



Wolls Lehr- und Handbücher der Wirtschafts- und Sozialwissenschaften

Herausgegeben von

Universitätsprofessor Professor h. c. Dr. Dr. h. c. Artur Woll

Bisher erschienene Werke:

- Aberle, Transportwirtschaft, 2. A.
- Assenmacher, Konjunkturtheorie, 8. A.
- Barro, Makroökonomie, 3. A.
- Barro · Grilli, Makroökonomie - Europäische Perspektive
- Barro · Sala-i-Martin, Wirtschaftswachstum
- Blum, Volkswirtschaftslehre, 2. A.
- Branson, Makroökonomie, 4. A.
- Bretschger, Wachstumstheorie, 2. A.
- Brösse, Industriepolitik, 2. A.
- Büschges · Abraham · Funk, Grundzüge der Soziologie, 3. A.
- Cezanne, Allgemeine Volkswirtschaftslehre, 3. A.
- Fischer · Wiswede, Grundlagen der Sozialpsychologie
- Glastetter, Außenwirtschaftspolitik, 3. A.
- Leydold, Mathematik für Ökonomen
- Müller, Angewandte Makroökonomik
- Rosen · Windisch, Finanzwissenschaft I
- Rush, Übungsbuch zu Barro, Makroökonomie, 3. A.
- Sachs Larrain, Makroökonomik - in globaler Sicht
- Schneider, Grundlagen der Volkswirtschaftslehre, 3. A.
- Tirole, Industrieökonomik, 2. A.
- Varian, Mikroökonomie, 3. A.
- Wachtel, Makroökonomik
- Wacker · Blank, Ressourcenökonomik I
- Wohltmann, Grundzüge der makroökonomischen Theorie, 2. A.

Industrieökonomik

Von

Jean Tirole

Professor am MIT
Cambridge, Massachusetts, U.S.A.

Aus dem
Amerikanischen
von

Dr. Roland Ladwig
Professor Dr. Bruno Schönfelder
und
Professor Dr. Peter Seidelmann

2., deutschsprachige Auflage

R. Oldenbourg Verlag München Wien

Titel der amerikanischen Originalausgabe: „The Theory of Industrial Organization“.
© 1988, 1992 (Fifth Printing) by The MIT Press, Cambridge, Massachusetts, U.S.A.

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Tirole, Jean:

Industrieökonomik / von Jean Tirole. Aus dem Amerikan. von
Roland Ladwig ... – 2., deutschsprachige Aufl. – München ; Wien :
Oldenbourg, 1999

(Wolls Lehr- und Handbücher der Wirtschafts- und
Sozialwissenschaften)

Einheitssacht.: The theory of industrial organization <dt.>

ISBN 3-486-25082-5

© 1999 R. Oldenbourg Verlag
Rosenheimer Straße 145, D-81671 München
Telefon: (089) 45051-0, Internet: <http://www.oldenbourg.de>

Das Werk einschließlich aller Abbildungen ist urheberrechtlich geschützt. Jede Verwertung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Bearbeitung in elektronischen Systemen.

Gedruckt auf säure- und chlorfreiem Papier
Gesamtherstellung: R. Oldenbourg Graphische Betriebe GmbH, München

ISBN 3-486-25082-5

Inhaltsverzeichnis

Vorwort der Übersetzer	XXIII
Vorwort	XXV
Einleitung	1
Warum Industrieökonomik	1
Theoretische und empirische Analyse	5
Zum Inhalt des Buches	7
Hinweise zur Benutzung des Buches - Gliederung und Aufbau	8
Zusammenhänge zwischen den Kapiteln	9
Untergliederung der Kapitel	9
Übungsaufgaben	10
Voraussetzungen	10
Definition eines Marktes, partielles Gleichgewicht und Wohlfahrtskriterien	12
Das Paradigma des vollkommenen Wettbewerbs	12
Partielles Gleichgewicht, fallende Nachfragekurve und Konsumentenrente	14
Was ist ein Markt?	25
Literaturhinweise	28
Theorie der Unternehmung	31
1 Was ist eine Unternehmung?	35
1.1 Die Unternehmung als Instrument, um den Wettbewerb zu beschränken	35
Preisdifferenzierung	36
Preisüberwachung bei Zwischenprodukten	36

1.2	Synergieeffekte als raison d'être des Unternehmens	38
	Kann die technologieorientierte Sicht des Unternehmens als eine "Theorie der Unternehmung" gelten?	44
1.3	Die Unternehmung als langfristige Beziehung	46
	Idiosynkratische Investitionen und "Asset Specificity"	46
	Preisbildung beim bilateralen Monopol und das Ex-Post- Volumen des Leistungsaustausches zwischen einem Monopolisten und einem Monopsonisten	48
	Bargaining	49
	Verträge	52
	Spezifische Investitionen und das Hold up-Problem	55
	Bargaining	55
	Verträge	58
	Die Grenzen langfristiger Beziehungen	62
	Eine Anwendung: Periodische Versteigerung einer Konzession	63
	Empirische Untersuchungen	65
1.4	Das Unternehmen als ein unvollständiger Vertrag	67
	Schiedsgerichte	68
	Entscheidungsbefugnis	70
	Der Erstreckungsbereich der Entscheidungsbefugnis	77
	Empirische Untersuchungen	78
	Reputation als Ersatz für Verträge oder Integration	79
	Dual Sourcing als Ersatz für Verträge	80
2	Die Gewinnmaximierungshypothese	81
2.1	Das grundlegende Anreizproblem	83
	Beobachtbarkeit, Verifizierbarkeit und Entscheidungsbefugnis	89
	Obergrenzen für Bestrafungen und Renten für den Manager	91

2.2	Die Entscheidungsspielräume des Managements und ihre Grenzen	96
	Leistungsvergleiche	96
	Unerwünschte Übernahmeangebote	100
	Leistungsanreize für das Management in dynamischer Betrachtung	103
	Aufsicht	105
	Wettbewerb auf dem Absatzmarkt des Unternehmens	109
	Eine Anwendung: Divisionalisierung	112
2.3	Zweifel am neoklassischen Forschungsansatz	114
	Optimierungsverhalten	114
	Kommunikation und Wissen	115
	Die Dynamik von Organisationen	117
	Gruppenverhalten	117
2.4	Die Gewinnmaximierungshypothese und die Industrieökonomik	118
	Anhang: Die Principal-Agent-Beziehung	120
	Hinweise zur Lösung der Übungsaufgaben	129
	Literaturhinweise	131

Teil I

	Die Ausübung von Monopolmacht	141
1	Das Monopol	143
1.1	Preispolitik	144
1.1.1	Ein Eingut-Monopolist	144
1.1.1.1	Die Regel von der inversen Elastizität	144
1.1.1.2	Der Wohlfahrtsverlust (Dead-Weight Loss)	145
1.1.1.3	Die Wirkung einer speziellen Verbrauchsteuer	149
1.1.1.4	Bedingungen zweiter Ordnung	152

1.1.2	Der Mehrgüter-Monopolist	152
1.1.2.1	Nachfrageverflechtung, zerlegbare Kosten	154
	Eine Anwendung: Intertemporale Preisbildung und Goodwill	155
1.1.2.2	Unabhängige Nachfragen, unabhängige Kosten	156
	Eine Anwendung: Lernkurveneffekte	157
1.1.3	Ein Monopol auf ein langlebiges Gut	159
1.1.4	Die "Ermittlung" der Nachfragekurve	162
1.1.5	Läger	163
1.2	Kostenauftrieb	165
1.3	Rent-Seeking-Verhalten	166
1.4	Einige abschließende Bemerkungen	170
1.5	Anhang: Langlebige Güter und Grenzen der Monopolmacht	172
1.5.1	Recycling	172
1.5.2	Langlebige Güter und intertemporale Preisdifferenzierung	176
1.5.2.1	Leasing oder Verkauf	176
1.5.2.2	Das Coase-Problem	179
1.5.2.3	Umgehung des Coase-Problems	182
1.5.2.4	Monopol und geplante Obsoleszenz	189
	Hinweise zur Lösung der Übungsaufgaben	192
	Nachtrag: Ein heuristischer Beweis der Coase-Hypothese	200
	Literaturhinweise	203
2	Produktgestaltung, Qualität und Werbung	209
2.1	Der Begriff des Produktraumes	210
2.1.1	Vertikale Differenzierung	210
2.1.2	Horizontale Differenzierung	213

2.1.3	Der Lancaster-Ansatz	217
2.1.4	Die traditionelle Konsumtheorie	217
2.2	Produktpolitik	219
2.2.1	Produktqualität	219
2.2.1.1	Anwendungen	223
	Swans (1970) Theorem von der optimalen Haltbarkeitsdauer	223
	Die Dorfman-Steiner-Bedingung (1954)	224
	Ein Beispiel für "Unterversorgung" mit Qualität	227
	Ein Beispiel für "Übersorgung" mit Qualität	228
2.2.2	Zu flaches oder zu tiefes Sortiment?	228
2.2.2.1	Nichtabschöpfbare Konsumentenrente - eingeschränkte Angebotsvielfalt	229
2.2.2.2	Mehrgütermonopol und überzogene Angebotsvielfalt	230
2.2.3	Produktgestaltung und Differenzierung	231
2.3	Qualität und Information	232
2.3.1	Laufkundschaft: Moralisches Risiko und "Zitronen"	234
2.3.1.1	Moralisches Risiko (Moral Hazard)	234
2.3.1.2	Das Zitronenproblem	238
2.3.2	Wiederholungskäufe	242
2.3.3	Qualität, Information und Verbraucherschutz	248
2.3.3.1	Ungültigkeit des Coaseschen Theorems und Produkthaftung	248
	Unvollkommene Information	249
	Transaktionskosten	249
2.3.3.2	Erzeugung von Information	250
2.3.3.3	Fehleinschätzungen seitens der Verbrauchern	251
2.4	Werbung	252
2.5	Abschließende Bemerkungen	253

2.6	Anhang: Wiederholungskäufe	255
2.6.1	Wiederholungskäufe ohne moralisches Risiko	255
2.6.1.1	Goodwill und Einführungsofferten	255
2.6.1.2	Signale für eine bereits vorhandene Qualität	260
2.6.2	Veränderliche Qualität und Reputation	267
2.6.2.1	Qualitätsprämien und Bootstrap-Gleichgewichte	268
2.6.2.2	Asymmetrische Information und Reputation	272
	Hinweise zur Lösung der Übungsaufgaben	278
	Literaturhinweise	286
3	Preisdifferenzierung	291
3.1	Perfekte Preisdifferenzierung	296
3.2	Preisdifferenzierung im Vielmärkte-Fall (Preisdifferenzierung dritten Grades)	298
3.2.1	Noch einmal die Regel von der inversen Elastizität	298
3.2.2	Wohlfahrtsaspekte	300
3.2.2.1	Eine Anwendung auf den Fall linearer Nachfragekurven	302
3.2.2.2	Ein Einwand	303
3.2.2.3	Resümee	304
3.2.3	Anwendungen	304
3.2.3.1	Anwendung 1: Räumliche Diskriminierung	305
3.2.3.2	Anwendung 2: Vertikale Bindungen als Instrument der Differenzierung	307
3.2.4	Preisdifferenzierung dritten Grades auf Märkten für Zwischenprodukte	309
3.3	Interpersonelle Arbitrage und Kunden-Screening (Preisdifferenzierung zweiten Grades)	311
3.3.1	Gespaltene Tarife	311
3.3.1.1	Perfekte Differenzierung	315

3.3.1.2	Der Monopolpreis	316
3.3.1.3	Gespaltener Tarif	316
3.3.1.4	Vergleich	317
3.3.1.5	Eine Variante: Koppelungsvereinbarungen als Instrument der Preisdifferenzierung	319
3.3.2	Nichtlineare Tarife und Mengendifferenzierung	322
3.3.2.1	Wohlfahrt	325
3.3.3	Qualitätsdifferenzierung	326
3.3.3.1	Anwendung 1: Differenzierung bei Versicherungspolicen	328
3.3.3.2	Anwendung 2: Differenzierung mittels Wartezeiten und Streuung der Preise	328
3.4	Abschließende Bemerkungen	331
3.5	Anhang: Nichtlineare Preise	333
3.5.1	Nichtlineare Tarife	333
3.5.1.1	Der Zwei-Typen-Fall	333
	Die Bedingungen individueller Rationalität	336
3.5.1.2	Der Fall einer kontinuierlichen Menge von Verbrauchertypen	337
	Allgemeinere Nachfragefunktionen	343
	Mengenrabatte und Mengenzuschläge	344
	Wohlfahrtsaspekte	345
3.5.1.3	Optimale nichtlineare Tarife und Ramsey-Preise	345
3.5.2	Mengenrabatte und Koppelung	346
3.5.2.1	Homogene Güter	346
	Heterogene Güter	349
3.5.3	Der Versicherungsmarkt	350
	Hinweise zur Lösung der Übungsaufgaben	355
	Literaturhinweise	363

4	Vertikale Bindung	369
4.1	Proportionale Preise oder vertikale Bindungen	371
4.1.1	Das Grundmodell	371
4.1.2	Wettbewerb zwischen den Einzelhändlern	374
4.1.3	Mehrere Inputs	375
4.1.4	Konkurrierende Produkte	376
4.1.5	Zur Rechtmäßigkeit vertikaler Bindungen	377
4.2	Externe Effekte und vertikale Bindung	377
4.2.1	Methodik	377
4.2.2	Der grundlegende vertikale externe Effekt	378
	Hinreichende vertikale Bindungen	384
4.3	Wettbewerb auf der nachgelagerten Stufe	396
4.3.1	Wettbewerb und Service	396
4.3.2	Der horizontale externe Effekt	398
4.3.3	Spezialisierung im Handel	401
4.3.4	Wettbewerb im (Einzel-)Handel als Anreizmechanismus	402
4.3.5	Händlerkartelle	403
4.4	Eine Mehrzahl von Produkten	404
4.4.1	Ausschließlichkeitsbindung und Effizienz	404
4.4.2	Vertikale Beschränkungen und strategisches Verhalten der Zulieferer	405
4.5	Schlußbemerkungen	406
4.6	Anhang: Bindungen mit wettbewerbsbeschränkender Wirkung	407
4.6.1	Die Bedeutung von Wettbewerb für die Erhöhung der Effizienz	407
4.6.1.1	Unsicherheit, Delegation von Entscheidungskompetenzen und Versicherung	408
	Risiko	408

Delegation von Entscheidungsbefugnissen	409
Versicherung	411
4.6.1.2 Ein Modell des Wettbewerbs im Handel	412
4.6.1.3 Wohlfahrtstheoretische Analyse	417
4.6.2 "Marktabstottung"	421
4.6.2.1 Allgemeines über Sperrn	421
Sperrn als effizientes Instrument zur Ausnutzung bereits bestehender Monopolmacht	422
Monopolisierung	424
4.6.2.2 Verträge als Marktzutrittsschranken	428
Hinweise zur Lösung der Übungsaufgaben	433
Literaturhinweise	441

Teil II

Oligopole und strategische Entscheidungssituationen	447
Einführung - Preis- und Nichtpreiswettbewerb	447
Nichtkooperative Spiele und strategisches Verhalten	448
Reaktionsfunktionen: Strategische Komplemente und Substitute	451
5 Kurzfristiger Preiswettbewerb	455
5.1 Das Bertrand-Paradox	455
5.2 Auflösungen des Bertrand-Paradoxes: Eine Einführung	459
5.2.1 Die Edgeworth-Lösung	459
5.2.2 Die zeitliche Dimension	460
5.2.3 Produktdifferenzierung	461
5.2.4 Was von der Bertrand-Analyse zu halten ist	461
5.3 Abnehmende Skalenerträge und Kapazitätsgrenzen	462

5.3.1	Rationierungsregeln	462
5.3.1.1	Die effiziente Rationierungsregel	463
5.3.1.2	Die Proportionale Rationierungsregel	464
5.3.2	Preiswettbewerb	466
5.3.2.1	Ein Beispiel für Kapazitätsgrenzen	467
5.3.3	Ex ante Investition und Ex post Preiswettbewerb	470
5.3.4	Diskussion	472
5.4	Traditionelle Cournot-Analyse	475
5.5	Konzentrations-Indizes und Rentabilitätsunterschiede zwischen Wirtschaftszweigen	482
5.6	Abschließende Bemerkungen	487
5.7	Anhang: Mengenwettbewerb	489
5.7.1	Traditionelle Cournot-Analyse: Existenz, Eindeutigkeit und Konvergenzeigenschaften	489
5.7.1.1	Existenz eines Gleichgewichtes in reinen Strategien	489
5.7.1.2	Eindeutigkeit	493
5.7.1.3	Konvergenz zum Gleichgewicht mit vollkommenem Wettbewerb	494
5.7.2	Preisspiele mit Kapazitätsgrenzen	498
5.7.2.1	Das Preisspiel	498
5.7.2.2	Die Entscheidung über die Kapazitäten	506
5.7.2.3	Diskussion der Rationierungsregel	507
5.7.2.4	Die Zeitstruktur des Spiels	509
5.7.2.5	Wettbewerb um Inputs	511
	Hinweise zur Lösung der Übungsaufgaben	513
	Literaturhinweise	520

6	Dynamischer Preiswettbewerb und spontanes Parallelverhalten	525
6.1	Überkommene Lehrmeinungen (Faktoren, die Kollusion fördern oder behindern)	527
6.1.1	Kollusion	527
6.1.2	Verspätete Entdeckung (detection lags)	528
6.1.3	Asymmetrien	530
6.1.4	Andere Einflußfaktoren	532
6.2	Statische Modelle eines dynamischen Preiswettbewerbs	533
6.2.1	Die geknickte Nachfragekurve	533
6.2.2	Konjekturale Variationen	536
6.2.3	Diskussion	537
6.3	Superspiele	537
6.3.1	Theoretische Grundlagen	537
6.3.2	Anwendungen	543
6.3.2.1	Anwendung 1: Konzentration	543
6.3.2.2	Anwendung 2: Lange Informationslags oder seltene Interaktionen	544
6.3.2.3	Anwendung 3: Schwankende Nachfrage	545
6.3.2.4	Anwendung 4: Unterschiedliche Kostenverläufe	549
6.3.2.5	Anwendung 5: Kontakt auf mehreren Märkten	550
6.3.3	Heimliche Preissenkungen	552
6.3.4	Zur Kritik des Modells	554
6.3.4.1	Synchrone Entscheidungen	555
6.3.4.2	Unendlicher Zeithorizont	555
6.3.4.3	Eine Vielzahl von Gleichgewichten	556
6.4	Preisstarreheiten	557
6.4.1	Das Beispiel einer geknickten Nachfragekurve	559

6.4.2	Diskussion	562
6.5	Eine Reputation der Freundlichkeit	564
6.5.1	Eine Metapher: Das wiederholte "Gefangenen-Dilemma"-Spiel	567
6.5.2	Diskussion	571
6.5.3	Ein evolutionstheoretischer Ansatz	574
6.6	Abschließende Bemerkungen	575
6.7	Anhang: Dynamische Spiele und spontanes Parallelverhalten	577
6.7.1	Heimliche Preissenkungen	577
6.7.1.1	Preiswettbewerb	577
6.7.1.2	Mengenwettbewerb	581
6.7.2	Preisstarrheiten und die geknickte Nachfragekurve	584
6.7.2.1	Das Ein-Perioden-Abweichungskriterium	584
6.7.2.2	Die Bellmann-Gleichungen	585
6.7.2.3	Die Gewinne sind stets größer als null	587
6.7.2.4	Ein gegen Nachverhandlungen immunes Gleichgewicht	590
6.7.3	Folk-Theoreme	590
6.7.3.1	Unendlich oft wiederholte Spiele mit vollständiger Information	591
6.7.3.2	Wiederholte Spiele mit vollständiger Information, deren Stufenspiele multiple Gleichgewichte aufweisen	594
6.7.3.3	Wiederholte Spiele mit unvollständiger Information	595
	Hinweise zur Lösung der Übungsaufgaben	597
	Literaturhinweise	606
7	Produktdifferenzierung: Preispolitik und die Gestaltung anderer Absatzeinflußgrößen	611
7.1	Räumlicher Wettbewerb	614
7.1.1	Die lineare Stadt	614

7.1.1.1	Preiswettbewerb	615
7.1.1.2	Produktpolitische Entscheidungen	619
7.1.2	Die kreisförmige Stadt	622
7.1.2.1	Das Modell	622
7.1.2.2	Diskussion	628
	Standortwahl	628
	Sequentieller Marktzutritt	629
	Sortimentsvertiefung	629
7.1.3	Maximale oder minimale Differenzierung?	630
7.2	Monopolistischer Wettbewerb	634
7.3	Werbung und Produktdifferenzierung durch Informationsvermittlung	637
7.3.1	Ansichten über die Werbung	639
7.3.2	Suchgüter und informative Werbung	640
7.3.2.1	Monopolistischer Wettbewerb	641
7.3.2.2	Oligopole	645
7.3.3	Erfahrungsgüter: Informationsbedingte Produktdifferenzierung und Goodwill	651
7.4	Abschließende Bemerkungen	653
7.5	Anhang: Vertikale Differenzierung und monopolistischer Wettbewerb	654
7.5.1	Vertikale Differenzierung	654
7.5.2	Ein symmetrisches Modell mit monopolistischem Wettbewerb	659
	Hinweise zur Lösung der Übungsaufgaben	664
	Literaturhinweise	667
8	Marktzutritt, Anpassung und Marktaustritt	671
8.1	Fixkosten: Natürliche Monopole und Wettbewerbsfähigkeit eines Marktes	675

8.1.1	Fixkosten und versunkene Kosten	676
8.1.2	Wettbewerbsfähigkeit von Märkten	678
8.1.3	Abnutzungskrieg	685
8.2	Versunkene Kosten und Marktzutrittsschranken	694
	Das Stackelberg-Spence-Dixit-Modell	694
8.2.1	Anpassung, Abschreckung und Blockade	695
8.2.2	Mögliche Erweiterungen und Modifikationen	701
8.2.2.1	Gewinnfunktionen in reduzierter Form	701
8.2.2.2	Eine Mehrzahl etablierter Unternehmen	707
8.2.2.3	Marktzutritt mit dem Ziel, das etablierte Unternehmen aufzukaufen	709
8.2.2.4	Unsicherheit	710
8.2.2.5	Verzögerungen beim Aufbau von Kapazitäten	711
8.2.3	Andere Arten von Kapitalgütern	711
8.3	Eine Taxonomie von Unternehmensstrategien	714
8.3.1	Abschreckung	717
8.3.2	Anpassung	722
8.3.3	Verdrängung	726
8.4	Anwendungen der Taxonomie	727
	Beispiel 1: Freiwillige Beschränkung der Produktionskapazität	728
	Beispiel 2: Produktdifferenzierung	729
	Beispiel 3: Erfahrungskurveneffekte	730
	Beispiel 4: Meistbegünstigungsklausel	732
	Beispiel 5: Ein Oligopol auf mehreren Märkten	738
	Beispiel 6: Zölle und Kontingente	739
	Beispiel 7: Vertikale Bindung	740
	Beispiel 8: Koppelung	741

Beispiel 9: System- und Produktkompatibilität	744
8.5 Nachwort: Preise und Mengen	748
8.6 Anhang: Strategisches Verhalten, Marktzutritts- und Mobilitätsschranken	750
8.6.1 Kapitalakkumulation	750
8.6.1.1 Kurzfristige Kapitalakkumulation und Wettbewerbsfähigkeit	750
Dissipation der Rente ohne positive Wohlfahrtseffekte	751
Wettbewerbsfähigkeit	755
Keine fixen Kosten: Die Dynamik des Cournot-Wettbewerbes	759
8.6.1.2 Langfristige Kapitalakkumulationsspiele	763
8.6.2 Sortimentvertiefung, Gewinnung eines Vorsprungs und die Dauerhaftigkeit einer Monopolstellung	769
8.6.2.1 Sortimentvertiefung	770
8.6.2.2 Kann ein Monopolist eine Marktnische besetzen und glaubhaft machen, daß er sie nie wieder verlassen wird?	776
8.6.2.3 Sind Monopole von Dauer?	780
Hinweise zur Lösung der Übungsaufgabe	784
Literaturhinweise	793
9 Information und strategisches Verhalten: Reputation, Limitpreise und Verdrängungspreise	803
9.1 Statischer Wettbewerb unter asymmetrischer Information	804
9.1.1 Ein einfaches Modell des Preiswettbewerbs	804
9.1.1.1 Informationsaustausch	808
9.1.1.2 Versteigerungen	810
9.1.2 Einige kritische Anmerkungen	810
9.2 Dynamik: Eine heuristische Betrachtung	812
9.3 Anpassung und Parallelverhalten	814
9.4 Das Milgrom-Roberts-Modell des höchsten eintrittsverhindernden Preises	819

9.4.1	Ein Modell	820
9.4.1.1	Ableitung des Intervalls der Trennungsgleichgewichtspreise	824
9.4.1.2	Analyse eines Trennungsgleichgewichts	826
9.4.1.3	Pooling-Gleichgewichte	828
9.4.2	Zur Kritik dieser Schlußfolgerungen	830
9.5	Verdrängungspreise und Unternehmenszusammenschlüsse	836
9.6	Präsenz auf einer Mehrzahl von Märkten und Reputation	840
9.7	Die Geschichte von der dicken Brieftasche	844
9.8	Abschließende Bemerkungen	850
9.9	Anhang: Darwinsche Zuchtwahl in einem Wirtschaftszweig	852
	Hinweise zur Lösung der Übungsaufgaben	862
	Literaturhinweise	865
10	Forschung und Entwicklung und die Einführung neuer Technologien	871
10.1	Innovationsanreize und Marktstruktur: Der Wert von Innovationen	874
10.1.1	Der wohlwollende Diktator	874
10.1.2	Monopol	875
10.1.3	Wettbewerb	876
10.1.4	Ein Monopol, das von einem potentiellen Konkurrenten bedroht wird	878
10.2	Patentwettläufe - eine Einführung	881
10.2.1	Ein Modell	882
10.2.2	Diskussion	887
10.2.2.1	Wahl der Technologie	887
10.2.2.2	Patentwettlauf und Erfahrungskurven	891
10.3	Wohlfahrtstheoretische Analyse des Patentwesens	894
10.4	Anreize, FuE zu betreiben	897

10.5	Die Markteinführung von Innovationen: strategischer Erwägungen	899
10.5.1	Eine Innovation, die zu imitieren sich nicht empfiehlt	901
10.5.2	Rasche Imitation und einhelliges Zögern bei der Markteinführung	903
10.6	Netzexternalitäten, Standardisierung und Kompatibilität	906
10.6.1	Die Nachfrageseite: Die Koordination der Erwartungen der Verbraucher	910
10.6.2	Die Angebotsseite: Förderer und strategisches Verhalten	915
10.7	Abschließende Bemerkungen	917
10.8	Anhang: Lizenzen und Forschungs-Joint-Ventures	919
10.8.1	Der Anreiz zur Erteilung von Lizenzen	919
10.8.2	Ex post-Lizenzvergabe und der Wert eines Patents	921
10.8.2.1	Eine große Innovation	922
10.8.2.2	Eine kleine Innovation	924
10.8.3	Lizenzvergabe ex ante	926
10.8.4	Forschungs-Joint-Ventures	927
	Hinweise zur Lösung der Übungsaufgaben	930
	Literaturhinweise	939
11	Nichtkooperative Spieltheorie: Eine Gebrauchsanweisung	945
11.1	Spiele und Strategien	945
11.2	Nash-Gleichgewicht	949
11.3	Perfektes Gleichgewicht	957
11.3.1	Spiele mit perfekter Information	959
11.3.2	Spiele mit "beinahe perfekter" Information	964
11.4	Bayessches Gleichgewicht	967
11.4.1	Spiele mit unvollkommener oder unvollständiger Information	967
11.4.2	Statische Spiele mit unvollständiger Information	968

11.5	Das perfekte Bayessche Gleichgewicht	975
11.5.1	Ein Zwei-Perioden-Reputationsspiel	980
11.5.2	Sequentielles Verhandlungsspiel mit unvollständiger Information	984
11.5.3	Garantie als ein Signal für Qualität	987
11.5.4	Signalstörung	990
11.6	Anhang	994
11.6.1	Existenz eines Gleichgewichtes	994
11.6.1.1	Existenz eines Bayesschen Gleichgewichtes	995
11.6.1.2	Existenz eines perfekten Gleichgewichtes	995
11.6.1.3	Existenz eines perfekten Bayesschen Gleichgewichtes	996
11.6.2	Verfeinerungen	998
11.6.2.1	Das Signalspiel	1000
	Eliminierung schwach dominierter Strategien	1001
	Eliminierung von schwach dominierten Gleichgewichtsstrategien	1003
11.6.2.2	Beispiele	1005
	Hinweise zur Lösung der Übungsaufgaben	1016
	Literaturhinweise	1024
	Wiederholungsaufgaben	1031
	Register	1059

Vorwort der Übersetzer

Die Industrieökonomik stellt heute eines der Kerngebiete der volkswirtschaftlichen Theorie dar. Aus unserer Sicht sollte sie nicht nur Studenten der Volkswirtschaftslehre gelehrt werden, sondern auch und gerade Studenten der Betriebswirtschaftslehre. Die sowieso stets fragwürdige Aufteilung der Wirtschaftswissenschaften in Volks- und Betriebswirtschaftslehre ist durch die neuere Entwicklung der Industrieökonomik vollends obsolet geworden. In Bereichen wie der Organisationstheorie und der Absatzwirtschaft gehen die Lehrgebiete der VWL und der BWL nun bruchlos ineinander über, und die Industrieökonomik befindet sich genau auf der Schnittstelle.

Das Angebot an deutschsprachigen Lehrbüchern der Industrieökonomik läßt bislang zu wünschen übrig. Deswegen haben wir uns entschlossen, ein führendes amerikanisches Lehrbuch ins Deutsche zu übertragen. Der Marktanteil dieses Buches am Weltmarkt für (anspruchsvolle) industrieökonomische Lehrbücher ist erheblich. Den Leser mag es interessieren, daß unlängst auch eine französische Übersetzung erschienen ist.

Bei der Übertragung ins Deutsche haben wir uns auf den Standpunkt gestellt, daß Abweichungen vom Original mitunter unvermeidlich sind, um eben diesem Original gerecht zu werden. Auch haben wir nicht versucht, gewaltsam für jeden amerikanischen Begriff ein deutsches Äquivalent zu prägen. Sprachliche Puristen werden an unserem Text infolgedessen wenig Gefallen finden.

Die Arbeit der Übertragung war folgendermaßen aufgeteilt: Peter Seidelmann übertrug die Kapitel 1 - 4, Roland Ladwig die Kapitel 5 - 7 sowie die Anhänge zu den Kapiteln 8 - 10, Bruno Schönfelder den Rest. Antje Klimczak hat sich an der Übersetzung des Kapitels 11 beteiligt. Die Gesamtedaktion oblag Schönfelder. Tatkräftig mitgewirkt haben eine ganze Reihe von Studenten der Technischen Universität Bergakademie Freiberg. Insbesondere wollen wir uns bei Annette Fischer, Eireen Junge, Antje Klimczak, Henry Kost, Diana Kunath, Anja Reuter, Ralf Sippel und Katrin Zöbisch bedanken. Unser besonderer Dank gilt Evelyn Richter, die den Text mit rastlosem Fleiß geschrieben hat.

Die Lösungshinweise, die Tirole zu seinen Übungsaufgaben gibt, zeichnen sich durch Kürze aus und wir vermuten, daß sie mitunter Fragen offenlassen. Henry Kost hat ausführlichere Lösungen ausgearbeitet. Diese können gegen einen Unkostenbeitrag von uns bezogen werden. Interessenten mögen sich wenden an Prof. Dr. B. Schönfelder, Fakultät für Wirtschaftswissenschaften, TU Bergakademie Freiberg, Lessingstraße 45, D-09596 Freiberg, Fax 03731/394092. e-mail: schoenfe@orion.hrz.tu-freiberg.de.

Roland Ladwig, Bruno Schönfelder, Peter Seidelmann

Vorwort

Die theoretische Industrieökonomik hat seit den frühen siebziger Jahren wesentliche Fortschritte erzielt und ist zu einem der wichtigsten Bestandteile der mikroökonomischen Theorie geworden. Dieses Buch stellt einen Versuch dar, die neuere Entwicklung darzustellen und den Zusammenhang zur traditionellen Industrieökonomik aufzuzeigen.

Für ihre Ratschläge, Ermutigung und Kritik danke ich insbesondere Philippe Aghion, Roland Bénabou, Patrick Bolton, Bernard Caillaud, Franklin Fisher, Paul Joskow, Bruno Jullien, Eric Maskin, Patrick Rey, Garth Saloner, Richard Schmalensee und Michael Whinston. Dilip Abreu, Kyle Bagwell, John Bonin, Joel Demski, Peter Diamond, Drew Fudenberg, Robert Gertner, Robert Gibbons, Roger Guesnerie, Oliver Hart, Bengt Holmström, Jean-Jacques Laffont, Ariel Rubinstein, Stephen Salant, Steve Salop, Carl Shapiro, Andrea Shepard, Marius Schwartz und Oliver Williamson gaben sehr nützliche Hinweise zu bestimmten Kapiteln.

Paul Joskow und Richard Schmalensee haben mich bestärkt, dieses Unternehmen zu wagen; sie lasen das ganze Manuskript und gaben mir viele kritische Hinweise. Ich stehe schon seit meiner Studienzeit am Massachusetts Institute of Technology (MIT) in ihrer Schuld. Sie unterrichteten mich in Industrieökonomik. Frühere MIT-Studenten werden in der Herangehensweise und der Wahl der Schwerpunkte den Einfluß der Lehrveranstaltung 14.271 wiedererkennen, die Paul und Richard hielten. Mit Drew Fudenberg und Eric Maskin fühle ich mich ebenfalls seit meiner Studienzeit verbunden und nicht nur wegen ihres Einflusses auf meinen Standpunkt zur Industrieökonomik. Insbesondere meine Sicht der spieltheoretischen Aspekte der Industrieökonomik wurde von unserer Zusammenarbeit geprägt. Eric lehrte mich die Spieltheorie und zeigte mir, wie ihr Instrumentarium nutzbringend auf unterschiedliche ökonomische Fragestellungen angewendet werden kann. Drews Beitrag zu diesem Buch ist beinahe so unübersehbar, daß sich ein Hinweis auf ihn erübrigt. Der Teil II und die Gebrauchsanweisung für die Spieltheorie beruhen zu einem großen Teil auf unseren gemeinsamen Forschungsarbeiten und Übersichtsaufsätzen. David Kreps, Paul Milgrom, John Roberts und Robert Wilson haben meine Ansichten zum strategischen Verhalten auf Märkten stark beeinflusst. Außerdem stehe ich in der Schuld zahlreicher herausragender Wissenschaftler, die die heutige Theorie der Industrieökonomik geschaffen haben und deren Arbeiten ich vielfach zitiere.

Den Stoff dieses Buches habe ich - in jeweils unterschiedlicher Form - an der Ecole Nationale de la Statistique et de l'Administration Economique und an der Universität von Lausanne, am MIT und der Ecole des Hautes Etudes en Sciences Sociales gelehrt. Dabei

handelte es sich sowohl um Vorlesungen für "undergraduates" als auch um solche für "graduate students". Die Studenten dieser Schulen machten nützliche Bemerkungen und gaben weiterführende Anregungen.

Ich möchte Benjamin Hermalin danken, der für mich ein hervorragender Assistent war. Er las das gesamte Manuskript, rechnete die Übungsaufgaben durch und gab mir viele klarsichtige Hinweise. Sowohl was seinen Fleiß als auch sein Talent anbelangt, hätte ich mir keinen besseren Assistenten wünschen können. Ich bin auch Bernice Soltysik dankbar für die Anfertigung des Registers und Bruno Jullien für die Korrektur der Druckfahnen.

Während eines Aufenthalts an der Ecole Nationale des Ponts et Chaussées begann ich dieses Buch zu schreiben. Aus meinem Manuskript für die Vorlesungen an der ENSAE über Industrieökonomik im Jahre 1983 entstand das Buch *Concurrence Imperfaite*, das 1985 von Editions Economica verlegt wurde. John und Hélène Bonin übersetzten dieses Buch kompetent und mit großem Eifer, und dies ermöglichte es mir, kurz nach seiner Veröffentlichung auf der Grundlage der französischen Version *The Theory of Industrial Organisation* in Angriff zu nehmen.

Dieses Buch ist das Ergebnis der Anstrengungen vieler, aber die Verantwortung für Fehler oder Lücken trage ich allein. Ich entschuldige mich bei meinen Ratgebern dafür, daß ich viele ihrer guten Anregungen zur Vervollständigung und Erweiterung der Darstellung ignoriert habe. Der einzige Grund dafür ist, daß sich der Umfang des Buches sonst verdoppelt hätte.

Ich war froh, daß Paul Bethge der Herausgeber war. Er hat das Manuskript mit viel Verstand, Sachkenntnis und Humor bearbeitet. Ich danke auch dem gesamten Team von The MIT Press - besonders dem Lektor für Wirtschaftswissenschaften Terry Vaughn für seine Hilfe und Unterstützung und der Designerin Rebecca Daw für ihre gute Bearbeitung einiger reichlich schwieriger Textstellen.

Ganz besonderer Dank gilt der erfahrenen, mehrsprachigen Emily Gallagher, die das Schreiben von vielen Entwürfen, sowohl in französisch als auch in englisch, ertrug. Mit guter Laune verbrachte sie unzählige Abende und Wochenenden, um das Material für die Studenten und die Publikation fertigzustellen. Sie leistete eine hervorragende Arbeit. Der erste französischsprachige Entwurf wurde von Patricia Maillebouis und Pierrette Vayssade geschrieben.

Ich bin dankbar für die großzügige Förderung durch die National Science Foundation, die Sloan Foundation, das Center for Energy Policy Research am MIT, den MIT Fund for the Arts, Humanities, and Social Sciences und das Commissariat Général au Plan.

EINLEITUNG

Warum Industrieökonomik?

Diese Frage klingt beinahe töricht. Industrieökonomik zu studieren bedeutet, sich mit der Funktionsweise von Märkten zu befassen, und dies ist in der Mikroökonomie doch ein zentrales Thema. Es dauerte nichtsdestoweniger eine lange Zeit und bedurfte zweier Anläufe, bevor die Industrieökonomik zu einem der Hauptgebiete der Volkswirtschaftslehre wurde.¹

Die erste Anlauf ist mit den Namen Joe Bain und Edward Mason verbunden und wird manchmal als die "Harvard-Tradition" bezeichnet. Sie war empirischer Natur. Daraus entwickelte sich das berühmte "structure-conduct-performance" Paradigma, nach dem die Marktstruktur (die Anzahl der Verkäufer auf dem Markt, der Grad der Produktdifferenzierung, die Kostenstruktur, der Grad an vertikaler Integration der Hersteller usw.) das Verhalten determiniert (zum Verhalten werden gezählt: Preisbildung, Forschung und Entwicklung, Investitionen, Werbung und ähnliches mehr) und dieses wiederum die Marktergebnisse bestimmt (Effizienz, Verhältnis zwischen Preis und Grenzkosten, Produktvielfalt, Innovationsrate, Gewinne und Absatz usw.). Dies klingt zwar plausibel, aber es mangelte an einer befriedigenden theoretischen Fundierung. Im Vordergrund der Forschungsbemühungen standen stattdessen empirische Untersuchungen. Beispielsweise fand man heraus, daß bestimmte Kennziffern für das Verhalten ("conduct") oder das Ergebnis ("performance") eng mit der Marktstruktur korreliert sind. Eine typische Regressionsgleichung hatte die Form $\Pi_i = f(CR_i, BE_i, \dots)$, wobei i die Branche bezeichnet, Π_i ein Maß der Unternehmens- bzw. Branchenrentabilität, CR_i ein Konzentrationsmaß (ein Maß, das zusammenfassend darstellen soll, wie viel oder wenig Wettbewerb in einer Branche herrscht) und BE_i ("barriers to entry") Variablen meint, mit denen man die Schwierigkeiten eines Marktzutritts messen wollte (man versuchte dies mit der mindestoptimalen Betriebsgröße, mit dem Verhältnis zwischen Werbungsausgaben und Verkaufseinnahmen und ähnlichen Variablen). Andere Variablen könnte man ebensogut in die Regression einbeziehen. Bei der Regressionsanalyse verwendete man normalerweise Querschnittsdaten, die aus einer großen Zahl von Branchen stammten.²

¹ Dieser Abschnitt beruht auf Vorlesungen von Paul Joskow und Richard Schmalensee.

² Schmalensee (1986) vermittelt einen ausgezeichneten und mit kritischen Anmerkungen versehenen Überblick über diesen Forschungsansatz.

Wenn man über die vielen Meßprobleme hinwegsieht, lieferten solche Regressionen eine Anzahl nützlicher stilisierter Fakten. Die Zusammenhänge zwischen den Variablen, die dabei deutlich wurden (oder auch das Fehlen solcher Zusammenhänge), muß man als Korrelation oder "beschreibende Statistik" interpretieren und nicht als kausale Beziehung. Im vorangegangenen Beispiel stellen die Rentabilität der Branchen, ihr Konzentrationsgrad und das Verhältnis zwischen Werbung und Absatz endogene Variablen dar. Sie werden simultan von den "Grundbedingungen", denen der Markt unterliegt (exogene Variable), und durch das Verhalten der Unternehmen bestimmt.

Die Unmöglichkeit einer kausalen Interpretation stellt den Analytiker vor Probleme. Was ist von einer Regression zu halten, die zeigt, daß die Rentabilität einer Branche mit ihrem Konzentrationsgrad wächst?³ Dies mag bedeuten, daß die Unternehmen in stärker konzentrierten Branchen Marktmacht besitzen und daß das Marktergebnis solcher Zweige möglicherweise nicht optimal ist. Jedoch sagt das wenig über die Ursachen von Konzentration oder Marktmacht aus, und unser Analytiker erhält keinerlei Hinweise, ob und in welcher Weise ein wirtschaftspolitischer Eingriff das Marktergebnis verbessern kann.

Sicherlich unternahm die empirische Forschungsrichtung den Versuch, diese (exogenen) Grundbedingungen zu erfassen: die Technologie (Skalenerträge, Marktzutrittskosten, den Anteil der versunkenen Kosten, die Existenz einer Erfahrungskurve, langlebige oder kurzlebige Güter etc.), Präferenzen und Konsumentenverhalten (Struktur der Informationen über die Produktqualität, Reputation und Markentreue etc.), "exogenen" technischen Fortschritt usw.⁴ Obwohl hierbei durchaus Fortschritte erzielt wurden, ist es oft schwierig, die Daten zusammenzutragen, die erforderlich sind, um die "Grundbedingungen" mit dem nötigen Maß an Genauigkeit zu messen und ihre Vergleichbarkeit über die Branchengrenzen hinweg sicherzustellen.

³ Tatsächlich ist dieser statistische Zusammenhang wenig ausgeprägt. Erweitert man die exogenen Variablen um eine, die den Marktanteil mißt, so verschwindet der Zusammenhang zwischen der Ertragsrate und dem Konzentrationsgrad fast vollständig. Eine Erklärung dafür könnte sein, daß in Branchen, in denen sich die Kosten der einzelnen Unternehmen relativ stark unterscheiden, der Konzentrationsgrad (einige wenige Unternehmen mit niedrigen Kosten erzeugen den größten Teil des Outputs) und die Gewinne zumeist hoch sind (die besagten Unternehmen sind dem Wettbewerb nur wenig ausgesetzt). Es kann also sein, daß die Variable Konzentrationsgrad nur deswegen als wesentlich erscheint, weil die Variable Marktanteil nicht verwendet wurde. Vergleiche Kapitel 5.

⁴ Die Skalenerträge beispielsweise wurden mit unterschiedlichen Methoden gemessen. Der größten Beliebtheit hat sich wohl ein Verfahren erfreut, bei dem die Kostenfunktion geschätzt wird (genauer die Parameter einer Kostenfunktion, wobei der Output und die Inputpreise als exogene Variablen fungieren). Bain ermittelte "ingenieurwissenschaftlich fundierte" Produktionsfunktionen, indem er auf Unternehmensebene technische Daten sammelte; ferner griff er auf ein Verfahren zurück, das er "engineering managerial analysis" nannte: er befragte beispielsweise Manager nach der optimalen Größe ihres Unternehmens. George Stigler verwendete eine andere Methode, er sammelte Daten über die Größe derjenigen Unternehmen, die im Wettbewerb überleben.

Das mit diesen Äußerungen angedeutete Urteil über die empirische Tradition ist sicherlich allzu streng - schließlich war sie es, die die Industrieökonomik überhaupt erst geschaffen hat. Außerdem ließen wir unerwähnt, daß als Zugabe zu den Regressionsanalysen eine Fülle nicht formalisierter Erklärungsversuche mitgeliefert wurden. (In der Tat ging den Regressionsanalysen eine Vielzahl von Branchenstudien voraus, und diese wurden durch Kartellverfahren veranlaßt und ermöglicht. Die Fallstudien führten zum Entstehen informaler Erklärungsversuche. Die Regressionen ergänzten die Branchenstudien.) Solche Erklärungsversuche und die Analysen und Fallstudien, die anläßlich von Kartellverfahren erstellt wurden, bildeten die Grundlage für die spätere Entwicklung der Theorie.

Es ist auch unfair zu behaupten, daß die strenge Theorie vollkommen links liegen gelassen wurde. Insbesondere die "Chicagoer Tradition", die mit Aaron Director und George Stigler einsetzte, betonte die Notwendigkeit einer gründlichen theoretischen Analyse und einer empirischen Identifikation der konkurrierenden Theorien. In methodologischer Hinsicht übte die Chicagoer Tradition einen wichtigen Einfluß auf die Entwicklung unseres Faches aus. Sie ist darüberhinaus für ihre sehr permissive Sicht des Marktverhaltens bekannt - z.B. hinsichtlich vertikaler Kontrolle und Verdrängungspreisen (Kollusion ist ihrer Ansicht nach das Hauptvergehen) - und für ihr Mißtrauen gegenüber staatlichen Eingriffen, das sie deutlich von der Harvard-Tradition unterschied. Nichtsdestoweniger herrschte in den frühen siebziger Jahren die Meinung vor, daß die industrieökonomische Theorie wohl zur Erklärung bestimmter statistischer Ergebnisse oder einer Abstützung weltanschaulicher Positionen beitrage, aber nicht den Charakter einer strengen und systematischen Analyse besitze. Paul Joskow zum Beispiel gab folgende Einschätzung:

In gewissem Sinn ist die eigentliche Probe für die Brauchbarkeit der verschiedenen Modelle, ob sie sich für Leute als nützlich erweisen, die sich mit der Analyse tatsächlicher Märkte beschäftigen. Ich behaupte, daß sie nicht nur nicht besonders nützlich sind, sondern daß sie auch nicht wirklich angewendet werden... In gewisser Hinsicht kann man sich des Eindrucks nicht erwehren, daß das Wesentliche eigentlich in den informalen Erklärungsversuchen, "Geschichten" und Beschreibungen von Verhalten steckt, und daß die formalen Modelle ex post aus dem Boden gestampft werden, um zu zeigen, daß ein formaler Apparat einiges von dem, was man tatsächlich beobachtet, erklären oder integrieren kann. (Joskow, 1975, S. 273)

Die zweite Welle industrieökonomischer Forschung, die vorwiegend theoretisch orientiert war, setzte in den siebziger Jahren ein. Sie läßt sich auf drei Faktoren zurückführen. Was die "Nachfrage" nach Theorie anbetrifft, so haben wir bereits auf den wachsenden Unmut über die begrenzte Leistungsfähigkeit der Querschnittsanalysen hingewiesen, die damals das gesamte Gebiet der Industrieökonomik prägten. Vielfach stand man unter dem Eindruck, daß

zwischen der empirischen Arbeit und den formalen Modellen oligopolistischer Märkte im Grunde kein Zusammenhang bestand. Was das "Angebot" an Theorie anbetrifft, so spielten zwei Faktoren eine Rolle. Erstens ließen die eigentlichen Wirtschaftstheoretiker (bis auf wenige Ausnahmen) die Industrieökonomik bis in die siebziger Jahre hinein links liegen, da sie sich im Unterschied zur Theorie des allgemeinen Gleichgewichts nicht für elegante und allgemeine Formulierungen eignete. Seitdem aber entwickelten eine Reihe maßgeblicher Theoretiker ein Interesse für die Industrieökonomik.⁵ Zweitens - und dies war für die im Teil II dieses Buches behandelten Themen von Bedeutung - bot sich die Theorie nicht-kooperativer Spiele (hierüber gab es aber durchaus Meinungsverschiedenheiten) als Instrumentarium für die Analyse strategischer Konflikte an; auf diese Weise gelang eine Vereinheitlichung der Methodologie. Der Spieltheorie gelangen auch auf zwei wichtigen Gebieten deutliche Fortschritte, nämlich bei der Analyse von dynamischen Spielen und von Spielen mit asymmetrischer Information. Damit war die Zeit reif für die Überarbeitung der vielfältigen informalen Erklärungsversuche, die damals im Umlauf waren.

In diesem Buch stehen die Ergebnisse der zweiten Welle im Vordergrund. Darüber soll aber nicht vergessen werden, daß die älteren Arbeiten die Fundamente für die stärker theoretisch ausgerichteten Analysen der siebziger und achtziger Jahre legten. Ich glaube, daß die Industrieökonomien sich soweit über methodologische Fragen geeinigt haben, daß man guten Gewissens ein Lehrbuch der Industrieökonomik schreiben kann.

Wir haben über die Dogmengeschichte im Bereich der Industrieökonomik gesprochen, ohne überhaupt zu definieren, was das ist oder darzulegen, warum sie wichtig ist. In der Tat möchte ich es eigentlich vermeiden, mich auf eine genaue Definition einzulassen, sind doch die Grenzen des Faches fließend. Die Industrieökonomik befaßt sich - soviel steht fest - mit der Struktur und Funktionsweise von Unternehmen (ihrer Strategie auf dem Markt und ihrer Organisation). Daß sich die Industrieökonomik mit Unternehmensstrategien befaßt, ist der Grund, warum einige Professoren der Betriebswirtschaftslehre wichtige Beiträge zur Industrieökonomik geleistet haben. Aber die Industrieökonomik reduziert sich nicht auf eine Lehre von den Unternehmensstrategien. Die andere Seite der Medaille ist, daß "Außenstehende" (z.B. Professoren, Bürokraten oder die Mitarbeiter des Kartellamtes) das Marktergebnis beurteilen wollen oder müssen. Es ist wenig wahrscheinlich, daß unvollkommene Märkte (und das sind die meisten realen Märkte) ein Pareto-optimales Ergebnis

⁵ Damit läßt sich aber schwerlich das "Marktergebnis" auf dem Markt für industrieökonomische Theorien erklären. Die "Struktur" dieses Marktes (z.B. die Zahl fähiger Theoretiker, die sich zeitweise oder auf Dauer auf diesem Markt tummeln) ist selbst endogen. Um eine Erklärung für den Zustrom von "Anbietern" zu finden, muß man nach den tieferliegenden Gründen fragen.

erbringen. Daraus folgt nun nicht notwendigerweise, daß der Staat (der "wohlwollende Diktator") durch sein Eingreifen etwas verbessern kann - gerade angesichts der (häufig ungenügenden) Information, über die er verfügt, ist das durchaus zweifelhaft - und die Beobachtung eines "Marktversagens" liefert auch noch keinen Hinweis, wann und wie der Staat intervenieren sollte. Die Analyse unvollkommener Märkte zeigt immerhin eines: die Schlußfolgerungen aus Arbeiten, die vom Modell des vollkommenen Wettbewerbs ausgehen, können sowohl in positiver als auch in normativer Hinsicht höchst fragwürdig sein.

Was als ein staatlicher Eingriff betrachtet werden sollte, versteht sich ebenfalls keineswegs von selbst. Wettbewerbsförderung mit Hilfe des Kartellrechts soll uns ebenso als Eingriff gelten wie verschiedene Mittel der "Marktregulierung" (Steuern und Subventionen, Vorgabe von Qualitätsstandards etc.); sie kommen im Laufe des Buches zur Sprache.⁶ Andere Instrumente, wie die Regulierung der Preise, des Marktzutritts oder Eingriffe, die den Status des betroffenen Unternehmens selbst berühren (Gewährung eines Monopols, Verstaatlichung oder auch das Verhalten des Staates auf Märkten, auf denen er der wichtigste Auftraggeber ist), werden nicht behandelt. Dies hat natürlich einige ungünstige Auswirkungen. Insbesondere werden in den betrachteten Modellen jeweils nicht sämtliche möglichen Eingriffe durchdiskutiert. Die einzige Entschuldigung hierfür (von dieser Entschuldigung mache ich reichlichen Gebrauch) ist, daß das Buch sowieso schon sehr dick ist. Die verschiedensten Regulierungsmöglichkeiten zu behandeln, hätte seinen Umfang leicht verdoppeln können.

Theoretische und empirische Analyse

Die Industrieökonomik hat sich in den letzten Jahren in ein ziemlich stark theoretisch orientiertes Fach verwandelt. Zweifellos sollte selbst ein Theoretiker bedauern, daß es nunmehr sehr viel Theorie, aber vergleichsweise wenig empirische Evidenz zugunsten der einen oder anderen Theorie gibt, und das auf einem Gebiet, wo es theoretischen Modellen oft an Allgemeingültigkeit mangelt und praktische Implikationen sehr wichtig sind. Obwohl ich die Auffassung teile, daß hier ein Ungleichgewicht entstanden ist, glaube ich trotzdem, daß die Entwicklung der Theorie sehr hilfreich war.

⁶ Mit "Marktregulierung" sind Regulierungen gemeint, die alle Unternehmen einer Branche (einschließlich potentieller Konkurrenten) gleichermaßen betreffen.

Die "neue theoretische Industrieökonomik", die auf der alten Tradition verbaler Verhaltensschilderung und den stilisierten Fakten⁷ aufbaut, die im Laufe der Zeit zusammengetragen wurden, leistet m. E. einen Beitrag zur Analyse realer Märkte. Sie hat nicht nur einige der überlieferten verbalen Erklärungsversuche formalisiert; sie hat es auch ermöglicht, andere zu verwerfen. Die praktische Bedeutung soll hier aber nicht überbetont werden. Möglicherweise hat man zu viel Wert auf die positive Analyse gelegt, zu Lasten der normativen, und zu wenig getan, um dem Anwender eine Entscheidung zwischen den konkurrierenden Theorien zu ermöglichen. Nichtsdestoweniger hat die industrieökonomische Theorie zweifellos einen empirischen Gehalt. Darüberhinaus ist damit zu rechnen, daß die theoretischen Beiträge der empirischen Analyse neue Anregungen vermitteln werden.⁸ Die Theorie gibt Hinweise, worüber empirisches Material gesammelt werden sollte, sie ermöglicht es, endogene von exogenen Variablen zu unterscheiden und liefert die Hypothesen, die getestet werden sollen.

Eine ökonometrische Analyse ist sicherlich nicht die einzige Möglichkeit, in der Industrieökonomik empirische Untersuchungen durchzuführen. Wegen des unzureichenden Datenmaterials stützt sich die angewandte Forschung bei der Untersuchung des Unternehmens- und Branchenverhaltens und der Marktergebnisse häufig auf detaillierte Fallstudien von Unternehmen oder Branchen (das Material, das für kartellrechtliche Verfahren zusammengetragen wird, läßt sich hierfür ebenfalls verwenden). Diese Studien sind zwar auch mit Problemen behaftet, aber sie liefern viele interessante Einblicke. In der Tat ziehen es Industrieökonomien häufig vor, mit solchen Fallstudien statt mit statistischen Analysen zu arbeiten - dies mag daran liegen, daß es vielleicht eher gelingt, die Grundbedingungen und Verhaltensweisen einer Branche aufgrund inhaltsreicher Fallstudien zu erfassen als mit Hilfe bruchstückhafter Statistiken über Gewinne, Konzentrationsgrad, Werbungsausgaben usw., die einer mehr oder minder großen Stichprobe entnommen wurden.

Eine weitere Methode zur Sammlung von Beweismaterial, die von den Entwicklung der Theorie profitieren kann, ist die Durchführung von Experimenten unter Laborbedingungen.⁹

⁷ Scherer (1980) liefert eine bemerkenswerte Sammlung von Fakten über das Verhalten von Unternehmen und hat damit maßgeblich zur Entwicklung der modernen Theorie beigetragen.

⁸ Neuere Branchenstudien vermitteln den Eindruck, daß sich eine solche Entwicklung tatsächlich bereits anbahnt. Einige hervorragende Beiträge hierzu sind Bresnahan (1987a), Joskow (1985) und Porter (1983). Einen guten Überblick verschafft Bresnahan (1987b).

⁹ Einen Überblick über diesen Ansatz in der Industrieökonomik vermittelt Plott (1982).

So ist also zu hoffen, daß diese drei Ansätze der empirischen Forschung von der neueren theoretischen Entwicklung befruchtet werden. Das Buch stellt nicht die empirischen Implikationen der einzelnen Modelle dar und erklärt auch nicht, wie man sich zwischen konkurrierenden Modellen entscheiden kann. Trotzdem hoffe ich, daß die Darstellung der Modelle einleuchtend genug ist, um zu verdeutlichen, welche ihrer Aussagen einer empirischen Überprüfung zugänglich sind.

Zum Inhalt des Buches

Die empirische Seite des Faches (die auch die kartellrechtliche Praxis umfaßt) wird in dem Buch nicht dargestellt. Vernachlässigt werden auch einige wichtige Bereiche der Theorie, so die Regulierungstheorie, Industrieökonomik in offenen Volkswirtschaften¹⁰, allgemeine Gleichgewichte mit unvollkommenem Wettbewerb¹¹ und der Zusammenhang zwischen Industrieökonomik und Makroökonomie.¹²

In methodologischer Hinsicht wird ein in sich geschlossener Ansatz verfolgt. Teil I unterstellt Optimierungsverhalten, und Teil II wendet eine Verallgemeinerung der Optimierung auf das Entscheidungsverhalten einer Mehrzahl von Personen an, nämlich die Theorie der nichtkooperativen Spiele. Andere Ansätze wie das sog. satisficing (begrenzte Rationalität) werden nicht berücksichtigt. Der Vorteil einer solchen Vorgehensweise ist, daß eine einheitliche, in sich geschlossene Darstellung des Stoffes möglich wird.

Der Einfachheit halber werden die Unternehmen als Ganzes als die Entscheidungsträger betrachtet, sie maximieren ihre Gewinne. Im größten Teil des Buches werden die Probleme der Kontrolle der Aktionäre, der Banken oder des Kapitalmarktes über die Manager per Annahme aus der Welt geschafft. Die unternehmensinterne Delegation und Kontrolle von Entscheidungsbefugnissen werden ebenfalls vernachlässigt. Im Einleitungskapitel zur Theorie der Unternehmung werden diese Annahmen erörtert. In den Kapiteln 4 und 9 verweisen wir gelegentlich auf das sogenannte Agency-Problem. Da eine angemessene Behandlung der hierfür erforderlichen Theorie ein eigenes Buch erfordern würde, begnüge ich mich in diesem Buch damit, diese Themen nur zu erwähnen, obwohl sie dringend einer Vertiefung bedürften. In der Tat bin ich der Auffassung, daß die Schnittstelle zwischen Organisationstheorie und

¹⁰ Eine neuere Arbeit über die Industrieökonomik einer offenen Wirtschaft ist Helpman und Krugman (1985).

¹¹ Einen guten Überblick über die einschlägige Literatur vermittelt Hart (1985).

¹² Vergleiche Carlton (1987).

Industrieökonomik eines der interessantesten Gebiete für die theoretische Forschung der kommenden Jahre sein wird.

Natürlich mußte ich mich entscheiden, auf welche Themen ich mich konzentrieren sollte - und dies war manchmal nicht leicht. Meine Präferenzen haben dabei eine Rolle gespielt, aber diese sind nicht mit meinen Werturteilen identisch. Sie sind vielmehr von meinem Wissensstand und dem Stand meiner Überlegungen geprägt. Ich entschuldige mich bei den Autoren, deren Beiträge ich zu wenig hervorgehoben oder aufgrund meiner Unkenntnis, Gedächtnislücken oder Fehleinschätzungen vernachlässigt habe. Darüberhinaus schlägt sich in der Auswahl des Stoffes auch eine bestimmte Vorgehensweise bei der Erklärung der Theorie nieder. Einige interessante Arbeiten, die eine lange und mathematisch aufwendige Erklärung erfordern, habe ich in Fußnoten, Bemerkungen oder Übungen verbannt.

Hinweise zur Benutzung des Buches - Gliederung und Aufbau

Im Teil I (Kapitel 1 - 4) werden diejenigen Aspekte des Marktverhaltens betrachtet, die nichts mit strategischem Verhalten zu tun haben (aber mit ihm sicherlich nicht unvereinbar sind). Untersucht wird die Entscheidung eines Monopolisten über Preis und Qualität, über sein Sortiment, seine Werbung und seinen Vertrieb. Die meisten der dort gewonnenen Schlußfolgerungen lassen sich auf Oligopole übertragen. Teil II analysiert die Entscheidung über den Preis, die Kapazität, die Positionierung des Produkts, Forschung und Entwicklung und andere strategische Variablen eines Oligopolisten. Er macht regen Gebrauch von einigen Grundbegriffen der Spieltheorie.

Ich hielt diese Unterteilung in zwei Teile aus pädagogischen Gründen für nützlich. Indem er strategische Gesichtspunkte für einige Wochen vernachlässigt, kann sich der Student mit einigen der wichtigsten Themenbereiche der Industrieökonomik vertraut machen, ohne durch eine gleichzeitige Einführung der Spieltheorie überfordert zu werden. Die Unterteilung erleichtert es ferner, klar herauszuarbeiten, welche Verhaltensweisen nur durch strategisches Verhalten zu erklären sind. Ich habe eine "Gebrauchsanweisung" der nicht-kooperativen Spieltheorie (Kapitel 11) eingefügt, um dem Leser zu helfen, sich auf den zweiten Teil vorzubereiten. Diese soll aber kein Substitut für einen stärker mathematisch ausgerichteten Kurs in Spieltheorie darstellen. Sie soll die wichtigsten Instrumente der Spieltheorie vorstellen und den Leser in ihre Anwendung einführen. Ein zusätzlicher Kurs in Spieltheorie wäre nützlich, aber er ist keineswegs Voraussetzung für eine Lektüre von Teil II. Ich schlage den Lesern, die mit der Spieltheorie nicht vertraut sind, vor, das Kapitel 11 bis zum Abschnitt 11.4 durchzuarbeiten, während sie sich mit Teil I des Buches beschäftigen, und die

Abschnitte 11.4 und 11.5 zu lesen, ehe sie in das Kapitel 9 einsteigen.

Wie bereits erwähnt, ist dem Teil I eine Abhandlung über die Theorie der Unternehmung vorangestellt. Die Unternehmung ist Hauptgegenstand des Buches, und deshalb sollten wir ihre Natur und ihre Ziele untersuchen. Die Untersuchung des Unternehmens liefert eine Grundlage für die Analyse von Märkten. Ich fürchte (und ich hoffe gleichzeitig), daß der Leser von meinen Ausführungen nicht sonderlich befriedigt sein wird. Das Kapitel erregt wohl den Verdacht, daß es mich vor Gewissensbissen schützen (nach dem Motto "Jetzt haben wir das Unternehmen besprochen und dürfen es im weiteren ignorieren und als einen schwarzen Kasten behandeln, der Gewinne maximiert") und es mir ermöglichen soll, ein Lippenbekenntnis zum untrennbaren Zusammenhang zwischen der Organisation des Unternehmens und der Marktstruktur abzulegen. In Wirklichkeit aber habe ich dieses Kapitel nicht deswegen geschrieben, weil es die nachfolgenden ergänzt, sondern weil meiner Überzeugung nach die Theorie der Unternehmung für die Volkswirtschaftslehre ein wichtiges Thema und darüberhinaus ein integraler Teil der Industrieökonomik ist.

Zusammenhänge zwischen den Kapiteln

Die einzelnen Kapitel sind weitgehend in sich abgeschlossen. Trotzdem lassen sich einige Zusammenhänge feststellen. Zum Beispiel verallgemeinern Kapitel 5 und 6 den Inhalt des Kapitels 1 auf strategisches Verhalten, und dasselbe Verhältnis weist das Kapitel 7 zum Kapitel 2 auf. Die Hypothese von der Dissipation der Rente wird, nachdem sie in Kapitel 1 nur erwähnt wird, im Teil II eingehender behandelt. Die Fähigkeit eines Unternehmens, sich die Konsumentenrente anzueignen, wird in Kapitel 2 zum ersten Mal thematisiert, und in den Kapiteln 7 und 10 kehren wir dann zu diesem Thema zurück. Analogien wie diese werden im Laufe des Buches herausgestellt, wo dies als angemessen erscheint.

Untergliederung der Kapitel

Jedes Kapitel gliedert sich in einen Hauptteil und einen Anhang. Undergraduates (Studenten im dritten und vierten Studienjahr gemäß amerikanischem Hochschulsystem, nach dem vierten Studienjahr erwirbt man dort den Titel des bachelor - baccalaureus), Graduates im ersten Jahr (dies ist im amerikanischen System das fünfte Studienjahr) und Wissenschaftlern, die mit der Industrieökonomik nicht vertraut sind, wird empfohlen, sich auf den Hauptteil zu konzentrieren. Die anderen Graduates in höheren Studienjahren und Wissenschaftler, die mit der Industrieökonomik vertraut sind, werden das weiterführende Material in den Anhängen möglicherweise nützlich finden.

Übungsaufgaben

Die Übungsaufgaben wurden eingearbeitet, um dem Leser eine Gelegenheit zur Anwendung der Modelle zu verschaffen und sein Wissen zu erweitern. Einige Übungsaufgaben erscheinen innerhalb eines Kapitels; ihre Lösungen sind am Ende des Kapitels skizziert. Weitere Wiederholungsübungen (ohne Lösungen) finden sich am Ende des Buches. Leser, denen es nicht gelingt, diese Aufgaben zu lösen, sollten nicht verzagen: Einige der Aufgaben sind schwierig. Diese Leser werden bei den Wiederholungsübungen auch einfachere Aufgaben finden. Ich habe die Aufgaben mit Sternchen indiziert, um ihren Schwierigkeitsgrad zu kennzeichnen:

- * einfache Anwendung der im Text entwickelten Konzepte
- ** schwieriger, erfordern mehr Nachdenken
- *** fortgeschritten, eine Herausforderung

Voraussetzungen

Volkswirtschaftslehre

Der Leser sollte an einem Kurs in Mikroökonomie für mäßig fortgeschrittene Studenten teilgenommen haben. Ansonsten setzt das Buch wenig voraus. Immer wenn dies als notwendig erschien, habe ich versucht, den Leser auf das jeweilige Modell hinzuführen. Kenntnisse der stilisierten Fakten würden einen guten Hintergrund für die Modelle abgeben. Dem Leser kann empfohlen werden, vorher oder gleichzeitig das Standardwerk von Scherer (1980) zu lesen. Die Lektüre von Büchern über strategische Unternehmensplanung (z.B. Porter 1980) und Kartellrecht (z.B. Areeda 1974, Areeda und Turner 1976, Blair und Kasperman 1985, Posner und Easterbrook 1981) kann sich ebenfalls als nützlich erweisen.

Mathematik

Ich habe versucht, die Theorien in einer "leserfreundlichen" Form zu präsentieren. Ich habe oft spezielle Funktionen (etwa lineare Nachfragefunktionen) statt der allgemeinen verwendet, Zwei-Perioden-Modelle statt dynamischer und Duopole statt Oligopole. Meine Hoffnung ist, daß der zugrundeliegende Zusammenhang hinreichend deutlich wird, um den Leser zu überzeugen, daß die Ergebnisse auch in allgemeineren Modellen gelten.

Es werden nur geringe Mathematikkennnisse vorausgesetzt. Für den größten Teil des Buches genügen Grundkenntnisse in Differentialrechnung (z.B. Optimierung ohne Nebenbedingungen). Der Leser sollte wissen, wie man die erste und zweite Ableitung bildet, das "Enveloppen-Theorem"¹³ und die Kettenregel der Differentiation kennen und etwas Ahnung von Konkavität¹⁴ haben. In einzelnen Unterabschnitten kommen noch andere Methoden (partielle Integration, Bayessche Regel, dynamische Programmierung) zur Anwendung. Die erforderliche Mathematik ist in einfacher Form in den mathematischen Anhängen von Dixit und Norman (1980) und von Varian (1978) dargestellt.¹⁵

¹³ Gemäß diesem Theorem ist die Ableitung des Wertes, der als Ergebnis eines Optimierungsproblems gefunden wurde (d.h. der die Zielfunktion maximiert), nach einem exogenen Parameters gleich der partiellen Ableitung der Zielfunktion nach diesem Parameter. Das heißt, nur der direkte Effekt, den dieser Parameter auf die Zielfunktion ausübt, braucht berücksichtigt zu werden, nicht der indirekte, mittelbare, der darauf zurückzuführen ist, daß die Änderung des Parameters einen Einfluß auf die endogene (Kontroll-) Variable ausübt. Letzteren Einfluß bezeichnen wir auch als zweitrangig. Mit anderen Worten, wenn

$$V(a) = \max_x f(x, a)$$

dann ist

$$\frac{dV}{da} = \frac{\partial f}{\partial a}(x^*(a), a),$$

wobei $x^*(a)$ die optimale Kontrollvariable darstellt.

¹⁴ Eine Funktion $f(x)$, wobei x ein Vektor in R^n ist, ist konkav, wenn für jedes λ im Intervall $[0, 1]$ und alle x und x' gilt:

$$f(\lambda x + (1 - \lambda)x') \geq \lambda f(x) + (1 - \lambda)f(x').$$

Bei einer differenzierbaren Funktion läßt sich Konkavität auch folgendermaßen bestimmen: Für alle x und x' gilt:

$$f(x) \leq f(x') + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x')(x_i - x'_i).$$

(Ist $x \in R$, kann sich der Leser hiervon überzeugen, indem er ein Diagramm zeichnet.)

Die Funktion f ist quasikonkav, wenn die Menge der R^n , die durch $\{x | f(x) \geq y\}$ bestimmt wird, für alle y konvex ist. Ist $x \in R$, dann ist die folgende etwas stärkere Bedingung hinreichend für Quasikonkavität: $f'' < 0$ immer dann, wenn $f' = 0$. Wie sich mit Hilfes eines Diagramms leicht zeigen läßt, benötigen wir nur Quasikonkavität im letzteren Sinne (dies ist eine schwächere Bedingung als die der Konkavität), damit die Bedingungen zweiter Ordnung in einem Optimierungsproblem erfüllt sind. Wir werden sie aber nicht sehr oft benötigen.

¹⁵ Detaillierte Darstellungen der Optimierungstheorie finden sich bei Dixit (1976) und Kamien und Schwartz (1981).

Definition eines Marktes, partielles Gleichgewicht und Wohlfahrtskriterien**Das Paradigma des vollkommenen Wettbewerbs**

Das am weitesten entwickelte und ästhetisch ansprechendste Modell der Volkswirtschaftslehre ist das Paradigma des allgemeinen Gleichgewichts von Arrow und Debreu.¹⁶ Es hat kurz gesagt die folgende Gestalt: Die Analyse beginnt mit einer sehr genauen Beschreibung der verfügbaren Güter. Ein ökonomisches Gut wird charakterisiert durch seine physischen Eigenschaften, den Zeitpunkt der Verfügbarkeit, die Beschaffenheit, in welcher es verfügbar ist, seine Lage usw. Die Konsumenten sind vollständig über alle Produkteigenschaften informiert und besitzen Präferenzen, die in bezug auf Güterbündel definiert sind. Die Hersteller (Unternehmen) - die Konsumenten sind Eigentümer dieser Unternehmen - verfügen über bestimmte Produktionsmöglichkeiten. Hinzu kommt ein Paradigma, das darstellt, wie die Märkte funktionieren. Alle Marktteilnehmer nehmen die Preise als vorgegeben hin. Die Konsumenten maximieren ihren Nutzen, wobei ihre Ausgaben nicht ihr Einkommen (welches von ihrer Ausstattung mit Produktionsfaktoren und dem Eigentum an Unternehmen herrührt) überschreiten dürfen. Dies ergibt Nachfragefunktionen (bzw. -korrespondenzen, wenn die Zahl der Güterbündel, die, gegeben einen Preisvektor, den Nutzen des Konsumenten maximieren, größer als 1 ist). Die Produzenten maximieren ihre Gewinne unter Berücksichtigung ihrer technologischen Möglichkeiten, und dies liefert die Angebotsfunktionen (bzw. Angebotskorrespondenzen). Ein Gleichgewicht mit vollkommenem Wettbewerb ist eine Menge von Preisen, die die Nachfrage und das Angebot bestimmen, und besitzt die Eigenschaft, daß sämtliche Märkte (einer für jedes Gut) geräumt werden (die Gesamtnachfrage nicht das Gesamtangebot übersteigt).

Schwache Annahmen über Präferenzen und technologische Möglichkeiten genügen für die Ableitung einiger allgemeiner Theoreme, die im allgemeinen Gleichgewicht Geltung besitzen. Die bekanntesten davon sind wohl die zwei Hauptsätze der Wohlfahrtsökonomik. Der erste besagt, daß ein allgemeines Gleichgewicht mit vollkommenem Wettbewerb (ein Walras-Gleichgewicht) Pareto-optimal ist (das bedeutet, ein wohlwollender und allwissender Diktator kann die Allokation, die der Markt hervorbringt, nicht durch eine andere zulässige Allokation ersetzen, die allen Konsumenten ein höheres Nutzenniveau stiftet). Der zweite besagt, daß unter einer zusätzlichen Annahme jede Pareto-optimale Allokation "dezentral"

¹⁶ Die Darstellungen dieser Theorie durch Debreu (1959), Arrow und Hahn (1970) und Mas-Colell (1985) sind von klassischer Schönheit. Eine vereinfachte Version dieser Theorie findet sich bei Varian (1978) (für fortgeschrittene Studenten) und in verschiedenen mikroökonomischen Lehrbüchern für Anfänger.

(d.h. vom Markt) herbeigeführt werden kann, indem die entsprechenden Preise festgesetzt werden und eine entsprechende Umverteilung des Einkommens zwischen den Konsumenten vorgenommen wird. Die erforderliche zusätzliche Annahme besagt, daß die Technologie konvex ist (dies schließt zunehmende Skalenerträge aus).

Eine wesentliche Eigenschaft des Walras-Gleichgewichts ist, daß jedes Gut zu seinen Grenzkosten verkauft wird. Würde der Preis eines Gutes seine Grenzkosten übersteigen, so würde der Hersteller die Produktion ausdehnen - dies würde nämlich zu einer Gewinnerhöhung führen. Übersteigen die Grenzkosten den Preis und findet dennoch eine Produktion statt, so reduziert der Hersteller seinen Ausstoß. Dieses an sich triviale Ergebnis hat wichtige Implikationen. Bei seiner Entscheidung, ob er von einem Gut eine Einheit mehr konsumieren soll, steht der Konsument dem "richtigen" Preis gegenüber; dieser Preis internalisiert die Produktionskosten dieser zusätzlichen Einheit. Dies ist einer der tieferen Gründe, warum das Wettbewerbsgleichgewicht Pareto-optimal ist.

Der erste Hauptsatz der Wohlfahrtsökonomik setzt den Bemühungen der Industrieökonom enge Grenzen. Wirtschaftszweige sind in solch einer Welt stets effizient organisiert. Der einzige mögliche Grund für wirtschaftspolitische Eingriffe ist die Einkommensverteilung - es kann sein, daß ein wohlwollender Diktator die Einkommensverteilung nicht für optimal hält.¹⁷

Das Paradigma des Walras-Gleichgewichts beruht, wie wir dargelegt hatten, auf relativ schwachen Annahmen über Präferenzen und Produktionsmöglichkeiten, aber bei näherem Hinsehen erweisen sich diese Annahmen als durchaus restriktiv. Zu den geforderten Voraussetzungen gehören das Fehlen von Externalitäten zwischen den Wirtschaftssubjekten¹⁸, daß es sich um private Güter handelt¹⁹ und die vollständige Information der

¹⁷ Das heißt nicht, daß das allgemeine Gleichgewicht mit vollkommenem Wettbewerb für die Analyse industrieökonomischer Probleme völlig unnütz ist. Einige positive Aussagen lassen sich mit Hilfe einfacher Modelle mit vollkommenem Wettbewerb ableiten. Beispielsweise gibt es ein interessantes Schrifttum, das mit Hilfe des Paradigmas des vollkommenen Wettbewerbs Aussagen über den Eintritt neuer Anbieter in einen dynamischen Wirtschaftszweig und den Austritt alter Anbieter ableitet und das Gibrat-Gesetz theoretisch zu fundieren versucht (gemäß dem die Wachstumsrate eines Unternehmens im Durchschnitt unabhängig von seiner Größe ist). Diese Literatur wird in diesem Buch nicht besprochen. Der interessierte Leser wird auf Lucas (1978), Jovanovic (1982), Lippman und Rummelt (1982) und Hopenhayn (1986) verwiesen.

¹⁸ Ein externer Effekt tritt auf, wenn die Konsumtion eines Gutes durch einen Endverbraucher einen direkten Einfluß auf das Nutzenniveau eines anderen Konsumenten ausübt oder wenn die Produktion eines Unternehmens andere Wirtschaftssubjekte beeinflusst. Ein Konsument, der die Größe des Telefonnetzes verändert, indem er einen Anschluß an dieses Netz erwirbt, löst einen positiven externen Effekt auf andere Konsumenten aus. Ein Unternehmen, das einen Fluß verschmutzt, erzeugt einen negativen externen Effekt, der sowohl Konsumenten als auch andere Unternehmen betrifft.

¹⁹ Ein öffentliches Gut ist ein Gut, das gleichzeitig von mehreren Konsumenten konsumiert werden kann (z.B. Landesverteidigung oder ein Fernsehprogramm).

Konsumenten über die Produkte. Externalitäten und öffentliche Güter gehören traditionell zum Forschungsobjekt der Finanzwissenschaft, wir werden uns aber dennoch mit einigen Situationen befassen, in denen externe Effekte (z.B. sogenannte Netzexternalitäten) und öffentliche Güter (z.B. Vorschriften, die den Hersteller eines neuen Gutes zur Bereitstellung von Informationen verpflichten) eine Rolle spielen. Wir werden auch die dritte Annahme aufgeben, wenn wir uns mit unvollkommener Information der Konsumenten über bestimmte Güter (z.B. über ihre Qualität) beschäftigen.

Möglicherweise springt von all den Annahmen, die dem Walras-Gleichgewicht zugrundeliegen, am stärksten die ins Auge, daß alle Wirtschaftssubjekte sämtliche Preise als vorgegeben betrachten. Sicherlich lassen sich Märkte benennen (z.B. Märkte für landwirtschaftliche Güter), für die diese Annahme in etwa zutrifft, aber auf den meisten Märkten gibt es nur eine kleine Anzahl von Anbietern und diese verfügen über eine nennenswerte Marktmacht.²⁰ Das Buch befaßt sich vor allem mit den Gründen für diese Marktmacht und mit ihren Folgen.

Partielles Gleichgewicht, fallende Nachfragekurve und Konsumentenrente

Sobald man einige der Annahmen, die dem Walras-Gleichgewicht zugrundeliegen, auflockert, läßt sich über die Allokation nicht mehr sehr viel aussagen, sofern man nicht bereit ist, andere spezifische Annahmen einzuführen. Dies hat uns die "Theorie des Zweitbesten" gelehrt. Einer der Nachteile, die in Kauf genommen werden müssen, um in der Industrieökonomik zu realitätsnäheren Modellen zu kommen, ist der Rückzug auf die Partialanalyse, die Analyse partieller Gleichgewichte. Ein Gut (oder eine Gruppe verwandter Güter) wird herausgegriffen und die Wechselwirkungen mit dem Rest der Wirtschaft werden vernachlässigt.

Wir werden in Kürze auf den Begriff des Marktes zurückkommen; zunächst wollen wir zwei Annahmen erörtern, die das ganze Buch hindurch getroffen werden: Erstens, daß die Nachfrage nach einem Gut mit seinem Preis fällt, und zweitens, daß die Änderungen der Wohlfahrt der Konsumenten mit Hilfe der sogenannten Konsumentenrente gemessen werden können.

²⁰ Einen geistvollen Versuch, die Differenz zwischen Preis und Grenzkosten in unterschiedlichen Wirtschaftszweigen zu schätzen, hat Hall (1986) unternommen.

Als erstes wenden wir uns dem Begriff der Konsumentenrente zu: Betrachten wir einen Markt für ein einziges Gut. Wir nehmen an, daß die Nachfrage nach diesem Gut mit dem Preis dieses Gutes fällt und daß sie unabhängig vom Preis anderer Güter und vom Einkommen der Konsumenten ist. Um dies zu präzisieren, unterstellen wir "quasi-lineare" Nutzenfunktionen:

$$U(q_0, q_1, \dots, q_m) = q_0 + \sum_{h=1}^m V_h(q_h),$$

wobei Gut 0 das Numéraire ist, und die Funktionen V_h positive Steigung besitzen und konkav sind. Wir maximieren U , wobei die Budgetrestriktion als Nebenbedingung zu beachten ist:

$$q_0 + \sum_{h=1}^m p_h q_h \leq I.$$

Hierbei bezeichnet I das Einkommen des Konsumenten. Die Lösung des Maximierungsproblems führt zur Bedingung $V'_h(q_h) = p_h$ für alle h . Die Nachfragefunktion eines jeden Konsumenten nach dem Gut h und folglich auch die Gesamtnachfragefunktion erfüllen also die obigen Bedingungen. (Im Falle der allgemeineren quasi-linearen Präferenzen $U(q_0, q_1, \dots, q_m) = q_0 + W(q_1, \dots, q_m)$ weisen die Nachfragefunktionen Kreuz-Preis-Effekte auf, aber keine Einkommenseffekte.)

Betrachten wir ein homogenes Gut. Dupuit (1844) führte als erster eine Messung des Wohlfahrtsniveaus durch. (Dupuits Konsumentenrente wird manchmal auch als Marshallsche Konsumentenrente bezeichnet, vergleiche Marshall [1920, S. 811]. Im weiteren Verlauf der Darstellung werden wir einfach von *der Konsumentenrente* sprechen.²¹) Im Bild 1 wird die Konsumentenrente durch die Fläche zwischen der Nachfragekurve und der horizontalen Linie beim Preisniveau p^0 dargestellt.

²¹ Auerbach (1986) vermittelt einen umfassenden Überblick über verschiedene Ansätze zur Messung der Konsumentenrente und der Zusatzlast. Vergleiche hierzu ferner die klassischen Arbeiten von Hicks (1941) und Samuelson (1947).

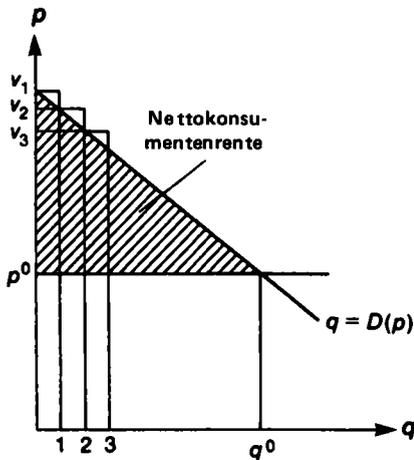


Abbildung 1

Dupuit behauptete, daß diese Fläche ein Maß abgibt, wieviel die Konsumenten statt dem von ihnen tatsächlich bezahlten Betrag ($p^0 q^0$) für die Möglichkeit, q^0 Einheiten des Gutes zu konsumieren, zu zahlen bereit gewesen wären. Die zugrundeliegende Überlegung läßt sich am einfachsten erläutern, indem man annimmt, daß die Nachfragekurve ein Aggregat aus einer Vielzahl von Nachfragefunktionen darstellt, wobei jeder dieser Nachfrager höchstens eine Einheit des Gutes kauft. Das bedeutet, es gibt viele Konsumenten, welche entweder 0 oder 1 Einheit des Gutes kaufen. Die Konsumenten sind heterogen und messen dem Gut einen unterschiedlichen Nutzen bei, ihre Zahlungsbereitschaft, ausgedrückt in Geldeinheiten, beträgt v_i . (Das bedeutet für die oben eingeführten quasi-linearen Nutzenfunktionen, daß $V_h(\cdot)$ für jeden Konsumenten eine Treppenfunktion ist, die den Wert 0 annimmt, wenn vom Gut h weniger als eine Einheit konsumiert wird, und den Wert v_i , wenn der Konsument i eine oder mehr als eine Einheit des Gutes konsumiert.) Ohne Einbuße der Allgemeingültigkeit können wir diese Konsumenten nach ihrer Zahlungsbereitschaft anordnen: $v_1 \geq v_2 \geq \dots$. Ein Konsument mit der Wertschätzung v_i kauft, wenn und nur wenn $v_i \geq p^0$. Der erste Konsument realisiert eine Rente von $v_1 - p^0$, da er bereit gewesen wäre, v_1 zu zahlen. Der zweite Konsument realisiert eine Rente von $v_2 - p^0$ und so geht das weiter bis zum Grenzkonsumenten (nennen wir ihn n), der keine Rente erzielt. Die gesamte Konsumentenrente ist dann:

$$(v_1 - p^0) + (v_2 - p^0) + \dots + (v_{n-1} - p^0).$$

Für große n kann die Treppenfunktion durch die stetige Gesamtnachfragefunktion $q = D(p)$ approximiert werden, und die Konsumentenrente beträgt dann:

$$S^n = \int_{p^0}^{\bar{p}} D(p) dp. \quad (1)$$

S^n ist die Nettokonsumentenrente. Die Bruttokonsumentenrente S^g ist gleich der Nettokonsumentenrente plus Ausgaben des Konsumenten $p^0 D(p^0)$. \bar{p} bezeichnet den Prohibitivpreis (den niedrigsten Preis, bei dem die Nachfrage null beträgt). In der diskreten Version des Modells ist er gleich v_1 , aber wir müssten an der Formel nichts ändern, wenn er unendlich wäre. Betrachten wir nun einen einzelnen Konsumenten mit einer fallenden Nachfragekurve $D(\cdot)$ nach einem Gut. Gemäß Dupuits Argumentation kann man sich diesen Konsumenten so vorstellen, als bestünde er aus vielen Konsumenten, von denen jeder nur eine einzige Einheit nachfragt. Das bedeutet, er ist bereit, für die erste gekaufte Einheit v_1 zu zahlen, v_2 für die zweite usw. Die gesamte Nettokonsumentenrente, die ihm der Konsum von q^0 Einheiten des Gutes verschafft, die er zum Preis p^0 gekauft hat, wird durch die Gleichung 1 dargestellt. Von jetzt an betrachten wir einen einzelnen Konsumenten. Erst später werden wir auf den Fall mehrerer Konsumenten zurückkommen.

Die Veränderungen der Netto- und Bruttokonsumentenrente, zu denen eine Erhöhung des Preises von p^0 auf p^1 führt, werden durch folgende Gleichungen beschrieben:

$$\Delta S^n = - \int_{p^0}^{p^1} D(p) dp, \quad (2)$$

$$\Delta S^g = - \int_{p^0}^{p^1} D(p) dp + [p^1 D(p^1) - p^0 D(p^0)].$$

Den Gewinn des Unternehmens bezeichnet man auch als Produzentenrente. Bild 2 zeigt die Grenzkostenkurve (welche bei vollkommenem Wettbewerb mit der Angebotskurve übereinstimmt). Der Gewinn ist gleich dem Erlös ($p^0 D(p^0)$) abzüglich der Kosten. Die Kosten sind das Integral der Grenzkosten.²² Der Gewinn, der in dem betreffenden Wirtschaftszweig erzielt wird, ist also gleich der Fläche zwischen der Grenzkostenkurve und der horizontalen Linie in Höhe des Preisniveaus p^0 (die vertikal schraffierte Fläche im Bild 2).

Die gesamte Rente, die der Wirtschaftszweig erzeugt, ist die Summe aus Konsumentenrente und Produzentenrente. Die Rente erreicht ihren größtmöglichen Wert, wenn der Preis, den die Konsumenten zahlen müssen, gleich den Grenzkosten ist (p^0 im Bild 2).

²² Wenn Fixkosten auftreten, müssen sie noch zusätzlich vom Gewinn abgezogen werden.

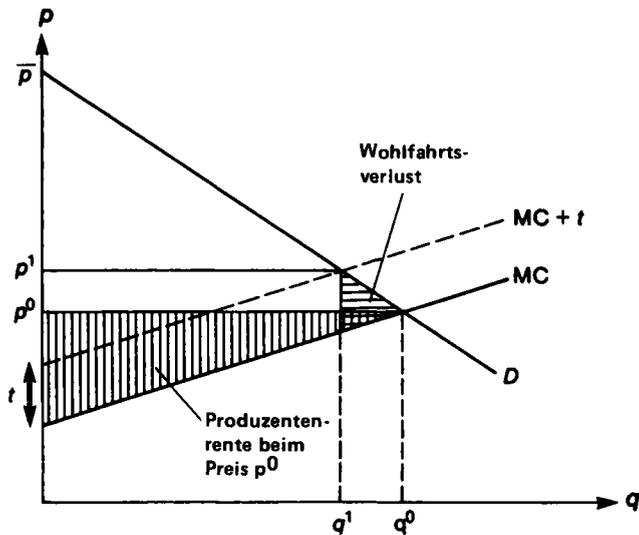


Abbildung 2

Eine berühmte Anwendung dieser Analyse ist die Messung des Wohlfahrtsverlustes (Zusatzlast, "dead-weight loss"), welcher durch eine spezielle Verbrauchsteuer hervorgerufen wird. Nehmen wir an, in der betreffenden Branche herrsche vollkommener Wettbewerb, so daß vor der Auflage der Steuer die Allokation Pareto-optimal ist. Wir erheben eine Mengensteuer in der Höhe t . Das neue Gleichgewicht wird dann beim Preis p^1 und der Konsummenge q^1 erreicht. Der Wohlfahrtsverlust ist gleich der Verringerung der (gesamten) Rente infolge der Steuer und läßt sich durch die Fläche des horizontal schraffierten Dreiecks in Bild 2 messen. (Die Rente umfaßt auch die Steuereinkünfte, die der Staat erzielt. Der einfachste Weg, die Rente zu berechnen, besteht darin, die Gesamtkosten von der Brutto-konsumentenrente abzuziehen. Für diese Berechnungen sind Transfers zwischen den Konsumenten, Produzenten und dem Staat irrelevant.) Gemäß dem Vorschlag von Dupuit läßt sich die Zusatzlast der Steuer mit dem Ausdruck

$$\frac{1}{2}t|q^1 - q^0| = \frac{1}{2}t^2|D'(p^1)|$$

approximieren, wenn die Steuer "klein" ist und die Grenzkosten konstant sind.²³

²³ Hotelling (1938) und Harberger (1964) lieferten einen mathematischen Beweis hierfür und verallgemeinerten die Aussage auf den Fall einer Mehrzahl von Gütern. Boiteux (1951) und Debreu (1951) entwickelten Methoden, um die Zusatzlast im allgemeinen Gleichgewicht zu messen.

Der Rest dieses Abschnittes ist mathematisch anspruchsvoller als der größte Teil der Einleitung und kann beim ersten Lesen übersprungen werden. Er diskutiert die Ausdehnung des Konzepts der Konsumentenrente auf den Fall mehrerer Güter und versucht dann zu klären, unter welchen Bedingungen die Nachfragefunktion nach jedem Gut negative Steigung aufweist und die Konsumentenrente eine gute Approximation darstellt.

Zwei Fragen zur Konsumentenrente drängen sich auf: Läßt sie sich auf den Fall einer Mehrzahl von Gütern verallgemeinern? Läßt sie sich in der Sprache der klassischen Nachfragetheorie darstellen? Um diese Fragen zu beantworten, betrachten wir die allgemeinen Nachfragefunktionen $q_h = D_h(\mathbf{p}, I)$, bei denen die Nachfrage nach Gut h vom Preisvektor \mathbf{p} und dem Einkommen I der Konsumenten abhängt.

Die Verallgemeinerung auf den Fall einer Mehrzahl verschiedener Güter scheint höchst einfach: Man braucht nur die Konsumentenrenten für die verschiedenen Güter zu addieren. Die Veränderung der Nettokonsumentenrente, wenn der Vektor \mathbf{p}^0 durch den Vektor \mathbf{p}^1 ersetzt wird, beträgt dann

$$\Delta S^n = \sum_h \Delta S_h = - \int_{\mathbf{p}^0}^{\mathbf{p}^1} \sum_h q_h dp_h. \tag{3}$$

Eine unerfreuliche Eigenschaft dieser Konsumentenrente ist, daß Gleichung 3 im Mehrproduktfall nicht immer eine eindeutige Lösung besitzt. Wie sich unschwer feststellen läßt, hängt der Wert des Integrals im allgemeinen davon ab, in welcher Reihenfolge die Integration vorgenommen wird. Diese Reihenfolge ist nur dann unwesentlich (und das Integral damit wohlbestimmt), wenn die Nachfragefunktionen keinen Einkommenseffekt aufweisen (oder allgemeiner, wenn die gemischten partiellen Ableitungen der Nachfragefunktionen gleich sind).²⁴

Zwischen diesem Problem und den Grundlagen der Nachfragetheorie besteht tatsächlich ein Zusammenhang. Hicks (1946) führte zwei weitere Maße für Veränderungen des Konsumentennutzens bei einer Preisveränderung ein: die "kompensierende Variation" (compensating variation) - der Geldbetrag, den man dem Konsumenten geben muß, damit sein

²⁴ Um dies zu erkennen, wollen wir zwei Preise betrachten: wir verändern erst den einen und dann den anderen, und im Anschluß verfahren wir gerade umgekehrt. Wir ermitteln die Differenz zwischen den beiden Maßen und schreiben die Nachfragefunktion jeweils als das Integral ihrer partiellen Ableitungen. Dann erhalten wir den Ausdruck $\partial D_2/\partial p_1 - \partial D_1/\partial p_2$. Dieser Ausdruck hätte den Wert 0, wenn es sich um die kompensierte Nachfrage handeln würde (dies folgt aus der Symmetrie der Slutsky-Matrix). Aber es handelt sich nicht um die kompensierte Nachfrage, sondern um die Marshallsche. Wer seine Kenntnisse über kompensierte Nachfragefunktionen auffrischen möchte, sei auf Varian (1978) verwiesen. Eine gute Darstellung der ganzen Problematik findet sich bei Auerbach (1986).

Nutzenniveau von der Preisveränderung unbeeinflusst bleibt - und die "äquivalente Variation" (equivalent variation) - der Einkommensbetrag, auf den der Konsument zu verzichten bereit ist, um eine Preisveränderung zu vermeiden.²⁵

Ein wesentlicher Begriff für die weitere Analyse ist die Hickssche oder kompensierte Nachfragefunktion $D^c(\mathbf{p}, u)$. Dies ist die Nachfrage, die man erhält, wenn das Einkommen bei Preisveränderungen so angepaßt wird, daß der Konsumentennutzen auf dem Niveau u konstant bleibt. Die kompensierte Nachfragefunktion hat für jedes Gut negative Steigung. Die Beziehung zwischen der "normalen" und der kompensierten Nachfragefunktion läßt sich mit Hilfe der Slutsky-Gleichung darstellen. Greifen wir ein Gut h heraus, dann lautet sie:

$$\frac{\partial D_h}{\partial p_h} = \frac{\partial D_h^c}{\partial p_h} - D_h \frac{\partial D_h}{\partial I}.$$

Dies bedeutet, daß die Änderung der Nachfrage nach einem Gut, die durch eine Veränderung des Preises dieses Gutes um eine Einheit hervorgerufen wird, gleich der Summe aus zwei Ausdrücken ist. Der erste ist die Ableitung der kompensierten Nachfrage nach dem Preis, und diese ist, wie wir bereits gesehen haben, negativ. Dieser Term wird als *Substitutionseffekt* bezeichnet. Der zweite Term, der *Einkommenseffekt*, ist für ein normales Gut negativ und für ein inferiores Gut positiv (und im Fall der quasi-linearen Nutzenfunktionen beträgt er null). Er rührt daher, daß eine Erhöhung von p_h um eine Einheit den Konsumenten D_h Einheiten seines Einkommens kostet, was die Nachfrage nach Gut h um $\partial D_h / \partial I$ verändert.

²⁵ Die Funktion, die das Maximierungsproblem

$$E(\mathbf{p}, u) = \min_{\mathbf{q}} \{\mathbf{p} \cdot \mathbf{q}\},$$

unter der Nebenbedingung $U(\mathbf{q}) \geq u$ löst, trägt den Namen Ausgabenfunktion des Konsumenten, wobei $U(\cdot)$ die Nutzenfunktion symbolisiert. Die Ausgabenfunktion gibt uns das Einkommen, das erforderlich ist, um das Nutzenniveau u zu erreichen, wenn der Preisvektor \mathbf{p} vorgegeben ist. Entsprechend bezeichnet

$$V(\mathbf{p}, I) = \max_{\mathbf{q}} U(\mathbf{q}),$$

die indirekte Nutzenfunktion, wobei die Maximierung unter der Nebenbedingung $\mathbf{p} \cdot \mathbf{q} \leq I$ erfolgt. Ändert sich der Preisvektor von \mathbf{p}^0 zu \mathbf{p}^1 , dann beträgt die kompensierende Variation

$$CV = E(\mathbf{p}^1, V(\mathbf{p}^0, I)) - I$$

und die äquivalente Variation

$$EV = I - E(\mathbf{p}^0, V(\mathbf{p}^1, I)).$$

Die äquivalente Variation ist insofern ein akzeptables Wohlfahrtsmaß, als der Vergleich zwischen den Wohlfahrtsniveaus bei den Preisen \mathbf{p}^1 und \mathbf{p}^2 - d.h. $V(\mathbf{p}^1, I)$ und $V(\mathbf{p}^2, I)$ - einem Vergleich zwischen den äquivalenten Variationen von \mathbf{p}^0 nach \mathbf{p}^1 einerseits und von \mathbf{p}^0 nach \mathbf{p}^2 andererseits äquivalent ist. Für die kompensierende Variation trifft dies im allgemeinen nicht zu.

Der Wert, den diese zwei Variationen bei multiplen Preisveränderungen annehmen, hängt nicht davon ab, in welcher Reihenfolge diese Preisveränderungen vorgenommen werden.

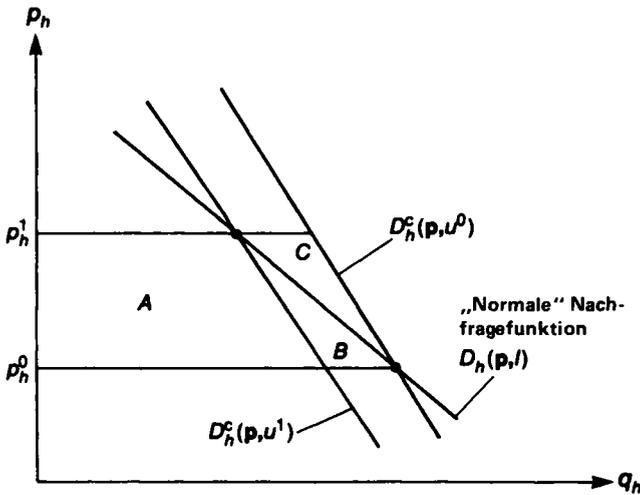


Abbildung 3

Betrachten wir eine Preisveränderung, die nur ein einziges Gut betrifft (sagen wir Gut h). Es läßt sich leicht erkennen, daß die "äquivalente" (beziehungsweise "kompensierende") Variation von Preis p^0 zu Preis p^1 mit der Fläche unter der "kompensierten" Nachfragekurve beim Nutzenniveau u^1 (beziehungsweise u^0) identisch ist.²⁶ Wenn wir davon ausgehen, daß $D(p^0, I) \equiv D^c(p^0, u^0)$ und $D(p^1, I) \equiv D^c(p^1, u^1)$, können wir die Konsumentenrente sowie die "äquivalente" und "kompensierende Variation" wie im Bild 3 (das den Fall eines normalen Gutes zeigt) darstellen. Die Veränderung der Konsumentenrente entspricht der Fläche $A + B$. Die "äquivalente Variation" entspricht der Fläche A und die "kompensierende" der Fläche $A + B + C$.²⁷

²⁶ Die Formel für die äquivalente Variation lautet

$$E(p^1, V(p^1, I)) - E(p^0, V(p^1, I)) = \int_{p^0}^{p^1} D_h^c(p, V(p^1, I)) dp_h,$$

wobei $I = E(p^1, V(p^1, I)) = E(p^0, V(p^0, I))$, das Enveloppentheorem und die Definition der Ausgabenfunktion benutzt worden sind, um die Ableitung der Ausgabenfunktion nach dem Preis zu ermitteln. Entsprechend kann man auch für die kompensierende Variation verfahren.

²⁷ In diesem Fall liefern die zwei Hicksschen Variationen Ober- und Untergrenzen für die Konsumentenrente. Wenn wir uns allerdings für die Zusatzlasten statt für die Renten und die Variationen interessieren, dann braucht dies nicht zuzutreffen; vgl. hierzu Hausman 1981.

Verallgemeinerungen der Theorie der Zusatzlast auf den Fall einer Mehrzahl von Preisveränderungen finden sich bei Mohring (1971) und Diamond und McFadden (1974). In diesen Abhandlungen wird jeweils auf die äquivalente und die kompensierende Variation zurückgegriffen.

Wie gut eignet sich die Konsumentenrente als Approximation an die Hicksschen ("äquivalenten und kompensierenden") Variationen? Da die Ableitungen dieser Variationen nach dem Preis nichts anderes sind als die normale und die kompensierte Nachfragefunktion, legt die Slutsky-Gleichung die Vermutung nahe, daß der Unterschied zwischen der Konsumentenrente und den Hicksschen Variationen gering ist, wenn die Einkommenseffekte klein sind. Willig (1976) hat eine Abschätzung des Fehlers vorgenommen, den man begeht, wenn man die Hicksschen Variationen mit Hilfe der Konsumentenrente approximiert, die ja von der Einkommenselastizität der Nachfrage und/oder dem Anteil der Ausgaben für das Gut am Budget des Haushalts abhängen.²⁸

Den Leser wollen wir bitten, sich davon zu überzeugen, daß der Einkommenseffekt wahrscheinlich klein ist, wenn auf das betreffende Gut nur ein kleiner Teil der Ausgaben entfällt. Die Ausdrücke

$$\varepsilon_h \equiv - \frac{\partial D_h}{\partial p_h} \bigg/ \frac{D_h}{p_h}$$

und

$$\varepsilon_h^c \equiv - \frac{\partial D_h^c}{\partial p_h} \bigg/ \frac{D_h^c}{p_h}$$

sollen die (direkte) Preiselastizität der "normalen" Nachfrage und die der kompensierten bezeichnen und

$$\varepsilon_h^I \equiv - \frac{\partial D_h}{\partial I} \bigg/ \frac{D_h}{I}$$

die Einkommenselastizität der "normalen" Nachfrage. Dann können wir die Slutsky-Gleichung auch folgendermaßen darstellen:

$$\varepsilon_h = \varepsilon_h^c + \left(\frac{p_h D_h}{I} \right) \varepsilon_h^I.$$

²⁸ Ist der prozentuale Fehler bei dieser Annäherung klein, dann ist damit - dies ist zu beachten - noch keineswegs gewährleistet, daß auch der prozentuale Fehler klein ist, der sich ergibt, wenn wir die Zusatzlast anhand der Konsumentenrente schätzen, statt sie mit Hilfe der Hicksschen Variation zu berechnen. (Um die Zusatzlast zu ermitteln, muß man das zusätzliche Einkommen für den Staat oder das Unternehmen abziehen.) Siehe Hausman (1981).

Hieraus wird deutlich, daß der Einkommenseffekt unerheblich ist, wenn die Ausgaben $p_h D_h$ der Konsumenten für das Gut h im Vergleich zu seinem Einkommen gering sind. Daraus ergeben sich zwei weitere nützliche Erkenntnisse:

- Die Nachfragekurve für Gut h hat negative Steigung, weil die kompensierte Nachfrage dies hat.
- Die Konsumentenrente und die Zusatzlast, die anhand der Konsumentenrente bestimmt wird, stellen eine gute Approximation dar, wenn es um die Messung des Wohlfahrtsverlustes geht.

Diese Einsichten gehen bis auf Marshall zurück (1920, S. 842), der die Auffassung vertrat, daß diese zwei Aussagen deswegen gelten, weil die "Ausgaben" des Konsumenten "für irgendein Gut, zum Beispiel Tee, nur einen kleinen Teil seines Gesamteinkommens darstellen." Vives (1987) bewies, daß Marshall im Grunde recht hatte. Unter einigen zusätzlichen Voraussetzungen²⁹ zeigte er, daß die folgenden Aussagen gelten, wenn der Konsument eine große Zahl von Gütern n konsumiert:

- Leitet man die Nachfragefunktion für ein Gut nach dem Einkommen ab, so findet man, daß der Wert dieses Differentialquotienten nicht groß ist (er hat die Größenordnung $1/\sqrt{n}$ oder sogar $1/n$, wenn es sich um additiv zerlegbare Präferenzen handelt), die Nachfragekurven weisen negative Steigung auf.³⁰
- Ändert sich nur ein einziger Preis, so ist der Fehler bei der Approximation der Hicksschen Variationen durch die Konsumentenrente gering - jedenfalls wenn wir den Fehler als Prozentsatz des Meßwertes ausdrücken (er liegt ebenfalls in der Größenordnung $1/\sqrt{n}$). Dasselbe gilt für die näherungsweise Bestimmung der Zusatzlast.

²⁹ Insbesondere unterstellt Vives, daß die Präferenzen für unterschiedliche Güter ausreichend symmetrisch sind (damit will er die Möglichkeit ausschließen, daß der größte Teil des Einkommenseffekts auf ein Gut entfällt), daß es keine zwei Güter gibt, die beinahe vollkommene Substitute darstellen (damit soll ausgeschlossen werden, daß sich die Nachfrage auf ein Gut konzentriert). Ferner nimmt er an, daß die Nutzenfunktion des Konsumenten eine bestimmte Krümmungseigenschaft aufweist. Dann verschwinden die Einkommenseffekte, wenn die Zahl der Güter gegen unendlich geht, wohingegen die Substitutionseffekte bedeutsam bleiben.

³⁰ Es gibt ein umfangreiches Schrifttum darüber, unter welchen Bedingungen die Nachfragekurve eine negative Steigung aufweist. Beispielsweise zeigt Hildenbrand (1983), daß dies für sämtliche Nachfragefunktionen der Fall ist, wenn die Konsumenten identische Nachfragefunktionen besitzen und die Einkommensverteilung sich durch eine monoton fallende Dichtefunktion beschreiben läßt. (Vergleiche hierzu auch Chiappori 1985.) Bei diesem Ansatz kann man auf die Annahme verzichten, daß die Zahl der Güter groß ist (und die anderen damit zusammenhängenden Annahmen - siehe Fußnote 29). Indes sind die zwei angeführten Annahmen (identische Präferenzen und abnehmende Dichte) reichlich restriktiv.

Daß restriktive Annahmen erforderlich sind, kann angesichts der Sätze von Sonnenschein (1973), Debreu (1974) und Mantel (1976) über Überschufnachfragefunktionen nicht allzusehr überraschen. Die genannten Autoren zeigten nämlich, daß abgesehen von der Homogenität vom Grade null hinsichtlich der Preise und dem Walraschen Gesetz über die Gesamtnachfragefunktionen keinerlei Aussagen möglich sind, wenn es wenigstens soviele Konsumenten wie Güter gibt.

- Ändern sich mehrere Preise, so lassen sich die Hicksschen Variationen ebenfalls gut durch die Konsumentenrente eines diese Güter konsumierenden Verbrauchers annähern. Diese Annäherung ist unabhängig von der Reihenfolge, in der die Integration stattfindet. (Der Leser möge sich daran erinnern, daß sich die Konsumentenrente nicht eindeutig definieren läßt, wenn sich die Preise von mehreren Gütern verändern; ihre Größe ist dann pfadabhängig.)

Von den Gütern und Branchen, die in diesem Buch betrachtet werden, wird stets unterstellt werden, daß auf sie nur ein kleiner Teil der Konsumausgaben entfällt. Preisveränderungen bewirken deshalb wahrscheinlich nur kleine Einkommenseffekte. Es ist infolgedessen wohl zulässig, davon auszugehen, daß die Nachfragefunktion negative Steigung aufweist und daß die Konsumentenrente bei Bemühungen, die Wohlfahrtswirkungen zu messen, als eine gute Approximation gelten darf.

Wenn wir vom Fall eines einzelnen auf den mehrerer Konsumenten übergehen, stoßen wir auf neuerliche Schwierigkeiten. Zwar läßt sich ohne weiteres eine aggregierte "äquivalente Variation" definieren, nämlich als die Summe der einzelnen "äquivalenten Variationen". Doch das Problem ist, daß der Wert dieser aggregierten Variation im allgemeinen auf eine Einkommensumverteilung zwischen den Konsumenten reagiert. Nur unter reichlich restriktiven Annahmen kann man von diesen Verteilungseffekten absehen.³¹ (Mehr darüber findet sich bei Auerbach [1986].)

In diesem Buch werden Probleme der Einkommensverteilung vernachlässigt. Mit anderen Worten, von einer Umverteilung von einem Konsumenten zum anderen wird angenommen, daß sie keinen Wohlfahrtseffekt auslöst. (Die sozialen Grenznutzen des Einkommens sind bei allen Konsumenten dieselben.) Damit soll aber nicht die Meinung zum Ausdruck gebracht werden, daß die bestehende Einkommensverteilung optimal ist, sie wäre es nicht einmal, wenn eine optimale Einkommensteuer erhoben würde (weil es Grenzen und Kosten der Einkommensbesteuerung gibt, wie sie in der Literatur zur optimalen Besteuerung hervorgehoben werden). Jeder Eingriff in den Markt hat erwünschte oder unerwünschte Verteilungseffekte. Wir werden uns auf die Effizienz der Märkte konzentrieren und damit gewissermaßen auf die alte Musgravesche Unterscheidung (1959) zurückgreifen, wonach man

³¹ Insbesondere muß die Nachfragefunktion des Konsumenten i die Gormansche polare Form aufweisen:

$$D^i(\mathbf{p}, I^i) = \Phi^i(\mathbf{p}) + \theta(\mathbf{p})I^i,$$

wobei Φ^i hinsichtlich der Preise homogen vom Grade 0 ist und θ homogen vom Grad -1.

sich den Staat so vorstellen soll, als wäre er in eine Distributionsabteilung und eine Allokationsabteilung unterteilt. Die Distributionsabteilung kümmert sich nur um Verteilungsfragen und die Allokationsabteilung nur um solche der Effizienz - und dieses Buch befaßt sich nur mit dem Aufgabenbereich der Allokationsabteilung. Das "Kompensationsprinzip" von Hicks (1940) und Kaldor (1939) behauptet, daß wir uns nur mit Effizienz beschäftigen müssen, wenn es um eine Erhöhung der Gesamtheit der Renten geht - die Gewinner können dann nämlich die Verlierer entschädigen und es geht allen besser. Diese Vorgehensweise hat die wohlbekannte Schwäche, daß in der Realität die Distributionsabteilung vielleicht nicht funktioniert und die Kompensation nicht stattfindet (Samuelson 1947). Diesen Vorbehalt müssen wir bei allen unseren Schlußfolgerungen über die gesellschaftliche Wohlfahrt im Auge behalten. Die Schlußfolgerung aus Kapitel 3, gemäß der vollkommene Preisdifferenzierung durch einen Monopolisten die gesellschaftliche Wohlfahrt erhöht und folglich nicht verboten werden sollte, läßt sich nicht mehr aufrechterhalten, wenn dem wohlwollenden Diktator die Einkommen der Konsumenten stärker am Herzen liegen als die der Anteilseigner der betroffenen Unternehmen.

Was ist ein Markt?

Die Begriff des Marktes ist auf keinen Fall einfach. Offensichtlich wollen wir uns nicht darauf beschränken, von homogenen Gütern zu sprechen. Sollen Güter nur dann zu demselben Markt gezählt werden, wenn es sich um vollkommene Substitute handelt, dann würden fast alle Märkte von "Monopolisten" versorgt - Unternehmen produzieren Güter, die wenigstens etwas unterschiedlich sind (entweder physisch oder hinsichtlich des Standorts, der Verfügbarkeit, des Informationsstandes der Konsumenten über sie oder anderer Aspekte). Aber die meisten Unternehmen besitzen keineswegs eine gewaltige Monopolmacht. Eine Preissteigerung veranlaßt die Konsumenten, auf eine möglicherweise kleine Anzahl von Alternativgütern auszuweichen. Deshalb sollte die Abgrenzung eines Marktes nicht zu eng ausfallen.³²

Sie sollte aber auch nicht zu weit sein. Jedes Gut ist potentiell ein Substitut für jedes andere, und sei es auch nur in einem infinitesimalen Sinn. Ein Markt sollte jedenfalls nicht so abgegrenzt werden, daß er mit der ganzen Volkswirtschaft identisch ist. Eine Partialanalyse sollte zulässig bleiben. Es sollte auch noch möglich sein, die wichtigsten Interaktionen zwischen den Unternehmen zusammenfassend zu charakterisieren. Ferner ist es auch

³² Dieser Unterabschnitt beruht auf Vorlesungen von Paul Joskow und Richard Schmalensee.

wichtig zu erkennen, daß die "richtige" Abgrenzung eines Marktes davon abhängt, womit der Analytiker sich befassen möchte. Denken wir zum Beispiel an Kohle. Sollen allgemeine Probleme der Energiepolitik im Vordergrund stehen (Wirkungen von Subventionen für bestimmte Energieträger), ist der Energiemarkt als ganzes der relevante Markt, und der umfaßt Kohle, Gas, Öl und Kernenergie. Soll eine Analyse der langfristigen Lieferbeziehungen und der vertikalen Integration zwischen den US-Kohleproduzenten und Elektrizitätsunternehmen angefertigt werden, so empfiehlt sich eine Abgrenzung nach regionalen Gesichtspunkten (z.B. Nordosten, mittlerer Westen und Westen; vergleiche hierzu das Kapitel Theorie der Unternehmung). Um die Auswirkungen eines Unternehmenszusammenschlusses zwischen zwei Kohlelieferanten auf den Wettbewerb einzuschätzen, empfiehlt sich wegen der hohen Beförderungskosten eine noch viel engere Abgrenzung eines Marktes.

Es gibt kein einfaches Rezept für die Abgrenzung eines Marktes. Dies zeigen auch die vielen Kontroversen, die zwischen Volkswirten und Rechtsanwälten über die Bedeutung von Marktmacht in bestimmten Branchen geführt werden. Immerhin hat man einige Kriterien entwickelt, die das Problem allerdings nur teilweise lösen. Robinson (1933) schlug vor, mit einem bestimmten Gut zu beginnen, dann seine Substitute durchzugehen, die Substitute für diese Substitute und so weiter, bis man eine ausgeprägte Lücke in dieser Kette der Substitute findet. Diese Lücken, behauptete sie, bestimmen die Grenzen des Marktes "rund um" das betrachtete Gut. Diese Definition hat mehrere Nachteile. Erstens behandelt sie die Güter vielleicht zu symmetrisch. Gut 1 und Gut 3 konkurrieren vielleicht intensiv mit Gut 2, aber nur wenig miteinander. Ein Auto von Hyundai und eines von Rolls-Royce gehören zweifellos zu derselben Kette von Substituten, aber gehören sie wirklich demselben Markt an? Zweitens berücksichtigt dieses Abgrenzungsverfahren nur den tatsächlich bestehenden, nicht aber den potentiellen Wettbewerb. Der dritte Nachteil betrifft die praktische Anwendung des Verfahrens: was soll uns als Lücke gelten?³³ Ein anderes Abgrenzungskriterium stellt auf die Korrelation zwischen den Preisen der Güter ab. Der Grundgedanke ist, daß Güter, die demselben Markt angehören, ähnlichen Kosten- und Nachfrageschocks ausgesetzt und ihre Preise deshalb tendenziell korreliert sind. Allerdings ist eine solche Korrelation höchstens eine

³³ Die Richtlinien des US-Justizministeriums sehen ein Entscheidungskriterium vor, das eine gewisse Verwandtschaft mit diesem Kriterium aufweist: Man beginnt mit einem bestimmten Produkt und einem bestimmten Anbieter und bezieht dann solange enge Substitute in die Betrachtung ein (diese müssen nicht notwendigerweise vom selben Anbieter erzeugt werden), bis die Preiselastizität der Nachfrage nach dieser Menge von Gütern so gering wird, daß die Anbieter im Falle der Kollusion einen Gewinnaufschlag durchsetzen könnten, der einen bestimmten Schwellenwert überschreitet. Eine solche Menge von Gütern wird als ein Markt betrachtet. Gemäß dieser Richtlinien muß man also nicht unbedingt nach Substitutionslücken suchen; in der Praxis tut man dies allerdings möglicherweise doch.

notwendige Bedingung für die Zugehörigkeit zum gleichen Markt. Boston Edison und Electricité de France verteilen elektrischen Strom, sind aber keinesfalls Konkurrenten, obwohl ihre Preise korrelieren - die Korrelation erklärt sich durch die Brennstoffpreise. Daraus zu schlußfolgern, daß sie dem gleichen Markt angehören, wäre falsch.

In diesem Buch wird über diese Schwierigkeit der Abgrenzung eines Marktes hinweggesehen. Wir werden annehmen, daß der Markt wohldefiniert ist. Er umfaßt entweder ein homogenes Gut oder eine Gruppe verschiedener Produkte, die für mindestens ein Gut der jeweiligen Gruppe enge Substitute (oder Komplemente) sind. Von der Interaktion mit "dem Rest" der Volkswirtschaft wird angenommen, daß sie sich in Grenzen hält.

Literaturverzeichnis

- Areeda, P. (1974), *Antitrust Analysis*, zweite Auflage, Boston: Little, Brown
- Areeda, P. und D. Turner, (1976), *Antitrust Law*, Boston: Little, Brown
- Arrow, K. und F. Hahn, (1970), *General Competitive Analysis*, San Francisco: Holden Day
- Auerbach, A. (1986), The Theory of Excess Burden and Optimal Taxation, in: *Handbook of Public Economics*, Band 1, Hrsg.: A. Auerbach und M. Feldstein, New York: Elsevier
- Blair, R. und D. Kaserman. (1985), *Antitrust Economics*, Homewood, Illinois: Irwin
- Boiteux, M. (1951), Le "Revenu Distribuable" et les Pertes Economiques, *Econometrica* 19: 291-309
- Bresnahan, T. 1987a, Competition and Collusion in the American Automobile Industry: The 1955 Price War, *Journal of Industrial Economics* 35: 457-482
- Bresnahan, T. (1987b), Empirical Studies of Industries with Market Power, in: *Handbook of Industrial Organization*, Hrsg.: R. Schmalensee und R. Willig, Amsterdam: North-Holland
- Carlton, D. (1987), The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?, in: *Handbook of Industrial Organization*, Hrsg.: R. Schmalensee und R. Willig, Amsterdam: North-Holland
- Chiappori, P.-A. (1985), Distribution of Income and the "Law of Demand", *Econometrica* 53: 109-127
- Debreu, G. (1951), The Coefficient of Resource Allocation, *Econometrica* 19: 273-292
- Debreu, G. (1959), *The Theory of Value*, New York: Wiley
- Debreu, G. (1974), Excess Demand Functions, *Journal of Mathematical Economics* 1: 15-21
- Diamond, P. und D. McFadden, (1974), Some Uses of the Expenditure Function in Public Finance, *Journal of Public Economics* 3: 3-21
- Dixit, A. (1976), *Optimization in Economic Theory*, Oxford University Press
- Dixit, A. und V. Norman, (1980), *Theory of International Trade*, Welwyn: Nisbet
- Dupuit, J. (1844), De la Mesure de l'Utilité des Travaux Publics, Translation in American Economic Association, *Readings in Welfare Economics*, Hrsg.: K. Arrow und T. Scitovsky
- Hall, R. (1986), The Relationship between Price and Marginal Cost in U.S. Industry Working Paper E-86-24, Hoover Institution, Stanford University
- Harberger, A. (1964), Taxation, Resource Allocation and Welfare, in: *The Role of Direct and Indirect Taxes in the Federal Reserve System*, Princeton University Press for NBER and Brookings Institution

-
- Hart, O. (1985), Imperfect Competition in General Equilibrium: An Overview of Recent Work, in: *Frontiers of Economics*, Hrsg.: K. Arrow und S. Honkapohja, Oxford: Blackwell
- Hausman, J. (1981), Exact Consumer's Surplus and Deadweight-Loss, *American Economic Review* 71: 662-676
- Helpman, E. und P. Krugman, (1985), *Market Structure and Foreign Trade*, Cambridge, Mass.: MIT Press
- Hicks, J. (1940), The Valuation of Social Income, *Economica* NS 7: 105-129
- Hicks, J. (1941), The Rehabilitation of Consumer's Surplus, *Review of Economic Studies* 9: 108-116
- Hicks, J. (1946), *Value and Capital*, zweite Auflage, Oxford University Press
- Hildenbrand, W. (1983), On the Law of Demand, *Econometrica* 51: 997-1019
- Hopenhayn, H. (1986), A Competitive Stochastic Model of Entry and Exit to an Industry
Unveröffentlichtes Manuskript, University of Minnesota
- Hotelling, H. (1938), The General Welfare in Relation to Problems of Tarification and of Railway and Utility Raises, *Econometrica* 6: 242-269
- Joskow, P. (1975), Firm Decision-Making Process and Oligopoly Theory, *American Economic Review, Papers and Proceedings* 65: 270-279
- Joskow, P. (1985), Vertical Integration and Long Term Contracts: The Case of Coal-Burning Electric Generating Plants, *Journal of Law, Economics and Organization* 1: 33-80
- Jovanovic, B. (1982), Selection and the Evolution of Industry, *Econometrica* 50: 649-670
- Kaldor, N. (1939), Welfare Propositions in Economics, *Economic Journal* 49: 549-552
- Kamien, M. und N. Schwartz, (1981), *Dynamic Optimization*, Amsterdam: North-Holland
- Lippman, C. und R. Rummelt, (1982), Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency under Competition, *Bell Journal of Economics* 13: 418-438
- Lucas, R. (1978), On the Size Distribution of Business Firms, *Bell Journal of Economics* 9: 508-523
- Mantel, R. (1976), Homothetic Preferences and Community Excess Demand Functions, *Journal of Economic Theory* 12: 197-201
- Marshall, A. (1920), *Principles of Economics*, London: Macmillan
- Mas-Colell, A. (1985), *The Theory of General Economic Equilibrium: A Differentiable Approach*, Cambridge University Press
- Mohring, H. (1971), Alternative Welfare Gain and Loss Measures, *Western Economic Journal* 9: 349-368

- Musgrave, R. (1959), *The Theory of Public Finance*, New York: McGraw-Hill
- Plott, C. (1982), Industrial Organization Theory and Experimental Economics, *Journal of Economic Literature* 20: 1485-1527
- Porter, M. (1980), *Competitive Strategy*, New York: Free Press
- Porter, R. (1983), A Study of Cartel Stability: The Joint Economic Committee: 1880-1886, *Bell Journal of Economics* 14: 301-314
- Posner, R. und F. Easterbrook, (1981), *Antitrust Cases, Economic Notes and Other Materials*, zweite Auflage, St. Paul: West
- Robinson, J. (1933), *The Economics of Imperfect Competition*, London: Macmillan
- Samuelson, P. (1947), *Foundation of Economic Analysis*, Cambridge, Mass.: Harvard University Press
- Scherer, F. (1980), *Industrial Market Structure and Economic Performance*, zweite Auflage, Chicago: Rand-McNally
- Schmalensee, R. (1986), Inter-Industry Studies of Structure and Performance, in: *Handbook of Industrial Organization*, Hrsg.: R. Schmalensee und R. Willig, Amsterdam: North-Holland
- Sonnenschein, H. (1973), Do Walras' Identity and Homogeneity Characterize the Class of Community Excess Demand Functions? *Journal of Economic Theory* 6: 345-354
- Varian, H. (1978), *Microeconomic Analysis*, New York: Norton
- Vives, X. (1987), Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand, *Review of Economic Studies* 54: 87-103
- Willig, R. (1976), Consumer's Surplus without Apology, *American Economic Review* 66: 589-597

Theorie der Unternehmung

Die Hauptakteure in diesem Buch sind Unternehmen. Aber was ist ein Unternehmen, und wie verhalten sich Unternehmen? Diese zwei Fragen sind Gegenstand dieses Einführungskapitels.

Der Inhalt der Begriffe "Unternehmen", "vertikale Verflechtung (Integration)" und "Entscheidungsbefugnis" versteht sich keineswegs von selbst. In der Literatur finden sich sehr unterschiedliche Begriffsbestimmungen. Wir werden hier einige der gängigsten Definitionen darstellen. Diese stellen jeweils unterschiedliche Aspekte in den Vordergrund und ergänzen sich in einem gewissen Sinne wechselseitig. Gemeinsam ist ihnen die Vorstellung, daß eine Unternehmung in der Lage ist, Güter bzw. Dienste effizienter zu produzieren (oder zu vermarkten) als es der Fall wäre, wenn ihre Bestandteile jeweils selbständig agieren würden. Hinter diesen Definitionen steckt also ein Optimierungsansatz.

Die Bestimmung der Zielfunktion eines Unternehmens ist ebenfalls keine triviale Angelegenheit. Das Verhalten der Entscheidungsträger mag in vielerlei Hinsicht und sehr unterschiedlicher Richtung von der klassischen Gewinnmaximierungshypothese abweichen. Andererseits sind vielfältige Vorkehrungen denkbar, die solchen Abweichungen Grenzen setzen, die die Entscheidungsspielräume des Managements verringern. Neuere Entwicklungen in der Theorie, die sich mit den Entscheidungen von Unternehmen befaßt, beruhen vorwiegend auf der Anreiztheorie. Diese kann hier nur unzureichend gewürdigt werden. Wir werden nur einige ihrer grundlegenden Aussagen ableiten.

Dieses Kapitel ist vielleicht nicht mehr als ein Eintopfgericht. Als Zutaten fungieren die Theorien, die sich derzeit bei den Wissenschaftlern besonderer Beliebtheit erfreuen. Wir hoffen aber, daß es zumindest als eine Einführung taugt, die dem Leser den Zugang zu den wichtigsten Beiträgen eröffnet. Zahlreiche Theorien werden kurz besprochen, und der Leser erhält die erforderlichen Literaturhinweise. Einen Überblick über das ganze Gebiet verschaffen auch die Aufsätze von Hart und Holmström (1987) und Holmström und Tirole (1987). Sie stellen eine gute Ergänzung zu diesem Kapitel dar¹.

¹ Hart und Holmström (1987) führen einige der in diesem Kapitel erörterten Gegenstände mehr im einzelnen aus. Holmström und Tirole (1987) konzentrieren sich vergleichsweise stärker auf die inhaltlichen, nicht so sehr die methodologischen Probleme und gehen auf einen weit größeren Themenkreis ein als das vorliegende Kapitel.

Bei der Messung der Größe eines Unternehmens unterscheidet man üblicherweise zwischen einer horizontalen und einer vertikalen Dimension. Bei einem Einproduktunternehmen ist mit der horizontalen Dimension das Prozeßniveau gemeint ("scale of production", man spricht auch von Produktionsniveau), bei einem Mehrproduktunternehmen eine Verallgemeinerung dieses Begriffs, die man als "scope of production" bezeichnet. Die vertikale Dimension stellt dar, in welchem Umfang das Unternehmen Vorprodukte, Komponenten oder Dienstleistungen selbst erstellt, statt sie von anderen Unternehmen zu beziehen. Es geht also um die Alternative Selbsterstellung oder Fremdbezug. Fusioniert ein Tapetenhersteller mit einem anderen Tapetenhersteller oder beispielsweise einem Fliesenproduzenten, so handelt es sich um horizontale Verflechtung. Schließt sich hingegen ein Weinhersteller mit einer Flaschen- oder Korkenfabrik zusammen, so sprechen wir von vertikaler Verflechtung oder Integration.

Das nachfolgende Kapitel unterteilt sich in zwei Abschnitte. Der erste erörtert verschiedene Definitionen des Begriffs Unternehmen. Dabei geht es uns aber nicht um Probleme wie die Rechtsformwahl (also beispielsweise, ob das jeweilige Unternehmen als Aktien- oder als Kommanditgesellschaft organisiert sein sollte) und die Faktoren, die dabei zu berücksichtigen sind (Haftungsbeschränkung, Möglichkeiten der Beschaffung von Eigenkapital usw.)². Vielmehr interessieren wir uns für die Frage, wovon die Größe des Unternehmens abhängt. Immerhin sollte ein rechtlicher Aspekt zu Beginn erwähnt werden: Manchmal werden bestimmte Transaktionen deswegen in einem Unternehmen internalisiert (damit meinen wir, daß diese Transaktionen und Aktivitäten innerhalb eines Unternehmens abgewickelt werden, statt daß sie die Gestalt von Transaktionen zwischen zwei oder mehreren verschiedenen Unternehmen annehmen), weil sich dadurch verschleiern läßt, daß sie gesetzeswidrig sind. Sie könnten etwa gegen das Gesetz gegen Wettbewerbsbeschränkungen verstoßen. Sie würden dann als Wettbewerbsbeschränkungen die Aufmerksamkeit des Kartellamtes auf sich ziehen, handelte es sich nicht um innerbetriebliche Angelegenheiten. Wir werden Beispiele anführen, in denen ein vertikaler oder horizontaler Zusammenschluß nur deswegen stattfindet, weil das Unternehmen dadurch in die Lage versetzt wird, den Wettbewerb zu beschränken, ohne gegen das Kartellrecht zu verstoßen.

² Alchian und Demsetz (1972) und Jensen und Meckling (1976) befassen sich mit der Rolle des Eigenkapitals.

Abschnitt 1 beschäftigt sich hauptsächlich mit Fällen, in denen eine Verflechtung oder Entflechtung um einer Erhöhung der Effizienz willen erfolgt (und nicht, um den Wettbewerb zu beschränken). Es geht also darum, wirtschaftliche Aktivitäten so zu organisieren, daß die Kosten minimiert werden. Solche Effizienzerwägungen sind auch und gerade dann relevant, wenn das Unternehmen über keine Marktmacht verfügt. Herrscht auf dem Gütermarkt vollkommene Konkurrenz, so ist die Organisationsform, die aus einzelwirtschaftlicher Sicht optimal ist, im Prinzip auch gesamtwirtschaftlich optimal, d.h. sie maximiert die soziale Wohlfahrt³.

Zuerst wollen wir das Unternehmen als ein Gebilde betrachten, das zu einem bestimmten Zeitpunkt aus verschiedenen Bestandteilen zusammengefügt wird, um zunehmende Skalenerträge und Verbundvorteile ("economies of scope") auszunutzen⁴. Diese technologieorientierte Sicht des Unternehmens wird uns dazu veranlassen, verschiedene Arten von Kosten- und Nachfragekomplementaritäten und den traditionellen Organisationstyp der Funktionalorganisation zu erörtern.

Eine zweite Sicht des Unternehmens betrachtet dieses als einen langfristigen Vertrag zwischen den Elementen, aus denen es sich zusammensetzt. Dies führt uns zur Williamson'schen Theorie über die Risiken des idiosynkratischen Austauschs in einer langfristigen Beziehung. Man stelle sich eine Kunden-Lieferanten-Beziehung vor, bei der die Parteien irreversible geschäftsspezifische Investitionen tätigen müssen (eine Investition ist geschäftsspezifisch, wenn sie speziell zur Abwicklung eines bestimmten Geschäfts getätigt wird), bevor sie tatsächlich beginnen können, Leistungen auszutauschen. Ex ante (bevor das Geschäft vereinbart wird) mag es auf diesem Markt eine Vielzahl von Anbietern und Nachfragern geben, aber ex post (nachdem die Investitionen getätigt worden sind) mag es sein, daß sich die beiden vertragschließenden Parteien in einem bilateralen Monopol wiederfinden. Möglicherweise kann weder der Lieferant rechtzeitig andere Abnehmer gewinnen, noch ist der Abnehmer in der Lage, rechtzeitig einen anderen Lieferanten ausfindig zu machen. Ex post gibt es keinen Wettbewerb mehr, und das kann einen der beiden

³ Sofern bei der Produktion keine externen Effekte zwischen den Unternehmen auftreten. Solche Effekte treten beispielsweise dann auf, wenn auf dem Inputmarkt kein vollkommener Wettbewerb herrscht, oder wenn, wie dies in dem Hartschen Modell der Anreize für Manager der Fall ist, die Fähigkeit der Eigentümer, Manager zu kontrollieren, (wegen Informationsbeschaffungsproblemen) davon abhängt, wie andere Unternehmen im betreffenden Wirtschaftszweig organisiert sind.

⁴ Etwas verkürzt kann man sagen: Zunehmende Skalenerträge liegen vor, wenn die Produktionskosten eines bestimmten Gutes mit der Zahl der erzeugten Einheiten dieses Gutes abnehmen. Verbundvorteile liegen vor, wenn mehrere verschiedene Güter erzeugt werden und zwischen ihnen externe Effekte auftreten, die die Produktionskosten verringern (z. B., wenn die Erzeugung des Gutes *A* zur Verringerung der Produktionskosten des Gutes *B* führt).

Vertragspartner in die Lage versetzen, den anderen "über den Tisch ziehen" zu können. Man bezeichnete dies auch als opportunistisches Verhalten (eine Vertragspartei eignet sich die Erträge an, die durch eine Investition der anderen geschaffen wurden). Ein langfristiger Vertrag muß gewährleisten, daß die beiden Vertragsparteien auch ex post einen angemessenen Ertrag erhalten, sonst fehlt es an einem ausreichenden Anreiz, solche spezifischen Investitionen vorzunehmen. Er muß also gewährleisten, daß auch ex post die effiziente Menge geliefert wird, und dies geschieht nur, wenn keine der beiden Parteien von der anderen einen Monopolpreis verlangen kann.

Diese beiden Ansätze und die Definitionen eines Unternehmens, zu denen sie führen, unterscheiden sich stark voneinander. Beide haben außerordentlich wenig mit dem handelsrechtlichen Begriff eines Kaufmanns oder einer Körperschaft zu tun. Sowohl von Unternehmen, die vereinbaren, bestimmte Fixkosten gemeinsam zu tragen (z.B. Marktforschung auf einem Exportmarkt zu betreiben, bestimmte Inputs zu beschaffen oder ein bestimmtes Forschungs- und Entwicklungsprojekt durchzuführen), als auch von einem Hersteller und einem Einzelhändler, die miteinander bestimmte vertikale Bindungen vereinbaren oder einem Kraftwerk und einem Kohlebergwerk, die benachbart sind und auf 30 Jahre einen detaillierten Liefervertrag abschließen, können wir sagen, daß sie zumindest partiell verflochten sind, obwohl sie rechtlich völlig selbständig bleiben. (Man stelle sich im Gegensatz dazu einen Angestellten vor, der nur für die Ausführung einer ganz bestimmten Aufgabe eingestellt wird, oder einen, der ohne Kündigungsfrist kündigen oder gekündigt werden kann, ohne daß eine Abfindung oder Vertragsstrafe gezahlt werden muß).

Ein dritter Ansatz faßt das Unternehmen als einen unvollständigen Vertrag auf. Dieser Ansatz hat eine gewisse Verwandtschaft mit der Betrachtungsweise der Juristen. Er stellt in den Vordergrund, daß Unternehmen einerseits und zivilrechtliche Verträge andererseits unterschiedliche Verfahren zur Regelung von Beziehungen darstellen ("governance modes"). Das Wesentliche am Unternehmen wird dabei in der Art und Weise gesehen, in der es auf den Eintritt eines unvorhergesehenen Ereignisses reagieren kann. Eben wegen dieser Unvorhersehbarkeit kann der Vertrag diesem Ereignis nicht von vornherein Rechnung tragen. Diesem Ansatz liegt die Vorstellung zugrunde, daß Verträge praktisch stets unvollständig sind, da fast immer unvorhersehbare Ereignisse eintreten können oder aber weil die Zahl der möglichen Ereignisse und Zufälle derart groß ist, daß es unzweckmäßig wäre, sie alle ausdrücklich in einem Vertrag zu berücksichtigen. Der Aufwand wäre einfach zu groß. Verträge regeln gewissermaßen nur die großen Linien einer geschäftlichen Beziehung. Das Eigentum an Maschinen und das Beschäftigungsverhältnis gibt dem Eigentümer des Unternehmens innerhalb gewisser Grenzen die Macht, d.h. die

Entscheidungsbefugnis, zu bestimmen, wofür die Maschinen eingesetzt werden und was das Personal tut. Inwieweit eine vertikale Verflechtung vorliegt, richtet sich gemäß diesem Ansatz danach, inwieweit diese Entscheidungsbefugnis bei einer der beiden vertragschließenden Parteien konzentriert ist.

In Abschnitt 1 werden auch die Implikationen dieser drei Theorien für die Faktoren diskutiert, die die Größe eines Unternehmens bestimmen. Zu diesen Faktoren gehören die Erschöpfung der zunehmenden Skalenerträge und der Verbundvorteile, unvermehrbar Produktionsfaktoren, die Nachteile langfristiger Beziehungen und eine übermäßige Konzentration der Entscheidungsbefugnisse.

Im Abschnitt 2 werden die Unternehmen demgegenüber als wohldefinierte Einheiten betrachtet. Analysiert wird nicht die Größe des Unternehmens, sondern seine Zielfunktion. Maximieren Unternehmen tatsächlich ihre Gewinne, wie wir es im Rest dieses Buches annehmen werden? Selbst wenn sie es nicht tun, könnte die Gewinnmaximierungshypothese trotzdem von Nutzen sein, um Einblicke in das Verhalten von Monopolisten oder in eine strategische Entscheidungssituation zu gewinnen, um Voraussagen über das Verhalten derer zu machen, die sich in einer solchen Situation befinden? Ferner wird untersucht, welche Faktoren die Entscheidungsspielräume des Managements begrenzen (Aufsicht sowie die Anreize, die von den Güter-, Arbeits- und Kapitalmärkten erzeugt werden). Dies führt zu einer nochmaligen Überprüfung der Gewinnmaximierungshypothese.

1 Was ist eine Unternehmung?

In den Unterpunkten 1.2 bis 1.4 werden wir drei Ansätze erörtern, die sämtlich das Unternehmen als ein Gebilde betrachten, das der Minimierung der Kosten dient. Aber bevor wir dies tun, werden wir uns noch mit einigen anderen Motiven für vertikale Verflechtung auseinandersetzen.

1.1 Die Unternehmung als Instrument, um den Wettbewerb zu beschränken

Sei es zu Recht oder zu Unrecht, das Kartellrecht untersagt eine ganze Reihe von Handelsgepflogenheiten und brandmarkt sie als wettbewerbswidrig. Durch eine Internalisierung der betreffenden Geschäfte können Unternehmen dieses Verbot umgehen und sich insgeheim die Möglichkeit verschaffen, Monopolmacht auszuüben (unternehmensinterne

Transaktionen bleiben Außenstehenden normalerweise unbekannt und entgehen infolgedessen auch der Aufmerksamkeit des Kartellamtes⁵).

Preisdifferenzierung

Wie im Kapitel 3 gezeigt werden wird, kann es für ein Unternehmen von Vorteil sein, ein bestimmtes Produkt auf verschiedenen Märkten zu verschiedenen Preisen zu verkaufen. Dadurch entstehen aber für die Einzelhändler, die auf diesen Märkten tätig sind, Möglichkeiten, Arbitragegeschäfte vorzunehmen. Um dem vorzubeugen, kann der Hersteller sich eine Verkaufsorganisation anlagern, die den Markt bedient, auf dem das Produkt zu einem niedrigen Preis verkauft wird. Ein ähnliches Phänomen hat seine Wurzeln in einer gesetzlichen Regelung statt in der Vorbeugung von Arbitrage: Die Gesetzgebung mancher Länder (wie z.B. die der USA oder Frankreichs) verpflichtet ein Unternehmen zur Gleichbehandlung von Kunden, die sich in einer "gleichartigen Lage" befinden. Es kann sein, daß das Unternehmen zwei solche Kunden keineswegs gleich behandeln will, möglicherweise will es einem der beiden die Lieferung verweigern. Durch einen Unternehmenszusammenschluß mit einem der beiden Kunden kann es eine solche Diskriminierung legalisieren. Es verschafft sich dadurch die Möglichkeit, dem anderen die Lieferung zu verweigern, indem es den Verkauf des Gutes einfach ganz einstellt.

Preisüberwachung bei Zwischenprodukten

Stellen wir uns vor, ein bestimmter Markt für ein Zwischenprodukt - auf diesem Markt soll es eine Vielzahl von Anbietern und Nachfragern geben - wird von einer Aufsichtsbehörde reguliert und diese setzt den Preis so fest, daß er unter dem Gleichgewichtspreis liegt. Die Nachfrage übertrifft infolgedessen das Angebot, und das Gut muß mit Hilfe irgendeines Rationierungsverfahrens verteilt werden. Dies schafft für die Anbieter einen Anreiz zu einem Unternehmenszusammenschluß mit einigen ihrer Kunden. Warum? Der (von der Aufsichtsbehörde festgesetzte) Preis, zu dem das Zwischenprodukt gehandelt wird, ist niedriger als der Schattenpreis, den die Nachfrager zu zahlen bereit sind. Hier gibt es also Renten zu verdienen und sie warten darauf, daß es jemandem gelingt, sie sich anzueignen. Als

⁵ Im Kartellrecht wird häufig auch ausdrücklich zwischen unternehmensinternen Transaktionen und Transaktionen mit anderen Unternehmen unterschieden.

Empfänger solcher Renten kommen die Anbieter und diejenigen Nachfrager in Betracht, deren Nachfrage infolge der Anwendung des Rationierungsverfahrens nicht voll befriedigt wird. Wäre der Markt unreguliert, so könnten sich die Anbieter ihren Teil dieser Renten aneignen, indem sie den Preis auf das Gleichgewichtsniveau anheben. Aber die Regulierung hindert sie daran. Vertikale Verflechtung ermöglicht es den Anbietern, die Preisüberwachung ihrer Wirksamkeit zu berauben. Aus den Liefergeschäften werden unternehmensinterne und damit für Außenstehende nicht mehr beobachtbare Vorgänge, die sich als solche der Kontrolle durch die Aufsichtsbehörde entziehen.

Auch das Steuerrecht kann Anreize für vertikale Unternehmenszusammenschlüsse erzeugen. Ein wichtiges Beispiel ist die Bruttoallphasenumsatzsteuer (ähnliche Effekte können auch bei einphasigen Steuern auf Zwischenprodukte auftreten). Findet die Lieferung des Zwischenprodukts innerhalb eines Unternehmens statt, ist dies kein Umsatz und folglich auch keine Steuer zu entrichten. (Dieses Phänomen tritt nicht auf, wenn die Steuerbemessungsgrundlage nur der Mehrwert ist, der dem Produkt in dem Unternehmen hinzugefügt wurde, nicht aber der Gesamtwert des Produktes, das Gegenstand der steuerpflichtigen Transaktion ist. Die Mehrwertsteuer schafft keinen Anreiz zu vertikalen Unternehmenszusammenschlüssen). Eine Ertragsratenregulierung ist ein weiterer möglicher Grund für vertikale Unternehmenszusammenschlüsse. Eine Unternehmung, die einer solchen Regulierung unterliegt, kann ihren Gewinn steigern, indem sie sich mit dem Lieferanten ihrer Ausrüstungen zusammenschließt - vorausgesetzt die unternehmensinternen Verrechnungspreise für diese Ausrüstungen unterliegen keiner Regulierung⁶.

Damit haben wir uns mit Motiven für Vertikalverflechtung befaßt, die ihre Wurzeln in der Rechtsordnung haben. Im Kapitel 4 befassen wir uns demgegenüber mit wirtschaftlichen Motiven für Vertikalverflechtung, die in der Ausübung von Monopolmacht begründet sind.

Bemerkung: Auch horizontale Unternehmenszusammenabschlüsse können die Absicht verfolgen, Monopolmacht zu erlangen. Das einfachste Beispiel ist natürlich ein Zusammenschluß zweier Unternehmen, die das gleiche Produkt erstellen. Solche Zusammenschlüsse spielten in der großen Konzentrationsbewegung, zu der es in den USA von 1887 bis

⁶ Ein Unternehmen, das von einer Regulierungsbehörde überwacht wird, kann die unternehmensinternen Verrechnungspreise aufblähen, um eine Ertragsratenregulierung zu umgehen. Vgl. Dayan (1972).

1904 kam, die maßgebliche Rolle⁷. Vielfach geht es hier darum, zu verhindern, daß die Monopolgewinne durch den Wettbewerb auf dem Gütermarkt verloren gehen (siehe Kapitel 5⁸).

1.2 Synergieeffekte als *raison d'être* des Unternehmens

In der Industrieökonomik hat man seit langem, insbesondere seit Viners (1932) klassischer Arbeit über Kostenkurven, einen Zusammenhang zwischen der Bedeutung zunehmender Skalenerträge innerhalb einer Branche einerseits und der Größe und Anzahl der Unternehmen andererseits hergestellt, die sich in ihr betätigen.

Eine der Hauptdeterminanten der Unternehmensgröße ist, in welchem Maße es Skalenerträge und Verbundvorteile ausnutzen kann. Ingenieure haben reichlich dokumentiert⁹, daß höhere Produktionsniveaus die Nutzung effizienterer Technologien zulassen. Bei höheren Produktionsniveaus werden Investitionen in kostenreduzierende Technologien wirtschaftlich, höhere Produktionsniveaus erlauben es den Arbeitnehmern, sich stärker zu spezialisieren. Die Stückkosten verringern sich. Solche zunehmenden Skalenerträge, deren Erzielung davon abhängt, wieviel von einem bestimmten Produkt erzeugt wird, werden auch produkt-spezifische Kostenersparnisse genannt. Obwohl dies nicht üblich ist, können wir in diese Kategorie auch die sogenannten "economies of massed reserves" (Vorteil größerer Reservekapazitäten) einordnen (Robinson 1958). Eine Fabrik mit einer großen Anzahl von Maschinen ist im Vergleich zu einer kleineren zu einem überproportional höheren Ausstoß fähig. Der Grund dafür ist, daß der Ausfall einer Maschine sich deswegen weniger ungünstig auf den Output des Unternehmens auswirkt, weil noch andere Maschinen zur Verfügung stehen, auf die die Produktion umgelenkt werden kann. (Bei stochastischer Unabhängigkeit der Maschinenschäden wird - wegen des Gesetzes der großen Zahl - der Output bei einem wachsenden Maschinenpark an der Grenze deterministisch). Ähnlich verhält es sich mit einem Unternehmen, das mehrere Märkte mit schwankender Nachfrage beliefert, sofern die

⁷ Vgl. hierzu z.B. Scherer (1980, S. 119-122). Derartige Unternehmenszusammenschlüsse wurden später durch die Entwicklung des Kartellrechts sehr erschwert.

⁸ Oft wird auch die Meinung vertreten, daß es für ein Unternehmen von Vorteil sein kann, auf mehreren unterschiedlichen Märkten tätig zu sein, zwischen denen möglicherweise kein Zusammenhang besteht. Dies mag der Grund für Zusammenschlüsse auf horizontaler Ebene oder Zusammenschlüsse anorganischer Art sein. Vgl. hierzu Kapitel 6.

⁹ Einige konkrete Beispiele finden sich bei Scherer (1980, S. 81-84).

Nachfrageveränderungen auf diesen Märkten nicht vollkommen korreliert sind: Wird die Gesamtheit dieser Märkte von einem Unternehmen versorgt, so operiert dieses unter weniger Unsicherheit als es bei einer Mehrzahl von Unternehmen der Fall wäre, von denen jedes einen dieser Märkte versorgt. Dies läßt sich in Kostenersparnisse umsetzen, weil der horizontale Zusammenschluß es ermöglicht, auf Investitionen zu verzichten, die sonst für die Bedienung von Nachfragespitzen erforderlich wären. "Economies of massed reserves" mögen auch bei Mehrproduktunternehmen eine Rolle spielen, wenn ihre Produkte mit Hilfe derselben Technologie erzeugt werden¹⁰.

Abgesehen hiervon kann es auch sein, daß die Zusammenfassung von Produktionsabläufen, ob sie sich nun auf ein einziges oder auf eine Mehrzahl von Produkten beziehen, bei der Produktion im engeren Sinne zwar keine Einsparungseffekte auslöst, aber dafür bei verschiedenen anderen betrieblichen Funktionen wie dem Marketing, dem personalwirtschaftlichen Bereich, der Finanzierung, Materialbeschaffung, Forschung und Entwicklung oder etwa auch bei der Rechnungsprüfung. Solch eine Zusammenfassung verhindert die Duplizierung von Fixkosten, die bei der Erbringung derartiger Dienstleistungen entstehen. Zumindest verringert sie im Durchschnitt diese Kosten. Auch Nachfragekomplementaritäten können einen Grund für einen Zusammenschluß darstellen¹¹.

Im späten 19. Jahrhundert entstanden große Unternehmen, die sich funktional organisierten, wie dies in Figur 1 gezeigt ist¹². Wir können diese Organisationsform als einen Versuch betrachten, zunehmende Skalenerträge zu erzielen. Die Konzentration aller Aufgaben, die im Rahmen einer Funktion anfallen (der Erstellung eines bestimmten Gutes oder Dienstes), innerhalb großer Organisationseinheiten soll zu einer Verringerung der Stückkosten führen.

Wie bedeutsam zunehmende Skalenerträge auch sein mögen, so sind sie doch nicht unerschöpflich. Eine Zusammenfassung von Maschinen oder von zwei Abteilungen, die

¹⁰ Arrow et al. (1972) haben einen aufschlußreichen Aufsatz über den Gegenstand dieses Unterabschnitts verfaßt. Die Autoren zeigen den Zusammenhang zwischen der Reparatur von Aggregaten und zunehmenden Skalenerträgen auf. Sie leiten eine interessante Produktionsfunktion mit zunehmenden Skalenerträgen ab, die gegen eine mit konstanten Skalenerträgen konvergiert, wenn die Zahl der Aggregate gegen unendlich geht.

¹¹ Man mag an Unternehmen denken, die linke Schuhe erzeugen, und solche, die rechte Schuhe erzeugen. Diese Unternehmen sollten sich wohl beim Design abstimmen. Die Notwendigkeit einer solchen Abstimmung begründet einen Synergieeffekt.

¹² Eine Erörterung der funktionalen Organisation und der Divisionalisierung findet sich bei Chandler (1966) und Williamson (1975). (Aus rein technischer Sicht erscheint die funktionale Organisation als naheliegend. Sie führt aber zu ernstern Koordinations- und Anreizproblemen und wurde deswegen in der Realität häufig durch eine divisionalisierte Organisation ersetzt. Vergleiche hierzu S. 113).

innerhalb zweier verschiedener produzierender Einheiten dieselbe Funktion wahrnehmen, ist nur von Vorteil, wenn sie nicht beide bereits bis zur Kapazitätsgrenze ausgelastet sind. Je größer das Unternehmen bereits ist, umso geringer werden die Einspareffekte, die sich im Fall einer weiteren Vergrößerung im Bereich der Bedienung von Nachfragespitzen erzielen lassen, weil die Möglichkeiten zur Konsolidierung (zum Ausgleich) von Risiko schon weitgehend ausgeschöpft sind, das Gesetz der großen Zahlen bereits ausgenutzt ist. Ein weiteres Argument verweist darauf, daß es knappe Produktionsfaktoren gibt wie z.B. das Talent des Managers, seine Führungsqualitäten, die sich bei einer Expansion des Unternehmens nicht ohne weiteres vermehren lassen¹³.

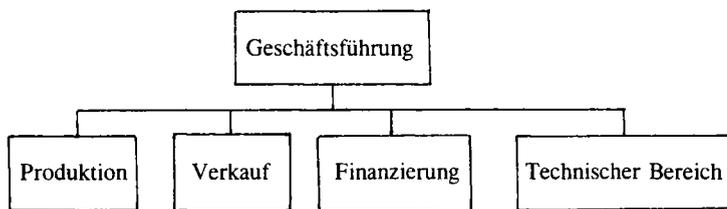


Abbildung 1

Beispiel der Funktionalorganisation

(Quelle: Williamson 1975, S. 134)

Wir wenden uns nun der formalen Analyse von zunehmenden Skalenerträgen und Verbundvorteilen zu. (Eine vollständigere Behandlung findet sich bei Baumol [1982] - siehe Kapitel 2 für den Einproduktfall und Kapitel 3 und 4 für den Mehrproduktfall.)

Wir beginnen mit der Einproduktunternehmung. $C(q)$ bezeichne die Gesamtkosten des Unternehmens, wenn es den Output q erzeugt; genauer, $C(q)$ sind die Kosten der Minimal-kostenkombination von Inputs, welche die Produktion von q Outputeinheiten ermöglichen.

¹³ Lucas (1967) und Prescott und Visscher (1980) entwickeln Modelle, in denen die Expansion eines Unternehmens durch Anpassungskosten gebremst wird. Prescott und Visscher bringen diese Anpassungskosten damit in Zusammenhang, daß der Arbeitgeber die Fähigkeiten und Eigenschaften neuereinstellter Arbeitnehmer erst allmählich kennenlernt. Erst nach längerer Zeit kann das Unternehmen beurteilen, für welche Aufgaben sie sich am besten eignen. Das Unternehmen kann dann zwischen einem raschen Wachstum, das mit Fehlentscheidungen im personalwirtschaftlichen Bereich verbunden ist, und einem langsamen Wachstum wählen, bei dem sich solche Fehlentscheidungen in weit höherem Maße vermeiden lassen. Bei Führungskräften in der Wirtschaft ist in der Tat - ganz im Einklang mit dieser Theorie - die Auffassung weit verbreitet, daß in Expansionsphasen das Humankapital zum wichtigsten Engpaßfaktor wird.

Lucas (1978) und Kihlstrom und Laffont (1979) analysieren Modellökonomien, in denen Manager knapp sind. Hierbei werden Manager mit Unternehmen gleichgesetzt.

Zur Vereinfachung wird vorausgesetzt, daß die Kostenfunktion zweimal differenzierbar ist (im Punkt $q = 0$ benötigen wir diese Annahme nicht). Es gilt also

$$C(q) = \begin{cases} F + \int_0^q C'(x)dx & \text{für } q > 0 \\ 0 & \text{ansonsten} \end{cases}$$

wobei $F \geq 0$ die Fixkosten bezeichnet.

Die Grenzkosten nehmen streng monoton ab, wenn $C''(q) < 0$ für alle q . Die Durchschnittskosten nehmen streng monoton ab, wenn für alle q_1 und q_2 mit $0 < q_1 < q_2$ gilt:

$$\frac{C(q_2)}{q_2} < \frac{C(q_1)}{q_1}.$$

Wir bezeichnen die Kostenfunktion als streng subadditiv, wenn für alle n -Tupel von Outputs q_1, \dots, q_n gilt:

$$\sum_{i=1}^n C(q_i) > C\left(\sum_{i=1}^n q_i\right).$$

Subadditivität bedeutet folglich, daß die Kosten geringer sind, wenn mehrere verschiedene Produkte zusammen erzeugt werden, als wenn jedes einzeln für sich erzeugt wird.

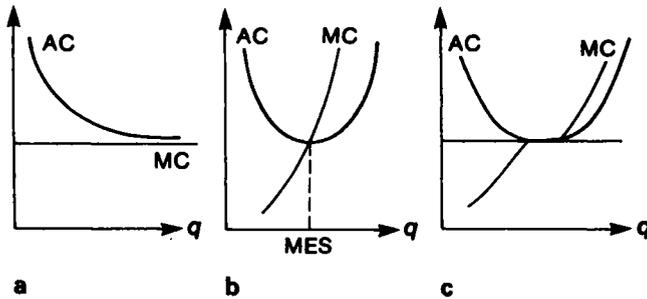


Abbildung 2

Abbildung 2 zeigt Funktionsgestalten der Durchschnittskosten (AC) und der Grenzkostenkurve (MC), die dem Leser aus Einführungs- und Fortgeschrittenenlehrbüchern für Mikroökonomik vertraut sind. Abbildung 2a zeigt die Kostenfunktion $C(q) = F + cq$ mit $q > 0$. Die Unternehmung hat die fixen Kosten F zu tragen, außerdem entstehen konstante Grenzkosten. Die Durchschnittskosten nehmen streng monoton ab (für $q = 0$ sind sie nicht definiert), aber mit abnehmender Rate. Abbildung 2b zeigt eine U-förmige Durchschnitts-

kostenkurve. Die Durchschnittskostenkurve fällt, bis das Betriebsoptimum (MES) erreicht ist. In diesem schneiden sich die Durchschnitts- und die Grenzkostenkurve, danach steigt die Durchschnittskostenkurve wieder an. Ein Beispiel für solch eine Kostenkurve ist $C(q) = F + aq^2$ (mit $a > 0$). In der Abbildung 2c ist eine wannenförmige Durchschnittskostenkurve dargestellt. Sie ist gewissermaßen ein Zwischending zwischen den Kurven in 2a und 2b. Die Durchschnittskosten bleiben in einem bestimmten Outputbereich konstant, bevor sie infolge abnehmender Skalenerträge wieder zunehmen¹⁴.

Monoton fallende Grenzkosten implizieren monoton fallende Durchschnittskosten¹⁵, und monoton fallende Durchschnittskosten implizieren Subadditivität¹⁶. Die Umkehrungen dieser zwei Sätze sind jedoch falsch. (Der in der Abbildung 2a dargestellte Fall ist ein Gegenbeispiel zur Umkehrung des ersten Satzes, Hauptsatz 2A1 in Baumol et alii [1982] liefert ein Gegenbeispiel gegen die Umkehrung des zweiten Satzes.)

Bemerkung: Der Begriff natürliches Monopol kann, je nach dem Fall, der zu untersuchen ist, mehrere Bedeutungen aufweisen. Stellen wir uns vor, daß eine jedermann zugängliche Technologie $C(q)$ existiert, mit der man den Output q erzeugen kann. Baumol et alii (1982) bezeichnen eine Branche als natürliches Monopol, wenn die Kostenfunktion im relevanten Outputbereich subadditiv ist. Diese Definition ist sinnvoll in einer Planwirtschaft mit einem allwissenden Planer (genauer: einem Planer, der die Kostenfunktion kennt). Für einen solchen Planer gibt es keinen Grund, dasselbe in mehreren Unternehmen produzieren zu lassen, wenn derselbe Gesamtpoutput billiger in einem einzigen Unternehmen erzeugt werden könnte.

Gilt das Interesse des Wissenschaftlers eher der positiven Analyse, so dürfte er sich für die Verhältnisse in einem unregulierten Wirtschaftszweig interessieren. $\Pi(n)$ soll den Gewinn eines einzelnen Unternehmens bezeichnen, wenn n Unternehmen in dieser Branche tätig

¹⁴ Bain (1954) und Scherer (1980, speziell das vierte Kapitel) erörtern empirisches Material über Durchschnittskostenkurven.

$$^{15} \frac{d}{dq} \left(\frac{C(q)}{q} \right) = \frac{d}{dq} \left(\frac{F}{q} \right) + \frac{d}{dq} \left(\int_0^q C'(x) dx / q \right) < 0,$$

da $C'(q) < C'(x)$ für alle $x \in (0, q)$ bedeutet, daß $C'(q) - \int_0^q C'(x) dx / q < 0$.

¹⁶ Es sei $q \equiv \sum_i q_i$ (wobei $q_i > 0$). Dann gilt $C(q_i)/q_i > C(q)/q$, und hieraus folgt

$$\sum_i C(q_i) > \sum_i q_i C(q)/q = C(q).$$

sind. (Zur Vereinfachung sei vorausgesetzt, daß alle Unternehmen den gleichen Gewinn erwirtschaften.) Dieser Gewinn wird durch den Wettbewerb zwischen den Unternehmen bestimmt, und wir können hier offenlassen, um welche Art von Wettbewerb es sich handelt. Er versteht sich als Nettogröße (die Fixkosten sind bereits in Abzug gebracht). Es ist nahelegend anzunehmen, daß $\Pi(n)$ mit n abnimmt. Eine Industrie ist ein natürliches Monopol, wenn $\Pi(1) > 0 > \Pi(2)$, das heißt, wenn nur ein Unternehmen überleben kann, nicht aber zwei oder mehr¹⁷.

Schließlich können wir auch noch den Fall einer Aufsichtsbehörde betrachten, die nur unvollständig über die Kosten- (oder Nachfrage-) Funktion informiert ist. Eine solche Aufsichtsbehörde steht vor dem folgenden Problem. Sie kann entweder versuchen, für eine vollständige Ausschöpfung eventuell vorhandener zunehmender Skalenerträge zu sorgen (die Konsequenz wäre, daß sie nur ein einziges Unternehmen für die Erzeugung des betreffenden Produktes lizenziert). Oder aber sie erachtet es als wichtiger, bessere Informationen über die von ihr regulierte Branche zu gewinnen - dann würde sie Wettbewerb vorziehen (wir kommen hierauf im Abschnitt 2 dieses Kapitels unter dem Stichwort Wettkämpfe zurück sowie im Anhang zum Kapitel 4). Zwischen technischer Effizienz im Sinne einer Ausschöpfung von zunehmenden Skalenerträgen und dem Informationsstand der Aufsichtsbehörde besteht also ein Spannungsverhältnis. Ein natürliches Monopol liegt vor, wenn die Aufsichtsbehörde sich für erstere Alternative entscheidet.

Der Begriff der Subadditivität läßt sich problemlos auf den Fall einer Mehrproduktunternehmung erweitern. \mathbf{q} sei nun der Outputvektor oder Produktionsplan $\mathbf{q} = (q_1, \dots, q_m)$, die Zahl der Outputs sei m . $\mathbf{q}^1, \dots, \mathbf{q}^n$ soll n solcher Vektoren bezeichnen. (Hochgestellte Indizes bezeichnen Produktionspläne, tiefgestellte Güter.) Die Kostenfunktion C ist streng subadditiv, wenn

$$\sum_{i=1}^n C(\mathbf{q}^i) > C\left(\sum_{i=1}^n \mathbf{q}^i\right)$$

für alle \mathbf{q} mit der Eigenschaft $\sum_i q^i \neq 0$. Diese Definition läßt sich auch auf den Spezialfall der Einprodukt-Kostenfunktion anwenden. (Hierfür brauchen Sie nur zu unterstellen, daß die m Güter homogen sind, d.h. es handelt sich um ein und dasselbe Gut.) Interessanter ist, daß sich mit Hilfe dieser Formel Verbundvorteile darstellen lassen. Ein Beispiel: q_1 und q_2 seien

¹⁷ Stattdessen könnte man einen Wirtschaftszweig auch dann als natürliches Monopol bezeichnen, wenn $\Pi(2) > 0 > \Pi(3), \dots$

Mengen zweier unterschiedlicher Güter. Bei einer streng subadditiven Kostenfunktion gilt dann

$$C(q_1, 0) + C(0, q_2) > C(q_1, q_2).$$

$C(q_1, 0)$ und $C(0, q_2)$ werden als "stand-alone costs" bezeichnet (dies sind die Kosten, die anfallen, wenn nur eines dieser zwei Güter erzeugt wird, nicht aber das andere). Betrachten wir den Fall einer Eisenbahngesellschaft. Sie produziert Verkehrsleistungen. Wenn sie sowohl im Personen- als auch im Gütertransport tätig ist (in wirtschaftlicher Betrachtung ist dies die Produktion zweier verschiedener Güter), dürfte dies technologisch effizienter sein als wenn zwei verschiedene Gesellschaften auftreten, von denen sich die eine auf den Personen- und die andere auf den Güterverkehr spezialisiert. Ähnlich dürfte es sich bei der Elektrizitätsversorgung verhalten. Die Stromerzeugung für die Schwachlast und für die Spitzenlast kann man als Produktion zweier unterschiedlicher Güter betrachten. Aber die Kosten sind geringer, wenn diese beiden Güter von ein und demselben Elektrizitätsversorgungsunternehmen erzeugt werden¹⁸.

Kann die technologieorientierte Sicht des Unternehmens als eine "Theorie der Unternehmung" gelten?

Der technologieorientierte Ansatz ist vor allem darauf gerichtet, die Größe des Unternehmens zu bestimmen. Zunehmende Skalenerträge liefern einen Anlaß, Aktivitäten zusammenzufassen. Das Unternehmen wächst nicht ins Unermeßliche, weil, wie dies in den

¹⁸ Das Argument der zunehmenden Skalenerträge stellt auf Interdependenzen zwischen den Kosten der Erzeugung verschiedener Güter oder zwischen der Nachfrage nach den verschiedenen Gütern ab. Es können aber auch Synergieeffekte zwischen Einheiten auftreten, die an sich wenig miteinander zu tun haben. Zwei solche Synergieeffekte beruhen auf Unsicherheit und der Diversifikation von Risiken und wurden bei dem Versuch herangezogen, Unternehmenszusammenschlüsse anorganischer Natur zu erklären (dies sind Zusammenschlüsse, die weder zu Kostensenkungen führen noch die Möglichkeit eröffnen, Interdependenzen bei der Nachfrage auszunutzen; solche Zusammenschlüsse erfolgten in den USA in den sechziger und achtziger Jahren in großer Zahl). Erstens wird angeführt, daß eine Diversifikation der Aktivitäten eines Unternehmens, die zur Aufnahme ganz unterschiedlicher Aktivitäten führt, das Risiko des Unternehmens vermindert. (Dieses Argument erklärt eine solche Diversifikation der Unternehmensaktivitäten dann nicht, wenn die Aktionäre sie auch ohne weiteres selbst vornehmen könnten, indem sie ihr Portfolio entsprechend diversifizieren.) Zweitens wird angeführt, daß in dem Fall, in dem die Führungsspitze des Unternehmens trotz der Diversifizierung alle seine Aktivitäten im Auge und unter ihrer Kontrolle zu behalten vermag, eine Messung der Leistung des Managements deswegen leichter fällt, weil Zufallseinflüsse, die auf einzelnen Märkten auftreten, aufgrund des Gesetzes der großen Zahl weitgehend ausgeschaltet sind. Das Anreizsystem für die Manager kann sich infolgedessen stärker an ihrem tatsächlichen Erfolg orientieren (vergleiche Aron 1984).

Abbildungen 2b und 2c zu sehen ist, die Durchschnittskosten ab einem gewissen Prozeßniveau wieder ansteigen. Es gibt zwei unterschiedliche Gründe, warum die Aussagen, zu denen dieser Ansatz führt, für eine befriedigende Erklärung der Unternehmensgröße nicht ausreichen.

Erstens erklärt dieser Ansatz nicht, warum zunehmende Skalenerträge nur innerhalb eines Unternehmens anfallen. A priori ist es nicht selbstverständlich, warum sie nicht auch durch Vertragsbeziehungen zwischen rechtlich selbständigen Einheiten realisiert werden können. Betrachten wir zum Beispiel das oben angeführte Argument der "economies of massed reserves". Wir hatten darauf hingewiesen, daß ein Unternehmen, das mehrere verschiedene Märkte bedient, weniger Unsicherheit ausgesetzt ist als eine Mehrzahl voneinander getrennter Unternehmen, von denen jedes einen dieser Märkte bedient. Auf kostspielige Investitionen, die letztere tätigen müssen, um Nachfragespitzen zu bedienen, kann im ersteren Fall verzichtet werden. Aber dasselbe Ergebnis ließe sich durch eine vertragliche Vereinbarung erzielen, die regelt, unter welchen Umständen und zu welchen Preisen welches dieser Unternehmen das betreffende Gut an welches andere verkauft. In der Elektrizitätswirtschaft beispielsweise schließen die verschiedenen Erzeuger Abkommen, in denen sie sich zu einem wechselseitigen Austausch von Strom verpflichten - auf diese Weise werden Nachfrageschwankungen ausgeglichen. Ähnlich verhält es sich mit dem "Reparaturproblem" von Arrow et alii (1972), mit dessen Hilfe sie eine Produktionsfunktion mit zunehmenden Skalenerträgen ableiten. Ihre Analyse impliziert nicht, daß die Produktion innerhalb eines einzigen Unternehmens stattfinden muß. Ebenso gut könnten mehrere selbständige Unternehmen ein Abkommen über die gemeinsame Nutzung eines Reparaturservice abschließen, oder aber sie könnten ein selbständiges Reparaturunternehmen errichten, das sie mit Reparaturdiensten versorgt.

Zweitens ist es nicht selbstverständlich, daß die Durchschnittskostenkurve ab einem bestimmten hohen Prozeßniveau wieder ansteigt. Kostet zum Beispiel die Produktion des Outputs $q_1 + q_2$ mehr als die separate Produktion von q_1 und q_2 kosten würde, dann könnte eine Unternehmung, die $q_1 + q_2$ produzieren wollte, zwei unabhängige Betriebsteile schaffen, die ähnlich wie selbständige Unternehmen betrieben werden. Die Produktionsfunktion des Gesamtunternehmens würde dann niemals abnehmende Skalenerträge aufweisen. Eine Umformulierung dieses Arguments führt uns zu Williamsons "Rätsel der ausgewählten Eingriffe" (puzzle of selective intervention): Warum sollte es nicht möglich sein, zwei Unternehmen zusammenzuschließen und immer noch die Produktionsleistung zu erbringen, die die beiden Unternehmen vor dem Zusammenschluß erbracht haben (oder vielleicht sogar mehr, nämlich indem man "ausgewählte Eingriffe" vornimmt, die eine Leistungssteigerung

ermöglichen)? Wenn dem aber so ist, gibt es überhaupt eine Obergrenze für die Unternehmensgröße? (Ein Teil der Antwort auf diese Frage findet sich im nachfolgenden Unterabschnitt 1.4).

1.3 Die Unternehmung als eine langfristige Beziehung

Im Unterabschnitt 1.2 untersuchten wir Gründe, warum sich bestimmte Organisationen zu einem bestimmten Zeitpunkt zusammenschließen oder ihre Aktivitäten mit Hilfe eines statischen Vertrages koordinieren. Wir befassen uns nun mit Problemen, die bei langfristigen Beziehungen auftauchen - insbesondere damit, warum die Spielregeln, nach denen morgen Geschäfte gemacht werden, wann immer möglich bereits heute vereinbart werden sollten¹⁹. Der Einfachheit halber werden wir uns nur mit der Beziehung zwischen einem Lieferanten und einem Kunden beschäftigen. Wenn Unsicherheit eine Rolle spielt, wird vorausgesetzt, daß beide Parteien risikoneutral sind.

Idiosynkratische Investitionen und "Asset Specificity"

Bei langfristigen Beziehungen spielen häufig zwei Arten von Kosten eine Rolle, die als Umstellungskosten und als spezifische Investitionen bezeichnet werden. Williamson (1976) hat Umstellungskosten im einzelnen analysiert. Dies geschah in seiner Analyse der Probleme, die auftreten, wenn der Staat natürliche Monopole zeitlich befristet vergibt und zwar mittels einer Versteigerung²⁰. Bei diesem Arrangement kann es natürlich vorkommen, daß sich nach Ablauf einer solchen Frist ein anderer Unternehmer (ein anderes Team) als vorher das Recht ersteigert, sich als natürlicher Monopolist zu betätigen. Zu den bekanntesten der dann auftretenden Umstellungskosten gehören Kosten, die entstehen, weil sich das neue Team erst einarbeiten muß und das alte wenig Neigung zeigt, ihm dabei behilflich zu sein und ihm Informationen zu übermitteln. Diese Kosten mögen eine Aufsichtsbehörde dazu veranlassen, auf solche wiederholten Versteigerungen lieber zu verzichten; dieselben Kosten treten aber auch bei langfristigen Lieferbeziehungen auf und mögen den Abnehmer davon abhalten, öfters seinen Lieferanten zu wechseln.

¹⁹ Wegen einer Erörterung unvollständiger Verträge vgl. Unterabschnitt 1.4.

²⁰ Williamson reagierte damit auf Beiträge von Demsetz (1968), Stigler (1968) and Posner (1972), die sich für die Versteigerung solcher befristeter Lizenzen mit dem Argument aussprachen, daß sich dadurch die sogenannte Mißbrauchsaufsicht erleichtern lasse, die das Wettbewerbsrecht für derartige Monopolisten vorsieht.

Umstellungskosten können in einem gewissen Sinne unter dem Begriff der idiosynkratischen Investition subsumiert werden. Sobald zwei Parteien miteinander in Geschäftsbeziehungen getreten sind, kann es sein, daß es für sie allemal vorteilhafter ist, zusammenzubleiben, statt den Geschäftspartner zu wechseln. Eine andere, noch wesentlichere Art von idiosynkratischen Investitionen liegt vor, wenn die Aufgabe nicht darin besteht, heute eine bestimmte Lieferung zu tätigen, sondern die Erwartung vorhanden ist, morgen einen Auftrag zu erhalten. Dies ist beispielsweise der Fall, wenn ein Lieferant sich "spezifische" Anlagen beschaffen muß, d.h. solche, die eine besondere Eignung für die Ausführung eines Auftrags eines bestimmten Abnehmers besitzen, die gewissermaßen auf die Ausführung solcher Aufträge zugeschnitten sind. Bei Werkzeugmaschinen ist dies häufig so. Oder ein anderer Fall: Einem Unternehmen entstehen Kosten für den Verkauf oder die Erstellung eines Finalproduktes und zwar bereits bevor ein Zwischenprodukt, das der Unternehmer für die Erstellung des Finalproduktes benötigt, geliefert wird. In dieser Situation befindet sich beispielsweise ein Konzertveranstalter, mietet er doch normalerweise den Konzertsaal noch bevor der Sänger auftritt. Ein dritter Fall: Ein weiterverarbeitender Betrieb erwirbt Maschinen, die sich nur für die Verarbeitung bestimmter Materialien eignen. Williamson (1975) führt noch zwei weitere Arten von Spezifität an: Standortspezifität und spezifische Investitionen in Humankapital. Standortspezifität tritt auf, wenn räumliche Nähe des Lieferanten oder des Kunden die Produktivität deutlich steigert. In der Stahlindustrie zum Beispiel empfiehlt es sich, Schmelz- und Preßoperationen auf relativ engem Raum zusammenzufassen, da dann die Transportkosten geringer ausfallen und es nicht erforderlich wird, den Stahl zum Pressen wieder zu erhitzen. Bei einem Kraftwerk sind die Transportkosten geringer, wenn es sich in der Nähe eines Kohlebergbaus befindet. Spezifische Humankapitalinvestitionen erfolgen beispielsweise, wenn das Personal eine besondere Vertrautheit mit den Eigenheiten eines bestimmten Produktionsprozesses erwirbt oder bei der Teamarbeit²¹.

Alle diese Arten von Spezifität ziehen dieselbe Folge nach sich: Die Parteien, die heute eine Geschäftsverbindung eingehen, wissen wohl, daß diese Verbindung in der Zukunft zum Entstehen einer Rente führen wird und daß es sich lohnen wird, sich diese Rente anzueignen. Es kommt darauf an, daß der Prozeß der Aneignung in die richtigen Bahnen gelenkt wird (insbesondere sollte Effizienz gewährleistet werden, es sollte also nicht ex post zu wenig

²¹ Wegen einer Analyse spezifischer Investitionen innerhalb eines Arbeitsverhältnisses vgl. Williamson et al. (1975).

Leistungsaustausch stattfinden). Es sollte eine derartige Verteilung der künftigen Erträge gewährleistet werden, daß ex ante der Anreiz besteht, spezifische Investitionen in dem Umfang vorzunehmen, der Pareto-effizient ist.

Der kritische Punkt bei spezifischen Investitionen ist, daß selbst dann, wenn Lieferant und Kunde einander ex ante nach Belieben aus einer Vielzahl von Konkurrenten auswählen können, die Lage sich ex post ganz anders darstellt. Ex post handelt es sich um ein bilaterales Monopol, in dem es für beide beteiligten Parteien vorteilhafter ist, miteinander Geschäfte zu treiben als mit Dritten. In dieser bilateralen Monopolsituation aber hat jede der beiden Parteien ein Interesse daran, sich die Erträge anzueignen, die sie gemeinsam hervor gebracht haben. Dieses Bestreben aber wird ex post zu einem Hindernis für die Zusammenarbeit, es kann bewirken, daß das Transaktionsvolumen ex post zu gering ausfällt, und daß in Vorausschau dieser Entwicklung ex ante spezifische Investitionen nicht im wünschenswerten, weil effizienten Ausmaß erfolgen²².

Preisbildung beim bilateralen Monopol und das Ex-Post-Volumen des Leistungsaustausches zwischen einem Monopolisten und einem Monopsonisten

Um eine konkrete Vorstellung zu gewinnen, wollen wir annehmen, daß es zwei Perioden gibt: $t = 1$ (ex ante) und $t = 2$ (ex post). Ein Lieferant und ein Kunde können in Periode 1 einen Vertrag schließen oder es auch bleiben lassen. Um das Ex-post-Problem zu analysieren, vernachlässigen wir für einen Moment die spezifischen Investitionen in der ersten Periode. Zu Beginn der Periode 2 erfahren die beiden Parteien, welche Erträge das Geschäft in Periode 2 stiften wird (in der Periode 1 weisen diese Erträge noch den Charakter von Zufallsvariablen auf). Gegenstand des Vertrages ist die Lieferung einer Einheit eines unteilbaren Gutes (oder der Lieferant realisiert ein "Projekt"). Die ausgetauschte Menge beträgt also entweder 0 oder 1 (in einer komplizierteren Version des Modells kann es sich auch um eine stetige Variable handeln). Der Wert des Gutes für den Käufer beträgt v , und die Produktionskosten des Lieferanten c ²³. Der Ertrag, der für eine Aufteilung zwischen

²² Die nachfolgende Diskussion beruht insbesondere auf Farrell (1985), Grossman und Hart (1984) sowie Hall und Lazear (1984).

²³ Dieser Wert und die Kosten können durch die anderweitigen Möglichkeiten bedingt sein, die sich für den Einsatz der betreffenden Ressourcen bieten. Beispielsweise mag v den Unterschied zwischen dem Wert, den das Gut für den Käufer besitzt, solange er in der betreffenden Beziehung verbleibt, und dem, den es in einer andere Beziehung besäße, bezeichnen.

den Parteien zur Verfügung steht, beläuft sich folglich auf $v - c$. Wird das unteilbare Gut zu dem Preis p verkauft, dann fließt dem Käufer der Ertrag $v - p$ zu und dem Lieferanten der Ertrag $p - c$. Kommt das Geschäft nicht zustande, so ist der Ertrag für beide Parteien null.

Bargaining

Nehmen wir an, daß in der Periode 1 kein Vertrag abgeschlossen wird. In Periode 2 treten die beiden Parteien in ein Bargaining ein, als dessen Ergebnis dann feststeht, ob ein Geschäft zustandekommt und zu welchem Preis.

Sind v und c allgemein bekannt (d.h. beide Parteien kennen v und c , bevor sie in Verhandlungen eintreten), sollten wir erwarten, daß das Geschäft genau den effizienten Umfang annimmt (in unserem Fall heißt dies, es findet nur statt, wenn $v \geq c$). Um dies klarzulegen, unterstellen wir, daß $v > c$, kein Geschäft vereinbart ist und den beiden Parteien infolgedessen auch kein Ertrag zufließt. Dann könnte eine der Parteien das Angebot unterbreiten, zu einem Preis p aus dem Intervall (c, v) eine Transaktion vorzunehmen, so daß beide einen Nettogewinn erzielen. Dies ist beiden Parteien lieber, als auf das Geschäft ganz zu verzichten. Wäre hingegen $v < c$ und käme das Geschäft dennoch zustande, so würde eine der Parteien einen Verlust erleiden und besser daran tun, auf das Geschäft zu verzichten. Allgemeiner läßt sich sagen, daß bei symmetrischer Information Verhandlungen zu einem effizienten Ergebnis führen, das Problem eines ineffizienten Ex-post-Geschäftsvolumens tritt nicht auf. (Dies ist eine Version des Coase-Theorems [1960].) Wie wir nun sehen werden, kann asymmetrische Information bewirken, daß die Verhandlungen nicht zu einem effizienten Ergebnis führen.

Oft stellen der Wert v , den das Gut für den Kunden besitzt, und die Kosten c des Lieferanten "private Information" dar. Nur der Käufer kennt diesen Wert, und nur der Hersteller die Kosten²⁴. Dann kann es geschehen, daß das effiziente Geschäftsvolumen wegen des (bilateralen) Monopolpreisproblems nicht erreicht wird. Die Ineffizienz rührt von dem Umstand her, daß beide Parteien sich den Ertrag aneignen wollen, den das Geschäft stiftet. Wegen der vorliegenden Informationsmängel, der asymmetrischen Information, gehen sie

²⁴ In einigen Fällen lassen sich diese Werte ex post aufgrund des betrieblichen Rechnungswesens ermitteln - im Handel mag dies beispielsweise der Fall sein. Diese Daten sind aber häufig nicht unbedingt verlässlich. Beispielsweise mag die Zurechnung der Kosten Schwierigkeiten verursachen, ferner kann moralisches Risiko auftreten (welche Anstrengungen zur Verringerung der Kosten der Anbieter ex post unternimmt, läßt sich möglicherweise nicht beobachten). Außerdem läßt es sich vielleicht nicht ohne weiteres feststellen, welchen Beitrag der Verkauf des betreffenden Zwischenprodukts zum Gewinn des Anbieters leistet.

das Risiko ein, daß das Geschäft nicht (allgemeiner: nicht im Pareto-optimalen Umfang) zustandekommt. Jede der beiden Parteien versucht nämlich, sich einen möglichst großen Teil des Kuchens zu sichern, riskiert damit aber auch, den Bogen zu überspannen.

Ein einfaches Beispiel eignet sich für eine Illustration, wie es zu Ineffizienz kommen kann, wenn die "falsche Partei" sich in einer "starken Verhandlungsposition" befindet. Nehmen wir an, daß die Kosten c beiden Parteien bekannt sind, der Wert v aber nur dem Nachfrager. Die Vorstellungen des Anbieters über v werden durch die Verteilungsfunktion $F(v)$ beschrieben; für die Dichtefunktion gilt $f(v) > 0$ im Intervall $[\underline{v}, \bar{v}]$, für die Verteilungsfunktion $F(\underline{v}) = 0$ und $F(\bar{v}) = 1$. Es wird vorausgesetzt, daß das Geschäft mit positiver Wahrscheinlichkeit einen positiven Ertrag stiftet (d.h. $\bar{v} > c$). Weiter wird der Einfachheit halber angenommen, daß diese Wahrscheinlichkeit kleiner als eins ist. Ferner wird unterstellt, daß sich der Anbieter in Periode 2 in einer "starken Verhandlungsposition" befindet. Damit ist gemeint, daß er den Preis p festsetzt und über diesen nicht zu verhandeln bereit ist. (Es mag als naheliegender erscheinen, daß in einem bilateralen Monopol beide Parteien ungefähr gleich stark sind, aber der von uns unterstellte Extremfall hat den Vorteil, daß er die Ableitung eines einfachen und dennoch aufschlußreichen Ergebnisses ermöglicht.) Wenn der Anbieter den Preis p festsetzt, kauft der Nachfrager nur, wenn $v \geq p$. Die Wahrscheinlichkeit, daß das Geschäft zustandekommt, beträgt also $1 - F(p)$, der Erwartungswert des Gewinns des Anbieters beläuft sich auf

$$(p - c)[1 - F(p)].$$

Leitet man diese Funktion nach p ab, so erhält man als Bedingung erster Ordnung²⁵

$$[1 - F(p)] - (p - c)f(p) = 0. \quad (1)$$

Gleichung 1 besagt, daß eine Preiserhöhung von p auf $p + dp$ mit der Wahrscheinlichkeit $1 - F(p)$ den zusätzlichen Gewinn dp stiftet und mit der Wahrscheinlichkeit $f(p)dp$ dazu führt, daß das Geschäft nicht zustandekommt. Der Gewinn des Anbieters geht infolgedessen um $p - c$ zurück. Im Optimum gleichen sich diese zwei Effekte gerade aus. Man beachte, daß das Geschäftsvolumen suboptimal ist, d.h. das Geschäft kommt auch in solchen Fällen nicht zustande, in denen sein Zustandekommen zu einer Effizienzerhöhung führen würde²⁶.

²⁵ Die (lokale) Bedingung zweiter Ordnung lautet $-2f(p) - (p - c)f'(p) \leq 0$. Unter Verwendung der Bedingung erster Ordnung kann sich der Leser davon überzeugen, daß diese Bedingung zweiter Ordnung dann und nur dann erfüllt ist, wenn die Verteilung $f/(1 - F)$ eine zunehmende Risikorate (hazard rate) besitzt.

²⁶ Ist es "gemeinsames Wissen", daß das Geschäft einen Gewinn stiftet (d.h. $\underline{v} > c$), so kann das Volumen der Transaktion effizient sein, muß es aber nicht sein.

Effizienz wäre nur gewährleistet, wenn $p = c$ ist (der Abnehmer bezahlt dem Anbieter einen Preis in Höhe seiner Kosten, ersterer trifft dann die "richtige Entscheidung", d.h. kauft immer dann, wenn der Kauf zum Entstehen eines nichtnegativen Nettoertrages führt). Aus Gleichung 1 aber folgt, daß stets $p > c$. Der Grund für dieses ineffiziente Ergebnis ist folgender: Betrachten wir zuerst den Fall, daß der Anbieter einen Preis in Höhe seiner Kosten verlangt. Dann erwirtschaftet er keinen Gewinn. Verlangt er stattdessen einen etwas höheren Preis, so erwirtschaftet er mit positiver Wahrscheinlichkeit einen Gewinn. Daß die Preiserhöhung mit einer gewissen Wahrscheinlichkeit dazu führt, daß das Geschäft nicht zustandekommt (also zu einem ineffizienten Transaktionsvolumen führt), fällt für ihn nicht ins Gewicht, weil dies für ihn mit keinerlei Gewinneinbuße verbunden ist. Ausgangspunkt war ja ein Gewinnaufschlag von null Prozent.

Gleichung 1 ist nichts anderes als die bekannte Formel für den Monopolpreis (siehe Kapitel 1) bei einer Nachfragekurve mit der Gestalt $p = D(p) = 1 - F(p)$. Ist der Anbieter risikoneutral, so ist für ihn das soeben dargestellte Entscheidungsproblem identisch mit einem zweiten Problem. In diesem zweiten Problem gibt es eine kontinuierliche Menge von Nachfragern, jeder von ihnen fragt höchstens eine Einheit des Gutes nach, und die Bewertung des Gutes durch diese Nachfrager wird durch die Verteilungsfunktion $F(\cdot)$ dargestellt.²⁷

Bemerkung: Dieses Ergebnis läßt sich verallgemeinern. So ist gezeigt worden, daß Bargaining nicht zu einem effizienten Ergebnis führt, wenn die folgenden drei Bedingungen erfüllt sind: 1) der Wert und die Produktionskosten des verhandelten Gegenstandes sind private Information, 2) es ist nicht sicher, daß das Geschäft einen positiven Ertrag stiftet (d. h. mit positiver Wahrscheinlichkeit ist $v < c$), 3) die Parteien können sich frei entscheiden, ob und wieviel sie austauschen wollen (und sind infolgedessen in der Lage, dafür zu sorgen, daß das Ergebnis nicht negativ wird - nämlich indem sie darauf verzichten, das Geschäft

²⁷ In ähnlicher Weise läßt sich dynamisches Bargaining zwischen einem Anbieter und einem Nachfrager, bei dem der Nachfrager sequentiell Angebote unterbreitet, in ein Problem intertemporaler Preisdifferenzierung durch einen Monopolisten uminterpretieren (vgl. den Anhang zum Kapitel 1). In der Tat ist das hier untersuchte Problem der Monopolpreisbildung nichts anderes als eine Einperiodenversion (bzw. eine Mehrperiodenversion mit bindenden Selbstverpflichtungen) des intertemporalen Preisdifferenzierungsproblems.

abzuschließen) (siehe Myerson und Satterthwaite 1983)²⁸. Bargaining führt zu Ineffizienz (im allgemeinen wird zu wenig ausgetauscht). Das zugrundeliegende Prinzip ist, daß die oben dargestellte mit Monopolpreisen verbundene Ineffizienz bei Problemen dieser Art regelmäßig wieder auftritt, weil jede der beiden Parteien unvollständig über die andere informiert ist.

Verträge

Der Umstand, daß es ex post zu Ineffizienz kommt, schafft für die zwei Parteien einen Anreiz, ex ante einen Vertrag abzuschließen, der dem vorbeugt oder die Ineffizienz zumindest verringert. In dem zuvor erörterten Fall, in dem nur der Wert, den das Produkt für den Kunden besitzt, eine private Information darstellt, läßt sich dies sehr einfach erreichen.

²⁸ Fortgeschrittene Leser können sich hiervon auf verschiedenere Weise überzeugen. Man könnte beispielsweise annehmen, daß die Kosten c des Anbieters über das Intervall $[\underline{c}, \bar{c}]$ gemäß der Verteilungsfunktion $G(\cdot)$ mit der Dichtefunktion $g(\cdot)$ verteilt sind. Die Zahlungsbereitschaft des Käufers v sei durch die Verteilungsfunktion $F(\cdot)$ und die Dichtefunktion $f(\cdot)$ gekennzeichnet. Damit ein effizientes Handelsvolumen zustandekommt, muß das Geschäft zwischen den beiden Parteien immer dann und nur dann zustandekommen, wenn $v \geq c$. Ex ante beträgt der Erwartungswert der Rente also

$$W = \int_{\underline{c}}^{\bar{c}} \left(\int_c^{\bar{v}} (v - c)f(v)dv \right) g(c)dc.$$

$V(v)$ soll den Erwartungswert der Gewinne bezeichnen, die der Käufer im Ergebnis des Verhandlungsprozesses zu erlangen hofft (die Zufallsgröße sind die Kosten des Verkäufers). Nun gilt $V(v) = G(v)$, d.h. der Erwartungswert des Gewinns wächst mit einer Rate, die gleich der Wahrscheinlichkeit ist, daß ein Käufer mit der Zahlungsbereitschaft v das Geschäft abschließt. (Um dies zu erkennen, mag sich der Leser vor Augen halten, daß sich ein Käufer mit der Zahlungsbereitschaft v stets ebenso wie einer mit der Zahlungsbereitschaft $v + dv$ verhalten kann, so daß $V(v + dv) - V(v) \leq G(v + dv)dv$. Entsprechendes gilt auch im umgekehrten Fall.) Analog gilt das folgende: $C(c)$ soll den Erwartungswert des Gewinns des Anbieters bezeichnen, wenn seine Kosten c betragen. Dann ist $C(c) = -[1 - F(c)]$. Ferner folgt daraus, daß die Teilnahme an den Verhandlungen keine Kosten verursacht, daß $V(\underline{v}) \geq 0$ und $C(\bar{c}) \geq 0$. Folglich

$$V(v) \geq \int_v^{\bar{v}} G(x)dx$$

und

$$C(c) \geq \int_c^{\bar{c}} [1 - F(x)]dx.$$

Mit Hilfe einer einfachen partiellen Integration läßt sich dann zeigen, daß

$$\int_{\underline{v}}^{\bar{v}} V(v)f(v)dv + \int_{\underline{c}}^{\bar{c}} C(c)g(c)dc$$

stets größer als W ist, wenn $\bar{c} > \underline{v}$ (die Träger überlappen sich) - und das ist ein Widerspruch.

Myerson und Satterthwaite (1983) untersuchen die Frage, welcher Bargainingmechanismus optimal ist, mit anderen Worten, wie der Verhandlungsprozeß "strukturiert" werden sollte (natürlich ist nicht gesagt, daß dieses optimale Verfahren in der Realität auch zur Anwendung gelangt). Cramton et al. (1987) zeigen auf, daß es nicht zu Ineffizienzen kommen muß, wenn das Eigentum an dem Gut anfänglich eher symmetrisch verteilt ist (wie dies innerhalb einer Handelsgesellschaft der Fall ist), und verallgemeinern die Ergebnisse von Myerson und Satterthwaite.

Es genügt, die "informierte Partei", den Kunden, in die Lage zu versetzen, den Preis festzulegen (d.h. die Verhandlungsposition des Kunden wesentlich zu stärken). Da c bekannt ist, entsteht dann keine Ineffizienz mehr. Der Nachfrager verhält sich als Monopsonist, er wählt den niedrigsten Preis, zu dem der Anbieter bereit ist, das Gut zu liefern (nämlich den, bei dem letzterer zwischen den Alternativen Abschluß des Geschäfts oder kein Abschluß gerade indifferent ist), und das heißt $p = c$. Der Kunde eignet sich ex post alle Erträge aus dem Geschäft an. Bestandteil des Vertrages kann sein, daß der Nachfrager ex ante eine Pauschale an den Anbieter zahlt. Durch entsprechende Dimensionierung dieser Pauschale läßt sich der Ertrag aus dem Geschäft (genauer: sein Erwartungswert) in jeder gewünschten Weise zwischen den beiden aufteilen. (Verallgemeinernd können wir sagen: Bei der Formulierung des Vertrages geht es darum, die Größe des "Kuchens" zu maximieren. Seine Aufteilung hängt davon ab, wie stark ex ante die Verhandlungsposition der beiden Geschäftspartner ist).

Ist der Wert, den das Gut für den Kunden besitzt, hingegen allgemein bekannt, und sind die Kosten des Lieferanten private Information, dann ist es effizient, dem Anbieter das Recht auf Festsetzung des Preises zu gewähren. Hier lassen sich Ähnlichkeiten mit einer "Beziehung mit sequentiellen Entscheidungsbefugnissen" (sequential authority relationship) feststellen (vergleiche Teilabschnitt 1.4). Die eine Partei hat die Entscheidungsbefugnis über den Preis, die andere nur darüber, ob ein Leistungsaustausch zustandekommt.

Ist entweder der Wert, den das Gut für den Kunden hat, oder sind seine Kosten bereits am Tage des Vertragsabschlusses allgemein bekannt (wir sagen dafür: gemeinsames Wissen, common knowledge), nimmt der effiziente Vertrag die besonders einfache Form eines Festpreisvertrages an. Ist beispielsweise c zum Zeitpunkt des Vertragsabschlusses allgemein bekannt, so kann das folgende vertraglich vereinbart werden: "Der Käufer entscheidet, wieviel geliefert werden soll (im obigen Beispiel kommen nur 0 und 1 in Frage). Der Preis beträgt c ."

Unter bilateral asymmetrischer Information ist es im allgemeinen nicht mehr effizient, der einen Partei das Recht auf die Festsetzung des Preises zu übertragen bzw. der Menge, wenn der Preis vorgegeben ist. Die Parteien dürften daher den Wunsch verspüren, über eine andere Vertragsgestaltung nachzudenken. Eine Möglichkeit besteht darin, sowohl das Transaktionsvolumen als auch den Preis im voraus fest zu vereinbaren. Der Leistungsaustausch kommt dann auf jeden Fall zustande, unabhängig davon, welche Werte c und v annehmen. Solch ein Vertrag ist natürlich nur dann effizient, wenn die Parteien ex ante wissen, daß der Leistungsaustausch auf jeden Fall wohlfahrtsvermehrend ist. (Im hier betrachteten Fall ist dieser einfache Vertrag im allgemeinen jedem Vertrag überlegen, der

einer der beiden Parteien das Recht überträgt, den Preis festzusetzen.²⁹⁾

Zusammenfassend können wir sagen, daß Ex-post-Bargaining möglicherweise nicht zu einem effizienten Transaktionsvolumen führt. Es empfiehlt sich daher, die Menge der Entscheidungsmöglichkeiten, die in der zweiten Periode bestehen, in der ersten durch möglichst einfache vertragliche Vereinbarungen zu verkleinern. Mit anderen Worten, kraft Vertrag sollten in den Entscheidungsprozeß der zweiten Periode zusätzliche Restriktionen eingeführt werden. Stellt die Größe, die die relevante Variable bei der einen Vertragspartei annimmt, gemeinsames Wissen dar, sollte sich die andere (die mit der privaten Information), die Entscheidung vorbehalten, ob die Transaktion zu einem vorgegebenen Preis zustandekommt; hört sie erst nach Vertragsabschluß auf, private Information der ersteren Partei zu sein, sollte die letztere sich das Recht vorbehalten, den Preis zu setzen. Allgemeiner gesprochen lautet die Regel, daß diejenige Partei, die einen Informationsvorsprung hat, auch die stärkere Verhandlungsposition besitzen sollte.

Um dieses Ergebnis auf ein Beispiel anzuwenden, betrachten wir die Beziehung zwischen einem Verleger und einer Druckerei. Der Verleger ist der Käufer, die Druckerei der Anbieter, die Transaktion eine zusätzliche Druckauflage. Sind die Kosten einer zusätzlichen Druckauflage bekannt und gleich c , dann sollte der Verleger entscheiden können, ob

²⁹⁾ Es kommt nicht zu Ineffizienz, wenn die Parteien einen optimalen Vertrag unterzeichnen, bevor sie ihre private Information erhalten. Dies wurde von Arrow (1979) und d'Aspremont und Gerard-Varet (1979) gezeigt. Fortgeschrittene Leser können dies wie folgt überprüfen: Nehmen wir an, daß v a priori über $[\underline{v}, \bar{v}]$ gemäß der Verteilungsfunktion F mit der Dichtefunktion f verteilt ist. c ist über das Intervall $[\underline{c}, \bar{c}]$ gemäß der Verteilungsfunktion G mit der Dichtefunktion g verteilt. Die beiden Verteilungen sind voneinander unabhängig. Nehmen wir weiter an, daß die beiden Parteien ex ante zu der Übereinkunft gelangen, gleichzeitig einer dritten Partei ihre private Information mitzuteilen, nachdem sie diese in der zweiten Periode erhalten haben. \bar{v} und \bar{c} sollen diese Mitteilungen bezeichnen (sie brauchen nicht der Wahrheit zu entsprechen). Die vereinbarte Regel sieht vor, daß das Geschäft dann und nur dann zustandekommt, wenn $\bar{v} \geq \bar{c}$. Wenn beide Parteien die Wahrheit sagen, ist Effizienz gewährleistet. Die Zahlung des Käufers an den Verkäufer beträgt $p(\bar{v}, \bar{c})$ und hängt nicht vom Zustandekommen oder Umfang des Geschäfts ab (sie erfolgt selbst, wenn $\bar{v} < \bar{c}$). Sie bestimmt sich gemäß der Gleichung

$$p(\bar{v}, \bar{c}) = \int_{\bar{c}}^{\bar{v}} w g(w) dw - \int_{\bar{v}}^{\bar{c}} b f(b) db + \text{Konstante}$$

Bei einer solchen Zahlung haben die beiden Parteien Veranlassung, die Wahrheit zu sagen. Man bezeichnet sie als Zahlung gemäß dem Erwartungswert des externen Effekts, weil beispielsweise der Käufer, wenn er \bar{v} statt $\bar{v} + dv$ mitteilt, die Zahlung um $\bar{v}g(\bar{v})dv$ vergrößert, und dies ist nichts anderes als der Erwartungswert der Kosten, mit denen die entsprechende Erweiterung des Geschäftsvolumens verbunden ist. (Dieses zusätzliche Volumen wird realisiert, wenn $c = \bar{c}$ im Intervall $[\bar{v}, \bar{v} + dv]$ liegt, und dies geschieht mit der Wahrscheinlichkeit $g(\bar{v})dv$.) Eine Verallgemeinerung dieses Mechanismus (er verletzt für manche Werte die Bedingung, daß ex post die Rente nicht negativ werden darf) findet sich bei Maskin (1985), Pratt und Zeckhauser (1985) sowie bei Johnson et al. (1986).

Ein solcher Mechanismus ist stets effizient, aber er läßt sich nicht so leicht in Gang setzen wie die im Text vorgesehenen. (Subjektive Wahrscheinlichkeitsverteilungen müssen beschrieben werden, private Information muß ex post an die dritte Partei - oder an ein Aufnahmegerät - weitergegeben werden. Ferner muß das Gut, das Gegenstand des Geschäfts werden soll, in der Periode 1 genau beschrieben werden.)

die zusätzliche Auflage gedruckt wird, der Preis sollte c betragen. Ist c hingegen private Information (etwa, weil die Druckerei noch andere Verwendungsmöglichkeiten für ihre Kapazität hat, die nur ihr bekannt sind) und v leicht meßbar (Beispiel: eine große Zahl von Exemplaren des Buches wird zu einem leicht feststellbaren Preis bestellt), sollte es der Druckerei überlassen sein, den Preis festzusetzen.

Spezifische Investitionen und das "Hold-up Problem"

Nehmen wir an, daß zum Zeitpunkt 1 ein Anbieter eine Investition vornimmt, die zu einer Verringerung der Produktionskosten c führt, und ein Nachfrager eine Investition vornimmt, die zu einer Erhöhung des Nutzens (Werts) führt, den ihm das Gut stiftet. Insbesondere soll angenommen werden, daß diese Investitionen in dem Sinne spezifisch sind, daß weder die Kostensenkung noch die Nutzenerhöhung eintritt, wenn die beiden statt miteinander mit anderen Partnern Leistungen austauschen.

Bargaining

Nehmen wir erneut zunächst an, daß in Periode 1 kein Vertrag zustandekommt. Folglich müssen die beiden Parteien in Periode 2 darüber verhandeln, ob eine Lieferung erfolgen soll und zu welchem Preis.

Erreicht der Leistungsaustausch ex post nicht den effizienten Umfang, so hat dies offenbar einen Einfluß auf die Ertragsrate der Investitionen. Ist der Leistungsaustausch zu gering, so haben Anbieter und Nachfrager Grund, ihre Investitionen zu verringern, ist doch die Wahrscheinlichkeit, daß sie einen entsprechenden Nutzen stiften, geringer als im Optimum³⁰. Um die Frage nach dem effizienten Transaktionsvolumen analytisch von der nach dem effizienten Investitionsvolumen zu trennen, nehmen wir an, daß ex post v und c gemeinsames Wissen darstellen. Ein Leistungsaustausch kommt infolgedessen nur zustande, wenn $v \geq c$. Wir können uns auf die Frage konzentrieren, welchen Einfluß die ex post erfolgende Aufteilung der Erträge auf die ex ante bestimmten spezifischen Investitionen ausübt. Welchen Teil seiner spezifischen Investition gewinnt der Nachfrager (Anbieter) zurück?

³⁰ Die Wirkung einer solchen Verringerung des Geschäftsvolumens wird in der Arbeit von Tirole (1986a) näher untersucht.

Nehmen wir an, es ist bekannt, daß zum Zeitpunkt des Vertragsabschlusses der Wert v , den das Gut für den Nachfrager besitzt, gleich 3 ist. Der Anbieter kann "investieren" ($I = 2$ ausgeben) oder "nicht investieren" ($I = 0$ ausgeben). Investiert er, verringern sich seine ex post Grenzkosten auf null. Tut er es nicht, so sind sie "hoch" (> 3). Ex-post-Bargaining soll zur Nash-Lösung führen³¹: Der Ertrag, der im Ergebnis der Transaktion entsteht, wird gleichmäßig verteilt. Gilt $c = 0$, dann ist er 3 und der Preis des Gutes infolgedessen 1,5 (so daß jede Partei einen Ertrag von 1,5 erzielt). Ist c hoch, dann stiftet ein Leistungsaustausch keinen Ertrag; er kommt folglich nicht zustande und keine der beiden Parteien erzielt einen Gewinn. Betrachten wir nun die spezifische Investition. Verzichtet der Anbieter auf die Investition, so erzielt er einen Gewinn von 0. Der Gewinn beträgt $-2 + 1,5 < 0$ (wir haben auf eine Abzinsung der Erträge der zweiten Periode verzichtet), wenn er investiert. Also unterbleibt die Investition. Die Investition würde aber die soziale Wohlfahrt erhöhen. Sie würde einen Nettonutzen von $3 - 2 > 0$ stiften.

Dieses Beispiel läßt sich leicht auf den Fall einer kontinuierlichen Menge von Investitionen verallgemeinern. Nehmen wir an, daß die Produktionskosten eine deterministische Funktion $c(I)$ der Größe der vom Anbieter vorgenommenen Investition sind, wobei $c'(I) < 0$ und $c''(I) > 0$ (die Investition senkt die Kosten, aber mit abnehmender Rate). v sei der (deterministische) Wert des Gutes. Zur Vereinfachung wird angenommen, daß $v \geq c(0)$. Der Preis werde ex post gemäß der Nash-Verhandlungslösung bestimmt: $p(I) = [c(I) + v]/2$ (so daß $v - p(I) = p(I) - c(I)$), sobald I investiert worden ist ($v \geq c(I)$). Der Gewinn des Anbieters beläuft sich dann auf:

$$\max_I [p(I) - c(I) - I] = \max_I [v/2 - c(I)/2 - I].$$

Mit anderen Worten: Eine Kosteneinsparung von \$1 führt für den Anbieter nur zu zusätzlichen Erlösen in Höhe von 50 Cent. Die anderen 50 Cent behält der Nachfrager. Aus seiner höchstpersönlichen Sicht (die nur seinen privaten, d.h. ihm persönlich verbleibenden Vorteil berücksichtigt) ist die optimale Investition $-c'(I) = 2$. Im Unterschied hierzu wird die gesamtgesellschaftlich optimale Investition durch das Ergebnis der Maximierungsaufgabe

$$\max_I [v - c(I) - I]$$

³¹ Vergleiche Nash (1953). Nash verwendete für die Ableitung seiner Verhandlungslösung einen axiomatischen Ansatz; er rechtfertigte dies mit dem Hinweis, es handle sich um das Ergebnis eines "Nachfragespiels" mit simultanen Zügen. Rubinstein (1982) und Binmore (1982) rechtfertigen die Nash-Lösung für den Fall sequentieller, nicht-kooperativer Verhandlungen, wobei sie unterstellen, daß die Parteien für die eingesetzten Ressourcen keine alternativen Verwendungsmöglichkeiten besitzen. (Vergleiche den Abschnitt über dynamische Spiele mit vollständiger Information in der Gebrauchsanweisung für die Spieltheorie am Ende dieses Buches.)

bestimmt, und die ist $-c'(I) = 1$. Da c konvex ist ($-c'$ ist fallend), ist das realisierte (weil aus der höchstpersönlichen Sicht des Investors optimale) Investitionsvolumen aus wohlfahrts-ökonomischer Sicht tatsächlich suboptimal³².

Das Problem ist natürlich, daß die investierende Partei sich nicht alle Kosteneinsparungen (Wertsteigerungen) aneignen kann, die sich dank der Investition realisieren lassen. Die andere Partei kann sich mit Hilfe der Drohung, die Geschäftsbeziehungen abubrechen, einen Teil dieser Einsparungen aneignen. Genau das ist es, was Williamson (1975) Opportunismus nennt. Er betont, daß es zu Unterinvestition in spezifische Kapitalgüter kommen muß, wenn die Situation darauf angelegt ist, daß ex post Verhandlungen (Bargaining) stattfinden und zwar unter den Bedingungen eines bilateralen Monopols. Ein Stromerzeugungsunternehmen wird davor zurückschrecken, in der Nachbarschaft eines Kohlelagebaus ein kapitalintensives Großprojekt zu realisieren, wenn es damit rechnen muß, daß andere sich die Erträge dieser Investition aneignen werden, sobald sie erst einmal stattgefunden hat und die Kosten (zumindest zum Teil) irreversibel geworden sind. Ebenso dürfte ein Arbeitnehmer zögern, in betriebsspezifisches Humankapital (Fertigkeiten und Kenntnisse, die sich anderweitig nur schlecht verwerten lassen) zu investieren, wenn er nicht die Gewähr hat, daß die Unternehmung später davor zurückschrecken wird, sich die Erträge dieser Investition anzueignen.

Dieses einfache Modell erlaubt es uns auch zu verdeutlichen, daß es auf den Grad der Spezifität und die Existenz alternativer Verwendungsmöglichkeiten (outside opportunities) ankommt. Wir werden hier von letzteren als den "alternativen Käufern" sprechen. Stellen wir uns vor, daß es sowohl ex ante als auch ex post eine große Anzahl solcher Käufer gibt und daß sie bereit sind, für das Gut den Preis v zu zahlen. Es soll aber immer noch ein bestimmtes Maß an Spezifität der Investition verbleiben, d.h. es besitzt eine besondere Eignung für die Bedürfnisse eines spezifischen Käufers. Wenn der Lieferant mit einem der alternativen Käufer in Austauschbeziehungen tritt, entstehen ihm Produktionskosten in einer solchen Höhe, als hätte er nicht I , sondern λI investiert, wobei λ aus dem Intervall $[0, 1]$ entnommen ist. ($\lambda = 0$ repräsentiert den Extremfall der totalen Spezifität und $\lambda = 1$ den entgegengesetzten Extremfall totaler Abwesenheit von Spezifität.) Angenommen der Lieferant hat I investiert. Wenn er nicht mit dem spezifischen Käufer handelt, erhält er den Preis v (dies ist der Preis, zu dem die "Alternativen" - es handelt sich um miteinander im

³² Eine Anwendung dieses Gedankengangs auf Arbeitsmärkte findet sich bei Grout (1984).

Wettbewerb stehende Unternehmen - zu kaufen bereit sind) und erzielt den Überschuß $v - c(\lambda I)$. Vorausgesetzt, daß die Verhandlungen zwischen dem Anbieter und seinem spezifischen Käufer erneut zu einer Nash-Verhandlungslösung führen, kommt es zur Vereinbarung eines Preises p , der durch die folgende Gleichung bestimmt wird:

$$v - p = [p - c(I)] - [v - c(\lambda I)].$$

Man erkennt sofort, daß der Anbieter einen stärkeren Anreiz hat zu investieren, als wenn es keine alternativen Käufer gäbe. Die Existenz dieser alternativen Verwendungsmöglichkeiten stärkt seine Verhandlungsposition, sie erhöht den Erlös, den er im Ergebnis der Verhandlungen erzielen kann, d.h. auch ohne gültigen Ex-ante-Vertrag. Der Anbieter erzielt dann bei Vornahme der Investition den Gewinn

$$v - \frac{1}{2}[c(I) + c(\lambda I)] - I.$$

Das für ihn ex ante optimale Investitionsvolumen wird dann durch die Gleichung

$$-[c'(I) + \lambda c'(\lambda I)] = 2$$

bestimmt. Ist $\lambda = 1$ (keine Spezifität), dann ist $p = v$, und das Investitionsvolumen gleich dem, das die soziale Wohlfahrt maximiert. Wenn $\lambda = 0$ (totale Spezifität der Investition), ist das Investitionsvolumen das gleiche wie im oben dargestellten Fall ohne alternative Verwendungsmöglichkeiten. Liegt λ im Intervall $(0, 1)$, ist die Investition eine zunehmende Funktion von λ (außer bei extremen Krümmungseigenschaften der Kostenfunktion).

Verträge

Wir nehmen nun an, daß die zwei Parteien ex ante Verträge abschließen können, die festlegen, wie ex post das Transaktionsvolumen und der Preis bestimmt werden. Wir wollen voraussetzen, daß spezifische Investitionen für beide Parteien beobachtbar sind, daß ihr Volumen aber nicht Gegenstand eines gerichtlich durchsetzbaren Vertrags sein kann. Das Gericht kann nämlich nicht nachprüfen, ob tatsächlich eine Investition dieses Umfangs stattgefunden hat (wir sagen auch, sie ist nicht belegbar oder verifizierbar). Es ergibt infolgedessen keinen Sinn, das Investitionsvolumen in den Vertrag hineinzuschreiben. Wären die Investitionen verifizierbar, so könnten sie natürlich ex ante bindend vereinbart und die Erfüllung des Vertrages ex post auch durchgesetzt werden. Dann wäre das Problem verschwunden, für das wir uns hier interessieren. (In diesem Zusammenhang mag es interessieren, daß General Motors oder auch andere Hersteller manchmal die Werkzeugmaschinen beschaffen, die ihre Lieferanten benötigen, und sie ihren Lieferanten dann beispielsweise

vermieten.) Auch von dem Wert, den das Gut für den Nachfrager besitzt, und den Kosten werden wir annehmen, daß sie an sich beobachtbar, die Beobachtungswerte aber nicht verifizierbar sind. Das heißt, sie sind den beteiligten Parteien bekannt, aber vor Gericht taugen sie nicht als Beweismittel.

Da die Drohung mit dem Abbruch der Geschäftsbeziehung nicht selten den Hebel darstellt, den der Geschäftspartner ansetzt, um sich in Verhandlungen einen mehr oder großen Teil des durch die Beziehung geschaffenen Mehrwerts zu sichern, mag es wünschenswert sein, ex ante Pönalen zu vereinbaren, die fällig werden, wenn einer der beiden nachträglich versucht, aus dem Geschäft wieder auszusteigen. Derartige Pönalen stabilisieren die Beziehung und verhindern opportunistisches Verhalten. Eine extreme Form einer solchen Pönale stellt ein Vertrag dar, bei dem die Parteien sowohl das Transaktionsvolumen als auch den Preis im voraus vollständig fixieren und sich verpflichten, diese Zusagen unter allen Umständen einzuhalten (damit dies "unter allen Umständen" durchsetzbar ist, muß die vereinbarte Pönale gewissermaßen unendlich groß sein). Da der Preis fest vereinbart ist, besteht dann keine Möglichkeit mehr, den Geschäftspartner um die Erträge aus spezifischen Investitionen zu bringen (oder in der Sprache der Mathematik: p ist nicht mehr von I abhängig).

Die Höhe der Pönalen, die beim Abbruch einer Beziehung fällig werden, stellt in einem gewissen Sinn ein Kriterium zur Beurteilung ihrer Langfristigkeit dar. Sie stabilisieren die Beziehung. Einer ihrer Nachteile ist natürlich, daß sie die Parteien zur Abwicklung von Transaktionen zwingen, auch wenn diese keinen Ertrag stiften (d.h. wenn sich plötzlich herausstellt, daß der Wert des Gutes für den Nachfrager gering oder die Kosten zu hoch sind, oder aber, daß sich bessere Alternativen auftun). Ist die Wahrscheinlichkeit, daß dies geschieht, erheblich, dann bedarf es flexiblerer Arrangements, die zumeist aber auch anfälliger für "Mißbräuche" (opportunistisches Verhalten) sind. Immerhin gibt es auch einige einfache Fälle. Stellen wir uns vor, daß in der ersten Periode die künftigen v und c noch Zufallsgrößen darstellen (und $v < c$ mit positiver Wahrscheinlichkeit eintreten kann), und erst zu Beginn der zweiten Periode zum gemeinsamen Wissen der beiden Parteien werden. Nehmen wir weiter an, daß nur eine der beiden Parteien - sagen wir, der Lieferant - eine Investition vornimmt (c ist eine stochastische Funktion dieser Investition I). In diesem Fall ist es optimal, wenn der Lieferant den Preis festsetzt und der Käufer dann nur mehr entscheiden kann, wieviel er bestellen will (dieses Arrangement gewährleistet das Paretoeffiziente Investitions- und Transaktionsvolumen). Dies ist deswegen der Fall, weil hier ex post keine Informationsasymmetrien vorliegen. Der Anbieter verlangt den Preis v , wenn $v \geq c$. Ist $v < c$, verlangt er irgendeinen Preis, der größer als v ist. Die Transaktion kommt nur zustande, wenn $v \geq c$. Der gesamte ex post Ertrag der Transaktion fließt dem Anbieter

zu. Da das Transaktionsvolumen optimal ist und der Preis nicht von der Investition abhängt, wählt der Lieferant das gesellschaftlich optimale Investitionsvolumen³³. Ist v bereits zum Zeitpunkt 1 bekannt, läuft dies darauf hinaus, den Preis v bereits zu diesem Zeitpunkt festzusetzen und es dem Anbieter zu überlassen, wieviel er in Periode 2 liefern will. Als Faustregel können wir festhalten, daß diejenige Partei den Preis festsetzen sollte, die die Investition vornimmt. Das gilt aber nicht, wenn bereits im voraus bekannt ist, welchen Wert die relevante Variable (Wert oder Kosten) des Geschäftspartners annehmen wird - dann sollte die Partei, die die Investition vornimmt, über das Transaktionsvolumen entscheiden.³⁴ (Hierbei kann die gewünschte Aufteilung der Erträge ex ante mit Hilfe eines Pauschaltransfers gewährleistet werden.)

*Übungsaufgabe I*** Im Text wird vorausgesetzt, daß eine Investition des Lieferanten (Unternehmens) seine Produktionskosten vermindert. Unterstellen Sie stattdessen, daß die Ex-ante-Investition die Qualität des Produktes und somit seinen Wert für den Käufer erhöht. Die Ex-post-Zahlungsbereitschaft des Käufers beträgt $v(I) = 3I - \frac{1}{2}I^2$. Folglich beläuft sich die Rente des Nachfragers, wenn er das Gut erwirbt, auf $v(I) - p$. Der Gewinn des Anbieters beträgt dann $p - c - I$ ($c < \frac{1}{2}$ seien seine konstanten Produktionskosten). Wir nehmen an, daß der Käufer I (und folglich v) beobachten kann, den Beobachtungswert aber nicht vor Gericht verifizieren kann, so daß es nicht möglich ist, einen bestimmten Wert von I vertraglich zu vereinbaren.

- (i) Bestimmen Sie die effiziente Investitionssumme!
- (ii) Angenommen es gibt keinen Vertrag, und die zwei Parteien verhandeln ex post. Dabei stellt sich die Nash-Verhandlungslösung ein. Ist das Investitionsvolumen optimal? Zeigen Sie den externen Effekt auf.
- (iii) Angenommen die Parteien unterschreiben einen Vertrag, der vorsieht, daß der Käufer das Recht besitzt, das Produkt zu einem bestimmten Preis p zu kaufen. Ist

³³ $F(v)$, $f(v)$, $G(c|I)$ und $g(c|I)$ sollen die Verteilungs- und Dichtefunktionen für v und c bezeichnen (wobei $\partial G/\partial I > 0$). Der Anbieter löst das Entscheidungsproblem

$$\max_I \left(\int_{\{v \geq c\}} [p(v, c) - c] f(v) g(c|I) dv dc - I \right),$$

das mit dem eines wohlwollenden Diktators identisch ist, weil $p(v, c) = c$, wenn $v \geq c$.

³⁴ Nehmen beide Parteien jeweils eine Investition vor oder nimmt der Anbieter bei asymmetrischer Information über v eine Investition vor, so ist ein noch komplizierterer Mechanismus erforderlich. Der in Fußnote 29 beschriebene Mechanismus mit Zahlungen, die auf den Erwartungswert eines externen Effekts abstellen, sorgt selbst dann für das effiziente Investitions- (und Transaktions-)volumen, wenn beide Parteien investieren und ex post bilaterale asymmetrische Information vorliegt.

dieser Vertrag effizient? Wie verhält es sich, wenn der Lieferant das Recht hat, zu einem gegebenen Preis zu verkaufen?

- (iv) Was geschieht, wenn der Anbieter das Recht erhält, *ex post* über den Preis zu entscheiden?

Bemerkung: Unsere Untersuchung legt den Schluß nahe, daß die Werte, die v und c annehmen, wann immer möglich, "geprüft" werden sollten, beispielsweise von Wirtschaftsprüfern. Lassen sie sich zweifelsfrei ermitteln und überprüfen, so stellen sie in den Verhandlungen gemeinsames Wissen dar (oder genauer: die Situation, in der Verhandlungen geführt werden, ist dann äquivalent zu einer, in der sie gemeinsames Wissen darstellen). Den Ineffizienzen, zu denen unvollkommene Information führen kann, ist damit vorgebeugt. Wird *ex ante* ein Vertrag geschlossen, dann ermöglicht es eine solche Überprüfung, die Fragen der Aufteilung des Ertrags und der Entscheidung über das Transaktionsvolumen voneinander zu trennen, und letzteres davon abhängig zu machen, welche Ergebnisse die Prüfung erbringt. Immer dann, wenn es ohne eine solche Prüfung zu Ineffizienzen käme, ermöglichen die zusätzlichen Informationen, die sie zu Tage fördert, eine Steigerung der Effizienz³⁵. Williamson (1975) argumentierte, daß in einem vertikal integrierten Unternehmen der Vornahme einer solchen Prüfung weniger Hindernisse entgegenstehen, als wenn sich zwei nicht miteinander verflochtene Unternehmen gegenüberstehen³⁶. Insbesondere vertrat er die Meinung, eine interne Revision, also eine Revision bei einer Abteilung des eigenen Unternehmens, stoße auf weniger Schwierigkeiten als eine externe, die erforderlich sei, wenn die Vorleistung von einem selbständigen Unternehmen bereitgestellt werde. Ein externer Prüfer werde mißtrauisch beäugt, die Bereitschaft der Mitarbeiter, ihm Informationen zu vermitteln, sei eher gering, sie kolludieren gegen ihn³⁷. Ein weiteres Argument, das diesem hinzugefügt werden kann: ein Unternehmen hat das Recht, seine eigenen Bestandteile zu prüfen, wann immer es der Unternehmensleitung

³⁵ Oft lassen sich natürlich die Werte von v und c nachträglich - *ex post* - feststellen, d.h. nachdem das Geschäft abgeschlossen ist. Zum Zeitpunkt des Einstiegs in das Geschäft handelt es sich hingegen oft um recht subjektive Schätzwerte. Eine *ex post* Überprüfung kann im Vertrag vorgesehen sein. Die Parteien können sich verpflichten, zu Beginn der Geschäftsbeziehung ihre Schätzwerte anzugeben und eine Konventionalstrafe zu zahlen, wenn diese sich später als falsch herausstellen. (Bereitet die Ermittlung einigermaßen verlässlicher Schätzwerte für die Zahlungsbereitschaft oder die Kosten große Schwierigkeiten, so empfiehlt sich eine solche Vorgehensweise nur, wenn die Parteien nicht allzu risikoscheu sind.)

³⁶ Wegen eines Modells vertikaler Integration, bei dem unterstellt wird, daß die Integration eine solche Prüfung erleichtert, siehe Arrow (1975).

³⁷ Damit soll nicht behauptet werden, daß beim Controlling innerhalb einer Unternehmung nicht auch Probleme mit Kollusion zwischen den Mitarbeitern auftreten. Vergleiche z.B. Dalton (1959, S. 206).

beliebt; gegenüber einem selbständigen Lieferanten besitzt es dieses Recht nicht (eine Revision ist nur unter speziellen Umständen möglich). Grossman und Hart (1986) argumentieren demgegenüber, daß Integration als solche schwerlich die Informationsstruktur verändern dürfte. Sie führen an, zwei selbständige Unternehmen könnten stets einen Vertrag über wechselseitige Informationspflichten schließen, der zum Entstehen einer Informationsstruktur führe, die genau identisch mit der des vertikal integrierten Unternehmens sei. Insbesondere könne sich jede der beiden Parteien dazu verpflichten, eine Prüfung durch den Geschäftspartner zu dulden, wann immer dieser eine solche vorzunehmen beliebe³⁸.

Die Grenzen langfristiger Beziehungen

Langfristigen Beziehungen sind offenbar insbesondere deswegen Grenzen gesetzt, weil sich Alternativen auftun. Wie oben erwähnt, kann es den Parteien schaden, wenn sie sich durch hohe Vertragsstrafen aneinander ketten. Dies ist dann der Fall, wenn sich herausstellt, daß die Geschäftsbeziehung keinen Ertrag abwirft, oder bessere Alternativen auftauchen. Dann wäre ein Abbruch der Beziehung wünschenswert. Bei der Formulierung des Vertrags sollte versucht werden, die optimale Mischung zwischen Flexibilität und Vorbeugung gegen Opportunismus zu finden³⁹.

Die Gefahr, günstige Gelegenheiten zu Geschäften mit anderen nicht wahrnehmen zu können, ist nicht der einzige Makel langfristiger Verträge. Das Bestehen einer langfristigen Beziehung fördert Kollusion zwischen den Mitarbeitern der Unternehmen, die an der Beziehung beteiligt sind (Tirole 1986b). Ihnen bieten sich genügend Gelegenheiten, Freundlichkeiten mit Freundlichkeiten zu vergelten, sie können Vertrauen ineinander fassen und die

³⁸ Diese Argumentation unterstellt, daß es möglich ist, einen Vertrag abzuschließen, der eine Prüfung speziell derjenigen Sachverhalte vorsieht, die ein bestimmtes Geschäft betreffen - im Unterschied zu einer Prüfung des ganzen Unternehmens; letzteres zu dulden wird ein Geschäftspartner nämlich nicht ohne weiteres bereit sein. Wenn eine solche Abgrenzung *ex ante* aber nicht ohne weiteres oder nur mit erheblichem zusätzlichem Aufwand möglich ist, erleichtert Integration die Vornahme einer Prüfung. Williamson (1975, S. 146) greift auf ein ähnliches Argument zurück, um zu zeigen, daß die Führung eines Unternehmens die Leistung seiner einzelner Organisationseinheiten (er unterstellt Divisionalisierung) besser beurteilen kann als der Kapitalmarkt. Der Zusammenhang zwischen unvollständigen Verträgen und Entscheidungsbefugnissen wird im Unterabschnitt 1.4 noch näher dargelegt. Eine eingehendere Erörterung der Frage, warum die Verteilung von Entscheidungsbefugnissen Bedeutung für die Informationsstruktur besitzt, findet der Leser bei Holmström und Tirole (1987).

³⁹ Ein suchtheoretisches Modell, bei dem es um die Suche nach vorteilhaften alternativen Verwendungsrichtungen geht, findet sich bei Harris und Holmström (1983).

Überzeugung gewinnen, daß die Kollusion von Dauer ist. Pettigrew (1972) stellt dar, wie sich die Manager eines Unternehmens im Laufe der Zeit mit einem bestimmten Zulieferer identifizieren können⁴⁰. Da eine solche Kollusion Ineffizienzen hervorbringen kann, empfiehlt es sich, die Mitarbeiter, die in einer Abteilung tätig sind, mit wechselnden Aufgaben zu betrauen oder, wenn dies zu kostspielig ist (sagen wir wegen des arbeitsplatzspezifischen Humankapitals), den Zulieferer oder den Abnehmer von Zeit zu Zeit auszuwechseln.

Eine andere Schwierigkeit mit langfristigen Beziehungen kann darauf zurückzuführen sein, daß eine kurzfristige Beziehung für eine Partei im allgemeinen die vorteilhaftere ist, wenn sie weiß, daß sich ihr in Zukunft gute Alternativen (outside opportunities) bieten werden. Inwieweit sich jemandem aber solche guten anderen Möglichkeiten eröffnen, hängt nicht zuletzt davon ab, wie tüchtig und erfolgreich er ist. Um eine stärkere Verhandlungsposition zu gewinnen, mag jemand also ein Signal aussenden, aus dem der Empfänger schließen kann, daß der Sender vermutlich in diese Kategorie der Tüchtigen und Erfolgreichen gehört, und als ein solches Signal eignet sich der Abschluß eines kurzfristigen Vertrages⁴¹.

Eine Anwendung: Periodische Versteigerung einer Konzession

An diesem Punkt wollen wir auf Williamsons Warnung vor den Gefahren zurückkommen, die entstehen, wenn man ein natürliches Monopol aufgrund einer Versteigerung befristet an einen Privaten vergibt, um es nach Ablauf dieser Frist erneut zu versteigern⁴². Er weist darauf hin, daß die spezifischen Investitionen des Unternehmens, das vor der neuerlichen Versteigerung das natürliche Monopol innehatte (wir wollen es kurz *Altsasse* nennen),

⁴⁰ Ganz ähnliche Probleme treten beispielsweise im Verteidigungsministerium auf - manche Beamte dieses Ministeriums entwickeln eine Zuneigung zu bestimmten Lieferanten, mit denen sie schon zahlreiche Geschäfte abgeschlossen haben. Unternehmensberater büßen die Fähigkeit zur Objektivität ein, wenn sie ein bestimmtes Unternehmen zu lange beraten. Wirtschaftsprüfungsunternehmen lassen aus demselben Grund einen bestimmten Klienten nicht jedes Jahr vom selben Mitarbeiter prüfen.

⁴¹ Hermalin (1986) hebt dies bei seiner Betrachtung von Arbeitsmärkten besonders hervor. Er untersucht, welche Konsequenzen aus einem solchen Signalisierungsverhalten für die Ausgestaltung von Weiterbildungsprogrammen gezogen werden können. Vergleiche hierzu ferner die Arbeit von Aghion und Bolton (1987); sie befassen sich mit der Abwehr neuer Anbieter. Sie betrachten den Fall eines Monopolisten, der die Botschaft vermitteln möchte, daß wegen der Eigentümlichkeiten der von ihm verwendeten Technologie ein Marktzutritt unwahrscheinlich ist. Er signalisiert dies, indem der Liefervertrag für den Fall, daß der Kunde sein Produkt nicht abnimmt, nur eine geringe Pönale vorsieht. Ein kurzfristiger Vertrag ist aber nichts anderes als ein langfristiger ohne Pönale.

⁴² Joskow und Schmalensee (1983) analysieren dies anhand eines Beispiels aus der Elektrizitätserzeugung.

meßbar und die Meßergebnisse gerichtlich nachprüfbar sein müssen. Sonst ist es nicht möglich, die Kapitalgüter, die durch diese spezifischen Investitionen geschaffen worden sind, an das Unternehmen zu überführen, das sich bei der Neuvergabe die Konzession ersteigert, und den Altsassen dafür angemessen zu entschädigen. Ohne Aussicht auf eine solche Entschädigung hat der Altsasse aber keinen Grund, die spezifischen Investitionen vorzunehmen. Diese Voraussetzung ist unglücklicherweise häufig nicht erfüllt. Für eine Bewertung dieser Kapitalgüter wären umfangreiche Aufzeichnungen erforderlich, der Werteverzehr durch Abnutzung müßte präzise erfaßt worden sein, die Anschaffungs- und Herstellungskosten dürften nicht überhöht ausgewiesen sein. Noch schwieriger wird es, wenn der Altsasse Investitionen in das Humankapital der Belegschaft vorgenommen hat, deren Wert nun gemessen werden müßte. Wie hoch die Entschädigung ausfallen muß, läßt sich nicht einfach aus der Buchhaltung des Unternehmens ableiten. Im Prinzip ließe sich das Problem lösen, indem das Vermögen des Unternehmens versteigert wird⁴³. Je mehr Investitionen in dieses Unternehmen geflossen sind, umso höher sollte das Gebot ausfallen. Aber auch ein solches Vorgehen stößt auf Probleme. Es kann sein, daß abgesehen vielleicht von dem einen, der die Konzession ersteigert hat, alle anderen Teilnehmer an einer solchen Auktion außerstande sind, den Wert des zur Versteigerung anstehenden Vermögens zu beurteilen. Oder es fehlt ihnen an adäquaten Verwendungsmöglichkeiten für derartige Kapitalgüter, wobei mit adäquat gemeint ist, daß der Erwartungswert der Erträge, die das jeweilige Kapitalgut in dieser Verwendung stiftet, derselbe ist, wie er es für den Altsassen gewesen wäre, hätte dieser erneut die Konzession erhalten. Dies mag insbesondere bei Investitionen in Humankapital eine Rolle spielen⁴⁴. Wenn es sich um ein natürliches Monopol handelt, das sich durch hohe spezifische Investitionen auszeichnet (dies ist beispielsweise in der Elektrizitäts- oder Gasversorgung, beim Telefon und beim Kabelfernsehen der Fall), so stellt eine periodisch wiederkehrende Versteigerung der Konzession folglich wohl ein unverhältnismäßig kostspieliges Allokationsverfahren dar. Eher zu

⁴³ Zu unterscheiden ist hier also zwischen einer Versteigerung des Vermögens des Unternehmens (der Erlös aus dieser Versteigerung fließt dem Altsassen zu) und der Versteigerung der Monopolstellung (der Erlös hieraus fließt der Regulierungsbehörde zu).

⁴⁴ Hinzu kommen noch Kosten, die mit dem Sperrklinkeneffekts zusammenhängen, der vor der Versteigerung auftritt. Der Altsasse hat wenig Veranlassung zu effizienter Produktion, solange nicht klar ist, wer bei der jeweils bevorstehenden Versteigerung der Monopolstellung den Zuschlag erhält. Effiziente Produktion würde den Wettbewerbern nämlich Informationen über die Technologie vermitteln, die sie veranlassen, ihr Gebot zu erhöhen. Sie würde auch die Gefahr heraufbeschwören, daß die Regulierungsbehörde das Mindestgebot heraufsetzt. Der Sperrklinkeneffekt wird auch von Freixas et al. (1985) sowie von Laffont und Tirole (1988) analysiert. Eine Anwendung auf den Arbeitsmarkt findet sich bei Gibbons (1987).

empfehlen ist sie im Fall von Wirtschaftszweigen, in denen nur ein kleiner Teil der Investitionen irreversibel ist (ein Beispiel aus dem Luftverkehr, für das diese Bedingung erfüllt sein könnte, ist eine Konzession, die ihrem Besitzer das Monopol auf die Bedienung einer bestimmten Strecke sichert; stattdessen könnte man auch an das hypothetische Beispiel von Demsetz denken, in dem mittels Universalmaschinen Nummernschilder für Autos mit zunehmenden Skalenerträgen erzeugt werden)⁴⁵.

Empirische Untersuchungen

Williamsons Theorie der langfristigen Beziehungen läßt vermuten, daß Unternehmen, wenn dies möglich und nicht übermäßig aufwendig ist, umfangreiche und detaillierte Vertragswerke ausarbeiten, und daß der Anreiz, dies zu tun, umso stärker ist, je spezifischer die Investition ist und je weniger Möglichkeiten zum Wechsel des Geschäftspartners ex post verbleiben. Joskow (1985, 1987) hat Verträge zwischen Kohlebergwerken und Elektrizitätswerken in den Vereinigten Staaten daraufhin untersucht. (Solche Beziehungen beruhen dort meist auf Verträgen, vertikale Integration ist selten.) Als spezifische Investitionen, bei denen dem Investor die Gefahr droht, übervorteilt zu werden, lassen sich im Fall der Kohlebergwerke (Anbieter) die Verbesserung der Fördertechnik, im Fall der Elektrizitätsunternehmen (Nachfrager) Kraftwerksanlagen und Dampfkessel benennen, vor allem wenn sie dem speziellen Typ der Kohle angepaßt sind⁴⁶.

Es lassen sich zwei Fälle unterscheiden, die in zwei verschiedenen Regionen auftreten (natürlich liegen die Dinge in der Praxis nicht immer so eindeutig). In der einen Region gibt es eine große Zahl von Elektrizitätsunternehmen und Kohlebergwerken sowie zahlreiche Transportmöglichkeiten für Kohle (Eisenbahn, Lastkähne), so daß auf dem Markt für Verkehrsleistungen Wettbewerb herrscht. Jedem Elektrizitätswerk und jedem Kohlebergwerk

⁴⁵ Laffont und Tirole (1987) enthält eine Analyse einer Versteigerung für den Fall nicht beobachtbarer übertragbarer und nicht übertragbarer Investitionen.

⁴⁶ Dampfkessel, die für eine bestimmte Art von Kohle ausgelegt sind, besitzen einen höheren Wirkungsgrad als solche, mit denen man verschiedenerelei Kohle verheizen kann.

bleibt infolgedessen auch ex post noch eine große Auswahl möglicher Partner. Außerdem ist die Kohle, die aus den verschiedenen Bergwerken kommt, ziemlich homogen. Eine spezielle Konstruktion der Dampfkessel, um sie den Eigenschaften der jeweiligen Kohlesorte anzupassen, ist folglich nicht erforderlich. In einer solchen Gegend ist die Gefahr opportunistischen Verhaltens gering. Spot-Märkte (kurzfristige Verträge) sind daher relativ effizient. In der anderen Region ist die Zahl der Kohlebergwerke gering, die Transportmöglichkeiten sind begrenzt und die Qualität der Kohle sehr unterschiedlich. Man würde dann erwarten, daß es entweder zu vertikaler Integration kommt oder komplizierte, auf langfristige Zusammenarbeit abzielende Vertragswerke ausgearbeitet werden.

Joskow zeigt, daß der Vergleich zwischen diesen beiden Regionen in den USA die Williamsonsche Theorie hervorragend bestätigt. Im Osten handelt es sich vorwiegend um untertägigen Abbau (60 Prozent der Produktion), bei dem zunehmende Skalenerträge keine größere Rolle spielen; infolgedessen gibt es viele kleine Bergwerke. Transportmöglichkeiten gibt es zuhauf, die Qualität der Kohle ist relativ homogen. Im Westen, wo Tagebau vorherrscht, sind zunehmende Skalenerträge typisch, es gibt dort nur einige wenige, aber große Tagebaue. Auf dem Verkehrsmarkt herrscht auch weniger Wettbewerb als im Osten, und die Qualität der Kohle ist sehr unterschiedlich. Joskow zeigt auch, daß die Verträge im Westen tatsächlich eine viel längere Laufzeit aufweisen als im Osten (und daß der Spotmarkt für Kohle im Osten eine erhebliche Rolle spielt, wohingegen es einen solchen Markt im Westen im Grunde genommen nicht gibt). Der Leser möge beachten, daß in diesem Beispiel das Fehlen von "outside opportunities" mit der Spezifität der Kapitalgüter zusammenhängt.

Die Williamsonsche Theorie läßt sich auch am Beispiel von Kraftwerken überprüfen, die unmittelbar neben einer Kohlengrube domiziliert sind. Hat sich neben einer solchen Kohlengrube noch kein anderes Kraftwerk niedergelassen und sind die Transportkosten hoch, so kann es für ein Elektrizitätsunternehmen von Vorteil sein, diesen Standort zu wählen, und dies ist offenbar eine spezifische Investition. Joskow stellte fest, daß solche Kraftwerke typischerweise durch langfristige Verträge mit der Kohlengrube verbunden oder mit ihr vertikal integriert sind; es handelt sich normalerweise um Verträge mit einer Laufzeit von 20 - 50 Jahren, die Kohlepreise für einen Zeitraum von 20 Jahren festlegen, eine detaillierte Beschreibung der Mengen enthalten, die über diesen Zeitraum hinweg zu liefern sind, die Qualität der Kohle spezifizieren und auf Indizes für die Kosten und die Preise der Substitute Bezug nehmen. (Sie sehen zumeist für den Fall eines Streits auch ein Schiedsgericht vor.) Die Vereinbarung solcher Konditionen ex ante mag verhindern, daß die Erträge spezifischer Investitionen einem anderen als dem Investor zufließen.

1.4 Das Unternehmen als ein unvollständiger Vertrag⁴⁷

In den Unterabschnitten 1.2 und 1.3 betrachteten wir Organisationen als effiziente kurz- und langfristige Verträge. In der Realität aber sind Verträge ziemlich unvollständig; der Grund dafür sind die "Transaktionskosten"⁴⁸. Coase (1937) und Williamson (1975) haben zwischen vier Typen von Transaktionskosten unterschieden. Zwei von ihnen fallen zum Zeitpunkt des Vertragsabschlusses an und die zwei anderen erst danach. Erstens kann es geschehen, daß während der Laufzeit des Vertrages Zufälle auftreten, die zum Zeitpunkt des Vertragsabschlusses nicht absehbar waren. Zweitens, selbst wenn sie vorausgesehen werden könnten, gäbe es zu viele denkbare Zufälle, als daß man sie im Vertragstext explizit berücksichtigen könnte⁴⁹. Drittens kann es ein aufwendiges Unterfangen sein, die Erfüllung des Vertrages durch den jeweils anderen Vertragspartner zu kontrollieren. Viertens kann es mit erheblichen Kosten verbunden sein, mit Hilfe der Gerichte die Einhaltung vertraglicher Verpflichtungen zu erzwingen, die der Partner eingegangen ist. Coase und Williamson sind der Meinung, daß die Minimierung der Transaktionskosten eines der wichtigsten Effizienzkriterien bei der Gestaltung von Organisationen darstellt.

Die zwei erstgenannten Gründe lassen sich nur schwer formalisieren. Wir besitzen keine ausgereiften Theorien über Komplexität und über Entscheidungsprobleme mit unvorhersehbaren Zufällen - beides spielt in der Wirtschaft aber eine bedeutsame Rolle. Reale Verträge verzichten meist darauf, bestimmte Zufälle zu berücksichtigen. Treten diese Zufälle dann aber doch ein, kommt es zwischen den beiden Parteien wahrscheinlich zum Konflikt. Wir können zwischen zwei Typen von ex post Vorgehensweisen beim Eintritt solcher unvorhergesehenen Ereignisse unterscheiden.

Die einfachste Vorgehensweise, nämlich Bargaining, wurde im Unterabschnitt 1.3 untersucht. Dort nahmen wir zuerst an, daß ex ante gar kein Vertrag unterzeichnet wurde und daß die Parteien über die Transaktion und die Aufteilung der Erträge verhandeln, nach-

⁴⁷ Der Leser, der noch mehr über unvollständige Verträge wissen will, mag Holmström und Tirole (1987) konsultieren.

⁴⁸ Ein "vollständiger Vertrag" ist einer, bei dem alle relevanten Variablen (Zahlungen, Geschäftsvolumen usw.) von verifizierbaren Größen abhängen - letztere mögen öffentliche Äußerungen der vertragsschließenden Parteien einschließen (z.B. über ihre Zahlungsbereitschaft, Kosten usw.). Schließlich mag man, obwohl dies nicht ganz korrekt ist, auch diejenigen Verträge als "vollständig" bezeichnen, bei denen die Vertragsparteien dieselbe Auszahlung erhalten wie bei einem optimalen und im eigentlichen Sinne vollständigen Vertrag.

⁴⁹ Die Unterscheidung zwischen diesen zwei Fällen mag "in der Praxis" Schwierigkeiten bereiten. Daß in Arbeitsverträgen mit Schreibkräften nichts über den Einsatz von Textverarbeitungssystemen festgelegt war, solange es diese nicht gab, lag möglicherweise nicht daran, daß niemand sich derartige Systeme vorstellen konnte, sondern daran, daß ein Einschluß aller möglichen denkbaren Arten von Textverarbeitung in solche Verträge zeitraubend und kostspielig gewesen wäre.

dem die Investition bereits vorgenommen worden ist und nachdem sie zusätzliche Informationen über Werte und Kosten erhalten haben. Weil es keinen ex ante Vertrag gab, stand dieses Bargaining auch unter keinerlei vertragsbedingten Restriktionen. Nur das Prinzip der Vertragsfreiheit begrenzte die Spielräume der beiden Parteien und zwar in dem Sinne, daß für beide der Abschluß eines Vertrags eine freie Entscheidung darstellte. Wir verglichen das Ergebnis, zu dem das Bargaining führte, mit demjenigen, das unter einem vollständigen ex ante Vertrag zustandegekommen wäre.

Man kann sich aber auch Mischformen denken. Damit meinen wir die Vereinbarungen, die ex post ausgehandelt werden, wenn ex ante ein Vertrag abgeschlossen wird, der zwar unvollständig ist, aber doch die ex post verbleibenden Verhandlungsspielräume eingrenzt. Bei diesen Mischformen könnte es, verglichen mit vollständigen Verträgen, zu einer Verringerung der Transaktionskosten kommen, aber ohne die Nachteile, mit denen der Verzicht auf solche Eingrenzungen verbunden ist. Wir unterscheiden zwei Alternativen. Die erste besteht darin, daß die beiden Parteien auf eine dritte zurückgreifen. Die dritte Partei hat die Aufgabe, zumindest annähernd die effizienten Entscheidungen zu treffen, die ein vollständiger Vertrag vorgesehen hätte. Ihre ex post Entscheidungen über das Transaktionsvolumen und die Aufteilung des Ertrags müssen also ex post ein effizientes Transaktionsvolumen sichern und damit den Anreiz schaffen, daß ex ante im richtigen, d.h. effizienten Umfang spezifische Investitionen vorgenommen werden. Bei der zweiten Alternative besitzt eine der beiden vertragsschließenden Parteien, nicht eine dritte, das Recht zu bestimmen, was in einem unvorhergesehenen Fall geschehen soll.

Schiedsgerichte

Der Rückgriff auf einen Dritten kann die Form annehmen, daß ein Schiedsgericht vereinbart wird. Eine Gewerkschaft und ein Unternehmen können sich zum Beispiel verpflichten, einen Schlichter einzuschalten, wenn sie sich nicht auf einen neuen Tarifvertrag einigen können. Ebenso können ein Lieferant und sein Abnehmer vereinbaren, im Konfliktfall einen Schiedsspruch eines unabhängigen Experten zu akzeptieren⁵⁰.

⁵⁰ Außerdem kann man noch die Gerichte zu Hilfe rufen. Im Unterschied zum Schiedsgericht steht diese Möglichkeit auch dann zur Verfügung, wenn sie nicht vertraglich vereinbart ist. Abgesehen davon, daß sie Verträgen zur Durchsetzung verhelfen (sofern diese nicht sittenwidrig sind), beschränken sich die Gerichte allerdings vielfach darauf, zu beurteilen, ob ein bestimmtes Verhalten mit dem Grundsatz von Treu und Glauben im Einklang steht (es gibt Verhaltensweisen, die nur dann mit diesem Grundsatz im Einklang stehen, wenn sie ausdrücklich im Vertrag vorgesehen sind). Schiedsrichter können weitergehende Befugnisse besitzen (z.B. festsetzen, was in einem im Vertrag nicht vorgesehenen Fall zu geschehen hat).

Schiedsgerichte sind meist kostspielig. Außenstehende besitzen häufig nicht die erforderlichen Informationen, um die effiziente Entscheidung treffen zu können. Sie müssen auf Experten zurückgreifen oder sich die Mühe machen, die Besonderheiten der Situation zu studieren. In dieser Hinsicht sind "interne" Schlichter, sofern welche zur Verfügung stehen, wahrscheinlich effizienter. Die Rolle eines internen Schlichters bei einem Konflikt zwischen Abteilungen eines Unternehmens oder Mitarbeitern kann beispielsweise das Topmanagement dieses Unternehmens übernehmen. Dies sind Personen, die Tag für Tag mit unternehmens-internen Vorgängen und Mitarbeitern befaßt sind und auf diese Weise Einblicke gewinnen, die externe Schlichter nicht besitzen können⁵¹. Williamson (1975, S. 29) ist der Überzeugung, daß interne Schlichtungsverfahren praktisch stets überlegen sind.

Ein Schlichter sollte befähigt sein, sich ohne allzu großen Aufwand soweit einzuarbeiten, daß er die wesentlichen Probleme erkennt, ferner sollte er unabhängig sein. Die erste Bedingung kann, wie wir gesehen haben, dazu führen, daß externe Schiedsrichter nicht sehr viel weiterhelfen. Aber auch interne Schiedsrichter erfüllen diese Bedingung möglicherweise nicht ohne weiteres. In großen Unternehmen ist der Vorstand vermutlich mit solchen Schlichtungsaufgaben überlastet. Wenn das der Fall und infolgedessen anzunehmen ist, daß er nicht sehr gut informiert ist, verfehlt eine Entscheidungsbefugnis des Vorstands ihren Zweck⁵². Die zweite Bedingung, Unabhängigkeit, erfordert, daß der Schiedsrichter nicht beides sein kann - Richter und Partei. Er muß Entscheidungen treffen, die sich nach den Gesamtinteressen der zwei Parteien richten und nicht die eine zum Nachteil der anderen begünstigen. Unabhängigkeit ist beispielsweise dann nicht mehr gewährleistet, wenn das betreffende Vorstandsmitglied enge Verbindungen mit einem Unternehmensbereich besitzt. Oder allgemeiner ausgedrückt, der interne Schiedsrichter muß ebenso wie der externe vertrauenswürdig erscheinen, oder doch zumindest die Reputation aufbauen, daß er Streitigkeiten "fair" schlichtet (wobei "fair" heißt, der Schiedsrichter muß den Erwartungen der Parteien entsprechen, daß er die effizienten Entscheidungen trifft - diejenigen also, die ein vollständiger Vertrag vorgesehen hätte).⁵³

⁵¹ Gerichte und Schiedsgerichte sind sich dieses Informationsvorsprungs der Insider wohl bewußt und versuchen, ihn zur Geltung zu bringen. Beispielsweise sind die Richter, die in einem französischen tribunaux de commerce Recht sprechen, zum Teil nicht Juristen, sondern Kaufleute und Ingenieure. Ähnlich verhält es sich mit den Schiedsrichtern, die bei der International Chamber of Commerce akkreditiert sind.

Andererseits mag ein internes Schiedsgericht auf Verfahrensweisen zurückgreifen, die in externen üblich sind. Beispielsweise besitzen manche Unternehmen ein ausgeklügeltes Regelwerk für die Entscheidung über Beschwerden.

⁵² Dann mag es sein, daß die Entscheidungsbefugnis an die nachfolgende Hierarchiestufe abgegeben werden muß.

⁵³ Daß das Management die Reputation besitzt, Konflikte "gerecht" (d.h. effizient) zu lösen, ist ein wichtiger Bestandteil einer "Unternehmenskultur".

Entscheidungsbefugnis

Mit Entscheidungsbefugnis meinen wir das Recht, die Lücken im Vertragswerk zu füllen, also die Entscheidung zu treffen, wenn ein Umstand eintritt, der im Vertrag nicht vorgesehen ist. Es ist durchaus möglich, eine der beiden Parteien mit dieser Entscheidungsbefugnis auszustatten statt einen unbeteiligten Dritten. Wie Grossman und Hart (1986) und Hart und Moore (1985) betonen, bedeutet Entscheidungsbefugnis nicht, daß die Beteiligten ex post nicht verhandeln. Die Entscheidung, die die entscheidungsbefugte Partei vorzieht, mag für die andere schwere Nachteile mit sich bringen. Es könnte sich noch eine Alternative finden lassen, die zu beiderseitigem Vorteil ist, und die Partei, die die Entscheidungsbefugnis besitzt, kann sich einen Verzicht auf die Ausübung dieser Befugnis vergolden lassen. Die Grossman-Hart-Moore-Analyse zeigt besonders deutlich, daß die Funktion von Entscheidungsbefugnis darin besteht, den Status quo beim Bargaining zu verschieben - sie stärkt die Verhandlungsposition der Partei, die sie besitzt. Die Aufteilung des Gewinns, die sich ex post ergibt, hat einen Einfluß auf die ex ante Investitionen.

Um uns den Zusammenhang zwischen der Verteilung der Entscheidungskompetenz und der des Gewinns zu verdeutlichen, nehmen wir mit Grossman und Hart an, daß ex post die Parteien (Käufer und Lieferant) eine Entscheidung d aus einer Menge D treffen müssen. Die Auszahlungen betragen ex post $B_i(d)$, wobei $i = 1, 2$.⁵⁴ Der Partei 1 (Käufer oder Lieferant) die Entscheidungsbefugnis zu geben, heißt, daß sie ex post dasjenige d wählen darf, welches sie bevorzugt. Wenn die zwei Parteien sich nicht auf etwas anderes einigen können, wird Partei 1 d_1^* wählen, um $B_1(d)$ zu maximieren. Aber wenn d_1^* die Auszahlung der Partei 2 nicht maximiert, besitzen im allgemeinen beide Parteien einen Anreiz, erneut zu verhandeln, um d^* zu implementieren (dies ist die Entscheidung, die die Summe der Auszahlungen $B_1(d) + B_2(d)$ maximiert). Nehmen wir an, daß Transfers t von Partei 2 zugunsten von Partei 1 in einem solchen Umfang erfolgen, daß die zusätzlichen Gewinne, die durch eine Wiederaufnahme der Verhandlungen erzielt werden können, gleichmäßig verteilt werden (d.h. wir gehen von der Nash-Verhandlungslösung aus), so erhalten wir

$$[B_1(d^*) + t] - B_1(d_1^*) = [B_2(d^*) - t] - B_2(d_1^*).$$

⁵⁴ Die Entscheidung d mag das Zustandekommen eines Geschäfts betreffen, B_i kann einen Wert oder die Kosten bezeichnen. Die Fragestellung ist aber allgemeinerer Natur. So könnte d ebensogut eine Qualitätsstufe oder eine Entscheidung über das Design symbolisieren, oder B_i die Anstrengungen, die die Partei i unternimmt. Der Leser mag beachten, daß d ein Vektor sein kann.

Mit B_1 und B_2 bezeichnen wir die Nettoauszahlung (also die Auszahlung, die übrigbleibt, nachdem die Transfers abgezogen bzw. gutgeschrieben worden sind und das Spiel zu Ende gespielt ist). Dann gilt

$$B_1 = B_1(d_1^*) + \frac{1}{2}[B_1(d^*) + B_2(d^*) - B_1(d_1^*) - B_2(d_1^*)]$$

und

$$B_2 = B_2(d_1^*) + \frac{1}{2}[B_1(d^*) + B_2(d^*) - B_1(d_1^*) - B_2(d_1^*)].$$

Es ist klar, daß Partei 1 davon profitiert, die Entscheidungsmacht zu besitzen, gilt doch per definitionem

$$B_1(d_1^*) \geq B_1(d_2^*),$$

$$B_2(d_1^*) \leq B_2(d_2^*),$$

so daß

$$B_1(d_1^*) - B_2(d_1^*) \geq B_1(d_2^*) - B_2(d_2^*).$$

Wichtiger noch: Wenn die Ex-post-Nettoauszahlungen sowohl von der Ex-ante-Investition abhängen (wie in Unterabschnitt 1.3) als auch von Ex-post-Entscheidung, die die beiden Parteien treffen, beeinflußt die Aufteilung der Entscheidungskompetenzen die Anreize der beiden Parteien, in spezifische Kapitalgüter zu investieren. Es ist unmittelbar einleuchtend, daß im Status quo-Punkt die Partei 1 nicht um die Früchte ihrer Investitionen geprellt werden kann, weil sie die Entscheidungsbefugnis besitzt. (Partei 2 kann sehr wohl geprellt werden.) Kommt es nicht zur Wiederaufnahme der Verhandlungen, so übt die Entscheidungsbefugnis der Partei 1 also einen Einfluß auf die Investitionsanreize für beide Parteien aus. Kommt es zur Wiederaufnahme, so entfaltet der Status quo selbst dann, wenn er nicht erhalten bleibt, einen Einfluß auf die Nettoauszahlungen; auch in diesem Fall wirkt sich also die Entscheidungsbefugnis auf die Investitionsanreize aus. Will man noch mehr und genaueres wissen, muß man beschreiben, wie sich die spezifischen Investitionen und die möglichen Entscheidungen auf die Auszahlungen auswirken (siehe auch das Beispiel unten).

Grossman und Hart (1986) sprechen von Kontrolle durch den Anbieter (bzw. Nachfrager), wenn der Anbieter (der Nachfrager) die Entscheidungsbefugnis innehat. Als Integration bezeichnen sie einen Zustand, in dem eine Partei das "Residuum" an Ver-

fügungsrechten besitzt, d.h. aller jene Verfügungsrechte, die nicht ausdrücklich durch Vertrag abgetreten sind. Das Gegenteil, die Nichtintegration, läßt sich dann als ein Zustand auffassen, bei dem die Menge der möglichen Entscheidungen mindestens zweidimensional ist und jede der beiden Parteien für mindestens eine dieser Dimensionen die Entscheidungsbefugnis besitzt (im Unterschied zur Integration, bei der eine Partei alle Dimensionen in der Hand hat)⁵⁵. Optimal ist diejenige Struktur des Entscheidungsprozesses, bei der die spezifischen Investitionen am besten geschützt sind (oder die gewährleistet, daß der durch die Transaktion gestiftete Nutzen maximiert wird, wenn es schon zu einem an sich ineffizienten Bargaining kommt). Ist es nicht möglich, einen vollständigen Vertrag abzuschließen, ist das Eigentumsrecht die zweitbeste Lösung, um die Sicherheit einer Investition zu gewährleisten. Bei einem Friseurgeschäft oder einer Rechtsanwaltskanzlei beispielsweise kann es vorkommen, daß die Geschäftsleitung in t einem anderen Mitarbeiter als in der Vorperiode den Auftrag erteilt, einen bestimmten Stammkunden des Geschäfts zu bedienen; die Kunden sind Kunden des Unternehmens (oder anders ausgedrückt, das Unternehmen besitzt "property rights" an diesen Kunden), nicht seiner einzelnen Mitarbeiter. Die Rechtsordnung verbietet es den Mitarbeitern normalerweise, sich unmittelbar nach einem Ausscheiden aus dem Geschäft selbständig zu machen und seinem vormaligen Arbeitgeber die Kunden abzuwerben; sie dürfen also die Kunden, die sie bei ihrem früheren Arbeitgeber betreut haben, nicht einfach in ihre neue selbständige Existenz mitnehmen. (Manchmal wird hier ein feiner Unterschied zwischen den Stammkunden des Unternehmens gemacht und jenen, die der ausgeschiedene Mitarbeiter selbst hinzugewonnen hat; diese Unterscheidung steht durchaus im Einklang mit der hier dargelegten Theorie.) Ebenso wenig kann ein Ingenieur, der beispielsweise in einer Forschungs- und Entwicklungsabteilung eines Unternehmens angestellt war, kündigen und im Anschluß an sein Ausscheiden ein Patent für eine Erfindung anmelden, die nur deswegen möglich war, weil die Forschungs- und Entwicklungsabteilung die erforderliche Forschung unternommen hat. Grossman und Hart weisen auch auf den Umstand hin, daß Versicherungsgesellschaften es immer dann vorziehen, ihre Policen von

⁵⁵ Beispielsweise mag es sein, daß der Anbieter und der Nachfrager über alle personalwirtschaftlichen Fragen selbst entscheiden, der Nachfrager aber Entscheidungsbefugnis darüber besitzt, ob die Qualität der gelieferten Ware ausreichend ist. Oder aber ein Vorarbeiter und ein Arbeiter entscheiden jeder für sich über ihre Arbeitskleidung, aber wenn es um Fragen der Arbeitsorganisation geht, so entscheidet der Vorarbeiter. Diese Beispiele zeigen auch, daß Entscheidungsbefugnisse derart viele Dimensionen besitzen, daß Nichtintegration die Regel darstellen muß. Soll die Unterscheidung zwischen Integration und Nichtintegration nicht weitgehend inhaltsleer sein, muß man sich infolgedessen auf einen Teilbereich aller möglichen Entscheidungsbefugnisse beschränken wie z.B. auf jene Entscheidungen, die im Eigentum über bestimmte materielle Vermögensgüter begründet sind. Grossman und Hart nehmen eine solche Eingrenzung vor.

Mitarbeitern vertreiben zu lassen (von Personen, die kein Eigentumsrecht an der Liste ihrer Kunden besitzen), wenn die spezifischen Investitionen, die der "Vertreter" vornehmen muß, um die Kunden zu halten, eher gering sind (wie z.B. bei Lebensversicherungen) und von unabhängigen Versicherungsmaklern (die ein Eigentumsrecht an ihren Listen besitzen), wenn sie beträchtlich sind.

Beispiel

Ganz im Sinne der Grossman-Hart-Analyse zeigt folgendes Beispiel, wie die Aufteilung der Entscheidungskompetenzen die Verteilung der Erträge und die Investitionsanreize beeinflusst.

Ein Anbieter und ein Nachfrager vereinbaren vertraglich, morgen ein Geschäft abzuwickeln. Das Geschäft ist derart vorteilhaft, daß beide sich einig sind, daß es auf jeden Fall zustandekommen sollte. Unsicher sind nur die genauen Eigenschaften des Gutes, das geliefert werden soll. In der Periode 1 werden seine wichtigsten Eigenschaften vertraglich vereinbart, in der Periode 2 ergeben sich dann aber "ungeahnte" (das heißt nicht im Vertrag stipulierte - wir können uns das so vorstellen, daß unendliche viele Verbesserungen denkbar sind, man aber in der Periode 1 nicht ahnt, welche dieser unendlich vielen denkbaren Varianten in der Periode 2 tatsächlich realisierbar sein wird) Möglichkeiten zur Verbesserung der Qualität. Daß diese Möglichkeiten tatsächlich wahrgenommen werden, kann man im Zeitpunkt 1 nicht ohne weiteres vertraglich vereinbaren. In der Periode 2 erfahren beide Parteien von der Möglichkeit der Qualitätserhöhung. Sie tatsächlich vorzunehmen, kostet den Anbieter in den zweiten Periode $c > 0$. Der Einfachheit halber stellen wir uns vor, daß c zum Zeitpunkt 1 bekannt ist und nicht davon abhängt, um welche Verbesserung es sich handelt. Der Nachfrager trifft in der Periode 1 eine Entscheidung über eine Investition. Der Wert, den die Qualitätserhöhung für ihn in der zweiten Periode besitzt, beläuft sich mit der Wahrscheinlichkeit x auf $v > c$ und mit der Wahrscheinlichkeit $1 - x$ auf 0; die Investition I kostet $x^2/2$. Außenstehende können nicht beobachten, wieviel der Nachfrager tatsächlich investiert, dies kann also kein Gegenstand des Vertrages sein. Wir erinnern daran, daß v und c sich auf die Qualitätserhöhung beziehen, nicht auf eine Variante des Gutes, die nur mit den "wichtigsten" Eigenschaften ausgestattet ist.

Diese Annahmen über die Investition lassen sich wie folgt interpretieren: Der Käufer investiert in Flexibilität. Investiert er mehr, so erhöht sich die Wahrscheinlichkeit, daß die Qualitätserhöhung einen Nutzen stiftet. Beispielsweise kann ein Unternehmer (ein Nachfrager) seine Arbeiter so ausbilden, daß sie sich Technologieveränderungen anpassen können; dies ist der Gegenstand der Theorie der Flexibilität von Piore und Sabels (1984).

Oder ein Kraftwerk kann kostspieligere und flexiblere Dampfkessel auswählen, die so konstruiert sind, daß Kohle unterschiedlicher Qualität eingesetzt werden kann.

Analysieren wir zuerst die Pareto-optimale Allokation in diesem Modell. Offensichtlich sollte die Qualitätsverbesserung vorgenommen werden, wenn und nur wenn sie für den Nachfrager den Wert v besitzt. Die optimale Investition ist dann gegeben durch die Bedingung

$$\max_x [x(v - c) - x^2/2].$$

Folglich ist $x^* = v - c$. Die Summe der Erträge der beiden ist $W^* = (v - c)^2/2$. (Hier und im folgenden wollen wir Parameterwerte unterstellen, für die die Bedingung $x \leq 1$ nicht als Gleichung erfüllt ist.)

Es wird angenommen, daß die Parteien eigennützig handeln. Über die Qualitätserhöhung läßt sich zwar im Zeitpunkt 1 nichts bestimmtes sagen, aber in der Periode 2 kann sie Gegenstand eines Vertrags werden. Wir werden drei unterschiedliche Vorgehensweisen untersuchen: erstens Verhandlungen (ohne Einschränkung der Verhandlungsspielräume einer Partei durch das Vertragswerk der ersten Periode, d.h. in diesem steht nichts von der Qualitätserhöhung; die Parteien verhandeln dann in der zweiten Periode darüber, ob die Qualitätserhöhung vorgenommen wird; einigen sie sich nicht, wird sie nicht vorgenommen), zweitens Entscheidungsbefugnis des Nachfragers (der Nachfrager hat das Recht zu entscheiden, ob die Verbesserung vorgenommen wird) und drittens Entscheidungsbefugnis des Anbieters (der Anbieter entscheidet, ob die Verbesserung vorgenommen wird). Wir unterstellen, daß in den zwei letztgenannten Fällen die entscheidungsbefugte Partei, statt von ihrer Entscheidungsbefugnis Gebrauch zu machen, ein Verhandlungsangebot unterbreiten kann und daß auch der Verzicht auf die Entscheidungsbefugnis Bestandteil dieses Angebots sein kann (beispielsweise kann der Nachfrager anbieten, daß er darauf verzichtet, den Anbieter zur Vornahme der Verbesserung zu zwingen, wenn dieser ihm dafür ein bestimmtes Abstandsgeld bezahlt). Ferner wollen wir annehmen, daß immer, wenn es zu Verhandlungen kommt, vereinbart wird, den Ertrag aus dem Geschäft zu teilen. Außerdem unterstellen wir noch, daß die Parteien sich für diejenige Vorgehensweise entscheiden, die die Summe aus ihren erwarteten Erträgen maximiert. Diese Annahme gründet sich auf der Erwägung, daß ja stets die Möglichkeit besteht, bereits in der ersten Periode zu einer Vereinbarung über die Aufteilung des Mehrertrags zu gelangen, den der Übergang von einer weniger effizienten zu dieser optimalen Vorgehensweise stiftet.

Bei der Vorgehensweise 1 (Verhandlungen mit weiten Verhandlungsspielräumen) kommt das Geschäft dann und nur dann zustande, wenn $v > c$. Jede der beiden Parteien erhält $(v - c)/2$. Wieviel der Nachfrager investiert, ergibt sich dann aus der Lösung des Maximierungsproblems

$$\max_x \left(\frac{x(v - c)}{2} - \frac{x^2}{2} \right).$$

Hieraus folgt

$$x^B = (v - c)/2 = x^*/2.$$

Dieses Ergebnis kennen wir schon. Wir hatten es weiter oben schon einmal abgeleitet und gezeigt, daß hier der Nachfrager weniger investiert als im Pareto-Optimum (Unterinvestition). Der Ertrag, den beide zusammengenommen erzielen, ist

$$W^B = x(v - c) - x^2/2 = 3(v - c)^2/8 = 3W^*/4.$$

Entscheidungsbefugnis des Anbieters ist in diesem einfachen Beispiel mit der Vorgehensweise 1 identisch, weil in beiden Fällen der Status quo derselbe ist: Wenn die Parteien verschiedener Meinung sind, entscheidet sich der Lieferant gegen die Vornahme der Qualitätserhöhung (dies ist für ihn die billigste Alternative). Seine Entscheidungsbefugnis versetzt ihn in die Lage, nicht mehr zu tun als im ursprünglichen Vertrag vorgesehen, und die Drohung, dies zu tun, verleiht ihm eine starke Verhandlungsposition. Dieser Fall ergibt sich nicht selten beispielsweise bei Verhandlungen über die Ergänzung von Vertragswerken, wie sie sich zwischen dem Verteidigungsministerium und Herstellern von Rüstungsgütern abspielen. Die Hersteller sind nicht verpflichtet, bestimmte Veränderungen vorzunehmen, und sie lassen sich dies teuer bezahlen. Wegen der Äquivalenz zwischen der Vorgehensweise 1 und 2 gilt $x^{SC} = x^B$ und $W^{SC} = W^B$ (SC steht hierbei für Entscheidungsbefugnis des Anbieters). Der Nachfrager kann sich die Erträge seiner Investition nur zur Hälfte aneignen und investiert folglich zu wenig.

Besitzt der Nachfrager die Entscheidungsbefugnis, kommt es stets zu der Qualitätserhöhung, es sei denn, er läßt sich dieses Recht abhandeln - wir sagen dann, daß über den Status quo neu verhandelt wird. (Genauer gesagt: Ist der Wert der Verbesserung für den Nachfrager null, so ist er zwischen ihrer Vornahme und dem Verzicht darauf indifferent; hier nehmen wir aber einfach an, daß er dennoch für ihre Vornahme optiert. Zur Rechtfertigung dieser Annahme läßt sich zweierlei vortragen: Erstens würde er für die Vornahme optieren, wenn ihr Wert zwar nicht sehr groß, aber doch von null verschieden wäre.

Zweitens: Würde er im dargelegten Fall verzichten, so würde dies nur unsere Schlußfolgerungen unterstützen). Wenn der Wert v betragt, so ist der Status quo effizient und es kommt nicht zu Verhandlungen. Der Nachfrager erhalt v , wenn er sein Recht wahrnimmt, den Anbieter zur Vornahme der Qualitatserhohung zu zwingen. Betragt der Wert 0, so ist der Status quo ineffizient und es kommt zu Verhandlungen. Das Verhandlungsergebnis besteht darin, auf die Vornahme der Qualitatserhohung zu verzichten und den Gewinn aus dem Verzicht (namlich die Kosteneinsparung!) gleichmaig zwischen den Parteien zu teilen. Der Nachfrager erhalt also $c/2$. Das optimale Investitionsvolumen fur den Nachfrager ergibt sich damit als Losung des Maximierungsproblems

$$\max_x \left(xv + (1 - x) \frac{c}{2} - \frac{x^2}{2} \right)$$

Es fuhrt zu den Gleichungen

$$x^{BC} = v - c/2 > x^*$$

und

$$W^{BC} = \frac{1}{2}(v - c/2)(v - 3c/2).$$

Das bemerkenswerte Ergebnis ist hierbei, da der Kauser nun zu viel investiert. Der Grund hierfür ist seine Entscheidungsbefugnis, die ihn in die Lage versetzt, seinem Lieferanten die Kosten c aufzuburden. Er berucksichtigt (internalisiert) diese Produktionskosten also nicht und investiert infolgedessen zu viel in Aktivitaten, die die Vornahme der Qualitatserhohung wahrscheinlicher machen.

Es lat sich leicht erkennen, da es je nach den Umstanden optimal sein kann, die Entscheidungsbefugnis dem Anbieter oder dem Nachfrager zu geben. Ist $c = 0$, so ist letzteres effizient (es gibt keine Produktionskosten, folglich kann auch das Problem einer fehlenden Internalisierung solcher Kosten nicht auftreten); eine Entscheidungsbefugnis des Anbieters ist dann suboptimal. Ist $v = c > 0$, ist es optimal, die Investition zu unterlassen. Dann empfiehlt es sich, da entweder der Anbieter die Entscheidungsbefugnis besitzt oder keiner (dies ist der Fall der Verhandlungen "mit weiten Verhandlungsspielraumen"). Entscheidungsbefugnis des Nachfragers fuhrt tendenziell zu hoheren Investitionen, und diese weisen eine negative Ertragsrate auf.

Der Erstreckungsbereich der Entscheidungsbefugnis

Wir haben vorausgesetzt, daß die Menge der möglichen Entscheidungen D , aus denen die entscheidungsbefugte Partei eine auswählt, ex ante wohl abgegrenzt ist. Diese Annahme steht möglicherweise im Widerspruch zur Transaktionskosten-Hypothese. Wenn die künftige Entwicklung unvorhersehbar ist oder die Zahl der möglichen Ereignisse derart groß, daß es nicht möglich ist, sie in einem Vertrag zu berücksichtigen, dann ist die Menge der Entscheidungen, die unter Berücksichtigung dieser künftigen Umstände in Frage kommen, wohl auch nicht vorhersehbar oder zu komplex, um einer Beschreibung zugänglich zu sein (tatsächlich mag es sein, daß diese Umstände nichts anderes sind als die möglichen Handlungen). Aber wie ist D abzugrenzen, wenn eine solche Abgrenzung nicht innerhalb des Vertragswerks geleistet werden kann? Grossman und Hart führen Entscheidungsbefugnis auf das Eigentumsrecht an physischen Kapitalgütern zurück; der grundlegende Gedanke ist, daß der Eigentümer einer Maschine das Recht hat, diese nach seinem Belieben zu nutzen. Kreps (1984) weist darauf hin, daß ein Eigentumsrecht auch in bezug auf immaterielle "Kapitalgüter" wie z.B. eine Reputation bestehen kann. Es kann sich auch auf eine Funktion innerhalb eines Unternehmens beziehen, nämlich dann, wenn Entscheidungsbefugnisse delegiert werden. Ein Vorarbeiter mag innerhalb seines Verantwortungsbereichs über einige Entscheidungsbefugnis verfügen.

Obwohl das Eigentumsrecht (und sei es ein delegiertes Eigentumsrecht) dazu beiträgt, die Menge D abzugrenzen, wird dieses Ziel doch bei weitem nicht vollständig erreicht. Es mag sein, daß der Eigentümer der Maschine sie höchst unterschiedlich einsetzen kann. Entfernt er allerdings den Lärmdämpfer, so dürften sowohl die Arbeiter als auch das Gericht zu dem Schluß kommen, daß er dazu keineswegs befugt war. Ein Manager, dem ein Unternehmensteil unterstellt ist, mag befugt sein zu entscheiden, ob die Sekretärinnen Schreibmaschinen oder Textverarbeitungsmaschinen benutzen sollen, aber er kann nicht verlangen, daß sie an Bildschirmen mit erhöhter Strahlenbelastung arbeiten. Das Verteidigungsministerium kann die Hersteller von Rüstungsgütern verpflichten, bestimmte auch anderweitig übliche Sicherheitsvorkehrungen zu treffen, sogar wenn dies nicht im Vertrag steht, aber es kann nicht verlangen, daß das Design des bestellten Produktes wesentlich verändert wird. Wie Kreps richtig bemerkt, muß es über den Erstreckungsbereich der Entscheidungsbefugnis ein Vorverständnis geben, das von den Parteien geteilt wird, obwohl seine Grenzen nicht im Vorhinein genau abgesteckt wurden; ein solches Vorverständnis (das auch einen Bestandteil einer Unternehmenskultur darstellt) kann sich leichter herausbilden, wenn die möglichen, aber nicht vorhersehbaren Ereignisse, auf die wir oben Bezug ge-

nommen haben, gewisse Muster erkennen lassen.

Da der Erstreckungsbereich der Entscheidungsbefugnis ex post geklärt werden muß, stoßen wir hier auf einen neuerlichen Grund, Schlichtungsverfahren vorzusehen. Das Problem ähnelt sehr dem oben erörterten. Ein "interner" oder "externer" Schlichter mag damit überfordert sein, die Entscheidungen selbst im einzelnen zu beurteilen und zu korrigieren, aber er mag imstande sein, die Menge der zulässigen Entscheidungen einzugrenzen. Tatsächlich wird in einem Unternehmen die Delegation von Entscheidungskompetenzen (beispielsweise auf Vorarbeiter) normalerweise dadurch ermöglicht, daß es noch eine höhere Ebene gibt (beispielsweise das Management), die eventuelle Konflikte schlichten kann.

Ein weiterer möglicher Schutz vor Machtmißbrauch besteht darin, daß es der nicht entscheidungsbefugten Partei erlaubt wird, die Beziehung zu verlassen. Hat sie diese Möglichkeit, dann bedeutet das nichts anderes, als daß sie ex post befugt ist, die Entscheidung der entscheidungsbefugten Partei zurückzuweisen.

Empirische Untersuchungen

Vertikale Integration wird ebenso wie langfristige Verträge umso wahrscheinlicher, je spezifischer die Investitionen werden. Dies kann es erschweren, die zwei Phänomene empirisch voneinander zu trennen. Die Hauptschlußfolgerung aus der Coase-Williamson-Analyse ist in dieser Hinsicht, daß vertikale Integration umso wahrscheinlicher (und ein langfristiger Vertrag umso unwahrscheinlicher) wird, je höher die "Transaktionskosten" ausfallen. Die Vermutung liegt natürlich nahe, daß die "Transaktionskosten" hoch sind, wenn erhebliche technologische Unsicherheit vorliegt. (Dies könnte die Erklärung dafür liefern, daß (in den USA) Kohlebergwerke und Elektrizitätswerke nur selten vertikal integriert sind, eine zweite mögliche Erklärung ist allerdings die, daß die Aufsichtsbehörden sich meistens gegen solche vertikalen Zusammenschlüsse wandten). Unglücklicherweise ist es bei empirischen Arbeiten kaum möglich, "Vorhersehbarkeit" oder Komplexität zu messen; man muß versuchen, brauchbare Ersatzgrößen zu finden.

Die meisten der vorliegenden Fallstudien und Regressionsanalysen befassen sich vor allem mit dem Zusammenhang zwischen der Spezifität der Investitionen und der Wahrscheinlichkeit, mit der vertikale Integration eintritt. Die Arbeit von Klein, Crawford und Alchian (1978) enthält eine besonders interessante Analyse dessen, wie Spezifitäten der Kapitalgüter zur vertikalen Integration zwischen General Motors und Fisher Body sowie zwischen Pipelinebetreibern, den Besitzern von Ölfeldern und denen von Raffinerien geführt

haben. Monteverde und Teece (1982) untersuchten die Beschaffung von Bauteilen durch amerikanische Autohersteller. Um zu erklären, warum einige Bauteile hinzugekauft, andere aber selbst erstellt werden, greifen sie insbesondere auf die folgenden Variablen zurück: das Ausmaß, in dem das Bauteil für den jeweiligen Autohersteller spezifisch ist (Handelt es sich um ein Bauteil, das speziell für diesen Autohersteller entwickelt wurde und deswegen beispielsweise ein besonderes Design aufweist? Kann es kurzfristig auch von anderen Lieferanten bereitgestellt werden? Gibt es also alternative Beschaffungsquellen?) und die Komplexität des Systems, in welches das Bauteil eingefügt werden muß (Motor, Fahrgestell, etc.)⁵⁶. Sie zeigen, daß insbesondere die erstere Variable einen erheblichen Einfluß auf die Integrationsentscheidung besitzt. Masten (1984) kommt in seiner Untersuchung über den Flugzeugbau zu ähnlichen Ergebnissen. Anderson und Schmittlein (1984) befassen sich mit der Organisation des Vertriebs (Vertrieb durch Angestellte des Herstellers oder durch unabhängige Vermittler). Sie zeigen, daß die Spezifität des Humankapitals des Vertriebspersonals negativ mit der Wahrscheinlichkeit korreliert ist, daß unabhängige Vermittler eingesetzt werden. Die Spezifität messen sie, indem sie Manager befragen, wie lange man braucht, um mit der Funktionsweise des Unternehmens im erforderlichen Maße vertraut zu werden, die Eigentümlichkeiten seiner Produkte und seiner Märkte kennenzulernen und ähnliches mehr.⁵⁷

Reputation als Ersatz für Verträge oder Integration

Hinter diesem und den vorherigen Unterabschnitten steht die folgende Vorstellung: Um künftigen Gefahren zu entgehen, sollten die Parteien vollständige Verträge unterzeichnen, oder, wenn dies zu kostspielig oder unmöglich ist, die Entscheidungskompetenzen so verteilen, daß ein möglichst effizientes Ergebnis erzielt wird. MaCaulay (1963) konnte demgegenüber zeigen, daß die Beziehungen zwischen Unternehmen in der Praxis tendenziell deutlich mehr informelle Elemente enthalten als die theoretische Analyse nahelegt. Dies ist sogar dann der Fall, wenn es sich um eine langfristige Beziehung handelt. Effizienz ist dann möglicherweise dennoch gewährleistet, und zwar dadurch, daß die Partner eine Reputation

⁵⁶ Der Grund für die Einführung dieser Variable ist der folgende: Sowohl Williamson (1975) als auch Scherer (1980, S. 90) haben das Argument vorgebracht, daß vertikale Integration ein höheres Maß an Koordination ermöglicht - nämlich durch Entscheidungen des jeweiligen Vorgesetzten (in einem unvollständigen Vertrag bleibt die genaue Zeitstruktur des Produktionsprozesses zum Teil ungeregelt). In einem komplexen System kann ein solches hohes Maß an Koordination erforderlich sein.

⁵⁷ Weitere interessante Beispiele hierfür finden sich bei Williamson (1985).

zu verlieren haben. Eine Unternehmung geht, wenn sie "unlautere" Geschäfte tätigt (genauer gesagt, eine Entscheidung trifft, die nicht die aggregierten Erträge der beiden Partner maximiert), das Risiko ein, daß der übervorteilte Partner davon Abstand nimmt, mit ihr noch weitere Geschäfte abzuschließen (vergleiche Williamson 1975, Kapitel 6, Kreps 1984, den Unterabschnitt 2.2 dieses Kapitels und die Kapitel 2 und 6 weiter unten). Eine Reputation versetzt ein Unternehmen in die Lage, die Kosten einzusparen, die die Ausarbeitung eines vollständigen Vertrags oder eine genaue Aufteilung der Entscheidungskompetenzen mit sich bringt. Auf der anderen Seite setzt das Fehlen detaillierter Vereinbarungen das Unternehmen der Gefahr des Opportunismus aus. Folglich würde man erwarten, daß informelle Beziehungen dann überwiegen, wenn sich die spezifischen Investitionen in Grenzen halten und die Geschäftspartner vergleichsweise häufig vor der Frage stehen, ob sie ein neues Geschäft abschließen wollen. Je häufiger sie vor dieser Entscheidung stehen, umso geringer ist der Anreiz, den anderen zu übervorteilen.

Dual Sourcing als Ersatz für Verträge

Ein weitere Möglichkeit, das Problem des "hold-up" zu vermeiden (daß man nach der Vornahme einer spezifischen Investition gewissermaßen in der Falle steckt), besteht darin, nach Möglichkeit dafür zu sorgen, daß auch ex post noch Wettbewerb vorhanden ist. Farrell und Gallini (1986) und Shepard (1986) haben Modelle von der Williamsonschen Art analysiert, in denen der Nachfrager in spezifische Kapitalgüter investiert und der Anbieter ex post über einige Variablen (sagen wir, es handelt sich um Qualitätsvariablen) entscheidet, die ex ante einer vertraglichen Vereinbarung nicht zugänglich waren⁵⁸. Ex post hat der Anbieter folglich den Anreiz, die Qualität zu senken⁵⁹; darum investiert der Abnehmer ex ante wenig in die Beziehung. Dual sourcing heißt, daß der Nachfrager dieselben Einsatzgüter von zwei oder mehreren Anbietern bezieht, so daß ihm auch ex post noch die Möglichkeit verbleibt, auszuwählen; die Anbieter liefern sich einen Qualitätswettbewerb. Es handelt sich gewissermaßen um eine Diversifizierung des Beschaffungsrisikos. Dies führt dazu, daß im Gleichgewicht sowohl ein höheres Qualitätsniveau als auch ein größeres Volumen der ex ante Investitionen gewährleistet ist. Wettbewerb löst oder lindert damit das Problem eines bilateralen Monopols ex post und führt zu mehr Effizienz. Farrell, Gallini und Shepard

⁵⁸ Bei Farrell und Gallini (1986) handelt es sich bei dieser Variable um einen Preis, bei Shepard (1986) um eine Lieferfrist.

⁵⁹ In den angeführten Modellen handelt es sich dann um einen hohen Preis oder eine lange Lieferfrist.

argumentieren, daß dies eine überzeugende Erklärung dafür liefert, warum Intel Lizenzen für seine Mikroprozessortechnologien erteilt oder warum IBM eine Politik der "offenen Architektur" in bezug auf seine Personalcomputer verfolgt.

2 Die Gewinnmaximierungshypothese

In diesem und den meisten anderen wirtschaftstheoretischen Büchern wird unterstellt, daß Unternehmen den Erwartungswert ihrer Gewinne maximieren. Viele Leute hegen demgegenüber den Verdacht, daß die Manager in Wirklichkeit andere Ziele verfolgen (z. B. die Unternehmensgröße maximieren oder das Wachstumstempo des Unternehmens oder die Vergünstigungen, die sie genießen)⁶⁰. Dieser Abschnitt stellt Argumente für und gegen die Gewinnmaximierungshypothese dar. Wir erörtern auch die Aussagekraft, die die industrieökonomische Theorie in einer Welt von Unternehmen, die nicht ihren Gewinn maximieren, für sich in Anspruch nehmen kann.

Die Aktionäre einer Unternehmung besitzen einen Anspruch auf die Gewinne des Unternehmens, wobei bei der Ermittlung dieser Gewinne die Aufwendungen für die verschiedenen Inputs von den Erträgen abzuziehen sind. Könnten sie selbst das Unternehmen leiten, würden sie Entscheidungen treffen, die die Kosten minimieren und den Gewinn maximieren⁶¹. Es muß folglich in erster Linie an der Trennung von Eigentum und Kontrolle (separation of ownership and control) liegen, wenn ein Unternehmen nicht danach strebt,

⁶⁰ Vergleiche die Modelle des Unternehmensverhaltens, die in den fünfziger und sechziger Jahren entwickelt wurden, z.B. Baumol (1962), Marris (1964).

⁶¹ Es gibt zwei Gründe, warum die Aktionäre vielleicht nicht den Erwartungswert der Gewinne maximieren wollen (oder den Wert des Unternehmens). Erstens sind sie möglicherweise risikoavers; dann ist ihnen unter Umständen an einer Unternehmenspolitik gelegen, die dazu führt, daß sich die Gewinne des Unternehmens antizyklisch verhalten (und damit die Risikohaltigkeit ihres Portfolios verringert wird) und zwar auch um den Preis, daß der Erwartungswert der Gewinne nicht maximiert wird. Der zweite Grund hat ebenfalls mit Wechselwirkungen zu tun, die bei der Betrachtung eines allgemeinen Gleichgewichts erkennbar werden. Liegt kein vollkommener Wettbewerb vor, übt der Preis, den ein Unternehmen für sein Produkt verlangt, nicht nur deswegen einen Einfluß auf das Nutzenniveau seiner Aktionäre aus, weil er Auswirkungen auf seinen Gewinn besitzt, sondern auch wegen seiner Auswirkungen auf den Konsum der Aktionäre. (Es kann dann durchaus der Fall sein, daß unter den Aktionären keine Einigkeit über die optimale Unternehmenspolitik erzielbar ist.) Diese zwei Effekte sind für die Theorie bedeutsam, aber ob sie tatsächlich von einer solchen Stärke sind, daß sie empirisch relevant sind, ist keineswegs klar. Das Portfolio von Aktionären ist meist ziemlich gut diversifiziert, und die Möglichkeiten, durch eine entsprechende Unternehmenspolitik für ein antizyklisches Verhalten der Gewinne des Unternehmens zu sorgen, sind eher beschränkt. Die Menge an Gütern ihres Unternehmens, die Aktionäre (zumindest die tatsächlich einflußreichen Aktionäre) konsumieren, ist normalerweise eher gering, so daß die Wirkungen des Preises auf den Konsum im Vergleich zu den Einkommenseffekten, die von den Gewinnen des Unternehmens ausgehen, nicht sonderlich ins Gewicht fallen.

seinen Gewinn zu maximieren⁶². Die Anhänger der sogenannten Principal-Agent-Theorie und ihrer Weiterentwicklungen haben sich seit den frühen siebziger Jahren auf den Standpunkt gestellt, daß Abweichungen vom Gewinnmaximierungsverhalten nicht einfach postuliert, sondern vielmehr erklärt werden sollten. Sie sollten darauf zurückgeführt werden, daß die Aktionäre nicht in der Lage sind, die Manager ausreichend zu kontrollieren und die Kosten- und Nachfragesituation ihrer Unternehmen zu erkennen. Beispielsweise besitzen die Manager einen Informationsvorsprung vor den Aktionären, und dies verschafft ihnen Entscheidungsspielräume (die sie z.B. nutzen können, um sich Vergünstigungen zu verschaffe.). Das Streben eines Unternehmens nach Expansion ist möglicherweise weder im Sinne der Aktionäre, noch entspricht es unmittelbar den Präferenzen der Manager. Erklären läßt es sich vielmehr aus dem Konflikt zwischen Aktionären und Managern. So mag der Umstand, daß die Aktionäre über die Technologie des Unternehmens unvollkommen informiert sind, den Managern die Möglichkeit verschaffen, den Personalbestand aufzublähen. Dies hat für die Manager den angenehmen Effekt, daß es den Leistungsdruck am Arbeitsplatz senkt (oder anders ausgedrückt, ihnen ein ruhigeres Leben verschafft). Oder aber die Manager erwarten sich von einer Expansion des Unternehmens zusätzliche Möglichkeiten für einen beruflichen Aufstieg.

Eine umfassende Würdigung der Principal-Agent-Literatur und alternativer Erklärungsansätze würde den Rahmen dieses Kapitels sprengen⁶³. Wir beschränken uns hier darauf, die wichtigsten Probleme zu erörtern. Zunächst werden wir das einfachste mögliche Modell eines Problems mit moralischem Risiko betrachten und erörtern, wie finanzielle Anreize, Vergleiche mit Konkurrenten, Übernahmeangebote, der Wettbewerb auf Gütermärkten und Aufsicht träge Manager in Trab bringen und ihre Entscheidungsspielräume einschränken. Die Grenzen dieser Kontrollmechanismen werden ebenfalls aufgezeigt. Es wird weiter dargestellt, warum selbst in Situationen, in denen die Gewinnmaximierungshypothese als falsifi-

⁶² Wegen einer Erörterung empirischen Materials über die Trennung von Eigentum und Kontrolle vergleiche Scherer (1980, S. 32 - 33).

⁶³ Arrow (1985) enthält eine verbale Einführung in die Problematik. Man unterscheidet zwischen Modellen, in denen eine bestimmte Handlung nicht beobachtbar ist (hidden-action) - dies ist der Fall des moralischen Risikos - und solchen, in denen bestimmte Informationen für eine Partei nicht zugänglich sind (hidden-information). Im ersteren Fall nimmt der Agent eine Handlung vor, die für den Principal nicht beobachtbar ist. Im letzteren ist er über eine bestimmte exogene Variable, die einen Einfluß auf die Umgebung besitzt, besser informiert. Bei den Modellen vom letzteren Typ lassen sich zwei Arten unterscheiden. Unterscheidungskriterium ist, ob der Agent vor oder nach Unterzeichnung des Vertrages in den Besitz der betreffenden Informationen gelangt (im letzteren Fall handelt es sich um den Fall adverser Selektion). Hart und Holmström (1986) enthalten eine sehr empfehlenswerte kritische Darstellung von Modellen, in denen keine adverse Selektion vorliegt (ferner behandeln sie die Theorie unvollständiger Verträge). Wegen eines Überblicks über Modelle adverser Selektion, die Beschaffungsvorgänge oder Fragen der Regulierung zum Gegenstand haben, vergleiche Baron (1986), Caillaud et al. (1988) sowie Sappington und Stiglitz (1987). Im nachfolgenden Abschnitt befassen wir uns vorwiegend mit dem Fall eines moralischen Risikos.

ziert gelten muß, die Implikationen, die sie für die industrieökonomische Theoriebildung besitzt, dennoch nicht falsch zu sein brauchen.

Wenn wir weiter unten von der Entlohnung des Managers sprechen, so fassen wir diese in einem weiten Sinne auf. Sie kann sowohl in Geld bestehen (dies wird in den nachfolgenden Modellen unterstellt) als auch in einer Beförderung, in Vergünstigungen und Annehmlichkeiten, Prestige oder darin, daß die Abteilung, die der Manager leitet, mit zusätzlichen Ressourcen ausgestattet wird. Obwohl wir das Principal-Agent-Modell hauptsächlich auf die Beziehung zwischen Aktionären und den Managern, also das Problem der Trennung von Eigentum und Kontrolle anwenden, sollte der Leser nicht verkennen, daß viele der Anreizmechanismen, die wir ableiten, sich ebenso auf anderen Hierarchieebenen anwenden lassen; manche davon besitzen für eine Anwendung auf den unteren Hierarchieebenen sogar eine größere Eignung als im Bereich der Spitzenführungskräfte.

2.1 Das grundlegende Anreizproblem

Wir erörtern nun eine Variante des Agency-Problems, bei der das moralische Risiko im Vordergrund steht (dies ist der Schwerpunkt dieses Unterabschnitts). Es hat mit dem Gegensatz zwischen Versicherung und Anreizen zu tun. Die Theorie der optimalen Versicherung zeigt, daß eine Zufallsgröße, wie z.B. der Gewinn, so zwischen einer risikoneutralen (Aktionäre) und einer risikoaversen Partei (Manager) aufgeteilt werden sollte, daß die risikoneutrale das gesamte Risiko trägt⁶⁴. Dies gilt, solange man von allen Anreizproblemen absieht. (Siehe Arrow 1970 und Borch 1963). Man stelle sich eine (diskrete) Zufallsgröße Π vor, die zwischen zwei Parteien aufgeteilt werden soll. Die Aktivitäten der beiden Parteien üben keinen Einfluß auf den Ausgang des Zufallsexperiments aus. Die möglichen Ausprägungen von Π bilden die Menge $\Pi_1 < \dots < \Pi_i < \dots < \Pi_n$, wobei den einzelnen Elementen die Wahrscheinlichkeiten $p_1, \dots, p_i, \dots, p_n$ zugeordnet sind (wobei $p_i > 0$ und $\sum_{i=1}^n p_i = 1$). $\Pi - w(\Pi)$ und $w(\Pi)$ beschreiben, wieviel die risikoneutrale und die risikoaverse Partei jeweils erhalten, wenn Π eintritt. Der Erwartungsnutzen der beiden Parteien beträgt dann

$$E_{\Pi} [\Pi - w(\Pi)] = \sum_i p_i (\Pi_i - w_i)$$

⁶⁴ Die Zielfunktionen lauten Ex und $Eu(x)$, wobei x das Einkommen bezeichnet, u eine zunehmende und streng konkave Nutzenfunktion und $E(\cdot)$ den Erwartungswert, x ist die Zufallsvariable.

und

$$E_{\Pi} u(w(\Pi)) = \sum_i p_i u(w_i),$$

wobei $w_i \equiv w(\Pi_i)$. Ein effizienter (oder Pareto-optimaler) Vertrag maximiert den Nutzen der einen Partei bei gegebenem Nutzenniveau der anderen. Er erfüllt die Bedingung

$$\max_{\{w_i\}} \sum_i p_i (\Pi_i - w_i) \quad \text{unter der Nebenbedingung} \quad \sum_i p_i u(w_i) \geq U_0,$$

wobei U_0 eine Konstante ist. Der Lagrange-Ansatz für diese Optimierungsaufgabe lautet

$$L = \sum_i p_i (\Pi_i - w_i) + \lambda \left(\sum_i p_i u(w_i) - U_0 \right).$$

Leitet man nach w_i ab, so erhält man für alle i

$$u'(w_i) = 1/\lambda.$$

w_i ist also unabhängig von i , wenn der Manager stets risikoavers ist ($u'' < 0$). Das gleiche Ergebnis läßt sich auch für eine stetige Verteilung von Π ableiten.

Die risikoaverse Partei sollte also in den Genuß einer vollständigen Versicherung kommen (d.h. sie sollte in allen Zuständen (states of nature) dasselbe Einkommen erhalten). Hier stoßen wir nun aber auf das Anreizproblem. Man stelle sich vor, die risikoaverse Partei unternimmt irgendeine nichtbeobachtbare Handlung, die die Größe des zu verteilenden Kuchens (Gewinns) beeinflusst (in einem stochastischem Sinne). Diese Handlung ist mit Kosten verbunden. Man kann sich vorstellen, daß sie darin besteht, sich in einem bestimmten Maße anzustrengen. Man nehme weiterhin an, daß die risikoneutrale Partei nur die tatsächliche Größe des Kuchens sieht (die Höhe des Gewinns), nicht aber diese Anstrengung. Erhält die risikoaverse Partei ein Einkommen, dessen Höhe nicht von dieser Größe abhängt, so bietet sich ihr kein Anreiz, sich anzustrengen, weil solche Anstrengungen keinen Einfluß auf ihr Einkommen ausüben. Vollständige Versicherung verträgt sich also nicht mit der Notwendigkeit, Anreize zu setzen. In der Tat gibt es ein Spannungsverhältnis zwischen Versicherung und Anreizen, und beide Parteien erhalten im allgemeinen nur eine suboptimale Versicherung und einen suboptimalen Gewinn.

Es gibt einen Fall, in dem kein Spannungsverhältnis vorliegt. Angenommen, beide Parteien sind risikoneutral (d.h. u' ist konstant), so benötigt die Partei, die die nicht beobachtbare Handlung vornimmt (der Agent), keine Versicherung. Die andere (der