bezug auf ihre Ärzte: Sie können die Arbeitnehmeraktivitäten nur unzureichend überwachen, da das Betriebsergebnis auch von exogenen und zufälligen Einflüssen (Wetter, Ausfall von Konkurrenten, staatlichen Maßnahmen) abhängt. Ihre Kenntnis im Umgang mit diesen Problemen könnten sie nun als Sachwalter der bei ihnen beschäftigten Patienten nutzbringend einsetzen, zumal sie ein eigenständiges Interesse am Gesundheitszustand ihrer Arbeitnehmer haben oder diesen dadurch einen lohn-analogen Anreiz schaffen können.

Ein wichtiges Problem, das verhindert, dass diese potentielle Sachwalterrolle öfter als in Ausnahme- und Spezialfällen genutzt wird (in Abbildung 7.1 taucht sie z. B. nirgends auf), ist jedoch, dass die Arbeitgeber die Informationen, die sie in einer solchen Rolle erhalten, auch gegen ihre Arbeitnehmer einsetzen könnten. So können Arbeitgeber die Grenzproduktivität ihrer Arbeitnehmer als Sachwalter für deren Gesundheit besser einschätzen und sich diesen Informationsvorteil bei Lohnverhandlungen zu nutze machen. Zumeist beschränkt sich die Rolle von Arbeitgebern daher darauf, für ihre Arbeitnehmer und deren Angehörige Verträge mit Krankenversicherern abzuschließen, soweit dies nicht bereits anderweitig – etwa durch gesetzliche Zwangsmitgliedschaften – geregelt ist, und diesen die weitere Sachwalterrolle zu überlassen.

**Staat:** Der Staat mit seiner Möglichkeit, markt- und verhandlungsgesteuerte Allokationsmechanismen mit Hilfe seiner hoheitlichen Zwangsbefugnis zu modifizieren oder zu ersetzen, kann in verschiedener Weise als Sachwalter der Patienten auftreten. Er kann die Versorgung mit medizinischen Leistungen selbst organisieren, indem er etwa die Finanzierung der Krankenhäuser sicherstellt oder das gesamte Gesundheitssystem über Steuern finanziert (wie in Großbritannien). Er

kann als monopolistischer Krankenversicherer auftreten wie in Kanada. Er kann sich jedoch auch auf – im Einzelfall u. U. sehr weit gehende – Regulierungen von Versorgung und Versicherungswesen beschränken, die das Verhalten von Leistungserbringern und formell mehr oder weniger selbständigen Versicherungsunternehmen begrenzen. Völlig unregulierte Gesundheitssysteme sind in den entwickelten Industrieländern nicht zu finden, so dass, abgesehen von extremen Lösungen wie in den Fällen Großbritanniens und Kanadas klare Grenzziehungen hier nicht immer möglich sind. Innerhalb der „konventionellen Versorgungssysteme“ z. B. kann der effektive „Public-private mix“ in den Sachwalterbeziehungen daher immer noch stark variieren.

Neben der Möglichkeit, Verträge im Gesundheitswesen durch gesetzliche Beschränkungen fast beliebig zu beeinflussen, kann der Staat als Anbieter von Gesundheitsleitungen und Versicherungen auch erhebliche Marktmacht gewinnen, durch die er die relevanten Preise direkt kontrolliert. In jedem Fall wird bei einer starken Rolle des Staates Wettbewerb innerhalb des Gesundheitssystems auf einen allgemeineren Wettbewerb um Wählerstimmen verlagert, so dass die Wettbewerbsintensität in Einzelfeldern wie dem Gesundheitswesen sinkt. Ferner kommen dabei Eigeninteressen von Politikern und Administrationen ins Spiel. So lassen sich im Gesundheitssystem relativ leicht Umverteilungen zugunsten bestimmter Wählergruppen bewirken, die von der zufallsgesteuerten Umverteilung zwischen Menschen mit verschiedenen Gesundheitsrisiken oder allseits intendierten Umverteilungen zwischen Beziehern verschieden hoher Einkommen nur schwer zu unterscheiden sind. Wenn Politiker in einer starken Rolle sich letztlich einseitig auf einzelne der für das Gesundheitssystem relevanten Ziele – z. B. Kostendämpfung, wegen der leichter oder kurzfristiger beobachtbaren Erfolge – konzentrieren, kann dies auf Kosten der Versorgungssicherheit oder Qualitätssicherung gehen, etwa weil das schon Markteintrittsentscheidungen potentieller Leistungsanbieter zurückwirken kann.

**Versicherer:** Insbesondere private Krankenversicherer, die selbst im Wettbewerb um Kunden stehen, können eine Rolle als ergänzende Sachwalter übernehmen, die Probleme staatlicher Instanzen als Sachwalter vermeiden. Allerdings gilt es hier, eine sinnvolle Balance zu finden zwischen Nutzung des Wettbewerbs als effizienzsteigerndem Allokationsmechanismus und möglichen Fehlfunktionen des Wettbewerbs, der für Versicherer zugleich einen starken Anreiz zur Risikoselektion auslöst (vgl. dazu die Abschnitte 6.4 und 7.2). Als starker Sachwalter können Versicherungen nur auftreten, wenn ihre Aktionsparameter eine direkte Beeinflussung des Verhaltens von Leistungserbringern und Patienten erlauben, z. B. durch die Möglichkeit, Vergütungssysteme und Honorare (mit einzelnen

Leistungserbringern) direkt auszuhandeln und ihre Versicherungsverträge mit gewisser Freiheit zu gestalten.

Eine in den USA verbreitete Form der Sachwalterbeziehung, in der Versicherer eine Schlüsselstellung einnehmen und konsequent *Managed care* betreiben, ist die *Health Maintenance Organization* (HMO). Der Versicherte zahlt dort eine fixe Prämie an die HMO. Im Gegenzug wird ihm im Krankheitsfall eine vom HMO-Arzt festgelegte Behandlung garantiert, an deren Kosten er sich nicht weiter beteiligen muss. Patienten müssen sich von einem der HMO angeschlossenen Ärzte behandeln und von diesen ggfs. an ausgewählte Krankenhäuser *etc.* weiter verweisen lassen, verzichten also auf die freie Arztwahl. Der HMO-Arzt wird nicht nach im Einzelfall erbrachten Leistungen bezahlt, sondern erhält ein festes Budget für die durch ihn und weitere Leistungserbringer erbrachte Behandlung der ihm zugewiesenen Patienten und ist überdies am finanziellen Erfolg der HMO beteiligt. Er trägt insofern hohe Kostenverantwortung.

Je nach Typ der HMO kann diese selbst Träger von Gruppenpraxen sein und die Ärzte im Angestelltenverhältnis (mit erfolgsabhängiger Einkommenskomponente) beschäftigen, ein „Netzwerk“ selbständiger Ärzte auf der Basis von Behandlungsverträgen unterhalten, die auch weiterhin konventionell versicherte Patienten behandeln, oder ihre Verträge mit lokalen Ärzteverbänden schließen und den Versicherten etwas größere Freiheit bei der Bestimmung ihres Hausarztes belassen, der die primäre Versorgung übernimmt und auch nach Überweisung in ein Krankenhaus seine Zuständigkeit behält. In jedem Fall stellt Wettbewerb um die Versicherten sicher, dass die Einsparungen durch HMOs in nennenswertem Umfang an die Versicherten weitergegeben werden. Allerdings gibt es Anhaltspunkte dafür, dass ein Teil der Kostensenkungen durch HMOs, die sich in den USA auf bis zu 30 % belaufen, letztlich auf Kostenverschiebungen – hin zu Patienten mit anderer, konventioneller Versicherung – basieren. In Ländern, in denen die Leistungsanbieter keine Vollkostenkalkulation durchführen müssen, etwa weil wie in Deutschland die Investitionskosten von Krankenhäusern öffentlich finanziert sind, dürften die Einsparpotentiale durch *Managed care* im Übrigen geringer ausfallen als im US-System.

## 7.2 Organisation des Krankenversicherungssystems

Die Herausbildung von Formen des *Managed care*, die sich letztlich nicht auf die USA beschränkt, verdeutlicht, dass Krankenversicherungen eine Schlüsselrolle bei der Organisation des Gesundheitswesens zukommen kann. Die zentrale Herausforderung liegt dabei darin, Wettbewerb unter den Versicherern zu installieren und – dies muss gleich hinzugesetzt werden – sinnvoll zu kanalisieren und von dort

aus kompetitives Verhalten auch in den Bereich der medizinischen Versorgung zu tragen.

Ein typisierender Vergleich existierender Gesundheitssysteme legt nahe, dass stärker durch Wettbewerb (unter Krankenversicherern) gesteuerte Gesundheitssysteme Vorteile im Hinblick auf eine im Umfang ausreichende, auch qualitativ hochstehende Versorgung haben, die mit dem medizin-technischen Fortschritt Schritt hält und differierenden Präferenzen der Individuen sowohl hinsichtlich der Behandlung als auch hinsichtlich Art und Umfang des Versicherungsschutzes besser entspricht. Nachteile weisen solche Systeme offenbar im Hinblick auf die generell höhere Kostendynamik, hohe Transaktionskosten (Suche nach geeigneten Versicherungen und Behandlungsarten, Vertragsabschlusskosten, Abrechnungskosten) auf. Ferner veranlasst Wettbewerb um Kunden die Versicherer zu Risikoselektion (aufgrund beobachtbarer Merkmale, im Unterschied zur adversen Selektion, vgl. Kapitel 6). Die Vor- und Nachteile stärker administrativ gesteuerter Systeme sind spiegelbildlich: Der leichteren Vermeidung von Risikoselektion, der besseren Implementierbarkeit von Maßnahmen zur generellen Kostendämpfung (z. B. durch eine feste Budgetierung) und geringeren Transaktionskosten stehen Einbußen bei Versorgungsumfang, -qualität sowie starre, stark vereinheitlichte Angebote hinsichtlich Versicherungsschutz und medizinischen Leistungen gegenüber.

Die Organisation des Krankenversicherungswesens in Deutschland basiert hinsichtlich der Alternative Wettbewerb vs. staatliche Administration in mehrfacher Hinsicht auf einem intermediären Design. Dominiert wird der Sektor durch staatliche Krankenkassen, die untereinander jedoch zumindest im Hinblick auf ein sehr enges Aktionsfeld im Wettbewerb stehen. Daneben gibt es auch private Krankenversicherer, die auf der Basis eines einkommensabhängig eingeräumten Wahlrechts nicht nur Zusatzversicherungen, sondern auch einen Vollversicherungsschutz anbieten können. Eine zusätzliche Mischform stellt das System für Beamte dar, die für 50 % ihrer Krankheitskosten (bei Angehörigen und Pensionären: bis zu 70 %) einen Beihilfeanspruch gegen ihren Arbeitgeber, den Staat, haben, den sie in der Regel durch eine private Teilkostenversicherung ergänzen. Schon der Wettbewerb innerhalb des Systems der gesetzlichen Krankenversicherungen (GKV), der in seiner heutigen Form erst durch das Gesundheitsstrukturgesetz von 1993 errichtet wurde, schafft Selektionsanreize und macht Maßnahmen zu seiner Eindämmung erforderlich.

Risikoselektion, d. h. die Auswahl möglichst guter Risiken durch die Versicherer, schafft in Gestalt niedrigerer Prämien und höherer Marktanteile Vorteile für einzelne Versicherungen und ihre Mitglieder und bewirkt damit Verteilungseffekte, bewirkt

aber letztlich keine Effizienzsteigerungen. Wären Krankenkassen in der Gestaltung ihrer Tarife vollkommen frei, so dass sie eine hinreichend genaue Erfahrungstarifizierung aufgrund beobachtbarer Merkmale ihrer (potentiellen) Kunden vornehmen könnten, gäbe es zwar keinen Anreiz zur Risikoselektion im Sinne einer Abweisung schlechter Risiken; diese müssten jedoch höhere – u. U. prohibitiv hohe – Prämien bezahlen. Ein Diskriminierungsverbot, das derart differenzierte Prämienbemessung ausschließt, lässt den Anreiz zur Risikoselektion im eigentlichen Sinn entstehen. Auch ein Kontrahierungszwang, der die Versicherer verpflichtet, alle potentiellen Kunden aufzunehmen, löst das Problem nicht, solange die Versicherer die Möglichkeit behalten, indirekt zu selektieren, indem sie den Umfang (hohe Selbstbehalte) und die Struktur des angebotenen Versicherungsschutzes (Ausschluss kostenträchtiger Behandlungen oder bestimmter Leistungen, die auf generell erhöhte Krankheitskostenrisiken schließen lassen) frei gestalten können. Selbst wenn die Versicherungsverträge daher zusätzlich gewissen (Mindest-)Standards genügen müssen, verbleiben den Versicherern in der Praxis jedoch weitere Wege zur versteckten, direkten oder indirekten Risikoselektion – verzögerte Bearbeitung von Aufnahmeanträgen, selektiv verschlechterter Service, etc.

Innerhalb eines Systems wie dem der deutschen GKV, das auf einer weitreichenden Versicherungspflicht beruht und auch allen weiteren, eben angesprochenen Einschränkungen der Vertragsfreiheit unterliegt, ist Risikoselektion im Wesentlichen ein reines Nullsummen-Spiel, das den dort zugleich verfolgten Verteilungszielen zuwiderläuft. Zentraler Parameter des Wettbewerbs um Mitglieder ist dort der jeweilige Beitragssatz, der von der Ausgaben-Einnahmen-Relation jeder Kasse bestimmt wird. Angesichts der rechtlichen Rahmenbedingungen des GKV-Systems ist diese insbesondere bestimmt durch folgende Merkmale: (1) Altersstruktur, (2) Geschlechtsstruktur, (3) Haushaltsstruktur, (4) durchschnittliche versicherungspflichtige Bruttoentgelte sowie (5) effektive Morbiditätsstruktur der Versicherten. Für den Bestand an Mitgliedern bei Einrichtung von Wettbewerb („Kassenwahlfreiheit“) sind die Ausprägungen dieser Merkmale für jede Kasse zunächst gegeben, und ein existierendes Versichertenkollektiv kann „Pech“ haben, weil es eine ungünstige Struktur aufweist, die mobile Mitglieder vertreibt und Neueintritte günstiger Risiken verhindert. Ausgehend von diesen Startbedingungen sind bei der Anwerbung neuer Mitglieder junge, männliche Versicherte mit geringer Zahl mitzuversichernder Angehöriger, hohem Einkommen und niedrigem Gesundheitsrisiko besonders attraktiv, während ungünstige Risiken tendenziell auf relativ teure „*Fall-back-Optionen*“ (z. B. die AOKs) verwiesen werden.

Um Risikoselektion einzudämmen und die damit verbundenen Verteilungseffekte zu vermeiden, kann ein finanzieller Ausgleich zwischen den untereinander im Wettbewerb stehenden Kassen eingerichtet werden, der zugleich, gezielt oder

ungezielt, den Wettbewerb auf bestimmte Felder lenkt. Ein reiner (Ex-post-)Ausgabenausgleich vernichtet dabei allerdings alle Wirtschaftlichkeitsanreize aus der Sicht jeder einzelnen Kasse, die sich – auf Kosten aller Versicherten des Gesamtsystems – vorrangig nur noch um möglichst große Marktanteile bemühen wird. (In Deutschland wird ein solcher Ausgabenausgleich im Bereich der gesetzlichen Pflegeversicherung praktiziert, dort allerdings vor dem Hintergrund eines fast perfekt standardisierten Leistungspakets und einer Einstufung als Leistungsberechtigter mit verschieden hohem Anspruch, der kaum diskretionäre Spielräume zulässt.) Der in der deutschen GKV praktizierte „Risikostrukturausgleich“ soll hingegen, alle kassenspezifischen Unterschiede von Einnahmen und Ausgaben ausgleichen, die sich aus den vergleichsweise leicht beobachtbaren Merkmalen (1) bis (4) der obigen Aufzählung ergeben. Ziel ist, den Kassenwettbewerb auf das Bemühen um möglichst geringe Verwaltungs- oder – allgemeiner ausgedrückt – Transaktionskosten einzuengen. Vernachlässigt werden dabei bislang morbiditätsbedingte Ausgabenunterschiede, einschließlich der Tatsache, dass diese auch auf die jeweils anfallenden Verwaltungskosten zurückwirken können. Mit dem derzeit vorbereiteten Übergang zu einem „morbiditätsorientierten Risikostrukturausgleich“ soll dieses Problem – auf der Basis differenzierter Kennziffernsysteme für die erwarteten Kosten einer sehr großen Zahl einzeln erfasster Krankheiten – gelöst werden. Die enormen Schwierigkeiten, zutreffende Prognosen für die gesundheitliche Entwicklung jedes Versichertenkollektivs und für die Entwicklung der (relativen) Kosten der Behandlung jedes möglichen Krankheitsbildes aufzustellen, lassen dabei jedoch eine nennenswerte Gefahr der Überregulierung und neuerlicher Fehlsteuerungen entstehen. Selbst wenn dadurch dem Ziel eines fairen Wettbewerbs um niedrige Verwaltungskosten näher gekommen wird, erhebt sich allerdings die Frage, ob dieses Wettbewerbsfeld nicht zu eng ist. Alternativen wären eine Erweiterung des Wettbewerbs auf differenziertere Vertragsangebote mit stärkeren Wahlmöglichkeiten der Versicherten hinsichtlich bestimmter Preis-Leistungs-Pakete und/oder mit mehr Spielräumen für die Versicherer hinsichtlich vielversprechender Prozessinnovationen (etwa mehr *Managed care* auf der Basis direkter Verträge mit einzelnen Leistungserbringern anstelle der heutigen, stark kollektivierten Vertragsbeziehungen) – wobei die Möglichkeit zur Risikoselektion allerdings zwangsläufig wieder größer wird.

## **Kapitel 8: Politische Ökonomie**

### **Literatur:**

Breyer, Zweifel, Kifmann (2005, Kapitel 13)

Fenge (2003)

### **8.1 Motivation**

Während in den vorangegangenen Kapiteln eine wohlfahrtsökonomische Perspektive für die Analyse gewählt worden war, soll jetzt eine gänzliche neue Sichtweise in den Vordergrund gestellt werden. Mit Hilfe des theoretischen Instrumentariums der Politischen Ökonomie geht es im Folgenden um die Frage, wie man das Zustandekommen und die Weiterexistenz der tatsächlich vorhandenen Institutionen im Gesundheitswesen erklären kann. Dabei steht zunächst der Wähler in einer direkten Demokratie im Mittelpunkt. Danach werden konkrete Strukturen politischer Systeme, insbesondere die Existenz von Interessensverbänden, berücksichtigt.

### **8.2 Direkte Demokratie**

#### **8.2.1 Existenz und Ausmaß eines kollektiv finanzierten System der Gesundheitsvorsorge**

In den letzten 100-120 Jahren sind in vielen Industrieländern kollektiv finanzierte Gesundheitssysteme eingerichtet worden, die zu einer gleichmäßigeren Verteilung medizinischer Leistungen auf die Gesellschaftsmitglieder als der Konsum anderer Güter oder Dienstleistungen geführt haben. Dabei erfolgt die kollektive Finanzierung in manchen Ländern im Rahmen eines Nationalen Gesundheitsdienstes durch allgemeine Steuern, z.B. in Großbritannien. In anderen Ländern dagegen findet man soziale Krankenversicherungen mit Zwangsmitgliedschaft, die über (Versicherungs-) Beiträge finanziert werden. Mögliche Formen sind einheitliche Pro-Kopf Prämien wie in der Schweiz oder einkommensproportionale Beiträge wie in Deutschland.

Das folgende Modell dient dazu, die wesentlichen Faktoren zu untersuchen, die dafür verantwortlich sind,

- 1) ob es in einer direkten Demokratie mit Mehrheitsentscheid überhaupt zur Gründung bzw. Aufrechterhaltung einer sozialen Krankenversicherung kommt
- 2) wovon der Umfang des Versicherungsschutzes abhängt und
- 3) ob die Versicherungsbeiträge als einheitliche Pro-Kopf Prämien oder einkommensproportionale Beiträge erhoben werden.

Bei Frage 3) soll es sich um eine übergeordnete Regime-Entscheidung handeln, die hinter dem Schleier des Nichtwissens bezüglich des individuellen Krankheitsrisikos

(„Veil of Ignorance“) getroffen wird und nur einstimmig geändert werden kann. Bei Frage 2) dagegen liegt eine Entscheidung innerhalb eines bestimmten Regimes vor. Sie wird bei voller Kenntnis der persönlichen Eigenschaften mit Mehrheitsentscheid getroffen.

Für das Modell wird angenommen, dass die Finanzierung der Krankenversicherung über einkommensproportionale Beiträge erfolgt und dass privater Zukauf von medizinischen Leistungen möglich ist.

Die Wähler können sich in zwei Charakteristika unterscheiden, die beide exogen und beobachtbar sind und stochastisch nicht voneinander abhängen: das Einkommen und das Krankheitsrisiko. Jeder Parameter kann genau zwei Werte annehmen. Ein bestimmter Anteil der Bevölkerung ist reich, während der Rest arm ist; einige Individuen haben eine hohe Krankheitswahrscheinlichkeit und andere eine niedrige. Es wird dabei unterstellt, dass es mehr Arme als Reiche in der Bevölkerung gibt und dass die Mehrheit Niedrigrisikotypen sind. Die Individuen maximieren den Erwartungsnutzen, der sich additiv aus dem Konsum von (staatlich bereitgestellter und privat zugekaufter) Krankenbehandlung und allgemeinen Konsumgütern zusammensetzt. Der Nutzen aus Krankenbehandlung wieder dabei nur im Zustand der Krankheit relevant.

Die Entscheidungen werden in folgender Reihenfolge getroffen:

1) Mehrheitsentscheidung über das Regime, d.h. wonach sich bei einer etwaigen Kollektivversicherung die Beiträge richten.

Im Anschluss daran erfahren die Individuen ihren Risikotyps.

2) Mehrheitsentscheidung über den Umfang der staatlichen Bereitstellung

3) Individuelle Entscheidung über den privaten Zukauf

Die Entscheidungen werden in umgekehrter Reihenfolge analysiert („Backward induction“), da im politischen Prozess die Entscheidungen der niedrigeren Ebenen jeweils antizipiert werden.

ad 3) Entscheidung über privaten Zukaufen

Individuelles Optimieren der Individuen führt dazu, dass so viel Versicherungsschutz gekauft wird, bis der Grenznutzen des allgemeinen Konsums mit dem Grenznutzen aus Gesundheitsleistungen insgesamt übereinstimmt. Diese Bedingung gilt also sowohl bei Abwesenheit einer kollektiv finanzierten Versicherung wie auch für den Fall mit staatlicher Krankenversicherung. Die gewünschte Versicherungsdeckung ist dabei bei reichen Individuen und bei niedrigen Risiken höher als im jeweils anderen Fall.

ad 2) Entscheidung über Umfang des staatl. Versicherungsschutzes bei gegebenem Regime

Mit einer Mehrheitsentscheidung wird der Umfang des kollektiv finanzierten Versicherungsschutzes entschieden. Dabei ist eine Pro-Kopf Prämie wie auch einkommensproportionale Beiträge möglich.

Einheitliche Kopfprämie:

Bei einheitlichen Kopfprämien ist der Preis einer Einheit staatlichen Versicherungsschutzes gleich der durchschnittlichen Krankheitswahrscheinlichkeit. Niedrigrisikotypen zahlen somit mehr für staatlichen Versicherungsschutz als für privaten, der sich nach ihrem individuellen Krankheitsrisiko richtet. Da diese Gruppe annahmegemäß in der Mehrheit ist, wird in diesem Regime der Umfang des kollektiv finanzierten Versicherungsschutzes bei einer Mehrheitsentscheidung bei null liegen.

Einkommensproportionale Beiträge:

Bei einkommensproportionalen Beiträgen unterscheidet sich der Beitrag zwischen armen und reichen Individuen, ohne jedoch – anders als auf dem privaten Markt – die unterschiedlichen Krankheitsrisiken widerzuspiegeln. Es stimmen deshalb nur die Wähler für ein positives Niveau staatlichen Versicherungsschutzes, für die gilt, dass das Verhältnis von individuellem zu durchschnittlichem Krankheitsrisiko größer ist als das Verhältnis von individuellem zu durchschnittlichem Einkommen. Das ist für arme Individuen mit einem hohen Krankheitsrisiko immer der Fall, während diese Bedingung für reiche Individuen mit einem niedrigen Krankheitsrisiko nie erfüllt ist. Entscheidend für den Ausgang ist deshalb, wie sich die armen „Niedrigrisiken“ entscheiden, die mit beiden Gruppen jeweils eine Mehrheit bilden können. Sie stimmen um so eher für einen positiven Umfang der staatlichen Versicherung, je höher entweder das Durchschnittseinkommen, d.h. je größer die Ungleichheit im Einkommen oder je größer der Anteil der Reichen, oder je kleiner das Krankheitsrisiko, d.h. je geringer die Ungleichheit im Krankheitsrisiko oder je geringer der Anteil der Hochrisiken. Wenn es eine Mehrheit für ein positives Versicherungsniveau gibt, dann entspricht der Umfang der Versicherung der Lösung des Optimierungsproblems der armen „Niedrigrisiken“. Andernfalls ist das Niveau null.

ad 1) Wahl der Finanzierungsform (Regime)

Schließlich ist zu untersuchen, welche Finanzierungsform sich auf der obersten Ebene der Entscheidung durchsetzt. Dazu wird zuerst die erstmalige Entscheidung hinter dem Schleier des Nichtwissens über den eigenen Risikotyp analysiert und anschließend die Frage untersucht, ob es bei Kenntnis des individuellen Risikos zu einer einstimmigen Revision kommen kann.

Wir beschränken uns auf den Fall, in dem es eine Mehrheit für eine positive Menge staatlichen Versicherungsschutz gibt. Diese Mehrheit wird gebildet von den armen Individuen mit hohem und niedrigem Krankheitsrisiko. Somit weiß ein armes Individuum ex ante, dass es ex post unter einkommensproportionalen Beiträgen nicht schlechter gestellt sein wird als mit einheitlichen Kopfprämien. Da die Armen annahmegemäß über eine Mehrheit verfügen, setzen sich bei der erstmaligen Wahl einkommensproportionale Beträge als Finanzierungsform durch.

Diese Entscheidung kann später – nach Aufdeckung des individuellen Risikotyps – nicht mehr rückgängig gemacht werden, wenn dazu Einstimmigkeit erforderlich ist. Die armen „Hochrisiken“ werden auf alle Fälle gegen einen Wechsel zu einheitlichen Kopfprämien sein.

In direkten Demokratien mit Mehrheitsentscheid ist also nach den Ergebnissen des theoretischen Modells zu erwarten, dass einerseits eine staatliche Krankenversicherung eher durch einkommensproportionale Beiträge als durch einheitliche Pro-Kopf Prämien finanziert wird und dass andererseits bei Entscheid für Pro-Kopf-Prämien der Umfang des staatlichen Versicherungsschutzes gegen null tendiert.

Auf den ersten Blick werden diese Ergebnisse durch empirische Fakten widerlegt. In der Schweiz existiert eine durch Pro-Kopf Prämien finanzierte staatliche Krankenversicherung, deren Umfang weit von null entfernt ist.

Es bleibt dann zu prüfen, ob das Modell empirische Beobachtungen zumindest besser erklären kann als konkurrierende Modelle. Die Ergebnisse müssen dafür entsprechend umformuliert werden. Sie lauten dann, dass staatliche Krankenversicherungen mit einheitlichen Pro-Kopf Prämien seltener sind als solche mit einkommensproportionalen Beiträgen und dass der erst genannte Typ durch einen geringeren Umfang gekennzeichnet ist als der zweite.

Tatsächlich gibt es einheitliche Pro-Kopf Prämien nur in der Schweiz und auch dort nur für Personen, die nicht der untersten Einkommensklasse angehören. Daneben zeigt ein Vergleich der Schweiz mit Deutschland als einem Land mit ähnlichem Pro-Kopf Einkommen, dass der Umfang sowohl der Ausgaben der jeweiligen staatlichen Krankenversicherungen als auch der BIP-Anteil der Sozialausgaben für Gesundheit in der Schweiz kleiner ist. Außerdem ist der Leistungskatalog in der Schweiz weniger umfangreich als in Deutschland und Versicherte müssen sich über einen Selbstbehalt an den Krankheitskosten beteiligen.

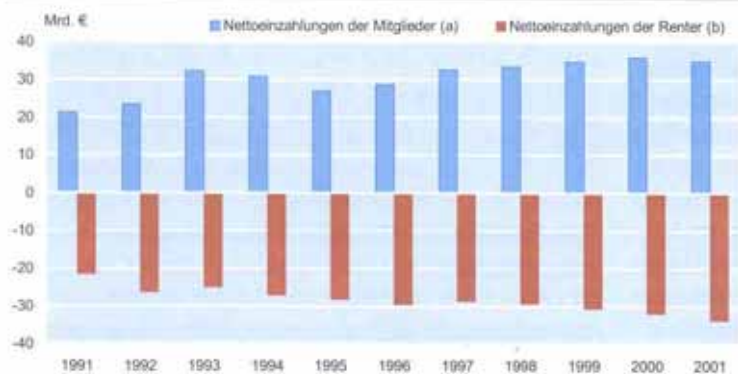
### **8.2.2 Krankenversicherung als partielles Umlageverfahren**

Aus politökonomischer Sicht lassen sich darüber hinaus auch Parallelen aufzeigen zwischen einem umlagefinanzierten Rentensystem und der gesetzlichen Krankenversicherung. Auch wenn es auf den ersten Blick nicht so erscheinen mag, da im Gegensatz zu Rentenversicherung Rentner wie Nicht-Rentner Beitragszahlungen für die Krankenversicherung leisten, so lässt sich doch eine intergenerative Umverteilung beobachten. Generationengerechtigkeit würde voraussetzen, dass jede Generation ihre eigenen Krankheitskosten trägt. Wenn für jüngere Mitglieder die Beiträge die Leistungsansprüche übersteigen, müssten diese Überschüsse gespart und zur Begleichung der höheren Krankheitskosten im Alter herangezogen werden. Jede Generation würde dann im Rahmen eines Kapitaldeckungsverfahrens für die eigenen Kosten aufkommen. Dies ist jedoch in der gesetzlichen Krankenversicherung in Deutschland nicht der Fall. Die Beitragszahlungen in einer Periode dienen dazu, die in der gleichen Periode anfallenden Kosten zu decken. Wie aus Abbildung 1 ersichtlich, übersteigen die Beiträge der Mitglieder ihre Leistungsausgaben um ca. 46%, während die Beiträge der Rentner nur ca. 43% ihrer Krankheitskosten decken.

Abbildung 1: Nettoeinzahlungen der Mitglieder in der GKV 1991-2001

Abb. 2

Nettoeinzahlungen der Mitglieder in der Gesetzlichen Krankenversicherung, 1991 - 2001



(a) Beiträge abzüglich Leistungsausgaben der pflichtversicherten und freiwilligen Mitglieder (ohne Rentner) in der GKV.

(b) Beiträge abzüglich Leistungsausgaben der Rentner in der GKV.

Quelle: BMG (2002), Sachverständigenrat (2002), Berechnungen des ifo Instituts.

ifo Schnelldienst 14/2003 – 56. Jahrgang

Fenge (2003)

Zu welchen Problemen dieses intergenerative Ungleichgewicht im Zusammenhang mit der demographischen Entwicklung in den nächsten Jahrzehnten führt, wird im nächsten Kapitel diskutiert. Im Folgenden sollen die politökonomischen Aspekte, die aus der teilweisen Finanzierung über ein Umlagesystem resultieren, im Hinblick auf die Einführung oder Erweiterung eines staatlichen Krankenversicherungssystems diskutiert werden.

Wie bei einer umlagefinanzierten Rentenversicherung bedeutet die Einführung einer Krankenversicherung einen Vorteil für die Einführungsgeneration. Dazu zählen in diesem Fall die Rentner und rentennahen Generationen. Der Einführungsgewinn besteht darin, dass sie Leistungen im Krankenfall erhalten, ohne von der ersten Lebensperiode an Beiträge gezahlt zu haben. Sie profitieren somit von den Einzahlungen der jüngeren Mitglieder – auch wenn in etwas geringerem Umfang als bei der Einführung eines Rentensystems wegen ihrer eigenen Beitragszahlungen.

In einem einfachen Modell mit drei gleich großen Generationen lässt sich das Abstimmungsverhalten analysieren. Die junge und mittlere Generation arbeitet und zahlt die vollen Beiträge, während die ältere Generation in Rente ist und nur noch den halben Beitragssatz leisten muss. Es kann nun davon ausgegangen werden, dass die älteren Individuen für die Einführung einer staatlichen Krankenversicherung stimmen werden. Sie zahlen lediglich in der dritten Periode Beiträge, ohne dass ihr Leistungsanspruch eingeschränkt werden würde, und erhalten so auf ihre Beiträge bei durchschnittlichen Krankheitskosten im Alter eine hohe Rendite. Die Individuen mittleren und jüngeren Alters werden nur dann dafür stimmen, wenn das Umlageverfahren eine höhere Rendite verspricht als eine Anlage auf dem Kapitalmarkt, wobei jeweils Risiko-Pooling für die Mitglieder einer Generation unterstellt wird. Für die mittlere Generation wird dies eher der Fall sein wegen ihrer kürzeren Beitragszeit als für die jüngere Generation. Eine Mehrheit für die Einführung existiert in diesem einfachen Modell dann, wenn die Individuen mittleren Alters (Mediengeneration) für die Einführung stimmen.

Das Ergebnis gilt in gleichem Maße auch für eine mögliche Ausweitung des Systems, was analog zu der Einführung einer (kleineren) zusätzlichen (teilweise) umlagefinanzierten Krankenversicherung gesehen werden kann.

Diese Überlegungen können erweitert werden hin zu einem Modell mit einer Generation pro Jahrgang mit unterschiedlicher Größe. Auch dann muss zunächst die Mediengeneration bestimmt und ihr Abstimmungsverhalten analysiert werden. Es ist nicht unwahrscheinlich, dass die Rendite der Krankenversicherung für die Mediengeneration wegen der nur noch für die zukünftigen Perioden zu leistenden Beiträge höher ist als die Rendite auf dem Kapitalmarkt. Es besteht also eine Tendenz zu einer Neu-Einführung von staatlichen Krankenversicherungssystemen bzw. zu Neu-Gründungen von Krankenkassen. Wenn letzteres in Deutschland nicht in großem Umfang zu beobachten ist, dann liegt das vor allem daran, dass die Krankenkassen in Deutschland nicht voneinander unabhängig sind, sondern über den Risikostrukturausgleich miteinander verbunden sind. Die Einführungsgewinne

können deshalb größtenteils nicht für die Mitglieder der neuen Kasse realisiert werden.

Eine Alterung der Gesellschaft führt dazu, dass auch die Mediengeneration älter wird. Ob dadurch eine Mehrheit für ein größeres System zustande kommt, hängt von der Bedeutung zweier Effekte ab. Zu einen bevorzugen ältere Individuen eine stärkere intergenerative Umverteilung, da sie nur noch relativ wenige Jahre Beitragszahlungen leisten müssen. Andererseits muss die Mediengeneration auch selbst einen Teil der höheren Beiträge, die dafür nötig sind, tragen. Wenn der erste Effekt überwiegt, dann wird die Tendenz hin zu Einführung oder Ausweitung eines staatlichen Krankenversicherungssystems bei Alterung sogar noch verstärkt.

### **8.3 Verbände im Gesundheitswesen**

Berufsverbände, in denen sich Ärzte oder Apotheker zusammenschließen, spielen im Gesundheitsmarkt eine außerordentliche Rolle. Dafür finden sich sowohl auf der Seite der Nachfrage nach Verbandstätigkeit als auch auf Seiten des Angebots von Verbandsunternehmen Gründe.

Auf der Nachfrageseite fühlt sich der einzelne Arzt veranlasst, einem Verband beizutreten, weil er ihm Schutz vor Wettbewerb bietet. Die Preis- und Qualitätskonkurrenz zwischen Mitgliedern wird eingeschränkt, um ihr Einkommen hochzuhalten. Der einzelne Arzt hat keine Veranlassung aus dieser Kartellposition auszubrechen, da die Patienten als Nachfrager in der Regel einen Versicherungsschutz genießen, der verhindert, dass sie nach dem jeweils preiswertesten Dienstleistungsangebot eines Arztes suchen. Auch ist ein Kostenvergleich zwischen verschiedenen Leistungsanbietern gerade im Gesundheitssektor sehr schwer zu bewerkstelligen. Außerdem verschaffen die Verbände dem einzelnen Mitglied Zugang zu öffentlichen Mitteln, die ein einzelner Anbieter nicht erlangen könnte.

Auf der Angebotsseite können die Kosten der Bildung und Funktionsfähigkeit eines Verbandes durch eine Reihe von Faktoren niedrig gehalten werden. Preisabsprachen lassen sich relativ leicht durchsetzen. So kann eine Preisdiskriminierung nicht etwa dadurch unterlaufen werden, dass reiche Patienten ärmere als Strohänner vorschicken. Das heißt, dass die personengebundene Dienstleistung des Arztes keine sekundären Märkte zulässt, die sonst eine Preisdiskriminierung erschweren würden. Außerdem ist der Zutritt zum Gesundheitsmarkt leicht zu kontrollieren. Über die Einschränkung der Niederlassungen wie auch durch die Überzeugungsarbeit staatlicher Stellen, einen Numerus clausus beim Medizinstudium einzuführen, ist der Zugang und die

Ausführung des Arztberufes weitgehend abgeschottet. Schließlich sind die Interessen der Verbandsmitglieder aufgrund der Überschaubarkeit der Zahl der Ärzte einer Fachrichtung und aufgrund der spezifischen Investitionen ziemlich homogen.

Diese besonderen Faktoren auf der Angebots- und Nachfrageseite des Gesundheitsmarktes können die Existenz von einflussreichen Verbänden erklären.

Die Funktion der Verbände können erstens in der Sicherung der Behandlungsqualität gesehen werden. Diese Funktion ist allerdings wenig glaubwürdig, da es kaum einen Verband gibt, der einem Arzt, selbst wenn er als unfähig erkannt wurde, die Approbation wieder entzieht. Außerdem weigern sich die Ärzteverbände, periodisch wiederkehrende Zulassungsprüfungen für Ärzte vorzunehmen, wie sie etwa von Fluggesellschaften bei Piloten durchgeführt werden. Zweitens kann der Verband als Wählerbeschaffer für Parteien fungieren, indem er an seine Mitglieder Wahlempfehlungen abgibt. Drittens und vorrangig besteht die Funktion des Verbandes darin, das Einkommen seiner Mitglieder zu sichern. Das wird erreicht, indem die Ärztedichte in einem Gebiet relative gering gehalten wird, um ihre Mitglieder vor Konkurrenz zu schützen.

Die Chancen für mehr Wettbewerb zwischen den Verbänden stehen deshalb schlecht, weil eigentlich nur die Patienten ein Interesse daran haben, dass der Wettbewerb das Entstehen neuer Organisationsformen der medizinischen Versorgung begünstigt. Für Ärzte besteht hingegen das Interesse gegenüber den Krankenkassen als Kollektivmonopol aufzutreten, um die Preise für ihre Leistungen hochzuhalten. Politiker wiederum sehen sich lieber einem einzelnen Verband gegenüber als einer Menge von konkurrierenden Verbänden, die dann für sie eher unüberschaubare Interessen vertreten. Allerdings werden sich Zwangsverbände langfristig nur mit Schwierigkeiten halten können, da sie im Widerspruch zur angestrebten Öffnung der Arbeitsmärkte innerhalb der Europäischen Union stehen.

## Kapitel 9: Herausforderungen

### Literatur:

Breyer, Zweifel, Kifmann (2005, Kap. 14 und 15)

Breyer und Ulrich (2000)

Fenge (2003)

Rürup Bericht (2003)

SVR (2003)

Wagner (2003)

### 9.1 Motivation

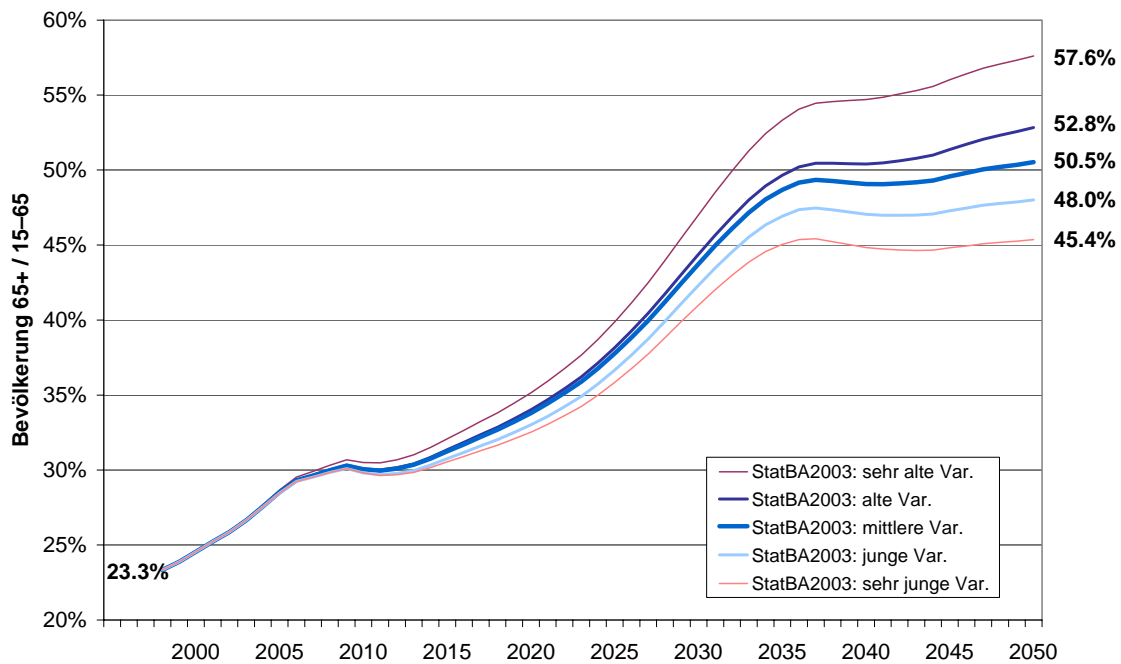
Zwei Aspekte, die die Finanzierung des Gesundheitssystems – unabhängig von dessen genauer Organisation – in den nächsten Jahrzehnten in praktisch ausnahmslos allen entwickelten Industrieländern erschweren werden, sind die demographische Entwicklung und der zu erwartende medizinisch-technische Fortschritt. Sie stellen das Gesundheitssystem vor große Herausforderungen und bilden den Hintergrund für Diskussionen über grundlegende, „strukturelle“ Reformen des Gesundheitswesens, wie sie z. B. derzeit in Deutschland geführt werden. Diese müssen zwar nicht zwangsläufig den Anstieg der Gesundheitsausgaben verringern, sollten aber die institutionellen Rahmenbedingungen ändern, unter denen er sich vollziehen wird.

### 9.2 Wirkungen von demographischer Entwicklung und technischem Fortschritt

#### 9.2.1 Demographischer Wandel

Die demographische Entwicklung in Deutschland ist geprägt durch den Umschwung von der „Baby-Boom“-Phase zu einem relativ ausgeprägten Geburtenrückgang in den 1960-er und 1970-er Jahren, einen säkularen Anstieg der Lebenserwartung, der internationalen Trends entspricht, und eine anhaltende Nettozuwanderung nennenswerten Umfangs. Auf der Basis dieser Einzelentwicklungen zeigen Vorausschätzungen des Statistischen Bundesamtes an, dass zwar eine gewisse Unsicherheit hinsichtlich der langfristigen Entwicklung der Bevölkerungszahl Deutschlands besteht, bis etwa 2050 aber mit einer massiven Verschiebung der Altersstruktur der Wohnbevölkerung zu rechnen ist (vgl. Abbildung 9.1). In allen Varianten der jüngsten Projektion der amtlichen Statistik wird sich der sogenannten „Altenquotient“, d. h. die Relation zwischen der Bevölkerung im Alter ab 65 Jahren zur („erwerbsfähigen“) Bevölkerung im Alter von 15 bis unter 65 Jahren mindestens annähernd verdoppeln.

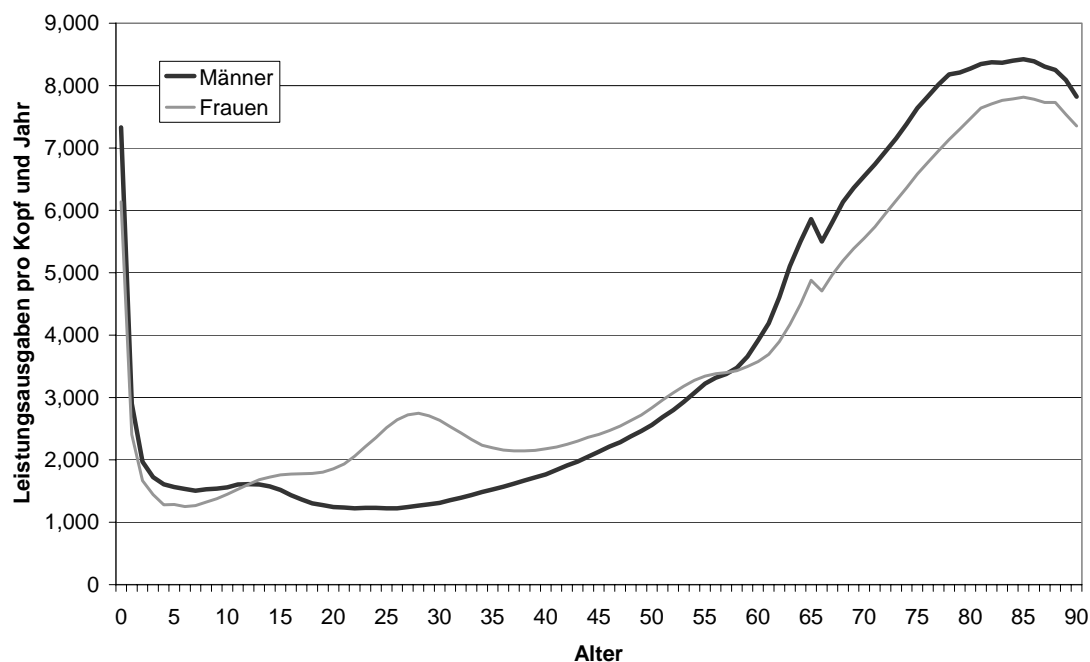
Abb. 9.1: Altenquotient der deutschen Wohnbevölkerung (1998–2050)



Quelle: Statistisches Bundesamt (10. koordinierte Bevölkerungsvorausschätzung, 2003)

Für das Gesundheitssystem ist diese Entwicklung deswegen bedeutsam, weil die Gesundheitsausgaben generell ein starkes Altersprofil aufweisen: Die Leistungsausgaben für Mitglieder der deutschen gesetzlichen Rentenversicherung belaufen sich beispielsweise für Männer im Alter über 75 Jahren auf das 6- bis 7-fache der Ausgaben für Männer im Alter von 20 bis 30 Jahren (vgl. Abbildung 9.2).

Abb. 9.1: Altenquotient der deutschen Wohnbevölkerung (1998–2050)



Quelle: Bundesministerium für Gesundheit und Soziale Sicherung

Für Prognosen hinsichtlich der zukünftigen Entwicklung der Gesundheitsausgaben muss allerdings nicht unbedingt damit gerechnet werden, dass das heute beobachtbare Altersprofil, korrigiert um reale Kostensteigerungen für alle Altersgruppen, unverändert fortzuschreiben ist und sich lediglich die Besetzung der Alterskohorten – namentlich mit einer sehr viel höheren Zahl von Personen im Alter über 80 Jahren als in der Vergangenheit – ändert. Die sogenannte „Kompressionshypothese“ besagt vielmehr, dass die weiter steigende Lebenserwartung zugleich die Zahl der Lebensjahre in Gesundheit erhöht, so dass die in Abbildung 9.2 gezeigten Kurven sich mit der Zeit immer weiter nach rechts dehnen würden. Eine Gegenthese – die sogenannte „Medikalisierungshypothese“ – geht jedoch davon aus, dass die höhere Lebenserwartung in erster Linie Ergebnis immer intensiverer medizinischer Behandlung ist, so dass die Kurven unverändert bleiben oder durch neue Formen von Multimorbidität am rechten Rand sogar noch weiter ansteigen werden. Empirisch ist der Widerspruch zwischen diesen beiden Hypothesen bis heute nicht ganz aufgeklärt.

### 9.2.2 Medizinisch-technischer Fortschritt

Es lässt sich beobachten, dass die Preissteigerungen im Gesundheitswesen regelmäßig höher ausfallen als die allgemeine Preisentwicklung. Üblicherweise wird diese Sonderentwicklung mit qualitativen Verbesserungen der Diagnose- und Behandlungsmethoden in Verbindung gebracht, für die es auch zahlreiche Beispiele gibt. Gleichwohl lässt sich diese Entwicklung in den Daten kaum gegen die Effekte

alternativer Erklärungsmöglichkeiten – inflationäre Tendenzen, steigende Monopolrenten der Leistungserbringer, Verschwendung der Versicherten oder der Versicherer – abgrenzen.

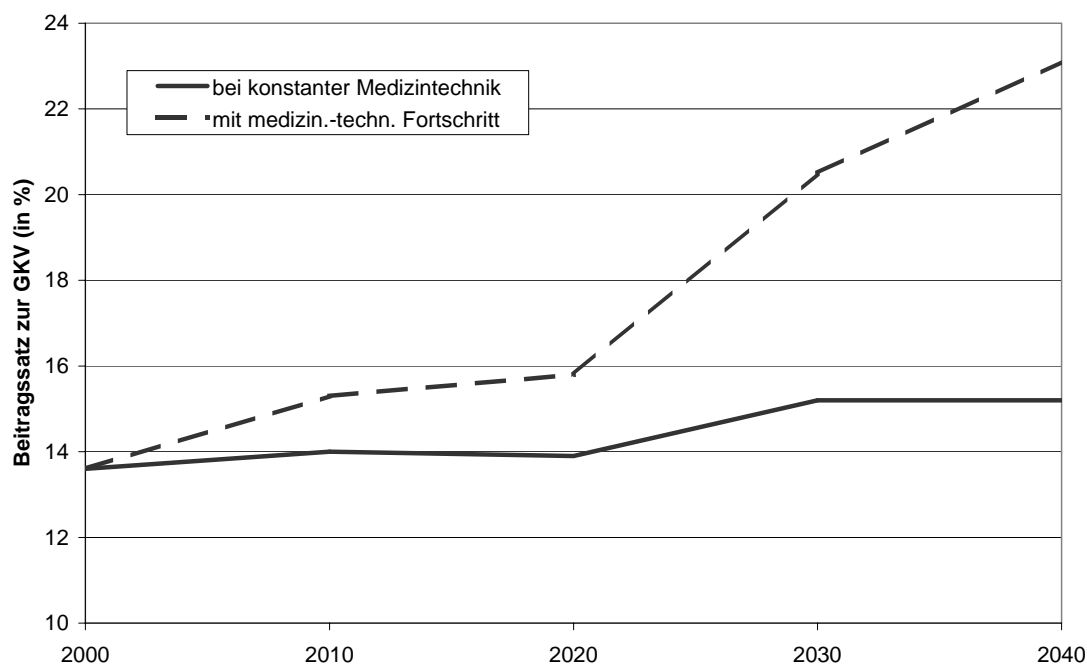
Soweit die auffallenden Preisentwicklungen im Gesundheitssektor in der Tat mit Qualitätsverbesserungen durch technischen Fortschritt begründet werden können, ist die zukünftige Richtung der medizinisch-technischen Entwicklung und ihrer Kostenwirkungen allerdings alles andere als klar. Dies liegt daran, dass dem technischen Fortschritt seiner Natur nach ein Element der Spontaneität und der Kreativität, kurz: des Unvorhersagbaren, innewohnt. Gleichwohl folgt die Richtung des technischen Fortschritts auch gewissen Anreizen, soweit er sich auf die Lösung aktueller Knappheitsprobleme richtet. Wenn im Gesundheitswesen in Zukunft die Suche nach ausgabendämpfenden Innovationen in den Vordergrund rückt, kann sich der ausgabensteigernde Effekt des bisherigen medizinisch-technischen Fortschritts auch durchaus umkehren (ohne dass heute klar sein könnte, wann das geschieht und wie stark der Effekt auf aggregierter Ebene ist).

Üblicherweise unterscheidet man „Produktinnovationen“ (im Gesundheitswesen: zur Verbesserung von Diagnosen oder zur Erzielung von Heilerfolgen, die vorher nicht möglich waren) und „Prozessinnovationen“ (bessere, häufig schnellere oder einfachere Verfahren für vorher schon erreichbare Therapieziele). Empirisch gesehen wirken erstere tendenziell preis- und kostentreibend, während letztere auch kostensenkend wirken können. Einen Spezialfall von Prozessinnovationen stellen „organisatorische Innovationen“ (Änderungen im Zusammenwirken verschiedener Instanzen und Handlungsträger) dar, die im Gesundheitswesen zumeist stark kostensenkend gewirkt haben (ein gutes Beispiel ist „*Managed care*“, vgl. Kapitel 7). Welcher dieser Typen von Innovationen in Zukunft dominiert, ist eine offene Frage, wobei dies – wie angedeutet – teilweise gesteuert werden kann.

### 9.2.3 Projektionen für die Ausgabenentwicklung im Gesundheitswesen

Auf der Basis von Vergangenheitsdaten (und mit dem Warnhinweis, dass daraus nur bedingt Schlussfolgerungen für die Zukunft gezogen werden können), lassen sich Art und Stärke des Einflusses von demographischer Entwicklung und technischem Fortschritt für die Gesundheitsausgaben mit Hilfe ökonometrischer Verfahren bestimmen und anschließend zu Zwecken der Simulation zukünftiger Entwicklungen verwenden. Da die Identifikation der Preis- und Kosteneffekte des technischen Fortschritts in den Daten, wie erläutert, allerdings nicht zweifelsfrei möglich ist, wird er in der Regel residual oder über reine Zeittrends vorgegebener Form (linear, exponentiell...) gemessen, was nicht ganz befriedigend sein kann.

Abb. 9.3 Beitragssätze zur deutschen gesetzlichen Krankenversicherung (2000–2040)



Quelle: Breyer und Ulrich (2000).

Die bis dato jüngste Arbeit dieses Typs stammt von Breyer und Ulrich (2000). Im Basismodell wird der technische Fortschritt dort als exponentieller Trend der nicht anderweitig kontrollierten, z. B. demographisch bedingten Ausgabenentwicklung gemessen, was seinen Effekt langfristig verstärkt. Abbildung 9.3 gibt die Ergebnisse der Simulation wieder, wobei die Ausgabenentwicklung in der deutschen GKV bei unveränderten rechtlichen Rahmenbedingungen hier in eine Projektion der zur Finanzierung dieser Ausgaben erforderlichen Beitragssätze übertragen wird.

Nach den Schätzungen von Breyer und Ulrich (2000) wird der Effekt des demographischen Wandels auf den GKV-Beitragssatz bis 2040, im Sinne der „Kompressionshypothese“, eher gering sein. In ihrer Studie ist es vor allem der medizinisch-technische Fortschritt, der einen Anstieg der Beitragssätze um rund 50 % innerhalb der nächsten vierzig Jahre nach sich zieht.

Dass das Gesundheitswesen im Zuge des laufenden demographischen Wandels ein Wachstums-„Markt“ sein wird, ist nicht *per se* schlecht. Vielmehr wird dies grundsätzlich den Präferenzen der im Durchschnitt alternden Individuen entsprechen und Felder für Forschung und Entwicklung, aber auch für zusätzliche Beschäftigung auf jedwedem Qualifikationsniveau, schaffen. Zu überdenken sind aber u. U. die institutionellen Rahmenbedingungen, unter denen sich diese Entwicklung vollziehen soll, namentlich die Modalitäten der – bisher lohnbezogenen – Finanzierung.

### 9.3 Reformdiskussionen

Die gegenwärtigen Reformdiskussionen im Gesundheitswesen finden vor dem Hintergrund des zu erwartenden technischen Fortschritts und der prognostizierten demographischen Entwicklung statt. Beide Faktoren werden zu einem steigenden Kostendruck führen, der auf der Einnahmeseite mehr Mittel erfordert und gleichzeitig auf der Ausgabenseite Leistungskürzungen notwendig macht und effizienzsteigernde Maßnahmen stärker in den Mittelpunkt rückt.

In den letzten 25 Jahren konnte dank ca. 260 Kostendämpfungsgesetzen eine Kostenexplosion weitgehend verhindert werden – zumindest wenn man dafür den Anteil der öffentlichen Gesundheitsausgaben am BIP betrachtet. Trotzdem ist der Beitragssatz in diesem Zeitraum von 8% auf 14,5% gestiegen bedingt durch – aber in gewissem Umfang auch ursächlich verantwortlich für – den Anstieg der Arbeitslosigkeit und der Frühverrentung. Die starke Verknüpfung von Arbeitskosten und Gesundheitssektor muss deshalb entkoppelt werden, um die negativen Effekte auf den Arbeitsmarkt und den Gesundheitssektor zu verringern.

#### 9.3.1 Gesundheitsreform 2003

Die Gesundheitsreform 2003 beinhaltet vor diesem Hintergrund folgende Maßnahmen:

Die Ausgaben sollen durch ein erhöhtes Kostenbewusstsein der Versicherten (Praxisgebühr und Zuzahlungsregeln) verringert werden. Außerdem werden die Vorschriften im Arzneimittelbereich gelockert und den Kassen die (kleine) Möglichkeit eröffnet, im Umfang von 1% ihres Budgets integrierte Versorgungsnetze zu initiieren. Die Einnahmen für die Krankheitsleistungen sollen gleichzeitig durch die Auslagerung bestimmter Leistungen aus der (paritätischen) Finanzierung erhöht werden. Dazu zählen die versicherungsfremden Leistungen, die durch einen pauschalen, steigenden Bundeszuschuss finanziert werden, sowie das Krankengeld und der Zahnersatz, die in Zukunft allein vom Arbeitnehmer zu tragen sind.

Trotz dieser Maßnahmen ändert sich jedoch nichts an der engen Verbindung von Arbeitskosten und Gesundheitsausgaben. Es ist deshalb notwendig, sich über andere Möglichkeiten der Finanzierung, insbesondere über konzeptionelle Alternativen zur lohnzentrierten Finanzierung der Gesetzlichen Krankenversicherung (GKV), Gedanken zu machen.

#### 9.3.2 Weitere notwendige Schritte bezüglich der Finanzierung

Wie schon erwähnt und wie auch im Rürup-Bericht betont, ist das gegenwärtige System lohnabhängiger Beiträge wegen der Koppelung des Gesundheitssystems mit Wachstum und Beschäftigung und wegen der konjunkturellen Abhängigkeit negativ zu beurteilen. Auch die Nachhaltigkeit in intergenerativer Sicht ist wegen der demographischen Herausforderung und dem medizinisch-technischen Fortschritt nicht gewährleistet.

Darüber hinaus ist es problematisch, ein System zu haben, das gesamtgesellschaftliche Verteilungsaufgaben erfüllt, ohne dass alle Einkommen aller Personen dafür herangezogen werden. In der gegenwärtigen Ausgestaltung findet in der GKV Umverteilung auf vielfältige Weise statt: von guten zu schlechten Risiken, aber eben auch von Alleinstehenden zu Familien und von hohen zu niedrigen Einkommen.

Eine Reform muss deshalb zum Ziel haben, die Finanzierung der GKV auf eine nachhaltigere Grundlage zu stellen und die negativen Beschäftigungswirkungen steigender Gesundheitsausgaben zu mindern. Im Rürup-Bericht und in der Öffentlichkeit sind als mögliche Richtungen Bürgerversicherung und Kopfpauschale diskutiert worden.

#### **a) Kopfpauschale** bzw. „Gesundheitsprämie“

Dieses Konzept ist angemessen, wenn Gesundheit als ein Gut angesehen wird, für das sich aufgrund von Angebot und Nachfrage Preise ergeben.

Auf der Grundlage des Äquivalenzprinzips umfasst die Kopfpauschale eine pauschale Belastung aller Bürger mit ca. €210. Somit erfolgt nur noch die Umverteilung von Gesunden zu Kranken innerhalb des Gesundheitssystems. Umverteilung von Alleinstehenden zu Familien und von Mitgliedern mit hohem Einkommen zu Mitgliedern mit niedrigem Einkommen werden dagegen aus dem Gesundheitssystem ausgegliedert und dem Steuer-Transfer System zugewiesen.

#### **b) Bürgerversicherung**

Dieses Konzept beruht auf der Vorstellung, dass Gesundheit ein Grundrecht ist, für das Marktmechanismen nicht von Bedeutung sein dürfen.

In Orientierung an dem Leistungsfähigkeitsprinzip werden alle Einkommen aller Bürger über einkommensabhängige Beiträge bis zu einer Grenze zur Finanzierung herangezogen. Trotz dieser stark verbreiteten Bemessungsgrundlage spielen aber weiterhin Arbeitseinkommen mit 95% die bei weitem wichtigste Rolle. Die personelle Einkommensumverteilung bleibt ein integrierter Bestandteil des System, d.h. es

kommt weiterhin zu Umverteilung von Gesunden zu Kranken, aber auch von Alleinstehenden zu Familien und von Besserverdienden zu Schlechterverdienden.

Beiden Vorschlägen ist in der konkreten Ausgestaltung gemein, dass nach ihnen das Gesundheitssystem nach wie vor als Umlageverfahren organisiert bleibt in Verbindung mit einem morbiditätsorientierten Risikostrukturausgleich. Kapitaldeckung wird nur eine ergänzende Rolle zugesprochen, da zwar die demographischen Risiken gut abgeschätzt werden können, es aber schwierig ist, den Kapitalstock zu bestimmen, der zum Ausgleich der durch den medizinisch-technischen Fortschritt bedingten Kostensteigerungen erforderlich ist.

Welche Argumente lassen sich nun für oder gegen die jeweilige Alternative finden? Es wird häufig vorgebracht, dass für die Kopfprämie die Effizienzgewinne sprechen, die durch die Umstellung realisiert werden können. Allerdings darf man nicht vergessen, dass bei Einführung einer Kopfpauschale auch eine klare Entscheidung für ein bestimmtes Umverteilungsniveau notwendig ist. Soll das Umverteilungsniveau so erhalten bleiben wie es ist, dann ist der Ausbau des Steuersystems nötig, soll jedoch Umverteilung reduziert werden, dann wäre auch eine Reform des Steuersystems ausreichend. Wenn man der Kopfpauschale positive Effekte zuspricht, dann liegen diese eher darin, dass Selbstbeteiligung in effizienterer und gerechterer Form einführbar ist, dass auch Kapitaleinkommen – wenn auch in quantitativ unbedeutendem Umfang – zur Umverteilung herangezogen wird, und dass die Transparenz durch die Konzentration der Umverteilungsaktivitäten im Steuersystem erhöht wird. Am letzten Punkt setzt aber auch die Kritik an. Eine vollständige oder teilweise Verlagerung der gesundheitsunabhängigen Umverteilung in das Steuer-Transfer System führt zu einem zusätzlichen Finanzierungsbedarf in Höhe von €27 Mrd. und damit entsprechend zu zusätzlicher Belastung der öffentlichen Haushalte. Auch wenn der Arbeitgeberanteil zur GKV ausgeschüttet und versteuert werden würde, bliebe immer noch eine Lücke von €10 Mrd. Diese Lücke ist vor allem dann immer schwieriger zu finanzieren, wenn das Steuersystem gleichzeitig immer weniger progressiv wird.

Von einer Bürgerversicherung verspricht man sich eine Stärkung des Solidarprinzips. Die Hauptkritik richtet sich dabei darauf, dass es noch weniger gelingt, die Beitragsentwicklung von der Entwicklung der Arbeitseinkommen abzukoppeln. Daneben ist zu erwarten, dass es administrative und konzeptionelle Schwierigkeiten wegen der konkurrierenden Abgrenzung von Einkommensbegriffen im Steuer- und Sozialrecht gibt.

Beide Konzepte überzeugen also nicht vollständig. Es ist deshalb nicht unwahrscheinlich, dass es zu einer Kombination von Kopfpauschale und

Bürgerversicherung kommt. Gert Wagner (2003) hat mit der „Bürgerprämie“ einen Systemvorschlag gemacht, der in diese Richtung zielt. Jeder Erwachsene zahlt danach eine pauschale Prämie für einen Standardkatalog an Leistungen. Die Prämien schwanken in Abhängigkeit der Wirtschaftlichkeit der Krankenkassen und Versicherungsunternehmen; sie werden jedoch nicht von der Risikostruktur der Mitglieder beeinflusst (Risikostrukturausgleich). Es kommt somit zu einem Wettbewerb zwischen den (gesetzlichen und privaten) Versicherungen. Ein Problem ist aber auch hier, wie der soziale Ausgleich erreicht werden kann.