

DIPLOMARBEIT

Titel der Diplomarbeit

Double Moral Hazard im Arzt-Patienten Verhältnis

Verfasserin / Verfasser (Nichtzutreffendes weglassen)

Sandra Dobrozemsky

angestrebter akademischer Grad

Magistra der Sozial- und Wirtschaftswissenschaften
(Mag. rer. soc. oec.)

Wien, 2012

Studienkennzahl lt. Studienblatt:
Studienrichtung lt. Studienblatt:
Betreuer / Betreuerin:

A 157
Diplomstudium Internationale Betriebswirtschaft
Univ.-Prof. Dr. Franz Wirl

Eidesstattliche Erklärung

„Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.“

Widmung und Danksagung

Ich widme diese Arbeit meiner Familie, vor allem meinen Eltern, Irmtraud und Peter, die mich nicht nur während des Studiums, sondern während meines gesamten Lebens bei meinen Entscheidungen immer unterstützt haben und stets Vertrauen in mich hatten.

Ich möchte mich an dieser Stelle bei meinem Studienkollegen Matthias Folkmann für die Motivation, Unterstützung und langjährige Freundschaft bedanken.

Mein Dank gilt außerdem Univ.-Prof. Dr. Franz Wirl für die Betreuung während des Verfassens dieser Diplomarbeit und während meines Studiums an seinem Lehrstuhl.

1 Inhaltsverzeichnis

1	<i>Inhaltsverzeichnis</i>	2
2	<i>Abbildungsverzeichnis</i>	4
3	<i>Introduction</i>	5
4	<i>Entwicklung der Informationsökologie</i>	5
4.1	Definition des ökonomischen Informationsbegriffs	6
4.2	Typen von asymmetrischer Information	7
4.2.1	adverse selection	7
4.2.2	moral hazard	7
4.2.3	holdup	8
4.3	Akerlof'sche „Market for Lemons“	9
5	<i>moral hazard</i>	12
5.1	moral hazard im Modell	13
5.1.1	The first-order approach	14
5.2	moral hazard im Gesundheitswesen	16
6	<i>Double moral hazard im Arzt-Patienten Verhältnis</i>	17
6.1	Informationsasymmetrien zwischen Arzt und Patient	18
6.2	Interdependenzen zwischen Leistung des Arztes und Verhalten des Patienten	19
6.3	Die Modelle	21
6.3.1	Udo Schneider	21
6.3.2	Paula González	30
6.3.3	David Bardey und Romain Lesur	36
6.3.4	Ching-To Albert Ma und Thomas G. McGuire	39
6.3.5	Kenneth L. Leonard und Joshua Zivin	44
6.3.6	Vergleich der Modelle	50
6.4	Strategisches Zusammenspiel von ärztlicher Leistung und Compliance	51
6.4.1	Strategische Substitute versus strategische Komplementäre	52
6.4.2	Interdependenzen ohne Einführung einer Selbstbeteiligung	53
6.4.3	Interdependenzen mit Einführung einer Selbstbeteiligung	59

7	<i>Empirische Studien</i>	67
7.1	Das Model.....	68
7.2	Ergebnisse.....	70
7.3	Weitere empirische Erkenntnisse.....	72
8	<i>Conclusio</i>	74
9	<i>Literaturverzeichnis</i>	77
10	<i>Anhänge</i>	80
10.1	Abstract	80
10.2	Lebenslauf.....	81

2 Abbildungsverzeichnis

Abbildung 1: Grundtypen asymmetrischer Information.....	9
Abbildung 2: Zeitliche Abfolge des moral hazard Problems	13
Abbildung 3: Vertragsstufe und Behandlungsstufe.....	21
Abbildung 4: Unabhängigkeit von ärztlicher Leistung und Compliance	54
Abbildung 5: Ärztliche Leistung und Compliance als strategische Komplemente	56
Abbildung 6: Ärztliche Leistung und Compliance als strategische Substitute mit gleichgerichteten Effekten	57
Abbildung 7: Ärztliche Leistung und Compliance als strategische Substitute mit gegenläufigen Effekten.....	58
Abbildung 8: Selbstbeteiligung bei strategischer Unabhängigkeit	61
Abbildung 9: Selbstbeteiligung bei strategischen Komplementen	63
Abbildung 10: Selbstbeteiligung bei strategischen Substituten und erhöhter Genesungswahrscheinlichkeit.....	64
Abbildung 11: Selbstbeteiligung bei strategischen Substituten und verminderter Genesungswahrscheinlichkeit.....	65

3 Introduction

Im Verhältnis zwischen Arzt und Patient kommt es unwillkürlich zu Informationsasymmetrien. Ähnlich wie im Produktgarantiefall verursachen diese Asymmetrien wirtschaftliche Nachteile, anders als in anderen informationsökologischen Fällen werden beim Arzt-Patienten Verhältnis allerdings auch medizinische und gesundheitliche Nachteile verursacht. Trotz der Schädigung der Gesundheit haben beide Parteien Anreize sich „unmoralisch“ zu verhalten. In dieser Arbeit soll die doppelseitige Informationsasymmetrie im Arzt-Patienten Verhältnis untersucht werden. Im ersten Teil der Arbeit sollen der Begriff der Informationsökologie und die unterschiedlichen Formen erklärt werden. Im zweiten Teil wird moral hazard und double moral hazard im Arzt-Patienten Verhältnis anhand von unterschiedlichen Modellen untersucht. Der dritte Teil dieser Arbeit widmet sich empirischen Studien, die die theoretischen Modelle für moral hazard im Gesundheitswesen belegen sollen.

4 Entwicklung der Informationsökologie

Informationsasymmetrien entstehen dadurch, dass von zwei kooperierenden Partnern, der eine vergleichbar besser informiert ist als der andere. (Spremann 1990). Das an sich würde noch kein großes Problem darstellen, wenn sich daraus nicht ökonomische Nachteile und auch Wohlfahrtsverluste ergeben würden. Weswegen sich die Frage stellt, mit welchen Modellen und Methoden man dieser ungleichmäßigen Verteilung der Information entgegenwirken kann.

Überraschenderweise gab es lange Jahre kaum bis keine Anstrengungen, die Asymmetrie der Information innerhalb der Beziehung zweier Parteien in einem ökonomischen Verhältnis zu untersuchen. Man ging sogar in der ökonomischen Theorie lange Zeit von der Prämisse der vollständigen Information der Beteiligten aus.

Erste Überlegungen zu informationsökonomischen Ansätzen machte Hayek (1945) in seiner Arbeit „The Use of Knowledge in Society“. Er sprach davon, dass das Grundproblem der Koordination wirtschaftlicher Aktivitäten gerade darin läge, dass niemand über alle Informationen in ihrer Gesamtheit verfüge. Damit schien der Stein ins Rollen gebracht worden zu sein, denn seither wurde dieser Bereich oft und intensiv behandelt.

Eines der bedeutendsten Werke auf dem Gebiet ist ohne Zweifel „The Market of Lemons“ von Akerlof (1970), auf das später in dieser Arbeit noch näher eingegangen wird.

4.1 Definition des ökonomischen Informationsbegriffs

Als Grundlage für weitere Ausführungen wird in diesem Abschnitt der Informationsbegriff in der Entscheidungstheorie näher beleuchtet.

Information ist ein Werkzeug zur Verringerung von Unsicherheiten, wobei unter Unsicherheit eine Situation z_0 zum Zeitpunkt t_0 verstanden wird, von der aus betrachtet zu einem bestimmten späteren Zeitpunkt t_1 einer von mehreren möglichen Umweltzuständen z_i mit $i \in \{1, 2, \dots, n\} = I$ eintreten kann. Die Wahrscheinlichkeiten $p_i \in [0, 1]$ sind bekannt, wobei $\sum_{i \in I} p_i = 1$ gelten muss. (Decker 2000)

Wird nun angenommen, dass den einzelnen Zuständen z_i Güterbündel $x_i \in \mathfrak{R}_m$ zugeordnet sind, und gibt es eine kardinale Nutzenfunktion mit $u(x_i) = \bar{u} \in \mathfrak{R}_+$, so ergibt sich aus Sicht von z_0 bei Unsicherheit der Erwartungsnutzen

$$\bar{u} = \sum_{i \in I} p_i \cdot u(x_i)$$

4.2 Typen von asymmetrischer Information

Nach Spremann (1990) lassen sich drei Grundtypen von asymmetrischer Information unterscheiden:

- adverse selection (Qualitätsunsicherheit)
- moral hazard
- holdup

Auch wenn in dieser Arbeit vorwiegend moral hazard-Fälle aufgezeigt werden, soll an dieser Stelle kurz auch auf die anderen Formen des Informationsungleichgewichts eingegangen werden.

4.2.1 adverse selection

Im Fall von adverse selection spricht man von *hidden information*. Der informierte Vertragspartner (**A**gent) weist bereits ex ante (vor Vertragsabschluss) eine gewisse unbeeinflussbare Eigenschaft auf, die der uninformierte Partner (**P**rinciple) erst ex post beobachten kann.

Ein Beispiel hierfür wäre der *Akerlof'sche* Gebrauchtwagenmarkt (vgl. 2.3) oder ein Arbeitgeber, der die einen Bewerber als fleißigen beziehungsweise faulen Arbeitnehmer einschätzen muss.

Lösungen für die adverse selection Problematik wäre zum Beispiel das Setzen von Signalen (Produktgarantien im Handel, Zusatzausbildung im Arbeitsmarkt).

4.2.2 moral hazard

Bei moral hazard haben es die Beteiligten mit *hidden action* zu tun. A kann sein Verhalten selbst beeinflussen (er wählt zum Beispiel im Arbeitsverhältnis einen hohen oder niedrigen Anstrengungslevel). Die Eigenschaft ist von P allerdings

nicht mal ex post direkt beobachtbar, er kann nur von einem erbrachten Ergebnis auf das Verhalten schließen.

moral hazard kommt häufig bei Versicherungen vor. Ein Fahrradbesitzer wird nach Abschluss einer Fahrradversicherung nicht mehr so sorgfältig beim Absperren und Abstellen seines Fahrrades umgehen, wie er das vor dem Abschluss der Versicherung getan hat. Als Problemlösung werden hierbei oft Anreize, wie zum Beispiel Selbstbehaltskosten, geschaffen.

4.2.3 holdup

Das Verhalten von A unterliegt dessen Willen und ist noch variabel. In diesem Fall wird es allerdings von P ex post erkannt. Oft treten holdup-Probleme auf, wenn eine Vertragsseite juristische Freiräume bei Vertragsinhalten ihre Gegenleistung betreffend hat.

Ein Beispiel im Mietverhältnis wäre, dass der Vermieter dem Mieter nicht sofort nach Einzug die Miete erhöht – man spricht in diesem Zusammenhang auch von Kulanz oder Fairness. Der Mieter könnte in dem Fall nicht gegen den Vermieter klagen, da dieser nicht gegen den Vertragsinhalt verstoßen hat. Er würde sich aber mit hohen *sunk costs* konfrontiert sehen, würde er aus dem Vertrag aussteigen wollen (Kosten, die ein erneuter Umzug verursachen würde).

Abbildung 1: Grundtypen asymmetrischer Information

Grundtypen asymm. Info.	adverse selection	moral hazard	holdup
Verhalten des A	exogen gegeben; wird P ex post bekannt	willensabhängig; bleibt P auch ex post verborgen	willensabhängig; wird P ex post bekannt
Interpretation	Qualifikation	Anstrengung	Fairness
Ursachen	Informationsineffizienz	Ressourcenplastizität	sunk costs
Adäquates Kooperationsdesign	Offenlegung	Anreizsysteme	Autorität

Quelle: Spremann, 1990 (leicht angepasste Form)

4.3 Akerlof'sche „Market for Lemons“

Von George A. Akerlof stammt eines der bedeutendsten Werke der Informationsökonomie „The Market for ‚Lemons‘: Quality Uncertainty and the Market Mechanism“ (1970).

Anhand des Automarkts hat er die Problematik der asymmetrischen Information in der Ökonomie aufgezeigt. Mit seinem ‚lemon‘-Markt zeigte er, dass schlechte Qualität auf Grund von Informationsasymmetrien gute Qualität tatsächlich verdrängen kann.

Gegeben sein ein Gebrauchtwagenmarkt, in dem ‚gute‘ Autos und Autos schlechter Qualität (im amerikanischen ‚lemons‘ genannt) existieren.

Die Nachfrage von Gebrauchtwagen hängt nun ab vom Preis p und der durchschnittlichen Qualität μ .

Es gilt außerdem:

$$\mu = \mu(p) \quad \text{sowie} \quad S = S(p)$$

Sowohl die durchschnittliche Qualität, als auch das Angebot sind abhängig vom Preis. Für ein Marktgleichgewicht sollte wie immer das Angebot gleich der Nachfrage sein. Es gibt nur zwei Gruppen von Händlern (1, 2).

Die Nutzenfunktionen wären dann:

$$U_1 = M + \sum_{i=1}^n x_i \quad U_2 = M + \sum_{i=1}^n \frac{3}{2} x_i$$

M = alle Güter, außer Autos, die konsumiert werden, x_i ist die Qualität des i -ten Autos und n ist die Anzahl der Autos.

Beide Gruppen sind Neumann-Morgenstern Nutzenmaximierer, das heißt, jeder Akteur ist homo oeconomicus und strebt danach, aus der Menge ihm zur Verfügung stehenden Alternativen diejenige zu wählen, die ihm den größten Nutzen bringt.

Gruppe 1 besitzt N Autos mit einer Qualitätsverteilung von $0 \leq x \leq 2$ und Gruppe 2 besitzt keine Autos. Der Preis von M wird mit 1 angenommen. Y_1 , Y_2 bezeichnet die das Einkommen der jeweiligen Gruppe.

Nachfrage der Gruppe 1 für Autos:

$$D_1 = Y_1/p \quad \mu/p > 1$$

$$D_1 = 0 \quad \mu/p < 1$$

Angebot von Gruppe 1 an Autos:

$$S_2 = pN/2; \quad p \leq 2$$

mit einer durchschnittlichen Qualität von

$$\mu = p/2$$

Nachfrage der Gruppe 2 für Autos:

$$D_2 = Y_2/p \quad 3\mu/2 > p$$

$$D_2 = 0 \quad 3\mu/2 < p$$

$$S_2 = 0$$

Die gesamte Nachfrage $D(p, \mu)$ ist demnach:

$$D(p, \mu) = (Y_2 + Y_1)/p \quad \text{if } p < \mu$$

$$D(p, \mu) = Y_2/p \quad \text{if } \mu < p < 3\mu/2$$

$$D(p, \mu) = 0 \quad \text{if } p > 3\mu/2$$

Bei Preis p ist demnach nach Akerlof die durchschnittliche Qualität $p/2$ und deswegen wird bei keinem Preis ein Handel stattfinden; und das, obwohl, bei jedem beliebigen Preis zwischen 0 und 3 Akteure der Gruppe 1 Autos verkaufen würden und gleichzeitig Mitglieder der Gruppe 2 bereit wären zu kaufen.

Das Problem ist, dass nur der Verkäufer über die Qualität seiner Autos Bescheid weiß. Da der Käufer den Zustand des Fahrzeuges nicht kennt, wird er nur den Durchschnitt der Preise für hohe und niedrige Qualität zahlen. Zu diesem Preis werden allerdings nur die Besitzer von ‚lemons‘ verkaufen, nicht allerdings die Verkäufer der guten Wagen. Somit verdrängt die schlechte Qualität die gute Qualität vom Markt.

Eine ähnliche Variante davon ist Gresham's Gesetz, das besagt, dass schlechtes Geld gutes Geld verdrängt. Dies stammt aus dem 16. Jahrhundert, in dem noch mit Silber- und Goldmünzen bezahlt wurde. Wobei die Goldmünzen gespart wurden, während bald tatsächlich nur noch mit Silbermünzen bezahlt wurde. Der Unterschied zu Akerlof's Entdeckungen liegt darin, dass bei Gresham beide Parteien den gleichen Informationslevel besaßen.

5 moral hazard

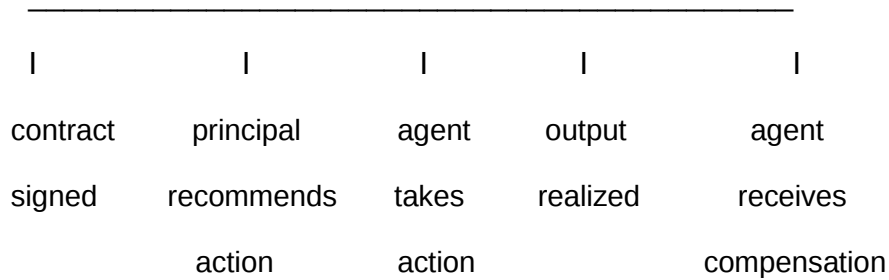
In einem Verhältnis zweier Vertragspartner, findet moral hazard statt, wenn nicht beobachtbare oder nicht verifizierbare Aktionen der einen Partei den Wohlstand der anderen Partei beeinflusst. (Müller 1997)

Zum Zeitpunkt des Vertragsabschlusses haben beide Parteien den gleichen Informationslevel. Eine Asymmetrie der Information ergibt sich meist erst nach dem Abschluss des Vertrages, da erst dann der Agent das Ausmaß seiner *hidden action* wählt.

Der Grad der Anstrengung ist dabei vom Prinzipal nicht ersichtlich, er kann nur das Ergebnis, den Payoff, als Indikator für die *hidden action* des Agenten verwenden. Härtere Arbeit und Anstrengung wird in den meisten Fällen die Wahrscheinlichkeit für höhere Payoffs mit sich bringen. (Theilen 1996, 9)

Prescott (1999) beschreibt in seiner Arbeit die fünf Stufen bei der Entstehung des moral hazard Problems. In der ersten Stufe einigen sich Prinzipal und Agent auf einen Vertrag und unterschreiben diesen. In Stufe 2 schlägt der Prinzipal dem Agenten eine Handlung vor. Im nächsten Schritt entscheidet sich der Agent für oder gegen diese Handlung. Wenn er mit dem Vorschlag des Prinzipals übereinstimmt, so führt er in dieser Phase die Handlung auch aus. Im vierten Schritt wird ein gewisser Output realisiert und im finalen Schritt wird der Agent entlohnt.

Abbildung 2: Zeitliche Abfolge des moral hazard Problems



Quelle: Prescott, 1999

Beispiele für moral hazard gefährdete Vertragverhältnisse sind das Verhältnis Arbeitgeber-Arbeitnehmer, Gesellschafter-Manager oder auch Versicherungsgesellschaft-Versicherter.

5.1 moral hazard im Modell

Den theoretischen Hintergrund zu moral hazard soll nun folgendes statisches Principal-Agent-Modell von Müller (1997) veranschaulichen.

a *hidden action* des Agenten $a \in A$

x Output

w wealth

θ exogene Variable (Umwelt)

$S_{(x)}$ optimale Risikoaufteilung

v Nutzen des Agenten

g negativer Nutzen des Agenten

Holmström (1979) interpretierte a als produktiven Input mit einem direkten Negativnutzen für den Agenten und nach Holmström entsteht dadurch der ausschlaggebende Unterschied in den Zielen des Agenten beziehungsweise des Prinzipals.

Die Handlung des Agenten a kann von θ beeinträchtigt werden, der Agent kann allerdings θ nicht oder nur zu einem Bruchteil beeinflussen.

5.1.1 The first-order approach

Das Verhältnis von a zu x wird als stochastisch angenommen, mit einer Verteilungsfunktion von $F(x | a)$ und einer Dichtefunktion von $f(x | a)$. Es wird weiters angenommen, dass der Prinzipal risikoneutral ist, während der Agent ein von Neumann-Morgenstern Nutzenmaximierer mit einer Nutzenfunktion

$$V_{(w,a)} = v_{(w)} - g_{(a)} \quad \text{mit} \quad v'_{(w)} > 0 \quad \text{und} \quad v''_{(w)} < 0$$

ist.

Dadurch, dass $S_{(x)}$ den Lohn des Agenten als Funktion von seinem Output festlegt, kann der Prinzipal sehr wohl die Wahl des Agenten, welchen Anstrengungslevel er wählt indirekt beeinflussen.

Könnte nun der Prinzipal a beobachten, so wäre sein Maximierungsproblem

$$(1.1) \quad \max_{S_{(x)}, a} \int (x - S_{(x)}) f(x | a) dx$$

$$(1.2) \quad \text{s.t.} \quad \int v(S_{(x)}) f(x | a) dx - g_{(a)} \geq V'$$

(participation constraint)

Voraussetzung für (1.2) ist allerdings, dass der erwartete Nutzen des Agenten, den Reservationsnutzen übersteigt $V' \geq 0$.

Durch Ableitung nach $S_{(x)}$ nach Lagrange mit dem Multiplikator λ erhält man die Bedingung erster Ordnung (first-order condition)

$$(1.3) \quad 1 / v'(S_{(x)}) = \lambda \quad \text{für alle } x$$

Dieses first-best Szenario ist wegen der Existenz von unvollständigen Verträgen sehr unrealistisch. Deswegen wird nun weiters angenommen, dass der Prinzipal a nicht beobachten kann, es entsteht dadurch ein moral hazard Szenario.

Es reicht jetzt nicht mehr aus, den Agenten dazu zu bringen, mitzumachen, also ihn besser als seinen Reservationsnutzen zu stellen. Der Prinzipal muss ihm zusätzlich Anreize bieten, damit sich der Agent für einen bestimmten Anstrengungslevel entscheidet. Denn ohne diese zusätzlichen Anreize wird sich der Agent eher für einen niedrigen Anstrengungslevel entscheiden. Um den Agenten für höhere Anstrengung zu motivieren, muss eine weitere Bedingung der Zielfunktion hinzugefügt werden, die Anreizkompatibilitätsbedingung (incentive compatibility constraint).

$$(1.4) \quad a \in \arg \max_{\hat{a} \in A} \int v(S_{(x)}) f(x | \hat{a}) dx - g(\hat{a})$$

Wobei “argmax” ein Set von Argumenten kennzeichnet, das die folgende Zielfunktion maximiert. (Holmström 1997)

Die vom Prinzipal gewünschte Anstrengung hat den Wert $\hat{a}$. Um $\hat{a}$ zu wählen, muss für den Agenten der Nutzen in diesem Punkt mindestens so hoch sein, wie für alle anderen möglichen Werte von a .

Da die Anreizkompatibilitätsbedingung selbst ein Optimierungsproblem ist, wird die Lösung des Optimierungsproblems selbst sehr kompliziert. In diesem Falle gäbe es bis zu unendlich viele Anstrengungsniveaus, aus denen der Agent wählen könnte, deswegen wird die Bedingung (1.4) durch die first-order condition (1.5) des Agenten ersetzt, um so das *relaxed program* des Prinzipalen zu erhalten.

$$(1.5) \quad \int v(S_{(x)}) f(x | \hat{a}) dx - g(\hat{a}) = 0$$

Wiederum mit Lagrange nach $S_{(x)}$ gelangt man zur notwendigen first-order condition

$$(1.6) \quad 1 / v'(S_{(x)}) = \lambda + \mu f_a(x | a) / f(x | a) \quad \text{für alle } x,$$

$\mu > 0$ und $v'(\cdot)$ ist strikt fallend, was bedeutet, dass das Einkommen des Agenten variiert. Dadurch wurde festgestellt, dass es sehr wohl eine wechselseitige Beziehung zwischen den Anreizen und der Risikoaufteilung gibt.

In dieser second-best Lösung kann der Output x als Signal für die nicht beobachtbare *hidden action* herangezogen werden. (Holmström 1997)

5.2 moral hazard im Gesundheitswesen

Wie schon oben erwähnt, tritt moral hazard häufig bei Vertragsverhältnissen im Versicherungswesen auf. Dies gilt vor allem bei Krankenversicherungen im Gesundheitswesen. Arrow (1963) war einer der ersten, der sich damit beschäftigt hat, wie sich eine Einführung von Krankenversicherung auf die Nachfrage nach sozialen Leistungen der Versicherten auswirkt. Wird mit einer Versicherung der Preis einer ärztlichen Behandlung gesenkt, so steigt bei elastischer Nachfrage beim Patienten die Nachfrage nach ärztlichen Leistungen. Zu beobachten ist dies zum Beispiel auch bei Patienten mit privaten Zusatzversicherungen. Diese Patienten haben nach Abschluss der Zusatzversicherung oft eine viel höhere Nachfrage nach medizinischen Leistungen als davor. Dies hängt allerdings nicht nur von der Nachfrage des Patienten ab. Auch Ärzte und Krankenhäuser schlagen teilweise Patienten mit großzügigeren Versicherungen mehr und teure Behandlungen vor, da diese Patienten nur einen kleinen Teilbetrag der Kosten tatsächlich selbst bezahlen müssen. (Pita Barros et.al. 2008)

Savage und Wright (2003) haben gezeigt, dass in Australien die Krankenhausaufenthalte der Patienten mit Zusatzversicherungen weitaus länger waren, als die Aufenthalte der herkömmlich versicherten Patienten.

6 Double moral hazard im Arzt-Patienten Verhältnis

Der theoretische Ansatz des double moral hazard soll in diesem Teil der Arbeit auf den praktischen Fall der Arzt-Patienten Beziehung übertragen werden. Die Anwendungen in diesem Teil der Arbeit stützen sich vor allem auf die Forschungen und Veröffentlichungen von Udo Schneider, der sich sehr intensiv mit dem Fall des double moral hazard in der Beziehung zwischen Arzt und Patient beschäftigt hat.

Der Erfolg einer medizinischen Behandlung hängt in großen Teilen von der Beziehung und Kommunikation zwischen Arzt und Patient ab. Bis jetzt wurde moral hazard im Gesundheitswesen zumeist beim Patienten als Agenten festgestellt, dessen Prinzipal das Versicherungsunternehmen ist. Hier lag also oft einfach ein klassisches Beispiel für moral hazard im Versicherungswesen vor.

Moral hazard besteht allerdings auch beim Verhältnis des Patienten zu seinem Arzt. Vergleichbar ist dieser Fall teilweise mit moral hazard im Produktgarantiefall. Die Behandlung des Arztes ist oft nur soweit erfolgreich, als sich der Patient auch vernünftig und vorsichtig verhält. Genauso wie eine Produktgarantie kein Freifahrtschein für die unvorsichtige Behandlung mit der Ware ist, muss auch der Patient vorsichtig mit seinem Körper umgehen. Er muss die medizinischen Ratschläge und Behandlungen – Gebrauchsanweisung – des Arztes befolgen.

Der Arzt-Patient-Fall ist dahingehend ein eher außergewöhnlicher, weil hierbei (wie auch vereinzelt im Produkthaftungsfall) double moral hazard vorliegt. Denn auch der Arzt ist versucht, sich nicht moralisch zu verhalten. Er ist in seiner Funktion nicht nur Anbieter der medizinischen Leistung, sondern auch, in seiner

Rolle als Berater für den Patient, Nachfrager der eigenen Leistungen. Bei beiden findet also eine *hidden action* statt.

6.1 Informationsasymmetrien zwischen Arzt und Patient

Der Arzt hat durch seine Ausbildung und seine Arbeitserfahrung einen Wissensvorsprung, gegenüber dem Patienten, der in den meisten Fällen zum Arzt geht, um dessen Wissensvorsprung zu nutzen beziehungsweise zu erwerben. Die Qualität der Handlungen des Arztes in der Diagnose und Therapie ist für den Patienten kaum bis gar nicht beobachtbar und verifizierbar. Der Arzt kann also seinen Handlungsspielraum zu seinen Gunsten ausnutzen und es liegt moral hazard auf Seiten des Arztes vor.

Um eine Behandlung so gut es geht erfolgreich abzuschließen, spielt neben den exogenen Umweltfaktoren, auch das Verhalten des Patienten eine große Rolle. In diesem Sinne werden auch vom Patienten Leistungen gefragt, die wiederum vom Arzt nicht beobachtbar beziehungsweise verifizierbar sind. Diese Leistung des Patienten ist sein behandlungsbegleitendes Verhalten, auch Compliance genannt. (Schneider 2002) Darunter fallen zum Beispiel seine Freizeitaktivitäten, sein Lebensstil, sein Trink- und Rauchverhalten, seine Ernährungsweise und das Befolgen von Therapievorschriften. Die Compliance kann sowohl gute als auch schlechte Auswirkungen auf den Genesungsprozess haben. Sie wird vom Patienten, im Gegensatz zu den exogenen Faktoren, selbst gewählt und durchgeführt. Der Patient nimmt daher sowohl die Rolle des Konsumenten als auch die eines Produktionsfaktors ein, so dass sich das Behandlungsergebnis als „joint product“ aus medizinischer Leistung und individueller Nutzung ergibt. (Wille et.al. 1991)

Außerdem ist die Compliance mit Kosten verbunden. Diese können monetärer Art in Form von Konsumausgaben, als auch nicht monetärer Art sein, so zum Beispiel Opportunitätskosten. Dies bedeutet also, dass der Patient vor seiner

Handlung unter Berücksichtigung seines Informationsstandes die zu erwartenden Folgen, Nutzen und Kosten abwägen muss. (Schneider 2002)

Der Patient kann in der Regel weder die Wirkung der Leistungen des Arztes genau einordnen, noch besitzt der Arzt genauere Informationen über das behandlungsbegleitende Verhalten (Compliance) des Patienten, das die gesundheitsfördernde Anstrengung des Patienten beschreibt, die er zusätzlich zu der medizinischen Behandlung durch den Arzt unternimmt. (Wille et.al. 1991, 27) Da es auf beiden Seiten zu moral hazard kommt, liegt hier der Fall von double moral hazard vor.

Die Frage ist nun, in wieweit Interdependenzen zwischen der ärztlichen Leistung und der Compliance des Patienten vorliegen. Dies hat Udo Schneider in seiner Veröffentlichung "Ärztliche Leistung und Compliance des Patienten – der Fall des Double Moral Hazard" (1999) untersucht. Auf dessen Erkenntnisse stützen sich die nächsten Abschnitte dieser Arbeit hauptsächlich.

6.2 Interdependenzen zwischen Leistung des Arztes und Verhalten des Patienten

Angenommen in der Versicherungsleistung des Patienten wären sämtliche ärztlichen Leistungen enthalten. Er kann also die Leistungen des Arztes zu Grenzkosten von null in Anspruch nehmen. Die Anreize, eigene Anstrengungen zu tätigen, um positiven Einfluss auf das Ergebnis zu erzielen sind somit nur sehr gering. Dies ist der sehr extreme Fall von einer substitutiven Beziehung zwischen den Leistungen der beiden Akteure, da die eigenen Anstrengungen mit Kosten verbunden wären und durch die kostenlosen medizinischen Leistungen ersetzt werden können. Ein Lösungsansatz hierbei wäre, wie so oft in Versicherungsfällen, die Einführung eines Selbstbehaltsbetrages. Dadurch wäre in den meisten Fällen eine höhere Compliance gesichert, da diese im Vergleich zur medizinischen Leistung und des Selbstbehaltes relativ billiger wird.

Es werden nun die Handlungen beider Parteien als Produktionsfaktoren in die Genesungsproduktion behandelt. Beide Akteure wählen simultan das Niveau ihrer Handlungen. Unter bestimmten Voraussetzungen lässt sich nachweisen, dass bei vollständiger Information und Kooperation von Arzt und Patient, sowohl die Menge der medizinischen Leistung als auch die Compliance über dem Niveau bei asymmetrischer Information liegen.

Klar ist, dass ein Zusammenspiel der beiden Faktoren, Leistung des Arztes und Compliance des Patienten, existiert, die Frage ist allerdings, zu welchen Teilen diese Faktoren Einfluss auf eine erfolgreiche Genesung nehmen und in welchem Verhältnis sie zueinander stehen.

Hierbei lässt sich zwischen strategischer Unabhängigkeit, strategischen Komplementen und strategischen Substituten unterscheiden.

Strategische Unabhängigkeit:

Die Erhöhung der Aktivität eines Akteurs hat keinen Einfluss auf die Grenzproduktivität des anderen.

Strategische Komplemente:

Die Erhöhung der Aktivität des einen, erhöht auch die Grenzproduktivität des anderen Akteurs.

Strategische Substitute:

Hierbei verringert sich die Grenzproduktivität des einen Akteurs, sobald der andere seine Aktivität erhöht.

Eine nähere Erläuterung zu dem strategischen Zusammenspiel der Einflussfaktoren wird unten aufgezeigt.

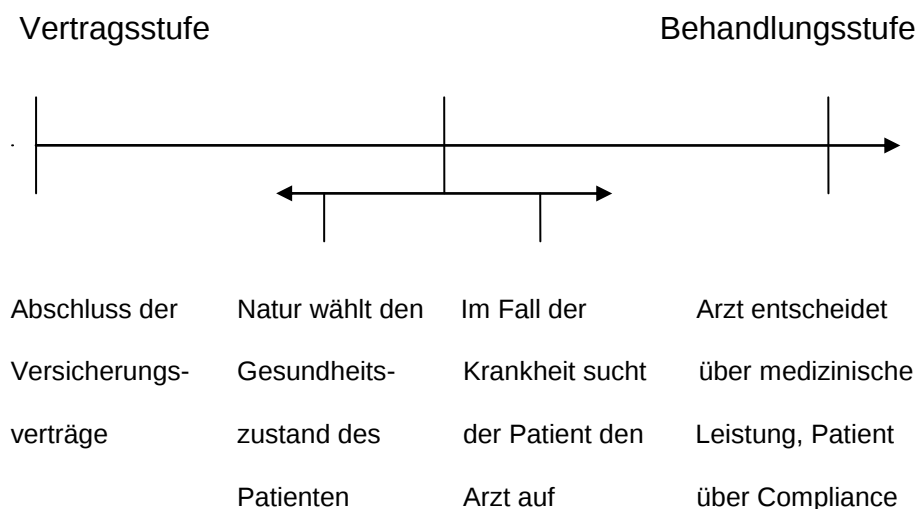
6.3 Die Modelle

In diesem Abschnitt sollen unterschiedliche Modelle aufgezeigt werden, die moral hazard im Gesundheitswesen näher betrachten und untersuchen.

6.3.1 Udo Schneider

In Schneiders Arbeit *“Ärztliche Leistung und Compliance des Patienten – der Fall des Double Moral Hazard”* (1999) werden zwei Stufen unterschieden, die Vertragsstufe und die Behandlungsstufe. Zu Beginn steht die Vertragsstufe, in der eine dritte Partei – die Versicherung – jeweils einen Vertrag mit dem Arzt und dem Patienten schließt. Danach ist der Patient versichert, erkrankt er nicht, kommt der Vertrag nicht zum Tragen, wenn allerdings die Natur als unabhängige exogene Variable Einfluss nimmt und entscheidet, dass der Patient erkrankt, kommt es, soweit sich der Patient zu einem Arztbesuch entschließt, zur Behandlungsstufe. In dieser Stufe entscheiden Arzt und Patient simultan über das Ausmaß und das Niveau ihrer Anstrengung.

Abbildung 3: Vertragsstufe und Behandlungsstufe



Quelle: Schneider, 1999

Im Folgenden wird die Behandlungsstufe näher beleuchtet und die Vertragsstufe, beziehungsweise die Tatsache, dass der Patient einen Arzt konsultiert, als bereits gegeben angenommen.

Parameter

y	Bruttoeinkommen von Patient
σ	Versicherungsprämie
m	medizinische Leistung
βm	prozentuale Selbstbeteiligung
	$0 < \beta < 1$
G	erreichter Gesundheitszustand
a	Compliance
$D(a)$	Aufwand der Compliance
$U(.)$	Nutzen aus Nettoeinkommen für den Patienten
	$U_1 = \text{Nutzen aus hohem Gesundheitszustand } G_1$
	$U_0 = \text{Nutzen aus niedrigem Gesundheitszustand } G_0$
w	Entlohnung in Form einer Fallpauschale für Arzt
	$w > 0$
δ	Kostenerstattung/Einheit der medizinischen Leistung
	$0 < \delta < 1$
$C(m)$	Aufwand der medizinischen Leistung
	$C' > 0, C'' > 0$

ε Intensität der Berufsethik

$$\varepsilon \in [0,1]$$

$E[.]$ Erwartungswert einer Auszahlung der Versicherung an den Arzt im Fall einer Konsultation

Das Nettoeinkommen des Patienten ist: $y - \sigma - \beta m$; σ muss er auf jeden Fall bezahlen, βm nur dann, wenn er auch tatsächlich einen Arztbesuch absolviert.

Da angenommen wird, dass der Patient mit besserem Gesundheitszustand auch ein höheres Einkommen erzielen kann und vice versa, besteht ein positiver Zusammenhang zwischen dem Bruttoeinkommen und dem erreichten Gesundheitszustand G . Das Bruttoeinkommen muss dabei allerdings hinreichend groß sein, damit das Nettoeinkommen > 0 bleibt.

Nachdem der Patient den Arzt aufgesucht hat, und von diesem behandelt wurde, kann sein Gesundheitszustand zwei Werte annehmen. G_0 für den Fall, dass der Zustand schlecht ist und G_1 für eine gute gesundheitliche Verfassung, dabei gilt $G_1 > G_0$. Mit einer Wahrscheinlichkeit von $p \in [0,1]$ wird G_1 erreicht, im Gegenzug dazu erfolgt G_0 mit $(1 - p)$. Die Wahrscheinlichkeit $p(a,m)$ hängt dabei von den zwei Faktoren Compliance a und medizinische Leistung m ab. Es wird dabei angenommen, dass eine Erhöhung einer der Inputs a und m auch die Wahrscheinlichkeit für einen guten Gesundheitszustand G_1 steigert. Da dies allerdings mit einer abnehmenden Rate passiert, weisen beide Inputs eine positive, wenn auch abnehmende Grenzproduktivität auf. Dabei ist in der Gesundheitsproduktion zu folgern, dass der Output G_1 ein Signal für hohe Compliance und hochwertige medizinische Leistung ist. Die Wahrscheinlichkeit dafür ist zwar sehr hoch, allerdings nicht gleich 1, weswegen das Ergebnis von G_1 mit vorhergegangenen kleinen a und/oder m nicht ganz auszuschließen ist. (Schneider 1998)

Je mehr also der Patient die Therapie unterstützt, desto höher ist die Genesungswahrscheinlichkeit, allerdings desto höher ist auch der damit

verbundene Aufwand. Somit ist die Erwartungsnutzenfunktion des Patienten additiv-separabel in Nutzen aus dem Nettoeinkommen und Aufwand der Compliance, das heißt, dass der Grad der Risikoaversion des Nutzens aus dem Einkommen U nicht mit der Anstrengung a variiert. (Macho-Stadler 1997, 19)

Durch die Risikoaversion des Patienten in seinem Nettoeinkommen ist dieser Teil der Nutzenfunktion streng konkav. Was allerdings den Aufwand für Compliance betrifft, so wird angenommen, dass der Disnutzen mit höherer Compliance weiter ansteigt und deswegen die Funktion $D(\cdot)$ konvex ist.

Der Erwartungsnutzen des Patienten ist demnach:

$$(1) \quad EU = p(a,m) U [y(G_1) - \sigma - \beta m] + \\ (1 - p(a,m)) U [y(G_0) - \sigma - \beta m] - D(a)$$

Demnach gilt:

$$U_1 = U[y(G_1) - \sigma - \beta m]$$

$$U_0 = U[y(G_0) - \sigma - \beta m]$$

Zusätzlich zu der Fallpauschale w erhält der Arzt auch eine Kostenerstattung δ pro Einheit der medizinischen Leistung. δ ist allerdings nur ein Anteil des wirklich bezahlten Betrages für die Leistung, der Arzt als Anbieter bekommt also nur eine Kostenbeteiligung. Die entstandenen Kosten für die medizinische Leistung werden in diesem Modell mit konstanten Grenzkosten, die auf eins normalisiert sind, angenommen.

Nach der Behandlung teilt der Arzt der dritten Partei – der Versicherung – die Kosten der Therapie mit.

Auch der Nutzen des Arztes ist additiv-separabel und zwar in Einkommen, Berufsethik und Aufwand.

Da der Arzt, im Gegensatz zum Patienten, in der Lage ist, sein Einkommensrisiko auf mehrere Patienten zu verteilen, ist der Arzt in seinem Einkommen risikoneutral.

Wegen des Faktors der Berufsethik, zu der sich der Arzt im Hippokratischen Eid verpflichtet, muss der erwartete Gesundheitszustand des Patienten positiv in die Erwartungsnutzenfunktion des Arztes integriert werden:

$$(2) \quad EV = (w + \delta m) + \varepsilon[p(a,m)G_1 + (1 - p(a,m))G_0] - C(m)$$

Eine mögliche Steigerung des guten Gesundheitszustandes wirkt nur dann umso stärker, je höher der Wert des Parameters ε ist.

Die außen stehende Partei der Versicherung erhält vom Patienten in jedem Fall eine Prämie σ und leistet bei Konsumation einer ärztlichen Leistung an den Arzt eine Zahlung in Höhe seines Honorars, abzüglich des Selbstbehalts des Behandelten:

$$(3) \quad \sigma = E[w + (\delta - \beta)m]$$

Weitere Annahmen im Zusammenhang mit der Versicherung sind, dass der erreichte Gesundheitszustand wohl dem Arzt und auch dem Patienten bekannt ist, nicht jedoch der Versicherung.

Kooperative Lösung

Sowohl die medizinische Leistung als auch die Anstrengung des Patienten beeinflussen den Wert des Arztbesuches für die jeweilige andere Partei – deswegen liegt hier double moral hazard vor. (Bhattacharyya 1995)

Es wird vom Fall der vollständigen Information ausgegangen, es wählen also beide Akteure ihre Handlungen derart, dass die Summe der zu Erwartungsnutzen maximiert wird.

First best Lösung:

$$(4) \quad \max_{(a,m)} S = p(a,m)[U_1 + \varepsilon G_1] + (1 - p(a,m)) [U_0 + \varepsilon G_0] + (\gamma + \delta m) - D(a) - C(m)$$

Bedingungen 1. Ordnung:

$$(5) \quad \partial S / \partial a: p_a[U_1 - U_0 + \varepsilon(G_1 - G_0)] = D'$$

$$(6) \quad \partial S / \partial m: p_m[U_1 - U_0 + \varepsilon(G_1 - G_0)] + \delta = C' + \beta[pU'_1 + (1-p)U'_0]$$

ad (5):

Die Summe aus den Erwartungsgrenznutzen von Arzt und Patient – resultierend aus einer Erhöhung der medizinischen Leistung oder der Compliance – muss gleich dem Grenzaufwand D' sein. Der Patient muss also den Level seiner Anstrengung so wählen, dass bei Erhöhung der Compliance der erwartete Grenznutzen bei Arzt und Patient im Optimum dem Grenzaufwand der Anstrengung des Patienten entspricht.

ad (6)

Hier steht auf der linken Seite der Gleichung wieder der gemeinsame erwartete Grenznutzen, der diesmal gleich den erwarteten Grenzkosten gleichen soll. Die Grenzkosten setzen sich zusammen aus dem Grenzaufwand der medizinischen Leistung beim Arzt und den Grenzkosten der Selbstbeteiligung auf Seiten des Patienten.

Beide Gleichungen sind abhängig von dem Level der medizinischen Leistung beziehungsweise der Compliance, sie spiegeln daher die Reaktionsfunktion des Arztes und des Patienten wider.

Nicht-kooperative Lösung

Da hier ein Fall von double moral hazard vorliegt, können weder Arzt noch Patient die Handlungen des jeweils anderen beobachten. Beide Akteure maximieren allerdings auch ihrer Seite ihren Nutzen, ohne dabei die Auswirkungen auf den jeweils anderen Vertragspartner zu berücksichtigen. Von dem Ergebnis – dem erreichten Gesundheitszustand – kann weder auf den unbekannten Level Compliance, noch dem der medizinischen Leistung geschlossen werden.

Das Maximierungsproblem des Patienten sieht demnach wie folgt aus:

$$(7) \max_{(a)} p(a,m)U_1 + (1 - p(a,m))U_0 - D(a)$$

Bedingung 1. Ordnung:

$$(8) p_a(a,m) (U_1 - U_0) = D'$$

Der Patient wählt also seine Compliance derart, dass der erwartete Grenznutzen der Anstrengung gleich dem Grenzaufwand ist.

Das Maximierungsproblem des Arztes wird mit einer Maximierung des Erwartungsnutzens über die medizinische Leistung derart ausgedrückt:

$$(9) \max_{(m)} (w + \delta m) + \varepsilon[p(a,m)G_1 + (1 - p(a,m))G_0] - C(m)$$

Bedingung 1. Ordnung:

$$(10) \quad \delta + p_m(a,m) \varepsilon(G_1 - G_0) = C'$$

Die optimale Menge an medizinischer Leistung ist für den Arzt erreicht, wenn der Grenznutzen der medizinischen Leistung, zusammengesetzt aus dem monetären Grenznutzen und dem Grenznutzen eines besseren Gesundheitszustandes (ethischer Einfluss), dem Grenzaufwand für die Bereitstellung der Leistung gleicht.

Vergleich kooperativer – nicht-kooperativer Lösung

Um einen Vergleich zwischen vollständiger und asymmetrischer Information anzustellen, werden im Folgenden jeweils die Bedingungen 1. Ordnung verglichen. Die Lösung im Optimum wird bei vollständiger Information durch (a^*, m^*) ausgedrückt, bei nicht-vollständiger Information durch (a^N, m^N) . Letztere Formulierung rührt daher, dass bei asymmetrischer Information die Handlungen eines Akteurs als Funktion der Aktionen des anderen Vertragspartners spezifiziert werden. (siehe (8) und (10)). Diese Reaktionsfunktionen ergeben ein Nash-Gleichgewicht für die nicht-kooperative Lösung (a^N, m^N) .

Durch den Vergleich der beiden unterschiedlichen Informationssituationen können Schlüsse über die Auswirkung der Informationsasymmetrie auf die Reaktionsfunktionen gezogen werden.

Auf der Seite des Patienten werden die Gleichungen (5) und (8) verglichen, die zeigen, dass der Aufwand bei symmetrischer Information genau dann höher ist, als im asymmetrischen Fall, wenn die Grenzkosten höher sind. Es gilt also genau dann

$$a^*(m) > a^N(m)$$

wenn aufgrund der Monotonie auch gilt, dass

$$D^*(a^*) > D^*(a^N)$$

Aus $a^*(m) > a^N(m)$ kann nun geschlossen werden, dass

$$(11) \quad p_a[U_1 - U_0 + \varepsilon(G_1 - G_0)] > p_a[U_1 - U_0]$$

Die linke Seite von (5) ist also ebenfalls größer als die linke Seite von (8), allerdings unter der Voraussetzung, dass $\varepsilon(G_1 - G_0) > 0$ ist. Da $\varepsilon > 0$ ist, ist diese Bedingung erfüllt.

Sobald der Patient erkennt, dass nicht nur er, sondern auch der Arzt von einem höheren Gesundheitszustand profitiert, erhöht er seine Anstrengungen für alle

Werte von m . Für eine gegebene medizinische Leistung wird der Patient also im kooperativen Fall mehr Anstrengung betreiben, als im nicht-kooperativen Fall, vorausgesetzt der Arzt handelt nach den Rahmenbedingungen seiner Berufsethik so, dass er tatsächlich von einem besseren Gesundheitszustand des Patienten profitiert.

Die Reaktionsfunktionen des Patienten verlaufen also bei symmetrischer Information für alle Werte von m über den Reaktionsfunktionen im Fall der asymmetrischen Information.

Vergleicht man auf Seiten des Arztes die kooperative mit der nicht-kooperativen Lösung, so muss man dafür die Gleichungen (6) und (10) heranziehen. Zuerst wird die Annahme getätigt, dass die medizinische Leistung im kooperativen Fall größer ist, als in der nicht-kooperativen Lösung

$$m^*(a) > m^N(a)$$

Daraus kann nun, gegeben die Monotonie der Funktion des Aufwandes, gefolgert werden, dass die Grenzkosten der Behandlung im Fall der asymmetrischen Information auch höher sein müssen. Wie schon beim Patienten, muss auch im Fall des Arztes die linke Seite von (6) abzüglich der Selbstbeteiligung größer sein als die linke Seite von (10)

$$(12) \quad p_m[U_1 - U_0 + \varepsilon(G_1 - G_0)] + \delta - \beta[pU'_1 + (1-p)U'_0] >$$

$$\delta + p_m \varepsilon(G_1 - G_0)$$

Um ex ante die optimale Menge an medizinischer Leistung aus der Sicht des Patienten zu bestimmen, dient diese Ungleichung.

$$p_m[U_1 - U_0] - \beta[pU'_1 + (1-p)U'_0] > 0$$

Wobei $p_m[U_1 - U_0]$ den Grenznutzen aus einer zusätzlichen Einheit von m bestimmt und $\beta[pU'_1 + (1-p)U'_0]$ die Grenzkosten durch die Selbstbeteiligung beschreibt.

Der Grenznutzen einer medizinischen Behandlung muss also die Grenzkosten für den Patienten übersteigen, damit die verordnete Menge an m im Fall der Kooperation (first-best case) über der bei asymmetrischer Information liegt.

Der Term ist allerdings nur dann immer größer null, wenn $\beta = 0$, der Patient also keinen Selbstbehalt zahlen muss. Dann wäre die medizinische Leistung bei Kooperation größer als im Fall der Nicht-Kooperation. Falls allerdings $\beta > 0$, kann es sein, dass die Reaktionsfunktionen sich schneiden. In diesem Fall wäre also m bei einer Informationsasymmetrie höher als bei Kooperation. Dies impliziert, dass dann die Behandlungsgrenzkosten höher wären als der Grenznutzen einer Behandlung. Die Selbstbeteiligung für die letzte Einheit medizinischer Leistung wäre dann höher als ihr Nutzen. Dieser Fall ist allerdings deswegen eher unrealistisch, da der Patient bei symmetrischer Information den Arzt dann nicht aufsuchen würde.

6.3.2 Paula González

In ihrer Arbeit „*Should Physicians' Dual Practice Be Limited? An Incentive Approach*“ (2003) hat Paula González moral hazard auf der Arzt Seite untersucht, wenn der Arzt neben einer öffentlichen Praxis auch noch eine Privatpraxis betreibt. Wann entscheidet sich der Arzt eine dem Signal entsprechende Behandlung zu verschreiben und wann wird er eher zu einer Überbehandlung tendieren, die auch für die Krankenkassen höhere Kosten bedeutet?

In ihrem Model gibt es drei Mitspieler, den Patienten, den Arzt und die Krankenkasse. In der ersten Phase bietet die Krankenkasse dem Arzt einen Vertrag an, er nimmt an oder lehnt ab. In der zweiten Phase stellt sich das Ausmaß der Krankheit des Patienten heraus. Er sucht einen Arzt auf, der mit einem gewissen Level an Anstrengung eine Diagnose stellt. Diese dient ihm als Signal über das Krankheitsausmaß. In Phase 3 entscheidet sich der Arzt dafür,

welche Behandlung er anwendet. Wird der Patient dadurch nicht gesund, verschreibt der Arzt eine andere Behandlung.

Parameter

s_n und s_h für niedrigen oder hohen Krankheitsgrad (beide gleich wahrscheinlich)

T_m und T_{st} für eine milde oder eine starke Behandlung

T_m kann s_n heilen, T_{st} kann s_n und s_h heilen, allerdings verursacht T_{st} einen gesundheitlichen Verlust von L beim Patienten, wenn er für eine leichte Krankheit eine starke Behandlung erhält. Sollte die Krankheit nicht in der ersten Runde geheilt werden und er erst nach der zweiten Behandlungsrunde gesund werden, wird beim Patienten ein Verlust von ℓ verursacht, da er zwei Runden durchmachen musste.

$e \in [0, 1]$ Level der Anstrengung des Arztes bei der Diagnose

s^e Signal der Diagnose über den Krankheitsgrad (je höher die Anstrengung, desto akkurater das Signal)

Das korrekte Signal wird mit folgender Wahrscheinlichkeit erreicht:

$$P(s^e = s_h | s_h) = P(s^e = s_n | s_n) = (1+e)/2$$

Durch die Anstrengung entsteht beim Arzt ein negativer Nutzen:

$$V(e) = k e^2/2 \quad k > 0$$

Ist der Arzt in der ersten Runde erfolgreich, so entsteht beim ihm ein Prestigefaktor, der sein Ansehen und somit auch Einkommen in seiner Privatpraxis steigert.

$\mu > 0$ Einfluss der ärztlichen Tätigkeit im öffentlichen Sektor auf seine Privatpraxis

c_m und c_{st}	Kosten für die jeweilige Behandlung, getragen von der Krankenkasse; $c_m < c_{st}$
C	Summe der finanziellen Kosten und des neg. Nutzens des Patienten $\approx$ Sozialkosten
w_m und w_{st}	Auszahlung an den Arzt für die jeweilige Behandlung
B	Bonus für Arzt, wenn er T_m verschreibt und der Patient nach der ersten Runde geheilt ist

Arzt wählt Behandlungsart T :

Die Wahrscheinlichkeit P , dass ein hoher Krankheitsgrad korrekt diagnostiziert wird ist:

$$P(s_h | s^e = s_h) = [P(s_h) P(s^e = s_h | s_h)] / [P(s_h) P(s^e = s_h | s_h) + P(s_n) P(s^e = s_h | s_n)] = (1+e) / 2$$

Die Wahrscheinlichkeit einen niedrigen Krankheitsgrad korrekt zu diagnostizieren liegt bei:

$$P(s_n | s^e = s_n) = (1+e) / 2$$

Daraus folgt:

$$P(s_n | s^e = s_h) = P(s_h | s^e = s_n) = (1-e) / 2$$

Wenn nun $s^e = s_n$ ist und der Arzt die T_m verschreibt, bekommt er w_m und sollte der Patient tatsächlich unter s_n leiden (mit $P(s_n | s^e = s_n)$), so bekommt der Arzt zusätzlich B dafür, dass er der Patient in nur einer Runde geheilt wurde, und μ weil er durch die erfolgreiche Behandlung Prestige für seine Privatpraxis gewinnt. Leidet der Patient unter s_h (mit $P(s_h | s^e = s_n)$), so bekommt der Arzt nur w_n . Verschreibt der Arzt T_{st} , bekommt er w_{st} und ebenfalls μ , da der Patient in Runde eins geheilt wurde.

Ist $s^e = s_h$ bekommt der Arzt $w_{st} + \mu$, wenn er T_{st} verschreibt. Bei T_m bekommt er w_m , falls die Diagnose korrekt war und $w_m + B + \mu$, falls die Diagnose falsch war (mit $P(s_n | s^e = s_h)$).

Bei

-) $s^e = s_n$ wird der Arzt immer dann T_m verschreiben, wenn

$$e \geq \ddot{e} = [2(w_{st} - w_m) + \mu - B] / (B + \mu) \text{ und sonst } T_{st}$$

-) $s^e = s_h$ wird der Arzt immer dann T_{st} verschreiben, wenn

$$e \geq \hat{e} = \{B - \mu - [2(w_{st} - w_m)]\} / (B + \mu) \text{ und sonst } T_m$$

Der Arzt sucht also die Art der Behandlung nach folgenden Kriterien aus:

- Falls $2(w_{st} - w_m) + \mu - B \geq 0$:
 -) falls $B < w_{st} - w_m$, wird der Arzt immer T_{st} verschreiben, ungeachtet des Signals
 -) falls $B \geq w_{st} - w_m$ und $e \geq \ddot{e}$ wird sich der Arzt für die bestgeeignete Behandlung entscheiden, bei $e \in [0, \ddot{e}]$ wird er ungeachtet des Signals immer T_{st} verschreiben
- Falls $B - \mu - [2(w_{st} - w_m)] \geq 0$:
 -) falls $\mu < w_m - w_{st}$, verschreibt der Arzt immer T_m ungeachtet des Signals
 -) falls $\mu \geq w_m - w_{st}$, und $e \geq \hat{e}$ wird der Arzt die bestgeeignete Behandlung wählen, bei $e \in [0, \hat{e}]$ wird er ungeachtet des Signals immer T_m verschreiben

Wenn seine Anstrengung niedrig ist, wird er immer eine Behandlung wählen, die unabhängig vom Signal der Diagnose ist, eher jedoch abhängig von der Art seines Vertrages ist.

Die Tatsache, dass ein Arzt nebenbei eine Privatpraxis betreibt, wird ihn eher dazu veranlassen T_{st} zu verschreiben unabhängig vom tatsächlichen Signal. $\ddot{e}$ ist steigend in μ , während $\hat{e}$ sinkend ist. Der Arzt ist daran interessiert den Patienten in der ersten Runde zu heilen, um an Prestige zu gewinnen.

Allerdings kann dies zu einer Überbehandlung führen. Der Arzt hat kann aus drei Behandlungsstrategien wählen: er kann immer T_{st} verschreiben, was zu C_{st} führt; immer T_m verschreiben, was C_m verursacht; oder immer die bestgeeignete Behandlung wählen, wodurch Sozialkosten von C^* entstehen. Beim Arzt entsteht je nach gewählter Behandlung Nutzen U_{st} , U_m , U^* .

Arzt wählt Level der Anstrengung e für die Diagnose:

- Bestgeeignete Behandlung T^*

Falls er sich für T^* entscheidet, so wird er seine Anstrengung so wählen, dass er U^* maximiert.

$$\mathcal{E} \equiv \max\{\hat{e}, \hat{e}\} = l2(w_{st} - w_m) + \mu + B l / (B + \mu)$$

Nur für $e \geq \mathcal{E}$ wäre es für den Arzt optimal sich für T^* zu entscheiden. Er ist mit folgendem Problem konfrontiert:

$$\max_{(e)} U^* = \frac{1}{2} [w_m + w_{st} + \mu + (1+e)/2 \cdot (B+\mu) - k e^2/2]$$

$$\text{s.t. } e \in [\mathcal{E}, 1]$$

$$\text{Lösung: } e^* = \min \{ \max \{ (B+\mu)/4k, \mathcal{E} \}, 1 \}$$

Entscheidet sich der Arzt für eine der anderen Strategien, so wird er jeweils immer T_{st} oder T_m wählen ungeachtet des Signals, wodurch sein optimaler Anstrengungslevel ein Minimum ist, da ein höherer Level an Anstrengung keinen Einfluss auf sein Einkommen hat und nur höhere Kosten für ihn verursacht.

Daraus folgt, dass der Arzt folgenden Anstrengungslevel wählt:

- T_{st} Strategie $\Rightarrow e_{st}=0$ und $U_{st} = w_{st} + \mu$
- T_m Strategie $\Rightarrow e_m=0$ und $U_m = w_m + \frac{1}{2} (\mu + B)$
- T^* Strategie $\Rightarrow e^* = \min \{ \max \{ (B+\mu)/4k, \mathcal{E} \}, 1 \}$ und

$$U^* = \frac{1}{2} [w_m + w_{st} + \mu + (1+e^*)/2 \cdot (B+\mu) - k e^{*2}/2]$$

Er wählt also die

- T_{st} Strategie, wenn
$$e^* [(B+\mu)/2 - ke^*] \leq w_{st} - w_m + (\mu-B)/2$$
- T_m Strategie, wenn
$$e^* [(B+\mu)/2 - ke^*] \leq w_m - w_{st} + (B-\mu)/2$$
- T^* Strategie, wenn
$$e^* [(B+\mu)/2 - ke^*] \geq |w_{st} - w_m + (\mu-B)/2|$$

mit $e^* = \min \{(B+\mu)/4k, 1\}$

Je höher w_{st} ist, desto eher wird der Arzt immer T_{st} verschreiben, je höher w_m ist, desto eher wird er die T_m Strategie wählen. Wenn B einen hohen Wert annimmt, wird der Arzt motiviert sein, die bestmögliche Behandlung zu einem gegebenen Signal zu wählen (T^*), erhält allerdings B einen zu hohen Wert, so kann es sein, dass sich der Arzt für die T_m Strategie entscheidet und immer die mildere Behandlung anwendet.

Dadurch, dass der Arzt neben seiner öffentlichen Arbeit auch als Privatarzt praktiziert, hat er auf Grund von μ weniger Anreiz die T_m Strategie zu wählen. Falls $e^* = (B+\mu)/4k$ ist, hat er einen außerdem einen höheren Anreiz die T_{st} Strategie zu wählen.

Wenn $e^* = (B+\mu)/4k < 1$, wenn also die Diagnose nicht perfekt ist, dann hat der Arzt eine höhere Neigung dazu „überzubehandeln“. Bei $e^* = 1$ ist die Entscheidung über die Strategie unabhängig von μ , da der Patient sowohl mit T^* als auch mit T_{st} immer nach der ersten Runde geheilt ist.

Dieses Modell von Paula González (2003) zeigt also, dass moral hazard auf der Seite des Arztes vorkommt. Der Arzt wählt die Art der Behandlung und den Anstrengungslevel für die Diagnose nicht immer so, wie es optimal für den Patienten wäre.

6.3.3 David Bardey und Romain Lesur

David Bardey und Romain Lesur haben in ihrer Arbeit „*Optimal Regulation of Health System with Induced Demand and Ex post Moral Hazard*“ (2006) unter anderem untersucht, wie im Gesundheitswesen eine Behandlung, die über das nötige Ausmaß hinausgeht, verursacht wird und mit welchen Entlohnungsstrategien man diese „Überbehandlung“ verhindern kann. Ursachen für Überbehandlungen ist zum Beispiel auf der Patientenseite die Kostendeckung für Behandlungen durch Versicherungen, welche die Nachfrage steigern und moral hazard verursachen. Bardey und Lesur machen den Arzt keinen reinen Gewinnmaximierer, sondern machen die Nutzenfunktion des Arztes sowohl von dem finanziellen Nutzen als auch von dem Wohlbefinden des Patienten abhängig. Der Arzt ist also als Anbieter gleichzeitig auch Nachfrager und profitiert von einem gewissen Maß an Selbstlosigkeit, da sein Ziel auch der gesunde Patient ist.

Parameter

w	Einkommen des Patienten
y	Anfangsvermögen des Patienten
h	konsumierte medizinische Einheiten
θ	Gesundheitszustand
$c(h)$	Selbstbeitrag
π	Premium, das der Versicherungsnehmer zahlt

Der Versicherungsnehmer

Er wird zum Patienten, sobald er krank wird und hat die Nutzenfunktion $v(w, h)$, welche in w steigend und konkav ist, zweiteres weist auf seine Risikoaversion hin ($v^w > 0$ und $v^{ww} < 0$). Außerdem erfüllt diese Funktion auch $\lim_{w \rightarrow 0} v'(w) = \infty$ und $\lim_{w \rightarrow \infty} v'(w) = 0$. Die Nutzenfunktion ist auch steigend und konkav in h ($v^h > 0$ und $v^{hh} < 0$).

Angenommen v ist additiv separabel ($v^{wh} = 0$), so ist der Nutzenlevel für den Versicherungsnehmer:

$$(1) v(\theta) = v(w, h - \theta)$$

$$(2) w = y - \pi - c(h)$$

Im ersten Schritt wählt der Patient seinen Bedarf und Konsum an medizinischer Leistung, daraus folgt ex post moral hazard:

$$(3) c'(h) v^w(y - \pi - c(h), h - \theta) = v^h(y - \pi - c(h), h - \theta)$$

In (3) stehen also die Grenzkosten $c(h)$ eines Arztbesuches dem Nutzen durch den verbesserten Gesundheitszustands v^h gegenüber.

Der Arzt

Die Nutzenfunktion des Arztes setzt sich zusammen aus dem eigenen finanziellen Nutzen und dem Wohlbefinden der Patienten und sieht folgendermaßen aus:

$$V(h) = \varphi \cdot T(h) + (1 - \varphi) \cdot v(y - \pi - c(h), h - \theta)$$

Wobei $T(h)$ für den finanziellen Vorteil abhängig von der Anzahl der konsumierten medizinischen Einheiten steht, gewichtet mit dem Koeffizienten φ . Der Nutzen des Patienten wird mit $(1 - \varphi)$ gewichtet.

FOC für das Maximierungsproblem des Arztes ist:

$$c'(h)v^w(.) = v^h(.) + \varphi/(1-\varphi) T'(h)$$

SOC $V''(h) < 0$ hält nur, wenn $\partial h/\partial \theta > 0$

Der Arzt wählt also h so, dass der Grenznutzen, bestehend aus $T'(h)$ und $(1-\varphi)v^h(.)$ ist gleich den Grenzkosten. Daran sieht man, dass die Überversorgung durch ex post moral hazard Verhalten des Patienten verursacht wird, allerdings auch dadurch, dass er Arzt in seiner Funktion als Anbieter auch selbst Nachfrage kreiert. Wäre $\varphi=0$, so würde der Arzt komplett selbstlos handeln. Bei einem Maximum von $\varphi=1$, versucht der Arzt einzig sein eigenes Einkommen zu maximieren. Eine Lösung gegen eine Überversorgung ist die Einführung eines Selbstbeitrags.

First-best Analyse:

Das Zielfunktion für den Versicherungsträger ist:

$$\max_{(c(.), T(.))} \int_{\theta} [v(y-\pi-c(h), h-\theta) f(\theta) d\theta]$$

Die drei Bedingungen sind wie folgt:

- Participation constraint Arzt

$$(4) \int_{\theta} T(h(\theta)) f(\theta) d\theta \geq w_0$$

$$w_0 \geq 0$$

- Budget constraint Versicherungsträger

$$(5) B = \int_{\theta} (\pi - T(h(\theta)) - h(\theta) + c(h(\theta))) f(\theta) d\theta \geq 0$$

- Verhalten des Arztes

$$(6) c'(h(\theta)) v^w(y-\pi-c(h(\theta))) = v^h(h(\theta) - \theta) + \varphi/(1-\varphi) T'(h(\theta))$$

Um diese Model als moral hazard Modell zu lösen, wird der Versicherungsträger als Prinzipal und der Arzt als Agent bezeichnet. Jeglicher Vertrag $(c(h), T(h))$ des Versicherungsträgers kann als $(c(\theta), T(\theta), h(\theta))$ formuliert werden.

Nach Langrange mit $\alpha = (4)$ und $\mu = (5)$ erhält man durch die FOC

$$v_w(y - \pi - c(\theta)) = v_h(h(\theta) - \theta) = \mu = \alpha$$

Der Selbstbeitrag von $c(\theta)$ ist konstant und der Endgesundheitszustand ist unabhängig von der Intensität der Krankheit: $h - \theta = H_0$. Gleichung (6) führt also zu $T'(h) = - (1-\phi)/\phi v_h(H_0)$ mit first best $c'(h) = 0$ ergibt das ein Entlohnungsmodell für den Arzt

$$T(h) = w_0 + (1-\phi)/\phi v_h(H_0) [H_0 + \int_{\theta} [\theta f(\theta) d\theta - h]$$

Unter der first-best Lösung ist der Nutzen des Versicherungsnehmers/Patienten konstant. Die first-best Lösung kann angewandt werden mit Kostenaufteilung nur auf der Angebotsseite. Die Steigung des Nutzen des Arztes ist $-(1-\phi)/\phi v_h(H_0)$ und ist steigend in ϕ . Je mehr sich der Arzt also selbstlos verhält, desto wichtiger wird Kostenteilung auf der Angebotsseite, um eine Überbehandlung zu verhindern. Allerdings kann h nicht immer perfekt beobachtet werden, weswegen es zu Kollusionen und Absprachen zwischen dem Arzt und dem Patienten geben, um diese zu verhindern muss dem eine Schranke in Form von der folgenden Bedingung vorgeschoben werden: $T'(h) \geq -1$

6.3.4 Ching-To Albert Ma und Thomas G. McGuire

Ma und McGuire haben in ihrem Artikel „Optimal Health Insurance and Provider Payment“ (1997) beim Arzt eine Art ethische Komponente untersucht. Bei der Interaktion zwischen Arzt und Patient wird medizinische Versorgung produziert durch einen Input des Patienten (Anzahl der Arztbesuche) und einem Input des Arztes (die Höhe seiner Anstrengung). Beide dieser Inputs sind ex ante nicht direkt von der Versicherung beobachtbar. Unter die Anstrengung des Arztes

fällt gute Kommunikation, die richtige Behandlung für eine spezifische Krankheit zu finden und Zeit. Allerdings wird der Arzt pro Patient entlohnt und obwohl mehr Zeit und Anstrengung zu einer besseren Heilungswahrscheinlichkeit führt, verursacht allerdings auch teurere Arztbesuche. Die Gesundheitsverbesserung, der Output, kann vertraglich nicht bestimmt werden. Es wird angenommen, dass die Feststellung der Anzahl der Behandlungseinheiten nur auf Grund von Patienten- und Arztberichten erfolgt. Der Arzt kann also nur anhand von der berichteten Anzahl bezahlt werden, nicht auf Grund seiner Anstrengung. In Phase 1 wählt die Versicherung ein Entlohnungssystem. In Phase 2 entscheidet die Natur darüber, ob der Patient krank wird (mit Wahrscheinlichkeit p) oder nicht. In Phase 3 wählt der Arzt den Level seiner Anstrengung e . In Phase 4 wählt der Patient auf Grund des Anstrengungslevels des Arztes die Anzahl der Behandlungseinheiten T , der Arzt kann die Nachfrage mit seinem Level an Anstrengung beeinflussen. In Phase 5 gibt es ein Subgame von Arzt und Patient über die Anzahl der konsumierten Einheiten, die der Versicherung gemeldet werden. Dabei schlägt der Arzt vor, dass die Menge T^r gemeldet wird, die nicht mit der tatsächlichen Menge T übereinstimmen muss. Der Patient kann T^r zustimmen oder nicht, im zweiten Fall würde T gemeldet werden. Der Arzt kann eine wahre Meldung erzwingen, indem er $T^r = T$ vorschlägt, der Patient kann im Gegenzug die Wahrheit erzwingen, indem er einfach dem Vorschlag des Arztes nicht zustimmt, in welchem Fall T der Versicherung gemeldet wird.

Der Patient hat ein Anfangseinkommen von w und wird mit der Wahrscheinlichkeit p krank. Im Falle einer Krankheit erleidet er einen negativen Gesundheitsschock s . Mit der strikt konkaven Funktion $F(T, e)$ kann die Gesundheit wieder hergestellt werden. F ist steigend in T und e . Der Patient zahlt der Versicherung einen Selbstbeitrag von β . Die Kosten c für die Behandlung sind konstant und es gilt: $0 \leq \beta \leq c$. Bevor der Krankheitszustand bekannt ist, muss der Patient der Versicherung außerdem eine Prämie $\alpha \geq 0$ zahlen. Er hat also, wenn er gesund ist ein Einkommen von $w - \alpha$. Wird er krank, so besteht sein Einkommen aus: $w - \alpha - \beta - T^r$. Er handelt außerdem risikoavers und seine Nutzenfunktion lautet:

$$(1) EU = pU(w - \alpha - \beta T^r - s + F(T, e)) + (1-p) U(w - \alpha)$$

Der Arzt ist risikoneutral. Er kann entweder pro Patient bezahlt werden mit φ oder für jede gemeldete Behandlungseinheit, wofür er $d+c$ bekommt. Wenn die tatsächlich konsumierte Einheiten T sind und der Arzt T^r meldet, dann verdient er $\varphi + (d+c)T^r - cT$. Er muss außerdem die Kosten der Anstrengung e tragen $G(e)$, wobei $G(0) > 0$. Seine Nutzenfunktion lautet:

$$(2) V = p[\varphi + (d+c)T^r - cT - G(e)]$$

Die Versicherung maximiert den Nutzen des Patienten, wobei die Prämie gleich dem erwarteten Wert der Zahlung der Versicherung an den Arzt ist.

$$(3) \alpha = p[\varphi + (d+c-\beta)T^r]$$

Da der Patient für jede Behandlung einen Selbstbehalt von $\beta > 0$ bezahlt, wird er nie einem Bericht zustimmen, der mehr als die tatsächlich konsumierten Einheiten angibt; er wird allerdings sehr wohl zustimmen, wenn der Arzt weniger Einheiten als T angeben möchte. Wenn der Arzt $(d+c) \geq 0$ erhält, wird er nie weniger als die tatsächlich konsumierten Besuche angeben. Er kann allerdings wegen des oben beschriebenen Verhaltens des Patienten nie mehr angeben. Wenn also $(d+c) \geq 0$ wird im Optimum immer die Wahrheit berichtet werden. Wenn $(d+c) < 0$ und der Arzt einen positiven Betrag pro Einheit an die Versicherung zahlen muss, dann wird er versucht sein, $T^r < T$ zu berichten, was dann auch im Sinne des Patienten ist.

Im nächsten Schritt wählt erst der Arzt den Level von e und dann der Patient die Quantität der medizinischen Leistung.

$$\max U(w - \alpha - \beta T - s + F(T, e))$$

$$(4) \text{ FOC: } \beta = F_T(T, e)$$

Da sowohl der Gesundheitsschock als auch die Produktion der Behandlung in Geldeinheiten dargestellt werden kann, hat die Prämie α keinen Einfluss auf die

Nachfrage. Der Arzt kann jedoch mit e die Nachfrage des Patienten beeinflussen, wobei das Einkommen des Arztes nach (4) wie folgt aussieht:

$$(5) \quad dT/de = - (F_{Te}/F_{TT})$$

Dies misst die Reaktion des Patienten auf die Wahl der Anstrengung des Arztes. Wenn e und T Substitute sind, $F_{Te} < 0$, dann ist ein niedrigeres e nötig, um den Patienten dazu zu bringen mehr T nachzufragen. Bei $F_{Te} > 0$ verhalten sich T und e wie Komplemente, der Patient wird also mehr T nachfragen, wenn der Arzt sein e erhöht.

Für ein Programm A mit $0 \leq \beta \leq c$ und $d \geq -c$ müssen e und T so gewählt werden, um Folgendes zu maximieren:

$$(6) \quad \varphi + dT - G(e); \text{ s.t. } \beta = F_T(T, e)$$

Das zukünftige Einkommen des Arztes $\varphi \geq 0$ beeinflusst nicht das angebotene e . Nur β und φ sind Parameter im Programm A . Angenommen:

$$\Omega = \{(T, e): \text{es gibt } (\beta, \varphi), \text{ mit}$$

$$0 \leq \beta \leq c \text{ und } d \geq -c, \text{ für die } (T, e)$$

Programm A lösen, wenn $(\beta, d)^t$ gegeben sind}

Das Set Ω = implementable set

Die FOC Ableitung der Zielfunktion nach e ergibt also:

$$(7) \quad d(dT/de) - G'(e) = -(F_{Te}(T, e)/F_{TT}(T, e))d - G'(e)$$

Wenn man nun eine Kombination von (T, e) wählt, die im Set Ω enthalten ist und es sich um Substitute handelt, dann ist e nur dann über dem Minimum, wenn ein Kostenbeitrag auf der Angebotsseite anfällt: $d < 0$. Handelt es sich um Komplemente, dann ist e nur dann über dem Minimum, wenn das Entlohnungssystem des Arztes keinen Kostenbeitrag auf der Angebotsseite enthält $d > 0$. Das Ausmaß d ist nur dann relevant, wenn die Anstrengung nicht

beobachtbar ist. Der Arzt kann nicht für mehr als die tatsächlichen Kosten der Behandlung aufkommen $d \geq -c$, wenn die wahre Anzahl der Behandlungen bei der Versicherung gemeldet wird. Im Fall von Substituten, wenn die FOC ausreichend ist um das Programm A zu lösen, dann ist die komparative Statik nach d gegeben durch:

$$(8) \quad \partial T / \partial d = - F_{Te}(T, e)^2 / H < 0$$

$$\partial e / \partial d = - [F_{Te}(T, e) F_{TT}(T, e)] / H < 0$$

Wenn e und T Substitute sind, bedarf ein höherer Level von e , dass der Wert von d immer weiter sinkt.

Einführung einer Ethik-Bedingung:

Der Arzt soll sich zusätzlich ethisch verhalten, weswegen eine Ethik-Bedingung eingeführt wird:

$$(9) \quad F(T, e) \geq \bar{F}, \text{ wobei } \bar{F} \text{ eine Konstante ist}$$

In Phase 3 muss der Arzt also nun seine Anstrengung so wählen, dass es zu einem Gesundheitsvorteil von $\bar{F}$ kommt, wenn der Patient seine Nachfrage wählt. Wenn β genügend niedrig ist, kann der Patient bereits mit einem gewissen Level von d einen Vorteil erreichen, der im Optimum in Phase 3 höher als $\bar{F}$ ist; (9) ist also nicht immer bindend.

Angenommen ohne Ethik-Bedingung wäre das Optimum gegeben durch α^* , β^* , d^* , e^* und T^* . EU^* wäre der erwartete Nutzen des Patienten im Gleichgewicht. Nun wird folgende Ethik-Bedingung eingeführt:

$$F(T, e) \geq F(T^*, e^*) \equiv F^*$$

Mit der neuen Ethik Beschränkung wird der Erwartungsnutzen des Patienten im Gleichgewicht mindestens EU^* sein. Wenn die Steigung von $(F(T, e) = \bar{F})$ höher ist als von der Reaktionsfunktion $F_T(T, e) = \beta$, dann wird der optimale

Erwartungsnutzen des Patienten durch den Anstieg von β zu β^* steigen. Ein höherer Selbstbeitrag führt allerdings zu einer geringeren Anzahl von nachgefragten Behandlungen. Um also den Gesundheitszustand F^* zu erhalten, muss der Arzt einen höheren Level von Anstrengung wählen.

Wenn $F(T,e) \geq F(T^*,e^*) \equiv F^*$ gilt, muss der Erwartungsnutzen des Patienten im Gleichgewicht höher sein als ohne Ethik-Bedingung (EU^*). und β muss über β^* steigen. Die Versicherung kann eine Ethik Beschränkung einführen, indem sie β ausreichend erhöht, womit die Nachfrage nach Behandlungen auf der Patientenseite sinkt und der Arzt seinen Anstrengungslevel so erhöhen muss, dass die Ethik-Bedingung wiederum erfüllt wird. Als Ergebnis erreicht der Patient einen höheren Erwartungsnutzen.

6.3.5 Kenneth L. Leonard und Joshua Zivin

Kenneth L. Leonard und Joshua Zivin haben „*How To Compensate Physicians When Both Patient and Physician Effort are Unobservable*“ (2000) festgestellt, dass sich für verschiedene Krankheitsarten unterschiedlich gute Bezahlungsformen für den Arzt ergeben. Sind die Anstrengungen von Arzt und Patient Komplemente, so ist es am besten den Arzt Ergebnis-orientiert zu bezahlen. Ist das Komplementärverhalten zwischen den Anstrengungslevels der beiden Parteien niedrig, so sollte der Arzt je nach gewählter Anstrengung bezahlt werden. Als Beispiel wird eine Operation genannt, bei der der Patient mit seiner Anstrengung/Compliance kaum Einfluss auf die Genesung hat, hier sollte der Arzt nach seiner eigenen Anstrengung bezahlt werden. Bei Rückenschmerzen hingegen, bei denen die Compliance des Patienten sehr wohl einen Beitrag zur Genesung beitragen kann, sollte der Arzt entsprechend dem Ergebnis entlohnt werden.

Wenn H die Gesundheit ist und ΔH für die Veränderung der Gesundheit steht, dann kann entweder diese Veränderung entweder gut oder schlecht sein. Entweder gibt es eine totale Genesung, dann ist $\Delta H = h^g$ oder es gibt keine

Veränderung im Gesundheitszustand nach einer Behandlung, dann gilt $\Delta H = h^S$. Wird der Patient richtig diagnostiziert, bekommt er die richtige Behandlung und befolgt er diese auch, dann wird er mit der Wahrscheinlichkeit q^* wieder gesund. Ist die Diagnose falsch, so wird er nur mit der Wahrscheinlichkeit q^0 gesund, die Verteilung ist Φ^* für die richtige Diagnose und Φ^0 für die falsche. Eine bessere Technik führt zu einer höheren Wahrscheinlichkeit für eine richtige Diagnose: e ($0 \leq e \leq 1$) mit

$$(1) \Delta H \sim e \Phi^* + (1 - e) \Phi^0$$

Die beste Technik $e=1$ führt mit q^* zu einer Genesung und vice versa. Technik e setzt sich zusammen aus sowohl der Compliance des Patienten und seinen Fähigkeiten als auch der Anstrengung und der Kompetenz des Arztes. Der Patient kann e nicht beobachten, sondern nur das Ergebnis h^g oder h^s .

Der Nutzen der Gesundheit ist $U = (H, I(H), c(p))$ und demnach ist der Erwartungswert:

$$(2) EU = eq^* U^g + e(1-q^*)U^s + (1-e)q^0U^g + (1-e)(1-q^*)U^s$$

$$U^g = U[h^g, (I(h^g) - C), c(p)]$$

$$U^s = U[h^s, (I(h^s) - C), c(p)]$$

$I(H)$ steht für das Einkommenspotential im gesunden Zustand, p ist die Compliance des Patienten und $c(p)$ sind die Kosten für diese Compliance. C sind die gesamten Kosten eines Arztbesuches.

$$U = U'[H, I(H) - C - c(p)]$$

Wenn $U'[h^g, I(h^g)] = U^{g'}$ und $U'[h^s, I(h^s)] = U^{s'}$, dann ergibt sich

$$(3) EU = (e(q^* - q^0) + q^0) U^{g'} + (1 - q^0 + e(q^0 - q^*)) U^{s'} - C - c(p)$$

$$(4) \Delta EU = e(q^* - q^0) (U^{g'} - U^{s'}) - C - c(p)$$

Annahmen:

- $U^S = 0$
- $h^g \cdot \omega^g$ ist konstant und verursacht den Gesundheitsnutzen (ω sind die Opportunitätskosten der gesunden Phase; $w = \omega_g + \omega$).

$$\Delta EU = e(q^* - q^0) w h^g - C - c(p)$$

$e = h/h^g$ und hat eine Schwankung zwischen 0 und 1 und h zwischen 0 und h^g .

$$\Delta EU = e(q^* - q^0) w h - C - c(p) \Rightarrow \text{im Folgenden } \Delta EU = U$$

Gesundheitsproduktionsfunktion

$h = \pi p^\alpha m^\beta$, wobei π der Produktionsfaktor ist und p für die Compliance des Patienten steht, m die Anstrengung des Arztes darstellt und β die Elastizität des Ergebnisses abhängig von m .

$$0 < \alpha < 1 \text{ und } 0 < \beta < 1 \text{ und } 0 < \alpha + \beta < 1$$

A steht für den Wert der erhaltenen Behandlung ($= \pi w (q^* - q^0)$)

Der negative Nutzen der Anstrengung ist eine lineare Funktion von Anstrengung und D bezeichnet den Koeffizienten für die Anstrengung des Arztes.

Gesundheitsproduktion unter symmetrischer Information

Es gilt

$$(7) A p^\alpha m^\beta - p - D m$$

nach p und m zu maximieren

$$(8a) \alpha A p^{\alpha-1} m^\beta - 1 = 0$$

$$(8b) \beta A p^\alpha m^{\beta-1} - D = 0$$

Die Grenzproduktivität jedes Inputs ist gleich der Grenzkosten des Inputs.

Mit der FOC wird der optimale Level an Anstrengung/Compliance für Arzt und Patient ermittelt (sy steht dabei für die Lösung mit symmetrischer Information).

$$(9a) p_{sy}^* = \alpha(A\alpha^\alpha (\beta/D)^\beta)^{1/1-\alpha-\beta}$$

$$(9b) m_{sy}^* = (\beta/D (A\alpha^\alpha (\beta/D)^\beta)^{1/1-\alpha-\beta}$$

$$(10a) U_{sy} = (1-\alpha) (A\alpha^\alpha (\beta/D)^\beta)^{1/1-\alpha-\beta}$$

$$(10b) W_{sy} = (1-\alpha-\beta) (A\alpha^\alpha (\beta/D)^\beta)^{1/1-\alpha-\beta}$$

Bezahlung nach Anstrengungslevel:

Eine dritte Partei, der Sozialplaner, kann die Anstrengung des Arztes beobachten und wählen, er kann allerdings nicht die Compliance auf Seiten des Patienten beobachten. Wenn die Anstrengungen des Arztes und des Patienten nun Komplemente sind, hängt der optimale Level der Anstrengung des Arztes von der Compliance ab, weswegen der Sozialplaner den Arzt nicht zum Optimum zwingen kann.

Der Sozialplaner maximiert den Wohlstand mit $p=P$

$$(11) \quad \max_{(m)} Ap^{-\alpha}m^\beta - P - Dm$$

$$(12) \quad FOC \Rightarrow \beta Ap^{-\alpha}m^{\beta-1} - D = 0$$

Die Grenzproduktivität der medizinischen Anstrengung, P , ist gleich der der Grenzkosten der Anstrengung. Der Patient reagiert auf die Wahl des Arztes von P mit seiner Compliance-Wahl. Die Reaktionsfunktion leitet sich von dem Maximierungsproblem abgeleitet nach p ab, wobei E für die Anstrengungsbasierende Entlohnung steht.

$$(13) \quad \alpha Ap^{\alpha-1}m^\beta - 1 = 0$$

Die Kombination der beiden FOC führt zu

$$(14a) \quad p^*_E = (A\alpha^{1-\beta}(\beta/D)^{\beta}p^{-\alpha\beta})^{1/(1-\beta)(1-\alpha)}$$

$$(14b) \quad m^*_E = (A(\beta/d)p^{-\alpha})^{1/(1-\beta)}$$

$$(14c) \quad U_E = U_{sy}(P/p^*_{sy})^{\alpha\beta/(1-\alpha)(1-\beta)}$$

$$(14d) \quad W_E = (1-\alpha-\beta)(P/p^*_{sy})^{\alpha(1-\alpha-\beta)/(1-\alpha)(1-\beta)}(A\alpha\alpha(\beta/D)^{\beta})^{1/(1-\alpha-\beta)}(P/p^*_{sy})^{\alpha\beta/(1-\alpha)(1-\beta)}$$

Unter Anstrengungsbasierender Entlohnung ist der Wohlstand im Vergleich zur symmetrischen Information gleich, wenn $P = p^*_{sy}$. Wenn $P < p^*_{sy}$ oder $P > p^*_{sy}$ dann ist der Wohlstand unter E geringer als unter symmetrischer Information.

Ergebnisorientierte Bezahlung:

Der Arzt und der Patient erhalten eine Zahlung als Funktion des Ergebnisses. Der Anteil des Patienten wird mit s_p bezeichnet, der Anteil des Mediziners mit s_m . Der Sozialplaner maximiert den sozialen Wohlstand, indem er die Ausmaße der Anteile wählt. Die perfekte Lösung entsteht, wenn beide Anteile gleich 1 sind und sowohl der Arzt als auch der Patient volle Anreize erhalten. Der Sozialplaner wird die Anteile so hoch wie möglich ansetzen, auch wenn sie normalerweise nicht gleich 1 sind.

Der Patient maximiert seinen Nutzen

$$(15) \quad s_p A p^{\alpha} m^{\beta} - p$$

Der Arzt maximiert ebenfalls seinen Nutzen

$$(16) \quad s_m A p^{\alpha} m^{\beta} - D m$$

Es wird ein Nash Gleichgewicht erreicht, wenn die Wahl des Anstrengungslevels bei beiden Mitspielern genau so hoch ist wie die Erwartung des anderen. Der Sozialplaner muss m nicht unbedingt wissen und kann p auf keinen Fall beobachten. Im Gleichgewicht sehen die Anstrengungslevel von

Patient und Arzt, sowie der Patientennutzen und der Wohlstand wie folgt aus, wobei O für die Ergebnisorientierte Bezahlung steht.

$$(17a) p_o^* = s_p \alpha (A \alpha^\alpha (\beta/D)^\beta)^{1/(1-\alpha-\beta)} (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)}$$

$$(17b) m_o^* = s_m (\beta/D) (A \alpha^\alpha (\beta/D)^\beta)^{1/(1-\alpha-\beta)} (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)}$$

$$(17c) U_o = U_{sy} s_p (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)}$$

$$(17d) W_o = (1 - s_p \alpha - s_m \beta) (A \alpha^\alpha (\beta/D)^\beta)^{1/(1-\alpha-\beta)} (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)}$$

$s_p (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)}$ ist 1, wenn s_p und s_m beide gleich 1 sind. Beide Mitspieler erhalten dann volle Entlohnung für ihre Anstrengung. Dies hält im Fall der symmetrischen Information.

Ist $s_m < 1$ oder $s_p < 1$ oder beide, dann ist $s_p (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)} < 1$.

Entweder der Patient oder der Arzt, oder beide, erhalten nicht die volle Entlohnung für ihre Anstrengung und werden den Level der Anstrengung senken.

Unter Ergebnisabhängiger Entlohnung sind p_o^* und W_o geringer als unter symmetrischer Information.

Anstrengungsabhängige versus Ergebnisabhängige Bezahlung:

Verglichen wird der Nutzen des Patienten:

$$(18) \quad U_o - U_E = U_{sy} (s_p (s_p^\alpha s_m^\beta)^{1/(1-\alpha-\beta)} - (P/p_{sy}^*)^{\alpha\beta/(1-\alpha)(1-\beta)})$$

Wenn P ausreichend klein ist, dann ist $U_o > U_E$. Wenn P allerdings groß ist, dann ist $U_o < U_E$. P ist a priori nicht bekannt. Angenommen $\acute{p}$ ist der Wert von P , wenn der Patientennutzen unter beiden Entlohnungsformen gleich ist.

Wenn $P > \acute{p}$, ist der Nutzen des Patienten unter der Anstrengungsabhängigen Entlohnung höher. Wenn $P < \acute{p}$, dann ist sein Nutzen unter der

Ergebnisabhängigen Entlohnungsform höher. Wenn nun die Veränderungen von $\hat{p}$ abhängig von α und β analysiert werden, erhält man:

$$(19a) \hat{p} = p_o^* (s_p (s_m/s_p)^\beta)^{1/\alpha\beta}$$

$$(19b) \partial \hat{p} / \partial \alpha = \hat{p} / (1-\alpha-\beta) (\ln p_o^* + [(1-\beta)/\alpha - (1-\alpha-\beta) \ln (s_p (s_m/s_p)^\beta)] / \alpha^2 \beta)$$

$$(19c) \partial \hat{p} / \partial \beta = \hat{p} / (1-\alpha-\beta) (\ln m_o^* - [(1-\alpha-\beta) \ln s_p] / \beta^2 \alpha + 1)$$

$$(19d) \partial^2 \hat{p} / \partial \alpha \partial \beta = \hat{p} / (1-\alpha-\beta)^2 \{ [\ln p_o^* - ((1-\alpha-\beta) \ln (s_p (s_m/s_p)^\beta) / \alpha^2 \beta) + 1] \\ [\ln m_o^* - ((1-\alpha-\beta) \ln s_p) / \beta^2 \alpha + 1] + (1-\beta)/\alpha (\ln m_o^* + 1) + \ln p_o^* + 1 \}$$

$s_p (s_m/s_p)^\beta$ und s_p sind immer kleiner 1. Falls p_o^* und m_o^* beide größer 1 sind, dann sind alle Ableitungen positiv. Damit sie größer 1 sind muss der Arztbesuch die Kosten für die Anstrengung wert sein.

Wenn α und β groß sind, dann ist es sehr wahrscheinlich, dass die Ergebnisorientierte Entlohnung besser ist als die Anstrengungsabhängige Entlohnung. Die Ergebnisorientierte Lösung ist ideal für Krankheiten, bei denen beide Anstrengungen, die des Arztes und die des Patienten, nötig für die Genesung und somit produktiv sind. Wenn nur die Anstrengung des Arztes produktiv ist und die Compliance des Patienten nicht, dann wird ein Anstrengungsorientierte Bezahlung auf Seiten des Arztes ideal sein.

Die Wahl für die beste Entlohnungsform hängt also von der Krankheitsform ab und ob die Krankheit einer Compliance des Patienten bedarf.

Ergebnisabhängige Bezahlung wird dann sinnvoll sein, wenn α und β beide hoch sind. Anstrengungsabhängige Entlohnung wird dann bevorzugt, wenn entweder α oder β sehr hoch ist.

6.3.6 Vergleich der Modelle

Diese Modelle untersuchen auf verschiedenen Weisen moral hazard im Arzt-Patienten Verhältnis, entweder auf der Arzt-, der Patientenseite oder auf beiden

Seiten gleichzeitig. Der Faktor der Ethik auf der Arztseite spielt vor allem bei Schmidt und bei Ma&McGuire eine entscheidende Rolle im Zusammenspiel der beiden Akteure. Ma&McGuire waren es auch, die eine eventuelle Absprache von Arzt und Patient eingearbeitet haben, die der Dritten Partei, der Versicherung, Nachteile bringen kann. Paula Gomez hat zwar vor allem die Arztseite untersucht, den ethischen Faktor allerdings außen vor gelassen, dafür allerdings die Tatsache verarbeitet, dass einige Ärzte neben ihrer Tätigkeit im öffentlichen Bereich auch eine Privatpraxis betreiben und herausgefunden, dass Ärzte mit Privatpraxis nicht immer die passende Behandlung für die Krankheit anbieten, sondern die, die ihr eigenes Prestige erhöht. Dieses egoistische Handeln hat nicht nur negative Auswirkungen auf den Patienten, sondern auch auf die Kosten, die die Versicherung beziehungsweise die Gesellschaft und der Staat tragen müssen. Sie hat außerdem auch ein Szenario dargestellt, in dem auch eine zweite Behandlungsrunde in Betracht gezogen wird.

Kenneth&Zivin haben nicht nur eine Entlohnungsart für den Arzt kalkuliert, sondern herausgefunden, dass je nach Krankheitsart jeweils entweder eine Entlohnung auf Basis des Anstrengungslevels oder des Ergebnisses bevorzugt wird beziehungsweise empfehlenswert ist.

6.4 Strategisches Zusammenspiel von ärztlicher Leistung und Compliance

Bis jetzt wurde nur der Unterschied in den Reaktionsfunktionen bei symmetrischer beziehungsweise asymmetrischer Information untersucht. Es wurde nun also aufgezeigt, dass unter bestimmten Voraussetzungen sowohl die medizinische Leistung des Arztes, als auch das Verhalten des Patienten bei vollständiger Information über dem Niveau im Fall der asymmetrischen Information liegt. Wie oben bereits erwähnt, bestehen Interdependenzen zwischen den beiden Produktionsfaktoren im Genesungsprozess. Was noch offen ist, ist die Frage, in welchem Verhältnis die Faktoren Einfluss auf das Gesamtergebnis nehmen, das heißt, die Auswirkungen des Verlaufs der

Reaktionsfunktionen auf das Niveau der medizinischen Leistung und der Compliance. Es wird hierbei unterschieden zwischen strategischer Unabhängigkeit, strategischen Komplementen und strategischen Substituten und untersucht, in wie weit die Verhältnisse der Faktoren zueinander das Endergebnis, die Genesung, beeinflussen.

Die Werte von ärztlicher Leistung und Compliance werden durch die Reaktionsfunktionen bestimmt. Hier wird nun angenommen, dass für die Reaktionsfunktionen im untersuchten Fall von Informationsasymmetrie ein eindeutiges und stabiles Nash-Gleichgewicht vorliegt.

Im folgenden Teil der Arbeit sollen diese Zusammenspiele bei Substitutions- und Komplementärsbeziehung näher betrachtet werden und gezeigt werden, inwieweit sich die Verhältnisse verändern, sobald eine Selbstbeteiligung eingeführt wird.

6.4.1 Strategische Substitute versus strategische Komplementäre

Nach Cooper und Ross (1985) wird auch hier angenommen, dass ein Akteur auch dann eine positive Aktion ($a, m > 0$) wählt, wenn der andere Akteur sich dazu entschließt, nichts zu unternehmen, da ansonsten unter Umständen kein Gleichgewicht mit positiven Aktionen existiert.

$$a(m=0) > 0 \text{ und } m(a=0) > 0$$

Das Ziel ist nun, den Verlauf der Reaktionsfunktionen zu ermitteln. Da hier double moral hazard existiert, werden dazu sowohl beim Arzt als auch auf der Seite des Patienten die Bedingungen erster Ordnung entsprechend des Impliziten Funktionentheorems total differenziert und nach da/dm aufgelöst. Der Satz des Impliziten Funktionentheorems besagt, dass in diesem Fall durch $F(a, m) = 0$ implizit eine Funktion $a(m)$ definiert wird, das heißt, es gilt $F(a(m), m) = 0$.

Nach der Differenzierung erhält man für den Patienten:

$$(13) da/dm = [\beta p_a(U_1 - U_0) - p_{am}(U_1 - U_0)] / p_{aa}(U_1 - U_0) - D''$$

Der erste Term des Zählers ist nur dann nicht null, wenn ein positiver Selbstbeteiligungssatz vorliegt. Der zweite Term des Zählers ist die strategische Komponente der Gleichung. Der Effekt der strategischen Komponente hängt vom Vorzeichen der Kreuzableitung der Wahrscheinlichkeit für einen hohen Gesundheitszustand (p_{am}) ab. Erhöht der Arzt seine medizinische Leistung, so zeigt die Kreuzableitung, wie sich die Grenzproduktivität der Compliance ändert.

Der Nenner ist die hinreichende Bedingung für ein Nutzenmaximum des Patienten und daher negativ.

Die Steigung der Reaktionsfunktion des Arztes ist:

$$(14) dm/da = - \varepsilon p_{am}(G_1 - G_0) / \varepsilon p_{mm}(G_1 - G_0) - C''$$

Auch im Fall des Arztes ist der Nenner als hinreichende Bedingung für ein Nutzenmaximum des Arztes negativ. In diesem Fall hängt die Steigung nur von einer Existenz von strategischen Komplementen oder Substituten ab. Sobald der Arzt den Gesundheitszustand des Patienten vernachlässigt, ist die Steigung null ($\varepsilon = 0$).

6.4.2 Interdependenzen ohne Einführung einer Selbstbeteiligung

Da nun der Effekt einer Selbstbeteiligung ignoriert wird, hängen die Steigungen der Reaktionsfunktionen davon ab, ob sich die Faktoren medizinische Leistung und Compliance wie strategische Komplemente oder Substitute verhalten.

Die Steigung der Reaktionsfunktion des Patienten sieht dann also folgend aus:

$$(13a) da/dm = - p_{am}(U_1 - U_0) / p_{aa}(U_1 - U_0) - D''$$

während die Steigung der Funktion auf Seiten des Arztes unverändert bleibt, da aufgrund der Risikoneutralität des Arztes die Kostenerstattung pro Einheit der

medizinischen Leistung (δ) keine Auswirkungen auf die Steigung seiner Reaktionsfunktion hat und somit ein Wegfall einer Selbstbeteiligung keinen Einfluss auf (14) hat.

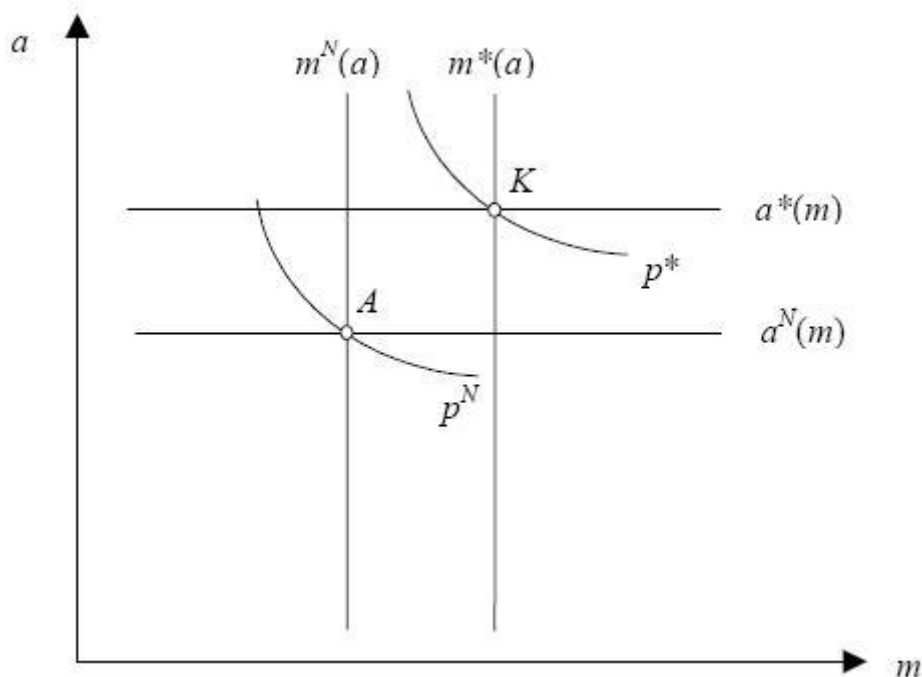
$$(14a) \quad dm/da = - \varepsilon_{p_{am}}(G_1 - G_0) / \varepsilon_{p_{mm}}(G_1 - G_0) - C''$$

Wie schon im Ausgangsfall ist der Nenner beider Gleichungen negativ.

Strategische Unabhängigkeit

Die Grenzproduktivität der beiden Faktoren ist jeweils vom anderen Produktionsfaktor (medizinische Leistung, Compliance) unabhängig ($p_{am} = 0$). Die beiden Reaktionsfunktionen sind somit unelastisch gegenüber den Aktionen der anderen Partei.

Abbildung 4: Unabhängigkeit von ärztlicher Leistung und Compliance



Quelle: Schneider, 1999

Die Abszisse zeigt die Menge der medizinischen Leistung (m), während die Ordinate die Compliance (a) abbildet. Die Reaktionsfunktionen (a^* , m^*) stellen

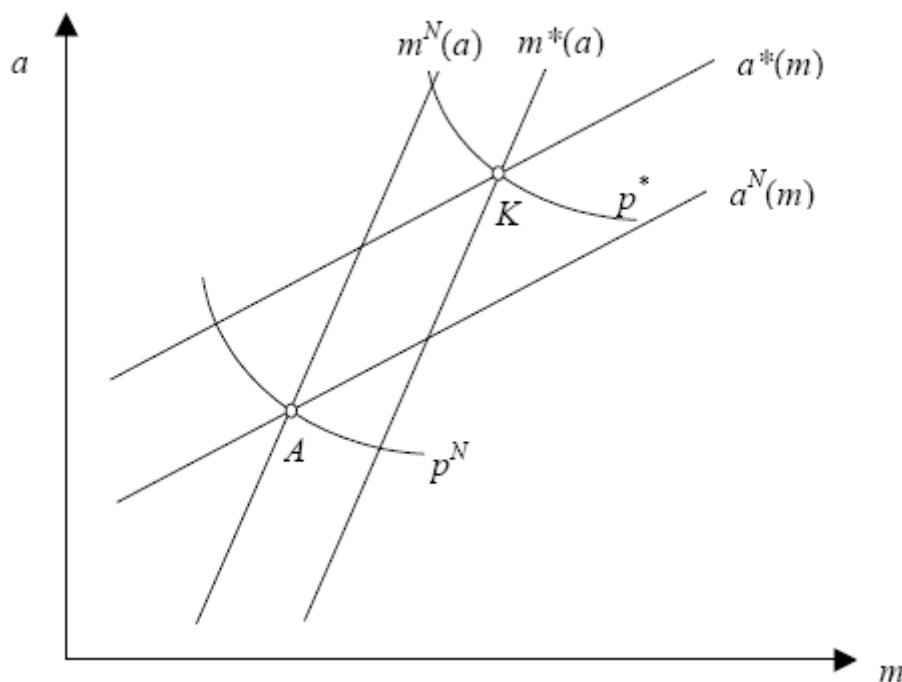
die first-best Lösung dar. Die Lösung bei Nicht-Kooperation wird durch a^N und m^N gekennzeichnet. Der Schnittpunkt der Reaktionsfunktionen im first-best Szenario wird durch K dargestellt. Vergleicht man nun die Punkte K für den first-best case und A für den Fall der asymmetrischen Information, so sieht man, dass sowohl die medizinische Leistung als auch die Compliance bei K auf einem höheren Niveau liegen. Die Iso-Wahrscheinlichkeitskurven sind in der Abbildung als p^* für den Fall der first-best und als p^N für den Fall der nicht-kooperativen Lösung zu sehen. Wie schon zuvor erwähnt, haben die Kurven durch totales Differenzieren von $p(a,m)$ eine negative Steigung. Lanoie (1991) fand heraus, dass je weiter die Kurve vom Ursprung entfernt ist, desto höher ist das Niveau mindestens eines Inputs und desto höher ist auch die Wahrscheinlichkeit, gesund zu werden.

Strategische Komplemente

In diesem Fall erwirkt die Erhöhung eines Faktors (a oder m) des einen Akteurs, eine Steigerung der Grenzproduktivität des anderen Vertragspartners ($p_{am} > 0$). Die Gleichungen (13a) und (14a) gestalten sich derart, dass der Zähler jeweils negativ ist, was zur Folge hat, dass sich bei einer Zunahme der medizinischen Leistung des Arztes, auch die Compliance auf Seiten des Patienten erhöht und vice versa.

Es gilt also $da/dm > 0$ beziehungsweise $dm/da > 0$.

Abbildung 5: Ärztliche Leistung und Compliance als strategische Komplemente



Quelle: Schneider, 1999

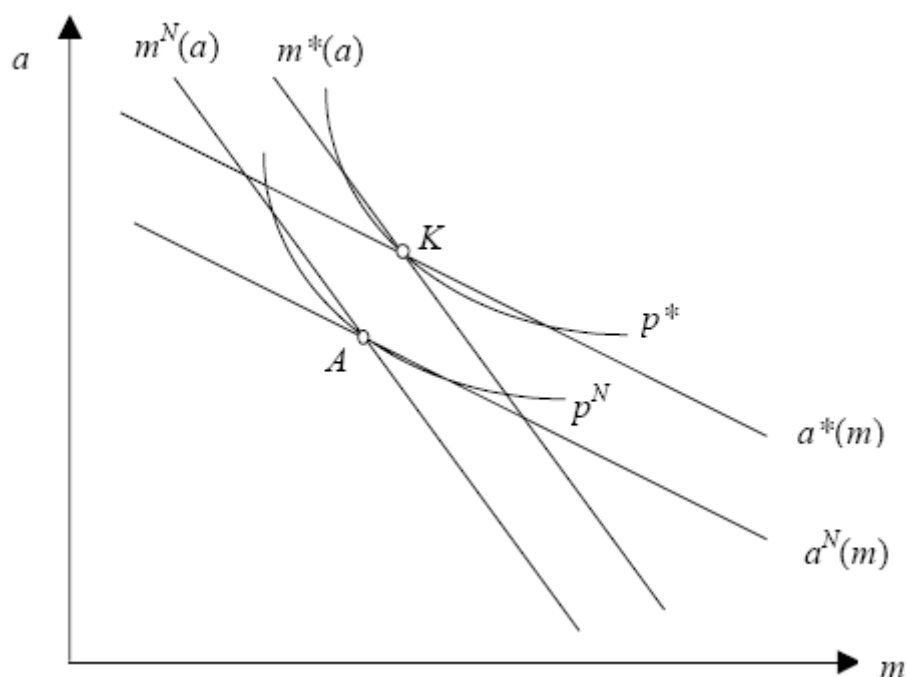
Es wird bei dieser Grafik zur Vereinfachung angenommen, dass die Reaktionsfunktionen die Form von Geraden annehmen, die bei symmetrischer und asymmetrischer Lösung parallel zueinander verlaufen. Der Verlauf hängt aber nach wie vor von der Nutzenfunktion und der Dichte ab.

Durch den Schnittpunkt der Geraden a^* und m^* erhält man die first-best Lösung im Punkt K, der verglichen mit der Situation bei asymmetrischer Information, und somit dem Schnittpunkt A, zeigt, dass durch das Informationsungleichgewicht bei der nicht-kooperativen Lösung der Arzt eine geringere medizinische Leistung anbietet und die Compliance des Patienten sinkt. Wenn die Akteure nicht wissen, wie sich ihre Handlungen auf die Grenzproduktivität des anderen auswirken, wählen sie, verglichen zur Situation bei vollkommener Information, ein zu niedriges Niveau.

Strategische Substitute

Eine Erhöhung der medizinischen Leistung erwirkt eine Senkung der Grenzproduktivität der Anstrengung des Patienten und umgekehrt. Ein Beispiel für strategische Substitute ist, wenn der Patient durch sportliche Betätigung den Heilungsprozess verlangsamt. In diesem Fall ist des strategischen Zusammenspiels ist der Zähler der Gleichungen (13a) und (14a) positiv, es entsteht also als Gesamteffekt $da/dm < 0$ und $dm/da < 0$. Die Reaktionsfunktionen sind in der Aktion des Mitspielers fallend. Der Unterschied zu den vorherigen zwei Fällen der Interdependenzen ist, dass bei strategischen Substituten die resultierenden Niveaus der medizinischen Leistung beziehungsweise der Compliance nicht mehr eindeutig sind. Es ergeben sich gleichgerichtete Effekte, sowie gegenläufige Effekte.

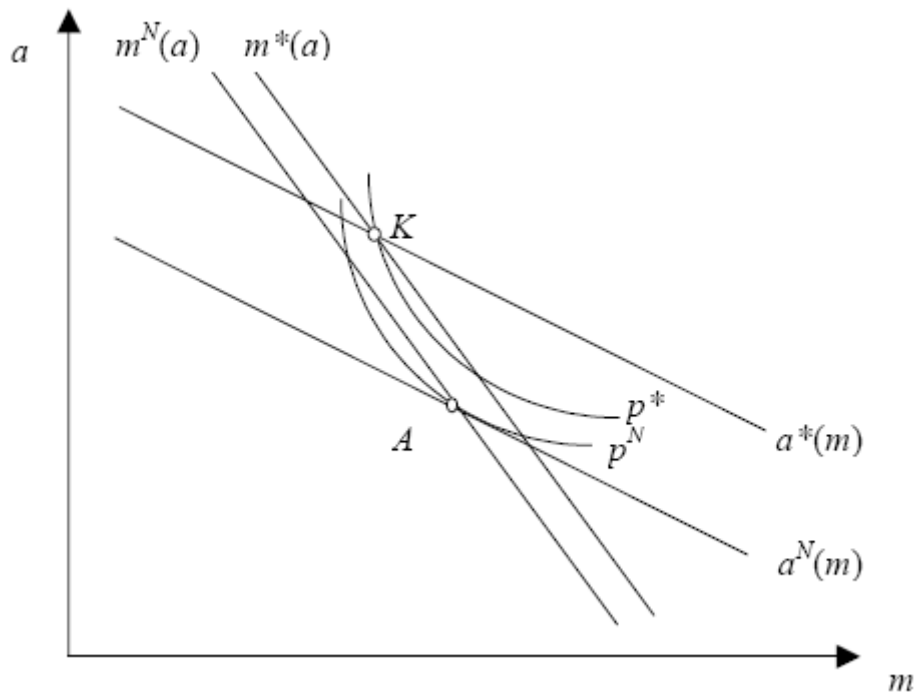
Abbildung 6: Ärztliche Leistung und Compliance als strategische Substitute mit gleichgerichteten Effekten



Quelle: Schneider, 1999

Wie in der Abbildung 6 zu sehen ist, senken sich, wie schon im Fall der strategischen Komplemente, bei Nicht-Kooperation beide Niveaus. Der Punkt A liegt sowohl auf der Abszisse, als auch auf der Ordinate auf einem niedrigeren Niveau als der Schnittpunkt K.

Abbildungung 7: Ärztliche Leistung und Compliance als strategische Substitute mit gegenläufigen Effekten



Quelle: Schneider, 1999

In Abbildung 7 ist zu sehen, dass zwar die Reaktionskurven bei Nicht-Kooperation unter denen bei Kooperation liegen, dennoch wird beim Vergleich der Punkte K und A ersichtlich, dass während es zu einer Absenkung der Compliance (a) kommt, das Niveau der medizinischen Leistung (m) über das first-best Niveau ansteigt. Erklären lässt sich diese Beobachtung dadurch, dass sich der Vorgang in zwei Effekte aufteilen lässt. Ausgegangen wird von der kooperativen Lösung, durch die Informationsasymmetrie entsteht ein Rückgang beider Inputs, gegeben die Anstrengungen der anderen Partei und damit verschieben sich die Reaktionsfunktionen. Der zweite Effekt ist, dass eine Senkung eines Produktionsfaktors (a oder m) dazu führt, dass sich der andere Faktor produktiver auf die Wahrscheinlichkeit eines höheren

Gesundheitszustandes auswirkt. Merkt zum Beispiel der Arzt, dass die Wahrscheinlichkeit für den Genesungszustand steigt, sobald er das Niveau seiner medizinischen Leistung erhört, so wird er die Leistung weiter erhöhen. Wenn nun der zweite Effekt der höheren Grenzproduktivität den ersten Effekt durch die Informationsasymmetrie für einen der Akteure dominiert und somit entstehen gegenläufige Effekte bei strategischen Substituten. Das first-best Niveau kann allerdings immer nur für eine Partei (Arzt oder Patient) überschritten werden.

6.4.3 Interdependenzen mit Einführung einer Selbstbeteiligung

Nach der Einführung einer Selbstbeteiligung kann der Patient die Leistungen des Arztes nicht mehr kostenfrei in Anspruch nehmen. Dadurch werden sein Nutzen und auch sein Grenznutzen beeinflusst. Es entstehen dem Patienten neben der Versicherungsprämie und neben den Kosten für die Compliance zusätzliche Kosten. In der Realität wird bei einer Versicherung mit Selbstbeteiligung die Prämienzahlung an das Versicherungsunternehmen geringer ausfallen, als bei einer Versicherung ohne Selbstbehalt. Man könnte also nicht einfach bei den Berechnungen den Selbstbehalt gänzlich vom Nettoeinkommen abziehen. Ohne Einkommenswirkung würde die Selbstbeteiligung allerdings keine Anreize auf die Nachfrage nach medizinischer Leistung haben, weswegen in dieser Arbeit zur Vereinfachung angenommen wird, dass das Nettoeinkommen des Patienten nach Einführung der Selbstbeteiligung unverändert bleibt. Dies erlaubt einen Vergleich mit den vorhergegangenen Ergebnissen.

Die Ableitung des Erwartungsnutzens des Patienten nach der medizinischen Leistung ergibt im Fall mit Selbstbeteiligung:

$$(15) \quad p_m(U_1 - U_0) = \beta[pU'_1 + (1 - p)U'_0]$$

Bei keiner Selbstbeteiligung ($\beta = 0$) ist der rechte Term gleich null und der Patient könnte nach Belieben die medizinischen Leistungen ausnutzen. Mit

Selbstbeteiligung wird dem ein Riegel vorgeschoben, durch die entstehenden Kosten, begrenzt sich die optimale Menge für den Patienten.

Nach der Einführung von β kann die Steigung der Reaktionsfunktion des Patienten nicht mehr in allen Fällen eindeutig bestimmt werden, da das resultierende Nash-Gleichgewicht beeinflusst wird. Wie schon oben festgestellt hängt die Steigung und somit das Vorzeichen des Zählers der Gleichung (13) von der Ausprägung der Selbstbeteiligungskomponente und der strategischen Komponente ab. Es stellt sich nun allerdings die Frage, in wie weit ein höherer Selbstbeteiligungssatz die Anstrengungen des Patienten beeinflusst.

Die Bedingung 1. Ordnung des Patienten zeigt dass der Patient seine Anstrengung derart wählt, dass der erwartete Grenznutzen der Compliance gleich dem Grenzaufwand ist.

$$(8) p_a(a, m) (U_1 - U_0) = D'$$

Mittels Impliziten Funktionstheorems kann nun anhand (8) gezeigt werden, wie sich eine Erhöhung des Selbstbeteiligungssatzes auf die Anstrengung des Patienten auswirkt:

$$(8a) da/d\beta = p_a(U_1 - U_0)(m + \partial\sigma/\partial\beta) / p_{aa}(U_1 - U_0) - D'' > 0$$

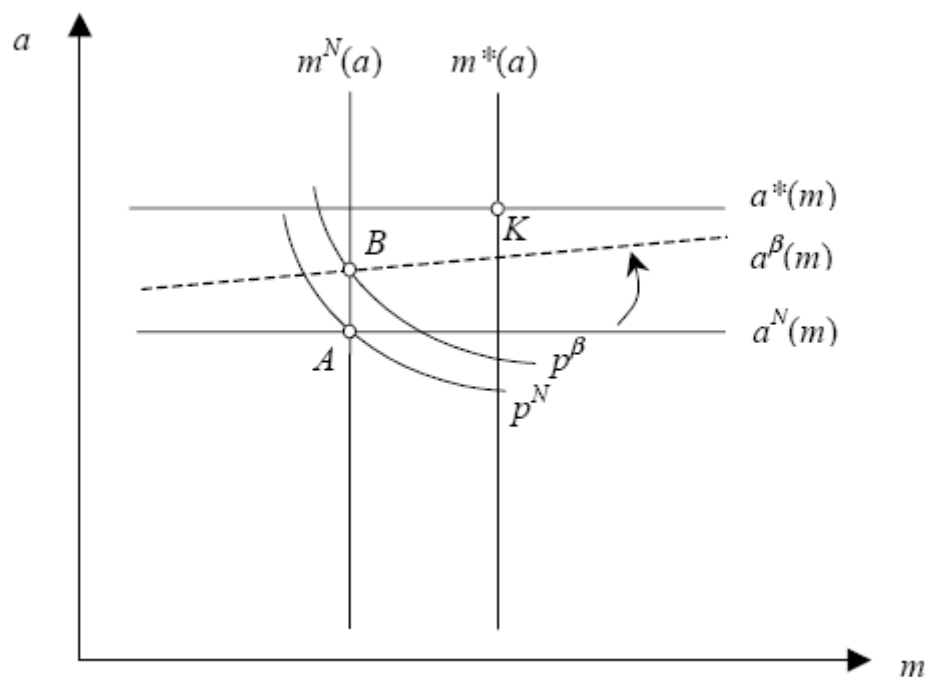
Hierbei ist sowohl der Nenner als auch der Zähler negativ. Der Gesamteffekt wäre demnach, dass eine Erhöhung des Selbstbeteiligungssatzes die Compliance des Patienten für eine gegebene Anstrengung des Arztes erhöht. Cutler und Zeckhauser (1999) haben außerdem mit empirischen Untersuchungen festgestellt, dass die Nachfrage nach medizinischer Leistung mit steigendem Preis sinkt. Durch dieses Zusammenspiel der Faktoren können medizinische Leistungen durch den gesunkenen Relativpreis für die eigene Anstrengung durch eine höhere Compliance des Patienten substituiert werden.

Für einen höheren Selbstbeteiligungssatz steigt also die Compliance des Patienten. Dadurch verläuft die neue Reaktionsfunktion oberhalb der Reaktionsfunktion, die sich vor der Einführung der Selbstbeteiligung ergeben hat.

Selbstbeteiligung bei strategischer Unabhängigkeit

Besteht vor der Einführung der Selbstbeteiligung strategische Unabhängigkeit zwischen der Compliance und der medizinischen Leistung ($p_{am}=0$), so bewirkt die Selbstbeteiligung eine positive Steigung der Reaktionsfunktion des Patienten. Aus der Sicht des Patienten nimmt die Interdependenz zwischen ärztlicher und seiner eigenen Anstrengung die Form einer komplementären Beziehung an.

Abbildung 8: Selbstbeteiligung bei strategischer Unabhängigkeit



Quelle: Schneider, 1999

Die positive Steigung der Reaktionsfunktion des Patienten lässt sich in Abbildung 8 an der gestrichelten Linie $a^\beta(m)$ erkennen. Durch die Wandlung von

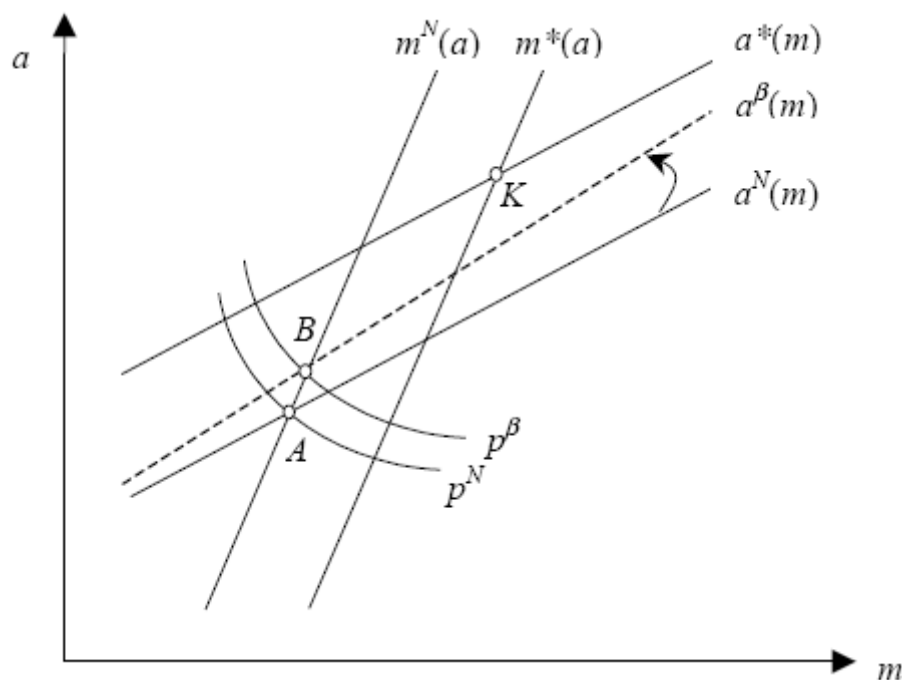
Unabhängigkeit zur komplementären Beziehung erbringt der Patient für jedes Niveau der ärztlichen Leistung höhere Anstrengungen. In weiterer Folge steigt damit auch die Wahrscheinlichkeit für einen besseren Genesungs- und Gesundheitszustand. Mit Hilfe der Iso-Wahrscheinlichkeitskurven, lässt sich veranschaulichen, dass, beim Vergleich von p^B und p^N , die Wahrscheinlichkeit der Genesung durch die höhere Compliance steigt. Die Einführung von einer Selbstbeteiligung kann sogar positive Folgen für den Erwartungsnutzen des Patienten haben. Für den Fall, dass der Erwartungsnutzen aus dem Einkommen nach der Einführung von Selbstbeteiligung stärker zunimmt als der Aufwand durch die höhere Anstrengung, steigt auch der gesamte Erwartungsnutzen an. Selbst wenn also die medizinische Leistung auf demselben Niveau bleibt, kann es sein, dass durch die Selbstbeteiligung die Anstrengung des Patienten zunimmt.

Selbstbeteiligung bei strategischen Komplementen

Soeben wurde gezeigt, dass bei vormals unabhängigem Zusammenspiel der Faktoren, durch die Einführung einer Selbstbeteiligung, eine komplementäre Beziehung entsteht.

Wenn die Faktoren ursprünglich schon ein komplementäres Verhältnis gebildet hatten ($p_{am} > 0$), so wird dieses nach der Einführung einer Selbstbeteiligung noch verstärkt. Die Reaktionsfunktion des Patienten liegt, wie schon im Fall der Unabhängigkeit, über derjenigen der nicht-kooperativen Lösung.

Abbildung 9: Selbstbeteiligung bei strategischen Komplementen



Quelle: Schneider, 1999

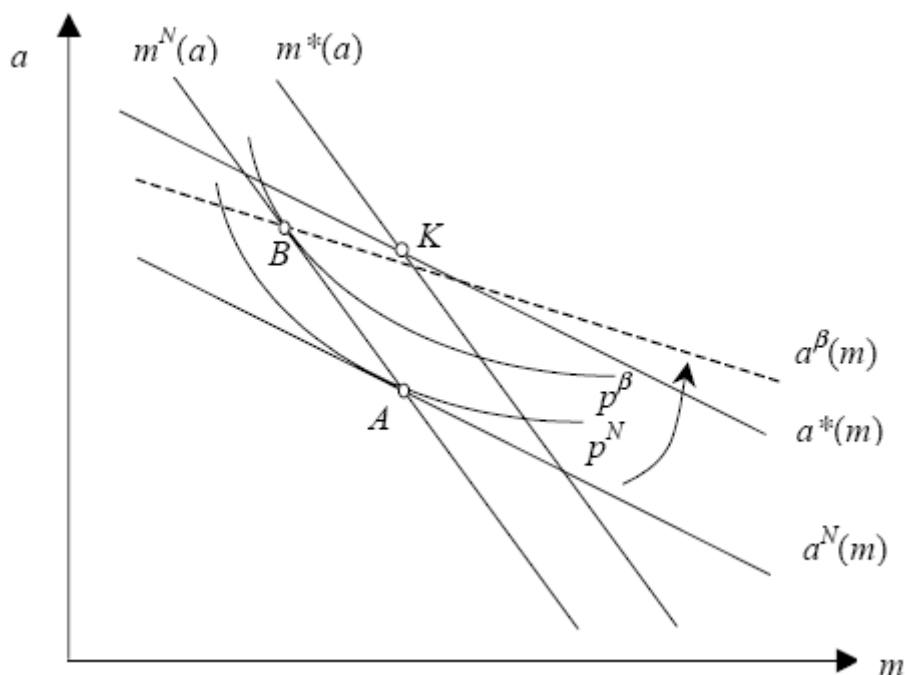
Am neuen Verlauf der Reaktionsfunktion des Patienten – gestrichelte Gerade $a^\beta(m)$ in Abbildung 9 – lässt sich erkennen, dass bei gegebenem Niveau der medizinischen Leistung die „neue“ Compliance über der Compliance ohne Selbstbeteiligung liegt. Am Schnittpunkt B kann man sehen, dass die Lösung mit Selbstbeteiligung und asymmetrischer Information zwar immer noch weit unter dem Niveau der symmetrischer Information liegt, aber doch über dem Niveau der Situation ohne Selbstbeteiligung bei Nicht-Kooperation. Somit ist ersichtlich, dass bei strategischen Komplementen sowohl die medizinische Leistung m als auch die Compliance a nach Einführung der Selbstbeteiligung steigen und somit eine Genesung wahrscheinlicher wird. Bewiesen wird dies auch anhand der Iso-Wahrscheinlichkeitskurven, da p^β über dem Nash-Gleichgewicht ohne Selbstbeteiligung p^N liegt. Wie schon im vorigen Fall steigt der gesamte Erwartungsnutzen des Patienten, wenn der Erwartungsnutzen des Nettoeinkommens hinreichend groß ist, um den Aufwand für die Compliance zu

übersteigen. Welcher der beiden Inputs der Genesungsproduktion stärker zunimmt, hängt dabei von den Steigungen der Reaktionsfunktionen ab.

Selbstbeteiligung bei strategischen Substituten

Bei strategischer Unabhängigkeit und strategischen Komplementen war das Verhältnis der Inputs zueinander nach der Einführung der Selbstbeteiligung wie jenes von strategischen Komplementen. Bei strategischen Substituten ist die Lage, wie schon im Fall ohne Selbstbeteiligung, nicht ganz so eindeutig ermittelbar. Die Einführung der Selbstbeteiligung kann für den Patienten zu einem flacheren Verlauf der Reaktionsfunktion $a^\beta(m)$ führen.

Abbildung 10: Selbstbeteiligung bei strategischen Substituten und erhöhter Genesungswahrscheinlichkeit



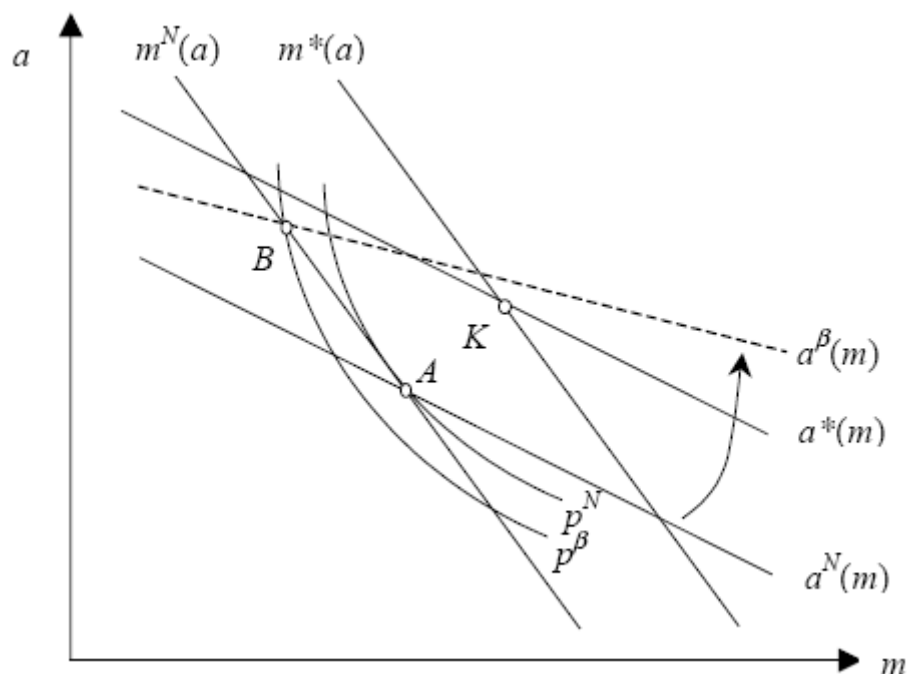
Quelle: Schneider, 1999

In diesem Fall wäre es sogar möglich, dass die Selbstbeteiligungskomponente die strategische Komponente überlagert und die Reaktionsfunktion einen steigenden Verlauf annimmt. Bei gegebener medizinischer Leistung und sonst

gleich bleibenden Voraussetzungen, ist die Compliance des Patienten höher. An Punkt B in der Abbildung 10 ist ersichtlich, dass durch die neue Lage der Reaktionsfunktion $a^\beta(m)$, die medizinische Leistung m zurückgeht, während sich die Anstrengung des Patienten a erhöht. Dabei kann es zu der Situation kommen, dass die Compliance des Patienten das first-best Niveau übersteigt, der Schnittpunkt B also auf der Ordinate über dem Punkt K liegt. Wird im Fall der strategischen Substitute eine Selbstbeteiligung eingeführt, so substituiert die Compliance die medizinische Leistung des Arztes. Der Wohlfahrtseffekt für den Patient ist durch die höhere Genesungswahrscheinlichkeit trotzdem gegeben. Wie schon in den vorigen Darstellungen ist der Erwartungsnutzen des Patienten höher, wenn der Erwartungsnutzen des Nettoeinkommens stärker steigt als die Kosten für den Aufwand des Patienten.

Die Einführung einer Selbstbeteiligung kann allerdings bei strategischen Substituten auch zu einer Senkung der Genesungswahrscheinlichkeit führen.

Abbildung 11: Selbstbeteiligung bei strategischen Substituten und verminderter Genesungswahrscheinlichkeit



Quelle: Schneider, 1999

Das Ergebnis des Erwartungsnutzens hängt im großen Maße von der Gesundheitsproduktionsfunktion ab. Die Grenzrate der Substitution zwischen ärztlicher Leistung und Anstrengung des Patienten hängt von den Grenzproduktivitäten der beiden Genesungsproduktionsfaktoren m und a ab. Wenn die Iso-Wahrscheinlichkeitskurve ausreichend steil verläuft, kann es dazu kommen, dass die Genesungswahrscheinlichkeit sinkt. In Abbildung 10 steigt die Compliance ausreichend an, um die Genesungswahrscheinlichkeit, trotz substituierter medizinischer Leistung, zu erhöhen.

In Abbildung 11 ist nun der Fall aufgezeigt, bei dem zwar die Compliance derart hoch ist, dass die Genesungswahrscheinlichkeit steigt, allerdings im Gesamteffekt nicht ausreichend hoch, um den Rückgang der Genesungswahrscheinlichkeit, verursacht durch die Substitution der medizinischen Leistung, wettzumachen. In diesem Fall verringert sich der Erwartungsnutzen des Patienten sogar.

Für alle strategischen Komponenten lässt sich sagen, dass nach Einführung der Selbstbeteiligung die Compliance angestiegen ist. In einer Situation ohne Selbstbeteiligung kann die medizinische Leistung zu Grenzkosten von null nachgefragt werden, durch die Selbstbeteiligung entstehen beim Patienten Kosten für die Konsumation von ärztlicher Behandlung. Die Compliance wird dadurch im Vergleich zur medizinischen Leistung relativ billiger.

Neben der Compliance und der medizinischen Leistung hängt der Erwartungsnutzen des Patienten auch von der Genesungswahrscheinlichkeit ab, die allerdings vorweg nicht genau bestimmt werden kann. Es wird allerdings gefolgert, dass eine steigende Genesungswahrscheinlichkeit auch zu einem höheren Erwartungsnutzen führt.

In welchem Fall welche strategische Komponente auftritt, hängt teilweise mit der Art und dem Ausmaß der Krankheit ab.

7 Empirische Studien

Die Theorie besagt also, dass durch die Einführung einer Versicherung die Compliance sinkt, weil die Kosten für medizinische Behandlung und dadurch auch die finanziellen Konsequenzen einer Krankheit sinken. Allerdings scheinen bisherige Studien in der Praxis keinen Beweis für diese Theorie gefunden zu haben.

Jedoch haben empirische Untersuchungen bereits bewiesen, dass die Einführung beziehungsweise der Erwerb von Autoversicherungen (Chiappori 2000) oder Versicherungen für Unfälle am Arbeitsplatz (Fortin und Lanoie 2000) das Verhalten des Agenten geändert haben. Der Agent verhält sich, sobald er versichert ist, weniger vorsichtig und durch moral hazard entstehen letztendlich mehr Unfälle als in Fällen ohne Versicherung. Wieso soll es bei moral hazard im Fall der Krankenversicherungen anders sein? Warum gibt es hier keinen empirischen Beweis für den theoretischen Ansatz? Diese Problematik haben Dhaval Dave und Robert Kaestner in ihrem Paper „Health Insurance and ex ante moral hazard: Evidence from Medicare“ (2006) untersucht. Auf diese Studie stützt sich der folgende Teil der Arbeit. Dave und Kaestner versuchen den Effekt der Einführung von Medicare, einer staatlichen Krankenversicherung für Menschen ab dem Alter von 65 Jahre, herauszuarbeiten, indem sie das Verhalten von älteren Männern vor und nach der Einführung der gratis Versicherung überprüfen. Einige ältere Studien konnten keine eindeutige Veränderung im Verhalten beobachten. Dave und Kaestner gehen davon aus, dass zwei Effekte sich gegenseitig aufheben und deswegen die Untersuchungen kein richtiges Ergebnis liefern. Sie spalten den Effekt von Medicare in einen direkten (ex ante moral hazard) und einen indirekten Effekt auf. Der direkte Effekt zeigt sich in einem veränderten Verhalten, also weniger Sport, mehr rauchen, höherer Alkoholkonsum. Der indirekte Effekt, der durch die Einführung von Medicare verursacht wird, ist jedoch der Anstieg der Arztbesuche und Kontakt mit Gesundheitsexperten, wodurch ein größeres Gesundheitsbewusstsein geschaffen wird. Während sich also der direkte Effekt

negativ auf die Compliance auswirkt, hat der indirekte Effekt einen positiven Einfluss auf die Compliance.

7.1 Das Model

Dave und Kaestner wählen für ihre Untersuchungen den Ansatz von Ehrlich und Becker aus dem Jahr 1972.

$$(1) EU = [1 - \pi(r)]U(I_1 - r) + \pi(r)U(I_0 - r)$$

Wobei (I) für Einkommen steht, 1 für gesund und 0 für krank und r sind die Kosten für Compliance.

[FOC(r)]

$$(2) FOC(r) = -\partial\pi / \partial r \cdot U_1 - [1 - \pi(r)]U'_1 + \partial\pi / \partial r U_0 - \pi(r)U'_0 = 0$$

$$- \partial\pi / \partial r (U_1 - U_0) = U'_1 + \pi(U'_0 - U'_1)$$

Das Einkommen ist also höher in einem gesunden Zustand als in einem kranken, was bedeutet, dass der Nutzen steigt, wenn die Wahrscheinlichkeit für Krankheit sinkt.

Nun wird eine gratis Krankenversicherung (q) in Form von Medicare eingeführt unter der Annahme, dass die Personen vor ihrem 65. Geburtstag nicht versichert waren und wir bekommen eine neue Nutzenfunktion.

$$(3) EU = [1 - \pi(r)]U(I_1 - r) + \pi(r)U(I_0 - r + q)$$

Durch die Einführung von Medicare wird der Einkommenseffekt zwischen „krank“ und „gesund“ verringert, was zu ex ante moral hazard führt. Ein Agent, der voll versichert ist ($q = I_1 - I_0$) hat keinen Vorteil, wenn er in Compliance investiert.

Aus (2) wissen wir, dass die Compliance Bemühungen des Patienten von der Produktivität der Compliance ($\partial\pi / \partial r$) abhängt, sowie von der Wahrscheinlichkeit (π) krank zu werden. Wie schon oben erwähnt ist es belegt, dass vermehrte Arztbesuche die Compliance Bemühungen erhöhen. Deswegen werden nun die Kontakte zu Ärzten (m) als Funktion von Versicherung (q) eingeführt: $m = m(q)$

$$(4) \pi = \pi[r, m(q)]$$

$$(\partial\pi / \partial r) = (\partial\pi / \partial r) [m(q)]$$

In (2) eingesetzt ergibt das:

$$(5) FOC(r) = - \partial\pi / \partial r [m(q)(U_1 - U_0) - \{1 - \pi(r, m(q))\} U'_1 + \partial\pi[r, m(q)] U'_0] = 0$$

Der Einfluss der Versicherung (q) auf die Vorsorge (r) ist gegeben durch:

$$(6) dr / dq = - \partial FOC(r) / \partial q / SOC(r)$$

und ist laut Annahme negativ.

$$(7) \partial FOC(r) / \partial q = - (\partial\pi)^2 / \partial r \partial m \cdot \partial m / \partial q (U_1 - U_0) + \partial\pi / \partial r [U'_0] - \partial\pi / \partial m \cdot \partial m / \partial q (U'_0 - U'_1) - \pi[r, m(q)] U''_0$$

Es wird angenommen, dass Ärzte das Grenzprodukt von Compliance bzw. Vorsorge steigern $[(\partial\pi)^2 / \partial r \partial m < 0]$ und dass der Einfluss des Arztes auf die Krankheitswahrscheinlichkeit nicht ganz eindeutig ist $\partial\pi / \partial m \leq / > 0$.

In Gleichung (7) ist der erste Teil also positiv, woraus man sehen kann, dass auf Grund der höheren Arztbesuche und der damit einhergehenden besseren Aufklärung die Compliance steigt, die Einführung einer Versicherung führt also zu einer besseren Vorsorge. Ausgeglichen wird dieser Effekt allerdings von dem zweiten Term der Gleichung, der negativ ist und somit den ex ante moral hazard Effekt beschreibt, womit die Einführung einer gratis Versicherung die Vorsorgebemühungen reduziert. Der dritte Gleichungsterm ist entweder positiv oder negativ, abhängig davon, ob ein Arzt feststellt, dass ein Patient gesünder

oder kranker ist als erwartet. Der letzte Teil ist positiv und zeigt, dass Medicare die Grenzkosten der Compliance senkt und somit die Vorsorge erhöht.

Für die Studie wurden die Daten der *Health and Retirement Study (HRS)* herangezogen. Untersucht wurde das Verhalten von Männern im Alter von 60 bis 69 Jahren und wie es sich mit der Medicare ab dem Alter von 65 verändert hat. Es wurde außerdem verglichen, wie sich das Verhalten von Männern unterscheidet, die vor dem Alter von 65 bereits versichert waren, zu denen, die vor Medicare keine Krankenversicherung hatten. Die Compliance wurde in drei Größen aufgeteilt: sportliche Betätigung, rauchen und Alkoholkonsum.

7.2 Ergebnisse

Frage 1: Wie wirkt sich die Einführung von Medicare für Männer ab dem Alter von 65 Jahren, die vorher nicht versichert waren, auf die Anzahl der Arztbesuche aus?

Die Wahrscheinlichkeit für einen Arztbesuch stieg um 12,4%, die tatsächlichen Arztbesuche stiegen um 30,9%, was sich dadurch erklären lässt, dass nicht versicherte Männer wegen der Kosten seltener zum Arzt gehen, als versicherte Personen.

Frage 2: Wie wirkt sich die Einführung von Medicare für Männer ab dem Alter von 65 Jahren, die vorher nicht versichert waren, auf die Compliance aus?

Aus den vorhergegangenen Annahmen wissen wir, dass Veränderungen im Verhalten auf ex ante moral hazard und auch auf die bessere Aufklärung durch einen höhere Wahrscheinlichkeit für Arztbesuche zurückzuführen sind. Demnach müssen beide Aspekte berücksichtigt werden.

Sportliche Betätigung:

Nach der Einführung von Medicare sank die Wahrscheinlichkeit für sportliche Betätigung um 24%. Dieser Wert erhöht sich allerdings, wenn man den

positiven Effekt von Besuchen beim Arzt berücksichtigt. Der reine ex ante moral hazard Effekt bewirkt eine Senkung um sogar 39,7%.

Rauchen:

Ohne die Berücksichtigung des positiven Einflusses durch Kontakte zu einem Arzt bewirkt die Einführung von Medicare nur eine geringe Senkung der Anzahl an Zigaretten pro Tag mit nur 1,3%. Dieser Wert allerdings um einiges signifikanter, wenn man sich nur auf den ex ante moral hazard Effekt konzentriert, dieser bewirkt eine Steigerung der Anzahl an pro Tag gerauchten Zigaretten um 22,7% nach der Einführung einer gratis Versicherung. Kontakt zu medizinischen Fachkräften hat also einen bedeutenden Einfluss auf das Rauchverhalten. Genauer betrachtet wurde im Zuge der Studie herausgefunden, dass jemand, der vor dem Alter von 65 bereits versichert war, nach einem Arztbesuch seine Zigarettenkonsum um 10% reduziert hat, während Männer, die vor Medicare keine Versicherung hatten, nach einem Arztbesuch ihren Tabakkonsum um sogar 42,1% verringert haben. Ähnliche Ergebnisse erhält man, wenn man sich ansieht, wie viele Männer nach Arztbesuchen ganz mit dem Rauchen aufgehört haben. Die Wahrscheinlichkeit mit dem Rauchen aufzuhören sinkt um 18% bei den Männern, die weder vor noch nach der Einführung von Medicare einen Arzt besucht haben.

Alkoholkonsum:

Wenn man nur den reinen ex ante moral hazard Effekt betrachtet und den Effekt durch Arztbesuche herausfiltert, so steigt die Wahrscheinlichkeit für Alkoholkonsum um 31,8% und die Wahrscheinlichkeit für täglichen Alkoholkonsum um 14,8% nach dem Erhalt von Medicare.

Aus diesen Untersuchungen geht also hervor, dass die Einführung einer gratis Versicherung, in diesem Fall Medicare, bei Männern zwischen 60 und 69 zu

einer Steigerung des gesundheitsschädlichen Verhaltens führt, ein Effekt, der nur durch den von der Versicherung bewirkte Anstieg der Arztbesuche minimiert wird.

Eine Erklärung, warum bei Krankenversicherungen der moral hazard Effekt nicht oder nur gering festgestellt werden konnte, könnte sich also dadurch erklären, dass durch die häufigeren Arztbesuche die Compliance tatsächlich steigt und teilweise so signifikant, dass der moral hazard Effekt im Endergebnis sogar aufgehoben wird. Kalkuliert man den Effekt eines Arztbesuches allerdings weg, so erhält man den tatsächlichen moral hazard Effekt, der durchaus beachtlich sein kann. Die Untersuchungen von Dave und Kaestner impliziert demnach, dass die Einführung von Versicherungen durchaus einen guten Einfluss auf die Patienten und deren Verhalten hat, wenn durch die Versicherung der Kontakt zu medizinischen Fachkräften ausreichend gefördert wird.

7.3 Weitere empirische Erkenntnisse

Ähnliche Erkenntnisse haben auch Jonathan Klick und Thomas Stratmann in ihrer Studie „Diabetes Treatment and Moral Hazard“ aus dem Jahr 2005 gewonnen. Sie haben untersucht, wie sich das Verhalten von Diabetes-Patienten und gesunden Menschen ändert, wenn eine staatliche Versicherung zur Verfügung gestellt wird. In Amerika wurden in den Jahren 1996 bis 2000 in mehreren Staaten Krankenversicherungen gesetzlich dazu verpflichtet, medizinische Leistungen, Medikamente und Behandlungen für Diabetiker zur Verfügung zu stellen. Auslöser für diese Neuerung war, dass Diabetes sich immer mehr zu einer Volkskrankheit entwickelt hat. Laut Schätzungen aus dem Jahr 2003 des Centers for Disease Control (CDC) haben mehr als 17 Millionen Menschen in den USA Diabetes. Die Krankheit kann unter anderem zu Blindheit, Nierenproblemen und Amputationen führen, die dem Staat sehr teuer zu stehen kommen können. Das CDC stuft Diabetes als fünfthöchste Todesursache in den USA ein. Die American Diabetes Association hat die

Kosten, die dem Staat durch Diabetes entstanden sind, auf 132 Milliarden Dollar geschätzt. Dies beinhaltet Kosten für medizinische Versorgung und auch durch Produktivitätsverluste. (ADA 2003)

Um diese Kosten zu reduzieren, haben einige Staaten die Versicherungen gesetzlich zu der Deckung von medizinischer Betreuung von Diabetes-Patienten vorgeschrieben. Eine bessere Vorsorge und Betreuung sollte helfen, die Krankheit besser zu behandeln und die Compliance der Patienten, beziehungsweise auch der noch gesunden Bevölkerung, zu erhöhen. Ob sich das Verhalten der Menschen tatsächlich nach der Einführung der erweiterten Versicherungsleistungen zum Positiven geändert hat, haben Klick und Stratmann in ihrem Paper untersucht, indem sie das Verhalten von Patienten in Staaten mit gesetzlicher Diabetes-Betreuung und ohne verglichen hat. Gemessen wurde die Verhaltensveränderung an dem Body-Mass-Index (BMI), der sich wie folgt berechnet:

$$\text{BMI} = \text{Körpergewicht in kg} / (\text{Körpergröße in m})^2$$

Der BMI ist deshalb eine sehr geeignete Größe, da Übergewicht sehr oft Ursache für Diabetes ist und weil sich an der Veränderung des BMI auch die Veränderung des Gesundheitsverhaltens des Patienten erkennen lässt. Vor allem Diabetes vom Typ 2 kann durch die richtige Diät und gesundheitsförderndes Verhalten verhindert oder zumindest die Beschwerden reduziert werden, wodurch sich auch die Kosten für den Staat in weiterer Folge senken.

Durch die Einführung einer erweiterten Versicherungsleistung für die gesamte Bevölkerung eines Staates kostet eine medizinische Behandlung für den einzelnen allerdings weniger, die Opportunitätskosten für Compliance, in diesem Fall Diät und sportliche Betätigung, steigen allerdings, wodurch moral hazard entsteht und der Versicherte sich gegen einen gesunden Lebensstil entscheiden könnte. Er kann ab nun wählen, ob er sich für medizinische Leistung oder für Diät und Sport, vor und während der Krankheit, entscheidet, um seinen Gesundheitszustand zu verbessern. Sie kamen zu dem Ergebnis,

dass sowohl bei Diabetes-Patienten als auch bei gesunde Menschen in den Staaten mit Gesetzen zur Deckung von Diabetes-Behandlungskosten nach der Einführung dieses Gesetzes der BMI gestiegen ist. Bei Diabetikern war gab es sogar einen Anstieg von bis zu 7%. Wie Klick und Stratmann feststellten wurde eine Einführung eines derartigen Gesetzes moral hazard veranlassen, da die Menschen statt gesunder Diät und sportlicher Betätigung zur Verbesserung ihres Gesundheitszustandes beziehungsweise Vorsorge eher zu den medizinischen Mitteln greifen würden, deren Kosten durch die neue, großzügigere Versicherung gedeckt wurden. (Klick und Stratmann 2005, 23). Auch in dieser Arbeit war die Erkenntnis, dass Kostenübernahme durch Versicherung oder Staat sich positiv auswirkt, wenn dadurch der Kontakt zu medizinischem Personal, Arzt, erhöht wird. Durch bessere Aufklärung und Erziehung steigt auch die Compliance.

8 Conclusio

Im Verhältnis zwischen Arzt und Patient kommt es zu einem double moral hazard, da beide hidden information haben, die die jeweils andere Partei nicht kennt.

Es wurde nachgewiesen, dass bei unter bestimmten Voraussetzungen sowohl die medizinische Leistung des Arztes als auch das Verhalten des Patienten bei vollständiger Information über dem Niveau im Fall der asymmetrischen Information liegt.

Der Arzt ist in seiner Funktion nicht nur Anbieter der medizinischen Leistung, sondern auch durch seine Rolle als Berater des Patienten, Nachfrager der eigenen Leistung. Merkt der Arzt, dass die Wahrscheinlichkeit für den Genesungsprozess mit der Erhöhung seiner medizinischen Leistung steigt, so wird er diese weiter erhöhen.

Der Patient hingegen kann den Genesungsprozess mit seiner Compliance beeinflussen. Er selbst bestimmt den Level seines gesundheitsfördernden oder

gesundheitsschädigenden Verhaltens. Unter Compliance fallen zum Beispiel sportliche Betätigung, Zigaretten- und Alkoholkonsum oder auch das Befolgen der ärztlichen Vorschriften. Die Compliance des Patienten ist natürlich mit Kosten verbunden, monetäre oder auch nicht-monetäre.

Wenn eine Selbstbeteiligung eingeführt wird, kann der Patient die ärztlichen Leistungen nicht mehr gratis in Anspruch nehmen, wodurch ihm neben den Kosten für die Compliance auch Kosten für medizinische Leistungen entstehen. Gleichzeitig sinkt die Nachfrage nach medizinischen Leistungen, sobald diese teurer werden. Dies verringert wiederum den Relativpreis für Compliance. Weswegen der Patient die medizinische Leistung teilweise oder gänzlich durch eine höhere Compliance substituiert.

In der Theorie kann dieser Effekt ausreichend nachgewiesen werden. Dennoch konnten viele empirische Studien keine signifikante Veränderung im Verhalten von Patienten feststellen, wenn eine gratis Versicherungsleistung eingeführt wurde. In Bereichen wie Produktgarantien war der empirische Nachweis von moral hazard eindeutig. Man könnte nun argumentieren, dass den Menschen ihre eigenen Gesundheit wichtiger ist, als die Lebensdauer eines Produkts. Leider ist der Mensch nicht gar so moralisch, wenn es um den eigenen Körper geht. Der moral hazard Effekt bei Krankenversicherungen war oft nicht eindeutig, weil sich Verhaltensveränderungen nach Einführung einer Versicherung aus zwei Effekten zusammensetzen. Durch gesunkene Kosten für medizinische Leistungen durch Versicherungen wird der Kontakt zu Ärzten erhöht, da ein Arztbesuch nun um einiges günstiger oder sogar gratis ist. Kontakt zu medizinischem Personal hat einen positiven Effekt auf das Verhalten der Patienten. So kam es mitunter vor, dass Patienten nach einem Arztbesuch ihren Tabakkonsum um rund 42% verringert haben. Betrachtet man die Verhaltensveränderung ohne den positiven Effekt eines Arztbesuches, kann man durchaus einen signifikanten moral hazard Effekt erkennen.

Patienten verhalten sich nach Einführung einer Versicherung schlecht, obwohl sie dadurch nicht nur der Versicherung und dem Staat schaden, sondern vor

allen sich selbst und der eigenen Gesundheit. Trotzdem entscheiden sie sich, nach Abschluss einer Versicherung, mit der sie medizinische Leistungen gratis oder günstiger erhalten, sich gesundheitsschädigend zu verhalten.

Eine Lösung für dieses Dilemma wäre eine Förderung des Kontakts zu medizinischem Personal. Da Arztbesuche einen derart positiven Effekt auf die Compliance haben, dass sie damit sogar den moral hazard Effekt aufwiegen, sollte mit einer Einführung einer Versicherung bessere Aufklärung und besserer Kontakt zu Gesundheitsexperten gefördert werden.

9 Literaturverzeichnis

Akerlof, G.A. (1970) – “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism”, The Quarterly Journal of Economics, Vol. 84, 448-500

Arrow, K. (1963) – “Uncertainty and the Welfare Economics of Medical Care”, The American Economic Review, Vol. 53/5, 941-973

Bardey, D. und Lesur, R. (2006), „Optimal Regulation of Health System with Induced Demand and Ex Post Moral Hazard“, Annals of Economic and Statistics, 83-84, 279-293

Bhattacharyya, S. und Lafontaine, F. (1995) – “Double-Sided Moral Hazard and the Nature of Share Contracts”, Rand Journal of Economics, 24, 761-781

Cooper, R. und Ross, T.W. (1985) – “Product Warranties and Double Moral Hazard” – Rand Journal of Economics, 16, 103-113

Cutler, D. M. und Zeckhauser, R.J. – “The Anatomy of Health Insurance”, NBER Working Paper No. 7176, Cambridge, Massachusetts

Decker, M. (2000) – „Die Behandlung der Information in der ökonomischen Theorie: Adverse Selektion und Moral Hazard“, GRIN Verlag, München

Dhaval, D. und Kaestner, R. (2006), „Health Insurance and Ex Ante Moral Hazard: Evidence from Medicare, National Bureau of Economic Research Working Paper 12764, 1-50

González, P. (2004), “Should Physicians’ Dual Practice Be Limited? An Incentive Approach”, Health Economics, 13/6, 505-524

Hayek, F.A. (1945) – „The Use of Knowledge in Society“, American Economic Review No.4, 519-530

Holmström, B. (1997) – “Moral hazard and observability”, The Bell Journal of Economics, Vol. 10, 74-91

Klick, J. und Stratmann, T. (2005), "Diabetes Treatments and Moral Hazard", Journal of Law and Economics, Vol. 50, 519-538

Lanoie, P. (1991) – "Occupational Safety and Health: a Problem of Double or Single Moral Hazard", Journal of Risk and Insurance, 58, 80-100

Leonard, K.L. und Zivin, J. (2000), „How To Compensate Physicians When Both Patient and Physician Effort are Unobservable“, Columbia University Discussion Paper Series No. 9900-1, 1-33

Ma, C-T.A. und McGuire, T.G. (1997), "Optimal Health Insurance and Provider Payment", The American Economic Review, 87/4, 685-702

Macho-Stadler, I, und Pérez-Castrillo, D. (1997) – „An Introduction to the Economics of Information – Incentives and Contracts“, Oxford University Press, USA

Müller, H.M. (1997) – „The theory of moral hazard“, Dissertation Universität St. Gallen

Pauly, M.V. (Jun., 1968) – "The Economics of Moral Hazard: Comment" – The American Economic Review, Vol. 58, No. 3, Part 1., 531-537

Pita Barros, P.L., Machado, M., Sanz de Galdeano, A. (2008) – "Moral Hazard and the demand for health services: A matching estimator approach", Journal of Health Economics, Vol. 27/4, 1006-1025

Prescott, E.S. (1999) – "A Primer on Moral-Hazard Models", Federal Reserve Bank of Richmond Economic Quarterly, Vol. 85/1, 47-77

Savage, E. und Wright, D.J. (2003) – "Moral Hazard and Adverse Selection in Australian Private Hospitals" – Journal of Health Economics, 22, 331-359.

Schneider, U. (1998) – „Der Arzt als Agent des Patienten – Zur Übertragbarkeit der Principal-Agent-Theorie auf die Arzt-Patient-Beziehung.“,

Wirtschaftswissenschaftliche Diskussionspapiere Nr. 2/98, Rechts- und Staatswissenschaftliche Fakultät, Ernst-Moritz-Arndt-Universität Greifswald

Schneider, U. (1999) – „Ärztliche Leistung und Compliance des Patienten – der Fall des Double Moral Hazard“, Wirtschaftswissenschaftliche Diskussionspapiere 7/99. Rechts- und Staatswissenschaftliche Fakultät, Ernst-Moritz-Arndt-Universität Greifswald

Schneider, U. (2002) – “Beidseitige Informationsasymmetrien in der Arzt-Patient-Beziehung: Implikationen für die GKV”, Vierteljahrshefte zur Wirtschaftsforschung 71/4, 447-458

Spremann, K. (1990) – „Asymmetrische Information“, Zeitschrift für Betriebswirtschaft, 60, 561-586

Theilen, B. (1996), “Optimal Contract and Organizational Design under Moral Hazard and Adverse Selection”, Verlag Josef Eul, Bergisch Gladbach – Köln

Varian, H.R. (1995) – “Grundzüge der Mikroökonomik“, Oldenbourg-Verlag, München

Wille, E. und Ulrich, V. (1991) - "Bestimmungsfaktoren der ausgabenentwicklung in der gesetzlichen Krankenversicherung (GKV)", Duncker & Humblot Verlag, Berlin

10 Anhänge

10.1 Abstract

Die Beziehung zwischen Arzt und Patient kann als Kernbeziehung des Gesundheitswesens bezeichnet werden. Ein zentraler Gegenstand zahlreicher ökonomischer Untersuchungen ist dabei der Wissensvorsprung des Arztes gegenüber dem Patienten in der Ausübung seiner Tätigkeit. Auch der Patient besitzt Informationen, die dem Arzt nicht ohne weiters zugänglich sind. Dies bezieht sich auf sein behandlungsbegleitendes und gesundheitsrelevantes Verhalten (Compliance), das heißt auf seine gesundheitsfördernden Anstrengungen, die der Arzt kaum oder nur ungenau beobachten kann, zum Beispiel Freizeitaktivitäten, individueller Lebensstil, Trink- und Rauchverhalten, Ernährungsweise. Wegen dieser beidseitigen Informationsasymmetrien entsteht im Arzt-Patienten Verhältnis double moral hazard.

The relationship between a doctor and his patient is a crucial relationship in healthcare. The focus of many economical studies has been the advantage in knowledge the physicians has when treating a patient. But the patient also has information, which is not accessible to the doctor. The patient's information refers to the patient's health relevant behavior during the treatment (compliance), i.e. his health enhancing effort, which cannot or hardly be observed by the doctor, like leisure activities, individual lifestyle, alcohol and cigarette consumption, nutrition. Because of these mutual information asymmetries there is double moral hazard in the doctor-patient relationship.

10.2 Lebenslauf

Persönliche Daten

Name: Sandra Dobrozemsky

Geburtsdatum, -ort: 11. November 1977, Wien

Nationalität: Österreich

Ausbildung

Oktober 2006 Letzte Diplomprüfung des Studiums der Internationalen Betriebswirtschaft

Februar 2004 Besuch der Sprachschule Eurocentres in Lausanne (Französisch)

Oktober 2003 Inskription am Institut für Romanistik der Universität Wien (Französisch)

Juli und August 2002 London School of Economics (Sommeruniversität; „Internationale Volkswirtschaft“)

Juli 1998 Besuch der Sprachschule Málaga Plus (Spanien)

Oktober 1996 Beginn des Studiums der Internationalen Betriebswirtschaft mit Schwerpunkt Industrielles Management und Marketing

1988-1996 Gymnasium
BG, BRG u. WK BRG Wien XV
Diefenbachgasse
(Matura, Juni 1996)