

Schriften des Vereins für Socialpolitik

Band 237

Elemente der ökonomischen Raumstruktur

Von

Friedel Bolle, Michael Fritsch, Dirk Schmelzer
Klaus Schöler, Helmut Seitz, Joachim Wagner,
Hans Wiesmeth, Reiner Wolff

Herausgegeben von Horst Todt



Duncker & Humblot · Berlin

Schriften des Vereins für Socialpolitik
Gesellschaft für Wirtschafts- und Sozialwissenschaften
Neue Folge Band 237

SCHRIFTEN DES VEREINS FÜR SOCIALPOLITIK
Gesellschaft für Wirtschafts- und Sozialwissenschaften
Neue Folge Band 237

**Elemente der
ökonomischen Raumstruktur**



Duncker & Humblot · Berlin

Elemente der ökonomischen Raumstruktur

Von

**Friedel Bolle, Michael Fritsch, Dirk Schmelzer
Klaus Schöler, Helmut Seitz, Joachim Wagner,
Hans Wiesmeth, Reiner Wolff**

Herausgegeben von Horst Todt



Duncker & Humblot · Berlin

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Elemente der ökonomischen Raumstruktur / von Friedel Bolle
... Hrsg. von Horst Todt. – Berlin : Duncker und Humblot,
1995

(Schriften des Vereins für Socialpolitik, Gesellschaft für Wirtschafts-
und Sozialwissenschaften ; N. F., Bd. 237)
ISBN 3-428-08103-X

NE: Bolle, Friedel; Todt, Horst [Hrsg.]; Gesellschaft für Wirtschafts-
und Sozialwissenschaften: Schriften des Vereins ...

Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen
Wiedergabe und der Übersetzung, für sämtliche Beiträge vorbehalten
© 1995 Duncker & Humblot GmbH, Berlin
Druck: Berliner Buchdruckerei Union GmbH, Berlin
Printed in Germany

ISSN 0505-2777
ISBN 3-428-08103-X

Gedruckt auf alterungsbeständigem (säurefreiem) Papier
gemäß der ANSI-Norm für Bibliotheken

Vorwort

Der Ausschuß traf sich vom 21.10 - 23.10.1993 in Freiberg in Sachsen zu seiner Jahrestagung.

Zu Beginn gedachte der Ausschuß der zwei verstorbenen Mitglieder Prof. Dr. Mählich und Prof. Dr. Müller von der Universität Freiburg. Beide hatten über Jahre hinweg Gesicht und Klima des Gremiums mitgeprägt. H. Müller war mit der Tätigkeit des Ausschusses besonders eng verbunden, weil er ihn als Vorsitzender mehrere Jahre gestaltet hat. Die Mitglieder erinnern sich gerne seines organisatorischen Talents und der lebhaften, anregenden Diskussionen in jener Zeit, die er geschickt in maßvolle Bahnen zu lenken verstand.

Der Ausschuß folgte auch in diesem Jahr der seit einiger Zeit geübten Praxis, die Ausschußarbeit nicht unter ein allgemeines Thema zu stellen. Gegenstand der Vorträge und entsprechend der Diskussionen war die laufende Arbeit der im deutschen Sprachbereich tätigen Wissenschaftler, die sich mit raumwirtschaftlichen Problemen befassen. So kann der Ausschuß seinen Mitgliedern einen Querschnitt der aktuellen Auseinandersetzung bieten. Entsprechend heterogen waren die Themen der Vorträge.

Friedel Bolle, Frankfurt/Oder, gab einen Überblick über den Stand der augenblicklichen Diskussion über die Preisbildung in Transportnetzen, einen Problemkreis, dessen Bedeutung für die Wettbewerbspolitik z.Z. intensiv erörtert wird; der Bezug zur regionalen Entwicklung wird jedoch noch immer nicht angemessen gewürdigt.

Eine Analyse regionaler und lokaler öffentlicher Güter führte *Hans Wiesmeth*, Dresden, durch. Die Vorgehensweise war strikt theoretisch im Rahmen eines geschlossenen Modells. Der Vortrag war insbesondere auch ein Plädoyer für Lindahl-Gleichgewichte.

Rainer Wolff, Paderborn, berichtete aus seiner Arbeit über genetische Algorithmen zur Lösung von Transport- und Rundreiseproblemen. Er vertrat und begründete die These, daß auf dieser Basis besonders effiziente Heuristiken entwickelt werden können. Er demonstrierte die von ihm entwickelte Variante eines solchen Algorithmus anhand eines Anwendungsbeispiels.

Aus einer ebenfalls noch nicht abgeschlossenen Untersuchung zum Thema räumlicher Preistechniken berichtete *Klaus Schöler*, Siegen. Einer

handfest praktischen Fragestellung näherte sich *Helmut Seitz*, Mannheim, von der theoretischen Seite her, nämlich dem Einfluß von städtischer Infrastruktur auf das verarbeitende Gewerbe.

Schließlich diskutierte *Michael Fritsch*, Freiberg, die regionale Beschäftigungsentwicklung im Zusammenhang mit Stillegungen und Neugründungen. Diese mehr empirische Untersuchung von großer praktischer Bedeutung war dem Ort der Tagung voll angemessen. Sie leitete zu den Fragestellungen über, welche die Wirtschaftspolitik der neuen Bundesländer prägen.

Neben diesen Arbeiten erscheint hier ein empirischer Beitrag von *Joachim Wagner* über die „Regionalen Unterschiede bei der Schließung von Industriebetrieben“, der während der Ausschußsitzung nicht Gegenstand der Beratungen war, weil der Autor kurzfristig seine Teilnahme absagen mußte. Der enge thematische Zusammenhang mit den Untersuchungen von Michael Fritsch lassen es jedoch sinnvoll erscheinen, die Arbeit dennoch an dieser Stelle zu veröffentlichen. Es ist zu wünschen, daß die Problematik der regional differenzierten ‚Todeshäufigkeit‘ von Industriebetrieben – nicht zuletzt wegen ihrer Aktualität – auf diese Weise doch noch die Diskussion erfährt, die sie verdient.

Vor diesem Hintergrund fand im Anschluß an die Referate eine Diskussionsrunde mit Herrn Dr. Walter Ort, Leiter des Referats ‚Grundsatzfragen der Wirtschaftspolitik‘ im ‚Sächsischen Staatsministerium für Wirtschaft und Arbeit‘ über die regionalen Probleme des Wiederaufbaus statt. Herr Dr. Ort informierte über die konkreten Probleme der Politik und über die Sichtweisen, die hier von der praktischen Seite her bestehen.

Horst Todt, Hamburg

Inhaltsverzeichnis

Die Preisbildung für Transporte in Netzen Von <i>Friedel Bolle und Dirk Schmelzer</i> , Frankfurt (Oder)	9
Regionale öffentliche Güter	
Von <i>Hans Wismeth</i> , Dresden	37
Zur näherungsweisen Lösung asymmetrischer Rundreiseprobleme mit Hilfe Genetischer Algorithmen	
Von <i>Reiner Wolff</i> , Paderborn	59
Räumliche Preistechniken im Wettbewerb	
Von <i>Klaus Schöler</i> , Kiel	77
Infrastrukturversorgung als Wettbewerbsinstrument zwischen Städten. Eine theoretische und empirische Untersuchung	
Von <i>Helmut Seitz</i> , Mannheim	101
Regionale Unterschiede bei Schließungen von Industriebetrieben. Eine Unter- suchung für Raumordnungsregionen in Niedersachsen (1979 - 1992)	
Von <i>Joachim Wagner</i> , Lüneburg	129
Betriebsgründungen und regionale Arbeitsplatzentwicklung	
Von <i>Michael Fritsch</i> , Freiberg	149

Die Preisbildung für Transporte in Netzen

Von *Friedel Bolle und Dirk Schmelzer*, Frankfurt (Oder)

Abstract

Transportleistungen werden auf jeder Stufe des Produktionsprozesses benötigt. Aufgrund der technischen und ökonomischen Eigenschaften von Transportnetzen werden Transportleistungen aber in der Regel nicht auf freien Märkten gehandelt, sondern von regulierten Unternehmen angeboten oder vom Staat bereitgestellt. Wir schlagen vor, einen Auktionsmechanismus für die Verteilung der Kapazitäten zu benutzen. Anhand eines Beispieles, in dem heterogene Güter über ein zentrales Netz transportiert werden, zeigt sich, daß in einer „first price auction“ eindeutige Gleichgewichte erreicht werden. Diese Aussage läßt sich verallgemeinern. Erhalten die Bieter die von ihnen geforderte Leistung und zahlen den von ihnen gebotenen Betrag, so ist jedes Gleichgewicht auch effizient. Experimentelle Ergebnisse lassen die Vermutung zu, daß solche Lösungen nicht nur theoretisch möglich sind, sondern auch tatsächlich (angenähert) erreicht werden können.

1. Vorbemerkung

Die verstärkte nationale und internationale Arbeitsteilung und die damit einhergehende Optimierung der Güterallokation führt zu überregionalen Tauschprozessen. Einer rationalen Organisation des Transportwesens kommt in diesem Prozeß entscheidende Bedeutung zu. Ein funktionierendes Transportwesen ist nicht nur Bedingung für den überregionalen Gütertausch, sondern selbst bereits Teil der Optimierung des Produktionsprozesses geworden (Just-in-Time Production).

Transporte in Netzen stellen somit das Rückgrat moderner Ökonomien dar. Transportleistungen werden dabei auf jeder Stufe des Produktionsprozesses, sei es für Rohstoffe, sei es für Halb- und Fertigprodukte oder für (Kommunikations-) Dienstleistungen benötigt. Sie sind vielfältig (Straßen, Bahnnetze, Binnenschiffahrt, Luftfverkehr, Leitungen) und komplex miteinander verknüpft. Sie weisen substitutive und komplementäre Beziehungen zueinander auf.¹

¹ Transporte auf Straßen, Wasserstraßen und Schienennetzen stehen für den Stückgutverkehr in Substitutionskonkurrenz. Komplementäre Beziehungen weisen sie aber im Roll-On-Roll-Off-Verkehr auf, so z.B. auf der geplanten Binnenschiffahrtsstrecke Dresden-Böhmen.

Aufgrund der technischen und ökonomischen Eigenschaften von Transportnetzen werden Transportleistungen dabei in der Regel von öffentlichen oder staatlich regulierten Monopolen angeboten. Sei es nur die Infrastruktur (Straßen), sei es der Transport selbst (Bahn, Telekommunikation) oder der Transport innerhalb eines Bündels von Leistungen (Strom, Gas, Telefonanlagen).

Obwohl die Auswirkungen der Monopolpreisbildung unter Ökonomen seit Cournot bekannt sind, blieben Forderungen nach einer Änderung der Industriestrukturen selten. Ein Grund dafür scheint darin zu liegen, daß sich die meisten Untersuchungen mit dem Einfluß der Industriestruktur auf allokativen Subeffizienzen befaßten. Beispielhaft sei hier auf Harberger verwiesen, der in einer empirischen Studie nachgewiesen hat, daß die Beseitigung von Monopolindustrien lediglich zu einem vernachlässigbaren Einkommensanstieg von 0,07% (für die Vereinigten Staaten) führen würde.² Der „Angriff“ der Ökonomie auf monopolistische Strukturen begann erst Mitte der sechziger Jahre und wurde an „zwei Fronten“ geführt. Zum einen wurde mit Harvey Leibenstein gleichnamigem Artikel neben der allokativen Effizienz die sogenannte „X-Effizienz“ als Maßstab für die Effizienz von Industriestrukturen herangezogen, zum anderen fragte Demsetz 1968 programmatisch „Why regulate Utilities?“³ Leibenstein kam zu dem Ergebnis, daß die Höhe der Stückkosten unter anderem vom Ausmaß der X-Effizienz bestimmt wird, die von der Intensität des Wettbewerbes abhängt. Mit steigendem Konkurrenzdruck, so seine These, steige auch die „X-Effizienz“ der Unternehmen. Das Potential zur Erhöhung der „X-Effizienz“ sah Leibenstein vornehmlich in unausgeschöpften Möglichkeiten zur Erhöhung der Arbeitsproduktivität. Im Unterschied zur allokativen Effizienz steckt in einer Erhöhung der „X-Effizienz“ jedoch ein weit größeres Einsparpotential. Demsetz hingegen kritisierte die Rationalität staatlicher Regulierung, die ausschließlich durch fallende Durchschnittskosten begründet wird. Seiner Meinung nach ist die bis dahin vorherrschende Theorie der Monopolpreisbildung unvollständig, da sie es nicht geschafft hat, den Zusammenhang zwischen economies of scale und Monopolpreisbildung herzustellen. Seine Unterscheidung von „competition for the field“ und „competition within the field“ (die auf Chadwick zurückgeht) mündete schließlich in der Theorie der bestreitbaren Märkte, wie sie von Baumol, Panzar und Willig ausformuliert wurde⁴. Seit Mitte der siebziger Jahre wurden Änderungsvorschläge unter den Schlagworten Deregulierung und Privatisierung in die Diskussion eingeführt. Teilweise sind bereits geänderte Industriestrukturen entstanden, die im wesentlichen zu einer vertikalen Desintegration der

² Zitiert nach Leibenstein (1966).

³ Demsetz (1968), Leibenstein (1966).

⁴ Baumol/Panzar/Willig (1982).

betroffenen Industrien geführt haben. In den hier besprochenen Sektoren kam es, wenn eine Änderung stattfand, meist zu einer Entbündelung des Transportsektors von komplementären Leistungen. Das Ziel war dabei im wesentlichen, die den Transporten vor- und nachgeordneten Produktionsstufen wettbewerbsfähig zu gestalten. Die Neustrukturierung der britischen Stromwirtschaft und die dortige Einrichtung einer wettbewerblichen Stromerzeugung ist ein populäres Beispiel.

Im Gegensatz zur (vertikalen) Neuordnung der Produktions- oder Handelsbereiche erweist sich die Neustrukturierung des Transportmarktes als besonders schwierig, da Transportnetze in vielfacher und komplizierter Weise economies of scope and scale aufweisen und damit natürliche Monopole darstellen.

Es wurden nun zahlreiche Organisationsformen und Preisbildungsmechanismen vorgeschlagen, die, wenn sie auch nicht zu einer Pareto-effizienten Allokation der Transportkapazitäten führen, zumindest im Sinne einer second-best Lösung der Monopolpreisbildung überlegen sind und „X-Ineffizienzen“ abbauen. Im weiteren wird daher zunächst ein (kurzer) Überblick über die ökonomischen Eigenschaften von Transportnetzen gegeben. Die Darstellung erfolgt anhand eines Beispieles aus der Gaswirtschaft, ist nach Meinung der Autoren jedoch mit geringen Änderungen auch auf andere Transportbereiche übertragbar. „Transport“ ist in diesem Zusammenhang streng von der Verteilung abzugrenzen. Der Begriff Verteilung bezieht sich vor allem auf die regionale Übertragung von Gas oder Strom, Transport hingegen auf die überregionale Übertragung. Vom Händler bis zur Haustür werden Strom und Gas also verteilt und vom Produzenten zum Händler transportiert. Die Implementierung der unten beschriebenen Ordnungsrahmen und Preisbildungsmechanismen kann für Verteilungsnetze teilweise mit sehr hohen Transaktionskosten verbunden sein. Es ist daher bei allen Änderungen der Industriestrukturen stets zu berücksichtigen, ob und inwie weit angesichts der geänderten Struktur Effizienzgewinne durch hohe Koordinationskosten kompensiert werden.

Allen Transportproblemen gemeinsam ist die Tatsache, daß in der statischen Betrachtung Nutzungsrechte für ein gegebenes Streckennetz unter der Berücksichtigung von Kapazitätsengpässen auf eine Anzahl von Nachfrägern mit unterschiedlichen Nutzenfunktionen zu einem „markträumenden“ Preis verteilt werden müssen. Daher werden in Kapitel 3 die unterschiedlichen Möglichkeiten der Organisation von Transportmärkten und in Kapitel 4 verschiedene Preisbildungsmechanismen dargestellt, um im Hauptteil der Arbeit ein „Auktionsmodell“ vorzustellen. In diesem Modell wird die Vergabe von Nutzungsrechten als „Team-Selection“-Problem modelliert. Dieses Verfahren erlaubt nicht nur eine theoretische Bearbeitung des Sys-

mes, sondern darüber hinaus den Rückgriff auf experimentelle Ergebnisse aus Untersuchungen verwandter Probleme.

2. Ökonomische Eigenschaften des Transportnetzes

Die Deregulierungskommission hat in ihrem Gutachten aus dem Jahre 1991⁵ festgestellt, daß die regulierten Monopole in Deutschland vielfältige Ineffizienzen aufweisen. Sie ist damit auf politischer Ebene Einsichten gefolgt, die in der ökonomischen Diskussion bereits seit Mitte der 60er Jahre thematisiert werden⁶. Im einzelnen kritisierte die Kommission, daß

- die staatlichen Instanzen mit der Regulierung überfordert seien,
- die Monopolunternehmen (hier insbes. EltvU) zu hohe Kapitalstöcke⁷ und
- Überkapazitäten aufweisen,
- zu hohe Abschreibungen und Rückstellungen getätigt werden,
- der Wirkungsgrad der Anlagen zu niedrig ist und
- zu hohe Löhne und Gehälter sowie Sozialleistungen gezahlt werden.

Es stellt sich die Frage, warum bei der offensichtlichen Ineffizienz der Monopole Güter – und dazu zählen eben auch Transportleistungen – weiterhin von Monopolindustrien bereitgestellt werden. Die Existenz, der Schutz und die Regulierung dieser Monopole stützt sich dabei auf drei Grundpfeiler. Zum einen weisen sie i.a. (a) steigende Skalenerträge, (b) Bündelungsvorteile und (c) hohe irreversible Kosten auf. Zum anderen wird argumentiert, diese Bereiche seien von großer Bedeutung für die Volkswirtschaft (Versorgungssicherheit, Schaffung der nötigen Infrastruktur und einer hohen Versorgungsdichte), so daß sie nicht dem Wettbewerb ausgesetzt werden dürften. Das letzte Argument ist der theoretischen Bearbeitung nur schwer zugänglich und seine Rechtfertigung findet sich eher im politökonomischen Bereich. Konzentrieren wir uns daher auf die ersten Argumente.

Steigende Skalenerträge (sinkende Durchschnittskosten⁸) sind Merkmal eines natürlichen Monopols. Eine gegebene Nachfrage kann so kostengün-

⁵ *Deregulierungskommission* (1991).

⁶ s.o.

⁷ Es handelt sich um den sogenannten Averch-Johnson Effekt, der bei der Rate-Of-Return Regulierung auftritt, wenn die zulässige Kapitalverzinsung über dem Marktzins angesetzt wird.

⁸ Sinkende Durchschnittskosten sind jedoch nur hinreichende, nicht notwendige Bedingung für ein natürliches Monopol. Es reicht die Subadditivität der Kostenfunktion. Vgl. dazu etwa *Bonbright, J. C.: Principles of Public Utility Rates*, New York 1969, S.14 - 15.

stiger von einem Anbieter befriedigt werden als durch mehrere Anbieter. Das folgende Beispiel aus der Versorgungswirtschaft illustriert diesen Sachverhalt.

Zwei, von der Produktion über den Transport bis zum Handel vertikal integrierte, Unternehmen A und B versorgen zwei Abnehmer C und D mit Gas. Unternehmen A beliefert Abnehmer D und Unternehmen B Abnehmer C. Anbieter und Abnehmer liegen in unterschiedlichen Regionen. Um die Belieferung aufnehmen zu können, müssen Versorgungsleitungen gebaut werden. Die Anbieter haben die Wahl zwischen zwei Alternativen. Jedes Unternehmen kann eine eigene Leitung zu seinem Kunden bauen (*Leitungssystem*), oder die Anbieter bauen eine gemeinsame Leitung, an die sie sich und die Abnehmer über Stichleitungen anschließen (*Netz*). Im ersten Fall entstehen für beide Unternehmen Kosten in Höhe von K_{l1} . Wenn beide Nachfrager ähnliche Maximalkapazitäten aufweisen gilt $K_{l1} = K_{l2}$. Die Gesamtkosten belaufen sich auf $\sum_{i=1}^m K_{li}$.

Die Kostenfunktion für die einzelne Leitung ergibt sich im wesentlichen aus

- den Baukosten,
- den Kosten von Leitungsverlusten (in modernen Gasnetzen relativ gering),
- den Kosten des Netzunterhaltes und Servicekosten, sowie
- den Koordinationskosten.

Die Höhe der variablen Kosten (i.e. der letzten drei Kostenarten) richtet sich nach der Transportleistung (-menge), nach der Transportkapazität, sowie nach der Temperatur, dem spezifischen Gewicht und der Dichte des Gases. Betrachtet man zunächst die Kosten der Leitungsverluste und die Baukosten, so lassen sich in einer Annäherung die Kostenverläufe in etwa wie in Abb. 1 darstellen.

Wenn beide Anbieter ein gemeinsames Netz errichten und Stichleitungen bauen, so liegen die Kosten des Netzes unter den Kosten getrennter Leitungen. Zwischen den Stichleitungen, mit denen die Anbieter und die Abnehmer an das Netz angeschlossen sind, treten Skalenerträge auf. Dies ist in der Tatsache begründet, daß eine Leitung, um die gleiche Transportkapazität wie zwei Leitungen bewältigen zu können, nicht den doppelten Durchmesser aufweisen muß (sog. 2/3 Regel). Darüber hinaus treten bei unregelmäßiger, nicht vollständig korrelierter Nachfrage Vorteile der „Durchmischung“ auf, denn die Kapazität einer gemeinsamen Leitung kann so unter der Kapazität der Summe zweier einzelner Leitungen liegen. Neben der Kostensparnis im Leitungsbau treten für die Anbieter auch Kosteneinsparpotentiale beim Unterhalt des Netzes auf. So können die Kosten der Besei-

tigung von Ausfallrisiken bei der Belieferung von Endabnehmern zwischen den Produzenten aufgeteilt werden. Außerdem besteht die Möglichkeit, die externen Kosten des Leitungsbaus zu verringern. Dies sind meist Veränderungen im Landschaftsbild (e.g. Freileitungen für den Stromtransport), Beeinträchtigungen in Naturschutzgebieten (e.g. Gasleitungen durch das Wattenmeer) und der Gesundheit (so werden bei Telekommunikationsnetzen gesundheitliche Schäden durch die Strahlung von Empfangs-/Sendestationen diskutiert).

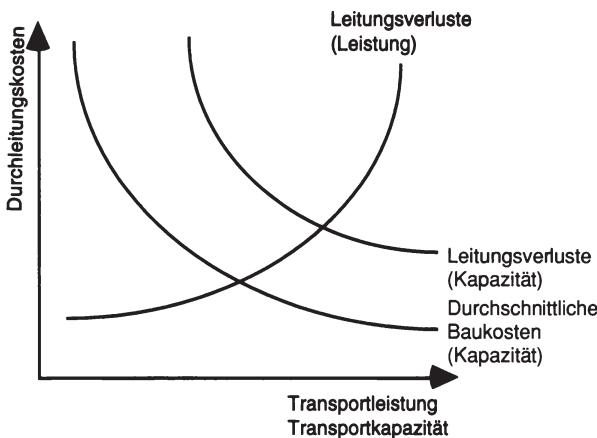


Abbildung 1: Kostenkomponenten beim Gastransport

Die dargestellten Kosteneinsparungen gelten sowohl für den Übergang von einem Leitungssystem zu einem Netzsystem, als auch jeweils im Netzsystem und im Leitungssystem beim Übergang zu einer höheren Kapazität.

Neben positiven Skalenerträgen weist der Gastransport Bündelungsvorteile mit den Bereichen Produktion und Handel auf. Ziel des Transporteurs ist es, große Gasmengen zu möglichst geringen Kosten und ohne Überschreiten von Druckrestriktionen rechtzeitig zum Abnehmer zu befördern. Zu diesem Zweck bedient er sich eines umfassenden Kontrollsystemes, welches Druck und Durchflußmenge im Netz überwacht. Dafür entstehen ihm Kosten der Koordination und Servicekosten. Der Rohrleitungsbetreiber kann durch dieses Kontrollsysteem unmittelbar auf Störungen im Netz reagieren. Dazu, ebenso wie zur sofortigen Bedienung von Nachfragespitzen, ist jedoch die Vorhaltung von Reservekapazitäten in Lagern notwendig. Gas wird aber in der Regel in alten Gasfeldern und Salzstockkavernen gelagert.

gert, die auch Produzenten zur Zwischenlagerung dienen, da die sofortige Anpassung der Produktion kostenintensiv ist. Außerdem kann von den Produzenten über die Anpassung von Druck und Durchfluß auch das Netz als Zwischenlager genutzt werden. Es treten also zwischen den vertikal nachgeordneten Bereichen Produktion, Transport und Handel zahlreiche Überschneidungen auf, die zu Bündelungsvorteilen führen. Die Höhe und damit die ökonomische Bedeutung der Bündelungsvorteile ist jedoch durch technische Entwicklungen, die die Kontrolle und Koordination zwischen den Bereichen Produktion, Transport, Speicherung und Handel erleichtern, gesunken.

Die Höhe des Investitionsrisikos sinkt mit einem Übergang vom Leitungssystem zum Netz. In einem Leitungssystem, wie es oben beschrieben wurde, kann nur Gas von A (B) nach C (D) transportiert werden. Bei einem Rückgang der Nachfrage in C wäre Anbieter A unter Umständen gezwungen, aus dem Markt auszutreten. Durch die starke Verbundenheit eines Netzes kann die ausgefallene Menge jedoch durch D oder durch eine weitere Stichleitung zu einem dritten Nachfrager E kompensiert werden. In einigen Fällen können so zwar einzelne Stichleitungen unbrauchbar werden, nicht jedoch das gesamte Netz.

Fraglich bleibt, ob die angesprochenen ökonomischen Eigenschaften der Transportnetze für Gas ausreichen, eine Regulierung der gesamten Industrie zu rechtfertigen. Die einfachste Form der Deregulierung, d.h. eine ersatzlose Streichung aller Regulierungsvorschriften, verbietet sich in den meisten Fällen. Denn bei freiem Marktz- und -austritt ist potentieller Wettbewerb im Sinne der Theorie der bestreitbaren Märkte auch angesichts vollständiger Information und kompetitiven Verhaltens der Anbieter nicht möglich. Der Aufbau von Netzen ist nämlich mit hohen irreversiblen Kosten verbunden, so daß der Marktaustritt nicht frei ist. Potentielle Neueintreter werden die gesamten Kosten des Netzaufbaus in ihre Kalkulation einbeziehen, der alteingesessene Anbieter wird jedoch nur die für ihn entscheidungsrelevanten variablen Kosten berücksichtigen. Um potentielle Neueintreter vom Marktzutritt abzuhalten, kann der etablierte Anbieter den Preis bis auf die kurzfristigen Grenzkosten herabsetzen.

In der Schaffung eines neuen Ordnungsrahmens wurden nun Wege gesucht, die zur Verbesserung der Situation beitragen können. Unter der Maxime „so viel Wettbewerb wie möglich“ soll durch vertikale Desintegration Wettbewerb in den Bereichen geschaffen werden, die sich auf das natürliche Monopol von Transportnetzen stützen, selber aber keine natürlichen Monopole sind. Es handelt sich also vor allem um die Schaffung von Wettbewerb bei der Produktion von Gütern und Leistungen. Die Transportnetze bleiben in ihrer monopolistischen Struktur erhalten. Jedem Produzenten und jedem Händler (jedem Nachfrager oder Anbieter von Transportlei-

stungen) wird jedoch freier Zugang zum (regulierten) Transportmarkt gewährleistet. Der Transportmarkt wird so zum zentralen Bestandteil des gesamten Systems und beeinflußt das Ergebnis auf den vertikal vor- und nachgelagerten Märkten⁹. Dort, aber vor allem auch im Transportsektor selbst, können erhebliche Leistungsverbesserungen erzielt werden. Dazu zählt die stärkere Kostenorientierung der Preise, verringerte Anpassungszeiten der Preise an geänderte Marktverhältnisse und die verbesserte Auslastung der bestehenden Transportkapazitäten sowie die Optimierung des Zubaues von Transportkapazitäten. Der monopolistische Transportsektor muß allerdings in den meisten Fällen weiter reguliert werden, und diese Regulierung ist entscheidend. Wir sollten uns bewußt sein, daß der Transportsektor nicht nur hohe Anreize braucht, Kosten zu sparen und optimal zu investieren (Kapazität zu erweitern), sondern daß der Transporteur eine „Market-Maker“-Funktion hat. Er bestimmt direkt oder indirekt über die Transportpreise die Güterflüsse. Somit gilt es, diesem Bereich ganz besondere Aufmerksamkeit zu widmen, damit die durch privatwirtschaftliche Konkurrenz erzielbaren Effizienzverbesserungen¹⁰ nicht durch subeffiziente Allokation der Transportleistungen wieder verloren gehen.

Wie aus der Theorie der Regulierung bekannt ist, resultieren allerdings oft nur „zweitbeste“ Lösungen. Neben diesen indirekten Kosten verursacht ein regulierter Markt zudem – wie jeder andere Markt auch – direkte Kosten. Solche Kosten abzuschätzen ist außerordentlich schwer. Sie legen aber die Frage nahe, ob es unter diesen Umständen nicht vielleicht – trotz Skalenerträgen und versunkenen Kosten – angezeigt wäre, auf Regulierung vollständig zu verzichten. In die folgende Diskussion wird auch diese Möglichkeit einzbezogen.

3. Unterschiedliche Ordnungsrahmen für Transportmärkte

Für die Gestaltung des Transportmarktes bieten sich unterschiedliche Möglichkeiten an, die im folgenden vorgestellt werden. Nicht erwähnt wird hier das herrschende Modell der vollständigen vertikalen Integration, da hier von einem gesonderten Transportmarkt nicht die Rede sein kann.

⁹ Auf europäischer Ebene wird die Debatte um eine wettbewerbliche Organisation der Versorgungswirtschaft spätestens seit dem Richtlinienvorschlag des Ministerrates vom Oktober 1990 geführt. Der Vorschlag wurde später im Hinblick auf die Diskriminierungsfreiheit beim Bau von Netzen, „unbundling“ und „Zwangsdurchleitung“ weiter konkretisiert. In seiner zunächst wettbewerblichen Orientierung wurde er jedoch Mitte 1993 durch die Vorschläge von Desama eingeschränkt.

¹⁰ Der Spielraum scheint trotz gegenläufiger Beteuerungen der betroffenen Industrien beträchtlich zu sein. So entließen die 1990 privatisierten britischen Stromversorger National Power und PowerGen die Hälfte ihrer Belegschaft, ohne daß die Produktion zurückgegangen wäre.

3.1 Freier Leitungsbau/Bypass

Für den Aufbau von Netzen ist zunächst einmal der Besitz von ausreichender Fläche oder die Genehmigung zur Benutzung von Fläche notwendig. Dies schließt den Aufbau eines privaten Straßennetzes aus und stellt sich sogar beim Aufbau eines Funknetzes (Vergabe von Frequenzen für das D2-Netz) als begrenzender Faktor für potentielle Anbieter dar. In der Regel schließen daher Versorgungsgesellschaften (Strom, Gas, Wasser) oder Verkehrsunternehmen mit Gebietskörperschaften sogenannte Konzessionsverträge über einen Zeitraum von mehreren Jahrzehnten ab. Diese Verträge geben ihnen das ausschließliche Wegerecht sowie damit verbunden das ausschließliche Recht zur Belieferung der Region (mit Ausnahmefällen). Als Gegenleistung erheben die Gebietskörperschaften von den Versorgungsunternehmen eine Konzessionsabgabe. Die Genehmigung für den Bau von Leitungen kann sowohl einer vertikal integrierten Versorgungsgesellschaft als auch einem reinen Transportunternehmen erteilt werden. Dies sei hier erwähnt, um dem Schluß vorzubeugen, freier Leitungsbau sei zwingend mit vertikaler Desintegration verbunden¹¹. Er ist vielmehr ein Ordnungsmodell, das sowohl in reinen Transportgesellschaften als auch bei vertikal integrierten Versorgungsgesellschaften anzuwenden ist.

Unter einem Regime des freien Leitungsbaus wird es jedem potentiellen Transporteur ermöglicht, sich nach geeigneten Projekten umzusehen und diese ohne eine Behinderung durch bestehende Transporteure auch durchzuführen. Dies schließt ein Genehmigungsverfahren und eine Gebührenerhebung nicht aus. Beides sollte aber diskriminierungsfrei jeden Anbieter von Transportleistung treffen. Wesentliches Wettbewerbselement ist dann nicht nur die bereits vorhandene, sondern auch die potentielle Konkurrenz, der sich die bestehenden Anbieter durch Neueintreter ausgesetzt sehen. Mit Ausnahme von Telekommunikationsnetzen leiden allerdings alle anderen Transportnetze unter langen Bauzeiten. Bestehende regionale und nationale Monopole sind deshalb nur mittel- oder langfristig einem Konkurrenzdruck ausgesetzt. Um den Wettbewerb sofort einzusetzen zu lassen, wäre in diesen Fällen eine Neuverteilung der Eigentumsrechte notwendig. Da es nicht immer parallele Leitungen gibt, müßte man sogar an die Verteilung von

¹¹ Vielmehr führt eine Kombination aus Konzessions- und sogenannten Demarkationsverträgen dazu, daß sich lediglich vertikal integrierte Unternehmen im Markt befinden. In der Bundesrepublik ist es verbreitete Praxis, daß die Versorgungsunternehmen untereinander sogenannte Demarkationsverträge mit Laufzeiten von mehreren Jahrzehnten abschließen. Diese Verträge sichern einem Unternehmen Absatzgebiete. „Konkurrenzunternehmen“ verpflichten sich, keine Abnehmer in diesem Gebiet zu beliefern. Diese auf der Handelsstufe greifenden Verträge machen es einer reinen Handelsgesellschaft unmöglich, in das bestehende Versorgungsgebiet einzudringen.

Anteilen an der Transportkapazität gewisser Strecken denken. Für Gasleitungen werden solche Anteile im amerikanischen Sprachgebrauch sehr anschaulich „Strohhalme“ (Straws) genannt.

Befürworter von Ausschließlichkeitsrechten im Leitungsbau kritisieren, daß unter dem Regime des freien Leitungsbaus auf die Realisierung von Economies of Scale and Scope weitgehend verzichtet wird. Mit diesem Argument geht der Vorwurf des „Cream Skimming“ einher. Den neueintretenden Transporteuren wird vorgeworfen, nur die profitablen Nachfrager zu bedienen, Nachfrager mit relativ geringen Deckungsbeiträgen jedoch nicht zu berücksichtigen. Der Grad des Verzichtes auf Skalenerträge ergibt sich aus der Gesamtnachfrage, der bestehenden Infrastruktur und der spezifischen Netzeigenschaft. So würde beim Bau von Parallelleitungen in der Gaswirtschaft unabhängig von der Gesamtnachfrage ein Verzicht auf die Realisierung von Skalenerträgen erfolgen. Liegen jedoch bereits Leitungen, so kann der Bau von Parallelleitungen durch andere Anbieter bei steigender Nachfrage aber durchaus sinnvoll sein. Die Möglichkeit des „Cream-Skimming“ hängt darüber hinaus von der Preisstruktur des (regulierten) Monopolisten ab. Außerdem müssen auch Nachfrager mit geringen Abnahmемengen und niedrigen Deckungsbeiträgen durch freien Leitungsbau nicht zwangsläufig schlechter gestellt werden.¹²

In der Bundesrepublik ist der überregionale Transport von Gas sowie der Aufbau konkurrierender Netze in der Telekommunikation (E-, D1-, D2-, C-Netze) ein Beispiel für „freien Leitungsbau“. Die Deutsche Verbundgesellschaft¹³ schlägt freien Leitungsbau als Alternative zum „Third Party Access“ (siehe unten) auch für die Elektrizitätswirtschaft vor, diesem Vorschlag werden allerdings in der sonstigen Diskussion nur geringe Chancen eingeräumt. In den Vereinigten Staaten wurden in den gleichen Branchen die Möglichkeiten zum Betreiben von Parallelleitungen schon sehr viel früher geschaffen. In der Telekommunikationsbranche wurde durch die Entwicklung von Kommunikationssatelliten in den 60er Jahren für unabhängige Telefongesellschaften die Möglichkeit eröffnet, den Hauptanbieter von Transportleistungen AT&T mit Parallelleitungen zu umgehen. Großen Gasabnehmern war es bereits seit 1978 gestattet, sich selbst über Stichleitungen mit Gas aus Interstate Pipelines zu versorgen.

¹² Laffont/Tirole (1993), S. 290ff.

¹³ Der Verband der überregional tätigen deutschen Elektrizitätsversorger (Verbundunternehmen), die Besitzer der Hochspannungsleitungen und somit von Durchleitungsansinnen am stärksten betroffen sind.

3.2 Open Access

Wirklich umstritten sind Transportverpflichtungen, die mit den Schlagworten „Common-Carrier“ (CC) und „Third-Party-Access“ (TPA) verbunden sind. Unter diesem Ordnungsrahmen ist generell jedes Transportunternehmen verpflichtet, für Dritte zu nicht diskriminierenden Preisen Gas oder Strom (TPA und CC sind in dieser Form bisher nur für Versorgungsunternehmen vorgeschlagen worden) durchzuleiten. Im Unterschied zum Common Carrier sind beim Third Party Access die Unternehmen nur bei freien Kapazitäten zur Durchleitung verpflichtet. Der Allokationsmechanismus ist also bei beiden Systemen „first-come-first-served“, jedoch genießen eigene Lieferungen im TPA-System Vorrang. Dies ist auch eines der wesentlichen Probleme des Open-Access-Regimes. Insbesondere wenn ein vertikal integriertes Unternehmen als Durchleiter mit Vorzugsrechten auftritt (TPA), kann die Vergabe von Transportrechten durch dieses Unternehmen strategisch genutzt werden. Konkurriert ein Transportunternehmen mit einem Händlerunternehmen um dieselben Nachfrager, so kann sich das Transportunternehmen unter Berufung auf Kapazitätsengpässe weigern, durchzuleiten. Dem Händlerunternehmen gehen so Abnehmer verloren, die dann unter Umständen wieder vom Transportunternehmen beliefert werden. Die strategische Nutzung des Transportmonopols ist einer der wesentlichen Gründe, warum mit der Forderung nach einem Open-Access-Regime auch häufig die Forderung nach Funktionentrennung einhergeht. Dies kann die rein rechnungsmäßige Trennung von Transport- und Handelsbereichen umfassen, kann aber auch bis zu der Bildung zweier verschiedener Gesellschaften gehen. Die mit der Umwandlung der Deutschen Bundesbahn in die Deutsche Bahn AG einhergehende Forderung nach der Bildung einer Fahrwege AG ist ein aktuelles Beispiel dafür. Der Zuständigkeitsbereich der Fahrwege AG bezieht sich nur auf die diskriminierungsfreie Vergabe von Transportkapazitäten. Die Bahn AG soll genau wie neu eintretende Transport- und Speditionsunternehmen als Nachfrager von Transportkapazitäten behandelt werden. In der amerikanischen Gaswirtschaft wurden mit der Order 436 (1985) der Federal Energy Regulatory Commission (FERC) amerikanische Pipelinegesellschaften verpflichtet, ihre Leitungen für Produzenten und Verbraucher von Gas zur Verfügung zu stellen.¹⁴ Bis zu diesem Zeitpunkt waren Pipelinegesellschaften Hauptanbieter von Gas bei großen Endverbrauchern und regionalen Handelsunternehmen. Sie schlossen direkte Verträge mit den Produzenten von Erdgas. Transportierte eine Pipelinegesellschaft 1984 nur 8% fremdes Gas, so waren es 1992 bereits 79%. Zur Zeit wird bei der FERC darüber nachgedacht, den Pipelinegesellschaften in

¹⁴ Zur Entwicklung der Durchleitungsverpflichtung in der amerikanischen Gaswirtschaft und den damit verbundenen Schwierigkeiten siehe *Funk* (1990).

Zukunft den Handel mit Gas ganz zu verbieten, so daß sie als reine Netzbetreibergesellschaften auftreten müßten.¹⁵ Auch in der Europäischen Union gibt es seit längerem Bestrebungen, Dritten den Zugang zu den Netzen zu ermöglichen. Betrachtet man den heutigen Stand der Diskussion, so soll es in Zukunft Abnehmern von großen Gasmengen möglich sein, sich aus alternativen Bezugsquellen zu versorgen und durch die Netze Dritter durchzuleiten. Eine Durchleitungspflicht besteht jedoch nur bei grenzüberschreitenden Lieferungen.

Der im letzten Unterabschnitt geschilderte freie Leitungsbau ist nicht inkompatibel mit einem Open-Access-Regime. Vielmehr kann unter bestimmten Voraussetzungen die Möglichkeit zu freiem Leitungsbau als eine glaubhafte Drohung gegenüber dem etablierten Transporteur, sollte dieser sich weigern, Kapazitäten für die Durchleitung zur Verfügung zu stellen, gesehen werden.

Die Schwierigkeiten, die mit allen Transportverpflichtungen verbunden und oben bereits angeklungen sind, werden in der Diskussion um die Auswirkung dieser beiden Konzepte besonders deutlich. Die Besitzer von Leitungen, die in Deutschland durchweg auch Handelsinteressen haben, wollen zunächst einmal jede Zwangsverpflichtung, für Konkurrenten zu transportieren, abwehren. Schon alleine die Diskussion auf Unternehmerseite um diese Transportrechte und die darauf folgenden zahlreichen, auf europäischer Ebene eingebrachten Änderungsvorschläge zeigen, daß CC und TPA auch nach ihrer Einführung auf beträchtlichen Widerstand stoßen werden. Einsicht in die Notwendigkeit, für Dritte zu transportieren, kann man nur von einem Unternehmen erwarten, das ein Monopol für den Transport im gesamten Wirtschaftsraum und – noch besser – das keine anderen Interessen als solche des Transportes hat.

3.3 Zentrale Netzgesellschaft

Um die mit den bisher erwähnten Ordnungsmodellen einhergehenden Schwierigkeiten zu umgehen, wird die Schaffung einer zentralen Netzgesellschaft vorgeschlagen. Diese Gesellschaft kann ein monopolistisches Staatsunternehmen sein, sie kann aber auch in den Händen eines „Club of Users“ liegen. Beide Modelle entbinden jedoch nicht von der Suche nach einem angemessenen Preismechanismus einerseits zur Kostendeckung des Netzbetreibers und andererseits zur effizienten Rationierung von Transportbestrebungen bei Kapazitätsengpässen.

¹⁵ Moyer (1993), S. 41.

4. Preisbildungsmechanismen

In diesem Abschnitt wird auf die unterschiedlichen Möglichkeiten der Preisbildung eingegangen. Die Mechanismen sind nicht streng an einen Ordnungsrahmen gebunden. Unterschiedliche Formen der Preisbildung können in Kombination mit unterschiedlichen Ordnungsrahmen auch zu unterschiedlichen Marktergebnissen führen. Häufig kommt dem Preis neben der Signalisierung von Knappheit (i.e. von Transportkapazitäten) auch noch die Aufgabe zu, Wettbewerb in den vor- und nachgelagerten Bereichen zu ermöglichen. Natürlich besteht zwischen diesen beiden Zielen kein strenger Konflikt. Anders ist es, wenn, gerade in so sensiblen Bereichen wie der Versorgungsindustrie, geringe interregionale Preisunterschiede aus Gründen der Verteilungspolitik angestrebt werden. Darüber hinaus führen Transportkosten dort, wo sie in signifikanter Höhe gefordert werden, zur sektoralen Aufspaltung von Märkten, die häufig nicht erwünscht ist. Hier kann es zu einem Konflikt zwischen ökonomischer Notwendigkeit und politischer Vorgabe kommen. Dies zeigt, daß Transportpreise ganz unterschiedlichen Zielen dienen. Konzentrieren wir uns jedoch auf die Suche nach einem effizienten Preis.

Ökonomischer Intuition folgend ließe sich in einer ersten Annäherung ein „effizienter“ Gleichgewichtspreis für den Transport als Differenz zwischen den Preisen zweier homogener Güter in verschiedenen Regionen ermitteln. Häufig ist die regionale Verteilung von Anbietern und Nachfragern nicht so eindeutig wie in dem in Abschnitt 2 dargestellten Beispiel. Es ist denkbar, daß sich direkt neben Abnehmer C ein Gasproduzent befindet. Der Gleichgewichtspreis für den Gastransport von A nach C ist dann die Differenz zwischen „Bohrlochpreis“ in C und „Bohrlochpreis“ in A. Ursprünglich wurde diese Bedingung für Stromnetze formuliert.¹⁶ Regionale Preisunterschiede zwischen den eingesetzten Kraftwerken entsprechen dann im „Economic Dispatch“¹⁷ genau den Transportkosten. Aufgrund der Homogenität dieser Güter ist zunächst auch irrelevant, daß die von dem Abnehmer entnommene Einheit nicht identisch mit der eingespeisten Einheit ist. Als problematisch erweist sich lediglich, daß der tatsächliche Transportweg aufgrund des Kirchhoffschen Gesetzes nicht der kürzesten bzw. kostenminimalen Verbindung zwischen Einspeiser und Abnehmer entsprechen muß. Obwohl auf Straßen-, Schifffahrts- und Schienennetzen nur selten homogene Güter befördert werden, ist die Bedingung, wie sie oben für Strom- und Gasnetze formuliert wurde, auch für Transportleistungen anderer Güter unmittelbar einsichtig. Dazu sei zunächst auf ein Beispiel aus der Erdöl-

¹⁶ Schweppé et al. (1988).

¹⁷ Die kostenminimale Anordnung der Kraftwerke zur Befriedigung der Gesamtnachfrage.

wirtschaft verwiesen. Mineralölunternehmen führen Swaps von Erdölprodukten durch, um Transportkosten zu umgehen. Hat ein konkurrierendes Unternehmen eine Raffinerie in einer Region, die das eigene Unternehmen beliefert, und hat dieses Unternehmen wiederum Abnehmer in der eigenen Region, so werden die Lieferverpflichtungen zwischen diesen beiden Unternehmen einfach ausgetauscht. Dies bedeutet aber nichts anderes, als daß bei identischen Preisen in unterschiedlichen Regionen nicht transportiert wird (bei Abwesenheit von Produktionsengpässen). Bei inhomogenen Gütern würde jedoch zu der Transportpreiskomponente noch ein „Qualitätsauf- oder -abschlag“ berechnet werden müssen.

Es sind unterschiedliche Preisbildungsverfahren vorgeschlagen worden, die im folgenden erläutert werden. Um ein effizientes Gleichgewicht zu erreichen, müssen die Preisbildungsmechanismen nicht nur in einem statischen Sinne die optimale Vergabe bestehender Kapazitäten sicherstellen, sondern auch den optimalen Kapazitätsauf- oder -abbau sicherstellen. Dabei kann das Problem der Kostendeckung für die Netzbetreibergesellschaft entstehen. Es sind auch hier nichtlineare Preise vorgeschlagen worden. Bei zweiteiligen Tarifen etwa sollen über das leistungsabhängige Preiselement die kurzfristigen Kosten der Netzbetreibergesellschaft gedeckt werden. Das fixe Preiselement sichert der Netzbetreibergesellschaft Gesamtkostendeckung. Sind fixe Preiselemente Gegenstand langfristiger Verträge, so sichern sich die Nachfrager den Netzzugang, werden auf der anderen Seite aber auch an den langfristigen Risiken der Netzinvestitionen beteiligt.¹⁸

4.1 Freie Verhandlungen

In einem System freier Verhandlungen bietet der Netzbetreiber freie Transportkapazitäten zu nicht regulierten Preisen an. Verhandlungsgegenstand sind Preise, Mengen und Kapazitäten (die die zeitliche Verteilung der gelieferten Mengen bestimmen). Der Netzbetreiber kann selbst Produzent und Händler sein, er kann aber auch als reiner Netzbetreiber agieren. Die Preisbildung auf der Basis freier Verhandlungen bietet den Vorteil, daß sie frei von Regulierungskosten ist. Der Markt, der so entsteht, wird zwischen zwei Extremen liegen. Auf der einen Seite können die Besitzer von Transportkapazitäten ihre Preise unabhängig voneinander setzen. Jeder Anbieter setzt seine Preise dann so, daß sein Durchschnittsgewinn bei gegebenen Preisen der anderen Anbieter maximiert wird. In Abhängigkeit von ihrer Position im Gesamtnetz können sie hohe Renten abschöpfen. Nur bei einem genügend dichten Transportnetz werden sich die Preise den (langfristigen) Grenzkostenpreisen annähern. Kooperieren die Anbieter oder gibt es nur

¹⁸ Hunt/Shuttleworth (1993).

einen Anbieter von Transportleistungen, so kommt es zu allgemeinen Verhandlungen, in denen jeder Teilnehmer bei Kenntnis seiner Marktposition seinen Gewinn maximiert. Das Postulat eines effizienten Verhandlungsergebnisses ist jedoch nur bei Abwesenheit von Verhandlungskosten aufrechtzuerhalten. Je größer die Anzahl der Verhandelnden ist, je häufiger die Verhandlungen wiederholt werden müssen, und je geringer die Handelsgewinne aus den Verhandlungen sind, desto stärker wird das Gewicht von Verhandlungskosten. Es spricht also vieles dafür, daß freie Verhandlungen nur bei geringer Wiederholungszahl und hohen Handelsgewinnen effizient sind. Dies sind i.d.R. langfristige Verträge, die mit Investitionsentscheidungen verbunden sind. Nun ist weder bei freien Verhandlungen noch bei unabhängiger Preisbildung gesichert, daß jedes gewinnbringende Transportgeschäft auch tatsächlich zustande kommt. Erhält der Transporteur jedoch einen Anteil vom Handelsgewinn der Transportleistung nachfragenden Unternehmen, so ist effiziente Rationierung auch in seinem Interesse. Je größer der Preisunterschied zwischen Gütern in zwei Regionen ist, desto stärker werden die Handelpartner von dem Transport des preiswerteren Gutes in die teurere Region profitieren. Je höher der Gewinn der Transportnachfrager ist, desto höher wird auch der Gewinn des Anbieters der Transportleistung sein. Den Autoren ist ein Beispiel für diese Verteilungsregel bisher nur aus der amerikanischen Elektrizitätswirtschaft bekannt.

Ein Beispiel für freie Verhandlungen aus dem Güterfernverkehr ist die Deutsche Bundesbahn, die zu frei verhandelbaren Preisen Transportleistungen anbieten kann.¹⁹ Im Western Systems Power Pool der amerikanischen Elektrizitätswirtschaft bieten die Stromversorgungsunternehmen neben der Belieferung mit Strom auch die reine Transportleistung an. An dieser Börse nehmen 24 Versorgungsunternehmen mit unterschiedlichen Erzeugungskapazitäten teil. Neben Preisen und Mengen für Angebot und Nachfrage von Strom werden dort über einen zentralen Broker auch freie Übertragungskapazitäten von Strom gehandelt. Die Informationen über die Nachfrager und Anbieter von Transportleistungen werden in Computersystemen gesammelt und sind jederzeit frei abrufbar. Die Preise für die Transportleistung sind dann zwischen den Parteien frei verhandelbar. Lediglich eine untere Preisgrenze wurde festgelegt.

Es ist jedoch nicht abzustreiten, daß sich vertikal integrierte Unternehmen, die auf ihren Beschaffungs- und Absatzmärkten in Konkurrenz zu den Unternehmen treten, für die sie Transportleistungen anbieten, sich in einem

¹⁹ Die ursprüngliche Tarifierungspflicht der Bundesbahn im Güterverkehr wurde durch die weitgehende Schaffung von Ausnahmebereichen immer weiter eingeschränkt, so daß die Bundesbahn einzelnen Nachfragern gegenüber Verhandlungsspielräume besitzt. Zu beachten ist allerdings, daß die Bundesbahn in scharfem Wettbewerb zu Straße und Wasserstraße steht.

Regime freier Verhandlungen strategisch verhalten. Man muß befürchten, daß sie ihre quasimonopolistische Stellung ausnutzen, indem sie entweder Transporte zu überhöhten Preisen anbieten, Transportkapazität verweigern oder notwendige Investitionen in die Erweiterung der Transportkapazität nicht vornehmen. Hier wird der Bezug zum zu Grunde liegenden Ordnungsrahmen besonders deutlich. Bei freiem Leitungsbau entsteht für den Netzbetreiber eine obere Preisgrenze, die er nicht überschreiten kann, ohne Gefahr zu laufen, daß sich Dritte im Markt als Transporteure engagieren.

4.2 Festgelegte/Regulierte Transporttarife

Ähnlich wie in einem System freier Verhandlungen sind die Transportmengen und die Transportkapazitäten in diesem Ordnungsrahmen weiterhin Gegenstand bilateraler Verhandlungen. Lediglich die Transportpreise unterliegen einer Tarifierungspflicht. Zwar wird dadurch die diskriminierungsfreie Bereitstellung von Transportkapazitäten sichergestellt, eine kurzfristige Anpassung der Preise an geänderte Marktgegebenheiten wird jedoch erschwert. Auch an dieser Stelle bietet sich ein Verweis auf die Deutsche Bundesbahn (Deutsche Bahn AG) an. Sie ist im Personenverkehr verpflichtet, jeden Nachfrager über ein beliebige Strecke von Transportkilometern zu befördern (Kontrahierungspflicht). Verhandlungsgegenstand ist lediglich die Transportstrecke. Die nötige Kapazität hat die Bahn bereitzustellen. Der Kunde kann jedoch zwischen verschiedenen Qualitäten wählen (InterCity, InterRegio, ICE etc.) über die er auch die Priorität seiner Beförderung (Lieferung) bestimmt. Eine den Bahnpreisen entsprechende Matrix von Preisen für den Transport zwischen verschiedenen Zonen wird von British Gas benutzt.

Die gesamte Transportgebühr setzt sich aus einem fixen und einem variablen Preiselement zusammen. Beide Elemente werden abhängig von der Transportstrecke berechnet. Die vom Durchleitenden zu zahlende Gebühr berechnet sich wie folgt:

Durchleitungsgebühr = Maximale Kapazität während der Kontraktperiode

- Capacity Element (fixe Gebühr)
- + durchgeleitete Gasmenge
- Commodity Element.

Zu beiden Elementen wird noch eine vom Konsumenten zu zahlende Anschlußgebühr addiert.

Ein weiteres Beispiel bietet der Florida Energy Broker, eine Organisation von Elektrizitätsunternehmen, die mittels kurzfristigen Handels Grenzko-

Tabelle 1
Transportpreiskomponenten für Gas von zwei der fünf Hauptanlandestationen zu ausgewählten Städten

	St Fergus		Easington	
	Commodity p/th	Capacity p/pdt/a	Commodity p/th	Capacity p/pdt/a
Aberdeen	0,4	93,4	0,9	217,6
Glasgow	0,9	212,9	0,8	190,7
Carlisle	1,2	283,3	0,7	156,1

p/th: pence per 100 thermal units

p/pdt/a: pence per peak day therm per annum (100 therms)

stenunterschiede in der Stromerzeugung ausnutzen wollen. Die Tarife der beiden transportierenden Unternehmen sind zwar freiwillig gesetzt bzw. ausgehandelt worden, sie bedürfen allerdings, ebenso wie die gesamte Marktorganisation, der Genehmigung der einzelstaatlichen und bundesweiten Regulierungsinstanzen.

4.3 Märkte für jede Transportstrecke

Im Unterschied zu den beiden bisher dargestellten Verfahren werden hier nicht komplette Transportstrecken zum Verhandlungsgegenstand, sondern Strecken zwischen Knotenpunkten. Diese Form der Preisbildung wird vor allem für Transporte in Stromnetzen gefordert, da hier aufgrund der Lastflüsse nach dem Kirchhoffsschen Gesetz die tatsächlichen Stromflüsse nur selten den kontrahierten Flüssen entsprechen. Durch eine Festlegung von Preisen an Knotenpunkten (i.e. an Einspeise- und Entnahmepunkten) wird den individuellen Netzbelastrungen entsprochen.²⁰ Allerdings ist diese Form der Preisbildung, so einleuchtend sie unter ökonomischen Aspekten sein mag, mit hohem Informationsaufwand und hohen Koordinationskosten verbunden. Sie ist daher nicht ohne weiteres auf andere Transportmärkte übertragbar. Die Zusatzkosten für diese Form der Preisbildung sind in Stromnetzen relativ gering, da jeder Netzbetreiber ohnehin die Stabilität seines Netzes jederzeit sicherstellen muß (Lastmanagement) und Kontrollkosten unabhängig vom Preisbildungsmechanismus anfallen.

²⁰ Hogan (1992).

5. Auktionen

Statt Preise zu regulieren, kann versucht werden, diese Preise von Märkten bestimmen zu lassen – wobei es allerdings wahrscheinlich ist, daß solche Märkte etabliert und organisiert werden müssen. Kurz: anstelle von Preisen können Märkte reguliert werden.

Hochorganisierte Märkte dieser Art sind Börsen (Double Auctions) und einseitige Auktionen. Die notwendige Frequenz der Auktionen bestimmt sich aus der Volatilität von Angebot und Nachfrage.

5.1 Marktformen

Transportauktionen können, abhängig von den zu transportierenden Gütern und der Zahl der Netzbetreiber, unterschiedlich organisiert sein. Als grobe Einteilung kann man von den in Tabelle 2 beschriebenen Fällen ausgehen.

Tabelle 2
Grobe Einteilung von Transportauktionen

	Eine Netzbetreibergesellschaft	Viele Leistungsbetreiber
Homogene Güter	(i) Anbieter und/oder Nachfrager bieten	(ii) Leistungsbetreiber und Anbieter und/oder Nachfrager bieten
Heterogene Güter	(iii) (Anbieter, Nachfrager)-Paare bieten	(iv) Leistungsbetreiber und (Anbieter, Nachfrager)-Paare bieten

Im Fall (i) kann die Netzbetreibergesellschaft gleichzeitig die Organisation der Auktionen vornehmen. Der Auktionator kann Anbieter und/oder Nachfrager nach dem Ziel der Maximierung der Renten gemäß Geboten und Transportkosten auswählen, i.e.

$$(M1) \quad \max \left(\text{Summe der Zahlungsbereitschaften der ausgewählten Nachfrager} - \text{Summe der Forderungen der Anbieter} - \text{Transportkosten} \right) \text{ mit Nebenbedingung Angebot} = \text{Nachfrage.}$$

Wenn die Nachfrage exogen gegeben ist, so kann die entsprechende Summe der Zahlungsbereitschaften einfach konstant gesetzt werden.²¹

²¹ Dies ist im Prinzip die Vorgehensweise im englischen Spotmarkt für Elektrizität. Die Entfernungskosten werden dort allerdings nicht explizit berücksichtigt.

Im Fall (iii) müssen sich Paare von Anbietern und Nachfragern bilden, die um Transporte von Punkt A nach Punkt B bieten. Es ist dabei gleich, wer von den beiden auf dem Transportmarkt agiert. Die Netzgesellschaft (der Auktionator) wird die Transporte auswählen, welche die

$$(M2) \quad \text{Summe der Zahlungsbereitschaften für ausgewählte Transporte} \\ - \text{Transportkosten}$$

maximieren.

Sind Netz und Transportmittel unterschiedlich (e.g. der Luftverkehr mit Slots und Start- und Landekapazitäten = Netz und mit Flugzeugen; die Eisenbahn mit Schienen und Zügen), so können die Besitzer der Fahrzeuge die Transportnachfrage bündeln und mit solchen Aggregaten in die Auktion gehen.

Im Fall (ii) nehmen nicht nur Anbieter und Nachfrager sondern auch die Leistungsbesitzer an der Auktion teil. Die Zielfunktion, die der Aktionator maximiert, ist

$$(M3) \quad \max \left(\begin{array}{l} \text{Summe der Zahlungsbereitschaften der ausgewählten Nachfrager} \\ - \text{Summe der Forderungen der ausgewählten Anbieter} \\ - \text{Transportkosten} \end{array} \right) \\ \text{mit Nebenbedingungen Güterangebot} = \text{Güternachfrage} \\ \text{Transportangebot} \geq \text{Transportnachfrage}.$$

Der Fall (iv) ist erkennbar eine Synthese aus (ii) und (iii).

5.2 Ein Beispiel

Nehmen wir an, die Situation (iii) läge vor. Die Preisregel (über die wir bisher noch nicht gesprochen haben) besagt: jeder akzeptierte Bieter zahlt gemäß seines Gebotes. Es seien Transporte von A nach D, von B nach E und von C nach F zu tätigen. Transportvolumen und Gewinne aus Transport werden angenommen als

Bieter	(A,D)	(B,E)	(C,F)
Transportvolumen	100	100	100
Gewinn aus Transport	1000	500	200.

Die kurzfristigen Transportkosten haben einen Verlauf, wie in Abbildungen 1 und 2 dargestellt. Im Gleichgewicht zahlen Bieter 1 = (A,D) und Bieter 2 = (B,E) je 400 = Zusatzkosten beim Übergang von 100 auf 200 transportierte Einheiten, Bieter 3 = (C,F) kommt nicht zum Zuge (siehe Abbildung 2).

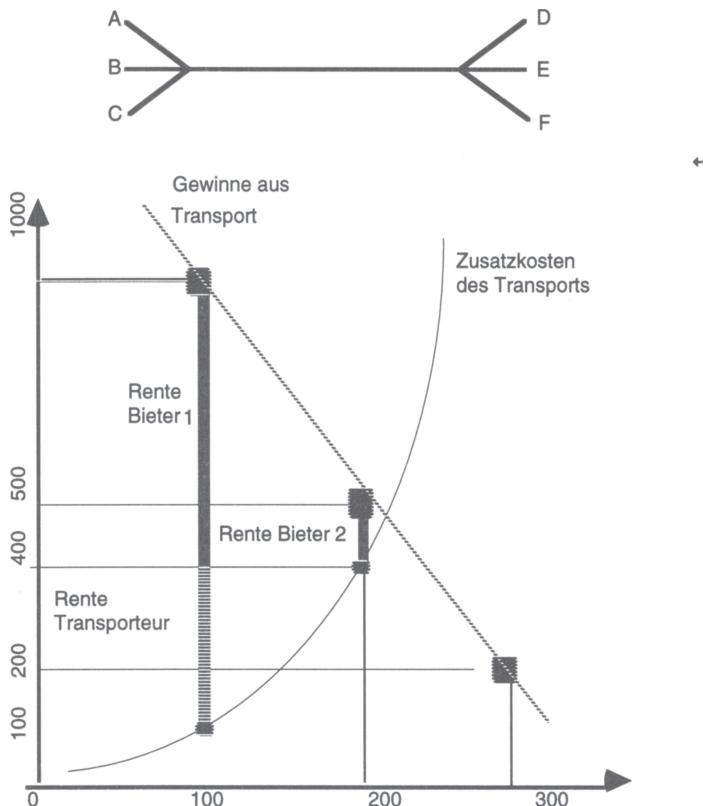


Abbildung 2: Das stilisierte Netz und die Gleichgewichtsgebote in einer Auktion mit der Preisregel „Die ausgewählten Bieter zahlen gemäß ihrer Gebote“, mit der Allokationsregel „Der Netzbetreiber wählt die Bieter aus, die seinen Gewinn maximieren“ und dem daraus resultierenden Gleichgewicht: „1 und 2 bieten 400, 3 bietet weniger als 400“.

5.3 Eine Verallgemeinerung für „kleine“ Bieter

Ein Bieter kann im folgenden der Anbieter oder Nachfrager eines homogenen Gutes oder ein Anbieter-Nachfrager-Paar oder der Besitzer eines Teilnetzes sein. „Kleine Bieter“ sind hier solche, die nur ein Gebot abgeben. Wir nehmen also an, daß die Besitzer von Teilnetzen nur ihr ganzes Teilnetz zu einem Gesamtpreis anbieten und nicht einzelne Strecken oder sogar Kapazitäten einzelner Strecken.

Die Preisregel besagt, daß jeder erfolgreiche Bieter den von ihm geforderten Betrag bekommt (den von ihm angebotenen Betrag zahlt). Der Auk-

tionator verfolge das Ziel, ein „Team“ T aus allen Bietern auszuwählen, das

$$(M4) \quad S(T) = \text{Summe der gezahlten Beträge}$$

- Summe der geforderten Beträge
- Summe der Transportkosten

maximiert unter den Nebenbedingungen „die Transporte sind technisch möglich“ und – bei homogenen Gütern – „Angebot = Nachfrage“.

Diese Struktur ist von einem der Autoren (*Bolle 1994*) untersucht worden.²² Das Hauptergebnis besagt, daß es immer Gleichgewichte in reinen (Biet-)Strategien gibt und daß alle Gleichgewichte effizient sind.

Die Transportkosten hängen vom ausgewählten Team von Anbietern und Nachfragern ab, womit zunehmende und abnehmende Grenzkosten außer in Spezialfällen nur noch intuitive Bedeutung haben. In solchen Spezialfällen läßt sich zeigen, daß zunehmende Grenzkosten zu eindeutigen Gleichgewichten führen, während abnehmende Grenzkosten ein Kontinuum von Gleichgewichten implizieren. Wir sollten uns an dieser Stelle daran erinnern, daß die kurzfristigen Grenzkosten des *Transportes auf einer Strecke* steigend sind (siehe Abbildungen 1 und 2).

Kann ein Bieter mehrere Gebote abgeben („großer“ Bieter), so haben wir es mit einer Multiple Bid Auction zu tun. Bei solchen Auktionen können wir im allgemeinen nicht mehr mit einem effizienten Ergebnis rechnen (*Bolle 1993*).

5.4 Experimente

Es gibt einige Experimente, die die *Struktur* der oben skizzierten Auktionen aufweisen, nicht notwendig allerdings deren *Einkleidung* als Transportprobleme. Zwei zentrale Fragen lassen sich mit Hilfe solcher Experimente angehen. Erstens: Wie ist es mit der Effizienz der aus den Auktionen ermittelten Allokationen bestellt? Zweitens: Was geschieht in Fällen, in denen einzelne Marktteilnehmer aufgrund ihrer strategischen Position den größten Teil der Effizienzgewinne für sich beanspruchen können? Letzteres tritt in Netzen sehr leicht auf.

Die Standardkritik an solchen und anderen Experimenten ist, daß die notwendig sehr beschränkten Anreize in Experimenten (i.a. Marktbeträge) wenig zur Erklärung „realen“ Verhaltens beisteuern können, bei denen es um Millionen gehen kann. Auch zeigt sich in Experimenten immer wieder, daß der Kontext (Einkleidung, „Framing“) einen ‚mal mehr ‚mal weniger großen Einfluß auf das Ergebnis hat. Dennoch wendet sich eine zuneh-

²² Dabei werden noch weitere Annahmen benötigt.

mende Anzahl von Ökonomen solchen Experimenten zu – schließlich sind die beobachteten Fähigkeiten, Verhaltensmuster und -normen nicht völlig losgelöst von ihren Entsprechungen im „wirklichen Leben“.

Wir wollen im weiteren kurz auf die Ergebnisse von vier Experimenten eingehen, von denen nur das letzte auch die Einkleidung als Transportmarkt aufweist.

Ultimatum Bargaining (*Güth et al. 1982*; viele Folgeexperimente). Zwei Spieler entscheiden über die Aufteilung eines gewissen Geldbetrages g . Spieler 1 fordert für sich a_1 mit $0 \leq a_1 \leq g$. Wenn Spieler 2 dieser Forderung zustimmt, dann erhält 1 den Betrag a_1 und 2 den Betrag $a_2 = g - a_1$. Für unseren Transportmarkt heißt das entweder, daß Spieler 1 der Netzbetreiber ist, der den Transportpreis festlegt, oder daß Spieler 1 den einzigen Nachfrager nach Transportleistung in der oben beschriebenen Auktion darstellt. g ist jeweils der Effizienzgewinn, wenn der Transport zustande kommt.

Das eindeutige Gleichgewicht in diesem Spiel ist die Forderung von g (oder $g - \varepsilon$, $\varepsilon = \text{kleinste Geldeinheit}$) durch Spieler 1. Spieler 2 akzeptiert die Forderung und erhält nichts oder fast nichts. Im Experiment stellt sich heraus, daß dieses Gleichgewicht für Spieler 2 (und vielleicht auch für Spieler 1) nicht akzeptabel ist. Hohe Forderungen von 1 werden abgelehnt. Deshalb(?) erhebt 1 i.a. auch nur Forderungen $a_1 = g/2$ bis höchstens $2g/3$. Die Akzeptanzraten (ungefähr gleich Effizienz) im ersten Ultimatum Bargaining Experiment (*Güth et al. 1982*) waren 79 - 83%. In einer Vielzahl von Folgeexperimenten (für einen Überblick siehe z.B. *Güth und Tietz 1982*) sind ähnliche Werte gefunden worden.

Team Selection mit zwei Bietern (*Bolle 1994*). Wir haben zwei Bieter 1 und 2, die z.B. ähnlich wie in Abbildung 2 bestimmte Beträge für einen Transport bieten. Der Auktionator kann dann zwischen den Teams \emptyset , $\{1\}$, $\{2\}$ und $\{1, 2\}$ wählen. Seine Zielfunktion ist (M4). Wie im Ultimatum Bargaining Experiment zeigt es sich, daß spieltheoretische Gleichgewichte nur geringen Einfluß auf das Verhalten der Versuchspersonen haben. Neben Gerechtigkeitsnormen spielt nun aber auch schon die Bewältigung der beträchtlich schwierigeren strategischen Situation eine Rolle. Vom maximal erreichbaren Effizienzgewinn werden 80 - 89% erreicht.

Space Station Problems (*Banks et al. 1989; Plott/Porter ohne Datum*). Die Einkleidung des Experiments sieht die Beschildung einer Raumstation mit Projekten vor, von denen jedes die knappen Ressourcen Arbeitszeit, Energie und Masse benötigt. Sechs bis sieben Bieter können jeweils ein Projekt vorschlagen. Offenbar ist dieses Problem von seiner Struktur her wieder ein Team Selection Problem. Übertragen auf den Transportmarkt könnte man die drei Ressourcen einfach als Leitungen oder allgemeiner als

Netzteile interpretieren. Man kann aber auch konstatieren, daß bei der Abwicklung realer Transporte neben dem reinen Transport auch Be- und Entladekapazitäten bereitstehen müssen und daß das Team der Transporte die Verwaltungskapazitäten des Transporteurs nicht übersteigen darf.

Es werden vier Marktstrukturen/Preisregeln getestet. Erstens die Etablierung von drei dezentralen Märkten für die drei Ressourcen; zweitens durch den Auktionator NASA festgelegte Preise (Posted Prices); drittens der sogenannte AUSM²³, der gleich dem oben beschriebenen Team-Selection-Spiel mit gewissen Revisionsmöglichkeiten ist; viertens (nur bei Banks et al.) eine Vickrey-Auktion.²⁴ Die Effizienz der Ergebnisse als Prozentsatz der Maximal erzielbaren Gewinne ist in Tabelle 3 angegeben.

Tabelle 3
Die Effizienz verschiedener Marktorganisationen in zwei Experimenten

	Effizienz	
	Banks et al. (1989)	Plott/Porter (ohne Datum)
Dezentrale Märkte	66 %	86 - 91 %
Posted Prices	64 %	60 - 63 %
AUSM	78 - 81 %	79 %
Vickrey Auction	78 %	-

Bemerkenswert ist, daß der bei sechs bis sieben Spielern doch schon recht komplizierte AUSM auf so gute Effizienzwerte kommt. Die sehr unterschiedlichen Ergebnisse im Fall dezentraler Märkte sind erklärbungsbedürftig.

Gas Transportation Network (*Rassenti* et al., 1991). In diesem Experiment geht es um Gastransporte in einem Netz mit neun Leitungen und acht Knotenpunkten. Im Basisexperiment gibt es sechs Käufer, drei Produzenten und drei Spieler, die sowohl Produktionskapazitäten als auch Leitungen besitzen. Die Auktion verläuft wie oben unter (ii) beschrieben. Es gibt aber wie bei AUSM Revisionsmöglichkeiten der Gebote. Die Effizienz des ex-

²³ Adaptive User Selection Mechanism.

²⁴ Dies ist eine Verallgemeinerung der von Vickrey (1961) vorgeschlagenen Preisregel. Sei T das gemäß der offebarten Zahlungsbereitschaften und Kosten wohlfahrtsmaximierende Team. Eine Wohlfahrtsdifferenz zwischen $T-\{i\}$ und dem optimalen Team, das aus der Erweiterung von $T-\{i\}$ entsteht. Unter dieser Preisregel ist es eine dominante Strategie für alle Bieter, die Wahrheit zu sagen.

perimentellen Marktes liegt bei 85 %, was angesichts des komplizierten Netzes ein erstaunlich guter Wert ist.

In einem veränderten Experiment wurde die Quasimonopolposition einiger Leitungsbetreiber dadurch geschwächt, daß die Kapazität ihrer Leitungen jeweils auf zwei Spieler aufgeteilt wurde. Dieser Zuwachs an Kompetitivität schlug sich in einer Erhöhung der Effizienz auf 89 % nieder.

Kann man aus den Experimenten für die Praxis lernen? Wie bereits in den Vorbemerkungen zu diesem Abschnitt gesagt, ist die Botschaft, die man aus Experimenten erhält, weder klar noch sicher. Dennoch ergeben sich einige Anhaltspunkte. Die geschilderten Experimente sprechen dafür, daß

- Auktionsmärkte auch für Transporte in Netzen funktionieren können, wenn auch mit (unvermeidlichen?) Reibungsverlusten,²⁵
- strategische Vorteile (Quasimonopole) oft nicht vollständig genutzt werden können.

6. Was tun?

Der Transport in Netzen stellt (beinahe) unbestritten ein natürliches Monopol dar. Traditionell ist der Transport und die Güterverteilung über Netze in integrierten Monopolen konzentriert. (Ausnahmen sind das Straßen- und Kanalnetz, die allen Nutzern steuer- und manchmal auch gebührenfinanziert zur Verfügung stehen.) Lohnt es sich angesichts der theoretisch wie empirisch unklaren Situation eigentlich, das sichere Regime solcher i.a. preisregulierten Monopole zu verlassen und sich auf Regime mit Güter- und möglicherweise auch Transportmärkten einzulassen? Der Erfolg ist nicht garantiert und (was die Vertreter der Monopole immer wieder betonen) die Risiken sind hoch. Hierzu kann eigentlich nur festgestellt werden, daß einer politischen Entscheidung selten ein verlässliches theoretisches und empirisches Fundament unterliegt.

Die vielleicht umfangreichste und radikalste Umgestaltung eines öffentlichen Versorgungssystems ist die Neuorganisation des Elektrizitätsmarktes in England und Wales. Diese Umgestaltung im Hinblick auf maximal möglichen Wettbewerb (unter wesentlicher Einbeziehung von Auktionen) ist praktisch ohne irgendein theoretisches Fundament erfolgt. Erst im Anschluß an die Umgestaltung hat man begonnen, adäquate Analyseinstrumente zu entwickeln. Der (empirisch zu ermittelnde) Erfolg des Systems ist von seiner Gesamtbewertung her noch unklar, große Produktivitätsfortschritte

²⁵ Wir wissen wenig über die Effizienz realer Märkte – insofern ist die Frage, ob 80% viel oder wenig sei, schwer zu entscheiden.

scheint es aber zumindest im Bereich der Erzeugung zu geben. Die Preisbildung für Transporte im Elektrizitätsnetz ist allerdings noch grob und unzureichend.

Im europäischen Rahmen strebt die Kommission Durchleitungsverpflichtungen für Eigentümer von Gas- und Elektrizitätsnetzen an. Wegen des starken Widerstandes aus den betroffenen Industrien ist allerdings bisher keine funktionsfähige Regelung in Sicht. Wenn man sich vor Augen hält, wie lange der entsprechende Prozeß unter wesentlich günstigeren Bedingungen in der britischen Gasindustrie gedauert hat und welche einschneidenden Maßnahmen letztlich ergriffen werden mußten, kann von der Vorgehensweise der Kommission kein großer Erfolg erwartet werden. Vielleicht sollte man sich speziell in der Gasindustrie auf den Aufbau eines Auktionsmarktes konzentrieren, bei dem allerdings zu sichern wäre, daß die Kapazität der vielen Rohrleitungen, die einem Konsortium gehören, entsprechend den Anteilen der Eigner *unabhängig* angeboten würde.

Die zentrale Aufgabe von Märkten besteht in der Verarbeitung von Informationen und deren Komprimierung in Form von Preisen. Transportmärkte, die diese Aufgabe schnell und effizient lösen, können sich wegen der Economies of Scope in Netzen nicht spontan bilden. Dieses Papier hat kein Patentrezept entwickelt, solche Märkte zu organisieren. Wir glauben allerdings, daß die besten Chancen in einer börsenähnlichen Institution liegen könnten.

Literatur

- Banks, J. S./Ledyard, J. O./Porter, D. P. (1989): Allocating Uncertain and Unresponsive Ressources: an Experimental Approach, in: RAND Journal of Economics, 1, 1989.
- Barker, J. U. Jr. (1982): Electric Energy Brokering, in: Public Utilities Fortnightly, Feb. 4th, 1982, S. 28 - 36.
- Baumol, William J./Panzar, John C./Willig, Robert D. (1982): Contestable Markets and the Theory of Industry Structure, New York, Harcourt Brace Jovonavich, 1982.
- Blankart, Charles, B./Knieps, Günther (1992): Netzökonomik, in: Jahrbuch für neue Politische Ökonomie, 11, 1992, S. 73 - 86.
- Bohn, Roger, E. (1990): Spot Prices and Spot Markets for Evaluation and Control of Power Systems, Euro-American Conference on Organizing and Regulating Electric Systems in the Nineties.
- Bolle, F. (1991): Necessary Conditions for Efficient Multiple Bid Auctions – Revised Version 1993, Frankfurt/Oder.
- (1992): Supply Function Equilibria and the Danger of Tacit Collusion – The Case of Spot Markets for Electricity, in: Energy Economics 1992, S. 94 - 102.

³ Todt

- (1993): Team Selection – Factor Pricing with Discrete and Inhomogenous Factors, in: Mathematical and Social Sciences, (forthcoming), 1993.
 - (1994): Team Selection – An Experimental Investigation, in: Journal of Economic Psychology, (forthcoming), 1994.
- DeCanio, Stephen J./Frech, H. E.* (1993): Vertical Contracts: A Natural Experiment in Gas Pipeline Regulation, in: JITE – Journal of Institutional and Theoretical Economics, 149(2), 1993, S. 370 - 392.
- Demsetz, H.* (1968): Why Regulate Utilities?, in: Journal of Land Economics, 11, 1968, S. 55 - 65.
- Deregulierungskommission* (1991): Marköffnung und Wettbewerb, Zweiter Bericht: Die Stromwirtschaft, Deregulierungskommission, 1991.
- Doyle, Chris/Maher, Maria* (1992): Common Carriage and the Pricing of Electricity Transmission, in: Energy Journal, 13, 3, 1992, S. 63 - 94.
- Ellig, Jerry/Giberson, Michael* (1993): Scale, Scope and Regulation in the Texas Gas Transmission Industry, in: Journal of Regulatory Economics, 5, 1993, S. 79 - 90.
- Falcone, Charles, A.* (1990): Transmission in the Nineties, in: Public Utilities Fortnightly, December 6th, 1990, S. 29 - 32.
- Funk, Cara* (1990): Durchleitung auf dem amerikanischen Erdgasmarkt. Zeitschrift für Energiewirtschaft 4, 1990.
- Güth, W./Schmittberger, R./Schwarze, B.* (1982): An Experimental Study of Ultimatum Bargaining, in: Journal of Economic Behaviour and Organizations, 3, 1982, S. 367 - 388.
- Hogan, William W.* (1990): An Efficient Concurrent Auction Model for Firm Natural Gas Transportation Capacity, Cambridge, Energy and Environmental Policy Center, John F. Kennedy School of Government Harvard University.
- (1992): Contract Networks for Electric Power Transmission, in: Journal of Regulatory Economics, 4, 1992, S. 211 - 242.
 - (1992): An Efficient Concurrent Auction Model for Firm Natural Gas Transportation Capacity, in: INFOR, 30, 3, 1992, S. 240 - 256.
- Hunt, Sally/Shuttleworth, Graham* (1993): Electricity Transmission Pricing, in: Utilities Policy, April, 1993, S. 98 - 110.
- Klemperer, P. D./Meyer, M. A.* (1989): Supply Function Equilibria in an Oligopoly under Price Uncertainty, in: Econometrica, 57, 1989, S. 1243 - 1277.
- Knieps, Günther* (1992): Konkurrenz auf europäischen Eisenbahnnetzen, in: Jahrbücher für Nationalökonomie und Statistik, 209, 3 - 4, 1992, S. 283 - 290.
- (1993): Competition, Coordination and Cooperation – A Disaggregated Approach to Transportation Regulation, in: Utilities Policy, July, 1993, S. 201 - 207.
- Laffont, Jean-Jacques/Tirole, Jean* (1993): A Theory of Incentives in Procurement and Regulation, Cambridge MA, London, MIT Press, 1993.

- Leibenstein, Harvey (1966): Allocative Efficiency vs. „X-efficiency“, in: American Economic Review, 56, June, 1966, S. 392 - 415.*
- Moyer, Charles R. (1993): The Impending Restructuring of the Electricity Supply Industry, in: Business Economics, October, (1993), S. 40 - 44.*
- Ofgas and British Gas (1993): Gas Transporation and Storage – Joint Consultation Document, Ofgas and British Gas, 1993.*
- Plott, Ch. R./Porter, D. P. (undated): An Experiment with Space Station Pricing Policies.*
- Rassenti, S. J./Reynolds, S. S./Smith, V. L. (1991): Cotenancy and Competition in an Experimental Auction Market for Natural Gas Pipeline Networks, University of Arizona, Discussion Paper.*
- Schweppes, Fred C. et al. (1988): Spot Pricing of Electricity, The Kluwer International Series in Engeneering and Computer Science – Power Electronics and Power Systems, Dordrecht, Kluwer, 1988.*
- Vaubel, Roland (1992): Bemerkungen zu dem Referat „Netzökonomik“ von Charles B. Blankart und Günther Knieps, in: Zeitschrift für die gesamte Staatswissenschaft, 11, 1992, S. 88.*
- Vickrey, W. (1961): Counterspeculation, Auctions, and Competitive Sealed Tenders. Journal of Finance 16, pp. 8 - 37.*
- Yarrow, George (undated): Simplified Procedures for Network Pricing, Oxford.*

Zusammenfassung der Diskussion

Die Diskussion konzentriert sich im wesentlichen auf die beschriebenen Auktionsexperimente und ihre Übertragbarkeit auf „reales“ Verhalten. Dabei dreht es sich um den Problemkreis theoretisch zu ermittelnder Gleichgewichtslösungen, Informationsstand der Bieter und deren Verhalten. Die Experimente haben gezeigt, daß eine große Lücke zwischen den theoretisch zu erwartenden Lösungen und dem beobachtbaren Verhalten besteht. Auch eine Diskussion der Gleichgewichtsstrategien mit Teilnehmern an Experimenten vor deren Beginn ändert daran nichts. Statt dessen spielen Normen gerechter Verteilungen und fairen Verhaltens bei den Entscheidungen eine wichtige Rolle.

Außerdem werden die Größenordnung der Auszahlung bzw. die Relevanz ihrer Differenzen als denkbare Einflußgröße auf das Verhalten angesprochen. Auch die Möglichkeit durch Anwendung der Lösungskonzeption Rubinstains zu ganz anderen Gleichgewichten zu gelangen wird diskutiert.

Die Frage nach möglichen Lerneffekten der Beteiligten wird dahingehend beantwortet, daß solche entweder gewollt sind, oder nicht auftreten können, da die betrachten Auktionen nur einmal durchgeführt werden.

3*

Viele der anderen bekannten Auktionsformen können nicht angewandt werden, da es für den Auktionator bei der Bestimmung der optimalen Lösung des Team Selection Problems unumgänglich ist, sämtliche Gebote zu kennen.

Regionale öffentliche Güter

Von *Hans Wiesmeth*, Dresden

Abstract

Öffentliche Güter spielen für die wirtschaftliche Entwicklung einer Region, aber auch für die überregionale Zusammenarbeit eine grundlegende Rolle. Probleme, die in diesem Zusammenhang auftreten und zu lösen sind, berühren insbesondere die Frage nach der optimalen Menge der bereitzustellenden öffentlichen Güter, sowie deren Finanzierung. In der Praxis werden diese Probleme auf unterschiedliche Art und Weise behandelt und gelöst, so daß sich die Frage nach der Möglichkeit einer tieferen theoretischen Bergründung der einen oder anderen Vorgehensweise stellt. Für sogenannte regionale öffentliche Güter, charakterisiert durch regionale Überfüllungseffekte, wird ein Kernäquivalenztheorem das Konzept des Lindahl-Gleichgewichts theoretisch untermauern.

1. Einleitung

Bei einem wirtschaftlichen oder politischen Zusammenschluß regional unterschiedlicher Gebiete sind öffentliche Güter oft in einem entscheidenden Ausmaß und an entscheidender Stelle beteiligt. In vielen Fällen führt so die Möglichkeit einer umfassenderen Bereitstellung eines öffentlichen Gutes zu einem Zweckbündnis. Erkennbar wird dies am Beispiel der NATO mit dem Angebot des öffentlichen Gutes „Verteidigung“, aber auch bei regionalen Zweckbündnissen, bei denen Güter mit wesentlichen Eigenschaften öffentlicher Güter erstellt werden sollen, wie etwa im Fall eines Abwasserzweckverbandes. Andererseits ermöglicht ein überregionaler Zusammenschluß häufig erst die Bereitstellung bestimmter öffentlicher Güter, ist in diesem Sinne also notwendig für das Angebot an diesen Gütern. Das Beispiel der EG zeigt dies in Hinblick auf die durch den Zusammenschluß ermöglichte Anpassung und Angleichung der wirtschaftlichen und gesetzlichen Rahmenbedingungen, in räumlich kleineren Gebieten kann etwa das Beispiel eines Verkehrsverbundes herangezogen werden. Zu berücksichtigen ist weiter, daß die Stabilität einer wirtschaftlichen oder politischen Union auch davon abhängt, inwieweit es gelingt, sich über ein geeignetes Angebot an öffentlichen Gütern einig zu werden. Mit einem einfach strukturierten Modell erörtert und untersucht so *Feinstein* (1992) diese Frage am Beispiel der EG. Insgesamt aber ist festzuhalten, daß dieser wich-

tige Aspekt bis jetzt kaum Gegenstand einer umfassenderen empirischen Analyse geworden ist.

Dieses Problem der Ausgestaltung und der Menge der für einen regionalen Zusammenschluß bereitzustellenden öffentlichen Güter sowie deren Finanzierung ist offensichtlich eng verknüpft mit dem Aspekt einer geeigneten, von den Mitgliedern der Union akzeptierten Vorgehensweise bei der Aggregation von Präferenzen bezüglich der relevanten öffentlichen Güter. Aus der Sicht der ökonomischen Theorie stellt sich somit die Frage nach einem effizienten und gerechten, einem „sozialverträglichen“ Allokationsmechanismus. Ein Blick auf die Praxis zeigt allerdings, daß dieses Allokationsproblem sehr unterschiedlich angegangen und „gelöst“ werden kann. Für den Bereich der NATO beispielsweise machen *Olson/Zeckhauser* (1966) auf eine empirisch nachprüfbare ungleiche Verteilung der Kosten zu Gunsten der kleineren Nationen aufmerksam. Unterstellt man nun ein nichtkooperatives Verhalten der Mitgliedsländer in bezug auf die Bereitstellung des öffentlichen Gutes „Verteidigung“, so zeigt der durch das Nash-Cournot-Gleichgewicht bestimmte Allokationsmechanismus, daß größere Mitgliedsländer, die unter den gegebenen Annahmen dem öffentlichen Gut Verteidigung einen höheren Wert zumessen, einen überdurchschnittlich hohen Anteil an den Kosten zu tragen haben. Eine ungleiche Verteilung der Kosten ist deshalb unvermeidlich, weil die größeren Nationen eine vergleichsweise schwache Verhandlungsposition haben, um die kleineren Nationen zu höheren Beiträgen zu veranlassen (vgl. auch *Weber/Wiesmeth* [1991a]). Anders verhält sich die Sachlage in der EG. Die gemeinsame Politik wird dort nach festen Regeln finanziert, über die zu produzierenden Güter, insbesondere auch die öffentlichen Güter, entscheiden jedoch im wesentlichen Kommission und Ministerrat der EG. Da in wichtigen Fragen einstimmige Beschlüsse zu fassen sind, werden gelegentlich „Pakete“ zur Abstimmung gestellt. Dieses „Package Dealing“ ermöglicht dann auch die Zustimmung derjenigen Mitgliedsländer, die ansonsten durch eine bestimmte einzelne Beschußvorlage benachteiligt würden. Eine ähnliche Überlegung gilt für die durch die Einheitliche Europäische Akte erweiterten Möglichkeiten einer Abstimmung mit qualifizierter Mehrheit (vgl. dazu auch *Weber/Wiesmeth* [1991b]).

Diese und andere in der Praxis erprobten Mechanismen führen in der Regel zu ökonomisch ineffizienten Ergebnissen (vgl. etwa *Atkinson/Stiglitz* [1980], Lecture 16). Weiter ist im Fall einer Mehrheitsabstimmung der durch das Verfahren festgelegte Allokationsmechanismus selbst gelegentlich Ziel der Kritik. Für die EG zeigt sich dies etwa am Beispiel der Diskussionen über den Vertrag von Maastricht. Damit genügen diese praxisrelevanten Mechanismen den oben erwähnten Anforderungen nur sehr bedingt.

Die ökonomische Theorie hat nun verschiedene Lösungskonzepte hervorgebracht, die in diesem Zusammenhang dem Effizienzkriterium genügen und die auch dem Gerechtigkeitsaspekt in gewisser Weise entgegenkommen, die aber aus verschiedenen anderen Gründen in der Regel als nicht praktikabel angesehen werden. Dazu zählt insbesondere das Konzept des *Lindahl-Gleichgewichts*, das vor allem wegen Problemen mit der Anreizverträglichkeit als „empirisch nicht relevant“ eingestuft wird (vgl. etwa *Roberts* [1974]). Vergleichbare Überlegungen gelten dann auch für *Bewertungsgleichgewichte* (vgl. *Mas-Colell* [1980]) und für *Kostenverteilungsgleichgewichte* (vgl. *Mas-Colell/Silvestre* [1989] und *Weber/Wiesmeth* [1992]), die in der Literatur als Verallgemeinerungen des Lindahl-Gleichgewichts analysiert werden und die auch zu effizienten Allokationen führen.

Trotz der eben erwähnten Unzulänglichkeiten des Lindahl-Gleichgewichts ist ein Aspekt zu berücksichtigen, der es von den anderen oben angesprochenen Lösungsverfahren unterscheidet. Der zugrundeliegende Allokationsmechanismus beruht auf dem Preismechanismus, wenn auch in einer gegenüber den privaten Gütern abgewandelten Form. Dieser Aspekt bringt das Lindahl-Gleichgewicht in eine gewisse Nähe zum herkömmlichen Konzept des Marktgleichgewichts. Um diese Parallele für das Lindahl-Gleichgewicht ausnützen zu können, stellt sich daher zunächst die Frage, was das Marktgleichgewicht, abgesehen von der bekannten Effizienz-eigenschaft, auszeichnet.

Diese Frage nach einer theoretischen Untermauerung des Marktmechanismus findet eine Antwort im *Kernäquivalenztheorem*, das im Fall einer „großen“ Ökonomie eine kooperative Charakterisierung des Konzepts eines kompetitiven Gleichgewichts ermöglicht (vgl. *Aumann* [1964], *Hildenbrand* [1968]). Genauer entsprechen die Gleichgewichtsallokationen den Elementen des Kerns, die nicht-kooperativen Marktlösungen stimmen mit den kooperativen Kernlösungen überein. In diesem Sinne kann der Preismechanismus dann als demokratisch, als sozialverträglich bezeichnet werden, obwohl natürlich zu beachten ist, daß das Ergebnis dieses Allokationsmechanismus, das kompetitive Gleichgewicht, im allgemeinen wesentlich von der vorgegebenen Verteilung der Erstausstattungen abhängt (vgl. *Hildenbrand/Kirman* [1976]).

Dieses Ergebnis kann nun zunächst nicht auf das Lindahl-Gleichgewicht übertragen werden. Beispiele von *Milleron* (1972) und *Muench* (1972) zeigen, daß der Kern einer Ökonomie mit öffentlichen Gütern im allgemeinen sehr viel größer ist als die Menge der Lindahl-Gleichgewichte, jedenfalls dann, wenn es keine Überfüllungseffekte gibt. Kleinere Koalitionen von Individuen sind dann benachteiligt und können eine vorgegebene Allokation nur in Ausnahmefällen verbessern. Eine kooperative Charakterisierung der Lindahl-Gleichgewichte scheitert demnach an der Kostendegres-

sion, die im Fall reiner öffentlicher Güter mit einer größeren Zahl an der Finanzierung beteiligter Individuen verbunden ist.

Allerdings vermindert sich dieser Nachteil kleinerer Koalitionen, wenn der Konsum der öffentlichen Güter durch Überfüllungseffekte beeinträchtigt wird. Dies ist der Fall für die hier zu betrachtenden *regionalen öffentlichen Güter*. Dazu ist ein System unterschiedlicher, jedoch in sich homogener Regionen vorgegeben, jeweils bewohnt von einer großen Zahl von Individuen. Ein öffentliches Gut, wie etwa ein öffentliches Schulsystem, ein Umweltgut oder ein Verteidigungssystem, das nun überregional, also für die Gemeinschaft dieser Regionen bereitgestellt werden soll, wird üblicherweise dezentral in den einzelnen Regionen produziert. Mit diesem lokalen, dezentralen Angebot sind aber häufig „Spillover-Effekte“ verbunden (vgl. etwa Williams [1966] oder Pauly [1970]), so daß die Bewohner einer Region in einem gewissen Umfang auch vom Angebot der anderen Regionen profitieren. Entscheidend ist aber nun, daß der Konsum der regionalen öffentlichen Güter durch *lokale* Überfüllungseffekte beeinträchtigt werden kann. Genauer bedeutet dies, daß eine Zunahme der Bevölkerung in einer gegebenen Region cet. par. den Konsum der öffentlichen Güter in dieser Region berührt, wohingegen eine Änderung der Bevölkerung in einer anderen Region cet. par. möglicherweise keinen oder nur einen geringen derartigen Effekt hat.

Es zeigt sich nun, daß man für diese regionalen öffentlichen Güter den Aspekt der Kernäquivalenz zur Charakterisierung der Lindahl-Gleichgewichte wieder aufgreifen kann. Grundlage für diese Überlegungen ist eine Arbeit von Vasil'ev/Weber/Wiesmeth (1992) zur Kernäquivalenz von Lindahl-Gleichgewichten in Ökonomien mit sogenannten *halb-öffentlichen Gütern*. In diesem Rahmen erhält man so über die Kernäquivalenz eine kooperative Charakterisierung der Lindahl-Gleichgewichte, durchaus vergleichbar mit den entsprechenden Ergebnissen für große Ökonomien mit privaten Gütern. Der mit dem Lindahl-Gleichgewicht verbundene Allokationsmechanismus gewinnt damit für regionale öffentliche Güter wieder an Bedeutung. Diese Bedeutung wird noch durch die Tatsache verstärkt, daß die oben erwähnte Anreizproblematik in diesen regional strukturierten Ökonomien nur eine untergeordnete Bedeutung besitzt.

Die ökonomische Relevanz der Kernäquivalenz wird in den folgenden Abschnitten immer wieder betont werden. Zunächst aber werden zusammen mit dem Begriff eines regionalen öffentlichen Gutes die grundlegenden Modellannahmen eingeführt, bevor im dann folgenden Abschnitt das Konzept des Lindahl-Gleichgewichts sowie das Kernkonzept genauer betrachtet werden. Anschließend wird das Äquivalenztheorem formal analysiert.

2. Regionale öffentliche Güter

Gegeben ist ein ökonomisches System E mit n verschiedenen, jedoch in sich homogenen Regionen. $N = \{1, \dots, n\}$ bezeichnet die Menge dieser Regionen. Eine Region i , $i \in N$, wird jeweils bewohnt von einem Kontinuum identischer Konsumenten. Die Menge T der Konsumenten wird genauer als Vereinigungsmenge $T_1 \cup \dots \cup T_n$ von abgeschlossenen, paarweise disjunkten Intervallen T_i der Länge μ_i mit $0 < \mu_i < \infty$ für $i \in N$ aufgefaßt. T_i wird folglich mit der Menge der Konsumenten assoziiert, die Region i bewohnen. Insgesamt entspricht die formale Beschreibung der Ökonomie E der einer *Typenökonomie* (vgl. Hildenbrand/Kirman [1976]), bei der die unterschiedlichen Typen von Individuen den verschiedenen Regionen zugeordnet sind. Im Sinne der Regionalwissenschaft ist die Ökonomie E dann nach dem *Homogenitätsprinzip* gegliedert. Dies entspricht einer gängigen Vorgehensweise in Modellen, die einen wirtschaftlichen oder politischen Zusammenschluß unterschiedlicher Regionen zur Grundlage haben (vgl. etwa Olson/Zeckhauser [1966], Weber/Wiesmeth [1991 b]).

Es gibt l private Güter und m öffentliche Güter, bezeichnet mit den Indizes $L = \{1, \dots, l\}$, bzw. $M = \{l + 1, \dots, l + m\}$. Die öffentlichen Güter werden zusätzlich klassifiziert nach der Region ihrer Herkunft. Bei einem Güterbündel $x = (x_L, x_M^1, \dots, x_M^n)$ entspricht also x_M^i für $i \in N$ der Menge der einzelnen öffentlichen Güter, die in Region i produziert und bereitgestellt werden. Durch die mögliche Existenz von *Spillover-Effekten* wird nun angenommen, daß einem Konsumenten der Region i , $i \in N$, öffentliche Güter im Umfang von

$$x_M^{(i)} = g_1^{(i)} \cdot x_M^1 + \dots + g_n^{(i)} \cdot x_M^n$$

zur Verfügung stehen, wobei die konstanten Koeffizienten $g_j^{(i)} \in [0, 1]$ für $i, j \in N$ die Stärke der Spillover-Effekte zum Ausdruck bringen¹ (vgl. auch Williams [1966]). Gilt etwa $g_j^{(i)} = 0$ für alle $i, j \in N$ mit $i \neq j$, und $g_i^{(i)} = 1$, $i \in N$, so ist die Wirkung von x_M^i offenbar auf die Region i beschränkt, man erhält so den Fall eines in bezug auf die Region i lokalen öffentlichen Gutes. Analog ergibt sich ein reines öffentliches Gut für $g_j^{(i)} = 1$ für alle $i, j \in N$.

Die exakte Definition des Begriffs eines regionalen öffentlichen Gutes beruht auf einigen weiteren grundlegenden Annahmen an die Struktur der Ökonomie: Ein Individuum $t \in T_i$, also ein Bewohner der Region i , besitzt die Erstausstattung $w^i = (w_L^i, w_M^i) \in \mathbb{R}_{++}^l \times \{0\}$. Es gibt folglich keine

¹ Der Einfachheit halber wird hier unterstellt, daß für alle hier betrachteten öffentlichen Güter dieselben Spillover-Effekte gelten. Die formale Analyse würde durch eine entsprechende Verallgemeinerung dieser Annahme nicht beeinträchtigt.

Erstausstattung bezüglich der öffentlichen Güter. Weiter ist Konsument $t \in T_i$ charakterisiert durch die Konsummenge $X^t = \mathbb{R}_+^{l+m}$ und durch die stetige, quasi-konkave und streng monoton wachsende Nutzenfunktion $v^i : X^t \rightarrow \mathbb{R}$.

Die Produktionsmöglichkeiten einer jeden Region der Ökonomie E werden durch die aggregierte Produktionsmenge $Y \subseteq \mathbb{R}^{l+m}$ angegeben. Y ist ein konvexer Kegel mit $0 \in Y$ und mit $Y_M \subset \mathbb{R}_+^m$. Es wird also angenommen, daß alle Regionen der Ökonomie gleichermaßen Zugriff auf diese Technologie haben. Man beachte, daß Y insbesondere auch durch konstante Skalenerträge charakterisiert ist. Diese Eigenschaft ist notwendig, um die Kerneigenschaft der Lindahl-Gleichgewichte sicherzustellen (vgl. etwa Kaneko [1976]).

Ein erstes grundlegendes Konzept ist das einer für eine Koalition D erreichbaren Allokation, einer sogenannten D -erreichbaren Allokation. Wie üblich ist dabei eine *Koalition* eine nichtleere, meßbare Teilmenge von T . Eine Koalition D wird dann beschrieben durch ihr *Profil* $\pi_D = (\pi_D^1, \dots, \pi_D^n)$ mit $\pi_D^i = \mu(D \cap T_i)$ der Menge der Individuen aus Region i , die der Koalition D angehören. Mit $s(D)$ wird schließlich der *Träger* von D bezeichnet, $s(D) := \{i \in N : \pi_D^i > 0\}$.

Definition 2.1 Ein Vektor $x^D = (x_L^1, \dots, x_L^n; x_M^1, \dots, x_M^n) \in \mathbb{R}_+^{n(l+m)}$ ist eine D -erreichbare Allokation², falls für jedes $i \in N$ ein $y^i \in Y$ existiert mit:

$$\sum_{i \in s(D)} \pi_D^i x_L^i = \sum_{i \in s(D)} \pi_D^i w_L^i + \sum_{i \in N} y_L^i, \quad \text{und} \quad x_M^i = y_M^i \quad \text{für alle } i \in N.$$

Zu beachten ist bei dieser Definition, daß etwa im Fall einer T -erreichbaren Allokation x^T die Bewohner der Region i , $i \in N$, die Mengen $\mu_i \cdot (w_L^i - x_L^i)$ an den privaten Gütern zur Produktion der öffentlichen Güter beitragen. Tatsächlich werden in Region i bei der Allokation x^T jedoch die Mengen $-y_L^i$ eingesetzt, die von $\mu_i(w_L^i - x_L^i)$ natürlich verschieden sein können. Die hier gegebene Definition einer erreichbaren Allokation läßt somit Verlagerungen der Inputfaktoren und der Produktion der öffentlichen Güter über die Regionen zu. Dies entspricht einer Vorgehensweise, die einer überregionalen Zusammenarbeit angemessen ist. Die ökonomische Relevanz dieser Verlagerung der Produktion liegt aber nicht in einem etwaigen komparativen Vorteil, der aufgrund der unterstellten konstanten Skalenerträge auch nicht gegeben ist, sondern allein in der Existenz der Spillover-Effekte, die durch die Bereitstellung der öffentlichen Güter in den einzelnen Regionen hervorgerufen werden.

² Der Einfachheit halber wird hier die *Gleichbehandlung* aller Individuen einer Region in dem Sinne unterstellt, daß alle dasselbe Güterbündel konsumieren. Eine allgemeinere Darstellung ist natürlich möglich.

Ist nun $x^D = (x_L^1, \dots, x_L^n; x_M^1, \dots, x_M^n) \in \mathbb{R}_+^{n(l+m)}$ eine D -erreichbare Allokation, so stehen den Individuen in Region i , $i \in N$, durch die mögliche Existenz von Spillover-Effekten öffentliche Güter im Umfang von $x_M^{(i)} = g_1^{(i)} \cdot x_M^1 + \dots + g_n^{(i)} \cdot x_M^n$ zur Verfügung. Häufig wird allerdings der Konsum dieser öffentlichen Güter durch *Überfüllungseffekte* beeinträchtigt, eine größere Zahl konsumierender Individuen reduziert ceteris paribus den individuellen Nutzen aus dem Konsum. Diese Beobachtung, die offenbar für viele öffentliche Güter zutrifft, war schon Gegenstand verschiedener theoretischer und empirischer Analysen (vgl. *Borcherding/Deacon* [1972], *Oakland* [1972], *Roberts* [1974], *Brueckner* [1981] und *Feinstein* [1992]).

Für eine formale Umsetzung dieses Aspekts der Überfüllungseffekte wird nun ein System von Funktionen $h = (h^1, \dots, h^n)$ betrachtet, wobei für jedes $i \in N$ gilt:

$$h^i : \left(\prod_{j=1}^n [0, \mu_j] \right) \setminus \{0\} \rightarrow \mathbb{R}_{++}$$

ist streng monoton wachsend in der i -ten Koordinate und ist normalisiert zu $h^i(\mu_1, \dots, \mu_n) = h(\pi_T) = \mu_i$. Ist also D eine Koalition mit Profil π_D , so beschreibt $h^i(\pi_D)$ die für die Region i relevanten Überfüllungseffekte. Offenbar impliziert diese Formulierung, daß eine Koalition D mit einer D -erreichbaren Allokation x^D Individuen, die nicht der Koalition angehören, auch vom Konsum der in x^D angebotenen öffentlichen Güter ausgeschlossen werden kann. Natürlich ist es eine andere Frage, ob von dieser Ausschlußmöglichkeit auch Gebrauch gemacht wird. Roberts (1974) spricht in einem verwandten Zusammenhang von einem *halb-öffentlichen Gut*.

Gegeben ist nun eine Koalition D und eine D -erreichbare Allokation x^D . Die durch das System h beschriebenen Überfüllungseffekte führen nun für einen Konsumenten der Region i , der auch der Koalition D angehört, zum Nutzen:

$$v^i \left(x_L, \frac{x_M^1}{h^i(\pi_D)}, \dots, \frac{x_M^n}{h^i(\pi_D)} \right).$$

Zur expliziten Berücksichtigung der Spillover-Effekte wird nun die Existenz einer stetigen, streng quasi-konkaven und streng monoton wachsenden Nutzenfunktion $u^i : \mathbb{R}_+^{l+m} \rightarrow \mathbb{R}$ unterstellt mit:

$$v^i \left(x_L, \frac{x_M^1}{h^i(\pi_D)}, \dots, \frac{x_M^n}{h^i(\pi_D)} \right) = u^i \left(x_L, \frac{x_M^{(i)}(x_M^1, \dots, x_M^n)}{h^i(\pi_D)} \right) = u^i \left(x_L, \frac{x_M^{(i)}}{h^i(\pi_D)} \right)$$

für jede D -erreichbare Allokation x^D bezüglich einer beliebigen Koalition D . Der Nutzen, der für ein Individuum in Region i aus dem Konsum des Güterbündels $(x_L, x_M^{(i)})$ zu erreichen ist, hängt folglich ab von der Zahl aller anderen Individuen, die ebenfalls am Konsum der öffentlichen Güter beteiligt sind.

Spezieller besteht die Möglichkeit, daß die Überfüllungseffekte lediglich von denjenigen Individuen ausgelöst werden, die in der betreffenden Region wohnen und einer vorgegebenen Koalition angehören. Dies trifft beispielsweise zu auf das öffentliche Schulsystem, falls für die in einer abgegrenzten Region wohnenden Kinder Schulpflicht in dieser Region besteht. Bei einer gegebenen Verteilung der Schulen, des hier zur Diskussion stehenden halb-öffentlichen Gutes, führt so eine Änderung in der Zahl der schulpflichtigen Kinder in einer bestimmten Region lediglich dort zu Überfüllungseffekten, nicht aber in den anderen Regionen. Ähnlich gilt dies etwa auch für die Agrarmarktordnung der EG. Geht man der Einfachheit von einer Gleichverteilung der insgesamt zur Verfügung stehenden Mittel aus, gilt also $x_M^{(i)} = (1/n)(x_M^1 + \dots + x_M^n)$ für das in diesem Sinne halb-öffentliche Gut Agrarbeihilfe, so hängt der Nutzen dieses Programms für den einzelnen Landwirt in einer bestimmten Region von der Zahl der insgesamt aus dieser Region an diesem Programm partizipierenden Landwirte ab.

Unterstellt man also, daß die Überfüllungseffekte in einer bestimmten Region nur von der Zahl der Individuen abhängen, die in dieser Region die betreffenden öffentlichen Güter konsumieren, so sind die öffentlichen Güter durch eine besondere regionale Eigenschaft ausgezeichnet, man erhält die *regionalen öffentlichen Güter*:

Definition 2.2 Gilt für die Funktionen h^i , $i \in N$, welche die Überfüllungseffekte in Region i beschreiben, $h^i(\pi_D) = h^i(\pi_D^i)$ für $i \in N$, so sind die öffentlichen Güter *regionale öffentliche Güter*.

Gilt etwa noch spezieller $h^i(\pi_D) = \pi_D^i$ für $i \in N$, so bedarf eine Verdopplung der Zahl der konsumierenden Individuen in einer Region demnach auch einer doppelten Menge der regionalen öffentlichen Güter, damit den Individuen kein Nutzenverlust entsteht. Die individuellen Beiträge zur Finanzierung der öffentlichen Güter ändern sich in diesem Fall nicht mit einer sich ändernden Zahl der regionalen Konsumenten. *Roberts* (1974) charakterisiert dies als „constant returns to group size“.

Abschließend wird zur Vereinfachung der Analyse noch unterstellt, daß die pareto-effizienten T -erreichbaren Allokationen der folgenden Randbedingung genügen:

Annahme 1 Sei $x^T = (x_L^1, \dots, x_L^n; x_M^1, \dots, x_M^n)$ eine pareto-effiziente Allokation. Dann gilt: $\sum_{i \in N} \mu_i x_L^i \in \mathcal{R}_{++}^I$.

Diese Annahme reguliert also das Randverhalten der Nutzenfunktionen. Es ist damit nicht optimal, die gesamte Erstausstattung bezüglich eines privaten Gutes in die Produktion der öffentlichen Güter zu stecken.

Unter dieser Annahme ist dann \mathcal{E}_n die Klasse aller Ökonomien E mit n Regionen, charakterisiert durch die Koeffizienten $(g_j^{(i)})$ für die Stärke der Spillover-Effekte und durch die Funktionen (h^1, \dots, h^n) für die Überfüllungseffekte.

Im nächsten Abschnitt werden zwei wichtige Konzepte betrachtet, die zumindest eine theoretische Lösung des Allokationsproblems bei öffentlichen Gütern erlauben, nämlich das Konzept des Lindahl-Gleichgewichts und das Konzept der Kernallokation. Die praktische Relevanz dieser beiden Begriffe wird dann im Kernäquivalenztheorem für Ökonomien mit regionalen öffentlichen Gütern zum Ausdruck kommen.

3. Kern und Lindahl-Gleichgewicht

Gegeben ist wieder die Ökonomie E mit n verschiedenen homogenen Regionen. Die Spillover-Effekte der in den einzelnen Regionen produzierten öffentlichen Güter sind wie oben beschrieben durch die Koeffizienten $(g_j^{(i)})_{i,j \in N}$, die Überfüllungseffekte sind in den Funktionen $h = (h^1, \dots, h^n)$ erfaßt. Zur Vereinfachung der Darstellung wird im folgenden nur der Fall regionaler öffentlicher Güter betrachtet, es gilt also $h^i : [0, \mu_i] \rightarrow [0, \mu_i]$ streng monoton steigend, $i \in N$.

Für den Begriff des Lindahl-Gleichgewichts wird nun ein Preisvektor $P = (p_L, (p_M^{i,j})_{i,j \in N})$ betrachtet mit Preisen $p_L \in \mathbb{R}_+^L$ für die privaten Güter und regionalen Preisen $p_M^{i,j} \in \mathbb{R}_+^M$ für die öffentlichen Güter in Region i . Zu beachten ist dabei, daß den Individuen in einer Region i über die Spillover-Effekte auch die in den anderen Regionen produzierten öffentlichen Güter zur Verfügung stehen. Diese Überlegung führt dann zum regionalen Preisvektor $P^i = (p_L; p_M^{i,1}, \dots, p_M^{i,n})$ für die verschiedenen Güter in Region i , $i \in N$. Schließlich erhält man für jeden Konsumenten $t \in T_i$ mit $i \in N$ die Budgetmenge $B^i(P^i)$:

$$B^i(P^i) = \{(x_L^i; x_M^1, \dots, x_M^n) \in X^i : p_L x_L^i + p_M^{i,1} x_M^1 + \dots + p_M^{i,n} x_M^n \leq p_L w_L^i\}.$$

Das Konzept eines *Lindahl-Gleichgewichts* kann dann wie gewöhnlich eingeführt werden, wobei die Darstellung hier allerdings eine Gleichbehandlung aller Individuen einer Region voraussetzt.³

³ Eine allgemeinere Darstellung ist natürlich möglich, wobei darauf hinzuweisen ist, daß diese Gleichbehandlung in einem Lindahl-Gleichgewicht nicht notwendigerweise gilt (vgl. Champsaur [1975]).

Definition 3.1 Eine T -erreichbare Allokation $\bar{x}^T = (\bar{x}_L^1, \dots, \bar{x}_L^n; \bar{x}_M^1, \dots, \bar{x}_M^n)$ mit zugeordneten Produktionsplänen $(\bar{y}_L^1, \dots, \bar{y}_L^n; \bar{y}_M^1, \dots, \bar{y}_M^n)$ ist ein Lindahl-Gleichgewicht, falls ein Preissystem $\bar{P} = (\bar{p}_L, (\bar{p}_M^{i,j})_{i,j \in N})$ existiert, so daß für alle $t \in T_i$ und alle $i \in N$ gilt:

$$\begin{aligned} & (\bar{x}_L^i; \bar{x}_M^1, \dots, \bar{x}_M^n) \in B^i(\bar{P}^i), \\ & u^i(\bar{x}_L^i, \bar{x}_M^{(i)} / \mu_i) \geq u^i(x_L^i, x_M^{(i)} / \mu_i) \text{ für alle } (x_L^i; x_M^1, \dots, x_M^n) \in B^i(\bar{P}^i), \\ & \sum_{j \in N} \left(\sum_{i \in N} \mu_i \bar{p}_M^{i,j} \right) \cdot \bar{y}_M^j + \bar{p}_L \cdot \sum_{i \in N} \bar{y}_L^i \geq \sum_{j \in N} \left(\sum_{i \in N} \mu_i \bar{p}_M^{i,j} \right) \cdot y_M^j \\ & \quad + \bar{p}_L \cdot \sum_{i \in N} y_L^i \\ & \text{für alle } (y_L^1, \dots, y_L^n; y_M^1, \dots, y_M^n) \text{ mit } (y_L^i, y_M^i) \in Y \text{ für } i \in N. \end{aligned}$$

Die Menge aller Lindahl-Gleichgewichte der Ökonomie E , die der Gleichbehandlungseigenschaft genügen, wird mit $\mathcal{L}^{eq}(E)$ bezeichnet.

Die obigen Bedingungen implizieren natürlich, daß das Güterbündel $(\bar{x}_L^i; \bar{x}_M^1, \dots, \bar{x}_M^n)$ die nutzenmaximale Wahl eines Konsumenten $t \in T_i$ in der Budgetmenge $B^i(\bar{P}^i)$ darstellt und daß der Vektor $\bar{y} = (\bar{y}_L^1, \dots, \bar{y}_L^n; \bar{y}_M^1, \dots, \bar{y}_M^n)$ den Gewinn auf Y bezüglich des gegebenen Preissystems maximiert. Gilt weiter $\bar{x}^T \in \Re_{++}^{n(l+m)}$ für die Lindahl-Allokation zum Preissystem $\bar{P} = (\bar{p}_L, (\bar{p}_M^{i,j})_{i,j \in N})$, so gilt offenbar:

$$\frac{\bar{p}_M^{i,j}}{\bar{p}_M^{i,k}} = \frac{g_j^{(i)}}{g_k^{(i)}} \quad \text{für alle } i, j, k \in N.$$

Deshalb existiert in diesem Fall ein diesem Lindahl-Gleichgewicht zugeordneter modifizierter Preisvektor $\bar{P} = (\bar{p}_L, (\bar{p}_M^i)_{i \in N})$ mit $\bar{p}_M^{i,j} = \bar{p}_M^i \cdot g_j^{(i)}$ für alle $i, j \in N$. Ist insbesondere $g_1^{(i)} = \dots = g_n^{(i)}$ für alle $i \in N$, so genügt der modifizierte Preisvektor $\bar{P} = (\bar{p}_L, (\bar{p}_M^i)_{i \in N})$ zur Beschreibung des gleichgewichtigen Preissystems.

Der Begriff des Lindahl-Gleichgewichts, das in seinen Ansätzen auf Überlegungen von *Wicksell* (1896) zurückgeht, beinhaltet über die im Gleichgewicht übereinstimmende Nachfrage nach den öffentlichen Gütern das *Einstimmigkeitsprinzip*. Insbesondere wird durch die individualisierten Preise niemand veranlaßt, mehr als gewünscht für die Bereitstellung der

öffentlichen Güter zu bezahlen. Somit ist das Lindahl-Gleichgewicht eine Art „Marktlösung“ für das Allokationsproblem in Ökonomien mit öffentlichen Gütern. Wie in einer Ökonomie mit privaten Gütern wird auf Seiten der Konsumenten und der Produzenten ein Preisnehmerverhalten unterstellt mit der Zielsetzung der Nutzen- und der Gewinnmaximierung. Über die vergleichbaren Anforderungen, die die Existenz eines Lindahl-Gleichgewichts sicherstellen (vgl. Foley [1970]), geht die Parallelität zum kompetitiven Gleichgewicht weiter zur Effizienz- und zur Kerneigenschaft der jeweiligen Gleichgewichtsallokationen, jedenfalls unter der Annahme konstanter Skalenerträge für die Technologie im Fall einer Ökonomie mit öffentlichen Gütern (vgl. Kaneko [1976]).

Theorem 3.2 Unter den gegebenen Annahmen an die regionalisierte Ökonomie E existiert ein Lindahl-Gleichgewicht, bei dem alle Individuen einer Region gleich behandelt werden.

Beweis: Für den Beweis wird folgende endliche Ökonomie E^n aus der vorgegebenen Ökonomie E abgeleitet: Der Güterraum der Ökonomie E^n entspricht dem der Ökonomie E , desgleichen die Technologie Y . Es gibt n Individuen in E^n , bezeichnet durch die Menge $N = \{1, \dots, n\}$, wobei Individuum i , $i \in N$, charakterisiert ist durch die Nutzenfunktion $f^i : \mathbb{R}_+^{l+n m} \rightarrow \mathbb{R}$ mit $f^i(x_L^i, x_M^{(i)} / \mu_i) := u^i(x_L^i / \mu_i, x_M^{(i)} / \mu_i)$. $u^i(x_L^i / \mu_i, x_M^{(i)} / \mu_i)$ gibt den Nutzen eines Konsumenten $t \in T_i$ der Ökonomie E in Bezug auf den Konsum des Güterbündels $(x_L^i / \mu_i; x_M^{(i)}) = (x_L^i; x_M^1, \dots, x_M^n) \in \mathbb{R}_+^{l+n m}$ an. Die Erstausstattung an den privaten Gütern ist $\omega_L^i := \mu_i w_L^i$. Eine Anwendung herkömmlicher Existenzsätze für Lindahl-Gleichgewichte auf diese Ökonomie E^n (vgl. etwa Foley (1970)) liefert nun unmittelbar die Existenz eines Lindahl-Gleichgewichts $\tilde{x} = (\tilde{x}_L^1, \dots, \tilde{x}_L^n; \tilde{x}_M^1, \dots, \tilde{x}_M^n)$ zum Preissystem $\tilde{P} = (\tilde{p}_L; (\tilde{p}_M^{i,j})_{i,j \in N})$ für die endliche Ökonomie E^n . Die Allokation $\bar{x}^T = (\bar{x}_L^1, \dots, \bar{x}_L^n; \bar{x}_M^1, \dots, \bar{x}_M^n)$ mit $\bar{x}_L^i := \tilde{x}_L^i / \mu_i$ und $\bar{x}_M^i := \tilde{x}_M^i / \mu_i$ für $i \in N$ ist dann eine T -erreichbare Allokation für die regionalisierte Ökonomie E . Weiter ist \bar{x}^T mit dem Preissystem $\bar{P} = (\bar{p}_L; (\bar{p}_M^{i,j})_{i,j \in N})$, definiert durch $\bar{p}_L := \tilde{p}_L$ und $\bar{p}_M^{i,j} := \tilde{p}_M^{i,j} / \mu_i$ für alle $i, j \in N$, ein Lindahl-Gleichgewicht für die Ökonomie E . Gemäß Konstruktion ist dieses Gleichgewicht durch die Gleichbehandlungseigenschaft ausgezeichnet.

Da sie das Ergebnis eines kooperativen Prozesses sind, werden die *Kernallokationen* oft als diejenigen Allokationen beschrieben, die über „rationale“ gesellschaftliche Entscheidungsmechanismen erreichbar sind (vgl. dazu etwa Muench [1972]). Wenn demnach ein bestimmter Allokationsmechanismus die *Kerneigenschaft* besitzt, also immer zu Kernallokationen führt, und wenn umgekehrt jede Allokation im Kern über diesen Mechanismus erreicht werden kann, so steht dann auch dem betreffenden Allokationsmechanismus diese Rationalitätseigenschaft zu. Die *Kernäquivalenz*, die Gleichheit des Kerns mit der Menge der möglichen Ergebnisse eines gege-

benen Allokationsmechanismus, hat demnach eine grundlegende Bedeutung für die „soziale Akzeptanz“ des betreffenden Allokationsmechanismus.

In bezug auf den Marktmechanismus gilt die Kernäquivalenz, jedenfalls für „große“ Ökonomen mit privaten Gütern (vgl. Aumann [1964], Hildenbrand [1974]). In diesem Sinne kann folglich der Preismechanismus als „erste Approximation“ an einen rationalen Mechanismus betrachtet werden (vgl. wieder Muench [1972]). Damit stellt sich natürlich sofort die Frage, ob auch der Lindahl-Mechanismus durch die Kernäquivalenz ausgezeichnet ist. Dies ist hier umso wichtiger als die Annahme eines Preisnehmerverhaltens bei den individualisierten Preisen für die öffentlichen Güter höchst problematisch ist. Vielfach wird dann die Gültigkeit eines Äquivalenztheorems als einzige Möglichkeit gesehen, dem Lindahl-Konzept aus ökonomischer Sicht Bedeutung beizumessen.

Die Gegenbeispiele von Muench (1972) und Milleron (1972) machen deutlich, daß mit dieser Äquivalenz im allgemeinen nicht gerechnet werden kann, jedenfalls solange man reine öffentliche Güter ohne Überfüllungseffekte betrachtet. In diesem Fall sinken nämlich die Pro-Kopf-Kosten der Bereitstellung einer bestimmten Menge an öffentlichen Gütern mit einer größeren Zahl sich an der Finanzierung beteiligender Individuen. Kleine Koalitionen werden damit vergleichsweise bedeutungslos, auch in Hinblick auf die Möglichkeit des „Verbesserns“ einer vorgegebenen Allokation, der Kern wird „groß“ (vgl. auch Foley [1970], Roberts [1974]).

Die Existenz von Überfüllungseffekten in bezug auf den Konsum der öffentlichen Güter kann aber die ökonomische Relevanz kleiner Koalitionen beachtlich steigern. Und diese Möglichkeit wird letztlich der Grund für das Äquivalenztheorem für Lindahl-Gleichgewichte bei speziellen regionalen öffentlichen Gütern sein, das im folgenden Abschnitt angeführt werden wird. Hier steht zunächst der formale Begriff des Kerns bei öffentlichen Gütern selbst zur Diskussion.

Sei dazu $x^T = (x_L^1, \dots, x_L^n; x_M^1, \dots, x_M^n) \in \mathbb{R}_+^{n(l+m)}$ eine T -erreichbare Allokation. Weiter sei ein System $h = (h^1, \dots, h^n)$ von Funktionen gegeben, die die Überfüllungseffekte in den einzelnen Regionen im Sinne regionaler öffentlicher Güter beschreiben.

Definition 3.3 Die Koalition D mit Profil $\pi_D = (\pi_D^1, \dots, \pi_D^n)$ kann die Allokation x^T über die D -erreichbare Allokation $\hat{x}^D = (\hat{x}_L^j; \hat{x}_M^j)_{j \in D}$ h -verbessern, falls für alle $i \in s(D)$ gilt:

$$u^i \left(\hat{x}_L^i, \frac{\hat{x}_M^{(i)}}{h^i(\pi_D^i)} \right) > u^i \left(x_L^i, \frac{x_M^{(i)}}{h^i(\mu_i)} \right).$$

Genauer beschreibt dann $h = (h^1, \dots, h^n)$ ein System *lokaler Überfüllungseffekte* in der regional differenzierten Ökonomie E . Der h -Kern wird damit folgendermaßen definiert:

Definition 3.4 Das System $h = (h^1, \dots, h^n)$ lokaler Überfüllungseffekte sei gegeben. Die Menge aller T -erreichbaren Allokationen die von keiner Koalition h -verbessert werden können, ist der *h -Kern* der Ökonomie E , bezeichnet mit $\mathcal{C}^h(E)$.

Die *Kerneigenschaft* für Lindahl-Gleichgewichte⁴ ergibt sich aus den nun folgenden Überlegungen. Zunächst wird der h -Kern der Ökonomie E bezüglich zweier Systeme h_o und h_1 lokaler Überfüllungseffekte betrachtet. Gilt etwa $h_o^i(\pi_i) \geq h_1^i(\pi)$ für alle $\pi \in [0, \mu_i]$ und für alle $i \in N$, so erhält man sofort das folgende Korollar:

Korollar 3.5 Gegeben seien zwei Systeme h_o und h_1 lokaler Überfüllungseffekte mit $h_o^i(\pi) \geq h_1^i(\pi)$ für alle $\pi \in [0, \mu_i]$ und für alle $i \in N$. Dann gilt $\mathcal{C}^{h_1}(E) \subseteq \mathcal{C}^{h_o}(E)$.

Insbesondere sei nun das System $h_* = (h_*^1, \dots, h_*^n)$ gegeben durch $h_*^i(\pi) := \pi$ für alle $\pi \in [0, \mu_i]$ und für alle $i \in N$. Dieser Fall charakterisiert „constant returns to group size“ (vgl. Roberts [1974]) lokal in jeder Region. Die Kerneigenschaft der Lindahl-Gleichgewichte ergibt sich damit folgendermaßen (vgl. auch Vasil'ev/Weber/Wiesmeth [1992] für ein entsprechendes Resultat in einem speziellen Fall):

Theorem 3.6 Für das System lokaler Überfüllungseffekte $h = (h^1, \dots, h^n)$ gelte $h^i(\pi) \geq h_*^i(\pi) = \pi$ für alle $\pi \in [0, \mu_i]$ und für alle $i \in N$. Dann gilt $\mathcal{L}^{eq}(E) \subseteq \mathcal{C}^h(E)$.

Beweis: Aufgrund des obigen Korollars genügt es, den Beweis für den Fall $h = h_*$ zu führen. Sei dazu \bar{x}^T ein Lindahl-Gleichgewicht zum Preissystem $\bar{P} = (\bar{p}_L, (\bar{p}_M^{i,j})_{i,j \in N})$. Angenommen, eine Koalition D mit Profil $\pi_D = (\pi_D^1, \dots, \pi_D^n)$ kann die Allokation \bar{x}^T über die D -erreichbare Allokation \hat{x}^D h^* -verbessern. Dann gilt:

$$\bar{p}_L \cdot \hat{x}_L^i + \sum_{j \in N} \bar{p}_M^{i,j} \cdot \frac{\mu_i}{\pi_D^i} \cdot \hat{x}_M^j - \bar{p}_L \cdot w_L^i > 0 \quad \text{für alle } i \in s(D).$$

Aufsummieren über alle Individuen in D liefert dann den folgenden Ausdruck mit den nicht-negativen Preisen \bar{P} :

⁴ Genauer kann die Kerneigenschaft nur für die hier betrachteten Lindahl-Gleichgewichte mit der Gleichbehandlungseigenschaft nachgewiesen werden, da unter den bestehenden Konvexitätsannahmen alle Kernallokationen durch diese Eigenschaft ausgezeichnet sind (vgl. Vasil'ev/Weber/Wiesmeth [1992]).

$$0 < \bar{p}_L \cdot \sum_{i \in s(D)} \pi_D^i \hat{x}_L^i + \sum_{i \in s(D)} \sum_{j \in N} \bar{p}_M^{i,j} \cdot \mu_i \cdot \hat{x}_M^j - \bar{p}_L \cdot \sum_{i \in s(D)} \pi_D^i w_L^i \\ \leq \bar{p}_L \cdot \sum_{i \in N} \hat{y}_L^i + \sum_{i \in N} \sum_{j \in N} \bar{p}_M^{i,j} \cdot \mu_i \cdot \hat{y}_M^j.$$

Das mit der Allokation \hat{x}^D verbundene System von Produktionsplänen $(\hat{y}_L^i, \hat{y}_M^i)_{i \in N}$ liefert somit einen positiven Gewinn bei den Lindahl-Preisen \bar{P} . Dies ist aber im Widerspruch zur gewinnmaximalen Wahl der mit dem Lindahl-Gleichgewicht verbundenen Produktionspläne.

Soviel zur Existenz und zur Kerneigenschaft von Lindahl-Gleichgewichten bei regionalen öffentlichen Gütern. Die *Kernäquivalenz* ist Gegenstand der Untersuchung des folgenden Abschnitts.

4. Das Kernäquivalenztheorem

In diesem Abschnitt soll der Frage nachgegangen werden, inwieweit Lindahl-Gleichgewichte nicht nur der Kerneigenschaft genügen, sondern vielmehr durch sie charakterisiert sind. Es geht also um die Frage der Äquivalenz von Kern und Lindahl-Gleichgewichten der regionalisierten Ökonomie E . Spätestens mit der Veröffentlichung der Gegenbeispiele von *Milleron* (1972) und *Muench* (1972) ist dazu bekannt, daß der Kern in Ökonomien mit reinen öffentlichen Gütern ohne die Existenz von Überfüllungseffekten im allgemeinen die Menge der Lindahl-Gleichgewichte echt umfassen wird, gerade auch im Fall einer Kontinuumsökonomie. Wie schon erwähnt, ist der Grund in der bei öffentlichen Gütern begrenzten ökonomischen Macht kleinerer Koalitionen zu suchen (vgl. auch die diesbezüglichen Hinweise von *Foley* [1970]).

Roberts (1974) gelang der Nachweis der Kernäquivalenz für eine Typenökonomie mit nur einem Typ von Konsumenten unter der Existenz von Überfüllungseffekten. Beschrieben wurden die Überfüllungseffekte dabei durch $h^1(\pi_1) = \pi_1$ für alle $\pi_1 \in (0, 1]$. An Hand des Beispiels einer Typenökonomie mit zwei Typen von Konsumenten wurde zugleich gezeigt, daß dieses Ergebnis nicht unmittelbar verallgemeinert werden kann. Eine genauere Analyse ergibt, daß die diesem Beispiel zugrundeliegenden Überfüllungseffekte durch $h^1(\pi_1, \pi_2) = h^2(\pi_1, \pi_2) = \pi_1 + \pi_2$ mit $\pi_1, \pi_2 \in (0, 1]$ beschrieben werden können, die betreffenden öffentlichen Güter sind demnach nicht regional öffentlich im Sinne dieser Ausführungen.

Zur weiteren Behandlung dieses Problemkreises, auch in bezug zu den hier unterstellten regionalen öffentlichen Gütern, sind einige Anmerkungen und Einschränkungen zu den Funktionen $h = (h^1, \dots, h^n)$ notwendig, die die lokalen Überfüllungseffekte beschreiben.

Dazu bezeichnet zunächst H die Familie aller Funktionen $h = (h^1, \dots, h^n)$, so daß $h^i : [0, \mu_i] \rightarrow [0, \mu_i]$ für jedes $i \in N$ stetig und monoton wachsend ist mit $h^i(\mu_i) = \mu_i$. Spezieller gehört h zu $\hat{H}_- \subset H$, falls ein $\tau^o \in (0,1)$ existiert mit $h^i(\tau^o \mu_i) < \tau^o \mu_i$ für alle $i \in N$. Entsprechend ist dann $\hat{H}_+ \subset H$ die Familie aller Funktionen $h = (h^1, \dots, h^n) \in H$ für die ein $\alpha \in (0,1)$ existiert mit $h^i(\tau \mu_i) \geq \mu_i \tau^\alpha$ für alle $\tau \in (0,1)$ und für alle $i \in N$. Schließlich wird noch $h^* = (h_1^*, \dots, h_n^*) \in H$ betrachtet mit $h_*(\pi_i) = \pi_i$ für alle $\pi_i \in [0, \mu_i]$ und für alle $i \in N$. Mit \hat{H} wird nun die Vereinigungsmenge dieser ausgezeichneten Familien von Funktionen benannt:

$$\hat{H} := \hat{H}_- \cup \{h^*\} \cup \hat{H}_+.$$

Mit diesen Bezeichnungen ergibt sich das folgende Kernäquivalenztheorem für Ökonomien mit regionalen öffentlichen Gütern:

Theorem 4.1 Für $h \in \hat{H}$ stimmt der h -Kern $C^h(E)$ mit der Menge der Lindahl-Gleichgewichte $\mathcal{L}^{eq}(E)$ dann überein, falls die lokalen Überfüllungseffekte durch $h = h^*$ beschrieben werden. Gilt $h \in \hat{H}_-$, so ist $C^h(E) = \emptyset$, gilt $h \in \hat{H}_+$, so gibt es eine Ökonomie $E \in \mathcal{E}_n$ derart, daß die Menge der Lindahl-Gleichgewichte $\mathcal{L}^{eq}(E)$ echt im h -Kern $C^h(E)$ der Ökonomie E enthalten ist.

Das Theorem berührt demnach die speziellen regionalen öffentlichen Güter, für die etwa bei einer Verdoppelung der Zahl der Konsumenten einer Region die doppelte Menge dieser öffentlichen Güter notwendig ist, um cet. par. das Nutzenniveau der Individuen dieser Region nicht zu verändern. Zu beachten ist ferner, daß die Menge der im Theorem angesprochenen Familien von Funktionen \hat{H} eine echte Teilmenge von H darstellt. Es ist momentan unklar, inwieweit eine allgemeinere Behandlung der Familien $h \in H \setminus \hat{H}$ im Sinne des Äquivalenztheorems möglich ist.

Eine ökonomische Würdigung des Äquivalenzresultats erfolgt im nächsten Abschnitt. Hier sollen noch einige Hinweise zum Beweis dieses Theorems gegeben werden.

Anmerkungen zum Beweis des Äquivalenztheorems: Der Beweis ist in mehrere Abschnitte gegliedert. Da der Beweis in seinen wesentlichen Überlegungen direkt an den entsprechenden Beweis in Vasil'ev/Weber/Wiesmeth (1992) anknüpft, sollen hier nur einige wichtige Beweisschritte angegeben oder angedeutet werden.⁵

⁵ Dabei ist zu berücksichtigen, daß das in Vasil'ev/Weber/Wiesmeth (1992) betrachtete Modell in bezug auf die Definition der Spillover-Effekte einen Spezialfall des dieser Arbeit zugrundeliegenden Modells darstellt.

Ein erster wichtiger Teil des Beweises ist der Nachweis, daß jede Kernallokation, genauer jede Allokation $\bar{x} \in \mathcal{C}^*(E)$ in einer Ökonomie E mit durch $h = h_*$ beschriebenen lokalen Überfüllungseffekten, auch ein Lindahl-Gleichgewicht der Ökonomie darstellt, daß also auch gilt $\bar{x} \in \mathcal{L}^{eq}(E)$.⁶

Sei dazu nun $\bar{x} \in \mathcal{C}^*(E)$ mit der zugehörigen Liste $\bar{y} = (\bar{y}_L^i, \bar{y}_M^i)_{i \in N}$ von Produktionsplänen. Definiere $Q := \prod_{k=0}^n Q_k$ mit $Q_0 = \mathfrak{R}^l$ und $Q_k = \mathfrak{R}^{nm}$ für $k = 1, \dots, n$. Für jedes $i \in N$ definiert man die linearen Operatoren $A_i^X : \mathfrak{R}^{l+nm} \rightarrow Q$ und $A_i^Y : \mathfrak{R}^{l+m} \rightarrow Q$ durch:

$$A_i^X [x_L, x_M^1, \dots, x_M^n] = \begin{cases} \mu_i x_L, & \text{falls } k = 0, \\ (x_M^1, \dots, x_M^n), & \text{falls } k = i, \\ 0, & \text{sonst} \end{cases}$$

und

$$A_i^Y [y_L^i, y_M^i] = \begin{cases} y_L^i, & \text{falls } k = 0, \\ (0, \dots, y_M^i, \dots, 0), & \text{falls } k = 1, \dots, n. \end{cases}$$

Für jedes $i \in N$ bezeichne nun $\mathcal{P}^i(\bar{x}^i)$ die Menge aller Güterbündel, die von den Konsumenten der Region i dem Güterbündel \bar{x}^i , definiert durch die Kernallokation \bar{x} , strikt vorgezogen werden. Dies führt zu den folgenden Definitionen:

$$V_i(\bar{x}) := \{A_i^X[\hat{x}^i - w^i] : \hat{x}^i \in \mathcal{P}^i(\bar{x}^i)\},$$

$$V(\bar{x}) := \left\{ \sum_{i \in N} \tau_i z^i : z^i \in V_i(\bar{x}), \tau = (\tau_1, \tau_2, \dots, \tau_n) \in I^n \right\}$$

mit $I^n := [0,1]^n \setminus \{0\}$,

$$\mathcal{K}(\bar{x}) := V(\bar{x}) - \sum_{i \in N} A_i^Y[Y].$$

Die Quasi-Konkavität der Nutzenfunktionen und die Linearität der verwendeten Operatoren implizieren die Konvexität der Menge $\mathcal{K}(\bar{x})$. Weiter gilt: Der Nullvektor ist nicht in $\mathcal{K}(\bar{x})$ enthalten. Wäre nämlich $0 \in \mathcal{K}(\bar{x})$, so existiert ein Vektor $\tau = (\tau_1, \dots, \tau_n) \in I^n$, sowie für jedes $i \in N$

⁶ Wie bisher wird auch hier von Kernallokationen ausgegangen, die der Gleichbehandlungseigenschaft genügen. Ein formaler Beweis dazu kann in enger Anlehnung an den entsprechenden Beweis in Vasil'ev/Weber/Wiesmeth (1992) geführt werden.

ein Güterbündel $\hat{x}^i = (\hat{x}_L^i, \hat{x}_M^{i,1}, \dots, \hat{x}_M^{i,n}) \in \mathcal{P}^i(\bar{x}^i)$ und ein Produktionsplan $(\hat{y}_L^i, \hat{y}_M^i) \in Y$ mit:

$$\sum_{i \in N} \mu_i \tau_i (\hat{x}_L^i - w_L^i) = \sum_{i \in N} \hat{y}_L^i$$

und

$$\tau_j (\hat{x}_M^{j,1}, \dots, \hat{x}_M^{j,n}) = (\hat{y}_M^1, \dots, \hat{y}_M^n) \quad \text{für alle } j \in N.$$

Betrachte nun eine Koalition D mit Profil $(\mu_1 \tau_1, \dots, \mu_n \tau_n)$. Die Allokation \tilde{x}^D mit $\tilde{x}_L^i := \hat{x}_L^i$ für alle $i \in N$ und $\tilde{x}_M^i := \tau_i \hat{x}_M^{i,i}$ ist dann eine D -erreichbare Allokation mit zugeordneten Produktionsplänen $\tilde{y} = (\tilde{y}_L^i, \tilde{y}_M^i)_{i \in N}$. Man weist nun unmittelbar nach, daß die Koalition D die Kernelallokation \bar{x} über \tilde{x}^D h^* -verbessern kann, ein Widerspruch. Folglich ist der Nullvektor nicht in $\mathcal{K}(\bar{x})$ enthalten.

Damit existiert ein Vektor $\hat{P} = (\hat{p}_L, (\hat{p}_M^{i,j})_{i,j \in N}) \in Q \setminus \{0\}$ mit $\hat{P} \cdot z \geq 0$ für alle $z \in \mathcal{K}(\bar{x})$. Sei $\bar{P} = (\bar{p}_L, (\bar{p}_M^{i,j})_{i,j \in N})$ definiert durch $\bar{p}_L := \hat{p}_L$ und $\bar{p}_M^{i,j} := \hat{p}_M^{i,j} / \mu_i$ für alle $i, j \in N$. Es ist nun nachzuweisen, daß die T -erreichbare Allokation \bar{x} ein Lindahl-Gleichgewicht mit Preissystem \bar{P} darstellt. Mit $0 \in Y$ gilt $\bar{p}_L \cdot \mu_i \hat{x}_L^i + \sum_{j \in N} \mu_i \bar{p}_M^{i,j} \cdot \hat{x}_M^j \geq \bar{p}_L \cdot \mu_i w_L^i$ für alle $i \in N$ und für jedes $\hat{x}^i \in \mathcal{P}^i(\bar{x}^i)$. Aufgrund der Monotonie und der Stetigkeit der Nutzenfunktionen gilt diese Ungleichung auch für \bar{x}^i , so daß man durch Addition über $i \in N$ $\bar{p}_L \cdot \sum_{i \in N} \bar{y}_L^i + \sum_{i \in N} \sum_{j \in N} \mu_i \bar{p}_M^{i,j} \cdot \bar{y}_M^j \geq 0$ erhält. Andererseits liegt der Nullvektor auch im Abschluß von $V(\bar{x})$. Folglich gilt auch $\bar{p}_L \cdot (-\sum_{i \in N} \bar{y}_L^i) + \sum_{i \in N} \sum_{j \in N} \mu_i \bar{p}_M^{i,j} \cdot (-\bar{y}_M^j) \geq 0$. Insgesamt ergibt sich demnach:

$$\bar{p}_L \cdot \sum_{i \in N} \bar{y}_L^i + \sum_{i \in N} \sum_{j \in N} \mu_i \bar{p}_M^{i,j} \cdot \bar{y}_M^j = 0$$

und für jedes $i \in N$:

$$\bar{p}_L \cdot \bar{x}_L^i + \sum_{j \in N} \bar{p}_M^{i,j} \cdot \bar{x}_M^j = \bar{p}_L \cdot w_L^i, \quad \text{also } \bar{x}^i \in B^i(\bar{P}^i).$$

Ist weiter $\bar{p}_L = 0$, so existiert wegen $\bar{P} \geq 0$ und $\bar{P} \neq 0$ ein $(y_L, y_M) \in Y$ und ein $j \in N$ mit $\bar{p}_L y_L + (\sum_{i \in N} \mu_i \bar{p}_M^{i,j}) \cdot y_M > 0$, ein Widerspruch. Damit gibt es also ein $r \in L$ mit $\bar{p}_r > 0$. Gibt es zusätzlich ein $k \in L$ mit $\bar{p}_k = 0$, so wähle $i \in N$ mit $\bar{x}_r^i > 0$. Beim Güterbündel $x^i(\delta_k, \delta_r)$ sei \bar{x}_k^i um eine Menge δ_k erhöht und \bar{x}_r^i um eine hinreichend kleine Menge δ_r reduziert. Offenbar kann man δ_k und δ_r so wählen, daß $x^i(\delta_k, \delta_r)$ zu $\mathcal{P}^i(\bar{x}^i)$ gehört und daß es bei Preisen \bar{P}^i echt billiger ist als \bar{x}^i , ein Widerspruch. Folglich gilt $\bar{p}_L \in \mathbb{R}_{++}^l$.

Annahme, es existiert ein $i \in N$ und ein Güterbündel $x^i \in \mathcal{P}^i(\bar{x}^i) \cap B^i(\bar{P}^i)$. Offenbar gilt dann $\bar{p}_L \cdot x_L^i + \sum_{j \in N} \bar{p}_M^{i,j} \cdot x_M^j = \bar{p}_L \cdot w_L^i$ aufgrund der Trennungseigenschaft für \hat{P} . Weiter kann $x_L^i \neq 0$ angenommen werden, da man ansonsten eine geeignete Konvexitätskombination von $(0, x_M^i)$ mit $(w_L^i, 0)$ betrachten kann. Dann wähle man eine beliebige Zahl $\gamma \in (0, 1)$. Das Güterbündel $\hat{x}^i = \gamma x^i + (1 - \gamma) \bar{x}^i = \bar{x}^i + \gamma(x^i - \bar{x}^i)$ gehört dann wegen der Quasi-Konkavität der Präferenzen ebenfalls zu $\mathcal{P}^i(\bar{x}^i) \cap B^i(\bar{P}^i)$. Wähle weiter $r \in L$ mit $x_r^i > 0$ und $\delta > 0$ so daß $\hat{x}^i - \delta e^r \in \mathcal{P}^i(\bar{x}^i)$, wobei e^r der r -te Einheitsvektor ist. Da jedoch $\bar{p}_r > 0$, ist $\hat{x}^i - \delta e^r$ beim Preissystem \bar{P} wieder echt billiger als \bar{x}^i , ein Widerspruch zur Trennungseigenschaft von \hat{P} . Damit ist nachgewiesen, daß jede Allokation $\bar{x} \in \mathcal{C}^*(E)$ bei mit $h = h^*$ beschriebenen lokalen Überfüllungseffekten auch ein Lindahl-Gleichgewicht der Ökonomie E darstellt.

Ein weiterer wichtiger Teil des Beweises behandelt den Fall $h \in \hat{H}_-$. Es ist zu zeigen, daß der h -Kern $\mathcal{C}^h(E)$ leer ist. Annahme, es existiert $x = (x_L^1, \dots, x_L^n; x_M^1, \dots, x_M^n) \in \mathcal{C}^h(E)$ mit zugehörigen Produktionsplänen $(y_L^i, y_M^i)_{i \in N}$. Wähle $\tau \in (0, 1)$ mit $h^i(\tau \mu_i) < \tau \mu_i$ für alle $i \in N$ und betrachte eine Koalition D mit Profil $\pi_D = (\tau \mu_1, \dots, \tau \mu_n)$. Definiere die D -erreichbare Allokation \bar{x}^D durch $\bar{x}_L^i := x_L^i$ und $\bar{x}_M^i := \tau x_M^i$ für $i \in N$. Die zu \bar{x}^D gehörenden Produktionspläne sind gegeben durch $(\bar{y}_L^i, \bar{y}_M^i) := (\tau y_L^i, \tau y_M^i)$ für $i \in N$. Man erhält dann für jede Region $i \in N$:

$$u^i \left(\bar{x}_L^i, \frac{\bar{x}_M^{(i)}}{h^i(\tau \mu_i)} \right) = u^i \left(\bar{x}_L^i, \frac{\tau x_M^{(i)}}{h^i(\tau \mu_i)} \right) > u^i \left(x_L^i, \frac{x_M^{(i)}}{\mu_i} \right).$$

Damit aber kann die Koalition D die Allokation x über die D -erreichbare Allokation \bar{x}^D h -verbessern, ein Widerspruch zur Kerneigenschaft von x .

Der letzte Teil des Beweises betrifft den Fall $h \in \hat{H}_+$. Hier ist explizit die Existenz einer Ökonomie $E \in \mathcal{E}^n$ nachzuweisen, so daß $\mathcal{L}^{eq}(E)$ eine echte Teilmenge von $\mathcal{C}^h(E)$ darstellt. Dies wird in Vasil'ev/Weber/Wiesmeth (1992) für den Fall $n = 2$ und mit $g_i^{(j)} = 1$, $i, j \in N$, gelöst. Dieses Beispiel kann auf den Fall $n > 2$ verallgemeinert werden, weiter können auch beliebige Spillover-Koeffizienten zugelassen werden.

Soviel zu den wesentlichen Beweisschritten. Für weitere Details sei auf die entsprechenden Überlegungen in Vasil'ev/Weber/Wiesmeth (1992) verwiesen, die auf den hier betrachteten Fall angepaßt werden können.

Im folgenden Abschnitt wird die ökonomische Bedeutung dieses speziellen Äquivalenzresultats gewürdigt. Dies betrifft vor allem die Problematik der überregionalen Bereitstellung öffentlicher Güter.

5. Würdigung des Äquivalenzresultats

Die vorangegangenen Überlegungen haben gezeigt, daß die Äquivalenz des Kerns mit der Menge der Lindahl-Gleichgewichte in einer regional strukturierten Ökonomie möglich ist, wenn die öffentlichen Güter durch bestimmte Überfüllungseffekte beschrieben werden können. Dem Konzept des Lindahl-Gleichgewichts kommt in diesem Rahmen damit eine neue Bedeutung zu, die nun abschließend zu bewerten ist.

Besonders wichtig ist in diesem Zusammenhang die Charakterisierung der Lindahl-Gleichgewichte durch die Kerneigenschaft. *Roberts* (1974) meint dazu, daß

„...no matter by what system [public] goods actually are allocated, if ... we assume that trade and production will take place among agents as long as it is advantageous, then any allocation that actually arose could have been achieved by the Lindahl price mechanism.“

Die auf dezentralen Entscheidungen gegründeten Lindahl-Gleichgewichte repräsentieren damit zugleich die „demokratischen“ (vgl. *Hildenbrand/Kirman* [1976]) Kernallokationen. Dies ist umso bedeutender als die mit dem Kernkonzept verbundenen Tauschvorgänge in einer großen Ökonomie nicht realisiert werden können. Die Dezentralisierung der wirtschaftlichen Entscheidungen über ein Preissystem bietet dagegen gerade in einer großen Ökonomie im Hinblick auf die Bedingungen des vollkommenen Wettbewerbs eine realisierbare Alternative.

Dies bedeutet offenbar eine Aufwertung für den mit dem Lindahl-Konzept verbundenen Mechanismus der Kostenverteilung für die bereitgestellten öffentlichen Güter. Für die Individuen einer jeden Region sind von einer überregionalen Institution die Preise für eine Einheit der regionalen öffentlichen Güter so festzulegen, daß die nachgefragten Mengen über alle Regionen hinweg gleich sind und daß die Produktionskosten insgesamt gedeckt sind. Die Frage nach der Anreizverträglichkeit dieses Allokationsmechanismus stellt sich dann in der Form eines Informationsproblems für die verantwortliche überregionale Institution in bezug auf die gleichgewichtigen Lindahl-Preise. Strategisches Verhalten einzelner Individuen, sei es durch Angabe falscher Präferenzen oder durch Angabe überhöhter oder zu niedriger Zahlungswilligkeiten, scheidet aus, da der ökonomische Einfluß einzelner Haushalte in einer großen Ökonomie vernachlässigbar gering ist.

Abschließend ist noch auf den Aspekt der Überfüllungseffekte einzugehen. Es ist sicherlich richtig, daß die meisten „öffentlichen“ Güter durch die eine oder andere Form von Überfüllungseffekten charakterisiert sind. Jedoch werden diese Staueffekte im allgemeinen Fall weder „lokal“, d.h. auf die Individuen einer bestimmten Region bezogen, noch „konstant“ sein

in dem Sinne, daß eine bestimmte prozentuale Erhöhung der Zahl der Konsumenten einer entsprechend größeren Menge der öffentlichen Güter bedarf, um zu keinen Nutzenverlusten zu führen. Jedoch sollte man hier nicht vergessen, daß staatlicherseits die Möglichkeit der „Regionalisierung“ eines öffentlichen Gutes gegeben ist. Ein Beispiel dieser Möglichkeit liefert das Vorgehen der Regierung von Ontario in Kanada, die vor etwa zwei Jahren eine Regionalisierung der Müllentsorgung vorschrieb. Anstelle einer zentralen Mülldeponie im nördlichen Ontario wurde jede Gemeinde mit der dezentralen Entsorgung des auf ihrem Gebiet anfallenden Mülls beauftragt. Die von der Regierung sicherlich beabsichtigte Folge ist, daß sich jede Gemeinde stärker des Müllproblems bewußt wird, daß die Überfüllungseffekte deutlicher werden als im Fall einer zentralen Entsorgung weit ab von jeder menschlichen Besiedlung. Entscheidend aber für das hier betrachtete Modell ist, daß die Überfüllungseffekte lokaler Natur sind, daß also nur die Bewohner einer jeden Gemeinde zur dortigen Überfüllung beitragen. Unterstellt man noch homogene Präferenzen in einer jeden Gemeinde in bezug auf das Abfallproblem, so sind die wesentlichen Annahmen des formalen Äquivalenzmodells erfüllt. Das Kernäquivalenztheorem bietet so eine theoretische Fundierung der Regionalisierung einer Ökonomie in bezug auf die Bereitstellung öffentlicher Güter. Dieser Aspekt ist interessant und sollte weiter verfolgt werden.

Literatur

- Atkinson, A./Stiglitz, J.* (1980), „Lectures on Public Economics“, McGraw-Hill, London.
- Aumann, R.* (1964), „Markets with a Continuum of Traders“, Econometrica, Vol. 32, S. 39 - 50.
- Borcherding, T. E./Deacon, R. T.* (1972), „The Demand for the Services of Non-federal Governments“, American Economic Review, Vol. 62, S. 891 - 901.
- Brueckner, J. K.* (1981), „Congested Public Goods: The Case of Fire Protection“, Journal of Public Economics, Vol. 15, S. 45 - 58.
- Champsaur, P.* (1975), „How to Share the Cost of a Public Good?“, International Journal of Game Theory, Vol. 4, S. 113 - 129.
- Feinstein, J.* (1992), „Public-Good Provision and Political Stability in Europe“, AEA Papers and Proceedings, Vol. 82, S. 323 - 329.
- Foley, D.* (1970), „Lindahl's Solution and the Core of an Economy with Public Goods“, Econometrica, Vol. 38, S. 66 - 72.
- Hildenbrand, W.* (1968), „The Core of an Economy with a Measure Space of Economic Agents“, Review of Economic Studies, Vol. 35, S. 443 - 452.
- Hildenbrand, W./Kirman, A.* (1976), „Introduction to Equilibrium Analysis“, North-Holland, Amsterdam.

- Kaneko, M.* (1976), „Examples that Lindahl Equilibria may not belong to the Core“, Discussion Paper.
- Mas-Colell, A.* (1980), „Efficiency and Decentralization in the Pure Theory of Public Goods“, *Quarterly Journal of Economics*, Vol. 94 , S. 625 - 641.
- Mas-Colell, A./Silvestre, J.* (1989), „Cost Share Equilibria: A Lindahlian Approach“, *Journal of Economic Theory*, Vol. 47, S. 239 - 256.
- Milleron, J.* (1972), „Theory of Value with Public Goods: A Survey Article“, *Journal of Economic Theory*, Vol. 5, S. 419 - 477.
- Muench, T.* (1972), „The Core and the Lindahl Equilibrium of an Economy with a Public Good“, *Journal of Economic Theory*, Vol. 4, S. 241 - 255.
- Oakland, W. H.* (1972), „Congestion, Public Goods and Welfare“, *Journal of Public Economics*, Vol. 1, S. 339 - 357.
- Olson, M./Zeckhauser, R.* (1966), „An Economic Theory of Alliances“, *Review of Economics and Statistics*, Vol. 48, S. 266 - 279.
- Pauly, M.* (1970), „Optimality, 'Public' Goods, and Local Government: A General Theoretical Analysis“, *Journal of Political Economy*, Vol. 78, S. 572 - 585.
- Roberts, D.* (1974), „A Note on Returns to Group Size and the Core with Public Goods“, *Journal of Economic Theory*, Vol. 9, S. 350 - 356.
- Vasil'ev, V./Weber, S./Wiesmeth, H.* (1992), „The Equivalence of Core and Lindahl Equilibria in an Economy with Semi-Public Goods“, *Economic Theory*, zur Veröffentlichung angenommen.
- Weber, S./Wiesmeth, H.* (1990), „On the Theory of Cost Sharing“, *Zeitschrift für Nationalökonomie/Journal of Economics*, Vol. 52, S. 71 - 82.
- Weber, S./Wiesmeth, H.* (1991a), „Economic Models of NATO“, *Journal of Public Economics*, Vol. 46 (1991), S. 181 - 197.
- Weber, S./Wiesmeth, H.* (1991b), „Issue Linkage in the European Community“, *Journal of Common Market Studies*, Vol. XXIX, S. 255 - 267.
- Williams, A.* (1966), „The Optimal Provision of Public Goods in a System of Local Government“, *Journal of Political Economy*, Vol. 74, S. 18 - 33.

Zusammenfassung der Diskussion

In der Diskussion wird mit der Berücksichtigung auftretender lokaler Überfüllungseffekte noch einmal der entscheidende neue Aspekt des vorgestellten Modells herausgestellt. Zusätzlich können normative Aspekte mittels der Einschränkung des Kerns allokativen Effekte ausüben.

Ferner wird klargestellt, daß Spillovereffekte entscheidend für den Nutzen sind, der aus einer Koalitionsbildung gezogen werden kann und nicht etwa, wie im Modell auch gar nicht angenommenen, positive Skalenerträge. Das Konzept des Lindahl Gleichgewichtes ist bei nicht konstanten

Skalenerträgen auch nicht anwendbar. Offen bleibt die Frage, ob konstante Skalenerträge bei regionalen Produktionsprozessen, wegen des Faktors Boden, überhaupt vorliegen können. Ebenso verhält es sich mit der Frage inwieweit es sich hier überhaupt um öffentliche Güter im engeren Sinne handelt.

Als weiterer interessanter Aspekt wird die Verträglichkeit der auftretenden finanzwissenschaftlichen Probleme mit dem Modellrahmen angesprochen.

Zur näherungsweisen Lösung asymmetrischer Rundreiseprobleme mit Hilfe Genetischer Algorithmen

Von *Reiner Wolff*, Paderborn

Abstract

Das Rundreiseproblem (traveling salesman problem) ist ein Beispiel eines ‚schwierigen‘ kombinatorischen Optimierungsproblems, für das gegenwärtig kein Lösungsalgorithmus mit polynomialer Zeitkomplexität bekannt ist. Wir untersuchen anhand einer größeren Anzahl von hypothetischen Testfällen mit 20, 40, 60 und 80 einbezogenen Orten, welchen Beitrag zufallsgesteuerte Suchverfahren in der Gestalt der Genetischen Algorithmen zur Ermittlung von Näherungslösungen für asymmetrische Rundreiseprobleme leisten können. Wir implementieren mehrere Crossover-Operatoren und konfrontieren die Ergebnisse der Genetischen Verfahren mit den Resultaten einer deterministischen Heuristik für asymmetrische Probleme.

1. Einführung

Das Rundreiseproblem oder Problem des Handlungsreisenden (traveling salesman problem) beinhaltet die Aufgabe, eine optimale Rundreise durch eine vorgegebene Anzahl endlich vieler Orte n (≥ 2) derart festzulegen, daß jeder Ort im Verlaufe der Tour genau einmal besucht wird. Dabei können unterschiedliche Kriterien wie etwa Reisezeit, Weglänge oder pekuniäre Aspekte für die Identifikation einer besten Route herangezogen werden. Wir wollen im weiteren generell von den ‚Kosten‘ der Rundreise sprechen, die minimiert werden sollen. Formal wird dadurch das Problem definiert, in einem bewerteten (möglicherweise gerichteten) Graphen $\tilde{G} = \langle V, \tilde{E}; C \rangle$ mit der Knotenmenge $V = \{1, \dots, n\}$, der Pfeilmenge \tilde{E} und der $n \times n$ -Kostenmatrix $C := (c_{ij})$ eine geschlossene Pfeilfolge zu finden, die jeden Knoten $i \in V$ genau einmal enthält und die Summe der zugehörigen Streckenkosten $c_{ij} \in \mathbb{R}_+$ minimal werden läßt. Die Existenz einer solchen Pfeilfolge ist gewährleistet, wenn \tilde{G} stark zusammenhängend ist, wenn also für je zwei Knoten $i, j \in V$ gilt, daß sowohl i von j aus als auch j von i aus erreicht werden kann. (Wenn der Graph einen Pfeil $\langle i, j \rangle$ nicht enthält, setzen wir $c_{ij} = \infty$. Im übrigen gilt $c_{ii} = \infty$ für alle $i \in V$.) In seiner allgemeinen Form stellt das Rundreiseproblem ein ‚schweres‘ kombinatorisches Optimie-

rungsproblem dar, für das gegenwärtig kein Lösungsalgorithmus mit polynomialer Zeitkomplexität $O(n^k)$, $k \in \mathbb{N}$, bekannt ist. (Auf spezielle Typen von Rundreiseproblemen mit polynomialem Rechenaufwand wollen wir hier nicht eingehen. Vgl. dazu *Gilmore/Lawler/Shmoys* [4].) Die exakte Lösung solcher Probleme vor allem mit Hilfe von Entscheidungsbaumverfahren (insbesondere branch and bound-Techniken) erfordert in der Regel einen exponentiellen Rechenaufwand und ist im übrigen oft sehr speicherintensiv.

Aus diesen Gründen hat man in der Literatur bereits frühzeitig begonnen, spezielle heuristische Lösungsmethoden für das Rundreiseproblem zu entwickeln, mit denen man in konkreten Fällen zufriedenstellende Näherungslösungen bei gleichzeitig vertretbarem Aufwand bestimmen kann. Da Rundreiseprobleme viele lokale Optima besitzen können, werden diese deterministischen Heuristiken häufig kombiniert als Eröffnungs- und Verbesserungsverfahren eingesetzt. Dazu zählen vor allem der Algorithmus von *Christofides* [3] (Eröffnungsverfahren) sowie die sogenannten *r*-opt-Algorithmen (Verbesserungsverfahren). (Ein besonders leistungsfähiger variabler *r*-opt-Algorithmus ist die Variante von *Lin/Kernighan* [13].) Die meisten dieser Verfahren sind ursprünglich für das symmetrische Rundreiseproblem konzipiert worden. In diesem Fall ist der Graph \tilde{G} symmetrisch, so daß es zu jedem Pfeil mit dem Anfangsknoten i und dem Endknoten j einen Pfeil mit dem Anfangsknoten j und dem Endknoten i gibt. Ferner gilt $c(i, j) = c(j, i)$ für alle $i, j \in V$. Wenn darüber hinaus die Knoten und (ungerichteten) Kanten eines solchen Graphen Punkten in der Euklidischen Ebene und ihren Euklidischen Entfernung entsprechen, ist die Dreiecksungleichung $c(i, j) \leq c(i, k) + c(k, j)$ für alle paarweise verschiedenen $i, j, k \in V$ erfüllt. Unter diesen Voraussetzungen stellt zum Beispiel das Verfahren von Christofides einen $\frac{1}{2}$ -approximativen Algorithmus dar, der stets Rundreisen ermittelt, deren Kosten höchstens um 50% über den Kosten einer besten Rundreise liegen (vgl. *Neumann/Morlock* [16, S. 450 - 451]).

In jüngerer Zeit wird zunehmend diskutiert, welchen Beitrag Genetische Algorithmen leisten können, um Lösungen von Rundreiseproblemen zu approximieren (vgl. *Goldberg/Lingle* [7], *Grefenstette* et al. [10], *Oliver/Smith/Holland* [17], *Whitley/Starkweather/Fuquay* [20], *Schöneburg/Heinzmann/Feddersen* [19, Kapitel 7]). Bei diesen Algorithmen handelt es sich um zufallsgesteuerte Heuristiken, die zuerst in den siebziger Jahren von *Holland* [11] den Ausleseprozessen in der belebten Natur nachempfunden worden sind und mittlerweile ein breites Anwendungsfeld in den Bereichen der Optimierung und Künstlichen Intelligenz besitzen (vgl. die Übersichten in *Goldberg* [5, S. 126 - 129 und S. 219 - 220]). In den Anwendungen auf das Rundreiseproblem dominiert bislang der symmetrische

Problemtyp, zumal hierfür am ehesten Testprobleme mit bekannten Lösungen oder Fehlerabschätzungen verfügbar sind. Demgegenüber werden wir in dieser Arbeit untersuchen, wie gut Genetische Algorithmen für die näherungsweise Lösung asymmetrischer Rundreiseprobleme geeignet sind, die im allgemeinen auch nicht der Dreiecksungleichung genügen. Derartige Probleme ergeben sich bei der Bestimmung bester Fahrtrouten, wenn in einem Verkehrsnetz eine Anzahl von Strecken nur in eine Richtung befahren werden darf (Einbahnstraßen), so daß die Hin- und Rückwege zwischen zwei Orten verschieden sind. Asymmetrien werden ferner auftreten, wenn die Wahl der Verkehrsmittel zwischen den Orten je nach den Zeittakten, Kapazitäten und Konditionen der verschiedenen Verkehrsträger richtungsabhängig ist. (*Müller-Merbach* [14, S. 65 - 66] skizziert vergleichbare Anwendungen aus dem Bereich der Planung von Maschinenbelegungen in der industriellen Fertigung.) Aus diesen Gründen wollen wir prüfen, ob die bisherigen guten Erfahrungen der verschiedenen Autoren mit der Behandlung symmetrischer Rundreiseprobleme durch Genetische Algorithmen auch für asymmetrische Rundreiseprobleme gelten.

Dieser Beitrag ist wie folgt aufgebaut. Zuerst legen wir im nächsten Abschnitt die informationstheoretischen Grundlagen der Genetischen Algorithmen im Überblick dar. Zu den Grundlagen gehören vor allem das Konzept der impliziten Parallelität und das Schema-Theorem von Holland. Anschließend stellen wir im dritten Abschnitt mehrere Genetische Operatoren für Rundreiseprobleme vor. Darunter befinden sich unterschiedliche sogenannte Crossover-Operatoren, die wir alternativ implementieren wollen. Im darauffolgenden vierten Abschnitt erläutern wir das Design von insgesamt 40 hypothetischen Rundreiseproblemen mit 20 - 80 zu besuchenden Orten. Für diese Probleme gewinnen wir eine Reihe von Näherungslösungen, denen wir anschließend die Resultate einer deterministischen Heuristik gegenüberstellen. (Bei dieser Heuristik handelt es sich um ein als Patching-Algorithmus bezeichnetes suboptimierendes Verfahren für asymmetrische Rundreiseprobleme.) Mit einer kritischen Einschätzung der Ergebnisse im letzten Abschnitt endet diese Arbeit.

2. Implizite Parallelität und Schema-Theorem

Die prinzipielle Vorgehensweise eines Genetischen Algorithmus und seine informationstheoretischen Grundlagen können wir bereits anhand eines einfachen Beispiels erläutern. Wir nehmen an, daß die Funktion $h : \mathbb{N} \rightarrow \mathbb{R}$ mit $h(x) = 1,2(50x - x^2)$ über dem Intervall $[10, 40]$ maximiert werden soll und daß vier Lösungsversuche $x_1 = 14$, $x_2 = 21$, $x_3 = 32$ und $x_4 = 39$ vorliegen. Sie bilden in der Terminologie der Genetischen Algorithmen eine Population von vier Individuen, die sich nacheinander als

sechsstellige binäre Zeichenketten „001110“, „010101“, „100000“ und „100111“ repräsentieren lassen. Jede Zeichenkette wird analog zu einem Chromosom als Träger der genetischen Informationen über das zugehörige Individuum interpretiert. (Entsprechend wird man die verschiedenen Positionen einer solchen Zeichenkette als Gene und ihre konkreten Ausprägungen als Allelen auffassen. Der Umstand, daß wir jedes Individuum lediglich durch ein einziges Chromosom charakterisieren und dessen Gene nur zwei Allelen besitzen, ist ohne weitere inhaltliche Bedeutung.) Mit Hilfe der Bewertungen $h(x_i)$ ($i = 1, \dots, 4$) werden nun Kennzahlen für die Fitneß der verschiedenen Individuen ermittelt. Zum Beispiel ergeben sich aufgrund der Vorschrift $f(x_i) := h(x_i) / \sum_{j=1}^4 h(x_j)$ ($i = 1, \dots, 4$) die Fitneßwerte 0,24, 0,29, 0,27 und 0,20 in Folge. Sie stellen den Ausgangspunkt für die Gewinnung neuer Lösungsversuche im Sinne einer neuen Population nach einem spezifischen Muster dar.

In einem ersten Schritt bildet man zunächst zufällig Paare von Individuen, und zwar derart, daß Individuen mit einer hohen Fitneß tendenziell (jedoch nicht zwangsläufig) eher als Individuen mit einer geringen Fitneß zum Zuge kommen. Zum Beispiel ist denkbar, daß man in jeweils zwei Auswahlrunden alle Fitneßwerte mit gleichverteilten Pseudo-Zufallszahlen gewichtet und in jeder Runde das Individuum mit der größten gewogenen Kennzahl aussucht. Wir wollen unterstellen, daß auf diesem Wege die Individuen x_1 und x_2 (das drittbeste Individuum und das beste Individuum der Population) als Paar neben anderen Paaren selektiert worden sind. Der nächste Schritt beinhaltet dann die Rekombination der genetischen Codes unter anderem von x_1 und x_2 , das sogenannte ‚crossover‘ oder ‚crossing over‘, das ebenfalls zufallsgesteuert abläuft. Eine mögliche Vorgehensweise ist, die zu x_1 und x_2 gehörenden binären Zeichenketten zufällig hinter einer der ersten fünf Positionen gemeinsam zu markieren, um anschließend die vor der Markierung plazierten Zeichen zwischen den beiden Zeichenketten zu tauschen. Bei einer Markierung nach der dritten Position gehen dann die Zeichenketten „001.110“ und „010.101“ in die Zeichenketten „010110“ und „001101“ über. Darin kann man im letzten Schritt einzelne Positionen (wiederum unter Zufallseinfluß) invertieren, was einer Mutation der jeweiligen Gene gleichkommt. Die Invertierung beispielsweise der niederwertigen sechsten beziehungsweise fünften Positionen führt zu den binären Code-Sequenzen „010111“ und „001111“, die Bestandteile einer neuen Population werden. Sie sind zu den Dezimalwerten $x_5 = 23$ und $x_6 = 15$ äquivalent und verbessern die ursprünglichen Lösungsversuche in Richtung auf die Optimalstelle $x^* = 25$. Man wird diesen Prozeß der Auswahl, Rekombination und Mutation binärer Zeichenketten abbrechen, wenn die Bewertungsfunktion h einen zureichend großen (gegebenenfalls den maximalen) Wert erreicht hat.

Für den soweit skizzierten Algorithmus ist charakteristisch, daß er Teile der genetischen Codes einzelner Individuen einer Population an spätere Generationen weitergibt, während er gleichzeitig andere Codestränge eliminiert. Das bedeutet, daß der Suchraum des Algorithmus faktisch aus Sequenzen der Form „*10*1*“ oder „110*1*“ mit dem Symbol * als Platzhalter besteht. Für diese Sequenzen hat *Holland* [11, Kapitel 4] den Begriff des ‚Schemas‘ geprägt. Man kann zeigen (vgl. *Goldberg* [5, S. 40 - 41]), daß die Anzahl der Schemata, die bei einer Population von n Individuen in einer Generation verarbeitet werden, von der Größenordnung $O(n^3)$ ist, obwohl lediglich n Auswertungen der Bewertungsfunktion erfolgen. Dieser bemerkenswerte Tatbestand wird als ‚implizite Parallelität‘ der Genetischen Algorithmen bezeichnet. Allerdings besitzen verschiedene Schemata im allgemeinen unterschiedliche Überlebenschancen, die sich jedoch abschätzen lassen. Dazu benötigen wir einige Begriffe und Symbole: Unter der Ordnung $o(H)$ eines gegebenen Schemas H versteht man in der Literatur die Anzahl seiner von * verschiedenen Positionen. Der Abstand zwischen der ersten und der letzten Position heißt definierende Länge $\delta(H)$ des Schemas. Eine sogenannte Instanz des Schemas H ist jede Zeichenkette, in der dieses Schema auftritt. Die Anzahl der Instanzen von H innerhalb einer Population (zu einem Zeitpunkt t) geben wir im weiteren mit $m(H, t)$ an. Die durchschnittliche Fitneß der Instanzen von H sei $f(H)$, und für die durchschnittliche Fitneß der gesamten Population schreiben wir kurz \bar{f} . Zuletzt sollen p_c und p_m die Wahrscheinlichkeiten sein, mit denen ein crossover zweier selektierter Individuen stattfindet beziehungsweise einzelne Positionen eines genetischen Codes mutieren. (In der Regel wird man $p_c = 1$ setzen.) Dann gilt die Abschätzung (zur Herleitung vgl. *Goldberg* [5, S. 28 - 33]):

$$m(H, t + 1) \geq m(H, t) \frac{f(H)}{\bar{f}} \left[1 - p_c \frac{\delta(H)}{n - 1} - p_m o(H) \right].$$

Folglich erzielen Schemata mit einer zugehörigen überdurchschnittlichen Fitneß $f(H)$, einer kurzen definierenden Länge $\delta(H)$ und einer niedrigen Ordnung $o(H)$ steigende Reproduktionsraten. Diese zentrale Aussage wird als Schema-Theorem bezeichnet. Sie charakterisiert in abstrakter Form die Bausteine („building blocks“, *Goldberg* [5, S. 41 - 45]) eines Genetischen Algorithmus.

Der nachstehende Pseudocode verdeutlicht, wie ein derartiger Algorithmus prinzipiell für eine elektronische Rechenanlage programmiert werden kann (ähnlich bei *Grefenstette* [8], *Grefenstette/Baker* [9]):

```

procedure GenAlg;
begin
    initialisiere Population  $P(0)$ ;
    bewerte  $P(0)$ ;
     $t := 0$ ;
repeat
     $t := t + 1$ ;
    selektiere  $P(t)$  aus  $P(t - 1)$ ;
    rekombiniere  $P(t)$ ;
    mutiere  $P(t)$ ;
    bewerte  $P(t)$ ;
until Lösung zufriedenstellend;
end.

```

Abbildung 1: Pseudocode eines Genetischen Algorithmus

Auf weiterführende Einzelaspekte der Programmierung Genetischer Verfahren wollen wir hier nicht eingehen.

3. Genetische Operatoren für das Rundreiseproblem

Bei der Implementation eines Genetischen Lösungsalgorithmus für das Rundreiseproblem müssen wir einige Besonderheiten des Rundreiseproblems beachten. Dazu zählt zunächst der Umstand, daß das Rundreiseproblem ein Minimierungsproblem ist. Das bedeutet, daß die Kennzahl für die Fitneß f_i eines Individuums i aus einer Population von m (≥ 2) Individuen umso höher anzusetzen ist, je niedriger die Kosten c_i der Route ausfallen, die durch das Individuum repräsentiert wird. Deshalb legen wir unseren Rechnungen die Fitneßfunktion $f_i = 1 - c_i / \sum_{j=1}^m c_j$ zugrunde, die wir allerdings noch modifizieren. Die Modifikation besteht in einer (affin) linearen Skalierung der Kennzahlen f_i mit dem Ergebnis, daß die kleinsten und die größten skalierten Fitneßwerte der verschiedenen Populationen stets den Wert 0,001 beziehungsweise 0,999 annehmen. Dadurch erreichen wir, daß der Selektionsdruck in jeder Generation von Individuen hoch bleibt. (Auf die theoretischen Implikationen alternativer Fitneßfunktionen und Skalierungsmethoden für das Konzept der impliziten Parallelität und das

Schema-Theorem wollen wir an dieser Stelle nicht eingehen. Vgl. hierzu *Grefenstette/Baker* [9].)

Eine weitere wichtige Besonderheit des Rundreiseproblems besteht darin, daß es sich dabei um ein Reihenfolgeproblem handelt. Somit darf der Fall nicht eintreten, daß die genetischen Operatoren der Rekombination und Mutation Reihenfolgen erzeugen, die ungültige Touren im Sinne des Rundreiseproblems sind. Unter diesen Aspekten erweisen sich binäre Zeichenketten als ineffiziente Codierungen genetischer Informationen (vgl. *Schöneburg/Heinzmann/Feddersen* [19, S. 272 - 273]). Deshalb ist man in der Literatur dazu übergegangen, Routen in Genetischen Algorithmen mit Hilfe von Zeigervektoren abzubilden, nachdem man zum Beispiel die n Orte eines Rundreiseproblems von 1 bis n durchnumeriert hat. In diesem Fall beschreibt der Vektor (1 5 3 2 4) eine gültige Rundreise durch fünf Orte in der durch seine Komponenten festgelegten Reihenfolge. (Man denkt sich die erste Komponente als Fortsetzung der letzten Komponente. Eine gleichwertige Tour wird durch den Vektor (5 3 2 4 1) festgelegt.) Dieser Vorgehensweise schließen wir uns an. Der Genetische Algorithmus transportiert dann Sequenzen der Form (* 5 3 * 4) oder (1 5 * * 4), die als Ordnungs-Schemata bezeichnet werden.

Für diese Form der Codierung gibt es zahlreiche Möglichkeiten, zwei Rundreisen durch Rekombination und Mutation zu einer neuen (gültigen) Rundreise zu verschmelzen. Wir verwenden in dieser Arbeit unter anderem Varianten zweier Crossover-Operatoren, die als besonders bewährt gelten. Dazu zählen das sogenannte ‚partially matched crossover‘ (PMX) von *Goldberg/Lingle* [7] und das ‚order crossover‘ (OX) von *Oliver/Smith/Holland* [17]. Konkret verfahren wir wie folgt, um aus je zwei Individuen einer Population einen Nachfolger zu generieren:

PMX: Zuerst kopieren wir den Inhalt eines zufällig positionierten zusammenhängenden Intervalls der Länge l ($1 \leq l \leq n - 1$) aus dem Zeigervektor des ersten Individuums an die gleiche Stelle im Zeigervektor des zweiten Individuums. Anschließend ersetzen wir darin alle duplizierten alten Einträge nacheinander durch diejenigen Komponenten, die als Folge des Kopiervorganges überschrieben worden sind. Auf diesem Weg entsteht zum Beispiel aus (1 6 . 5 3 . 2 4) und (5 1 . 4 2 . 3 6) der Nachfolger (4 1 5 3 2 6).

OX: Zunächst kopieren wir wie beim PMX-Operator. Im Anschluß daran löschen wir im Zeigervektor des zweiten Individuums alle duplizierten alten Einträge. Wir beginnen an der ersten Position rechts des Intervalls und lassen die verbleibenden alten Einträge von rechts aufrücken. Die somit links des Intervalls frei werdenden Stellen füllen wir mit den überschriebenen Intervalleinträgen in der

ursprünglichen Reihenfolge auf. Im Beispiel erhalten wir dann den Nachfolger (4 2 5 3 6 1).

In beiden Fällen starten wir mit $l = 1$. Solange die in einer Population repräsentierte kostenminimale Tour nicht verbessert werden kann, erhöhen wir l nach jeweils 50 Generationen um 1, falls noch nicht $l = n - 1$ gilt. Wenn eine bessere Tour gefunden wird, setzen wir l auf seinen Startwert zurück.

Die Operatoren PMX und OX ziehen für die Verschmelzung zweier Individuen einer Population keinerlei Informationen über die Kosten von Teilstrecken des zugrunde liegenden Graphen heran. Daher sind sie blind für die Güte einer einzelnen Ersetzung oder Vertauschung, die sie vornehmen. Güteabwägungen erfolgen ausschließlich global anhand der Gesamtkosten neu gebildeter Routen. Alternativ dazu kann man bereits in der Phase der Rekombination die Kosten der Teilstrecken als problemspezifische Informationen auswerten. Deshalb haben wir zu Vergleichszwecken den nachstehenden Greedy-Operator für das crossover zweier Individuen neben den Operatoren PMX und OX implementiert:

GX: Am Anfang bestimmen wir zufällig einen beliebigen Ort. Sodann suchen wir in den Zeigervektoren der beiden Individuen die sich jeweils unmittelbar an diesen Ort anschließenden Reisestationen. Wir wählen diejenige zulässige Station mit den geringsten zugehörigen Streckenkosten als nächsten Ort, und so weiter. Wenn zwei Stationen zu den gleichen Streckenkosten führen oder beide schon in der neu zu bildenden Tour enthalten sind, wählen wir zufällig einen Ort, der noch nicht berücksichtigt worden ist. Nach diesem Prinzip verfahren wir so lange, bis eine vollständige Rundreise vorliegt. Im Fall $n = 4$ mit der Kostenmatrix

$$\mathbf{C} = \begin{pmatrix} 0 & 10 & 21 & 12 \\ 15 & 0 & 5 & 8 \\ 17 & 14 & 0 & 20 \\ 13 & 9 & 7 & 0 \end{pmatrix}$$

und dem zweiten Ort als Ausgangspunkt einer neuen Rundreise erzeugt GX aus (3 2 4 1) und (4 3 2 1) die Route (2 4 3 1).

Damit ist dieser Operator eine stochastische Variante des deterministischen sogenannten Verfahrens des besten Nachfolgers (vgl. Müller-Merbach [15], S. 294 - 295). (Eine Version von GX für das symmetrische Rundreiseproblem und andere Greedy-Operatoren für symmetrische Probleme testet *Grennstette* [8, S. 49 - 53].)

Zuletzt wollen wir uns kurz dem Vorgang der Mutation im Zusammenhang mit Reihenfolgen zuwenden. Auch hierfür sind in der Literatur meh-

rere Operatoren diskutiert worden. Zum Beispiel kann man in einer vorgegebenen Rundreise die Reiserichtung entlang zufällig gewählter zusammenhängender Teilstrecken umkehren (die nach *Holland* [11, S. 106 - 109] sogenannte Inversion) oder zufällig zwei Orte miteinander vertauschen (vgl. etwa *Oliver/Smith/Holland* [17, S. 227]). Wir haben uns in dieser Arbeit durchgängig für eine dritte Alternative entschieden [vgl. *Schöneburg/Heinzmann/Feddersen* [19, S. 277 - 278]: In einem gegebenen Zeigervektor wählen wir fruestens nach der ersten Position zufällig ein Intervall der Länge v ($1 \leq v \leq n - 1$). Dieses Intervall verschieben wir um eine Stelle nach links und plazieren das linksseitig überschriebene Element unmittelbar hinter dem Intervall. Nach diesem Muster mutiert etwa (1 6 . 5 3 2 . 4) zu (1 5 3 2 6 4). Die Wahrscheinlichkeit dafür, daß ein Individuum einer Population einer derartigen Mutation unterworfen wird, sei p_m . Wir beginnen in allen Testrechnungen mit dem Startwert $p_m = 1$. Wird die kostenminimale Tour einer Population nicht verbessert, reduzieren wir p_m nach jeder weiteren 50. Generation um einen konstanten Betrag, so daß nach $(n - 1) / 50$ Generationen $p_m = 0$ ist. Diese Marke stellt das Abbruchkriterium unserer Genetischen Algorithmen dar. Für den Fall, daß eine günstigere Route ermittelt werden kann, starten wir neu bei $p_m = 1$.

4. Ergebnisse für 40 hypothetische Rundreiseprobleme

Wir werden nunmehr eine Reihe von Anwendungen Genetischer Algorithmen auf 40 hypothetische Rundreiseprobleme mit jeweils 10 Testfällen für 20, 40, 60 und 80 einbezogene Orte studieren. Dazu wählen wir die Elemente c_{ij} ($i \neq j$) der zugehörigen nicht-symmetrischen Kostenmatrizen C als ganzzahlige auf den Mengen $\{1, 2, \dots, n\}$ ($n = 20, 40, 60, 80$) gleichverteilte Pseudo-Zufallszahlen. (Den Zufall simulieren wir in allen Testläufen mit Hilfe der Routine RANO von *Press et al.* [18, S. 192 - 195].) Die Populationen in den verschiedenen Testrechnungen begrenzen wir für jedes gegebene n auf jeweils $m = 3n$ (für $n = 20, 40, 60$) und $m = 100$ (für $n = 80$) Individuen, die aus derselben zufällig initialisierten Ausgangspopulation durch die Genetischen Operatoren der Selektion, Rekombination und Mutation hervorgehen. Gleichzeitig wenden wir eine elitistische Strategie an. Sie besteht darin, daß wir stets ein bestes Individuum einer Population in die nachfolgende Generation übernehmen. Auf diese Weise gewinnen wir durch den Einsatz der Crossover-Operatoren PMX, OX und GX eine größere Anzahl von Einzelergebnissen, die wir zuletzt zu Vergleichszwecken mit den Resultaten einer deterministischen Heuristik konfrontieren.

Bei dieser Heuristik handelt es sich um den sogenannten ‚Patching-Algorithmus‘ (PA) (vgl. *Neumann/Morlock* [16, S. 456 - 459]). Er macht sich

den Tatbestand zunutze, daß man das Rundreiseproblem auch als ein Problem der wechselseitigen Zuordnung von Orten auffassen kann. Tatsächlich geht das Rundreiseproblem in das Zuordnungsproblem der Linearen Optimierung über, wenn man die Forderung aufgibt, daß die gesuchte optimale Zuordnung einen geschlossenen Zyklus bildet. Somit stellt das Zuordnungsproblem faktisch eine Relaxation des Rundreiseproblems dar. Daher können wir wie folgt vorgehen: Zunächst lösen wir das zu einer gegebenen Kostenmatrix C gehörende Zuordnungsproblem mit Hilfe der Ungarischen Methode. (Der Rechenaufwand hierfür beträgt $O(n^3)$.) Wir erhalten dann in der Regel ein System von Kurzzyklen, deren Anzahl bei nicht-symmetrischen Kostenmatrizen jedoch im allgemeinen klein ist. (Solche Zyklen berücksichtigen lediglich einen Teil der n Orte. In der Gesamtheit der Kurzzyklen kommen jedoch alle Orte vor.) Anschließend verschmelzen wir die gefundenen Kurzzyklen nacheinander (möglichst zuerst die Zyklen mit vielen Orten), so daß der entstehende Kostenzuwachs in jedem Schritt kleinstmöglich wird. Der damit beschriebene Patching-Algorithmus terminiert, wenn alle Kurzzyklen verschmolzen sind. (Der Gesamtaufwand für das Verfahren besitzt ebenfalls die Zeitkomplexität $O(n^3)$.)

Die Tabellen 1 - 3 im Appendix 1 geben die mit den Genetischen Algorithmen und der Patching-Heuristik bestimmten Näherungswerte für die Kostenminima in den insgesamt 30 Problemen mit 20, 40 und 60 zu bereisenden Orten im einzelnen wieder. Es zeigt sich, daß die Genetischen Algorithmen in knapp zwei Dritteln aller Fälle die günstigeren Touren finden. Diese Aussage beruht vor allem auf denjenigen Genetischen Verfahren, die die Crossover-Operatoren OX und GX (bei leichten Vorteilen für GX) verwenden. Die von den Genetischen Algorithmen erzeugten Entwicklungsverläufe der durchschnittlichen minimalen Rundreisekosten können wir anhand der Abbildungen des Appendix 2 (in den interessanteren Bereichen) verfolgen. Darin beobachten wir für die OX-basierten Verfahren ein gleichmäßiges durchschnittliches Konvergenzverhalten, während die mit PMX und GX arbeitenden Varianten bereits (für GX nicht unerwartet) frühzeitig stagnieren. Die Tabellen 1 - 3 dokumentieren noch eine weitere bemerkenswerte Tatsache: Der Patching-Algorithmus schneidet umso besser ab, je größer die Anzahl der Orte in den Rundreiseproblemen ist (vgl. dazu auch *Karp/Steele* [12, S. 195 - 203]). Bei 60 berücksichtigten Orten fallen seine Ergebnisse bereits in den meisten Einzelrechnungen und im Durchschnitt erheblich besser als die Resultate des PMX-gestützten Genetischen Verfahrens aus. Dabei benötigt der Patching-Algorithmus auf der benutzten Rechenanlage (Personal-Computer mit EISA-Bus, 80486-CPU, 256 KByte 2nd level cache und 40 MHz Taktfrequenz) nur wenige Sekunden, um ein Einzelergebnis zu ermitteln. Demgegenüber entsteht hierfür bei den Genetischen Verfahren ein Zeitbedarf im Umfang von bis zu einigen Stunden (in der

angegebenen Parametrisierung bei 40 und mehr Orten). Deshalb beschränken wir uns für die Fälle von 80 einbezogenen Orten bei den Genetischen Algorithmen auf die OX-basierte Variante mit einer Population von 100 Individuen. Tabelle 4 weist für diese Fälle die deterministische Patching-Heuristik in der größten Zahl der Einzelrechnungen ebenso wie im Durchschnitt als die deutlich überlegene Lösungsstrategie aus.

Auf die Diskussion weiterer Resultate wollen wir verzichten. An dieser Stelle soll der Hinweis genügen, daß zusätzliche Testrechnungen mit einem als „edge recombination crossover“ (vgl. *Whitley/Starkweather/Fuquay* [20]) bezeichneten kantenorientierten Crossover-Operator zu keinen qualitativ anderen Resultaten geführt haben. (Für $n = 20$ und $n = 40$ stellen sich Touren mit geringfügig niedrigeren Kosten als unter PMX nach wesentlich längeren Laufzeiten des Genetischen Algorithmus ein.) Wir haben auch mit kleinen Populationen nach der von *Goldberg* [6] vorgeschlagenen Methode experimentiert und vergleichbare Ergebnisse erzielt.

5. Schlußfolgerungen

Unsere Tests hinterlassen einen zwiespältigen Eindruck. Einerseits haben sich die Genetischen Verfahren im Vergleich mit dem Patching-Algorithmus bei den asymmetrischen Rundreiseproblemen mit 20, 40 und 60 vertretenen Orten gut bewährt. Probleme dieser Größenordnung lassen sich häufig aber noch exakt lösen (zum Beispiel mit Hilfe von branch and bound-Verfahren). Andererseits deuten unsere übrigen Ergebnisse darauf hin, daß man in Fällen vieler zu bereisender Orte mit der deterministischen Heuristik kostengünstigere Routen in sehr viel kürzerer Zeit erhält. Dabei haben wir noch nicht berücksichtigt, daß bereits die Parametrisierung der Genetischen Verfahren große Sorgfalt erfordert und in der Regel zeitaufwendig ist. In diesem Zusammenhang müssen wir jedoch einräumen, daß die Genetischen Algorithmen massiv parallelisierbar sind. Prinzipiell lassen sich nämlich die selektierten Individuen einer Genetischen Population gleichzeitig rekombinieren. Das gilt entsprechend für die Mutation (und bereits für die Selektion) der Individuen. Insofern ist das kritische Laufzeitverhalten der Genetischen Algorithmen auch die Folge ihrer prozeduralen Implementation. Auf Parallelrechnern (zum Beispiel auf Transputer-Basis) sollten wesentlich kürzere als die von uns festgestellten Laufzeiten zu realisieren sein.

Im übrigen dürfen wir uns den Blick auf die eigentlichen Stärken der Genetischen Verfahren nicht verstellen lassen. Diese Stärken liegen nach unserer Auffassung in der Flexibilität insbesondere der „blind“ Varianten, die sich ausschließlich an globalen Bewertungen der Genetischen Indivi-

duen orientieren. Für diese Varianten stellt das eigentliche Rundreiseproblem eine black box dar. Somit können wir ohne größere Schwierigkeiten zusätzliche Restriktionen berücksichtigen, indem wir zum Beispiel bestimmte Folgen von Teilstrecken in der Bewertungsfunktion mit Strafpunkten (penalties) belegen, falls eine konkret vorliegende Problemstellung das erfordert. Auf diese Weise können wir auch mehrfache (und einander widerstrebende) Zielsetzungen von Rundreisen erfassen. (Das ist auch ein großer Vorteil der sogenannten simulated annealing-Methode von *Aarts/Korst* [1]. Vgl. *Press* et al. [18, S. 329].) Damit dürften die Genetischen Verfahren besonders für solche praktischen Anwendungen interessant sein, die sich nicht in das starre Korsett des Standard-Rundreiseproblems und seiner deterministischen Lösungsmethoden zwängen lassen. Hier sehen wir noch Bedarf an weiterer Forschung.

Appendix 1: Ergebnistabellen

Tabelle 1
Ergebnisse der Tests für 20 Orte

$m = 60$	1	2	3	4	5	6	7	8	9	10	Durchschnitt
PMX	51	63	51	62	60	46	59	58	57	54	56,1
OX	47	39	47	55	43	46	47	57	56	52	48,9
GX	53	39	47	50	44	41	53	51	51	51	48,0
PA	63	55	39	53	83	73	71	90	51	45	62,3

Tabelle 2
Ergebnisse der Tests für 40 Orte

$m = 120$	1	2	3	4	5	6	7	8	9	10	Durchschnitt
PMX	157	161	183	159	140	133	171	142	166	163	157,5
OX	123	127	131	132	141	141	143	114	123	128	130,3
GX	129	128	113	139	110	126	126	116	123	136	124,6
PA	140	162	117	189	146	148	184	126	215	205	163,2

Tabelle 3
Ergebnisse der Tests für 60 Orte

$m = 180$	1	2	3	4	5	6	7	8	9	10	Durchschnitt
PMX	347	326	284	283	313	288	321	268	339	298	306,7
OX	255	247	271	278	218	218	232	239	269	248	247,5
GX	238	227	221	277	257	261	268	223	193	264	242,9
PA	155	272	167	374	238	384	260	225	171	299	254,5

Tabelle 4
Ergebnisse der Tests für 80 Orte

$m = 100$	1	2	3	4	5	6	7	8	9	10	Durchschnitt
OX	371	399	413	369	338	329	354	306	363	367	360,9
PA	334	295	335	325	252	250	376	406	346	234	315,3

Appendix 2: Kostenverläufe

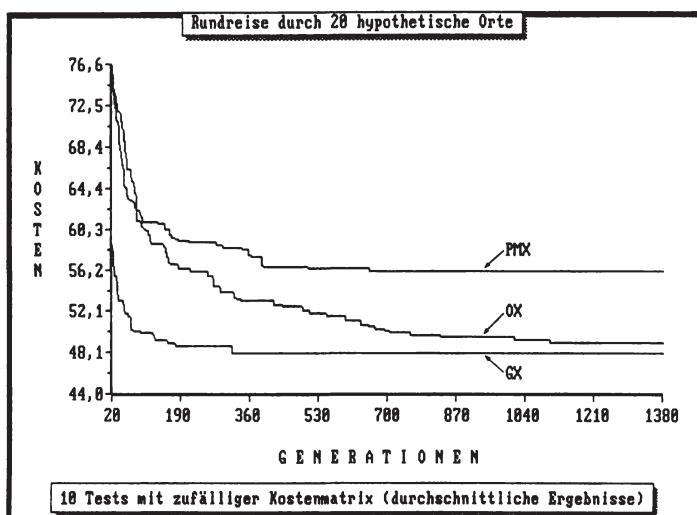


Abbildung 2: Ergebnisse der Tests für 20 Orte

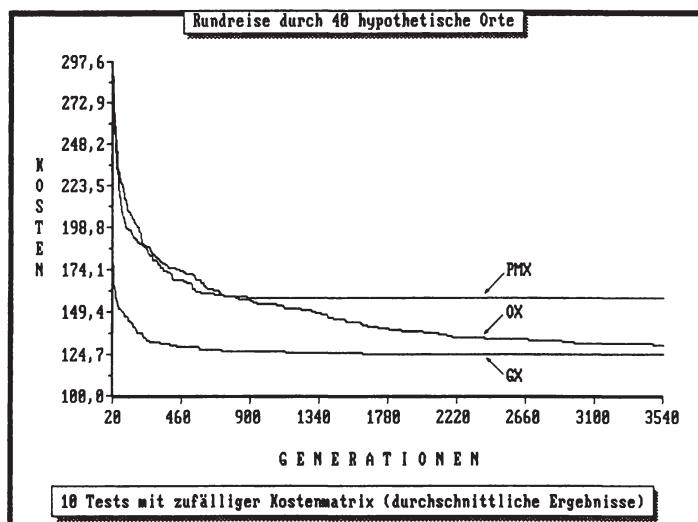


Abbildung 3: Ergebnisse der Tests für 40 Orte

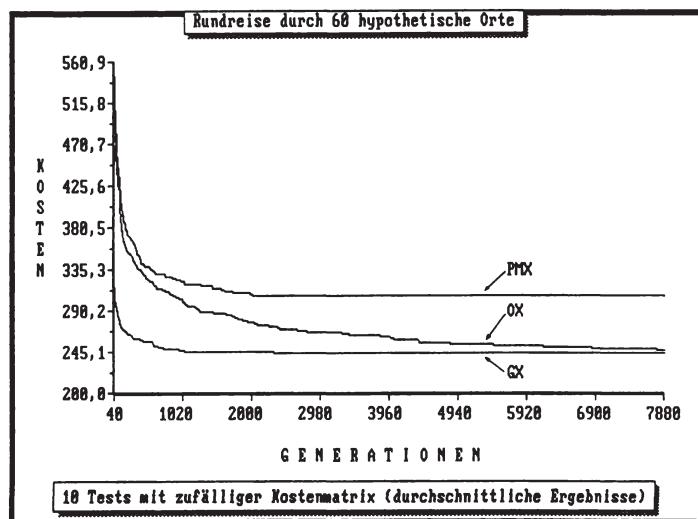


Abbildung 4: Ergebnisse der Tests für 60 Orte

Literatur

1. *Aarts, E. H. L./Korst, J.*, „Simulated Annealing and Boltzmann Machines, A Stochastic Approach to Combinatorial Optimization and Neural Computing,“ John Wiley & Sons, Chichester, 1989.
2. *Bridges, C. L./Goldberg, D. E.*, An Analysis of Reproduction and Crossover in a Binary-Coded Genetic Algorithm, „Genetic Algorithms and Their Applications: Proceedings of the Second Conference on Genetic Algorithms“ (J. J. Grefenstette, Ed.), 9 - 13, Lawrence Erlbaum, Hillsdale/N. J., 1987.
3. *Christofides, N.*, Worst-Case Analysis of a New Heuristic for the Traveling Salesman Problem, Management Science Report No. 388, Carnegie-Mellon University, 1976.
4. *Gilmore, P. C./Lawler, E. L./Shmoys, D. B.*, Well-Solved Special Cases, „The Traveling Salesman Problem“ (E. L. Lawler/J. K. Lenstra/A. H. G. Rinnooy Kan/D. B. Shmoys, Eds.), 87 - 143, John Wiley & Sons, New York, 1985.
5. *Goldberg, D. E.*, „Genetic Algorithms in Search, Optimization, and Machine Learning,“ Addison-Wesley, Reading/Mass., 1989.
6. — Sizing Populations for Serial and Parallel Genetic Algorithms, „Proceedings of the Third International Conference on Genetic Algorithms,“ (J. D. Schaffer, Ed.), 70 - 79, Morgan Kaufmann, San Mateo/Cal., 1989. Wiederabgedruckt in B. P. Buckles/F. E. Petry, Eds., „Genetic Algorithms,“ 20 - 29, IEEE Computer Society Press, Los Alamitos/Cal., 1992.
7. *Goldberg, D. E./Lingle, R.*, Alleles, Loci, and the Traveling Salesman Problem, „Proceedings of an International Conference on Genetic Algorithms and Their Applications“ (J. J. Grefenstette, Ed.), 154 - 159, Lawrence Erlbaum, Hillsdale/N. J., 1985.
8. *Grefenstette, J. J.*, Incorporating Problem Specific Knowledge into Genetic Algorithms, „Genetic Algorithms and Simulated Annealing“ (L. Davis, Ed.), 42 - 60, Pitman, London, 1987.
9. *Grefenstette, J. J./Baker, J. E.*, How Genetic Algorithms Work: A Critical Look at Implicit Parallelism, „Proceedings of the Third International Conference on Genetic Algorithms,“ (J. D. Schaffer, Ed.), 20 - 27, Morgan Kaufmann, San Mateo/Cal., 1989. Wiederabgedruckt in B. P. Buckles/F. E. Petry, Eds., „Genetic Algorithms,“ 12 - 19, IEEE Computer Society Press, Los Alamitos/Cal., 1992.
10. *Grefenstette, J. J./Gopal, R./Rosmaita, B./Gucht, D. van*, Genetic Algorithms for the Traveling Salesman Problem, „Proceedings of an International Conference on Genetic Algorithms and Their Applications“ (J. J. Grefenstette, Ed.), 160 - 168, Lawrence Erlbaum, Hillsdale/N. J., 1985.
11. *Holland, J. H.*, „Adaptation in Natural and Artificial Systems, An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence,“ MIT Press, Cambridge/Mass., 1992. (Erste Ausgabe: University of Michigan Press, Ann Arbor, 1975.)

12. *Karp, R. M./Steele, J. M.*, Probabilistic Analysis of Heuristics, „The Traveling Salesman Problem“ (E. L. Lawler/J. K. Lenstra/A. H. G. Rinnooy Kan/D. B. Shmoys, Eds.), 181 - 205, John Wiley & Sons, New York, 1985.
13. *Lin, S./Kernighan, B. W.*, An Effective Heuristic Algorithm for the Traveling Salesman Problem, Operations Research 21 (1973), 498 - 516.
14. *Müller-Merbach, H.*, „Optimale Reihenfolgen,“ Springer, Berlin, 1970.
15. — „Operations Research,“ 3. Aufl., Vahlen, München, 1973.
16. *Neumann, K./Morlock, M.*, „Operations Research,“ Carl Hanser, München, 1993.
17. *Oliver, I. M./Smith, D. J./Holland, J. R. C.*, A Study of Permutation Crossover Operators on the Traveling Salesman Problem, „Genetic Algorithms and Their Applications: Proceedings of the Second Conference on Genetic Algorithms“ (J. J. Grefenstette, Ed.), 224 - 230, Lawrence Erlbaum, Hillsdale/N. J., 1987.
18. *Press, W. H./Flannery, B. P./Teukolsky, S. A./Vetterling, W. T.*, „Numerical Recipes: The Art of Scientific Computing,“ 3rd ptg., Cambridge University Press, Cambridge, 1987.
19. *Schöneburg, E./Heinzmann, F./Feddersen, S.*, „Genetische Algorithmen und Evolutionsstrategien, Eine Einführung in Theorie und Praxis der simulierten Evolution,“ Addison-Wesley, Bonn, 1994.
20. *Whitley, D./Starkweather, T./Fuquay, D.*, Scheduling Problems and Traveling Salesman: The Genetic Edge Recombination Crossover, „Proceedings of the Third International Conference on Genetic Algorithms,“ (J. D. Schaffer, Ed.), 133 - 140, Morgan Kaufmann, San Mateo/Cal., 1989.

Zusammenfassung der Diskussion

Im Mittelpunkt der Diskussion stehen das Verhalten und die Arbeitsweise des genetischen Näherungsverfahrens. Es wird festgestellt, daß die erfolgreiche Anwendung des Algorithmus stark von einer günstigen Wahl der Rahmenparameter (Größe der Population, Mutationshäufigkeit und dem Vererbungsintervall) abhängt. Werden die ungünstig gewählt, so kommt es entweder zu einem unbefriedigenden Annäherung oder das Verfahren kommt nur sehr langsam voran. Ein geübter Anwender, so Herr Wolf, entdeckt allerdings schon nach wenigen Probeläufen eine günstige Parameterkonstellation. Interessant und noch ungeklärt ist, was der Anwender hierbei eigentlich besser macht als der Computer, da dieser die geeigneten Parameterkonstellationen selbst bestimmen könnten müßte.

Eine Verwendung von exakten Lösungen für Teilprobleme wäre, unter zusätzlichem Speicher- und Rechenzeitbedarf, zwar prinzipiell möglich, kann aber häufig zur Entwertung der eigentlichen Stärke des Verfahrens führen.

Bemerkenswert ist auch, daß das Verfahren über die einzelnen Iterationsstufen sehr unterschiedliche Fortschritte bzgl. der Minimierung der Zielgröße macht. Ob diese beobachteten Sprünge evtl. eher im Verfahren provoziert werden könnten, ist noch nicht geklärt.

Bei größeren Problemen (mehr als 20 Orte), existieren meistens keine exakten Lösungen, so daß eine Beurteilung der Güte des Verfahrens häufig schwierig ist. Allerdings sollen solche für die USA und die alten Bundesländer inzwischen vorliegen, wodurch neue Testmöglichkeiten bestehen.

Räumliche Preistechniken im Wettbewerb

Von *Klaus Schöler*, Kiel

Abstract

In diesem Beitrag werden einfache Modelle des räumlichen Wettbewerbs entwickelt, um die Marktergebnisse bei Verwendung alternativer Preistechniken zu diskutieren. Modell 1 ist gekennzeichnet durch identische Standorte der Anbieter und vollständig überlappende Marktgebiete; in Modell 2 werden unterschiedliche Standorte und nichtüberlappende Marktgebiete der Firmen unterstellt. In Modell 1 zeigt sich, daß nicht nur jenes Unternehmen seine Marktsituation verbessert, das zur überlegenen Strategie wechselt, sondern daß auch die Marktsituation des Unternehmens eine positive Änderung erfährt, das – aus welchen Gründen auch immer – beabsichtigt, bei der suboptimalen Strategie zu verbleiben. Für Modell 2 werden die Resultate bei Bertrand-Wettbewerb und bei monopolistischer Konkurrenz untersucht. Während im ersten Fall sich eine eindeutige Rangfolge der Preistechniken hinsichtlich des Gewinns ergibt, kann im zweiten Fall gezeigt werden, daß die Rangfolge entscheidend von der Standortentfernung der Anbieter abhängt.

1. Einführung

Die Diskussion räumlicher Preistechniken ist in der Literatur seit vielen Jahren ausführlich geführt worden (vgl. z.B. *M. Beckmann* [1976], *J. G. Greenhut* [1977], *M. L. Greenhut/G. Norman/C. Hung* [1987], *H. Ohta* [1988]). Neben den weniger beachteten Formen „uniform pricing“ und „Zonenpreise“ werden vor allem zwei Techniken miteinander verglichen: fob pricing und discriminatory pricing. Im ersten Fall setzt der Anbieter einen Verkaufspreis (mill price, Ab-Werk-Preis) am Produktionsstandort fest, und die Endpreise der räumlich gestreuten Nachfrager ergeben sich aus dem Verkaufspreis und den tatsächlich anfallenden Transportkosten zwischen Produktions- und Konsumtionsstandort. Dabei ist es unerheblich – unterstellt man einen vollkommenen Markt für Transportleistungen und die Abwesenheit von scale-Effekten im Transport –, ob Käufer oder Verkäufer den Transport durchführen. Im zweiten Fall – discriminatory pricing – setzt der Anbieter die Endpreise (Ortspreise) an jedem Verbrauchsort fest, wobei es sich als gewinnmaximierend erweist, den vom Produktionsstandort weiter entfernten Käufern nicht die gesamten Transportkosten anzulasten und zum Ausgleich dafür von den Käufern in unmittelbarer Standortnähe

einen die Transportkosten übersteigenden Aufschlag zu verlangen. Den jeweiligen Konsumenten ist entweder nur der eigene Endpreis oder aber die Menge aller Endpreise bekannt. Die Ergebnisse eines Vergleiches zwischen fob pricing und discriminatory pricing können wie folgt zusammengefaßt werden: discriminatory pricing führt unter sonst gleichen Bedingungen zu höheren Gewinnen, niedrigeren Konsumentenrenten und höheren Wohlfahrtseffekten als fob pricing. Vergleicht man jedoch die Pro-Kopf-Wohlfahrtseffekte, so ist fob pricing der räumlichen Preisdiskriminierung gegenüber überlegen (vgl. *K. Schöler* [1988], S. 144ff., *R. J. Claycombe* [1991]).

Bei den nachfolgenden Überlegungen soll der unterschiedliche Einsatz der Preistechniken im Mittelpunkt stehen (vgl. *J. F. Thisse/X. Vives* [1988], *S. P. Anderson/J. F. Thisse* [1988] sowie *K. Schöler* [1990]). Zu diesem Zweck werden die denkbaren Kombinationen für zwei Firmen und zwei Preistechniken – fob und discriminatory pricing – in einfachen Modellen untersucht. Die Kombination, die eine Nash-Lösung erzeugt, wird mit jenen Kombinationen verglichen, bei denen eine Firma oder beide Firmen Preistechniken anwenden, die nicht zum Nash-Gleichgewicht führen. Für dieses Verhalten können exogene Restriktionen verantwortlich sein, wie etwa ein staatliches Verbot diskriminierender Preistechniken oder die Einschätzung dieser Preistechnik als „unfaire“ Marktverhaltensweise durch die Käufer. Die Analyse wird im Rahmen unterschiedlicher Modellannahmen durchgeführt, wobei im ersten Modell Substitutionsgüter und im zweiten Modell – sieht man der räumlichen Komponente ab – homogene Güter unterstellt werden.

Der Beitrag ist wie folgt aufgebaut: In Abschnitt 2 wird das erste Modell des räumlichen Wettbewerbs dargestellt. In diesem Modell haben die beiden betrachteten Firmen einen gemeinsamen Standort in der Mitte des Gesamtmarktgebietes und vollständig überlappende firmenindividuelle Marktgebiete. In Abschnitt 3 werden die Marktergebnisse bei unterschiedlichen Preistechniken und Substitutionsbeziehungen diskutiert. Abschnitt 4 stellt das zweite Modell vor, in dem unterschiedliche Standorte und nicht-overlappende firmenindividuelle Marktgebiete angenommen werden. Dieses Modell wird in Abschnitt 5 unter der Annahme des Bertrand-Wettbewerbs und in Abschnitt 6 unter der Annahme der monopolistischen Konkurrenz untersucht. Der Vergleich der Ergebnisse aus Abschnitt 5 und 6 steht im Mittelpunkt des Abschnitts 7. In Abschnitt 8 werden einige allgemeine Schlüssefolgerungen gezogen.

2. Annahmen und Modell 1

Für die nachfolgenden Überlegungen ist es sinnvoll, einige vereinfachende Annahmen einzuführen und eine Reihe von Variablen auf 1 zu stan-

dardisieren, ohne daß dadurch die Allgemeinheit der Modellaussagen eingeschränkt wird.

A1: Zwei Anbieter sind im Zentrum eines linearen Marktgebietes mit der Ausdehnung L angesiedelt. Innerhalb des Marktgebietes beträgt die Nachfragerdichte je Entfernungseinheit r genau 1 und außerhalb 0. Die firmenindividuellen Marktgebiete überlappen sich in allen Fällen vollständig und weisen die Ausdehnung $2R (= L)$ auf.

A2: Die Nachfrage je Konsument q_i nach dem Gut des Anbieters i beträgt

$$(1) \quad q_i = 1 - m_i - r + \beta m_j + \beta r$$

beziehungsweise nach dem Gut des Anbieters j

$$(2) \quad q_j = 1 - m_j - r + \beta m_i + \beta r$$

wobei m_i und m_j die Ab-Werk-Preise der Anbieter i und j , r die Entfernung zwischen Produktions- und Konsumtionsstandort und β als symmetrischer Koeffizient in den Nachfragefunktionen (1) und (2) die Stärke der Substitutionsbeziehungen zwischen beiden Gütern angeben. Es liegt nahe, β auf den Bereich $(0,1)$ zu begrenzen; je größer β ist, um so stärker wird von den Nachfragern die Substitution zwischen beiden Gütern erwogen. Die Transportkosten je Mengen- und Entfernungseinheit zwischen Produktions- und Konsumtionsstandort sollen auf 1 standardisiert werden.

A3: Aus Gründen der Vereinfachung des Modells soll angenommen werden, daß die Kosten der Produktion 0 betragen. Positive Produktionskosten haben keinen Einfluß auf die Rangfolge der Marktergebnisse.

A4: Die Firmen verfolgen das Ziel der Gewinnmaximierung. Dabei sollen ihre Reaktionen auf Preisvariationen des jeweiligen Konkurrenten durch den Hotelling-Smithies-Wettbewerb ($dm_i/dm_j = dm_j/dm_i = 0$) gekennzeichnet sein.

A5: Den Firmen stehen prinzipiell zwei räumliche Preistechniken zur Verfügung. Sie können eine nicht-diskriminierende Preistechnik (fob pricing) und eine diskriminierende Preissetzung (discriminatory pricing) anwenden.

Aus der Forderung der Nichtnegativität der Nachfrage für das gesamte Marktgebiet ergeben sich für die Ab-Werk-Preise die nachfolgenden Bedingungen:

$$(3) \quad 0 \leq m_i \leq 1 - R + \beta(m_j + R), \quad 0 \leq m_j \leq 1 - R + \beta(m_i + R)$$

Der Gewinn (oder Erlös bei Vernachlässigung der Produktionskosten) beträgt unabhängig von der angewandten Preistechnik an einem beliebigen Ort im Marktgebiet, der durch r gekennzeichnet sei:

$$\pi_i(r) = m_i(1 - m_i + \beta m_j - r + \beta r)$$

bzw.

$$(4) \quad \pi_j(r) = m_j(1 - m_j + \beta m_i - r + \beta r)$$

Nunmehr ist es möglich, die Gewinne (Erlöse) der beiden Anbieter i und j für die nachfolgenden Kombinationen zu ermitteln:

Fall 1: Beide Anbieter wenden fob pricing an.

Fall 2: Beide Anbieter wenden discriminatory pricing an.

Fall 3: Anbieter i wendet discriminatory pricing, Anbieter j fob pricing an.

Fall 4: Anbieter i wendet fob pricing, Anbieter j discriminatory pricing an.

Da im Fall 4 lediglich die Indizes gegenüber Fall 3 vertauscht werden, ist es nicht notwendig, diesen Fall ausführlich zu diskutieren. Weitere Einzelheiten zur Bestimmung der Gewinne finden sich in Anhang 1.

Fall 1: Da beide Anbieter fob pricing anwenden, erhält man aus den Gewinngleichungen

$$(5) \quad \Pi_i = 2 \int_0^R \pi_i(r) dr, \quad \Pi_j = 2 \int_0^R \pi_j(r) dr$$

unter Verwendung der Gleichgewichtspreise

$$(6) \quad \Pi_i^* = \Pi_j^* = (R(\beta R + 2 - R)^2) / (2(\beta - 2)^2)$$

Die Höhe des Gewinns wird von der Stärke der Substitution und der Ausdehnung des gemeinsamen Gesamtmarktgebietes bestimmt.

Fall 2: Beide Anbieter verwenden die Strategie, an jedem Ort im Marktgebiet einen anderen Lieferpreis (Ortspreis) von den Käufern zu verlangen, der auf den Ort r bezogen gewinnmaximal ist. Aus Gleichung (4) erhält man die Preise

$$(7) \quad m_i = (\beta(m_j + r) + 1 - r)/2, \quad m_j = (\beta(m_i + r) + 1 - r)/2$$

und im Gleichgewicht Gewinne von:

$$(8) \quad \Pi_i^* = \Pi_j^* = (2R(\beta^2 R^2 - \beta R(2R - 3) + R^2 - 3R + 3)) / (3(\beta - 2)^2)$$

Auch im Fall 2 liegen symmetrische Gewinne im Gleichgewicht vor.

Fall 3: Die Firma i verfolgt die Strategie der optimalen Preisdiskriminierung, und Anbieter j verwendet fob pricing. Die entsprechenden Preise lauten:

$$m_i = (\beta(m_j + r) + 1 - r) / 2 \quad \text{und} \quad m_j = (\beta(2m_i + R) - R + 2) / 4$$

woraus sich für Firma i ein Gleichgewichtsgewinn von

$$(9) \quad \Pi_i^* = (R^3(3\beta^4 + 6\beta^3 - 5\beta^2 - 20\beta + 16) + 12R^2 \\ (\beta - 1)(\beta^2 + 4\beta + 4) + 12R(\beta^2 + 4\beta + 4)) / (6(\beta^2 - 4)^2)$$

und für Firma j von

$$(10) \quad \Pi_j^* = (R^3(-\beta^5 + 6\beta^4 + 7\beta^3 - 16\beta^2 - 8\beta + 12) + 12R^2 \\ (\beta - 1)(\beta^2 + 4\beta + 4) + 12R(\beta^2 + 4\beta + 4)) / (6(\beta^2 - 4)^2)$$

ergibt.

Fall 4: Die Firma i wendet die Strategie fob pricing, die Firma j discriminatory pricing an. Die entsprechenden Indizes der Gleichungen (9) und (10) werden vertauscht.

Ein Vergleich der Gewinnresultate zeigt für die Annahmen $0 < \beta < 1$ und $0 < R < 2/3$ die nachstehende Rangfolge $\Pi_{dd} > \Pi_{df} > \Pi_{fd} > \Pi_{ff}$, wobei der erste Index die Preisstrategie der betrachteten Firma und der zweite Index die Preisstrategie des Konkurrenten verdeutlichen sowie f für „fob pricing“ und d für „discriminatory pricing“ stehen. Als erstes Resultat kann festgehalten werden (s. Anhang 2):

Satz 1: Wenden beide Firmen die Strategie discriminatory pricing an, so stellt sich ein Nash-Gleichgewicht ein; keine der beiden Firmen sieht sich – bei Abwesenheit von Restriktionen für die Anwendung von Preistechniken – veranlaßt, eine andere Preistechnik zu wählen. Geht eine Firma von fob pricing zu discriminatory pricing über, so erzielt nicht nur dieses Unternehmen einen höheren Gewinn, sondern auch das Unternehmen, das bei der suboptimalen Preisstrategie fob pricing verbleibt.

6 Todt

3. Gewinne bei unterschiedlichen Substitutionsbeziehungen

Um die Wirkungen unterschiedlicher Substitutionsbeziehungen zwischen den Gütern der Firmen i und j untersuchen zu können, werden die Ableitungen der Gewinn bei unterschiedlichen Preistechniken bzw. der Gewinndifferenzen zwischen zwei Preistechniken nach β betrachtet (s. Anhang 3).

Satz 2: Je stärker die Substitutionsbeziehungen zwischen den gehandelten Gütern sind (je größer β ist), um so höher ist der Gewinn bei allen Preistechniken und um so kleiner sind die Gewinndifferenzen zwischen Π_{dd} , Π_{df} , Π_{fd} und Π_{ff} . Je stärker die Substitutionsbeziehung zwischen den Gütern sind (je größer β ist), um so mehr nähern sich die Gewinne Π_{df} und Π_{fd} in der Mitte zwischen Π_{dd} und Π_{ff} aneinander an.

Das letzte Ergebnis kann auch mit Hilfe ausgewählter numerischer Werte für β in den Gewinngleichungen verdeutlicht werden. Der Wert $\beta = 0,9$ steht für eine starke Substitution, $\beta = 0,5$ für eine mittlere und $\beta = 0,1$ für eine schwache Substitution. Die Gewinne lauten nunmehr:

Tabelle 1
Gewinne bei alternativen Preistechniken und Substitutionsbeziehungen

$\beta = 0,9$			
Strategien	fob pricing j	dis. pricing j	
.	$2R(0,75R^2 - 30R + 300)/363$	$2R(0,869R^2 - 30R + 300)/363$	
fob pricing i	$2R(0,75R^2 - 30R + 300)/363$	$2R(0,859R^2 - 30R + 300)/363$	
.	$2R(0,859R^2 - 30R + 300)/363$	$2R(R^2 - 30R + 300)/363$	
dis. pricing i	$2R(0,869R^2 - 30R + 300)/363$	$2R(R^2 - 30R + 300)/363$	
$\beta = 0,5$			
Strategien	fob pricing j	dis. pricing j	
.	$R(1,5R^2 - 12R + 24)/27$	$R(1,82R^2 - 12R + 24)/27$	
fob pricing i	$R(1,5R^2 - 12R + 24)/27$	$R(1,67R^2 - 12R + 24)/27$	
.	$R(1,67R^2 - 12R + 24)/27$	$R(2R^2 - 12R + 24)/27$	
dis. pricing i	$R(1,82R^2 - 12R + 24)/27$	$R(2R^2 - 12R + 24)/27$	
$\beta = 0,1$			
Strategien	fob pricing j	dis. pricing j	
.	$2R(20,25R^2 - 90R + 100)/361$	$2R(26,37R^2 - 90R + 100)/361$	
fob pricing i	$2R(20,25R^2 - 90R + 100)/361$	$2R(20,88R^2 - 90R + 100)/361$	
.	$2R(20,88R^2 - 90R + 100)/361$	$2R(27R^2 - 90R + 100)/361$	
dis. pricing i	$2R(26,37R^2 - 90R + 100)/361$	$2R(27R^2 - 90R + 100)/361$	

Mit anderen Worten kann der in Satz 2 ausgedrückte Sachverhalt auch wie folgt formuliert werden: Je stärker die Substitutionsbeziehungen zwischen den beiden von den Konkurrenten angebotenen Gütern sind, um so geringer ist auch der Gewinnzuwachs jenes Anbieters, der die suboptimale Preistechnik (fob pricing) anwendet und dessen Gewinn durch den Übergang von fob pricing zu discriminatory pricing im Gleichgewicht angehoben wird.

4. Annahmen und Modell 2

Im zweiten Modell wird die Ausdehnung der firmenindividuellen Marktgebiete in einem linearen, eindimensionalen Gesamtmarktgebiet durch die Gleichheit der Ortspreise benachbarter Firmen bestimmt. Die Annahmen für die Grundstruktur des Modells lauten:

- A1: Zwei Anbieter (i und j) sind an den jeweiligen Enden eines linearen Marktgebietes mit der Ausdehnung $L \in (0, 1]$ angesiedelt. Der Standort der Firma i möge bei $y_i = 0$ und der Standort der Firma j bei $y_j = L$ angesiedelt sein.
- A2: Die Nachfrage je Konsument $q(r)$ nach dem gehandelten Gut am Ort r lautet

$$(11) \quad q(r) = 1 - m - fr, \text{ mit } r \in L$$

wobei m den Ab-Werk-Preis, f die Transportkosten je Mengen- und Entfernungseinheit sowie r die Entfernung zwischen Anbieter- und Nachfragerstandort bzw. zwischen Produktions- und Konsumtionsstandort darstellen. Die Transportkosten werden auf 1 standardisiert.

- A3: Innerhalb des Marktgebietes sind die Nachfrager kontinuierlich und gleichmäßig verteilt; die Nachfragerdichte je Entfernungseinheit beträgt genau 1.
- A4: Aus Gründen der Vereinfachung des Modells soll angenommen werden, daß die Kosten der Produktion Null betragen.
- A5: Die Firmen i und j verfolgen das Ziel der Gewinnmaximierung. Zwei unterschiedliche Verhaltensweisen gegenüber dem jeweiligen Konkurrenten sollen diskutiert werden: (1) Die Anbieter wenden einen Preiswettbewerb (Bertrand-Wettbewerb) an oder (2) sie verhalten sich wie regionale Monopolisten (Monopolistische Konkurrenz).
- A6: Den Firmen stehen prinzipiell zwei räumliche Preistechniken zur Verfügung. Sie können (1) eine nicht-diskriminierende Preistechnik (fob pricing) oder (2) eine diskriminierende Preissetzung ('discriminatory pricing) anwenden.

6*

A7: Die Konsumenten kaufen bei der Firma, die das gehandelte Gut am Konsumtionsstandort zum niedrigsten Ortspreis $p(r)$ anbietet
 $p(r) = \min\{p_i(r), p_j(r)\}$ für alle $r \in L$.

Aus A7 folgt unmittelbar, daß sich die Marktgebiete nicht überlappen und daß die Grenze zwischen den Marktgebieten durch die Gleichheit der Ortspreise $p_i = p_j$ genau bestimmbar ist. Bei Ab-Werk-Preisen der Firmen i und j von m_i und m_j und bei einer Entfernung von L zwischen den Firmenstandorten gilt an der Wettbewerbsgrenze R für die gleichen Ortspreise

$$p_i(R) = p_j(R)$$

oder

$$(12) \quad m_i + R = m_j + L - R$$

und für die Ausdehnung des Marktgebietes der Firma i

$$(13) \quad R(m_i, m_j) = (m_j - m_i + L)/2$$

sowie für die Ausdehnung des Marktgebietes der Firma j

$$L - R(m_i, m_j) = (m_i - m_j + L)/2$$

Der Gewinn an einem beliebigen Ort im Marktgebiet lautet:

$$(14) \quad \pi_i(r) = m_i(1 - m_i - r) \text{ bzw. } \pi_j(r) = m_j(1 - m_j - r)$$

und der gesamte Gewinn

$$\Pi_i = \int_0^{R(m_i, m_j)} \pi_i(r) dr$$

und

$$(15) \quad \Pi_j = \int_0^{L - R(m_i, m_j)} \pi_j(r) dr$$

Da annahmegemäß die Produktionskosten Null betragen, werden lediglich die Transportkosten als Preisuntergrenze berücksichtigt.

Ebenso wie bei Modell 1 können wiederum vier unterschiedliche Strategiekombinationen aus räumlichen Preistechniken betrachtet werden:

Fall 1: Beide Anbieter wenden fob pricing an.

Fall 2: Beide Anbieter wenden discriminatory pricing an.

Fall 3: Anbieter i wendet discriminatory pricing, Anbieter j fob pricing an.

Fall 4: Anbieter i wendet fob pricing, Anbieter j discriminatory pricing an.

Da sich die beiden Firmen i und j weder durch unterschiedliche Kosten noch durch unterschiedliche Nachfragesituationen unterscheiden, ist es nicht notwendig, den Fall 4 gesondert zu diskutieren, da lediglich die Indizes ausgetauscht werden.

5. Bertrand-Wettbewerb

Zunächst wird von der Annahme eines Preiswettbewerbs ausgegangen. Grundsätzlich versucht jedes Unternehmen den Ortspreis seines Konkurrenten zu unterbieten, um somit sein firmenindividuelles Marktgebiet zu vergrößern und den Konkurrenten vom Markt zu verdrängen. Die Kosten – unter den getroffenen Modellannahmen lediglich die Transportkosten – stellen dabei die Preisuntergrenze dar.

Fall 1: Beide Firmen, die fob pricing anwenden, reduzieren im Zuge des Preiswettbewerbs ihre Ab-Werk-Preise, bis gilt $m_i = m_j = 0$ bzw. bis ihre Ortspreise gleich den Transportkosten sind $p_i^*(r) = p_j^*(r) = r$. Gleichung (3) reduziert sich zu $R = L/2$ und der Gewinn beträgt

$$(16) \quad \Pi_i = \Pi_j = 0$$

Dieses Ergebnis und der Verlauf der Ortspreislinien sind in Abbildung 1 festgehalten.

Fall 2: Beide Firmen verlangen bei discriminatory pricing an jedem Ort im Marktgebiet einen Lieferpreis p , der hinsichtlich dieses Ortes gewinnmaximal ist. Die gewinnmaximalen Ortspreise aus den Gleichungen (4) lauten $p_i^*(r) = p_j^*(r) = (1+r)/2$ und die entsprechenden Ab-Werk-Preise $m_i^*(r) = m_j^*(r) = (1-r)/2$. Die Nichtnegativitätsbedingung für die Nachfrage begrenzt die maximale Standortentfernung auf $L=2$. Es lassen sich für jeden Anbieter drei grundsätzlich unterschiedliche Bereiche entlang der Ortspreislinie feststellen (vgl. *Thisse/Vives [1988]*, *Anderson/Thisse [1988]*): (1) Zwischen dem Standort und dem Schnittpunkt der Ortspreislinie $p_i^*(r) = (1/2) + (1/2)r$ mit der Kostenkurve (Transportkostenlinie) des Konkurrenten $L-r$ verfügt die Firma i über einen monopolistischen Bereich; ein Konkurrent kann die gewinnmaximalen Ortspreise der anderen Firma nur unterbieten, und somit sein Absatzgebiet vergrößern, wenn er an diesen Orten Verluste in Kauf nimmt. (2) Zwischen diesem Schnittpunkt $R=(2L-1)/3$ und $L/2$ existiert ein Wettbewerbsbereich, in dem die

Firma i von ihren gewinnmaximalen Ortspreisen nach unten abweichen und die Ortspreise des Konkurrenten unterschreiten kann, ohne Verluste zu erleiden. Die Untergrenze für den Ortspreis wird durch die Transportkosten des Konkurrenten gebildet, da dieser bei Unterschreiten des Ortspreises Verluste erleiden würde $p_i^*(r) = L - r$. Es soll daher angenommen werden, daß dieser Teil des Marktgebiets von der Firma beliefert wird, die die geringeren Transportkosten in diesem Gebiet aufwenden muß. (3) Zwischen $L/2$ und L beträgt der niedrigste Ortspreis der Firma i , der keine Verluste entstehen läßt, genau $p_i^*(r) = r$. Diese drei Bereiche gelten ebenso für den Konkurrenten j . Der Gewinn der Firma i lautet bei discriminatory pricing und Preiswettbewerb unter Berücksichtigung der gewinnmaximalen Ortspreise

$$\begin{aligned} p_i^*(r) &= (1 + r)/2 \quad \text{für } 0 \leq r < (2L - 1)/3 \\ p_i^*(r) &= (L - r) \quad \text{für } (2L - 1)/3 \leq r \leq L/2 \\ p_i^*(r) &= r \quad \text{für } L/2 < r \leq L \end{aligned}$$

und der Gewinn der Firma j unter Berücksichtigung der entsprechenden Ortspreise $p_j^*(r)$

$$(17) \quad \Pi_i^* = \Pi_j^* = \int_0^{(2L-1)/3} ((1 - r)^2 / 4) dr + \int_{(2L-1)/3}^{L/2} (L - 2r)(1 - L + r) dr$$

oder

$$\Pi_i^* = \Pi_j^* = (L^3 - 6L^2 + 12L - 2)/72$$

Der Gewinn ist für alle $L > 2 - \sqrt[3]{6} \approx 0,18288$ positiv. Der Verlauf der Ortspreislinsien ist in Abbildung 2 wiedergegeben.

Fall 3: Wir wollen nun annehmen, daß Firma i discriminatory pricing und Firma j fob pricing anwenden. Wie in Fall 1 läßt der Preiswettbewerb den Ortspreis des Anbieters j auf $p_j^*(r) = r$ fallen. Unter Berücksichtigung der in Fall 2 diskutierten drei Bereiche entstehen die folgenden Gewinne:

$$(18) \quad \Pi_i^* = (L^3 - 6L^2 + 12L - 2)/72, \quad \Pi_j^* = 0$$

Die entsprechenden Ortspreislinsien sind in Abbildung 3 dargestellt.

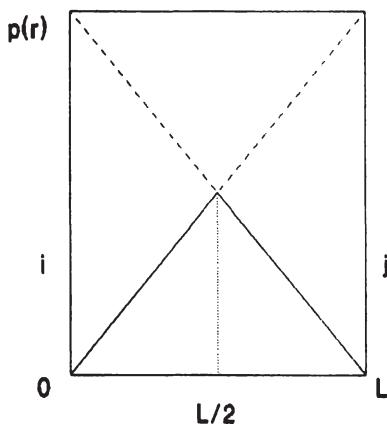


Abb. 1: fob pricing

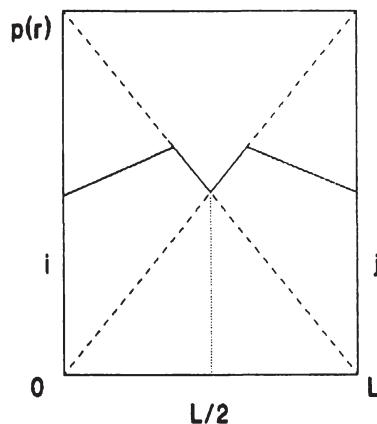


Abb. 2: discriminatory pricing

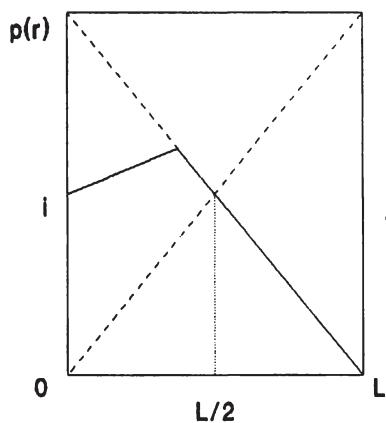


Abb. 3: discriminatory/fob pricing

Die drei diskutierten Fälle zeigen eindeutig, daß discriminatory pricing die überlegene Preistechnik darstellt.

Satz 3: Wenden beide Firmen die Strategie discriminatory pricing an, so stellt sich ein Nash-Gleichgewicht ein; keine der beiden Firmen sieht sich – bei Abwesenheit von Restriktionen für die Anwendung von Preistechniken – veranlaßt, eine andere Preistechnik zu wählen. Geht eine Firma von fob pricing zu discriminatory pricing über, so bleibt der Gewinn des Konkurrenten davon unberührt.

6. Monopolistische Konkurrenz

Es wird nunmehr von der Annahme ausgegangen, daß die beiden Anbieter sich wie räumliche Monopolisten verhalten. Daher verzichten sie auf eine Unterschreitung der Konkurrenzpreise und setzen ihre Ortspreise bzw. Ab-Werk-Preise ohne Rücksicht auf die Preistechnik und -höhe des Konkurrenten. Ein derartiges Verhalten kann verschiedene Gründe haben. Zum einen ist es denkbar, daß die Firmen erkannt haben, daß Preiswettbewerb zu geringeren Gewinnen führt. Zum anderen ist ein solches Verhalten in traditionell geprägten Märkten anzutreffen, in denen ein Preiskampf – nicht jedoch die Wahl alternativer Preistechniken – von der öffentlichen Meinung sanktioniert wird. Auch unter dieser Annahme können die oben beschriebenen Fälle unterschieden werden.

Fall 1: Da beide Firmen, die die Preistechnik fob pricing anwenden, hinsichtlich Nachfrage- und Kostenstruktur sowie hinsichtlich der Preistechniken völlig identisch sind, sind auch die Gleichgewichtspreise beider Firmen identisch $m_i^* = m_j^* = (4 - L)/8$. Damit reduziert sich Gleichung (13) zu $R = L/2$; die Nichtnegativitätsbedingung für die nachgefragte Gütermenge reduziert die Standortentfernung auf $L \leq 4/3$. Die Gewinne beider Firmen lauten:

$$(19) \quad \Pi_i^* = \Pi_j^* = (L(L - 4)^2)/128$$

Dieses Ergebnis zeigt, daß die Höhe des Gewinns von der Entfernung der Standorte bestimmt wird, wenn man gleichartige Preisreaktionen und Preistechniken der Anbieter unterstellt. Die Ortspreislinien sind in Abbildung 4 verdeutlicht.

Fall 2: Beide Firmen verlangen bei discriminatory pricing an jedem Ort im Marktgebiet einen Lieferpreis p^* , der hinsichtlich dieses Ortes gewinnmaximal ist. Die Nichtnegativitätsbedingung für die Nachfrage begrenzt die maximale Standortentfernung auf $L = 2$. Da wiederum angenommen wird, daß die Nachfrage- und Kostenstruktur sowie die Preistechniken beider Firmen gleich sind, reduziert sich Gleichung (13) zu $R = L/2$. Die Gleichgewichtsgewinne beider Unternehmen betragen somit

$$(20) \quad \Pi_i^* = \Pi_j^* = \int_0^{L/2} (1 - r)^2 / 4 \ dr = (L(L^2 - 6L + 12)) / 96$$

Vergleicht man die Gewinne (19) und (20), so kann für alle $L \in [0, 1]$ ein um $L^3/384$ höherer Gewinn bei discriminatory pricing festgestellt werden.

Fall 3: Wir wollen nun annehmen, daß Firma i discriminatory pricing und Firma j fob pricing anwenden. Unter Berücksichtigung des optimalen Preises m_i^* aus Gleichung (14) in (13) erhält man als obere Integrationsgrenze für das Marktgebiet der Firma j

$$(21) \quad L - R = L - (2L + 2m_j - 1)/3$$

und die Gewinnfunktion

$$(21) \quad \Pi_j = \int_0^{(L-2m_j+1)/3} m_j (1 - m_j - r) dr = \\ - (m_j (L - 2m_j + 1) (L - 4m_j - 5)) / 18$$

Der gewinnmaximale Preis des fob pricing anwendenden Unternehmens j beträgt:

$$(23) \quad m_j^* = - ((7L^2 - 10L + 19)^{0.5} - L - 7) / 12$$

Setzt man den Preis m_j^* in die Gewinnfunktion ein, so erhält man

$$(24) \quad \Pi_j^* = - (((7L^2 - 10L + 19)^{0.5} - 4L + 8) ((7L^2 - 10L + 19)^{0.5} - L - 7) \\ ((7L^2 - 10L + 19)^{0.5} + 5L - 1)) / 3888$$

Betrachten wir nunmehr das konkurrierende Unternehmen i , das unabhängig von Firma j discriminatory pricing verfolgt. Aus Fall 2 sind die optimalen Ab-Werk-Preise m_i^* für einen Ort r im Marktgebiet, und damit auch für $L - R$, bekannt. Da nunmehr auch m_j^* bekannt ist (Gleichung (23)), kann auch die Ausdehnung des Marktgebietes der Firma i durch

$$(25) \quad R = - ((7L^2 - 10L + 19)^{0.5} - 13L - 1) / 18$$

angegeben werden. Für $L < 0,240253$ wird R negativ, m.a.W. bei Standortentfernungen, die geringer als 0,240253 sind, fällt das gesamte Marktgebiet an den Konkurrenten j , der die Preistechnik fob pricing verfolgt. Der Gewinn der Firma i lautet:

$$(26) \quad \Pi_i^* = \int_0^{-((7L^2 - 10L + 19)^{0.5} - 13L - 1) / 18} (1 - r)^2 / 4 dr$$

oder

$$\Pi_i^* = -((7L^2 - 10L + 19)^{0.5} (257L^2 - 668L + 443) \\ - 1235L^3 + 4683L^2 - 6261L + 25)/34992$$

Der Verlauf der Ortspreislinien ist in Abbildung 6 dargestellt.

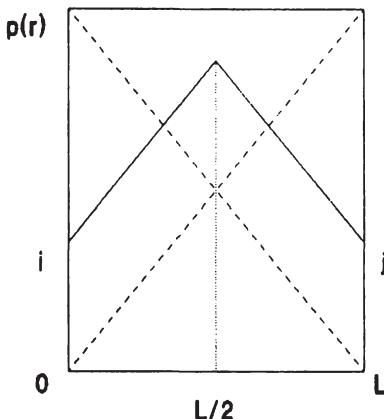


Abb. 4: fob pricing

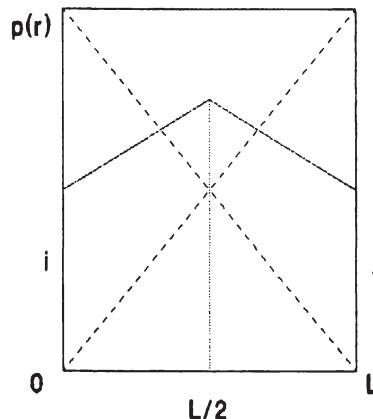
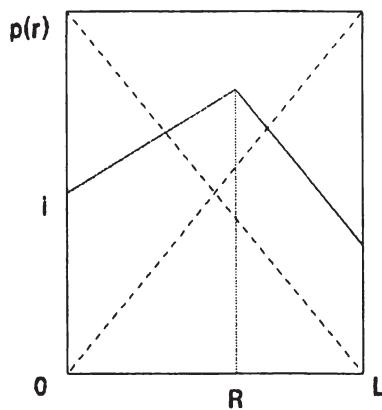


Abb. 5: discriminatory pricing



$$R = -(\sqrt{7LL - 10L + 19} - 13L - 1)/18$$

Abb. 6: discriminatory/fob pricing

7. Vergleich der Ergebnisse

Für den Vergleich der Ergebnisse wird ein Index zur Bezeichnung der Gewinne eingeführt: Π_{dd} , Π_{df} , Π_{fd} und Π_{ff} , wobei der erste Index die Preisstrategie der betrachteten Firma und der zweite Index die Preisstrategie des Konkurrenten verdeutlichen sowie f für „fob pricing“ und d für „discriminatory pricing“ stehen. Zeichnet man die vier Gewinngleichungen (19), (20), (24) und (26) aus Abschnitt 6 für den Bereich $0 < L < 1$ in Abbildung 7 ein, so zeigt sich, daß die Höhe der Gewinne von zwei Einflußgrößen bestimmt wird: (1) Zum einen ist der Gewinn der Firma i nicht nur von der gewählten eigenen Preistechnik, sondern auch von der angewendeten Preistechnik des konkurrierenden Unternehmens j abhängig. (2) Zum anderen wird die Höhe des Gewinns von der Entfernung der Standorte beider Anbieter i und j bestimmt und die gewinnoptimale Preistechnik – unter Berücksichtigung der Preistechnik des Konkurrenten – ebenfalls von der Entfernung der Standorte determiniert. Eine genauere Betrachtung der Abbildung 7 zeigt sechs Schnittpunkte der Gewinnfunktionen.

$$\begin{aligned}\Pi_{dd} = \Pi_{ff} &\quad L = 0 \\ \Pi_{dd} = \Pi_{fd} &\quad L = 0,545697 \\ \Pi_{ff} = \Pi_{fd} &\quad L = 0,562834 \\ \Pi_{fd} = \Pi_{df} &\quad L = 0,642747 \\ \Pi_{ff} = \Pi_{df} &\quad L = 0,695455 \\ \Pi_{dd} = \Pi_{df} &\quad L = 0,732053\end{aligned}$$

Für die nachstehenden Standortentfernungen erhält man folgende Rangfolgen der Gewinne:

$$\begin{array}{ll}L < 0,5457 & \Pi_{fd} > \Pi_{dd} > \Pi_{ff} > \Pi_{df} \\0,5457 < L < 0,5628 & \Pi_{dd} > \Pi_{fd} > \Pi_{ff} > \Pi_{df} \\0,5628 < L < 0,6427 & \Pi_{dd} > \Pi_{ff} > \Pi_{fd} > \Pi_{df} \\0,6427 < L < 0,6955 & \Pi_{dd} > \Pi_{ff} > \Pi_{df} > \Pi_{fd} \\0,6955 < L < 0,7321 & \Pi_{dd} > \Pi_{df} > \Pi_{ff} > \Pi_{fd} \\L > 0,7321 & \Pi_{df} > \Pi_{dd} > \Pi_{ff} > \Pi_{fd}\end{array}$$

Für alle Entfernungen zwischen den beiden Standorten zeigt sich bei gleichartigen Preistechniken der Firmen – wie auch schon aus der Literatur bekannt ist – ein höherer Gewinn bei discriminatory pricing und ein geringerer Gewinn bei fob pricing. Bei unterschiedlichen Preistechniken ist der Gewinn im Bereich $0 < L < 0,6424$ für fob pricing und im Bereich $0,6424 < L < 1$ für discriminatory pricing höher. Vernachlässigt man den Bereich $0,5456 < L < 0,7321$, so kann folgende Aussage formuliert werden:

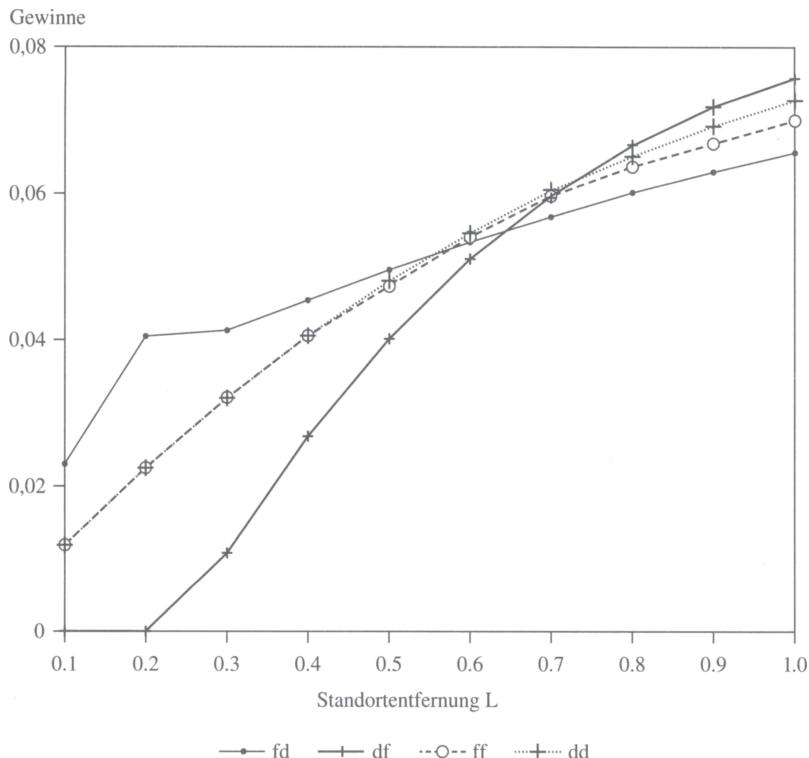


Abb. 7: Gewinne bei unterschiedlichen Preistechniken, Standortentfernungen und monopolistischer Konkurrenz

Satz 4: Wenden beide Anbieter bei monopolistischer Konkurrenz und geringen Standortentfernungen ($L < 0,5456$) diskriminierende Preistechniken an, so erhöht sich der Gewinn der Firma, die zu fob pricing übergeht, und der Gewinn der Firma, die bei der diskriminierenden Preistechnik verbleibt, wird reduziert. Wenden beide Anbieter bei großen Standortentfernungen ($L > 0,7321$) fob pricing an, so erhöht sich der Gewinn der Firma, die zur diskriminierenden Preistechnik übergeht, und der Gewinn der Firma, die bei fob pricing verbleibt, wird verringert.

Bei sehr kleinen Standortentfernungen ($0 < L < 0,2402$) wird der diskriminierende Anbieter von der unter fob pricing anbietenden Firma vom Markt verdrängt. In einem mittleren Bereich ($0,5457 < L < 0,7321$) führt die Anwendung der Preisdiskriminierung durch beide Anbieter zum höchsten Gewinn.

Zum Vergleich der Gewinne aus Abschnitt 6 mit denen aus Abschnitt 5 sollen neben den Kurzbezeichnungen Π_{dd} , Π_{df} , Π_{fd} und Π_{ff} die Bezeichnungen Π_{dd}^c , Π_{df}^c und Π_{fd}^c für die Gewinne aus dem Bertrand-Wettbewerb eingeführt werden. Zusammengefaßt erhält man folgende Ergebnisse:

$$\begin{aligned}\Pi_{dd}^c &= \Pi_{df}^c < \Pi_{fd}^c = \Pi_{ff}^c \quad \forall L \in [0, 0,18288) \\ \Pi_{dd}^c &= \Pi_{df}^c > \Pi_{fd}^c = \Pi_{ff}^c \quad \forall L \in (0,18288, 1]\end{aligned}$$

$$\begin{aligned}\Pi_{dd} &> \Pi_{dd}^c \quad \forall L \in [0,1] \\ \Pi_{dd} - \Pi_{dd}^c &= -(L^3 - 6L^2 + 12L - 8)/288\end{aligned}$$

$$\Pi_{ff} > \Pi_{ff}^c \quad \forall L \in [0,1]$$

$$\begin{aligned}\Pi_{df} &< \Pi_{df}^c \quad \forall L \in [0, 0,39061) \\ \Pi_{df} &> \Pi_{df}^c \quad \forall L \in (0,39061, 1)\end{aligned}$$

$$\begin{aligned}\Pi_{df} - \Pi_{df}^c &= -((7L^2 - 10L + 19)^{0.5} (257L^2 - 688L + 443) \\ &\quad - 749L^3 + 1767L^2 - 429L - 947)/34992\end{aligned}$$

$$\Pi_{fd} > \Pi_{fd}^c \quad \forall L \in [0,1]$$

Jenes Unternehmen, das discriminatory pricing und Preiswettbewerb anwendet, kann unabhängig von der Marktausdehnung einen höheren Gewinn erzielen als das Unternehmen, das bei fob pricing verbleibt $\Pi_{df}^c > \Pi_{fd}^c$. Da die unterlegene Firma die erfolgreiche Preistechnik des Konkurrenten imitiert, ist dieser Zustand aber nur kurzfristig denkbar. Geht das Unternehmen, das bisher seine Güter zu fob pricing anbot, auch zu discriminatory pricing über, so kann es seinen Gewinn bei unverändertem Gewinn des Konkurrenten steigern. Ferner ist hervorzuheben, daß bei fob pricing des Konkurrenten und kleinen Marktgebieten bis zu $L = 0,39061$, ein Unternehmen, das discriminatory pricing anwendet, seinen Gewinn durch Preiswettbewerb maximieren kann. Jenseits von $L = 0,39061$ führt jedoch der Verzicht auf strategische Preissetzungen zum maximalen Gewinn. Damit hängt die Anwendung des Preiswettbewerbs von der Ausdehnung des Marktgebietes und bei gegebenem Marktgebiet von der Anzahl der Anbieter ab. Allgemein kann gesagt werden, daß mit Ausnahme des df-Falls die Gewinne bei Preiswettbewerb für alle $L \in [0,1]$ niedriger sind als bei einem „monopolistischen“ Verhalten der Anbieter. Diese

Ergebnisse sind nicht überraschend und entsprechen den Resultaten in vergleichbaren Punktmarktmodellen.

8. Schlußfolgerungen

Unter den Annahmen eines gemeinsamen Standortes beider Konkurrenten, gemeinsamer vollständig überlappender Marktgebiete und der völligen Symmetrie hinsichtlich Nachfragestruktur und Verhaltensweisen (Modell 1) zeigt sich discriminatory pricing als die dominante Strategie; unabhängig davon, ob der Konkurrent beabsichtigt, fob pricing oder discriminatory pricing zu wählen, kann das betrachtete Unternehmen seinen jeweils höchsten Gewinn realisieren, wenn es seinerseits immer discriminatory pricing anwendet. Bietet auch der Konkurrent sein Gut unter discriminatory pricing an, so stellt sich ein Nash-Gleichgewicht ein.

Ein für das strategische Verhalten wichtiges Resultat zeigt sich, wenn man von der Tatsache ausgeht, daß beide Firmen ihre Güter zunächst unter fob pricing anbieten. Geht nun ein Unternehmen – sagen wir Unternehmen *i* – zu discriminatory pricing über, so ist es in der Lage, seinen Gewinn zu steigern. Im Gleichgewicht erhöht sich aber auch – wenn auch nicht im gleichen Ausmaß wie bei Unternehmen *i* – der Gewinn des Konkurrenten *j*. Im umgekehrten Fall – beide Anbieter wenden discriminatory pricing an – kann die Firma *i* zu fob pricing übergehen, um somit den Gleichgewichtsgewinn des Konkurrenten *j* zu verringern, allerdings um den Preis, selbst einen noch geringeren Gewinn zu erzielen. Dieses strategische Verhalten ist immerhin denkbar, wenn man von unterschiedlichen finanziellen Ausstattungen der Firmen ausgeht und von der Absicht der finanziell stärkeren Firma, angenommen *i*, ihren Konkurrenten *j* zum Marktaustritt zu veranlassen.

Die Spielräume für strategisches Verhalten werden entscheidend durch die Substitutionsbeziehungen zwischen den gehandelten Gütern bestimmt. Je stärker diese Beziehung ist, um so näher liegt der Gewinn der Firma *i*, die bei geplantem fob pricing des Konkurrenten *j* discriminatory pricing verfolgt, bei jenem Gewinn, der bei Preisdiskriminierung beider Anbieter entsteht, und um so geringer ist der dadurch verursachte Gewinnzuwachs bei dem Konkurrenten *j*. Je schwächer die Substitutionsbeziehung ist, um so stärker wird bei dem gleichen Vorgang der Gewinn des Konkurrenten *j* abgehoben. Allerdings ist auch in diesem Fall der Gewinn des Anbieters *i* höher als der des Anbieters *j*. Ungeachtet dessen kann der Schluß gezogen werden, daß mit abnehmender Substitutionsbeziehung zwischen den Gütern auch der Anreiz steigt, das gleiche Ergebnis durch Imitation der Preistechnik des erfolgreicheren Anbieters herbeizuführen.

Unter den Annahmen verschiedener Standorte, nichtüberlappender Marktgebiete und vollständiger Symmetrie hinsichtlich Nachfragestruktur und Verhaltensweisen (Modell 2) zeigt sich folgendes Bild: Der Vergleich der Resultate bei Bertrand-Wettbewerb verdeutlicht, daß discriminatory pricing – unabhängig von der durch den Konkurrenten gewählten Preistechnik – zu den höchsten Gewinnen führt. Folglich werden beide Firmen bei Abwesenheit von Restriktionen hinsichtlich Preistechnik und Verhaltensweise diese Preissetzung wählen. Bei monopolistischer Konkurrenz zeigt sich ein weniger einheitliches Bild. In Abhängigkeit von den Standortentfernungen lassen sich vier Bereiche für unterschiedliches strategisches Verhalten feststellen. (1) Im ersten Bereich ($0 < L < 0,2402$) würde zwar der Übergang eines Anbieters – sagen wir i – von der gemeinsamen Strategie der Preisdiskriminierung zu fob pricing den Konkurrenten j vom Markt verdrängen, jedoch ist kaum anzunehmen, daß die Firma j nicht auch zu fob pricing übergehen wird, um weiterhin am Markt zu verbleiben. Danach befinden sich beide Anbieter auf einem – wenn auch geringfügig – niedrigeren Gewinniveau. (2) Für den zweiten Bereich ($0,2402 < L < 0,5457$) gilt genau die gleiche Analyse der Situation wie in Bereich 1, jedoch mit der Ausnahme, daß der Konkurrent j nicht vom Markt verdrängt wird. (3) Im dritten Bereich ($0,5457 < L < 0,7321$) ist bei gleichartigen Preistechniken beider Firmen ein Wechsel eines Anbieters zu einer alternativen Preistechnik unwahrscheinlich, da er seine Gewinnsituation in jedem Fall verschlechtern würde. (4) Schließlich kann für den vierten Bereich ($0,7321 < L < 1,3333$) der folgende Prozeß für die Variation der Preistechniken angenommen werden: In der Ausgangssituation sollen beide Unternehmen das Gut unter fob pricing anbieten, und somit auch gleiche Gewinne erzielen. Wendet sich beispielsweise Firma i nun der Preisdiskriminierung zu, so kann sie ihren Gewinn steigern und gleichzeitig den Gewinn der konkurrierenden und bei fob pricing verbleibenden Firma j senken. Unternehmen j wird nunmehr die erfolgreiche Preistechnik ihres Konkurrenten imitieren und ebenfalls zur Preisdiskriminierung übergehen. Durch dieses Verhalten der Firma j steigt ihr Gewinn in einem erheblichen Ausmaß und der Gewinn der Firma i wird in einem geringen Umfang gesenkt. Die Gewinne beider Anbieter sind nunmehr gleich und liegen über den Gewinnen in der Ausgangssituation.

Wenn die optimale Wahl einer Preistechnik – sei es auch nur, um temporäre Gewinnvorteile zu erlangen – bei monopolistischer Konkurrenz von der Standortentfernung abhängt, dann bedeutet dies aber auch, daß sich durch Marktzutritte die Verhaltensweisen ändern können. Marktzutritte verringern die Standortentfernungen und verändern damit die als optimal anzusehende Wahl der Preistechnik. Anders gesagt: Im kurzfristigen Gleichgewicht können durchaus andere Strategien angewendet werden als im langfristigen Gleichgewicht.

Literatur

- Anderson, S. P./Thisse, J. F.*, (1988), Price Discrimination in Spatial Competitive Markets, in: European Economic Review, Bd. 32, S. 578 - 590.
- Beckmann, M.* (1976), Spatial Price Policies Revisited, in: Bell Journal of Economics, Bd. 7, S. 619 - 630.
- Claycombe, R. J.* (1991), Spatial Price Discrimination and Mill Pricing: Hypermobile Multiplant Monopoly, in: Journal of Industrial Economics, Bd. 39, S. 723 - 728.
- Greenhut, J. G.* (1977), On the Economic Advantages of Spatially Discriminatory Prices Compared with F.O.B. Prices, in: Southern Economic Journal, Bd. 44, S. 161 - 165.
- Greenhut, M. L./Norman, G./Hung, C.*, (1987), The Economics of Imperfect Competition: A Spatial Approach, Cambridge usw.
- Ohta, H.* (1988), Spatial Price Theory of Imperfect Competition, College Station.
- Schöler, K.* (1988), Räumliche Preistheorie, Berlin.
- (1990), Heterogene Güter und alternative Preistechniken in einem räumlichen Oligopol, in: Jahrbücher für Nationalökonomie und Statistik, Bd. 207, S. 143 - 156.
- (1992), Zum strategischen Einsatz räumlicher Preistechniken, Volkswirtschaftliche Diskussionsbeiträge, Nr. 31, Siegen.
- Thisse, J. F./Vives, X.* (1988), On the Strategic Choice of Spatial Price Policy, in: American Economic Review, Bd. 78, S. 122 - 137.

Anhang

Anhang 1

Fall 1: Da beide Anbieter fob pricing anwenden, sind die beiden Gleichgewichtspreise aus dem Gesamtgewinn der Firmen (5) unter Verwendung von (4) abzuleiten. Man erhält man (unter der Bedingung 2. Ordnung $-4R < 0$)

$$(A1) \quad m_i = (\beta(2m_j + R) - R + 2)/4, \quad m_j = (\beta(2m_i + R) - R + 2)/4$$

und die Gleichgewichtspreise

$$(A2) \quad m_i^* = m_j^* = (\beta R - R + 2)/(2(2 - \beta))$$

die in die Gewinngleichung (5) eingesetzt werden. Daraus ergibt sich Gleichung (6)

Fall 2: Da beide Anbieter die Strategie anwenden, an jedem Ort im Marktgebiet einen anderen Lieferpreis (Ortspreis) von den Käufern zu verlangen, der auf den Ort r bezogen gewinnmaximal ist, erhält man aus Gleichung (4) (unter der Bedingung 2. Ordnung $-2 < 0$) und unter Hinzunahme der Transportkosten je Mengen- und Entfernungseinheit die Ortspreise von .

$$(A3) \quad p_i = (\beta(m_j + r) + 1 + r)/2, \quad p_j = (\beta(m_i + r) + 1 + r)/2$$

Die Gleichgewichtspreise betragen, ausgedrückt als Ab-Werk-Preise:

$$(A4) \quad m_i^* = m_j^* = (\beta r - r + 1)/(2 - \beta)$$

und die Gewinne im Gleichgewicht lauten unter Verwendung von (4) und (A4):

$$\Pi_i^* = 2 \int_0^R m_i^* (1 - m_i^* + \beta m_j^* - r + \beta r) dr$$

bzw.

$$(A5) \quad \Pi_j^* = 2 \int_0^R m_j^* (1 - m_j^* + \beta m_i^* - r + \beta r) dr$$

Daraus erhält man die Gewinngleichung (8).

Fall 3: Die Firma i verfolgt die Strategie der optimalen Preisdiskriminierung, und Anbieter j verwendet fob pricing. Die entsprechenden Preise aus den Gleichungen (7) und (9) lauten:

$$m_i = (\beta(m_j + r) + 1 - r)/2 \text{ und } m_j = (\beta(2m_i + R) - R + 2)/4$$

woraus sich für Firma i ein Gleichgewichtspreis von

$$(A6) \quad m_i^* = (4r(1 - \beta) - \beta^2 R - \beta(2 - R) - 4)/(2(\beta^2 - 4))$$

ergibt. Für den Anbieter j erhält man den Gleichgewichtspreis von

$$(A7) \quad m_j^* = (\beta r(1 - \beta) - \beta(R + 1) + R - 2)/(\beta^2 - 4)$$

Die Gleichgewichtspreise sind im Fall 3 nicht mehr symmetrisch und die asymmetrischen Gewinne der Firmen lauten unter Verwendung von (A6) bzw. (A7):

⁷ Todt

$$\Pi_i^* = 2 \int_0^R m_i^* (1 - m_i^* + \beta m_j^* - r + \beta r) dr$$

oder Gleichung (9) für Firma i und

$$\Pi_j^* = 2 \int_0^R m_j^* (1 - m_j^* + \beta m_i^* - r + \beta r) dr$$

oder Gleichung (10) für Firma j .

Anhang 2

Für den Vergleich der Gewinne muß gelten, daß die Rangfolge der Gewinne unter den Annahmen ($0 < R$ und $0 < \beta < 1$) die folgenden Ergebnisse zeigt:

1. Aus $\Pi_{dd} - \Pi_{df}$ folgt $(\beta R^3 (\beta^3 + 2\beta^2 - 7\beta + 4)) / (6(\beta^2 - 4)^2) > 0$.
2. Aus $\Pi_{df} - \Pi_{fd}$ folgt $(R^3 (\beta^3 - 3\beta^2 + 3\beta - 1)) / (6(\beta^2 - 4)^2) > 0$.
3. Aus $\Pi_{fd} - \Pi_{ff}$ folgt $-(\beta R^2 (\beta^4 - 3\beta^3 - \beta^2 + 7\beta - 4)) / (6(\beta^2 - 4\beta)^2) > 0$.

Alle Ergebnisse sind mit den Annahmen vereinbar, woraus sich die Rangfolge $\Pi_{dd} > \Pi_{df} > \Pi_{fd} > \Pi_{ff}$ ergibt.

Anhang 3

Die nachfolgenden Überlegungen hinsichtlich β begründen Satz 2. Die Ableitungen der Gewinnfunktionen nach β ergeben:

$$\partial \Pi_{dd} / \partial \beta = -(2R(\beta R(2R + 3) - 2(R^2 - 3)) / (3(\beta - 2)^3) > 0$$

$$\begin{aligned} \partial \Pi_{df} / \partial \beta = & -(R(3\beta^4 R(R + 2) + \beta^3(19R^2 + 36R + 12) \\ & + 6\beta^2(R^2 - 12R + 12) + 12\beta(R^2 + 4R + 12) \\ & - 8(5R^2 - 12))) / (3(\beta^2 - 4)^3) > 0 \end{aligned}$$

$$\begin{aligned} \partial \Pi_{fd} / \partial \beta = & -(R(\beta^6 R^2 - \beta^4 R(13R - 12) \\ & + 8\beta^3(8R^2 + 9R + 3) + 12\beta^2(5R^2 + 12R + 12) \\ & - 16\beta(5R^2 - 6R - 18) - 32(R^2 - 6))) / (6(\beta^2 - 4)^3) > 0 \end{aligned}$$

$$\partial \Pi_{ff} / \partial \beta = - (R(R+2)(\beta R - R + 2)) / ((\beta - 2)^3) > 0$$

Die Ableitungen der Gewinndifferenzen nach β ergeben:

Differenz $\Pi_{dd} - \Pi_{ff}$:

$$\partial (\Pi_{dd} - \Pi_{ff}) / \partial \beta = (R^3(1 - \beta)) / (3(\beta - 2)^3) < 0$$

Differenz $\Pi_{df} - \Pi_{fd}$:

$$\begin{aligned} \partial (\Pi_{df} - \Pi_{fd}) / \partial \beta &= (R^3(\beta^4 - 15\beta^2 + 26\beta - 12)) / \\ &(6(\beta^2 - 4)^2) < 0 \end{aligned}$$

Differenz $\Pi_{dd} - \Pi_{df}$:

$$\begin{aligned} \partial (\Pi_{dd} - \Pi_{df}) / \partial \beta &= -(R^3(\beta^4 + \beta^3 + 18\beta^2 - 28\beta + 8)) / \\ &(3(\beta^2 - 4)^3) < 0 \end{aligned}$$

Differenz $\Pi_{fd} - \Pi_{ff}$:

$$\begin{aligned} \partial (\Pi_{fd} - \Pi_{ff}) / \partial \beta &= -(R^3(\beta^6 - 19\beta^4 + 34\beta^3 + 24\beta^2 - 56\beta + 16)) / \\ &(6(\beta^2 - 4)^3) < 0 \end{aligned}$$

Zusammenfassung der Diskussion

Es wird die Frage gestellt, warum ein Anbieter seine Preisstrategie ändern sollte, wenn er seinen Konkurrenten auch mit einer bloßen Änderung seines Preises, bei unveränderter Strategie, massiv angreifen kann.

Die Diskussion macht auch deutlich, daß bestimmte Preisstrategien nur unter bestimmten Rahmenbedingungen zum Zuge kommen können und daher nur in bestimmten Branchen zur Anwendung gelangen. In der Praxis spielt die räumliche Preisdifferenzierung z.B. bei Handelsketten eine große Rolle. Dort besteht aber ein großer zeitlicher Abstand zwischen Planung der räumlichen Preisstrategie und ihrer tatsächlichen Realisation.

Eine Ursache für die Nichtanwendung von räumlicher Preisdifferenzierung dort wo diese theoretisch möglich wäre, kann darin gesehen werden, daß diese von den Nachfragern als äußerst unfair empfunden wird. Auch läßt sich ein Übergang auf die Fläche nicht problemlos durchführen.

Als interessant wird empfunden, wie sich die Einführung unvollständiger Information auf die Lösungen auswirken würde.

Infrastrukturversorgung als Wettbewerbsinstrument zwischen Städten

Eine theoretische und empirische Untersuchung

Von *Helmut Seitz*, Mannheim*

Abstract

Im theoretischen Teil der Arbeit wird mit Hilfe eines einfachen mehrstufigen Modells die Bedeutung der Bereitstellung öffentlicher Infrastruktur für den Wettbewerb zwischen Firmen und Regionen untersucht. Hierbei wird berücksichtigt, daß zur Finanzierung der öffentlichen Infrastrukturinvestitionen bei den Firmen Steuern erhoben werden. Der empirische Teil der Arbeit präsentiert ökonometrische Schätzungen der Produktivität der kommunalen Infrastruktur sowie eine empirische Analyse der Bestimmungsgründe für die Anpassung der Gewerbesteuerhebesätze und der kommunalen Infrastrukturkapitalstöcke. Die empirischen Analysen beruhen auf einem Städte-Paneldatensatz, der für die Jahre 1980 - 1990 für die 100 größten Städte der Bundesrepublik erstellt wurde.

1. Einleitung

In jüngster Zeit gibt es eine zunehmende Anzahl empirischer Studien, in denen der Einfluß der Bereitstellung öffentlicher Infrastrukturdienstleistungen auf die private Wirtschaft untersucht wird. *Berndt/Hansson* (1991) sowie *Conrad/Seitz* (1992) haben diesen Zusammenhang für das gesamte Verarbeitende Gewerbe untersucht; Analysen auf der Ebene der 2-stelligen Industrieklassifikation wurden von *Nadiri/Mamuneas* (1991) sowie *Seitz* (1993a) vorgelegt; regional disaggregierte Daten liegen den Schätzungen von *Costa/Ellson/Martin* (1987), *Morrison/Schwartz* (1992) sowie *Seitz/Licht* (1993) zugrunde, wobei als Regionalabgrenzung die U.S. Bundesstaaten bzw. die (west)deutschen Bundesländer herangezogen werden. Alle erwähnten Analysen vernachlässigen die Tatsache, daß der mit Abstand größte Teil der öffentlichen Investitionen von Städten und Gemeinden getä-

* Mein Dank gehört Herrn Prof. Dr. Klaus Conrad sowie den Teilnehmern auf der Sitzung des ‚Regionalwissenschaftlichen Ausschusses‘ des Verein für Socialpolitik im Oktober 1993 in Freiberg/Sachsen. Die Forschungsarbeiten wurden aus Mitteln der Emil-Kümmerling-Stiftung unterstützt.

tigt wird. So entfallen in der Bundesrepublik ca 70% der gesamten staatlichen Investitionsausgaben und ca 65% des öffentlichen Infrastrukturkapitalstocks auf die Städte und Gemeinden.¹ Somit hat der größte Teil der Infrastruktureinrichtungen einen betont räumlichen Charakter und verursacht nur in sehr begrenztem Umfang überregionale Spillover Effekte, d.h. die räumliche ‚Reichweite‘ der mit Infrastruktureinrichtungen produzierten öffentlichen Güter ist relativ gering, so daß Infrastrukturdienstleistungen als ‚local public goods‘ anzusehen sind, siehe z.B. *Oakland* (1987). Trotz dieser hohen lokalen Komponente gibt es nur eine sehr kleine Anzahl von Studien in denen die Effekte der öffentlichen Infrastruktur auf lokaler, d.h. insbesonderer kommunaler, Ebene untersucht werden. Als einige der wenigen Arbeiten seien hier die Untersuchungen von *Duffy-Deno* (1991) sowie *Eberts/Fox* (1992) genannt. Im Bereich der finanzwissenschaftlichen Theorie, insbesondere in der ‚tax competition‘-Literatur, wurden Arbeiten vorgelegt, in denen die Bereitstellung lokaler Infrastrukturdienstleistungen in Modellen des interregionalen Wettbewerbs einbezogen wird, siehe z.B. die Beiträge von *Zodrow/Mieszkowski* (1986), *Oates/Schwab* (1988) und für einen Überblick *Wildasin/Wilson* (1991).

Der vorliegende Beitrag ist der Versuch, eine enge Verbindung zwischen der jüngsten empirischen Infrastrukturforschung und den theoretischen Ansätzen aus dem Bereich der Steuerkonkurrenzliteratur herzustellen. In Abschnitt 2 stellen wir ein einfaches theoretisches Modell vor, in dem zwei wohlfahrtsmaximierende Städte, in denen jeweils nur eine Firma angesiedelt ist, mittels Infrastrukturdienstleistungen konkurrieren. In den nachfolgenden Abschnitten werden die empirischen Implikationen des Modells in einem zweistufigen Ansatz untersucht. Hierbei wird in einem ersten Schritt in Abschnitt 3 getestet, ob von der öffentlichen Infrastruktur überhaupt signifikante Produktivitätseffekte ausgehen, und in einem zweiten Schritt wird in Abschnitt 4 ein einfaches Regressionsmodell zur Erklärung kommunaler Steuer- und Investitionspolitik vorgestellt. Abschnitt 5 faßt unsere Ergebnisse zusammen, diskutiert die Schwächen unseres Ansatzes und weist auf zukünftige Forschungsfelder hin.

2. Ein theoretisches Modell der Bereitstellung kommunaler Infrastruktur bei unvollkommenem Wettbewerb

Unser Modell stellt ein sehr stark vereinfachtes Abbild des privaten Sektors innerhalb von Städten dar. Aus Vereinfachungsgründen nehmen wir an,

¹ Allerdings sei angemerkt, daß ein erheblicher Anteil der kommunalen Investitionsausgaben durch Investitionszuweisungen von Bund und Ländern finanziert wird.

daß es nur zwei Gebietskörperschaften, $i = 1, 2$, gibt, in denen jeweils nur eine Firma beheimatet ist; Firma 1 in Kommune 1 und Firma 2 in Kommune 2. Beide Firmen produzieren ein homogenes Gut, Y^i , das auf dem nationalen Markt, auf dem die Nachfragefunktion $P(Y^1 + Y^2)$ mit $P' < 0$ gilt, verkauft wird. Firma i erzielt hierbei den Erlös $R^i = P(Y^1 + Y^2)Y^i$ wobei wir wie üblich unterstellen, daß $R_i^i > 0$ und $R_{ij}^i < 0$ gilt.² Beide Firmen haben die kurzfristige Kostenfunktion $C^i(Y^i, K^i, G^i)$, wobei G^i den städtischen Infrastruktukkapitalstock und K^i den privaten Kapitalstock der Firma i in Kommune i symbolisiert. G^i geht in die private Kostenfunktion als fixer, unbezahlter Inputfaktor ein, d.h. es wird unterstellt, daß die Firmen für Infrastrukturdienstleistungen keine direkten Gebühren zu bezahlen haben.³ Zu beachten wäre, daß nicht der Bestand an Infrastruktukkapital, sondern der Dienstleistungsstrom, der mit Hilfe von G^i produziert wird, in die Produktionsfunktion der Firmen eingeht; wir unterstellen daher implizit, daß dieser Leistungsstrom proportional zum Infrastruktukkapitalstock ist. Ferner unterstellen wir, daß es sich bei G^i um ein reines öffentliches lokales Gut handelt, d.h. es gibt keine interregionalen Nutzen-Spillover. Für die Kostenfunktion unterstellen wird:

$$\begin{aligned} C_{Y_i}^i &> 0, \quad C_{Y_i, Y_i}^i > 0, \quad C_{K_i}^i < 0, \quad C_{K_i, K_i}^i > 0, \\ C_{G_i}^i &< 0, \quad C_{G_i, K_i}^i < 0, \quad C_{G_i, Y_i}^i < 0, \quad C_{G_i, G_i}^i > 0, \end{aligned}$$

wobei einige dieser Eigenschaften später in unserer empirischen Untersuchung auf deren Validität geprüft wird.

Besteuerungsaspekte werden in die Analyse einbezogen, indem wir eine sehr stark vereinfachte Struktur der kommunalen Besteuerung in der Bundesrepublik unterstellen und hierbei alle Aspekte der nationalen Besteuerung vernachlässigen, da diese alle Firmen in allen Gebietskörperschaften gleichermaßen betrifft.⁴ Die Kommunen erheben in der Bundesrepublik im wesentlichen drei Arten von Gewerbesteuern:

- die Gewerbeertragsteuer, $t^{e,i}$,
- die Gewerbekapitalsteuer, $t^{K,i}$,
- sowie die Grundsteuer B .

² Indices verdeutlichen partielle Ableitungen.

³ Für einige Infrastrukturdienstleistungen müssen Firmen direkte Gebühren bezahlen, z.B. für die Abwasserentsorgung. Die Einbeziehung solcher Zahlungen würde das Modell verkomplizieren, ohne uns wesentlich neue Erkenntnisse zu liefern.

⁴ Diese Feststellung ist nicht ganz korrekt, da die Gewerbesteuerzahlungen bei der Ermittlung der Einkommen- bzw. Körperschaftsteuerschuld berücksichtigt werden.

Unser Modell berücksichtigt nur die ersten beiden Steuern, die Grundsteuer B wird ausgeklammert. Die örtlichen Gewerbesteuersätze $t^{e,i}$ und $t^{K,i}$ sind bestimmt durch die Beziehungen:

$$t^{e,i} = t^e \cdot M^i \quad \text{und:} \quad t^{K,i} = t^K \cdot M^i,$$

wobei t^e und t^K die bundeseinheitlichen Steuermeßzahlen für die Ertrag- und Kapitalsteuer bezeichnen, die schon seit Jahren bei $t^e = 0.05$ und $t^K = 0.002$ liegen, und M^i der von den Kommunen festzulegende Gewerbesteuerhebesatz bezeichnet, der zur Zeit bei bundesdurchschnittlich ca 3.62 liegt.

Bezüglich des Optimiererverhaltens der Firmen unterstellen wir, daß diese in einem ersten Schritt eine Entscheidung über den einzusetzenden optimalen privaten Kapitalstock treffen und erst in einem zweiten Schritt über ihr Produktionsvolumen entscheiden. Unter den oben gemachten Annahmen können wir daher die kurzfristige Gewinnfunktion der Firmen schreiben als:

$$(1) \quad \pi^i = [R^i(Y^1, Y^2) - C^i(Y^i, K^i, G^i) - vK^i] \cdot (1 - t^e M^i) - t^K M^i K^i$$

wobei v den national einheitlichen Preis für private Kapitalgüter bezeichnet.

Bevor wir mit der Modellanalyse fortfahren, seien einige Anmerkungen zur unterstellten Informationsstruktur sowie der Lösungstechnik gemacht. Die Kommunen, die perfekte Informationen über die privaten Firmen und deren Optimiererverhalten haben, entscheiden vorab über das Infrastrukturangebot und legen damit gleichzeitig Gewerbesteuerhebesätze fest, die ihnen ein ausgeglichenes Budget garantieren.⁵ Die Firmen nehmen diese Politikvariablen als exogen gegeben hin und entscheiden nach Bekanntgabe von G^i und M^i zunächst über den privaten Kapitaleinsatz und im Anschluß daran über das optimale Outputvolumen. Unser Ansatz korrespondiert daher mit den Modellen von *Spencer/Brander* (1983) – im Kontext eines F&E-Wettbewerbmodells –, *Hands/Mann* (1987) – im Kontext eines Steuerwettbewerbmodells – und *Conrad* (1993) – im Kontext eines Modells strategischer Umweltpolitik. Wie üblich in einem solchen mehrstufigen spieltheoretischen Ansatz, erfolgt die Lösung „rückwärts“, d.h. zunächst bestimmen wir den optimalen Output (gegeben fester privater und öffentlicher Kapitalstöcke), bestimmen in einem nächsten Schritt den optimalen privaten Kapitaleinsatz (gegeben ein bestimmtes Infrastrukturangebot) und bestimmen in einer letzten Rechnung die wohlfahrtsoptimalen Politikparameter.

⁵ Dies impliziert, daß sich die Gebietskörperschaften gegenüber den Firmen in einer Stackelberg-Führer Rolle befinden.

Hierbei wird unterstellt, daß sowohl Firmen als auch die Regierungen ein Cournot-Verhalten aufweisen, d.h. daß Städte und Firmen nicht mit einer Reaktion ihrer Kontrahenten rechnen.

i) Outputbestimmung

Die Optimiererbedingung für den Output der Firmen erhalten wir, indem wir (1) nach Y^i differenzieren:

$$(2) \quad \pi_i^i = R_i^i - C_i^i = 0$$

Die Bedingung zweiter Ordnung lautet:

$$(3) \quad \pi_{ii}^i = R_{ii}^i - C_{ii}^i < 0.$$

Die Lösung der Gleichung (2) für $i = 1, 2$ gibt uns die optimalen Produktionsmengen in Abhängigkeit der vorab bestimmten privaten Kapitalstöcke. Diese Lösungen können wir kompakt schreiben als:

$$(4) \quad Y^i = Y^i(K^1, K^2).$$

wobei gilt:

$$(5) \quad Y_i^i = \frac{dY^i}{dK^i} > 0 \quad \text{sowie:} \quad Y_j^i = \frac{dY^j}{dK^i} < 0,$$

D.h. der Output Y_i der Firma i steigt (sinkt), wenn Firma i (Firma j) den Kapitalstock erhöht.⁶

ii) Der private Kapitaleinsatz

Mit Hilfe von (4) können wir (1) umschreiben zu:

$$(1') \quad \begin{aligned} \mu^i &= \pi^i(Y^1(K^1, K^2), Y^2(K^1, K^2), K^i, M^i, G^i) \\ &= [R^i(Y^1(\dots); Y^2(\dots)) - C^i(Y^i(\dots), K^i, G^i) \\ &\quad - vK^i](1 - t^e M^i) - t^K M^i K^i \end{aligned}$$

und uns der zweiten Stufe zuwenden. Aus (1') erhalten wir die Bedingung erster Ordnung für den Kapitaleinsatz:

⁶ Für den Beweis von (5) und einige der nachfolgenden Resultate sei auf den Anhang verwiesen, der vom Autor angefordert werden kann.

$$(2') \quad \mu_i^i = \left[\pi_i^i Y_i^i + \pi_j^i Y_i^j - \frac{\partial C^i}{\partial K^i} - v \right] (1 - t^e M^i) - t^K M^i = 0,$$

wobei wegen (2) die Restriktion $\pi_i^i = 0$ zu beachten ist, so daß sich (2') vereinfachend schreiben läßt als:

$$(6) \quad \mu_i^i = \left[R_j^i Y_i^j - C_{K_i}^i - v \right] (1 - t^e M^i) - t^K M^i = 0$$

Gleichung (6) ist die Optimalitätsbedingung für den privaten Kapitaleinsatz, die in der Form:

$$(7) \quad -C_{K_i}^i = v - R_j^i Y_i^j + \frac{t^K M^i}{1 - t^e M^i}$$

einer leichteren ökonomischen Interpretation zugänglich ist. Im Vergleich zu einer Firma, die keiner Besteuerung unterliegt und unter vollkommener Konkurrenz agiert, verdeutlicht (7), daß Firma i infolge der Marktunvollkommenheit strategische Überinvestitionen tätigt ($R_j^i \cdot Y_i^j > 0$), während die Kapitalbesteuerung in die entgegengesetzte Richtung wirkt. Der erste Effekt wird in der Literatur als ‚rent-shifting‘-Effekt bezeichnet während der zweite Effekt einen ‚Besteuerungseffekt‘ darstellt. Der Nettoeffekt aus den beiden entgegengesetzten Effekten hängt von der relativen Stärke der beiden Einzeleffekte ab.

Bevor wir uns der Bestimmung der optimalen Regierungspolitik zuwenden, wollen wir noch untersuchen, wie die Firmen auf Änderungen der Politikparameter reagieren. Differenzieren wir (2') für beide Firmen total, so erhalten wir unter der üblichen Annahme, daß die direkten die indirekten Effekte dominieren, d.h., daß gilt $|D| = \mu_{11}^1 \mu_{22}^2 - \mu_{12}^1 \mu_{21}^2 > 0$:

$$(8) \quad \frac{dK^i}{dG^i} = K_{G_i}^i > 0 \text{ sowie: } \frac{dK^j}{dG^i} = K_{G_i}^j < 0$$

$$(9) \quad \frac{dK^i}{dM^i} = K_{M_i}^i < 0 \text{ sowie: } \frac{dK^j}{dM^i} = K_{M_i}^j < 0$$

Die Gleichungen (8) - (9) zeigen die von Steueränderungen bzw. Infrastrukturinvestitionen induzierten privaten Kapitalbestandsanpassungen. Eine Erhöhung des lokalen Infrastrukturangebots (des Hebesatzes) in der Gebietskörperschaft 1 induziert eine Mehrnachfrage (Mindernachfrage) nach privatem Kapital der Firma 1 (und wegen (5) damit auch ein größeres Outputvolumen der Firma 1) während die Firma 2 ein gegensätzliches Reaktionsmuster aufweist. Mehr öffentliche Infrastruktur in der Region 1 verschafft der Firma in dieser Region Kostenvorteile (die Grenzkosten

sinken) gegenüber der Firma in der Region 2 und führt zu einem Marktanteilsgewinn für die Firma 1. Eine Erhöhung des Gewerbesteuerhebesatzes wirkt hingegen in die umgekehrte Richtung. Infrastrukturinvestitionen bzw. Hebesatzerhöhungen sind somit mit strategischen Vorteilen bzw. Nachteilen für das ortsansässige Gewerbe verbunden.

Abbildung 1 verdeutlicht den Zusammenhang zwischen dem Politikvariableneinsatz und dem ‚Kapitalstockspiel‘ der beiden Firmen. Steigt G^1 (M^1), verschiebt sich die Kapitalstockreaktionsfunktion der Firma 1 nach rechts (links) mit dem Resultat, daß Firma 1 einen höheren (niedrigeren) privaten Kapitalstock realisiert, während Firma 2 ihre Investitionen reduziert (erhöht).⁷

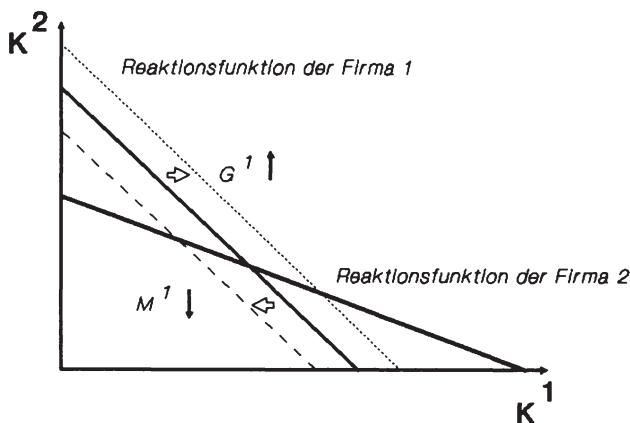


Abbildung 1: Die Reaktionsfunktionen der Firmen in den privaten Kapitalstöcken

Die Steigung der Reaktionsfunktion der Firma j ist:

$$(10) \quad \frac{dK^j}{dK^i} = - \frac{\mu_{ji}^j}{\mu_{jj}^j} < 0,$$

so daß die Reaktion der Firma j auf Veränderungen von G^i geschrieben werden kann als:

⁷ In Anlehnung an die Steuerkonkurrenzliteratur, siehe z.B. Wildasin (1988), lassen sich diese strategischen Effekte auf das Marktergebnis als fiskalische Externalitäten der Besteuerung bzw. der öffentlichen Infrastruktur interpretieren.

$$(11\text{a}) \quad K_{G_i}^j = K_{G_i}^i \frac{dK^j}{dK^i}$$

und analog:

$$(11\text{b}) \quad K_{M_i}^j = K_{M_i}^i \frac{dK^j}{dK^i}$$

Ferner halten wir die Bedingung:

$$(12) \quad K_{M_i}^i = \frac{t^K}{\frac{\partial^2 C^i}{\partial K^i \partial G^i}} (1 - t^\epsilon M^i) \cdot K_{G_i}^i = -\alpha^i K_{G_i}^i > 0$$

fest, die wir an späterer Stelle benötigen werden, wobei α^i im Intervall $0 \leq \alpha^i \leq 1$ liegt, wenn die Bedingung:

$$\left| C_{K_i, G_i}^i \right| \geq t^K / (1 - t^\epsilon M^i)$$

erfüllt ist.

iii) Das Verhalten der örtlichen Regierungen

In einem letzten Schritt analysieren wir die erste Stufe des Spiels, d.h. die optimale Wahl der Politikinstrumente G und M , durch Maximierung der lokalen Wohlfahrtsfunktion⁸:

$$(13) \quad W^i = \pi^i + T^i - G^i q (1 - zq^i)$$

die sich auch schreiben lässt als:

$$(13') \quad W^i = \pi^{i,g} - G^i q (1 - zq^i)$$

wobei $\pi^{i,g}$ den Bruttogewinn der Firma bezeichnet. (13') ist unter Beachtung der Budgetrestriktion:

$$(14) \quad T^i = M^i (t^K K^i + t^\epsilon \pi^{i,g}) = G^i q (1 - zq^i)$$

zu maximieren, wobei q den überregional fixierten Preis staatlicher Infrastrukturgüter und zq^i die exogen gegebene Investitionszuweisungsquote seitens der Landes- bzw. Bundesregierung angibt. Lokale Infrastrukturinvesti-

⁸ Wir nehmen hierbei an, daß der nationale Markt auf dem die Güter der beiden Firmen abgesetzt wird im Vergleich zum lokalen Gütermarkt sehr groß ist, so daß Effekte auf die Konsumentenrente in den Regionen von vernachlässigbarer Bedeutung sind.

tionen müssen somit – nach Abzug der Investitionszuweisungen – vollständig durch das örtliche Gewerbesteueraufkommen finanziert werden. Die Forderung nach einem ausgeglichenen Budget macht es daher erforderlich, zu jedem Wert von G^i einen das Budget ausgleichenden Hebesatz M^i zu bestimmen. Da wir G^i als die Entscheidungsvariable wählen⁹, müssen wir zunächst eine Gleichung für die Beziehung zwischen Änderungen des lokalen Infrastrukturkapitalstocks und der hierfür erforderlichen budgetausgleichenden Anpassung von M^i ableiten. Dies erreichen wir, indem wir die Budgetrestriktion (14) total differenzieren:

$$(15) \quad \frac{dM^i}{dG^i} = \frac{q(1 - zq^i) - M^i[t^e\pi_{G_i}^{i,g} + t^K K_{G_i}^i]}{\frac{G^i}{M^i} + M^i[t^e\pi_{M_i}^i + t^K K_{M_i}^i]} < 0.$$

Die Annahme $dM^i/dG^i > 0$ ist sicherlich plausibel, da man ansonsten sich selbst finanzierende Infrastrukturinvestitionen unterstellen würde. Wird (13') total nach G^i differenziert, erhalten wir:

$$(16) \quad \begin{aligned} W_{G_i}^i &= \frac{t^K M^i}{1 - t^e M^i} \left[K_{G_i}^i + K_{M_i}^i \frac{dM_i}{dG_i} \right] + R_j^i Y_j^i \left[K_{G_i}^j + K_{M_i}^j \frac{dM_i}{dG_i} \right] \\ &- q(1 - zq^i) - \frac{\partial C^i}{\partial G^i} = 0. \end{aligned}$$

Bei der Ableitung von (16) haben wir Cournot-Verhalten der Regierungen unterstellt, d.h. $dG^i/dG^j = 0$; ferner wurde der Umstand mißachtet, daß jede Anpassung von G^i (bei Konstanz von G^j) in der Kommune j zu einer Änderung des Steueraufkommens führt, und somit dort ein unausgeglichenes Budget verursacht. Die Annahme keiner Reaktion der Kommune j mit G^j bei einer Anpassung der Infrastrukturausgaben G^i würde bei gleichzeitiger Annahme ausgeglichener Budgets beider Regierungen zu einer passiven Anpassung des Hebesatzes M^j führen, um dort den Budgetausgleich zu gewährleisten, siehe hierzu Wildasin (1991).¹⁰ Die Berücksichtigung dieses indirekten Effektes würde aber die Analyse erheblich komplizieren, so daß wir diesen außer acht lassen.¹¹ (16) ist eine implizite Reak-

⁹ Alternativ könnte man auch den Hebesatz M^i als Entscheidungsvariable wählen und G^i die Rolle einer sich ‚passiv anpassenden‘ Instrumentvariable zuweisen. Man kann zeigen, siehe z.B. Wildasin (1991), daß die Wahl der Entscheidungsvariable nicht ergebnisneutral ist.

¹⁰ Steigt G^i (bei Konstanz von G^j), so reduziert sich der Gewinn der Firma j was in der Kommune j zu einer Erhöhung des Hebesatzes M^j führen müßte, um dort den Budgetausgleich zu erreichen, was wiederum von Vorteil für die Firma 1 wäre.

¹¹ Im Gleichgewicht sind die Budgets beider Städte selbstverständlich ausgeglichen.

tionsfunktion der lokalen Regierung, die sich in der Form $G^i = g_i(G^j, zq^i, zq^j)$ schreiben läßt. Wegen der Komplexität des Modells, die sich insbesondere aus der Einbeziehung der Finanzierungsrestriktion ergibt, ist es jedoch ohne weitere einschränkende Annahmen nicht möglich, die Steigung der Reaktionsfunktion zu bestimmen. Unter Vernachlässigung der Finanzierungsrestriktion läßt sich relativ einfach nachweisen, daß gilt: $dG^i/dG^j < 0$, $dG^i/dzq^i > 0$, sowie $dG^i/dzq^j < 0$.

Mit Hilfe von (11) und (12) läßt sich Gleichung (16) umschreiben zu:

$$(16') \quad q(1 - zq^i) = \frac{t^K M^i}{1 - t^\epsilon M^i} Z^i + R_j^i Y_j^i \frac{dK^j}{dK^i} Z^i + \left(-\frac{\partial C^i}{\partial G^i} \right)$$

wobei Z^i definiert ist als:

$$(20) \quad Z^i = \left[K_{G_i}^i + K_{M_i}^i \frac{dM^i}{dG^i} \right] = K_{G_i}^i \left[1 - \alpha_i \frac{dM^i}{dG^i} \right] \leq 0.$$

Die Gleichung (16') ist einer einfachen ökonomischen Interpretation zugänglich. Auf der linken Seite haben wir den lokalen Nettopreis (d.h. nach Abzug der Investitionszuschüsse) der kommunalen Infrastrukturgüter. Auf der rechten Seite haben wir drei Effekte:

- i) $(t^K M^i / 1 - t^\epsilon M^i) \cdot Z^i = R_j^i Y_j^i - C_{K^j}^i + v > 0$ ist die Änderung des Bruttogewinns der Firma i infolge der durch die zusätzliche Bereitstellung öffentlicher Infrastrukturdiensleistungen induzierten Anpassung des privaten Kapitalstocks,
- ii) $R_j^i Y_j^i (dK^j / dK^i) > 0$ ist ein Rent-Shifting Effekt der dadurch zu stande kommt, daß infolge des höheren Infrastrukturkapitalstocks die Firma i gegenüber Firma j einen Kostenvorteil hat, der ihr einen Marktanteilsgewinn bringt,
- iii) $\partial C^i / \partial G^i < 0$ ist der direkte Kostensenkungseffekt, der durch die Bereitstellung von mehr Infrastruktur eintritt.

Die ersten beiden Effekte werden mit dem Term Z^i multipliziert, der sowohl positiv als auch negativ sein kann. Z^i mißt den Nettoeffekt der Infrastruktur- und Steuerpolitik auf die private Kapitalakkumulation. Gilt $Z^i > 0$, so dominiert der kostensenkende und somit investitionsanregende Infrastruktureffekt die Belastung aus der zusätzlichen Besteuerung, so daß die Regierung einen Anreiz hat, G^i auszuweiten und hierfür eine Hebesatz erhöhung in Kauf zu nehmen. Im umgekehrten Fall, d.h. $Z^i < 0$, überkom pensiert die zusätzliche Steuerbelastung die positiven Effekte der Infrastruktur und führt somit zu einer Einschränkung der öffentlichen und privaten Investitionsaktivität aber auch einer Senkung des Hebesatzes. Eine „gute“

Infrastrukturpolitik ist somit nicht unbedingt gleichbedeutend mit einem höheren Infrastrukturangebot; vielmehr kann eine ‚gute‘ Infrastrukturpolitik auch in der Reduktion der Infrastrukturinvestitionsaktivität bestehen. Letzteres wäre dann der Fall, wenn die mit einer Ausweitung der öffentlichen Investitionsaktivität verbundene zusätzlich Steuerbelastung größer ist als die mit diesen Investitionen verbundenen direkten und indirekten ‚Benefits‘.

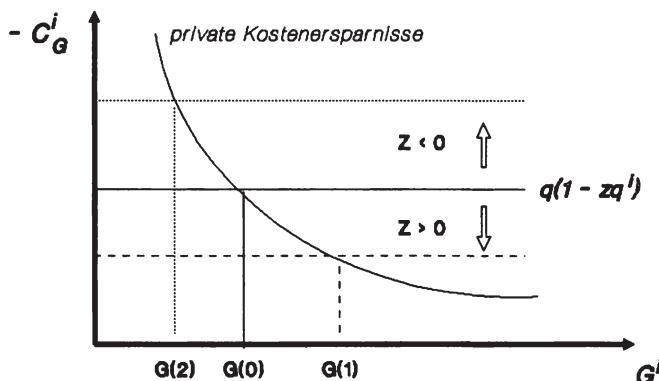


Abbildung 2: Die Optimalitätsbedingung für das lokale Infrastrukturangebot

Abbildung 2 verdeutlicht die Optimalitätsbedingung (16'). Der Schnittpunkt zwischen der Investitionsgüterpreisgerade $q(1 - zq^i)$ und der Kurve der privaten Kostenersparnisse, $-C_G^i$, gibt den optimalen Infrastrukturmobilstock $G(0)$ an, der bei nicht-strategischem Verhalten der Regierung und unter Vernachlässigung der Steuerfinanzierung realisiert wird.¹² Gilt $Z < 0$, wird die $q(1 - zq^i)$ -Linie nach oben verschoben, da die zusätzliche Steuerbelastung die Vorteile einer Erhöhung von G^i überkompenziert; entsprechend ist das Infrastrukturangebot mit $G(2)$ kleiner als der Kapitalstock $G(1)$. Das Umgekehrte trifft zu, wenn $Z > 0$ ist; in diesem Fall wird die Menge $G(1)$ angeboten.

Unser theoretisches Modell ist mangels Daten wohl kaum einem unmittelbaren empirischen Test zugänglich, aber es gibt uns einige wertvolle Hinweise für das empirische Arbeiten. Zunächst einmal sind die Produktivitätseffekte der öffentlichen Infrastrukturdiensleistungen eine zentrale Varia-

¹² Diese Infrastrukturvolumen wird auch gewählt wenn gilt $Z = 0$, d.h. die Effekte der Besteuerung und der Infrastrukturbereitstellung gleichen sich gerade aus.

ble unseres Modells. Daher untersuchen wir in einem ersten Schritt, ob von kommunalen Infrastrukturinvestitionen überhaupt nachweisbare Produktivitätseffekte ausgehen. Ferner legt unser Modell einen engen Zusammenhang zwischen den Produktivitätseffekten der öffentlichen Infrastruktur, der kommunalen Investitionsaktivität sowie der lokalen Besteuerung nahe. Mit Hilfe eines einfachen Regressionsmodells wird daher in einem zweiten Schritt der Zusammenhang zwischen diesen Variablen analysiert.

3. Die Schätzung der Produktivitätseffekte kommunaler Infrastruktur

Wir haben bereits in der Einleitung darauf hingewiesen, daß der größte Teil des öffentlichen Infrastrukturangebots auf lokaler Ebene von den Städten und Gemeinden zur Verfügung gestellt wird, und somit Infrastrukturdienstleistungen in einem hohen Maße den Charakter lokaler öffentlicher Güter haben. Dies impliziert, daß Infrastruktureffekte am besten auf kommunaler Ebene zu analysieren sind. Wir verwenden daher für unsere empirischen Untersuchungen Daten für das Verarbeitende Gewerbe in Städten.

Um die Produktivitätseffekte der Infrastruktur auf lokaler Ebene aufzudecken, unterstellen wir dem örtlichen Gewerbe in der Stadt i die Produktionsfunktion:

$$(17) \quad Y_i = f_i(L_i, K_i, t, B_i, G_i).$$

Y_i bezeichnet den Output, L_i den Arbeitseinsatz, K_i den privaten Kapiteleinsatz und der Zeitindex t dient als ‚Proxy‘-Variable für die Erfassung von technischem Fortschritt. Die Infrastrukturdienstleistungen, approximiert durch den öffentlichen Infrastrukturkapitalstock, G^i , gehen in den privaten Produktionsprozeß als ein fixer, unbezahlter Produktionsfaktor ein.¹³ Die Variable B^i ist eine Indikatorvariable für die Größe der Stadt, mit deren Hilfe wir Agglomerationseffekte erfassen. Der Index i bezieht sich auf das Verarbeitende Gewerbe in der Stadt i .

Die Minimierung der privaten Produktionskosten $w_i L_i + r_i K_i$ gegeben der Produktionsfunktion (17) resultiert in der Kostenfunktion:

$$(18) \quad C_i = C_i(w_i, r_i, t, B_i, G_i, Y_i),$$

wobei w_i den Lohnsatz und r_i die Nutzungskosten des privaten Kapitals bezeichnet. Differenzieren wir (18) nach G_i , so erhalten wir die privaten

¹³ Wir unterstellen also, daß die in die Produktion eingehenden Infrastrukturdienstleistungen proportional dem städtischen Infrastrukturkapitalstock sind.

Kostenersparnisse, s_i , die aus der zusätzlichen Bereitstellung einer Einheit Infrastrukturkapitals resultieren:

$$(19) \quad s_i = -\frac{\partial C_i}{\partial G_i} = s_i(w_i, r_i, t, B_i, G, Y_i).$$

Wegen der Annahme, daß es sich bei G^i um ein „freies“ Gut handelt, sollte man erwarten, daß $s_i \geq 0$ gilt. Neben dieser dualen Messung der Produktivitätseffekte der öffentlichen Infrastruktur können wir aus der Kostenfunktion auch Informationen darüber ableiten in welcher Beziehung G_i zu den privaten Inputs steht. Hierzu differenzieren wir die bedingten Faktornachfragefunktionen nach G_i :

$$(20) \quad s_{i,K} = \frac{\partial K_i}{\partial G_i} = \frac{\partial^2 C_i}{\partial r_i \partial G_i}; \quad s_{i,L} = \frac{\partial L_i}{\partial G_i} = \frac{\partial^2 C_i}{\partial w_i \partial G_i}.$$

Ist $s_{i,L}(s_{i,K}) > 0$ ist die öffentliche Infrastruktur arbeits(kapital-)sparend wohingegen wir im Fall $s_{i,L}(s_{i,K}) < 0$ einen substitutiven Effekt zwischen Infrastruktur und Arbeitsinput (Kapitalinput) erhalten.

Die ökonometrische Schätzung der Kostenfunktion macht es erforderlich, daß wir der Kostenfunktion eine explizite funktionale Form zuweisen. Um der zugrundeliegenden Technologie möglichst wenig restriktive Annahmen aufzuzwingen haben wir eine flexible funktionale Form in Gestalt einer Translog-Kostenfunktion ausgewählt, die geschrieben werden kann als:

$$(21) \quad \ln C_i = \sum_{i=1}^N \alpha_{0i} \cdot D_i + \sum_{i=1}^N \alpha_{Li} \ln \frac{w_i}{r_i} \cdot D_i + \sum_{k=1}^K \left(\gamma_{0k} Z_{ki} + \gamma_{1k} Z_{ki} \ln \frac{w_i}{r_i} \right) \\ + \ln r_i + \beta_x \ln X_i + \beta_t t + \beta_B \ln B_i + \beta_G \ln G_i \\ + 0.5 (\beta_{LL} \ln^2 \frac{w_i}{r_i} + \beta_{XX} \ln^2 X_i + \beta_{tt} t^2 + \beta_{BB} \ln^2 B_i \\ + \beta_{GG} \ln^2 G_i) + \beta_{LX} \ln X_i \ln \frac{w_i}{r_i} + \beta_{Lt} t \cdot \ln \frac{w_i}{r_i} + \beta_{LB} \ln B_i \ln \frac{w_i}{r_i} \\ + \beta_{LG} \ln G_i \ln \frac{w_i}{r_i} + \beta_{xt} t \ln X_i + \beta_{XB} \ln B_i \ln X_i \\ + \beta_{xG} \ln G_i \ln X_i + \beta_{tB} \ln B_i t + \beta_{tG} \ln G_i t + \beta_{BG} \ln G_i \ln B_i + U_C$$

und die Arbeitskostenanteilsgleichung:

$$(22) \quad \frac{\partial \ln C_i}{\partial \ln w_i} = \frac{w_i L_i}{C_i} = \sum_{i=1}^N \alpha_{Li} \cdot D_i + \alpha_{LL} \ln \frac{w_i}{r_i} + \sum_{k=1}^K \gamma_{1k} Z_{ki} \\ + \beta_{LX} \ln X_i + \beta_{Lt} \cdot t + \beta_{LB} \ln B_i + \beta_{LG} \ln G_i + U_L$$

impliziert. Die Variablen D_i bezeichnen Dummy-Variablen, die den Wert 1 für Stadt i und 0 sonst annehmen; die Variablen Z_{ki} bezeichnen $k = 3$ Industriestrukturvariablen, die wir an späterer Stelle konkretisieren. Unsere Spezifikation stellt sicher, daß die Arbeitsnachfragefunktion konsistent aus der Kostenfunktion abgeleitet werden kann. Der Parameter α_{0i} erfaßt stadt-spezifische Absolutglieder in der Kostenfunktion und α_{Li} stadspezifische Effekte in der Arbeitsnachfragegleichung.¹⁴ Für die Schätzung wurde die Kostenfunktion und die Arbeitsangleichung um die Störterme U_C bzw. U_L ergänzt. Bei dem zu schätzenden Modell handelt es sich somit um ein simultanes Gleichungssystem mit ‚fixed effects‘, das wir iterativ mit der Zellner Methode schätzen, wobei wir die stadspezifischen ‚fixed-effects‘ zunächst durch Anwendung der ‚within-Transformationen‘ eliminieren, siehe Hsiao (1986). Man beachte, daß unser Schätzansatz implizit unterstellt, daß die Technologieparameter – abgesehen von den Absoluttermen – über die Städte identisch sind. Die Schätzung städtespezifischer Parameter würde die Verfügbarkeit einer großen Datenmenge auf der Ebene von Städten erforderlich machen.¹⁵

Die Spezifikation (21) und (22) beinhaltet als Spezialfälle ein sehr breites Spektrum von in der Literatur häufig verwendeten Ansätzen: a) Hicks-neutrale technischer Fortschritt impliziert $\beta_{Lt} = \beta_{Xt} = \beta_{tB} = \beta_{tG} = 0$; b) Hicks-neutrale Agglomerationseffekte implizieren $\beta_{LB} = \beta_{XB} = \beta_{tB} = \beta_{BG} = 0$, und c) konstante Skalenerträge implizieren $\beta_{LX} = \beta_{Xt} = \beta_{XB} = \beta_{XG} = 0$, und $\beta_X = 1$. Bei der Schätzung können natürlich auch Kombinationen dieser Restriktionen verwendet werden.

Die Schätzung wurde mit einem Paneldatensatz des Verarbeitenden Gewerbes in 85 westdeutschen kreisfreien Städten für die Jahre 1980 - 1989 durchgeführt.¹⁶ Der Arbeitsinput ist gemessen durch die insgesamt geleisteten Beschäftigungsstunden¹⁷, und der Lohnsatz wurde berechnet durch Division der Bruttolohn- und -gehaltssumme (einschließlich der Sozialabgaben der Arbeitgeber) durch die gesamten Arbeitsstunden von Arbeitern und Angestellten. Städtespezifische Daten über das reale Nettoanlagevermögen des Verarbeitenden Gewerbes wurden an der Universität Münster berechnet, siehe Deitmer (1993). Die privaten Kapitalnutzungskosten basieren auf dem ‚user-

¹⁴ Die ebenfalls aus (21) ableitbare Kostenanteilsgleichung für privates Kapital kann bei der Schätzung nicht berücksichtigt werden da sich die beiden Anteilsgleichungen zu 1 addieren.

¹⁵ Zur Notationsvereinfachung haben wir bei allen Variablen in den Gleichungen (21) und (22) den Zeitindex t unterdrückt; selbstverständlich sind alle Daten – mit Ausnahme der Städte-Dummys D_i – zeitpunktspezifisch gemessen.

¹⁶ Siehe Seitz (1993b) für eine ausführliche Darstellung des Städtepanels.

¹⁷ Unsere Datenquelle, das ‚Statistische Jahrbuch Deutscher Gemeinden‘ weist nur die Arbeiterstunden aus. Wir haben daher die übliche Annahme gemacht, daß Angestellte die gleiche Jahresarbeitszeit wie Arbeiter haben.

cost'-Konzept von Jorgenson und berücksichtigen Unterschiede in der lokalen Besteuerung sowie der Investitionszuschüsse und -zulagen, siehe Deitmer (1993)¹⁸. Der Output wird durch die Bruttowertschöpfung des Verarbeitenden Gewerbes gemessen. Da das Kostenminimierermodell die Exogenität des Outputs unterstellt, haben wir bei der Schätzung die Outputvariable instrumentiert, um einen möglichen ‚simultaneity bias‘ zu berücksichtigen.¹⁹ Der Bestand an realem (Preisbasis 1980) städtischen Infrastrukturskapital wurde aus den Investitionsausgabedaten der Städte der Periode 1969 - 1989 durch Anwendung der ‚perpetual inventory method‘ geschätzt. Startwerte für die städtischen Infrastrukturskapitalstöcke im Jahr 1969 wurden aus bundeslandspezifischen Kapitalstockdaten abgeleitet; siehe Seitz (1993b) für Details. Für die Schätzung wurde der städtische Infrastrukturskapitalstock durch die Gesamtbeschäftigung in der Stadt dividiert, um unterschiedlichen Größenordnungen Rechnung zu tragen. Ferner geht der Infrastrukturskapitalstock mit einer einperiodigen Verzögerung in die Schätzung ein da von neuerrstellten Anlagen im laufenden Jahr wohl kaum Produktivitätseffekte zu erwarten sind. Als Agglomerationsvariable haben wir die Gesamtzahl der ‚sozialversicherungspflichtig Beschäftigten‘ in der Stadt verwendet. Die Industriestrukturvariablen, Z_{ki} , in unseren Schätzgleichungen (21) und (22) sind die Beschäftigungsanteile (gemessen an der Gesamtbeschäftigung im Verarbeitenden Gewerbe in der Stadt) in der i) Grundstoff- und Produktionsgüterindustrie, Z_{1i} , ii) Investitionsgüterindustrie, Z_{2i} , und iii) Konsumgüterindustrie, Z_{3i} . Der Beschäftigungsanteil in der Industrie ‚Nahrungs-, Genußmittel und Getränke‘ muß unberücksichtigt bleiben, da sich sonst die Beschäftigungsanteile zu 1 addieren.

Bei der Schätzung haben wir folgende Strategie verfolgt: Zunächst wurde das sehr allgemeine System (21) und (22) ohne weitere Restriktionen geschätzt (Modell 1), d.h. alle Interaktionsterme mit t und B_i wurden berücksichtigt. In weiteren Schätzungen haben wir systematisch Restriktionen bezüglich des technischen Fortschritts sowie der Wirkung der Agglomerationsvariable B_i eingeführt. Modell 2 unterstellt, daß die Variable B_i in die Kostenfunktion hicks-neutral eingeht, wohingegen t mit allen Interaktionstermen erscheint. Modell 3 vertauscht die Rolle von t und B_i in Modell 2. In Modell 4 wird sowohl die Variable B_i als auch die Variable t hicks-neutral spezifiziert. Ferner wurde das Modell 5 geschätzt bei dem alle Parameter bezüglich B_i und t auf Null gesetzt wurden. Aus Platzgründen verzichten wir

¹⁸ Wir schulden Hans-Joachim Schalk und Ingo Deitmer (beide Universität Münster) großen Dank für die Bereitstellung der Kapitalstock- und Kapitalnutzungskostendaten.

¹⁹ Hierbei wurden folgende Instrumente verwendet: Dummies für die Städte, der Zinssatz, die Stadtbevölkerung sowie deren quadrierten Wert, der Infrastrukturskapitalstock, die Beschäftigung außerhalb des Verarbeitenden Gewerbes sowie die Arbeitslosenquote in der Stadt.

Tabelle 1: Schätzergebnisse für alternativ spezifizierte Kostenfunktionen*)

A: Ge testete Modelle:							
	Zeitvariable		Agglomerationssvariable		$\log L$		η_{CG}
1:	alle Interaktionsterme		alle Interaktionsterme		4197.07		0.139
2:	alle Interaktionsterme		Hicks-neutral		4165.13		0.123
3:	Hicks-neutral		alle Interaktionsterme		4105.85		0.136
4:	Hicks-neutral		Hicks-neutral		4082.74		0.127
5:	unterdrückt		unterdrückt		4042.51		0.162
6:	Hicks-neutral und Annahme konstante Skalenerträge		Hicks-neutral		3541.19		0.345

B: Geschätzte Parameter: Model 4**)							
	α_L	(11.7)	γ_{01}	0.0013	(1.2)	γ_{11}	- 0.0009
γ_{12}	0.0010	(2.7)	γ_{03}	0.0007	(0.7)	γ_{13}	0.0009
β_r	- 0.0290	(11.5)	β_B	- 4.6687	(- 4.2)	β_G	2.8493
β_{xx}	0.0539	(- 3.6)	$\beta_{L,L}$	0.0961	(31.1)	β_{GG}	- 0.7975
β_{LX}	0.0115	(2.1)	β_{LG}	- 0.0550	(- 4.4)	β_{xG}	- 0.0515
R^2	Kostenfunktion: 0.76; Arbeitsnachfragefunktion: 0.79.						

*)
 $\log L$ = Log der Likelihood-Funktion; η_{CG} = durchschnittliche Infrastrukturelastizität.
**) Zahlen in Klammern sind t -Werte. α_L und α_{01} sind Durchschnittswerte über die 85 Städte.

auf die Dokumentation der detaillierten Schätzergebnisse und weisen im oberen Teil der *Tabelle 1* für alle Modellvarianten lediglich den Log der Likelihood-Funktion (Log L) sowie den Durchschnittswert der geschätzten Infrastrukturelastizität der Kostenfunktion, η_{CG} , die definiert ist als:

$$(23) \quad \eta_{CG} = s_i \cdot \frac{G_i}{C_i} = - \frac{\partial C_i}{\partial G_i} \cdot \frac{G_i}{C_i}.$$

Zieht man Likelihood-Ratio Tests für den Modellvergleich heran, so zeigt Modell 1 die beste Anpassungsgüte. Hier waren jedoch alle Interaktionsterme zwischen den Variablen t und B_i sowie den anderen Modellvariablen nicht signifikant. Da wir die Bestimmung der Produktivitätseffekte der Infrastruktur nicht auf insignifikante Parameter basieren wollten, haben wir Modell 4 als Basis für unsere weitere Auswertung herangezogen. In der Spezifikation 4 erscheinen keine insignifikanten Interaktionsparameter, die einen Einfluß auf die Schätzung der Infrastrukturelastizität der privaten Kosten haben. Wie aus Tabelle 1 hervorgeht, weichen die Durchschnittswerte der geschätzten Infrastrukturelastizitäten η_{CG} bei den unterschiedlichen Modellvarianten nur unwesentlich voneinander ab. Man beachte, daß das Modell 4 die Produktionsfunktion:

$$(17') \quad Y_i = g_t(t) \cdot g_B(B_i) \cdot f_i(L_i, K_i, G_i)$$

impliziert, wobei $g_t(..)$ bzw. $g_B(..)$ Hicks-neutrale Shift-Funktionen für den technischen Fortschritt bzw. Agglomerationseffekte darstellen. Produktionsfunktionen in der Form (17') werden häufig in stadtökonometrischen Untersuchungen postuliert, siehe z.B. *Henderson* (1986, 1988), und unsere Schätzungen belegen, daß diese Spezifikation auch recht gut mit den Daten kompatibel ist.

Als Nächstes wenden wir uns den ökonomischen Implikationen der Schätzungen des Modells 4 zu, dessen Parameter und T -Werte im unteren Teil der Tabelle 1 ausgewiesen sind. Für die bedingte Arbeitsnachfrage- bzw. Kapitalnachfragefunktion errechnen sich durchschnittliche direkte Faktorpreiselastizitäten von -0.21 bzw. -0.11 . Die Schätzungen $\beta_B < 0$ und $\beta_{BB} > 0$ deuten auf signifikante aber abnehmende Agglomerationseffekte und für den technischen Fortschritt erhalten wir kostensenkende Effekte ($\beta_t < 0$) die aber über den Beobachtungszeitraum leicht zurückgehen ($\beta_{tt} > 0$). Der Parameter β_{LG} ist negativ, was darauf hindeutet, daß bei festem Output und festen Faktorpreisen, Infrastruktur arbeitssparend, aber kapitalintensivierend wirkt. Die geschätzten durchschnittlichen Infrastrukturelastizitäten der bedingten Faktornachfrage nach Arbeit bzw. Kapital betragen -0.20 bzw. 0.13 . Die Infrastrukturelastizität der Kostenfunktion im Verarbeitenden Gewerbe, η_{CG} , hat einen Durchschnittswert von 0.13 , das

bedeutet, daß (im Durchschnitt) ein 1%-iger Anstieg des realen Nettoinfrastrukturkapitalstocks eine Kostenreduktion im Verarbeitenden Gewerbe von ca 0.13% induziert.²⁰ Wir haben das Modell 4 auch ohne die Infrastruktivariablen geschätzt, aber sowohl die Likelihood-Ratio-Statistik als auch die einzelnen *T*-Werte deuten klar auf die Signifikanz der Infrastruktur als Erklärungsvariable im Kostenminimierermodell hin. Auch wurde auf das Vorliegen konstanter Skalenerträge getestet, siehe Modell 6 in Tabelle 1; diese Restriktion wird aber von den Daten nicht gestützt.

Aus unserer Schätzung lassen sich drei wichtige Implikationen über den Einfluß der lokalen Infrastruktur auf das Verarbeitende Gewerbe in den Städten ableiten: 1.) gibt es signifikante kostenreduzierende Effekte, 2.) öffentliches und privates Kapital sind komplementär, d.h. öffentliche Infrastrukturinvestitionen induzieren zusätzlich private Investitionen und 3.) haben Infrastrukturdienstleistungen arbeitssparende Effekte. Letzteres ist teilweise bedingt durch die substitutive Beziehung zwischen privatem Kapital und privatem Arbeitseinsatz: Eine Ausweitung des öffentlichen Infrastrukturangebots induziert zusätzliche private Investitionen, die ihrerseits wieder arbeitssparend wirken, da privater Arbeits- und Kapitalinput substitutiv sind.²¹ Tabelle A1 im Anhang präsentiert stadspezifische Schätzungen der Infrastrukturelastizität der Kostenfunktion und gibt Auskunft über die Signifikanz der Schätzungen. Für die Mehrzahl der Städte sind die geschätzten Elastizitäten signifikant von Null verschieden, und nahezu alle negativen Elastizitäten sind insignifikant. Wir haben die geschätzten Elastizitäten auf die Stadtbevölkerung und einen Zeitindex regressiert. Hierbei zeigte sich keine Korrelation zwischen der Größe der Stadt, gemessen an der Bevölkerungszahl, und η_{CG} . Jedoch zeigt sich, daß η_{CG} in der Zeit leicht ansteigend ist, so daß man hieraus die Schlußfolgerung ziehen könnte, daß durch zusätzliche kommunale Infrastrukturinvestitionen der Produktivitätsanstieg in der Industrie unterstützt werden könnte. Dieses Ergebnis sollte auch im Licht der Tatsache gesehen werden, daß die Wachstumsrate des gesamten kommunalen Infrastrukturkapitalstocks in den letzten Jahren ständig zurückgegangen ist; so ist die jährliche Wachstumsrate der städtischen Infrastrukturkapitalstücke in unserem Datensatz von 4.1% im Jahr 1980 stetig auf 1.4% im Jahr 1989 zurückgegangen.

²⁰ Man beachte, daß G_i in die Schätzung als Infrastrukturkapitalbestand pro Beschäftigten eingeht, so daß sich die Elastizitätsinterpretation eigentlich auf diese Infrastrukturkennziffer beziehen müßte. Implizit wird mit unserer Variablendefinition auch ein Rivalitätseffekt unterstellt, d.h. bei konstantem Infrastrukturkapitalbestand und steigender Beschäftigung in der Stadt gehen die Kostenvorteile aus der Infrastruktur zurück.

²¹ Zu beachten ist auch, daß wir keine Outputeffekte der Infrastruktur mit dem Kostenminimierermodell identifizieren können, da wir lediglich konditionale Faktornachfragefunktionen schätzen, siehe hierzu Seitz (1994).

4. Ein einfaches Regressionsmodell zur Erklärung der städtischen Investitionsaktivität und der Gewerbesteuerhebesätze

In einem zweiten Analyseschritt formulieren wir ein einfaches Regressionsmodell zur Erklärung der Anpassung der städtischen Infrastrukturkapitalstöcke und Gewerbesteuerhebesätze. Hierzu leiten wir zunächst aus unserem theoretischen Modell das Konzept eines optimalen lokalen Infrastrukturkapitalstocks ab, G_i^* , also jener Kapitalstock, der aus dem Optimierermodell in Abschnitt 2 resultiert. Um die Anpassung des tatsächlichen an den optimalen Kapitalstock abzubilden, unterstellen wir, daß es Finanzierungsrestriktionen sowie Anpassungskosten gibt, die eine sofortige Anpassung des tatsächlichen an den gewünschten Bestand verhindern. Hierzu führen wir ein partielles Anpassungsmodell für die lokalen Infrastrukturkapitalstöcke in der folgenden Form ein:

$$(22a) \quad G_{i,t} - G_{i,t-1} = (1 - \lambda_G) \cdot (G_{i,t}^* - G_{i,t-1}).$$

λ_G ist der partielle Anpassungskoeffizient, aus dem sich ein ‚mean lag‘ von $\lambda_G/(1 - \lambda_G)$ ergibt. Dieser gibt an, wieviel Jahre es dauert bis 50% der Diskrepanz zwischen tatsächlichem und optimalem Kapitalstock geschlossen ist. Für $\lambda_G \rightarrow 1 (\rightarrow 0)$ gibt es überhaupt keine (eine sofortige) Anpassung. Ein analoges Anpassungsmodell formulieren wir für den optimalen Gewerbesteuerhebesatz M_i^* mit dem Anpassungsparameter λ_M :

$$(23a) \quad M_{i,t} - M_{i,t-1} = (1 - \lambda_M) \cdot (M_{i,t}^* - M_{i,t-1}).$$

Um ‚Konkurrenzaspekte‘ in die Schätzung einzubeziehen, haben wir zwei ‚Konkurrenzvariablen‘ konstruiert. Die erste ‚Konkurrenzvariable‘ $G_{D,i}$ ist der durchschnittliche Pro-Kopf-Infrastrukturkapitalstock aller Städte, die von der Stadt i innerhalb einer Stunde zu erreichen sind (gemessen an der schnellsten Zugverbindung). Eine zweite ‚Konkurrenzvariable‘ $G_{I,i}$ haben wir dadurch konstruiert, daß wir aus unserem Datensatz zu jeder Stadt i jene fünf Städte herausgesucht haben, deren Wirtschaftsstruktur der Stadt i ‚am nächsten kommt‘.²² Auch für diese Städte haben wir den durchschnitt-

²² Um ‚strukturähnliche‘ Städte zu identifizieren, haben wir das Abweichungsmaß:

$$\frac{1}{12} \cdot \sum_{l=1}^{12} (s_{i,l} - s_{j,l})^2 \quad \forall j \neq i$$

berechnet, wobei l den Index von 12 Wirtschaftssektoren bezeichnet und s die Beschäftigungsanteile dieser Sektoren im Jahr 1985. Dieses Maß wurde jeweils für alle Städte im Vergleich zu allen anderen Städten berechnet, so daß jede Stadt i mit allen anderen 84 Städten j verglichen wurde. Die berechneten Maße wurden sortiert und für jede Stadt i wurden jene 5 Städte ausgesucht, die das geringste Abweichungsmaß zur Stadt i aufweisen.

lichen Pro-Kopf-Infrastrukturkapitalstock errechnet. G_{Di} mißt den Einfluß räumlich benachbarter Städte auf die Investitionspolitik der Stadt i (räumliche Konkurrenz), während G_{Ii} den Einfluß der Infrastrukturpolitik wirtschaftlich ähnlich oder gleich strukturierter Städte auf die Politik der Stadt i messen soll (Marktwettbewerb). Ferner haben wir die korrespondierenden Variablen T_{Di} und T_{Ii} unter Verwendung der Gewerbesteuerhebesätze konstruiert. T_{Di} ist der durchschnittliche Hebesatz der benachbarten Städte, und T_{Ii} ist der durchschnittliche Hebesatz der 5 strukturgleichen Städte. Für die Schätzung haben wir einfache Gleichungen für den optimalen Kapitalstock und Hebesatz formuliert²³:

$$(22b) \quad \ln \left(\frac{G}{P} \right)_{it}^* = \sum_{i=1}^{85} \alpha_i D_i + \alpha_{IG} \ln \left(\frac{IG}{P} \right)_{it} + \alpha_{PI} \ln PI_{it} + \alpha_t \text{time} \\ + \alpha_E \ln \left(\frac{E}{P} \right)_{it} + \alpha_T \ln \left(\frac{T}{P} \right)_{it} + \alpha_\eta \eta_{CG,i,t-1} + \alpha_D \ln \left(\frac{\text{Debt}}{P} \right)_{t-1,i} \\ + \alpha_{CD} \ln G_{Dit} + \alpha_{CI} \ln G_{Iit}$$

$$(23b) \quad \ln M_{it}^* = \sum_{i=1}^{85} \beta_i D_i + \beta_B \ln \left(\frac{BT}{T} \right)_{i,t-1} + \beta_t \text{time} \\ + \beta_D \ln \left(\frac{\text{Debt}}{P} \right)_{t-1,i} + \beta_{CD} \ln M_{Dit} + \beta_{CI} \ln M_{Iit}$$

PI ist der Preisindex der städtischen Investitionsgüter, der ein (stadtspezifischer) gewichteter Durchschnitt des (nationalen) Preisindex für staatliche Bau- und Ausrüstungsinvestitionen ist. P bezeichnet die Bevölkerungszahl, IG die realen Investitionszuschüsse, E die Gesamtzahl ‚sozialversicherungspflichtig Beschäftigter‘ in der Stadt, T die gesamten Steuereinnahmen, BT die Gewerbesteuererinnahmen und ‚Debt‘ die Verschuldung der Stadt. Somit wird in der Gleichung (22b) der optimale Infrastrukturkapitalstand erklärt durch die gewährten Investitionszuschüsse, den Preisindex kommunaler Investitionen, der Beschäftigungsdichte, einem Zeitindex sowie der Kommunalverschuldung. Zusätzlich wurden die Konkurrenzvariablen G_{Di} und G_{Ii} sowie die geschätzte Kostenelastizität der kommunalen Infrastruktur, η_{CG} , als Indikator für die mit Infrastrukturinvestitionen verbundenen Produktivitätseffekte in die Gleichung einbezogen. Berücksichtigt sind wiederum stadtspezifische Dummies, D_i , um unbeobachtbare stadtspezifische

²³ Um Heteroskedastieproblem zu vermeiden wurden die Variablen, soweit dies sinnvoll ist, in Gleichung (22a) und (22b) in Pro-Kopf-Größen transformiert. Man beachte ferner, daß die Gleichungen in Logs spezifiziert sind, so daß die geschätzten Parameter unmittelbar als Elastizitäten interpretiert werden können.

Effekte aufzufangen. Der optimale Gewerbesteuerhebesatz wird in Gleichung (22b) durch die (verzögerte) Pro-Kopf-Verschuldung der Stadt, den Anteil der Gewerbesteuereinnahmen an den gesamten Steuereinnahmen, einem Zeitindex sowie fixed effects für die einzelnen Städte bestimmt. Ferner sind die beiden ‚Konkurrenzvariablen‘ T_{D_i} und T_{I_i} einbezogen, um den Einfluß der Steuerpolitik benachbarter und gleich strukturierter Städte zu erfassen.

Tabelle 2: Schätzergebnisse für die Infrastruktur- und Hebesatzanpassungsgleichungen¹⁾

<i>Infrastrukturgleichung</i>			<i>Hebesatzgleichung</i>		
α_i	3.339	(2.2)	β_i	1.442	(2.1)
α_{IG}	0.028	(5.1)	β_B	- 0.027	(2.1)
α_{PI}	- 0.399	(1.8)	β_t	0.001	(0.7)
α_t	- 0.015	(2.3)	β_{CI}	0.175	(1.5)
α_E	0.289	(2.9)	β_{DD}	0.553	(7.3)
α_T	0.365	(3.9)	λ_M	0.495	(18.2)
α_η	0.475	(5.0)			
α_D	- 0.130	(5.7)			
α_{CI}	0.404	(2.7)			
α_{CD}	0.097	(0.7)			
λ_G	0.810	(8.2)			
R^2	0.996		R^2	0.920	

¹⁾ Schätzperiode: 1981 - 1989. Die Parameterwerte für α_i und β_i sind Durchschnittswerte über die 85 Städte. Die Zahlen in Klammern sind T-Werte.

Gleichung (22b) wurde in die logarithmierte Form der Gleichung (22a) eingesetzt, wobei der Infrastrukturkapitalstock auf die Stadtbevölkerung bezogen wurde; Gleichung (23b) wurde in die logarithmierte Form der Gleichung (23a) eingesetzt. Die beiden resultierenden Gleichungen wurden simultan mit der iterativen Zellnermethode geschätzt.²⁴ Tabelle 2 präsentiert die Schätzergebnisse. Aus Platzgründen wollen wir uns bei der Interpretation auf die in unserem Kontext relevanten Variablen beschränken. Der geschätzte Anpassungskoeffizient in der Hebesatzgleichung, λ_M , impli-

²⁴ Die Variable $\ln(T/P)$ in der Infrastrukturgleichung wurde instrumentiert, da das Steueraufkommen in der Periode t direkt von dem endogenen Hebesatz M abhängig ist. Nickel (1981) hat darauf hingewiesen, daß es in dynamischen Modellen mit ‚fixed effects‘ einen Bias gibt und daher Instrumentvariablenmethoden anzuwenden wären. Dies wurde in der gegenwärtigen Schätzung noch nicht berücksichtigt.

ziert eine sehr schnelle Anpassung, während der Anpassungsparameter in der Infrastrukturgleichung, λ_G , eine sehr langsame Anpassung des Infrastrukturkapitalstocks impliziert; die ‚mean-lags‘ betragen ca. 1 Jahr für den Hebesatz und etwas über 4 Jahre bei der Infrastruktur. Die geschätzte Infrastrukturelastizität der Produktionskosten im Verarbeitenden Gewerbe, η_{CG} , die wir als eine ‚Proxy‘ für den Produktivitätsbeitrag der städtischen Infrastruktur betrachten, geht signifikant in die Infrastrukturgleichung ein, was ein Hinweis darauf ist, daß die Produktivitätseffekte der Infrastruktur bei der Investitionsentscheidung eine Rolle spielen: Je größer η_{CG} ist, desto größere Produktivitätsreserven können durch zusätzliche Infrastrukturinvestitionen noch mobilisiert werden, und umso größer sind nach unseren Schätzungen auch die Infrastrukturinvestitionen.²⁵ Wenden wir uns nun den ‚Konkurrenzvariablen‘ zu: Erhöhen Städte mit einer ‚ähnlichen‘ Sektorenstruktur ihren Kapitalstock um 1%, so erhöht die Stadt i ihren Kapitalstock im Durchschnitt um 0.4%, wohingegen die Hebesatzerhöhungen dieser Städte durch die Stadt i nicht mitgemacht werden, was durch die Insignifikanz des Parameters β_{CI} angezeigt wird. Bezuglich der räumlichen ‚Konkurrenzvariablen‘ beobachten wir genau das entgegengesetzte Muster: Die Stadt i reagiert nicht auf erhöhte Infrastrukturinvestitionsaktivitäten der benachbarten Städte, ‚kopiert‘ allerdings deren Hebesatzpolitik: Erhöhen die benachbarten Städte den Hebesatz um 1%, so folgt die Stadt i mit einer Erhöhung von ca. 0.5%. Die Daten deuten somit darauf hin, daß Städte mit ähnlicher Sektorenstruktur bestrebt sind, eine vergleichbare Infrastrukturversorgung anzubieten, und daß Städte bei der Festlegung ihrer Gewerbesteuerhebesätze sich an den Sätzen in unmittelbarer Nachbarschaft orientieren. Dieses Ergebnis erscheint uns nicht unplausibel.

5. Zusammenfassung und Schlußfolgerungen

Mit Hilfe eines einfachen theoretischen Modells wurde gezeigt, daß Infrastrukturinvestitionen im interregionalen Wettbewerb eine handelsstrategische Rolle spielen können.²⁶ Hierbei sind jedoch die mit einer besseren Infrastrukturversorgung einhergehenden Vorteile mit den Nachteilen aufzuwägen, die aus der zusätzlichen Steuerlast, resultieren die mit der Finanzierung zusätzlicher öffentlicher Infrastrukturinvestitionen verbunden ist. Im empirischen Teil des Beitrages wurde gezeigt, daß von der Infrastrukturver-

²⁵ Wir haben das Modell auch mit dem Subsample der Städte geschätzt, für die der T -Wert von η_{CG} größer als 1.75 war, siehe Tabelle A1. Die dabei erzielten Ergebnisse weichen nur unwesentlich von denen in Tabelle 2 ab.

²⁶ Man sollte nicht versäumen darauf hinzuweisen, daß Ramser bereits in einem Beitrag von 1968 ein Modell dezentraler Infrastrukturversorgung unter dem Aspekt des interregionalen Wettbewerbs vorgestellt hat, siehe Ramser (1968).

sorgung in den Städten signifikante produktivitätsfördernde und damit kostensenkende Effekte ausgehen. Ferner zeigte die Analyse der kommunalen Investitionsaktivität, daß die Kommunen die potentiellen Produktivitätsgewinne der Infrastruktur in ihr Investitionskalkül einbeziehen. Die Untersuchung zeigt ferner, daß Städte mit ähnlicher Sektorstruktur bestrebt sind dem ortsansässigen Gewerbe eine vergleichbare Infrastrukturversorgung zu bieten, und daß Städte bei der Festsetzung ihrer Gewerbesteuerhebesätze sich sehr stark an den Hebesätzen der Städte in unmittelbarer Nachbarschaft orientieren.

Sicherlich ist das von uns präsentierte theoretische Modell in seiner Struktur zu sehr vereinfachend angelegt, so daß über andere wichtige Variablen des interstädtischen Wettbewerbs, wie z.B. Wirtschaftsförderung oder Gewerbeflächenbereitstellung, keine Aussagen gemacht werden können. Auch ist der 2-stufige empirische Ansatz in der Arbeit etwas unbefriedigend; wünschenswert wäre die Integration beider Stufen in einem gemeinsamen Schätzansatz. Bezüglich der verwendeten Schätzmethoden sollten zukünftige Forschungsarbeiten mit regional disaggregierten Daten auch Ansätze aus dem Bereich ‚spatial econometrics‘, siehe z.B. Anselin (1988), einbeziehen. Trotz dieser Unzulänglichkeiten glauben wir jedoch, daß das vorgestellte Modell sowie die vorgelegten Schätzergebnisse eine gute Grundlage für weitere Forschungsarbeiten in dieser Richtung darstellen.

Anhang

Tabelle A1

Ausgewählte Daten für die 85 Städte des Paneldatensatzes sowie durchschnittliche Infrastrukturelastizitäten der Kostenfunktion für die einzelnen Städte

Stadt	(1)	(2)	(2a)	(3)	(3a)	(4)	(4a)	(5)
München	1 212	18 345	2.50	674	0.81	167	-1.51	0.241*
Köln	937	12 594	1.29	424	0.46	99	-2.44	0.311*
Frankfurt am Main	625	11 331	3.33	471	1.28	101	-0.83	0.090*
Essen	621	7 671	0.68	218	0.00	41	-2.41	0.394*
Dortmund	587	7 449	1.52	205	-0.68	45	-3.40	0.352*
Düsseldorf	570	10 797	2.09	344	0.23	73	-2.46	0.308*
Stuttgart	563	9 012	1.98	371	0.84	125	-0.68	0.147*
Duisburg	527	6 955	0.79	186	-0.97	67	-3.67	0.421*
Hannover	498	7 141	0.77	290	-0.19	74	-3.28	0.152*
Nürnberg	480	6 193	1.42	270	0.64	85	-1.65	0.068
Bochum	389	5 261	1.24	135	1.13	48	-1.68	0.517*
Wuppertal	371	4 072	0.34	144	0.56	56	-1.36	0.263*
Bielefeld	312	4 633	1.73	132	0.79	46	-0.98	0.375*
Mannheim	300	4 630	1.53	172	-0.06	63	-1.90	0.187*
Gelsenkirchen	287	3 629	1.02	93	-1.55	23	-3.53	0.452*
Bonn	282	5 198	1.07	125	0.64	14	-1.84	0.483*
Karlsruhe	265	4 108	2.31	146	0.79	35	-1.27	0.222*
Wiesbaden	254	3 431	1.36	118	0.99	22	-2.13	0.233*
Braunschweig	254	2 563	0.96	104	-0.32	26	-2.33	0.094*
Mönchengladbach	253	2 935	0.85	89	0.78	31	-0.89	0.369*
Münster (Westf.)	249	3 412	1.21	109	1.07	11	0.31	0.248*
Augsburg	248	2 705	1.41	131	0.98	48	-1.04	-0.037
Kiel	241	2 120	0.82	108	0.42	22	-2.30	-0.089
Krefeld	235	2 683	1.07	96	0.35	41	-1.42	0.213*
Aachen	233	3 110	1.40	99	1.32	22	-1.46	0.263*
Oberhausen	221	2 128	0.38	63	-1.77	19	-3.68	0.285*
Lübeck	211	2 118	1.27	79	-0.28	20	-2.89	0.107*
Hagen	211	2 532	0.90	78	0.14	28	-1.73	0.315*
Saarbrücken	188	1 952	1.21	150	-0.43	37	-2.25	-0.448
Kassel	189	2 078	1.55	95	-0.34	24	-1.60	-0.038
Freiburg/Breisgau	184	2 299	1.45	85	0.68	15	-0.28	0.132*
Mülheim a.d.Ruhr	175	2 196	0.18	62	-0.52	21	-2.05	0.346*
Herne	175	1 849	0.38	50	-0.65	11	-1.88	0.369*
Mainz	175	2 162	0.72	96	1.23	24	0.14	0.063
Hamm	174	1 617	0.86	52	-0.21	11	-2.51	0.224*

(Tabelle A1 – Fortsetzung)

Stadt	(1)	(2)	(2a)	(3)	(3a)	(4)	(4a)	(5)
Solingen	161	1863	0.56	55	0.41	22	-2.31	0.336*
Ludwigshafen	158	2443	0.88	102	0.68	64	0.04	0.147*
Leverkusen	157	1969	1.33	76	0.00	50	-0.82	0.115*
Osnabrück	155	1663	0.68	75	0.61	19	-1.97	0.014
Oldenburg	141	1461	1.18	57	0.60	8	-3.62	0.086*
Darmstadt	136	1746	1.60	92	1.15	31	0.34	-0.116*
Heidelberg	131	1693	2.06	70	1.36	16	0.35	0.030
Würzburg	126	1618	2.38	76	0.45	13	-1.28	-0.134
Remscheid	121	1348	0.86	53	1.83	27	-0.39	0.153*
Regensburg	119	1831	3.70	75	1.26	21	1.23	0.040*
Bottrop	116	956	0.35	28	0.40	4	-1.36	0.301*
Offenbach am Main	112	1333	0.39	51	-0.22	20	-0.41	0.126*
Heilbronn	112	1849	2.73	62	1.34	19	-1.34	0.187*
Pforzheim	109	1520	2.00	61	1.16	24	-0.75	0.075
Koblenz	107	1646	2.00	61	1.36	10	1.13	0.094*
Ulm	107	1870	3.25	74	0.62	28	-2.18	0.035**
Erlangen	101	1207	1.75	68	1.59	40	1.24	-0.147
Fürth	99	1115	1.15	44	1.06	17	-1.91	0.101*
Ingolstadt	98	1492	2.57	61	2.23	34	1.43	0.089**
Kaiserslautern	97	1119	1.21	53	0.00	17	-0.21	-0.076
Trier	96	1204	1.36	47	0.24	10	-1.45	0.097*
Wilhelmshaven	90	1026	0.72	32	0.35	5	-1.86	0.275*
Flensburg	86	682	0.92	39	0.59	9	-0.49	-0.206
Neumünster	80	981	1.52	32	1.49	10	-0.50	0.240*
Worms	75	897	0.73	27	0.42	10	0.15	0.328*
Delmenhorst	73	737	1.41	19	0.00	5	-3.38	0.325*
Bayreuth	71	905	1.36	37	1.28	8	-1.50	0.083*
Bamberg	70	1187	4.23	46	2.15	17	0.79	0.043*
Aschaffenburg	62	892	3.91	38	0.60	12	-0.95	-0.107
Kempten (Allgäu)	60	688	2.12	30	0.77	7	0.21	-0.075
Landshut	57	1043	5.51	31	1.97	8	-1.54	0.132*
Rosenheim	54	773	3.44	28	1.27	6	-0.86	0.059*
Schweinfurt	53	980	2.58	49	1.20	29	0.12	-0.123*
Hof	51	614	2.11	23	0.00	8	-1.19	0.031**
Baden-Baden	51	860	3.68	26	2.40	4	0.55	0.160*
Neustadt/Weinstr.	50	634	0.70	16	-0.67	2	-3.36	0.348*
Passau	49	835	3.85	30	2.51	9	-0.14	0.069*
Pirmasens	47	570	1.82	26	0.00	10	-1.33	-0.130
Frankenthal/Pfalz	45	603	1.85	17	-0.63	9	-2.14	0.251*

(Tabelle A1 – Fortsetzung)

Stadt	(1)	(2)	(2a)	(3)	(3a)	(4)	(4a)	(5)
Speyer	45	472	2.32	21	0.54	6	-2.69	-0.060
Coburg	43	553	3.86	26	1.87	8	0.79	-0.173
Amberg	42	529	2.75	24	2.63	8	0.03	-0.093
Weiden i. d. Opf.	42	741	4.11	21	1.73	5	-0.67	0.202*
Straubing	41	477	2.77	20	1.18	6	1.55	-0.078
Kaufbeuren	39	521	5.36	18	2.05	4	0.08	-0.034
Memmingen	38	623	2.79	21	-0.52	8	-0.30	0.062*
Ansbach	37	448	1.85	21	2.38	6	2.46	-0.090
Landau/Pfalz	36	452	2.80	16	-1.30	3	-0.51	-0.042
Schwabach	34	340	1.32	13	0.89	5	-1.44	0.034
Zweibrücken	33	408	1.83	13	1.87	5	-2.26	0.197*

Variablenbezeichnungen:

- (1) Bevölkerung 1989 in 1000.
- (2) Realer Nettoinfrastrukturkapitalstock Ende 1989 zu Preisen von 1980; Mrd. DM.
- (2a) Jahresdurchschnittliche Wachstumsrate des Infrastrukturkapitalstocks 1980 - 1989 in %.
- (3) Gesamte Beschäftigung 1989 in 1000.
- (3a) Jahresdurchschnittliches Beschäftigungswachstum 1980 - 1989 in %.
- (4) Beschäftigung im Verarbeitenden Gewerbe 1989 in 1000.
- (4a) Jahresdurchschnittliches Beschäftigungswachstums im Verarbeitenden Gewerbe 1980 - 1989 in %.
- (5) Geschätzte durchschnittliche Infrastrukturelastizität der Kostenfunktion. * (**) = T-Werte der geschätzten Elastizitäten größer als 2.00 (1.75).

Literatur

- Anselin, L., 1988, Spatial Econometrics, Methods and Models, (Kluwer Academic Publisher, Amsterdam).*
- Bennet, R./Krebs, G., 1988, Local Business Taxes in Britain and Germany, Schriften zur öffentlichen Verwaltung und öffentlichen Wirtschaft, Band 113, Nomos-Verlagsgesellschaft, Baden-Baden.*
- Berndt, E./Bengt, H., 1991, Measuring the Contribution of Public Infrastructure, Capital in Sweden', NBER Working Paper, # 3842.*
- Conrad, K., 1993, Taxes und Subsidies for Pollution Intensive Industries as Trade Policy, erscheint in: Journal of Environmental Economics und Management.*
- Conrad, K./Seitz, H., 1992, The Public Capital Hypothesis: The Case of Germany, Recherches Economiques de Louvain, Band 58, 309 - 328.*
- Costa J./Ellson, R./Martin, R., 1987, Public Capital, Regional Output and Development: Some Empirical Evidence, Journal of Regional Science, Band 27, 419 - 435.*
- Deitmer, I., 1993, Effekte der Regionalen Strukturpolitik auf Investitionen, Beschäftigung und Wachstum (Diss. Universität Münster).*

- Duffy-Deno*, K. T., 1991, Public Capital and the Factor Intensity of the Manufacturing Sector, *Urban Studies*, Band 28, 3 - 14.
- Eberts*, R. W./*Fox*, W. F., 1992, The Effect of Federal Policies on Local Public Infrastructure Investment, *Public Finance Quarterly*, Band 20, 557 - 571.
- Henderson*, V., 1986, Efficiency of Resource Usage and City Size, *Journal of Urban Economics*, Band 19, 47 - 70.
- 1988, *Urban Development: Theory, Facts und Illusion*, Oxford University Press.
- Hsiao*, C., 1986, *Analysis of panel data*, Cambridge University Press, Cambrigde USA.
- Hunds*, W. D./*Mann*, B. D., 1987, Urban Industrial Tax Subsidies, *Regional Science und Urban Economics*, Band 17, 179 - 190.
- Morrison*, C./*Schwartz*, A., 1992, State Infrastructure and Productive Performance, NBER Working Paper # 3981.
- Nadiri*, Ishaq/*Mamuneas*, Theofanis, 1991, The Effect of Public Infrastructure and R&D Capital on the Cost Structure and Performance of U.S. Manufacturing Industries, NBER Working Paper # 3887.
- Nickel*, S., 1981, Biases in Dynamic Models with Fixed Effects, *Econometrica*, Band 49, 1417 - 1426.
- Oaklund*, W., 1987, Theory of Public Goods, in: A. J. Auerbach und M. Feldstein (eds.), *Handbook of Public Economics*, Band II, 485 - 535, Elsevier Science Publishers, North-Hollund.
- Oates*, W. E./*Schwab*, R. M., 1988, Economic Competition Among Jurisdictions: Efficiency Enhancing or Distortion Inducing?, *Journal of Public Economics*, Band 35, 333 - 354.
- Ramser*, H. J., 1968, Dezentrale Planung der Infrastrukturausgaben, *Finanzarchiv*, Band 27, 484 - 503.
- Seitz*, H., 1993a, Public Capital and the Demund for Privat Inputs, *Journal of Public Economics*, Band 54, 287–307.
- 1993b, Stadt und Industrie: Die Performance der deutschen Städte, Manuscript Universität Mannheim.
- Seitz*, H./*Licht*, G., 1993, The Impact of the Provision of Public Infrastructures on Regional Economic Development in Germany, ZEW-Discussion Paper No. 93 - 14, erscheint in: *Regional Studies*.
- Seitz*, H., 1994, Infrastructure Investment, Employment and Private Capital Formation, *Arbeitspapier für die 'Task Force for the Employment/Unemployment Study' der OECD*.
- Spencer*, B./*Brunder*, J., 1983, International R&D Rivalry and Industrial Strategy, *The Review of Economic Studies*, Band 4, 707 - 722.
- Wildasin*, D. E., 1988, Nash Equilibria in Models of Fiscal Competition, *Journal of Public Economics*, Band 35, 229 - 240.

— 1991, Some rudimentary „duoply“ theory, *Regional Science and Urban Economics*, Band 21, 393 - 421.

Wildasin, D./Wilson, J., 1991, Theoretical Issues in local public economics, *Regional Science and Urban Economics*, Band 21, 317 - 332.

Zodrow, G. R./Mieszkowski, P., 1986, Pigou, Tiebout, Property Taxation, and the Underprovision of Local Public Goods, *Journal of Urban Economics*, Band 19, 356 - 370.

Zusammenfassung der Diskussion

In der Diskussion wird festgestellt, daß die Datenqualität, wie bei vielen anderen Untersuchungen mit Regionaldaten auch, nicht unproblematisch ist, da diese im wesentlichen aus Länderdaten hochgerechnet werden.

Außerdem unterliegen Infrastrukturmaßnahmen einer relativ langen Wirkungsverzögerung (ca. 5 Jahre), welche mit dem z.Zt. verfügbaren Datenmaterial schwierig zu identifizieren ist. Letzteres soll in näherer Zukunft durch eine verbesserte Datenbasis mittels Paneldaten auf EG-Ebene besser berücksichtigt werden können.

Es wird auch angemerkt, daß solche Fälle auftreten können, in denen der Kapitalzins vom Grenzprodukt der öffentlichen Infrastruktur überschritten wird.

Als zu beachtende Varianten werden jene Szenarien angesprochen, in welchen entweder die Kommunalverwaltungen selbst die betrachteten Unternehmen betreiben, oder sich ihrer eigenen strategischen Handlungsmacht nicht bewußt sind. Im letzteren Fall kommt es zu einer Optimalitätsbedingung, welche die Gleichheit des (Netto-)Preises der städtischen Infrastruktur mit der durch ihre Bereitstellung erfolgten Senkung der Kosten im nichtöffentlichen Bereich verlangt.

Regionale Unterschiede bei Schließungen von Industriebetrieben

**Eine Untersuchung für Raumordnungsregionen
in Niedersachsen (1979 - 1992)***

Von *Joachim Wagner*, Lüneburg

Abstract

In dieser Untersuchung werden regionale Unterschiede bei Schließungen von Industriebetrieben und die damit verbundenen Arbeitsplatzverluste für 16 niedersächsische Raumordnungsregionen in den Jahren von 1979 bis 1992 ermittelt. Im Durchschnitt schließen etwa zwischen vier und fünf Prozent aller Industriebetriebe pro Jahr, wodurch zwischen einem und zwei Prozent aller Arbeitsplätze verloren gehen. Hierbei gibt es ausgeprägte Unterschiede zwischen den Regionen und zahlreiche extreme Werte in einzelnen Regionen und Jahren. Eine ökonometrische Untersuchung der intertemporalen und interregionalen Unterschiede der Schließungsintensitäten unter Einschluß fixer Regionaleffekte und bei Verwendung eines robusten Schätzverfahrens zeigt, daß Schließungen besonders bedeutend sind in Regionen mit einer hohen Gründungsintensität, während sich kein Zusammenhang mit dem Anteil der Beschäftigten in Kleinbetrieben, dem regionalen Lohnniveau oder der Höhe der Gewerbesteuersätze erkennen läßt.

1. Einleitung

Gründungen und Schließungen von Firmen sind die neben dem Wachsen und Schrumpfen von Betrieben zentralen Komponenten der Arbeitsplatzdynamik auf sektoraler und regionaler Ebene. Empirische Analysen zu Bedeutung, Ursachen und Folgen dieser Komponenten setzen Mikrodaten auf Firmenebene aus einer Reihe von Jahren voraus. Solche umfangreichen Firmenpanel-Daten sind, obwohl sie aus Erhebungen der amtlichen Statistik seit langer Zeit und für viele Länder vorliegen, für die wissenschaftliche

* Die Berechnungen auf der Basis von Betriebsdaten aus Erhebungen der amtlichen Statistik wurden durch ein Gemeinschaftsprojekt mit dem Niedersächsischen Landesamt für Statistik (NLS) ermöglicht und im NLS durchgeführt. Geheimhaltungsvorschriften werden durch die hier veröffentlichten Angaben nicht verletzt. Ich danke Uwe Rode vom NLS für die Erstellung der Längsschnittdatei mit Angaben zu sämtlichen niedersächsischen Industriebetrieben.

Forschung nur in Ausnahmefällen und erst seit kurzem verfügbar (vgl. für die Bundesrepublik Deutschland *Gerlach/Wagner* 1992). Auswertungen dieser Datensätze haben sich bisher vor allem darauf konzentriert, die relative Bedeutung der einzelnen Komponenten der Arbeitsplatzdynamik zu quantifizieren (vgl. z.B. *Gerlach/Wagner* 1993a) und die Ursachen interindustrieller bzw. interregionaler Unterschiede von Betriebsgründungen sowie deren längerfristige Beschäftigungseffekte zu untersuchen (vgl. z.B. *Wagner* 1992a, *Wagner* 1992b, *Gerlach/Wagner* 1993b).

Entsprechende Studien über Betriebsschließungen mit deutschen Daten sind ausgesprochen selten: *Posselt* (1993) untersucht Marktaustritte für eine Stichprobe von 83 *Industriezweigen* und die Jahre 1983 bis 1985 auf der Basis von Unternehmensdaten, die die Monopolkommission publiziert hat. *Audretsch/Fritsch* (1992, 1993) analysieren mit zu Betriebsnummern zusammengefaßten Angaben aus der Statistik der sozialversicherungspflichtig Beschäftigten die Stillegungsraten (definiert als Anzahl geschlossener Betriebe in Prozent bestehender Betriebe) in den *Raumordnungsregionen* der Bundesrepublik Deutschland 1986 bzw. 1986 - 1989, wobei sie allerdings die hiermit verbundenen Beschäftigungseffekte nicht betrachten.

Angesichts der einerseits anhaltend hohen und andererseits regional sowohl in ihrem Niveau als auch in ihrer Dynamik ausgeprägt unterschiedlichen Arbeitslosigkeit stellt dies Fehlen von Informationen über die Bedeutung von Betriebsschließungen für die regionale Arbeitsplatzdynamik und die Ursachen für Unterschiede zwischen den Regionen hierbei eine Forschungslücke dar, die es zu schließen gilt. Die vorliegende Untersuchung möchte hierzu einen Beitrag leisten, indem regionale Unterschiede bei Schließungen von Industriebetrieben und ihre Arbeitsplatzeffekte für die Raumordnungsregionen Niedersachsens zwischen 1979 und 1992 ermittelt und ihren Ursachen in einer ökonometrischen Untersuchung nachgespürt wird. In Abschnitt 2 wird die Datenbasis vorgestellt, Abschnitt 3 präsentiert die empirischen Befunde zur Bedeutung von Betriebsschließungen in niedersächsischen Raumordnungsregionen, Abschnitt 4 diskutiert den Ansatz und Abschnitt 5 die Ergebnisse einer ökonometrischen Untersuchung zu den Determinanten interregionaler und intertemporaler Unterschiede bei Betriebsschließungen, in Abschnitt 6 wird ein Fazit gezogen.

2. Datenbasis

Die Grundlage für die Identifikation von Betriebsschließungen und für die Messung der mit ihnen verbundenen Arbeitsplatzverluste bildet eine Datei sämtlicher Industriebetriebe, die in mindestens einem Jahr zwischen 1978 und 1992 im Bereich Bergbau und Verarbeitendes Gewerbe in Niedersachsen existierten. Diese Datei beruht auf Angaben aus der amtlichen Sta-

tistik, wobei Daten aus zwei Erhebungen, dem *Monatsbericht für Betriebe im Bergbau und Verarbeitenden Gewerbe* (zu dem alle Industriebetriebe von Unternehmen mit in der Regel 20 oder mehr tätigen Personen meldepflichtig sind) und der jährlichen *Erhebung für industrielle Kleinbetriebe im Bergbau und Verarbeitenden Gewerbe* (die die nicht im Monatsbericht erfaßten Industriebetriebe umfaßt) kombiniert wurden, um eine Totalerhebungsdatei zu erstellen. Die Verknüpfung der Angaben zwischen den beiden Erhebungen und über die Zeit ist möglich, da jedem Betrieb vom Statistischen Landesamt eine unveränderliche Betriebsnummer zugeordnet wird, sobald er zum ersten Mal in einem der beiden genannten Berichtskreise auftaucht. Für jeden der 11 272 niedersächsischen Industriebetriebe ist in der Datei neben der Zuordnung zur einem Industriezweig und dem Ort seines Sitzes (Kreis bzw. kreisfreie Stadt) der Jahresumsatz und die Anzahl tätiger Personen für jedes Jahr zwischen 1978 und 1992 verzeichnet, wobei die Personenzahl entweder als Durchschnitt aus den Angaben in den Monatsmeldungen im Berichtsjahr gebildet wird oder (bei den Kleinbetrieben) auf den Angaben für Ende September beruht.

Weist ein Betrieb im Jahr t tätige Personen auf und ist dies im Jahr $t+1$ nicht mehr der Fall, so kann es sich um eine Betriebsschließung handeln – dies muß aber nicht der Fall sein: Der Betrieb kann auch seine Produktionsaktivität vorübergehend eingestellt haben, er kann (zeitweilig oder dauerhaft) seinen Tätigkeitsenschwerpunkt aus dem Bergbau und Verarbeitenden Gewerbe heraus (z.B. in den Dienstleistungssektor) verlegt haben, er kann schließlich seinen Sitz (zeitweilig oder auf Dauer) in ein anderes Bundesland oder ins Ausland verlegt haben. Alle genannten Gründe bewirken, daß der Betrieb im Jahr $t+1$ nicht mehr im Berichtskreis „Niedersächsische Industriebetriebe“ enthalten ist. Was für das Verlassen des Berichtskreises ausschlaggebend war, kann ohne Rückgriff auf die Akten im Statistischen Landesamt und eine Einzelfallprüfung nur für den Fall einer temporären Unterbrechung oder Verlagerung ermittelt werden, da in solchen Fällen in den betreffenden Betrieben tätige Personen in der Datei erneut für Jahre nach $t+1$ enthalten sind.

Aus Datenschutzgründen (und wegen fehlender finanzieller Mittel) war eine Einzelfallprüfung der Ursachen, die bei Betrieben jeweils zum Verlassen des Berichtskreises geführt haben, im Rahmen dieser Untersuchung nicht möglich. Es müssen daher bewußt Unschärfen bei der Abgrenzung von Betriebsschließungen und (tätigkeitsbezogenen oder räumlichen) Verlagerungen in Kauf genommen werden. Diese konnten nur dadurch reduziert werden, daß jeweils bei Betrieben, die im Jahr t aber nicht im Jahr $t+1$ tätige Personen aufweisen, geprüft wurde, ob sie in den Jahren $t+2$ und $t+3$ ebenfalls keine tätigen Personen gemeldet haben – nur dann wurden sie als „Schließung“ erfaßt und ansonsten als „Unterbrecher“ behandelt. Da

9*

die Datenbasis Angaben bis zum Jahr 1992 einschließlich umfaßt, war dies Korrektur- bzw. Prüfverfahren nur bis zum Jahr 1990 möglich; die Daten für 1991 konnten nur anhand der Angaben für 1992 überprüft werden, für 1992 war keine Prüfung durchführbar.

Da im Zentrum dieser Arbeit die interregionalen Unterschiede bei Betriebsschließungen stehen und da es keine Hinweise darauf gibt, daß die genannten unvermeidbaren Unschärfen ungleich auf die unten betrachteten Regionen verteilt sind, ist jedoch zu erwarten, daß hierdurch lediglich die Niveauwerte (in unbekanntem Ausmaß) und nicht auch die regionale Struktur der Schließungen verzerrt werden.

3. Schließungen von Industriebetrieben in niedersächsischen Raumordnungsregionen: Der empirische Befund für 1979 bis 1992

Mit den im vorigen Abschnitt beschriebenen Daten und Methoden wurden alle Schließungen niedersächsischer Industriebetriebe zwischen 1979 und 1992 ermittelt und regional zugeordnet. Die hierbei verwendete Regionalgliederung ist die nach *Raumordnungsregionen*; dies ist die kleinste räumliche Gliederung, die ohne die Gefahr einer Verletzung von Datenschutzvorschriften herangezogen werden konnte. Eine Definition der sechzehn niedersächsischen Raumordnungsregionen enthält die *Tabelle 1*.

Für einen interregionalen Vergleich ist es nicht sinnvoll, lediglich die Anzahl der in einem Jahr geschlossenen Industriebetriebe zu betrachten, da die Regionen unterschiedlich groß sind und unterschiedlich viele Betriebe aufweisen. Daher wird einerseits die Anzahl geschlossener Betriebe zur Anzahl der in einer Region bestehenden Betriebe in Beziehung gesetzt (*Betriebsschließungsrate*, definiert als Anzahl geschlossener Betriebe in Prozent aller bestehender Betriebe im Jahr t). Hierbei werden alle Betriebe unabhängig von ihrer Größe gleich gewichtet. Dies Vorgehen ist unbefriedigend, wenn sich geschlossene Betriebe hinsichtlich der Zahl der in ihnen durchschnittlich (vor ihrer Schließung) tätigen Personen von den weiter bestehenden Betrieben unterscheiden. Daher wird als weitere Kennziffer eine *Betriebsschließungsintensität*, definiert als Anzahl tätiger Personen in geschlossenen Betrieben im Jahr vor der Schließung in Prozent aller tätigen Personen im Jahr t , ausgewiesen. Angaben zu Betriebsschließungsralten und -intensitäten für die sechzehn niedersächsischen Raumordnungsregionen in den Jahren von 1979 bis 1992 enthalten *Tabelle 2* und *Tabelle 3*.

Während der Prozentsatz geschlossener Betriebe im Durchschnitt des Untersuchungszeitraums zwischen den Regionen nur geringfügig zwischen rund 4 und 5 variiert, gibt es in einzelnen Jahren erhebliche interregionale

Tabelle 1
Definition der Raumordnungsregionen

Nr.	Name	Kreise [Kreisnummer der amtlichen Statistik]
1	Unterelbe	Harburg [353], Stade [359]
2	Lüneburg	Lüchow-Dannenberg [354], Lüneburg [355], Uelzen [360]
3	Cuxhaven	Cuxhaven [352]
4	Wilhelmshaven	Wilhelmshaven Stadt [405], Friesland [455], Wittmund [462]
5	Ostfriesland	Emden Stadt [402], Aurich [452], Leer [457]
6	Oldenburg	Oldenburg Stadt [403], Ammerland [451], Cloppenburg [453], Wesermarsch [461]
7	Emsland	Emsland [454], Grafschaft Bentheim [456]
8	Osnabrück	Osnabrück Stadt [404], Osnabrück [459], Vechta [460]
9	Bremen Umland	Diepholz [251], Osterholz [356], Rothenburg (Wümme) [357], Verden [361], Delmenhorst Stadt [401], Oldenburg [458]
10	Südheide	Celle [351], Soltau-Fallingbostel [358]
11	Großraum Hannover	Hannover Stadt [201], Hannover Land [253]
12	Weser	Nienburg (Weser) [256], Schaumburg [257]
13	Hildesheim	Hameln-Pyrmont [262], Hildesheim [254], Holzminden [255]
14	Wolfsburg	Wolfsburg [103], Gifhorn [151], Helmstedt [154]
15	Braunschweig	Braunschweig Stadt [101], Salzgitter Stadt [102], Peine [157], Wolfenbüttel [158], Goslar [153]
16	Göttingen	Göttingen [152], Northeim [155], Osterode [156]

Unterschiede bei den Schließungsraten. Für die Schließungsintensitäten ist eine solche ausgeprägte Variation zwischen den Raumordnungsregionen auch für die Durchschnittswerte über alle betrachteten Jahre festzustellen (die vom Großbetrieb Volkswagenwerk dominierte Region 14 – Wolfsburg – hat den sehr niedrigen Wert 0.20, während der Wert in der Region 9 – Bremen Umland – bei 2.27 Prozent lag). Im (ungewichteten) Durchschnitt gingen in den niedersächsischen Raumordnungsregionen zwischen 1979 und 1992 pro Jahr rund 1.4% aller Industriearbeitsplätze durch die Schließung von Betrieben verloren.

Tabelle 2: Betriebsschließungsraten in der niedersächsischen Industrie nach Raumordnungsregionen (1978 - 1992)*

Region	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	Mittelwert	Stand.-abw.
1	4.68	2.43	4.91	2.10	6.09	6.72	4.46	6.74	4.60	8.50	5.71	5.37	3.77	7.59	5.26	1.82
2	6.05	3.15	6.34	4.21	9.16	5.45	3.26	7.33	3.25	4.76	3.62	3.23	2.46	5.57	4.85	1.91
3	5.56	2.42	5.60	5.69	3.25	7.63	7.14	3.64	2.73	1.82	2.65	2.65	3.57	3.45	4.13	1.85
4	7.30	5.29	6.40	6.79	4.29	5.03	3.73	5.10	5.30	4.70	5.92	3.29	7.43	9.79	5.74	1.71
5	8.44	2.23	5.41	4.19	1.83	5.02	8.61	6.40	6.28	5.24	5.79	3.66	4.21	7.37	5.33	2.04
6	5.82	2.70	3.98	4.79	6.68	4.83	5.71	6.62	6.54	4.45	5.43	4.18	2.36	6.28	5.03	1.39
7	5.80	4.90	3.69	4.43	4.74	4.96	5.81	6.02	2.86	2.51	4.31	5.99	4.24	3.91	4.58	1.11
8	6.60	3.46	5.20	6.43	5.04	5.24	3.46	4.57	4.92	4.35	4.25	2.83	3.81	3.63	4.56	1.10
9	6.12	2.05	3.91	3.81	6.98	5.64	6.41	7.04	5.13	4.92	4.08	4.39	4.13	4.86	4.96	1.39
10	6.21	4.45	5.23	6.38	6.12	3.25	3.57	3.93	5.28	3.83	3.86	5.04	4.80	6.85	4.91	1.15
11	6.00	3.85	5.28	5.94	6.47	4.74	5.38	4.58	3.56	5.07	4.34	3.32	4.77	5.40	4.91	0.93
12	4.27	3.13	5.01	7.19	7.80	5.37	4.15	5.62	4.56	5.21	0.87	3.24	4.42	3.51	4.45	1.57
13	5.18	4.64	2.88	4.57	6.57	5.56	6.09	4.95	4.46	3.22	5.50	4.05	4.49	6.79	4.93	1.14
14	8.41	1.40	6.19	4.35	5.45	3.50	7.33	4.74	6.56	2.75	2.09	3.65	2.04	2.51	4.35	2.17
15	5.63	4.63	4.56	5.59	4.58	5.81	6.97	5.66	4.62	3.41	3.79	2.99	6.46	5.77	5.03	1.15
16	6.61	4.88	4.61	4.23	4.19	5.15	4.66	4.37	5.57	3.76	4.16	5.19	3.74	4.69	4.70	0.76
Mittelwert	6.17	3.48	4.95	4.92	5.58	5.24	5.42	5.46	4.76	4.28	4.15	3.94	4.17	5.50		
Stand.-abw.	1.15	1.21	0.99	1.20	1.80	1.04	1.61	1.14	1.20	1.53	1.40	1.00	1.38	1.91		

*) Anzahl geschlossener Betriebe im Jahr t in Prozent aller bestehender Betriebe im Jahr t
 (zur Definition der Regionen vgl. Tabelle 1)

Tabelle 3: Betriebsschließungsintensitäten in der niedersächsischen Industrie nach Raumordnungsregionen (1979 - 1992)*

Region	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	Mittelwert	Stand.-abw.
1	0.98	0.55	1.65	1.03	1.49	1.08	3.04	2.86	0.89	2.18	1.77	2.10	1.48	1.18	1.59	0.74
2	2.30	0.48	1.26	1.92	3.31	2.57	0.61	1.64	0.51	1.79	0.62	0.70	0.83	1.13	1.41	0.88
3	1.40	0.63	1.11	1.11	1.96	1.84	6.94	2.14	0.30	0.22	2.61	2.64	0.70	1.60	2.05	1.82
4	2.63	1.62	1.72	0.77	0.45	1.38	1.46	0.72	0.71	3.50	0.88	3.72	0.84	2.33	1.62	1.06
5	2.13	0.20	0.56	0.86	1.03	2.38	1.80	1.51	1.07	0.63	1.30	0.70	1.46	0.71	1.17	0.63
6	0.89	0.69	1.17	2.37	2.16	2.86	1.62	4.88	1.94	0.70	2.54	0.85	0.85	1.41	1.78	1.15
7	0.88	0.82	0.72	0.93	1.60	1.72	2.09	2.33	0.79	0.46	1.20	1.04	2.08	0.45	1.22	0.63
8	1.77	0.41	1.23	2.42	1.31	1.36	4.37	0.89	2.39	1.28	1.72	0.96	0.87	0.49	1.53	1.02
9	3.25	0.93	1.42	1.23	4.08	2.33	3.62	3.61	2.93	2.73	1.15	2.22	1.37	0.83	2.27	1.12
10	1.58	0.76	0.83	1.10	1.29	0.45	0.82	1.25	1.07	0.44	0.64	1.78	0.98	1.13	1.01	0.39
11	0.64	0.92	0.91	1.21	1.26	0.72	0.80	0.59	3.47	1.05	0.84	0.78	1.25	2.47	1.21	0.80
12	1.79	0.83	0.93	1.46	3.42	2.01	1.33	2.11	1.68	1.03	0.17	0.45	1.92	0.43	1.40	0.86
13	0.66	1.44	0.37	1.51	2.20	1.96	2.12	0.80	2.05	0.78	1.07	0.72	0.80	0.82	1.31	0.58
14	0.26	0.03	0.21	0.12	0.17	0.48	0.26	0.28	0.06	0.10	0.38	0.11	0.14	0.20	0.12	
15	2.68	1.62	0.51	0.74	0.69	1.22	1.96	0.49	0.79	1.06	1.79	0.78	0.88	0.86	1.15	0.64
16	1.91	1.09	2.06	0.63	0.89	1.49	1.70	1.03	1.53	1.06	0.90	0.63	1.29	0.82	1.22	0.46
Mittelwert	1.61	0.81	1.10	1.43	1.71	1.60	2.17	1.69	1.40	1.19	1.21	1.10	1.29	1.05		
Stand.-abw.	0.85	0.46	0.49	1.06	1.11	0.76	1.66	1.26	0.95	0.94	0.73	0.69	0.81	0.65		

*) Anzahl tätige Personen in geschlossenen Betrieben im Jahr vor der Schließung in Prozent aller tätigen Personen im Jahr *t*
 (zur Definition der Regionen vgl. Tabelle 1)

Wie man aus *Tabelle 4* ersehen kann, sind die beiden Indikatoren für die Bedeutung von Betriebsschließungen nur in wenigen Jahren stark positiv korreliert; der Korrelationskoeffizient weist in einigen Jahren Werte von nahe bei Null auf und wird 1981 sogar negativ.

Tabelle 4
Korrelation von Schließungsraten und Schließungsintensitäten

Jahr	Korrelation	Jahr	Korrelation
1979	+0.08	1986	+0.57
1980	+0.77	1987	+0.05
1981	-0.13	1988	+0.55
1982	+0.23	1989	+0.38
1983	+0.65	1990	+0.18
1984	+0.41	1991	+0.68
1985	+0.20	1992	+0.55

Die in *Tabelle 5* abgedruckten Werte der Korrelation von Schließungsraten bzw. Schließungsintensitäten zwischen den Jahren des Analysezeitraums belegen, daß es auch im intertemporalen Vergleich ausgeprägte Unterschiede bei der relativen Bedeutung von Betriebsschließungen zwischen den Regionen gab. Raten bzw. Intensitäten sind für aufeinander folgende Jahre in vielen Fällen nur schwach positiv und oft sogar negativ korreliert.

Damit zeigt sich, daß die Bedeutung von Betriebsschließungen sowohl zwischen Raumordnungsregionen als auch über die Zeit stark variiert. Den Ursachen dieser Unterschiede soll im folgenden nachgespürt werden. Ange- sichts der anhaltend hohen Arbeitslosigkeit konzentriert sich die Untersuchung dabei auf die Erklärung der Unterschiede bei Betriebsschließungsintensitäten, d.h. auf die interregionale und intertemporale Variation der relativen Bedeutung von Arbeitsplatzverlusten durch geschlossene Industrie- betriebe.

4. Untersuchungsansatz: Das empirische Modell

Die ökonometrische Untersuchung der Ursachen interregionaler und intertemporaler Unterschiede des Ausmaßes von Arbeitsplatzverlusten durch Betriebsschließungen stellt keinen empirischen Test von Folgerungen aus einem formalen theoretischen Modell dar – ein entsprechend umsetzbares Modell fehlt bisher. Ziel ist es vielmehr, aufgrund theoretischer Über-

Tabelle 5: Intertemporale Korrelation von Schließungsräten und Schließungsintensitäten

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
1980 Rate	-.19												
Intensität	+.34												
1981 Rate	+.46	-.23											
Intensität	+.27	+.37											
1982 Rate	+.14	+.46	+.28										
Intensität	-.04	-.15	+.14										
1983 Rate	-.51	-.08	-.09	-.16									
Intensität	+.29	-.03	+.21	+.36									
1984 Rate	-.51	-.14	-.13	-.27	-.13								
Intensität	+.35	-.02	+.26	+.39	+.66								
1985 Rate	+.35	-.41	-.21	-.25	-.54	+.14							
Intensität	+.13	-.06	+.22	+.75	+.18	+.20							
1986 Rate	-.12	-.31	-.16	-.63	-.38	+.15	+.04						
Intensität	+.05	-.18	+.21	+.34	+.59	+.63	+.26						
1987 Rate	+.54	-.26	+.04	-.08	-.18	-.52	+.19	+.07					
Intensität	+.01	+.16	+.22	-.01	+.34	+.12	-.01	+.16					
1988 Rate	-.22	-.22	+.10	-.45	+.28	+.14	-.33	+.53	+.18				
Intensität	+.61	+.41	+.62	-.19	+.21	+.16	+.00	+.12	+.17				
1989 Rate	+.18	+.27	-.25	-.15	-.28	+.05	+.02	+.27	+.21	+.36			
Intensität	+.00	+.05	+.15	+.64	-.01	+.36	+.70	+.49	-.01	-.07			
1990 Rate	-.13	-.21	-.48	-.52	+.02	-.22	-.10	+.19	+.07	+.17	+.30		
Intensität	+.15	-.06	+.23	+.51	+.27	+.00	+.71	+.42	-.04	+.15	+.50		
1991 Rate	-.07	+.67	-.00	+.49	-.29	+.06	-.11	-.20	-.11	+.05	+.27	-.12	
Intensität	+.37	+.47	+.38	-.25	-.06	+.10	-.11	+.02	-.03	+.66	-.15	-.05	
1992 Rate	+.07	+.36	+.04	+.03	-.10	-.01	-.18	+.23	+.22	+.47	+.78	+.13	.55
Intensität	+.05	+.38	+.35	+.23	-.12	-.05	+.02	+.00	+.20	+.40	+.23	+.21	+.39

legungen ausgewählte regionale Charakteristika darauf zu überprüfen, ob sie mit hohen oder niedrigen Schließungsintensitäten verbunden sind und welche relative Bedeutung sie hierbei haben. Ein solches Vorgehen einer theoriegeleiteten ökonometrischen Analyse ist in der empirischen Industrieökonomie weit verbreitet (vgl. Schmalensee 1989 für eine Diskussion dieses Ansatzes); es findet sich z.B. explizit auch bei Untersuchungen zu interindustriellen bzw. interregionalen Unterschieden von Betriebsgründungen (vgl. Wagner 1992a sowie Gerlach/Wagner 1993b).

In die ökonometrischen Modelle werden folgende erklärenden Variablen einbezogen (zu Einzelheiten vgl. den Anhang):

Zahlreiche Studien kommen zu dem Ergebnis, daß die Wahrscheinlichkeit einer Betriebsschließung mit steigender Firmengröße abnimmt (vgl. Audretsch/Fritsch [1992, 4]). Demnach ist zu erwarten, daß *ceteris paribus* die Schließungsraten in den Regionen hoch sind, in denen der *Anteil der in Kleinbetrieben tätigen Personen* groß ist. Die *ceteris paribus* Klausel ist hierbei besonders hervorzuheben, denn es schließen sicherlich nicht nur kleine Betriebe, und daher kann z.B. in Regionen in einigen Jahren das Schließungsgeschehen durch den Marktaustritt einer großen Firma oder einiger Mittelbetriebe dominierend geprägt sein.

Es ist ferner aus zahlreichen Untersuchungen bekannt, daß das Sterblichkeitsrisiko gerade in den ersten Jahren der Existenz eines Betriebes sehr hoch ist (vgl. zu dem hohen Risiko der „Säuglingssterblichkeit“ bei Firmengründungen Wagner [1992b]). Die Altersstruktur der Firmenpopulation einer Region kann demnach mit dazu beitragen, Unterschiede in den Schließungsintensitäten zu erklären. Leider liegen über diese Altersstruktur für die niedersächsischen Raumordnungsregionen keine Angaben vor, und die Werte lassen sich aus den zur Verfügung stehenden Daten auch nicht berechnen, da die Gründungsjahre der 1978 bereits bestehenden Betriebe nicht zu ermitteln sind. Für die Jahre ab 1979 lassen sich jedoch (mit einigen Unschärfen behaftete) Gründungsintensitäten, d.h. Anteile von in Betrieben eines Gründungsjahrgangs tätigen Personen an allen Beschäftigten in einer Region, ermitteln (vgl. Gerlach/Wagner 1993b). Ein Blick auf die intertemporale Korrelation von Gründungs- und Schließungsintensitäten, die in Tabelle 6 dokumentiert ist, zeigt, daß die Schließungsintensitäten oft – aber nicht immer und nicht immer hoch – positiv mit den Gründungsintensitäten der Vorjahre korreliert sind, wobei jedoch kein klares Muster etwa in Form eines zunächst geringeren, dann starken und nach einer bestimmten Anzahl von Jahren sich abschwächenden Zusammenhangs zu beobachten ist.

Dies verwundert auch nicht, denn ein Zusammenhang zwischen Schließungsintensität und den Gründungsintensitäten der Vorjahre ist zwar nach

Tabelle 6: Intertemporale Korrelation von Gründungsintensitäten und Schließungsintensitäten

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Schließungsintensitäten														
1979	+.48													
1980	-.18	-.15												
1981	+.31	+.33	+.10											
1982	-.04	+.27	+.49	+.00										
1983	+.63	+.65	+.71	+.70	+.47									
1984	+.67	+.26	+.70	+.35	+.43	+.44								
1985	-.16	+.27	+.43	-.05	+.52	+.25	-.07							
1986	+.48	+.37	+.34	+.35	+.35	+.36	+.49	+.50						
1987	+.22	+.07	-.02	.26	-.18	-.03	+.25	+.29	+.39					
1988	+.44	+.19	+.22	+.41	-.01	+.43	+.66	+.29	+.30	+.24				
1989	-.10	-.05	+.15	-.18	+.25	+.02	-.20	-.11	+.06	+.06	+.01			
1990	-.19	+.27	+.31	+.14	+.54	+.28	+.11	+.17	+.07	+.10	+.26	+.31		
1991	+.15	+.03	-.01	+.08	+.05	+.35	+.43	+.09	+.05	+.41	+.20	+.18	+.38	
1992	-.23	-.25	-.06	-.17	+.01	-.14	-.06	-.17	-.14	-.17	-.01	+.19	-.14	

den oben angeführten Überlegungen *ceteris paribus* zu erwarten – die Korrelationskoeffizienten sind aber nicht Ergebnis einer *ceteris paribus-Analyse*, die erst in einer ökonometrischen Schätzung geleistet werden kann. Andererseits sind die Gründungsintensitäten in den Regionen intertemporal recht hoch positiv korreliert (vgl. Gerlach/Wagner 1993b), so daß der Wert dieses Indikators für das Jahr t auch die Gründungsintensität in den Vorjahren recht gut abbildet. Um die Bedeutung der Altersstruktur der Betriebspopulation einer Region für die Schließungsintensität zu überprüfen, wird daher die *Gründungsintensität* des entsprechenden Jahres in das ökonometrische Modell als erklärende Variable einbezogen.

Im interregionalen Vergleich hohe Kostenbelastungen der Betriebe werden, so ist zu erwarten, das Risiko von Betriebsschließungen erhöhen, wenn es den Firmen nicht gelingt, diese zusätzlichen Kosten auf ihre Kunden zu überwälzen und damit die negativen Einflüsse auf die Rendite auszugleichen. In das ökonometrische Modell werden zwei Variablen einzbezogen, die solche interregionalen Kostenunterschiede messen: *Gewerbesteuerhebesätze* und *regionales Lohnniveau*. Für beide Variablen gilt jedoch, daß *ceteris paribus* ein positiver Zusammenhang mit der Schließungsintensität nicht zwangsläufig zu erwarten ist. Selbst wenn man (was man angesichts fehlender Informationen tun muß, aber wohl mit einem gewissen Recht auch tun kann) unterstellt, daß es keine verzerrnden interregionalen Unterschiede bezüglich des Spielraums bei den genannten Überwälzungsmöglichkeiten gibt, so kann ein hoher Steuersatz einhergehen mit der Bereitstellung einer entsprechend gut ausgebauten öffentlich finanzierten Infrastruktur und damit mit Kostenentlastungen bei Betrieben, und ein hohes Lohnniveau kann auch Folge eines hohen Qualifikationsniveaus der tätigen Personen sein. Beides kann aus Mangel an entsprechenden Daten hier nicht direkt berücksichtigt werden (auf eine Möglichkeit einer indirekten Berücksichtigung im gewissen Umfang wird unten noch eingegangen).

Das empirische Modell wird mit einem über die 16 Regionen und 13 Jahre zwischen 1979 und 1991 gepoolten Datensatz geschätzt; das Jahr 1992 konnte aus Mangel an Daten für die erklärenden Variablen hierbei nicht berücksichtigt werden.

In Modellen für gepoolte Daten können Methoden der Paneldaten-Analyse (vgl. Hsiao 1986) verwendet werden, um *fixe Zeiteffekte* zu berücksichtigen. Damit ist es möglich, die Bedeutung makroökonomischer Einflüsse, die über die Zeit, aber nicht über die Regionen variieren, abzuschätzen und so für diese zu kontrollieren. Statt die Zeit in Form von Dummy-Variablen für die einzelnen Jahre einzubeziehen, wird hierbei ein „direkter“ Weg beschritten, indem ein Indikator für die gesamtwirtschaftliche konjunkturelle Entwicklung in das empirische Modell aufgenommen wird. Herangezogen wird die *Wachstumsrate des realen Bruttosozialprodukts des*

Vorjahrs; es ist zu erwarten, daß die Betriebsschließungsintensitäten um so höher ausfallen, je schlechter die Konjunkturlage ist. Die Verzögerung um ein Jahr in Form der Wachstumsrate des Vorjahres soll einerseits abbilden, daß diese Prozesse Zeit brauchen, um zu wirken; andererseits ist damit sichergestellt, daß die gemessene Wachstumsrate nicht selbst Folge der zu erklärenden Bestriebsschließungen sein kann.

Neben den für alle Regionen konstanten über die Jahre variierten Zeiteffekten können in Modellen mit gepoolten Regions-Jahres-Daten auch über die Zeit konstante zwischen den Regionen variierte Effekte berücksichtigt werden. Solche *fixen Regionseffekte* bilden unbeobachtete, nicht im empirischen Modell erfaßte Besonderheiten einzelner Regionen (etwa die geographische Lage mit Einflüssen auf die Qualität der Verkehrsanbindung oder die relativ gute Infrastrukturausstattung) ab. Die Kontrolle solcher fixer Regionaleffekte ist besonders wichtig, da immer dann, wenn diese mit den explizit im empirischen Modell berücksichtigten Einflußfaktoren korreliert sind, ein Verzicht auf ihre Einbeziehung zu verzerrten Koeffizientenschätzungen führt. Die oben angestellten Überlegungen zu den (möglicherweise) unterschiedlichen Konsequenzen eines vergleichsweise hohen Gewerbesteuerhebesatzes in Abhängigkeit von der damit gegebenenfalls verbundenen unterschiedlich guten Ausstattung der Regionen mit (aus der Sicht eines Betriebes kostensenkenden) öffentlichen Infrastruktureinrichtungen, für die im empirischen Modell aus Mangel an entsprechenden Informationen nicht explizit kontrolliert werden kann, verdeutlichen die Relevanz dieser methodischen Überlegungen im hier betrachteten Kontext. Das empirische Modell wird daher um einen Satz entsprechender Regionsdummies ergänzt, wobei auch über Schätzungen ohne diese Dummies berichtet wird, um zu zeigen, in welchem Ausmaß sich Vorzeichen, Größenordnung und ausgewiesene statistische Signifikanz der übrigen geschätzten Koeffizienten durch die Kontrolle für unbeobachtete Regionsheterogenität verändern.

5. Ergebnisse der ökonometrischen Untersuchung

Für die empirische Untersuchung der Bestimmungsgründe der interregionalen und intertemporalen Variation der Betriebsschließungsintensitäten in niedersächsischen Raumordnungsregionen wird ein dreistufiges Vorgehen gewählt:

In einem ersten Schritt wird ein Modell auf der Basis der über die 16 Regionen und 13 Jahre gepoolten Daten geschätzt. Hierfür wird die einfache Methode der kleinsten Quadrate (OLS) verwendet, wobei wegen der bei Arbeiten mit Querschnittsdaten häufig auftretenden Heteroskedastie die *t*-Werte auf der Basis einer heteroskedastie-konsistent geschätzten Kovarianzmatrix (White-Methode) berechnet werden. Außerdem wird mit drei

Varianten des RESET-Tests geprüft, ob die Nullhypothese einer korrekten linearen Spezifikation des Modells abzulehnen ist. Hierbei wird das Modell um die zweite, zweite und dritte bzw. zweite, dritte und vierte Potenz der geschätzten endogenen Variable erweitert und geprüft, ob die geschätzten Koeffizienten dieser Testgrößen signifikant von Null verschieden sind, was auf das Vorliegen nicht-linearer Zusammenhänge hinweist. Um zu prüfen, ob unbeobachtete regionale Heterogenität eine Rolle spielt und wie sich die Ergebnisse durch die Einbeziehung von für diese kontrollierende Regions-Dummies auf die Schätzergebnisse auswirkt, wird das empirische Modell sowohl mit als auch ohne einen Satz Regionsdummies geschätzt. Die Ergebnisse finden sich in den Spalten 1 und 2 der Tabelle 7.

Tabelle 7
Ergebnisse der ökonometrischen Untersuchung

Exogene Variable	Endogene Variable [Methode]	Schließungsintensität			
		[OLS]	[OLS]	[LMS/RLS]	[LMS/RLS]
Konstante	β	-0.730	1.663	0.190	-0.074
	<i>t</i>	-0.74	0.37	0.30	-0.01
Anteil Beschäftigte in Kleinbetrieben	β	0.060	0.075	0.031	-0.009
	<i>t</i>	3.38**	1.74	2.86**	-0.31
Gründungsintensität	β	0.221	0.386	0.614	0.619
	<i>t</i>	0.58	1.28	3.09**	3.14**
Lohnniveau	β	-0.019	0.029	-0.009	-0.006
	<i>t</i>	-1.12	1.15	-0.73	-0.38
Gewerbesteuersatz	β	0.005	-0.007	0.002	0.001
	<i>t</i>	1.93	-0.48	1.04	0.17
Wachstumsrate des realen BSP (Lag: 1)	β	-0.017	-0.044	-0.019	-0.039
	<i>t</i>	-0.41	-1.08	-0.74	-1.36
Fixe Regionaleffekte		nein	ja	nein	ja
Bestimmtheitsmaß R ²		0.156	0.254	0.235	0.332
korrigiertes R ²		0.136	0.174	0.214	0.254
Fallzahl		208	208	191	191
RESET Spezifikat.-Test					
RESET (2)		3.165	0.279	5.681*	2.037
RESET (3)		2.120	0.156	3.397*	1.170
RESET (4)		2.233	0.106	3.171*	1.031

Anm.: β = geschätzter Regressionskoeffizient

t = *t*-Wert, berechnet mit White's heteroskedastie-konsistentem Kovarianzmatrix-Schätzer

*(**) = signifikant für $\alpha = 0.05$ (0.01)

Die geschätzten Koeffizienten der fixen Regionaleffekte sind aus Platzgründen nicht mit abgedruckt. Zur Erläuterung der Variablen sowie der Schätz- und Testverfahren vgl. Text.

Ein *F*-Test der Nullhypothese, daß die Regionseffekte keine Rolle spielen, lehnt diese auf einem Fehlerniveau von deutlich unter ein Prozent ab; das Modell ohne die Regionsdummies ist also fehlspezifiziert, seine Ergebnisse können nicht interpretiert werden und sollen lediglich als Vergleichsgrößen dafür herangezogen werden, welche Bedeutung eine Kontrolle für unbeobachtete Regionsheterogenität bzw. ein (unberechtigter!) Verzicht hierauf hat.

Nach den Ergebnissen der OLS-Schätzung des Modells mit fixen Regionseffekten, dessen lineare Spezifikation von keiner der drei Varianten des RESET-Tests auf einem üblichen Signifikanzniveau von fünf Prozent abgelehnt wird, haben sowohl der Anteil der Beschäftigten in Kleinbetrieben als auch die Gründungsintensität den erwarteten die Schließungsintensität erhöhenden Einfluß, während eine bessere gesamtwirtschaftliche konjunkturelle Situation erwartungsgemäß die Schließungsintensität senkt. Allerdings ist von diesen drei genannten Koeffizienten auf einem üblichen Fehlerniveau keiner statistisch von Null verschieden, so daß diese Einflüsse nicht signifikant sind. Insignifikant sind auch die geschätzten Koeffizienten für das Lohnniveau und den Gewerbesteuersatz.

Ein Vergleich der Ergebnisse der Schätzungen aus dem Modell mit fixen Regionaleffekten und aus dem (fehlspezifizierten!) Modell ohne diese Kontrolle für unbeobachtete Heterogenität zeigt, daß ein Verzicht auf Regionsdummies in die Irre führt: Sowohl die aus den Ergebnissen der Spalte 1 der Tabelle 7 abzuleitende Folgerung eines signifikant positiven Einflusses des Anteils der in Kleinbetrieben tätigen Personen als auch der hiernach (nahezu auf einem Fehlerniveau von fünf Prozent statistisch signifikante) festzustellende positive Einfluß des Gewerbesteuerbesatzes auf die Schließungsintensität ist nicht haltbar. Dies zeigt deutlich, daß in Untersuchungen interregionaler Unterschiede von Schließungsintensitäten die Kontrolle unbeobachteter Regionsheterogenität durch fixe Effekte nicht vernachlässigt werden darf. Die Ergebnisse der von Audretsch/Fritsch (1992, 1993) für Schließungsraten in Raumordnungsregionen der Bundesrepublik Deutschland vorgelegten Studien können daher solange nicht akzeptiert werden, wie nicht gezeigt ist, daß sie auch auf der Basis gepoolter Daten und bei Kontrolle für fixe Regionaleffekte Bestand haben.

Insgesamt gesehen kann aber auch das hier geschätzte Modell mit fixen Regionaleffekten nicht überzeugen, da keiner der explizit berücksichtigten Einflußfaktoren statistisch signifikant ist. Ein erneuter Blick auf die in Tabelle 3 abgedruckten Werte der Schließungsintensitäten zeigt einen hierfür möglicherweise zumindest mit verantwortlichen Grund: Es fällt auf, daß in einzelnen Regionen für einzelne Jahre sehr hohe Werte auftreten, wobei sich diese weder auf bestimmte Regionen noch auf bestimmte Jahre konzentrieren. So beträgt z.B. die Schließungsintensität in der Region 3 im

Jahr 1982 4.67%, was mehr als das Dreifache des Mittelwertes der Intensitäten über alle Regionen dieses Jahres und mehr als das Doppelte des Durchschnittswertes der für diese Region ermittelten Intensitäten über die betrachteten Jahre ist. Ähnliche extreme Beobachtungen finden sich z.B. für die Region 6 im Jahr 1986, 1991 für Region 4 oder 1983 für Region 9. Weder die für alle Regionen identische gesamtwirtschaftliche konjunkturelle Entwicklung eines Jahres noch die über die Zeit konstanten fixen Regionseffekte können ihrer Konstruktion nach diese Besonderheiten in einzelnen Jahren in einzelnen Regionen erklären; dies gilt auch für die zwar über die Jahre und die Regionen variierenden, sich jedoch von Jahr zu Jahr eher langsam ändernden übrigen Variablen. Anders formuliert: Die Modellvariablen sind über Zeit und Raum nicht variabel genug, um diese starken Ausschläge der Schließungsintensitäten zu erklären, hinter denen sich z.B. spektakuläre Zusammenbrüche eines Großbetriebes verbergen können. Dies kann auch gar nicht Zweck eines empirischen Modells sein, bei dem es ja um die Identifikation der für die Masse der Jahre und Regionen zur Erklärung der Unterschiede in den Schließungsintensitäten relevanten Einflußfaktoren geht. Spektakuläre Einzelfälle wie die Schließung eines großen Werkes lassen sich nur durch Rückgriff auf Einzelinformationen in Fallstudien untersuchen, wobei es hier aus Datenschutzgründen allerdings noch nicht einmal möglich ist zu prüfen, ob die beobachteten hohen Schließungsintensitäten durch solchen Ereignisse verursacht wurden.

Selbst wenn man aber akzeptiert, daß es nicht Aufgabe eines empirischen Modells sein kann, neben dem Druckschnitt auch die Besonderheiten einzelner Regionen und Jahre zu erklären, so bleibt zu beachten, daß die Ergebnisse einer mit OLS vorgenommenen Modellschätzung von dem Vorhandensein solcher „Ausreißer“ im Datensatz in beliebigem Umfang beeinflußt werden können. Anders formuliert: Das geschätzte Modell kann für den Großteil der Fälle geltende Zusammenhänge auch dann nicht identifizieren, wenn es diese Zusammenhänge gibt, da die Schätzergebnisse durch die Sonderfälle dominierend geprägt werden. Eine einzige stark von anderen Jahren und Regionen abweichende Beobachtung kann dazu führen, daß die Modellschätzung „zusammenbricht“. Dies ist eine Eigenschaft des Schätzverfahrens OLS, deren negative Konsequenzen durch die Verwendung eines gegen die Präsenz von Ausreißern robusten Schätzverfahrens vermieden werden können.

In einem zweiten Schritt wird daher das empirische Modell (zunächst in der Variante ohne die fixen Regionaleffekte) mit Least Median of Squares und Reweighted Least Squares (LMS/RLS) geschätzt. Hierbei wird zunächst durch die Minimierung des Medians der quadrierten Residuen (anstelle ihrer Summe wie bei OLS) eine Identifikation von Ausreißern ermöglicht, die „weit weg“ von der Masse der Beobachtungen liegen, um

dann anschließend das Modell ohne diese Ausreißerwerte erneut zu schätzen (vgl. *Rousseeuw/Leroy* 1987). LMS identifiziert 17 der 208 Werte der Schließungsintensitäten als Ausreißer (dies sind die Werte 1979/15, 1982/3, 1983/2,9,12, 1984/6, 1985/3,8,9, 1986/6,9, 1987/11, 1988/4, 1989/3,6, 1990/3, 1991/4); viele – aber nicht alle! – dieser Werte sind vergleichsweise hoch, aber nicht alle vergleichsweise hohen Werte stellen Ausreißer dar. Eine Häufung von Ausreißern in einzelnen Regionen oder Jahren ist hierbei nicht festzustellen. Die Ergebnisse der Schätzung des Modells auf der Basis der um die Ausreißer bereinigten Daten findet sich in Spalte 3 der Tabelle 7. Dieses Modell ist jedoch fehlspezifiziert, wie ein Test auf fixe Regionaleffekte und die Tests auf eine korrekte lineare Spezifikation (RESET-Tests) zeigen. Auf die Ergebnisse soll daher hier auch nicht weiter eingegangen werden.

In einem dritten und letzten Schritt wird daher das Modell mit fixen Regionaleffekten auf der Basis der um die von LMS identifizierten Ausreißer bereinigten Daten geschätzt. Die Ergebnisse sind in der vierten Spalte der Tabelle 7 abgedruckt. Es zeigt sich, daß die Nullhypothese einer korrekten linearen Spezifikation von keiner der drei Varianten der RESET-Prozedur auf einem üblichen Fehlerniveau abgelehnt wird. Aus den Schätzergebnissen läßt sich folgern, daß die Schließungsintensität um so höher ist, je höher die Gründungsintensität in einer Region ist, während weder für den Anteil der Beschäftigten in Kleinbetrieben noch für das Lohnniveau noch für den Gewerbesteuersatz ein Einfluß auf die Schließungsintensität zu erkennen ist; wie erwartet fällt in Zeiten einer positiven gesamtwirtschaftlichen konjunkturellen Tendenz der Arbeitsplatzabbau durch Betriebsschließungen niedriger aus, wobei dieser Zusammenhang aber statistisch nicht auf einem üblichen Fehlerniveau als gesichert gelten kann.

6. Schlußfolgerungen

In der vorliegenden Untersuchung werden regionale Unterschiede bei Schließungen von Industriebetrieben und die damit verbundenen Arbeitsplatzverluste für 16 niedersächsische Raumordnungsregionen in den Jahren von 1979 bis 1992 ermittelt. Im Durchschnitt schließen etwa zwischen vier und fünf Prozent aller Industriebetriebe pro Jahr in jeder Region, wobei es in einzelnen Jahren erhebliche interregionale Unterschiede in diesen Schließungsraten gibt. Die Schließungsintensitäten, die die Anzahl der in geschlossenen Betrieben verlorenen Arbeitsplätze als Prozentsatz aller Industriearbeitsplätze messen, liegen im Mittel über die Regionen in den betrachteten Jahren zwischen rund einem und zwei Prozent, wobei es ausgeprägte Unterschiede zwischen den Regionen und zahlreiche extreme Werte in einzelnen Regionen und Jahren gibt.

Die interregionalen und intertemporalen Unterschiede der Schließungsintensitäten werden in einem ökonometrischen Modell untersucht, wobei sich zeigt, daß sowohl die Kontrolle für unbeobachtete regionale Heterogenität durch fixe Regionaleffekte als auch die Identifikation von Ausreißern mit Hilfe eines robusten Schätzverfahrens erforderlich ist. Dies verdeutlicht, daß Untersuchungen zu Schließungsintensitäten nicht auf Querschnittsdaten aus einem Jahr, sondern auf über Regionen und Jahre gepoolte Paneldaten zurückgreifen müssen, wie dies hier zum ersten Mal für bundesrepublikanische Daten getan wird.

Als Ergebnis läßt sich festhalten, daß die Schließungsintensität in einer Region um so höher ausfällt, je höher die Gründungsintensität, d.h. der Beschäftigtenanteil in neuen Betrieben mit weniger als 50 tätigen Personen, ist. Dies kann einerseits mit der hohen „Säuglingssterblichkeit“ neugegründeter Betriebe und andererseits mit einer Verdrängung von alten durch neue Produzenten zusammenhängen. Weder für den Anteil von Beschäftigten in Kleinbetrieben noch für das regionale Lohnniveau noch für die Höhe der Gewerbesteuersätze ist ein Zusammenhang mit den Schließungsintensitäten erkennbar. In Zeiten einer vergleichsweise guten gesamtwirtschaftlichen Konjunktur fallen die Schließungsintensitäten erwartungsgemäß geringer aus; allerdings kann dieser Zusammenhang nicht als signifikant angesehen werden.

Die ökonometrische Untersuchung läßt eine ganze Reihe von Fragen unbeantwortet. Zunächst einmal bleibt offen, was sich hinter den „unbeobachteten fixen Regionaleffekten“ verbirgt, die eine wichtige Rolle spielen. Hierbei kann es sich um schwer zu erfassende und eher vage Konstrukte wie ein „wirtschaftsfreundliches Klima“ handeln, es können aber auch im empirischen Modell nicht erfaßte weitaus konkretere Größen wie die Industriestruktur sein. Darüber hinaus können Unterschiede zwischen den einzelnen Firmen von entscheidender Bedeutung dafür sein, ob ein Betrieb schließt oder nicht. Zu denken ist hierbei z.B. an zwischenbetriebliche Unterschiede bei der Entwicklung neuer Produkte oder dem Einsatz neuer Produktionsverfahren und an unterschiedliche Managementstrategien. Es ist daher die Frage „Who exits and why?“ (Acs/Audretsch 1993) zu untersuchen und dabei zu klären, welche regionalen Einflußfaktoren neben betriebs- und industriespezifischen Ursachen hierbei eine Rolle spielen, warum sich relative Arbeitsplatzverluste durch schließende Betriebe also nicht gleichmäßig im Raum verteilen.

Zur Untersuchung dieser weiterführenden Fragen können allerdings die Daten aus den Erhebungen der amtlichen Industrieberichterstattung nicht herangezogen werden: Einerseits ist eine gleichzeitige Aufgliederung der Schließungen nach Regionen und Industriezweigen wegen der damit in vielen Fällen zwangsläufig verbundenen Möglichkeit einer Identifikation

von Einzelangaben aus Datenschutzgründen unzulässig, andererseits fehlen in diesen Erhebungen Informationen zu (vermutlich) mit entscheidenden Bereichen der Firmenstrategie wie z.B. dem Innovationsverhalten. Um diese Informationslücken zu schließen, müssen entsprechende Firmenpanel-daten erhoben werden. Projekte hierzu beginnen zur Zeit in Deutschland – bis jedoch Informationen über einen hinreichend langen Zeitraum vorliegen, werden noch sehr viele Arbeitsplätze durch Betriebsschließungen verloren gehen, ohne daß abgesicherte Aussagen über die hierbei zu beobachtenden regionalen Unterschiede möglich sind.

Literatur

- Acs, Zoltan J./Audretsch, David B.* (1993), Who Exits and Why? Discussion Paper FS IV 93 - 20, Wissenschaftszentrum Berlin.
- Audretsch, David B./Fritsch, Michael* (1992), Market Dynamics and Regional Development in the Federal Republic of Germany. Discussion Paper FS IV 92 - 6, Wissenschaftszentrum Berlin.
- (1993), Betriebsgründungen, Stillegungen, Turbulenz und Regionalentwicklung. Discussion Paper FS IV 93 - 17, Wissenschaftszentrum Berlin.
- Gerlach, Knut/Wagner, Joachim* (1992), Betriebsdaten der amtlichen Statistik – eine (fast) ungenutzte Informationsquelle. *Staatswissenschaften und Staatspraxis* 3, 578 - 596.
- (1993a), Gross and Net Employment Flows in Manufacturing Industries. *Zeitschrift für Wirtschafts- und Sozialwissenschaften* 113, 17 - 28.
- (1993b), Regional Differences in Small Firm Entry in Manufacturing Industries: Lower Saxony, 1979 - 1991. Universität Hannover, mimeo (erscheint in: Entrepreneurship and Regional Development).
- Hsiao, Cheng* (1986), Analysis of Panel Data. Cambridge etc.: Cambridge University Press.
- Posselt, Thomas* (1993), Mobilitätsverhalten von Unternehmen. Eine industrieökonomische Analyse. Wiesbaden: Gabler.
- Rousseeuw, Peter J./Leroy, Annik M.* (1987), Robust Regression and Outlier Detection. New York: Wiley.
- Schmalensee, Richard* (1989), Inter-Industry Studies of Structure and Performance, in: Richard Schmalensee und Robert Willig (Eds.), *Handbook of Industrial Organization*, Vol. 2, Amsterdam etc.: North-Holland.
- Wagner, Joachim* (1992a), Small Firm Entry in Manufacturing Industries: Lower Saxony, 1979 - 1989. Universität Hannover, mimeo (erscheint in: Small Business Economics).
- (1992b), Success or Failure? The Post-Entry Performance of New Small Firms in Manufacturing Industries, Lower Saxony, 1979 - 1990. Universität Hannover, mimeo.

Anhang: Zu den Daten

Schließungsraten und Schließungsintensitäten wurden mit Betriebsdaten aus Erhebungen der amtlichen Industriebreicherstattung berechnet. Die Daten liegen in Form eines sämtliche niedersächsische Industriebetriebe für den Zeitraum 1978 bis 1992 umfassenden Paneldatensatzes im Niedersächsischen Landesamt für Statistik vor.

Der Anteil der in Kleinbetrieben tätigen Personen wurde aus dem o.a. Paneldatensatz berechnet, wobei „Kleinbetriebe“ als Betriebe mit weniger als 50 tätigen Personen definiert sind.

Die Gründungsintensität wurde ebenfalls auf der Basis des o.a. Paneldatensatzes berechnet; eine ausführliche Darstellung der Methodik findet sich in *Gerlach/Wagner* (1993b).

Gewerbesteuerbesätze wurden nach Angaben in verschiedenen Ausgaben des vom Niedersächsischen Institut für Wirtschaftsforschung (NIW) publizierten Regionalberichts berechnet.

Das regionale Lohnniveau ist der durchschnittliche Stundenlohnsatz für Arbeiter, der sehr hoch positiv mit den entsprechenden durchschnittlichen Angestelltengehältern korreliert ist; die Werte wurden nach den Angaben in den NIW-Regionalberichten berechnet. Da in der Region 14 das Volkswagenwerk von dominierender Bedeutung ist, sind aus Datenschutzgründen keine publizierten Lohnangaben verfügbar; hier wurde ein Wert verwendet, der nach Angaben des Volkswagenwerkes Wolfsburg geschätzt wurde.

Die Werte der Wachstumsrate des realen Bruttonsozialprodukts der Bundesrepublik Deutschland wurden aus der Tabelle 26 des Jahresgutachtens 1992/93 des Sachverständigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung (Deutscher Bundestag, Drucksache 12/3744) entnommen.

Betriebsgründungen und regionale Arbeitsplatzentwicklung

Von *Michael Fritsch*, Freiberg

Abstract

In dem Aufsatz wird der Zusammenhang zwischen den regionalen Neugründungsaktivitäten und der Arbeitsplatzentwicklung während der nachfolgenden Jahre für die Raumordnungsregionen Westdeutschlands analysiert. Es zeigt sich, daß die Anzahl der in einer Region stattfindenden Gründungen relativ stark durch die dort vorhandene Branchenstruktur bestimmt ist, so daß es sich als sinnvoll erweist, eine um diesen Brancheneinfluß bereinigte Gründungsrate zu bilden. Die bereinigten Gründungsraten weisen ein klares Zentrum-Peripherie-Gefälle auf. Insgesamt ergibt sich ein nur schwach ausgeprägter positiver Zusammenhang zwischen dem Niveau der regionalen Gründungsaktivitäten und der Beschäftigungsentwicklung, wobei offenbar insbesondere die Gründungen im Industriesektor von Bedeutung sind.

1. Problemstellung*

Es erscheint unmittelbar plausibel, daß Betriebsgründungen zur wirtschaftlichen Expansion beitragen und somit auch die Entwicklung der betreffenden Region in positiver Weise beeinflussen. Schließlich bedeutet die Errichtung eines Betriebes ja die Aufnahme wirtschaftlicher Aktivität, was sich in der einen oder anderen Form in der Regionalentwicklung niederschlagen müßte. So wird denn auch die Förderung von Gründungen allgemein als ein wesentliches Mittel zur Stimulierung wirtschaftlicher Entwicklung angesehen. Erstaunlicherweise ist der Beitrag von Gründungen zur Regionalentwicklung bisher allerdings kaum empirisch untersucht worden, so daß nur sehr spärliche Informationen darüber vorliegen, welcher Stellenwert ihnen tatsächlich zukommt.

Im folgenden wird der Bedeutung von Betriebsgründungen für die Regionalentwicklung in der Bundesrepublik Deutschland nachgegangen. Zunächst resümiert Abschnitt 2 wesentliche Hypothesen und die bisherige empirische Evidenz. Aufbauend auf einer Beschreibung der Datenbasis (Abschnitt 3)

* Ich danke Kerstin Ulbricht für ihre engagierte und kompetente Unterstützung bei der Bewältigung der umfangreichen EDV-Arbeiten.

werden in Abschnitt 4 alternative Vorgehensweisen bei der Bildung einer regionalen Gründungsrate diskutiert. Abschnitt 5 vergleicht die verschiedenen Gründungsraten. Nach einem kurzen Überblick über räumliche Unterschiede des Gründungsgeschehens in Westdeutschland im Zeitraum 1986 - 89 (Abschnitt 6) wird dann der Zusammenhang zwischen Gründungsaktivitäten und Regionalentwicklung analysiert (Abschnitt 7).

2. Hypothesen und bisherige empirische Evidenz

Daß zwischen Gründungen und Regionalentwicklung ein positiver Zusammenhang besteht, ist weitgehend unbestritten. Unklar ist allerdings, wie groß dieser Beitrag der Gründungen im Vergleich zu anderen Komponenten der Arbeitsplatzdynamik – etwa der Entwicklung etablierter Betriebe – ausfällt. Es gibt eine Reihe von Argumenten, die an einem hohen Stellenwert der Gründungen für die Regionalentwicklung zweifeln lassen:

- *Erstens* scheitern viele Unternehmensgründungen bereits nach relativ kurzer Zeit, so daß sie nur einen temporären Beitrag zur Regionalentwicklung leisten.¹
- *Zweitens* weist eine Reihe empirischer Untersuchungen darauf hin, daß nur ein sehr geringer Anteil der Gründungen in wesentlichem Ausmaß zusätzliche Arbeitsplätze schafft (vgl. *Fritsch* 1990a, *König* 1994, *Storey/Johnson* 1987).
- *Drittens* ist der Einfluß von Gründungen auf die Regionalentwicklung auch davon abhängig, ob ein erfolgreicher Marktzutritt von Newcomern den Marktaustritt etablierter Konkurrenten erzwingt bzw. inwieweit der Marktzutritt vor allem durch die Auslagerung bestimmter Aktivitäten (und somit Arbeitsplätze) aus etablierten Unternehmen auf die neuen Anbieter ermöglicht wird (ausführlicher hierzu *Fritsch* 1990b). Stellen die Gründungen primär Ausfluß einer solchen Reorganisation dar und sind derartige Veränderungen der zwischenbetrieblichen Arbeitsteilung mit Einsparungseffekten verbunden, dann könnte (kurz- und mittelfristig) auch ein negativer Zusammenhang zwischen Gründungen und Arbeits-

¹ Interessant sind in diesem Zusammenhang etwa die Ergebnisse einer Untersuchung von *Boeri/Cramer* (1991) für die Bundesrepublik Deutschland. *Boeri/Cramer* verfolgen die Beschäftigungsentwicklung in Gründungskohorten über einen Zeitraum von bis zu zehn Jahren und stellen fest, daß die Zahl der Arbeitsplätze in den neugegründeten Betrieben am Ende der Analyseperiode z.T. geringer ist als im Jahr der Gründung; im ersten Jahr nach der Gründung erreicht die Zahl der Arbeitsplätze ihr Maximum. Die Abnahme der Beschäftigtenzahl in der Gründungskohorte ist vor allem durch Stilllegungen bedingt. Nach den Ergebnissen von *Boeri/Cramer* (1991) überleben nur gut 40% der Betriebe einen Zeitraum von zehn Jahren; die überlebenden Neugründungen haben dann im Durchschnitt 6,1 Beschäftigte.

platzentwicklung bestehen. Langfristig sind eher positive Wirkungen eines derartigen Reorganisationsprozesses zu erwarten, da die Rationalisierungseffekte zu einer Verbesserung der Wettbewerbsfähigkeit und, sei es, zu einer Vermeidung von (noch) stärkeren Arbeitsplatzverlusten führen.

- *Viertens* schließlich kann es für das Ausmaß der durch Gründungen ausgelösten regionalen Arbeitsplatzeffekte von entscheidender Bedeutung sein,
 - inwieweit die neuen Betriebe Vorleistungen aus der Region beziehen,
 - in welchem Umfang sie Inputs für die regionale Wirtschaft bereitstellen, die vorher von regionsexternen Anbietern bezogen wurden bzw.
 - ob sich ihre Existenz auf andere Weise – etwa durch Verstärkung von Agglomerationsvor- oder -nachteilen – auf die Regionalwirtschaft auswirkt.²

Während die ersten beiden Einwände die *direkten* Beschäftigungseffekte, nämlich die in den neugegründeten Betrieben entstehenden Arbeitsplätze, betreffen, beziehen sich die beiden letztgenannten Argumente auf Wirkungen, die *indirekter* Natur sind. Alle diese Überlegungen deuten darauf hin, daß es wohl sehr stark von der Art der neuen Unternehmen abhängt, inwieweit sie positiv zur regionalen Beschäftigungsentwicklung beitragen.³

Bislang liegen nur relativ wenige empirische Analysen des Beitrages von Gründungen zur Regionalentwicklung vor. Dies ist sicherlich zu einem ganz wesentlichen Teil durch die Datenlage bedingt, denn nur für wenige Länder sind (häufig erst seit kurzer Zeit) einigermaßen vollständige Informationen über die Anzahl der Gründungen differenziert nach Regionen verfügbar. Ein schwerwiegendes Defizit fast aller dieser Untersuchungen besteht darin, daß sie sich darauf beschränken, den *direkten* Beitrag der Gründungen zur Beschäftigungsentwicklung zu ermitteln, dabei aber die *indirekten* Effekte vollkommen vernachlässigen. Die verschiedenen Untersuchungen zu den direkten Beschäftigungseffekten von Gründungen kommen – je nach Untersuchungsregion, sektoraler Abgrenzung und Analysezeitraum – zu teilweise recht unterschiedlichen Ergebnissen. Während beispielsweise Birch (1979) in seiner berühmten Untersuchung für die USA

² Anhand der interregionalen Lieferverflechtung läßt sich auch ein Fall konstruieren, in dem die Gründungen zu regionalen Arbeitsplatzverlusten führen. Insgesamt negative Arbeitsplatzeffekte können dann eintreten, wenn die Newcomer eine relativ geringe Fertigungstiefe aufweisen, große Mengen von Vorleistungen aus anderen Regionen beziehen und gerade solche etablierten Betriebe verdrängen, die durch eine vergleichsweise hohe Fertigungstiefe gekennzeichnet sind und ihre Vorleistungen überwiegend aus der Region erhalten.

³ Siehe zu entsprechenden empirischen Befunden Brüderl/Bühler/Ziegler (1993) sowie auch Baumol (1990).

einen deutlich ausgeprägten positiven Zusammenhang zwischen den durch Betriebsgründungen geschaffenen Arbeitsplätzen und der regionalen Beschäftigungsentwicklung ermittelte, deuten andere Analysen (etwa Boeri/*Cramer* 1992, *Cramer/Koller* 1988, *Dunne/Roberts/Samuelson* 1989, *Gerrach/Wagner* 1993 und 1994, *König* 1994) darauf hin, daß die direkten Arbeitsplatzeffekte derjenigen Betriebe, die über die gesamte Untersuchungsperiode vorhanden sind (sogenannte „survivor“-Betriebe), häufig stärker ausfallen als die Arbeitsplatzeffekte der Neugründungen. Meist steht der Anzahl der Gründungen während einer Periode eine ähnlich hohe Anzahl an Stilllegungen gegenüber, so daß der entsprechende Saldo relativ gering ausfällt.⁴ Auch die Arbeitsplatzzuwächse in neugegründeten Betrieben werden vielfach durch negative Beschäftigungswirkungen von Stilllegungen in etwa kompensiert.

Eingehende Analysen des Zusammenhangs zwischen Betriebsgründungen und der regionalen Beschäftigungsentwicklung unter Berücksichtigung möglicher indirekter Effekte wurden erst vor kurzem, und m.W. erstmalig, von *Reynolds* (1994) bzw. *Reynolds/Maki* (1990) vorgelegt. Diese Untersuchung für die USA kommt zu dem Ergebnis, daß die Gründungsrate (Anzahl der Neugründungen dividiert durch die Anzahl der bereits ansässigen Betriebe) in einem positiven statistischen Zusammenhang mit dem Wachstum während der nachfolgenden Jahre steht. Allerdings sind nicht unerhebliche Unterschiede in der Stärke dieses Zusammenhangs zwischen den einzelnen Analyseperioden zu verzeichnen.

3. Datenbasis

Den nachfolgend präsentierten Analysen liegen Daten über die Betriebsgründungen in den Raumordnungsregionen Westdeutschlands während der Jahre 1986 - 89 zugrunde. Diese Angaben wurden aus der in eine Betriebsdatei transformierten Statistik der sozialversicherungspflichtig Beschäftigten (im folgenden: Beschäftigtenstatistik) generiert. Die Beschäftigtenstatistik enthält für jeden Betrieb eine Betriebsnummer; Informationen über Gründungen lassen sich aus dieser Statistik insofern generieren, als man neu erscheinende Betriebsnummern als neue Betriebe interpretiert (ausführlicher hierzu *Fritsch/König/Weißhuhn* 1994).

Eine solche Interpretation ist allerdings nicht ganz unproblematisch. So kann etwa der Fall eintreten, daß ein Betrieb im Rahmen einer Reorganisation, einer Übernahme oder einer Umgründung eine neue Nummer erhält, obwohl er mehr oder weniger unverändert weiterbesteht. Eine gewisse

⁴ Siehe zum Zusammenhang zwischen Gründungen und Stilllegungen *Cramer/Koller* (1988) sowie *Audretsch/Fritsch* (1994c).

Ungenauigkeit ist auch dadurch bedingt, daß die Statistik nur Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten erfaßt und daher keine Firmen enthält, die ausschließlich aus dem Gründer und eventuell mithelfenden Familienangehörigen bestehen. Somit fehlt in dem Datensatz zwar ein, gemessen an der Anzahl sämtlicher Gründungen, vermutlich nicht unerheblicher Anteil der Gründungen; allerdings handelt es sich dabei um genau diejenigen Kleinstunternehmen, die nur in sehr geringem Maße zur Arbeitsplatzentwicklung beitragen. Daraus, daß ein Betrieb erst dann in der Statistik erscheint, sobald der erste sozialversicherungspflichtig Beschäftigte dort tätig ist (und somit ein zusätzlicher Arbeitsplatz erzeugt wird), resultiert u.U. auch eine falsche zeitliche Zuordnung der Gründung.

Zur Identifikation der Gründungen wurden zunächst diejenigen Betriebe ermittelt, welche während des Zeitraums 1986 - 89 neu in der Statistik erscheinen. Es gibt nun eine Reihe von Fällen, in denen ein bestimmter Betrieb nur ein einziges Mal auftaucht. Hierbei dürfte es sich im wesentlichen entweder um fehlerhafte Angaben oder um solche Gründungen handeln, welche bereits während der allerersten Phase scheitern und denen das Etablieren dauerhafter Kontakte zu Abnehmern bzw. Zulieferern mißlingt. Um zumindest grobe Verzerrungen durch Fehler bzw. zweifelhafte Fälle auszuschließen, wurden nur solche in einem bestimmten Jahr erstmalig auftauchenden Betriebsnummern als „Gründung“ gewertet, welche auch noch im jeweils darauffolgenden Jahr im Datensatz enthalten waren.

Betrachtet man die Größenstruktur (Beschäftigtenzahl) der so ermittelten Gründungen, so zeigt sich, daß der ganz überwiegende Teil dieser Betriebe außerordentlich klein ist: Deutlich mehr als 80% haben weniger als zehn und gut 90% haben weniger als 50 Beschäftigte. Die Beschäftigtenstatistik enthält aber auch eine Anzahl von neuen Betrieben, welche bereits von Beginn an mehrere hundert Beschäftigte aufweisen. Hierbei dürfte es sich ganz überwiegend entweder um Umgründungen, um die Errichtung von Zweigbetrieben durch Großunternehmen oder um solche (seltenen) Fälle handeln, in denen vorhandene Betriebe aus „buchungstechnischen“ Gründen aufgespalten und unter separaten Nummern geführt werden. Da es hier weniger um Zweigbetriebs-Gründungen, sondern vor allem um die Gründung *neuer Unternehmen* gehen soll, und neue Unternehmen bekanntlich in der Regel klein beginnen, bleiben im folgenden neue Betriebe, die zwei Jahre nach dem erstmaligen Erscheinen in der Beschäftigtenstatistik 50 und mehr Beschäftigte aufweisen, ausgeschlossen. Neugründungen sind hier also Betriebe, die erstmalig in der Beschäftigtenstatistik erscheinen, auch noch im darauffolgenden Jahr enthalten sind und im zweiten Jahr weniger als 50 Beschäftigte aufweisen. Das räumliche Raster für die nachfolgend dargestellten Analysen stellen „Raumordnungsregionen“ in der Abgrenzung

und Typisierung der Bundesforschungsanstalt für Landeskunde und Raumordnung (Bonn) dar.

4. Alternative Vorgehensweisen bei der Bildung einer Gründungsrate

Da Regionen in der Regel deutliche Unterschiede hinsichtlich ihres jeweiligen wirtschaftlichen Potentials (z.B. Zahl der Betriebe und Beschäftigten) aufweisen, stellt die absolute Anzahl der Gründungen keinen sinnvollen Indikator für interregionale Vergleiche des Gründungsgeschehens bzw. der Gründungsneigung dar, sondern es muß eine Gründungsrate gebildet werden. Für die Bildung einer solchen Gründungsrate bestehen nun eine ganze Reihe von Alternativen, die einmal die Erfassung der Gründung selbst (Punkt 4.1), zum zweiten die Wahl der Bezugsgröße (Punkt 4.2) sowie drittens schließlich die Berücksichtigung von Einflüssen der regionalen Branchenstruktur betreffen (Punkt 4.3).

4.1 Die Erfassung der Gründungen

Was die Gründungen selbst angeht, so kann man sie einmal mit ihrer Anzahl, zum anderen auch mit ihrer Größe (i.d.R. Beschäftigtenzahl) erfassen. Eine Gewichtung der Gründungen mit ihrer jeweiligen Größe bringt verschiedene Probleme mit sich. Erstens ist es mit sehr hohen Anforderungen an die Datenbasis verbunden, nicht nur Gründungen eindeutig zu identifizieren, sondern auch noch zuverlässige Informationen über deren Größe zu bieten. Da eine Reihe von Neugründungen sich während der ersten Zeit ihrer Existenz relativ dynamisch entwickeln, ist zweitens der Zeitpunkt, zu dem die Größe gemessen wird, in Relation zum Zeitpunkt der Gründung festzulegen. Um solche Probleme zu umgehen, wird im folgenden immer auf die Anzahl der Gründungen und nicht auf ihre Größe abgestellt.⁵

4.2 Die Bezugsgröße: „Arbeitsmarkt“ und „betriebsökologischer“ Ansatz

Auch was den Nenner der Gründungsrate angeht, bestehen Alternativen. Zum einen kann man die Anzahl der Gründungen bzw. Stilllegungen durch

⁵ Dies nicht zuletzt auch deshalb, weil es sich bei den auf der Grundlage der Beschäftigtenstatistik irrtümlich als Gründungen identifizierten Fällen eher um relativ große Betriebe handeln dürfte. Gewichtet man die Gründungen mit ihrer Größe, so ist die durch Fehler oder durch im Datenmaterial enthaltene Zweifelsfalle bewirkte Verfälschung vermutlich sehr viel stärker, als wenn man lediglich auf die Anzahl der Gründungen abstellt.

die Anzahl der in der betreffenden Region vorhandenen Betriebe dividieren, wobei die Gründungsrate dann gewissermaßen einen Indikator für die „Fertilität“ des vorhandenen Betriebsbestandes darstellt („betriebsökologischer Ansatz“). Eine andere Möglichkeit besteht darin, die Gründungen auf das regionale Arbeitskräftepotential zu beziehen, was man als „Arbeitsmarktansatz“ kennzeichnen könnte. Geht man davon aus, daß jeder nicht freiwillig arbeitslose Erwerbsfähige vor der Frage steht, entweder einer „abhängigen“ Beschäftigung nachzugehen (bzw. eine solche „abhängige“ Beschäftigung zu suchen) oder sich selbstständig zu machen (d.h. ein Unternehmen zu gründen), dann stellen die in einer Region vorhandenen Erwerbspersonen einen wesentlichen Teil des Potentials an möglichen Gründern dar; dies insbesondere auch deshalb, weil – wie diverse empirische Untersuchungen belegen – für den Schritt in die Selbstständigkeit fast immer ein Standort in der Nähe des Wohnsitzes des Gründers gewählt wird (vgl. etwa *Gudgin* 1978, *O'Farrell* 1986). Die entsprechend dem „Arbeitsmarktansatz“ formulierte Gründungsrate läßt sich dann als „Gründungsnierung“ interpretieren; je höher der Wert des Indikators „Anzahl der Gründungen/Anzahl der Erwerbspersonen“, desto eher ist damit zu rechnen, daß eine Erwerbsperson in der betreffenden Region den Schritt in die Selbstständigkeit wagt.

Ob man die Anzahl der Gründungen in einer Region mit der Anzahl der bereits ansässigen Betriebe oder mit der Anzahl der Erwerbspersonen standardisiert, hat nicht nur Auswirkungen auf die Ergebnisse von Analysen interregionaler Unterschiede des Gründungsgeschehens (siehe hierzu *Audretsch/Fritsch* 1994a, b), sondern auch auf die Möglichkeiten für sektorspezifische Analysen. Während sich nämlich die Gründungsrate entsprechend dem „betriebsökologischen“ Ansatz ohne weiteres für bestimmte Branchen berechnen läßt, indem man die Anzahl der Gründungen in einer Branche durch die Anzahl der ansässigen Betriebe der betreffenden Branche dividiert, ist die Branchendifferenzierung der Gründungsrate entsprechend dem „Arbeitsmarktansatz“ weitaus problematischer. Zwar ist es möglich, die Anzahl der Gründungen in einem bestimmten Wirtschaftszweig auf die Anzahl der Beschäftigten in der betreffenden Branche zu beziehen; allerdings stellt sich dann das kaum lösbare Problem, die arbeitslosen Erwerbsfähigen eindeutig den verschiedenen Branchen zuzuordnen.⁶

⁶ Es könnte daran gedacht werden, diesen Personenkreis im Sinne einer Näherungslösung den Branchen entsprechend ihrer Beschäftigtenanteile in der betreffenden Region zuzuordnen. Da aber regelmäßig erhebliche Unterschiede in der Beschäftigungsentwicklung einzelner Branchen bestehen, die Wirtschaftszweige also in ganz unterschiedlichem Ausmaß zur Arbeitslosigkeit beitragen, dürfte ein solches Verfahren im Zweifel mit ganz erheblichen Fehlern behaftet sein.

4.3 Bereinigung um den Brancheneinfluß

Die Gründungsraten sind in den einzelnen Branchen in der Regel unterschiedlich hoch, was auf eine Reihe von Ursachen wie z.B. divergierende allgemeine Nachfrageentwicklung nach den betreffenden Gütern und/oder unterschiedlich hohe Marktzugangsbeschränkungen zurückzuführen ist. Aus diesem Grunde dürfte auch die Gründungsrate in einer Region insgesamt durch die dort vorhandene sektorale Wirtschaftsstruktur geprägt sein: Dementsprechend wäre bei einem relativ hohen Anteil von Betrieben in Branchen, die durch eine hohe Gründungsrate gekennzeichnet sind, auch eine überdurchschnittlich hohe regionale Gründungsrate zu erwarten; umgekehrt geht ein hoher Anteil an Betrieben in Branchen, die eine unterdurchschnittlich geringe Gründungsrate aufweisen, vermutlich mit einem relativ niedrigen Wert der Gründungsrate für die Region insgesamt einher.

Zur Berücksichtigung solcher Branchenstruktureffekte bietet sich die Bildung einer *bereinigten Gründungsrate* an, die im Sinne eines Shift-Share-Ansatzes den Regionalfaktor in bezug auf die Gründungsaktivitäten angibt. Um die um den Branchenstruktureffekt bereinigte Gründungsrate zu bestimmen, müssen zunächst die durchschnittlichen Gründungsraten in den einzelnen Wirtschaftszweigen auf gesamtwirtschaftlicher Ebene gebildet werden. Da nur der „betriebsökologische“ Ansatz eine sinnvolle Branchendifferenzierung der Gründungsraten zuläßt, ist diese durchschnittliche Gründungsrate als „Anzahl der Gründungen/Anzahl der Betriebe“ definiert. Anhand dieser durchschnittlichen Gründungsraten für die einzelnen Branchen lässt sich dann für jede Region eine *hypothetische Anzahl der Gründungen* berechnen, indem man die Anzahl der Betriebe in jeder Branche mit der durchschnittlichen Gründungsrate in dieser Branche multipliziert. Die hypothetische Anzahl der Gründungen gibt an, mit wieviel Neugründungen allein aufgrund der in der betreffenden Region bestehenden Branchenstruktur zu rechnen ist, wenn die Anzahl der Gründungen pro Branche dem gesamtwirtschaftlichen Durchschnitt entspräche. Die um den Branchenstruktureinfluss bereinigte Anzahl der Gründungen erhält man, indem man von der tatsächlichen Anzahl der Gründungen die hypothetische Anzahl an Gründungen subtrahiert. Dieser Wert lässt sich dann auf unterschiedliche Größen (z.B. Arbeitskräftepotential, Anzahl der vorhandenen Betriebe) beziehen. Anhand dieser bereinigten Gründungsrate lässt sich dann beurteilen, inwieweit die Gründungsaktivitäten in einer Region unter Berücksichtigung des Einflusses der regionalen Branchenstruktur über- oder unterdurchschnittlich ausfallen.

In die nachfolgend präsentierten Analysen gehen vier Varianten der regionalen Gründungsrate ein, wobei zum einen die Anzahl der Gründungen und zum anderen die um den Brancheneinfluss bereinigte Anzahl an Gründungen jeweils auf das regionale Arbeitskräftepotential bzw. auf die Anzahl der vorhandenen Betriebe bezogen wird. Es ergeben sich somit:

Gründungsrate A (Arbeitsmarktansatz):

$$\frac{\text{Anzahl der Gründungen}}{\text{Arbeitskräftepotential}}$$

Gründungsrate B (betriebsökologischer Ansatz):

$$\frac{\text{Anzahl der Gründungen}}{\text{Anzahl vorhandener Betriebe}}$$

Gründungsrate A bereinigt:

$$\frac{\text{Anzahl tatsächlicher Gründungen} - \text{Anzahl hypothetischer Gründungen}}{\text{Arbeitskräftepotential}}$$

Gründungsrate B bereinigt:

$$\frac{\text{Anzahl tatsächlicher Gründungen} - \text{Anzahl hypothetischer Gründungen}}{\text{Anzahl vorhandener Betriebe}}$$

5. Vergleich alternativer Gründungsraten

Zur Bildung einer um den Brancheneinfluß bereinigten regionalen Gründungsrate (siehe hierzu *Abschnitt 4.3*) muß zunächst die hypothetische Anzahl der Gründungen berechnet werden, die sich durch Multiplikation der Anzahl der in einer Region vorhandenen Betriebe einer bestimmten Branche mit der branchendurchschnittlichen Gründungsrate ergibt. Bei den hier dargestellten Berechnungen wurde eine Differenzierung der Gründungen und der vorhandenen Betriebe in 58 Branchen zugrundegelegt.

Ein Vergleich der tatsächlichen Anzahl der Gründungen mit der anhand der Branchenstruktur berechneten hypothetischen Anzahl neuer Betriebe zeigt ein hohes Maß an Übereinstimmung (vgl. *Abbildung 1* und *Tabelle 1*). So weicht bei Berechnungen über sämtliche Sektoren die hypothetische Anzahl der Gründungen von der tatsächlichen Anzahl in fast 70% der Fälle (Regionen) um weniger als 10% ab; in gut 85% der Fälle beträgt die Abweichung weniger als 15%. Die maximale Abweichung der tatsächlichen Zahl der Gründungen von der aufgrund der Branchenstruktur zu erwartenden Anzahl beträgt hier –30% bzw. +35%, wobei die Fälle mit großen Differenzen – wie *Abbildung 1* zeigt – rare Ausreißer darstellen. Ganz ähnlich auch die Ergebnisse der auf den Bereich der Industrie bzw. auf den Dienstleistungssektor beschränkten Berechnungen. Es zeigt sich also, daß die Anzahl der in einer Region stattfindenden Gründungen sehr weitgehend durch die in dieser Region gegebene Branchenstruktur bestimmt ist. Dies belegt die

Bedeutung einer Bereinigung der Gründungsrate um den Brancheneffekt im Rahmen interregionaler Vergleiche. Unterbleibt eine solche Bereinigung, dann besteht immer die Gefahr, daß sich in den Ergebnissen vorwiegend die bestehenden Unterschiede der regionalen Branchenstruktur und nicht die Unterschiede der Gründungsaktivitäten niederschlagen.

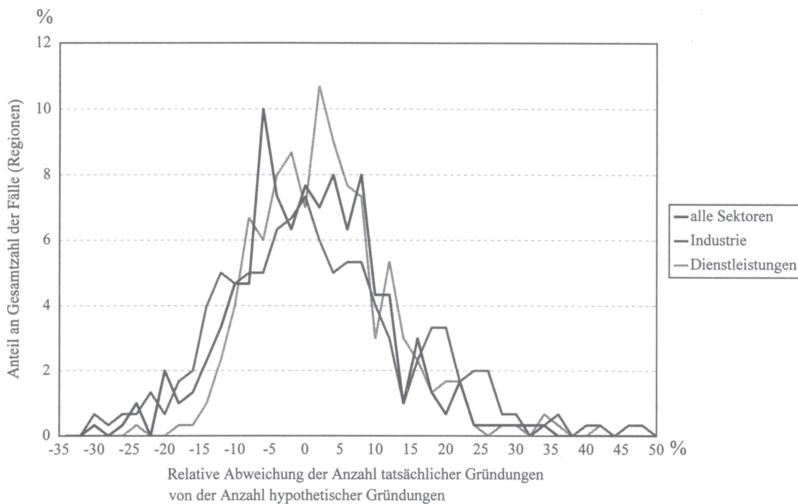


Abbildung 1: Prozentuale Abweichungen der tatsächlichen Anzahl der Gründungen von der hypothetischen Anzahl der Gründungen in den westdeutschen Raumordnungsregionen 1986 - 89

Tabelle 1
Kennzahlen zur Streuung der Differenz zwischen tatsächlichen und hypothetischen Gründungen*

Abweichung kleiner als	Alle Sektoren	Industrie	Dienstleistungen
± 1 %	7,6	7,3	7,0
± 5 %	36,3	31,3	43,3
± 10 %	69,6	57,6	75,0
± 15 %	85,3	73,6	89,6

* Gepoolte Gründungsraten für die Jahre 1986 - 89 (N=300). Angegeben ist jeweils der Anteil der Fälle (Regionen) in v.H.

In diesem Zusammenhang ist der statistische Zusammenhang zwischen den verschiedenen Gründungsraten aufschlußreich. *Tabelle 2* zeigt die entsprechenden Korrelationskoeffizienten der Gründungsraten über alle Sektoren sowie für die beiden großen Bereiche Industrie und Dienstleistungen. Die über sämtliche Sektoren gebildeten Gründungsraten sind sämtlich signifikant positiv miteinander korreliert, wobei zwischen der unbereinigten Gründungsrate A (Arbeitsmarktansatz) und den um den Brancheneinfluß bereinigten Gründungsraten nur ein vergleichsweise schwacher Zusammenhang besteht. Die weitaus höchste Korrelation ist für die Beziehung zwischen den beiden bereinigten Gründungsraten feststellbar.

Wie aus *Tabelle 3* hervorgeht, besteht ein enger statistischer Zusammenhang zwischen den Gründungsraten für die verschiedenen Jahre des Untersuchungszeitraums: Für Regionen, die in einem bestimmten Jahr eine relativ hohe Gründungsrate aufweisen, ist in der Regel auch in den darauffolgenden Jahren eine entsprechend hohe Rate zu verzeichnen. Sämtliche Korrelationskoeffizienten für die regionalen Gründungsraten in den verschiedenen Perioden weisen relativ hohe Werte auf und sind auf dem 1%-Niveau statistisch signifikant. Dies zeigt an, daß regionale Besonderheiten des Gründungsgeschehens nicht vorwiegend kurzfristigen Schwankungen unterliegen, sondern im Zeitablauf einigermaßen konstant sind und wohl in relativ starkem Maße auf den jeweiligen regionalen Kontextbedingungen beruhen.

6. Überblick über regionale Unterschiede des Gründungsgeschehens 1986 - 89

Entsprechend der hier gewählten Definition einer Gründung (vgl. Abschnitt 3) fanden in Westdeutschland während des Zeitraums 1986 - 89 zwischen 121 und 135 Tsd. Betriebsgründungen pro Jahr statt. Dies entspricht rund acht bis zehn Prozent des Betriebsbestandes, wobei – wie *Abbildung 2* zeigt – deutliche sektorale Unterschiede bestehen: Die Anzahl der Gründungen in Relation zur Anzahl der bestehenden Betriebe (Gründungsrate B) ist im Dienstleistungssektor mit neun bis elf Prozent deutlich höher als im Verarbeitenden Gewerbe, wo der entsprechende Wert lediglich bei fünf bis sechs Prozent liegt.⁷ Pro 1.000 Erwerbspersonen waren in jedem Jahr ca. 5,3 Gründungen zu verzeichnen.

Die Gründungsaktivitäten waren auf die Agglomerationen konzentriert, was offenbar durch das dort vorhandene, relativ große ökonomische Poten-

⁷ Deutlich mehr als die Hälfte aller Gründungen finden in den Raumordnungsregionen mit großen Verdichtungsräumen statt und nur etwa jede siebte Gründung im ländlich-peripheren Raum.

Tabelle 2
Korrelationen zwischen den alternativen Gründungsraten*

a: Alle Sektoren

Gründungsrate	B	A bereinigt	B bereinigt
A (Arbeitsmarkt- Ansatz)	0,341 ⁺⁺	0,178 ⁺⁺	0,176 ⁺⁺
B (betriebsökolo- gischer Ansatz)	-	0,709 ⁺⁺	0,730 ⁺⁺
A bereinigt	-	-	0,984 ⁺⁺

b: Industrie

Gründungsrate	B	A bereinigt	B bereinigt
A (Arbeitsmarkt- Ansatz)	0,106	0,117 ⁺	0,113 ⁺
B (betriebsökolo- gischer Ansatz)	-	0,923 ⁺⁺	0,963 ⁺⁺
A bereinigt	-	-	0,964 ⁺⁺

c: Dienstleistungen

Gründungsrate	B	A bereinigt	B bereinigt
A (Arbeitsmarkt- Ansatz)	0,523 ⁺⁺	0,253 ⁺⁺	0,327 ⁺⁺
B (betriebsökolo- gischer Ansatz)	-	0,752 ⁺⁺	0,774 ⁺⁺
A bereinigt	-	-	0,985 ⁺⁺

* Gepoolte Daten (N=300). ⁺⁺: statistisch signifikant auf dem 1%-Niveau; ⁺: statistisch signifikant auf dem 5%-Niveau.

Tabelle 3

Korrelationskoeffizienten für den Zusammenhang zwischen regionalen Gründungsraten in verschiedenen Jahren*

a: Alle Sektoren

Gründungsrate	Jahr t + 1	Jahr t + 2	Jahr t + 3
A (Arbeitsmarkt)	0,847	0,871	0,877
B (ökologisch)	0,582	0,610	0,679
A bereinigt	0,545	0,511	0,458
B bereinigt	0,564	0,555	0,495

b: Industrie

Gründungsrate	Jahr t + 1	Jahr t + 2	Jahr t + 3
A (Arbeitsmarkt)	0,794	0,762	0,762
B (ökologisch)	0,622	0,546	0,586
A bereinigt	0,569	0,498	0,514
B bereinigt	0,573	0,465	0,469

c: Dienstleistungen

Gründungsrate	Jahr t + 1	Jahr t + 2	Jahr t + 3
A (Arbeitsmarkt)	0,716	0,867	0,857
B (ökologisch)	0,201	0,298	0,494
A bereinigt	0,533	0,571	0,456
B bereinigt	0,539	0,581	0,490

* Gepoolte Daten. Jahr t + 1: 225 Fälle; Jahr t + 2: 150 Fälle; Jahr t + 3: 75 Fälle. Sämtliche Korrelationskoeffizienten sind auf dem 1 %-Niveau statistisch signifikant.

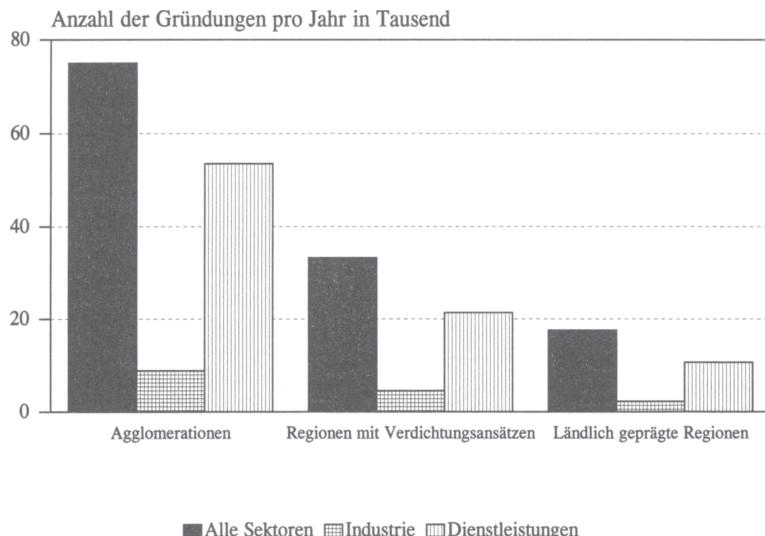
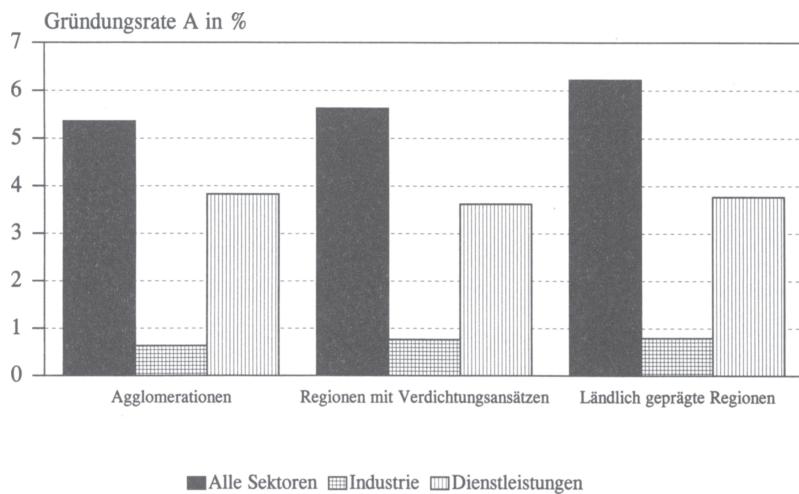


Abbildung 2: Durchschnittliche jährliche Anzahl an Gründungen nach Regionstypen

tial bedingt ist (*Abbildung 2*). Mehr als die Hälfte dieser Gründungen (59,7%) fand in Raumordnungsregionen „mit großen Verdichtungsräumen“ statt; der Anteil der Gründungen in „ländlich geprägten“ Raumordnungsregionen betrug lediglich 14 %. In sektoraler Hinsicht entfällt der ganz überwiegende Teil (ca. 68 %) der Neugründungen auf den Dienstleistungssektor. Der Anteil der Gründungen im Dienstleistungssektor ist in den Raumordnungsregionen „mit großen Verdichtungsräumen“ mit 71,4 % relativ hoch und in den „ländlich geprägten“ Raumordnungsregionen mit 60,6 % vergleichsweise niedrig; gut 62 % aller Dienstleistungs-Neugründungen fanden in Raumordnungsregionen „mit großen Verdichtungsräumen“ statt und lediglich 12 % haben einen Standort in „ländlich geprägten“ Raumordnungsregionen.

Bezieht man die Anzahl der Gründungen auf das in den Regionen vorhandene Arbeitskräftepotential (Gründungsrate A), dann ergibt sich für die Agglomerationen ein unterdurchschnittliches Niveau der Gründungsaktivität, während die Gründungsraten für den ländlich-peripheren Raum relativ hoch ausfallen (*Abbildung 3a*). Standardisiert man die Gründungen hingegen mit der Anzahl der vorhandenen Betriebe (Gründungsrate B), so ergibt sich das umgekehrte Muster, denn nun weisen die Verdichtungsgebiete jeweils die höchste und die ländlich geprägten Regionen weisen die niedrig-

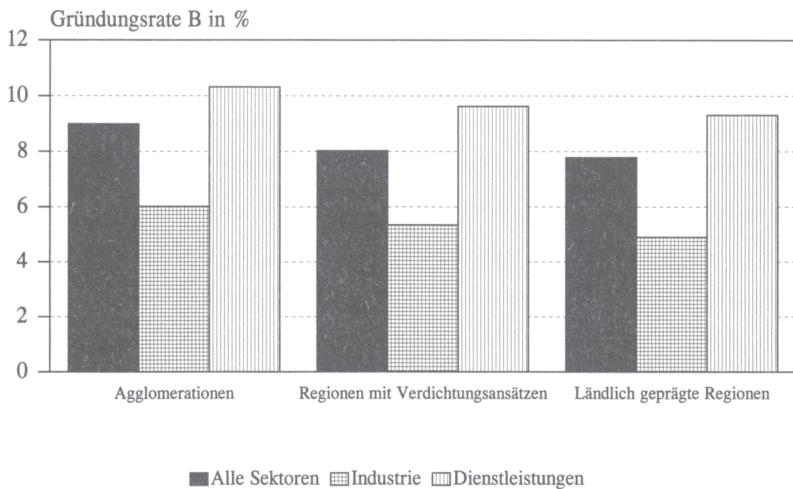
ste Gründungsrate auf (*Abbildung 3b*).⁸ Die höchsten Werte für die um den Brancheneinfluß bereinigten Gründungsraten sind für die Verdichtungsgebiete zu verzeichnen; hier sind die Gründungsaktivitäten in Anbetracht der Branchenstruktur also überdurchschnittlich hoch (siehe *Abbildungen 3c und d*).⁹ Der ländliche Raum ist hingegen durch negative Abweichungen von der aufgrund der Branchenstruktur zu erwartenden Anzahl der Gründungen gekennzeichnet; es zeigt sich also ein deutliches Zentrum-Peripherie-Gefälle hinsichtlich der Gründungsaktivitäten.



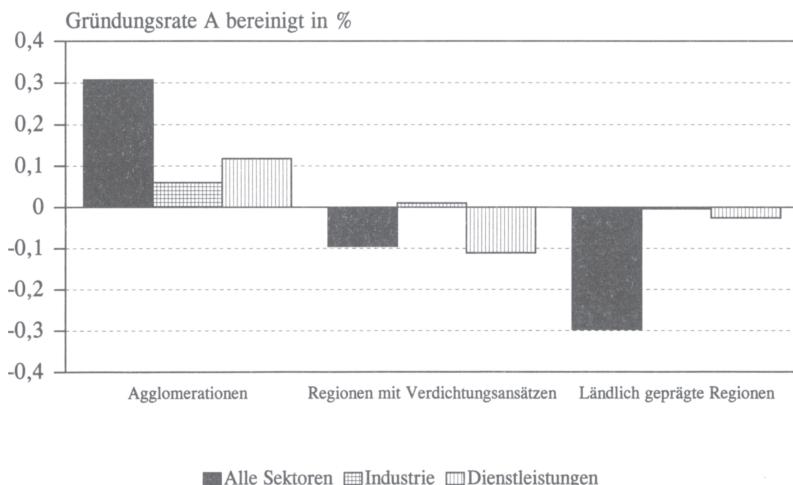
**Abbildung 3: Gründungsraten nach Regionstypen
a: Gründungsrate A (Anzahl der Gründungen/Arbeitskräftepotential)**

⁸ Die Ursache für diesen Unterschied besteht offensichtlich darin, daß die Verdichtungsregionen eine höhere durchschnittliche Betriebsgröße aufweisen und dort somit pro Betrieb eine größere Anzahl potentieller Gründer tätig ist. Ausführlicher hierzu *Audretsch/Fritsch* (1994a).

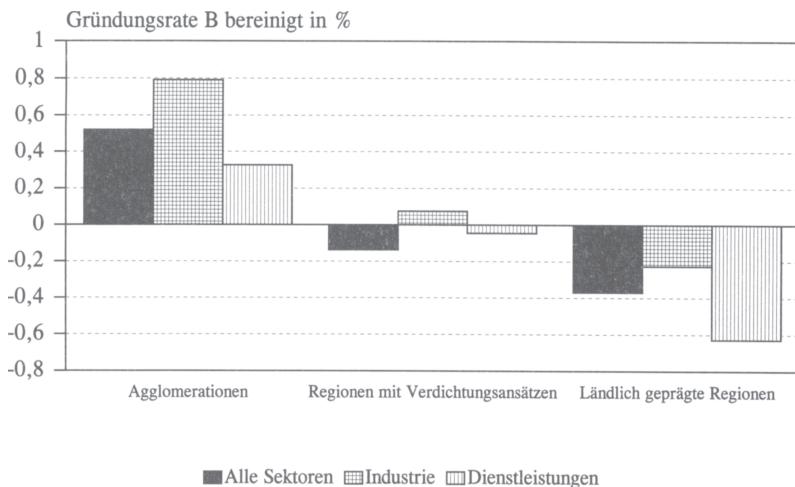
⁹ Wie kaum anders zu erwarten, schlagen sich die verschiedenen Arten der Formulierung einer Gründungsrate auch in den Ergebnissen von Analysen hinsichtlich der Bestimmungsgründe für regionale Unterschiede in den Gründungsraten nieder. Siehe zu entsprechenden Analysen *Fritsch* (1992, 1994) sowie *Audretsch/Fritsch* (1994a und b).



b: Gründungsrate B (Anzahl der Gründungen/Anzahl vorhandener Betriebe)



c: Gründungsrate A bereinigt (Bereinigte Anzahl der Gründungen/Arbeitskräftepotential)



d: Gründungsrate B bereinigt (Bereinigte Anzahl der Gründungen/Anzahl vorhandener Betriebe)

7. Der Einfluß des Gründungsgeschehens auf die regionale Arbeitsplatzentwicklung

Welcher Zusammenhang besteht nun zwischen den Gründungsaktivitäten und der regionalen Arbeitsplatzentwicklung in den nachfolgenden Jahren? Tabelle 4 zeigt Korrelationskoeffizienten für die Beziehung zwischen den verschiedenen Gründungsraten und der prozentualen Veränderung der Beschäftigtenzahl sowohl für alle Sektoren als auch jeweils für den Industrie- und den Dienstleistungssektor. Betrachten wir zunächst die Ergebnisse für sämtliche Wirtschaftsbereiche. Zunächst fällt auf, daß die Korrelationskoeffizienten für den Zusammenhang zwischen den nicht um den Brancheneinfluß bereinigten Gründungsraten und der Beschäftigungsentwicklung entweder statistisch nicht signifikant sind oder in der Mehrzahl der Fälle signifikant negative Werte aufweisen. Für die bereinigten Gründungsraten sind – abgesehen vom Jahr der Gründung selbst – sogar ausschließlich negative Werte zu verzeichnen.

Deutlich anders stellen sich die Zusammenhänge bei den auf den Industriebereich beschränkten Berechnungen dar. Hinsichtlich der nicht um den Brancheneinfluß bereinigten Gründungsraten zeigt sich ein diffuses und widersprüchliches Bild; teilweise ergeben sich für die Gründungsraten A und B jeweils signifikante Koeffizientenwerte mit entgegengesetzten Vorzeichen. Demgegenüber sind die Ergebnisse für die bereinigten Gründungs-

Tabelle 4

Korrelationskoeffizienten für den Zusammenhang zwischen regionaler Gründungsrate und Beschäftigungsentwicklung (gepoolte Daten)*

a: alle Sektoren

Grün-dungs-rate	Beschäftigungsentwicklung (%)					
	Jahr 0	Jahr t+1	Jahr t+2	Jahr t+3	Jahr t+4	Jahr t+5
A	0,119*	-0,054	-0,056	0,087	0,059	0,083
B	-0,095	-0,343**	-0,264**	-0,031	-0,205*	-0,138
A ber.	0,035	-0,094	-0,234**	-0,372**	-0,210**	-0,158
B ber.	0,020	-0,120*	-0,255**	-0,389**	-0,249**	-0,188

b: Industrie

Grün-dungs-rate	Beschäftigungsentwicklung (%)					
	Jahr 0	Jahr t+1	Jahr t+2	Jahr t+3	Jahr t+4	Jahr t+5
A	0,217**	0,134*	-0,074	-0,077	-0,106	-0,196
B	-0,202**	-0,194**	0,033	0,140*	0,225**	0,388**
A ber.	-0,105	-0,039	0,139*	0,184**	0,184*	0,350**
B ber.	-0,128*	-0,060	0,153**	0,204**	0,212**	0,400**

c: Dienstleistungen

Grün-dungs-rate	Beschäftigungsentwicklung (%)					
	Jahr 0	Jahr t+1	Jahr t+2	Jahr t+3	Jahr t+4	Jahr t+5
A	0,130*	-0,120*	-0,240**	-0,043	-0,209*	-0,236*
B	0,256**	-0,110	-0,229**	0,156*	-0,084	-0,169
A ber.	0,173**	0,134*	0,035	-0,018	-0,052	-0,149
B ber.	0,161**	0,114*	0,021	-0,024	-0,065	-0,173

* **: statistisch signifikant auf dem 1%-Niveau; *: statistisch signifikant auf dem 5%-Niveau. Jahr 0 bis Jahr t+2: 300 Fälle; Jahr t+3: 225 Fälle; Jahr t+4: 150 Fälle; Jahr +5: 75 Fälle.

raten ziemlich eindeutig: Hier ist der Zusammenhang mit der Arbeitsplatzentwicklung im Jahr der Gründung sowie im darauffolgenden Jahr zunächst negativ; für die Zeit ab dem zweiten Jahr nach der Gründung ist dann ein deutlicher und im Zeitablauf zunehmender positiver Zusammenhang mit der Arbeitsplatzentwicklung im Industriesektor zu verzeichnen. Besonders bemerkenswert ist, daß sich positive Wirkungen der Gründungsaktivitäten auf die Beschäftigungsentwicklung in der Regel erst zwei bis drei Jahre nach der Gründung zeigen. Dieses Muster läßt sich als Hinweis darauf auf-fassen, daß die längerfristigen bzw. die indirekten Beschäftigungseffekte der Gründungen im Industriesektor bedeutender sind als die direkten Arbeitsplatzwirkungen im näheren zeitlichen Umfeld der Gründung.

Im Dienstleistungssektor ist der positive Zusammenhang zwischen Gründungen und regionaler Beschäftigungsentwicklung vor allem auf das Jahr der Gründung und die darauffolgende Periode beschränkt. Während die nicht um den Brancheneinfluß bereinigten Gründungsraten schon ab dem auf die Gründung folgenden Jahr überwiegend einen negativen Zusammenhang mit der Beschäftigungsentwicklung aufweisen, zeigt sich für die bereinigten Gründungsraten und der regionalen Arbeitsplatzentwicklung im Dienstleistungssektor auch noch für das Jahr nach der Gründung ein signifikant positiver Zusammenhang. Der nur schwach ausgeprägte und lediglich kurzfristige Einfluß von Gründungen im Dienstleistungssektor auf die Arbeitsplatzentwicklung kann als Reflex der Tatsache angesehen werden, daß ein großer Teil der Dienstleistungsbetriebe relativ klein bleibt und nicht in wesentlichem Ausmaß Arbeitsplätze schafft. Eine weitere Erklärung bestünde darin, daß viele Dienstleistungsaktivitäten dem Nichtbasis-Bereich der Regionalwirtschaft (Betriebe, deren Absatz auf die Region beschränkt ist) zuzuordnen sind, weshalb ihr Einfluß auf die Regionalentwicklung deutlich geringer ausfällt als der von Gründungen im (über die Regionsgrenzen liefernden) Basis-Bereich, dem vermutlich ein großer Anteil der Gründungen im Industriesektor angehört.

Vor allem die zu verzeichnenden negativen Werte der Korrelationskoeffizienten für den Zusammenhang zwischen Gründungsrate und Arbeitsplatzentwicklung während der darauffolgenden Jahre sind bemerkenswert; widerspricht ein solcher negativer Zusammenhang doch der Tatsache, daß durch neugegründete Betriebe in der Regel zusätzliche Arbeitsplätze entstehen. Die Erklärung für diesen vermeintlichen Widerspruch besteht wohl darin, daß während des hier betrachteten Zeitraumes der positive Einfluß von Betriebsgründungen auf die Arbeitsplatzentwicklung von anderen Faktoren überkompensiert wurde. Zu vermuten ist, daß insbesondere die Gründungen im Dienstleistungssektor Teil eines Reorganisationsprozesses darstellen, bei dem Aktivitäten und Arbeitsplätze primär in anderer Weise auf Betriebe aufgeteilt werden.

Nicht zuletzt belegen die Ergebnisse, wie wichtig die Bereinigung um den Brancheneinfluß ist, denn erst die bereinigten Gründungsraten zeigen klar interpretierbare Muster des Zusammenhangs zwischen Gründungsaktivitäten und regionaler Arbeitsplatzentwicklung.

8. Schlußfolgerungen

Die hier dargestellten Analysen legen Schlußfolgerungen sowohl für zukünftige empirische Analysen als auch für die Wirtschaftspolitik nahe.

Zunächst einmal zeigte sich, daß die Anzahl der in einer Region stattfindenden Gründungen sehr stark durch die dort vorhandene Branchenstruktur geprägt ist, so daß der ganz überwiegende Anteil der Gründungen mit der regionalen Branchenstruktur erklärt werden kann. Für interregionale Vergleiche von Gründungsraten ist daher eine Bereinigung um diesen Brancheneffekt erforderlich.

Die Ergebnisse zum Zusammenhang zwischen Gründungsaktivitäten und regionaler Arbeitsplatzentwicklung sind geeignet, die Gründungsförderung als wachstums- und beschäftigungspolitische Strategie in Frage zu stellen. Dies betrifft vor allem eine unspezifische Förderung von Unternehmensgründungen, die gewissermaßen „mit der Gießkanne“ vorgeht. Eine bei den Gründungen ansetzende Strategie der Wirtschaftsförderung sollte selektiv sein und versuchen, die Hilfen möglichst auf „arbeitsplatzzeugende“ Gründungen zu konzentrieren. Dies gilt insbesondere für Neugründungen im Dienstleistungssektor. Daneben darf die Pflege des „endogenen“ Potentials der Regionen, also der bereits vorhandenen Betriebe, nicht vernachlässigt werden. Denn alles in allem deuten die Ergebnisse darauf hin, daß die Bedeutung der Gründungen für die Regionalentwicklung – zumindest während der hier betrachteten Periode – geringer ausfiel, als gemeinhin angenommen wird.

Literatur

- Audretsch, D. B./Fritsch, M. (1994a), On the Measurement of Entry Rates, in: Empirica, Vol. 21, S. 105 - 113.*
- (1994b), The Geography of Firm Births in Germany, in: Regional Studies, Vol. 27.
- (1994c), Neugründungen, Stillegungen, Turbulenz und regionale Beschäftigungsentwicklung, in: Frick, B./Semlinger, K. (Hrsg.), *Betriebliche Modernisierung und Fluktuation*, Berlin (edition sigma).
- Baumol, W. J. (1990), Entrepreneurship: Productive, Unproductive, and Destructive, in: Journal of Political Economy, Vol. 98, S. 893 - 921.*

- Birch, D.* (1979), *The Job Generation Process*, Cambridge/Mass. (mimeo).
- Boeri, T. M./Cramer, U.* (1991), Betriebliche Wachstumsprozesse – eine statistische Analyse mit der Beschäftigtenstatistik 1977 - 1987, in: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 24. Jg., S. 70 - 80.
- Brüderl, J./Bühler, Ch./Ziegler, R.* (1993), Beschäftigungswirkungen neugegründeter Betriebe, in: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 26. Jg., S. 521 - 528.
- Cramer, U./Koller, M.* (1988), Gewinne und Verluste von Arbeitsplätzen in Betrieben – der „Job-Turnover“-Ansatz, in: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 21. Jg., S. 361 - 377.
- Dunne, P./Roberts, M. J./Samuelson, L.* (1989), Plant turnover and gross employment flows in the US manufacturing sector, in: *Journal of Labor Economics*, Vol. 7, S. 48 - 71.
- Fritsch, M.* (1990a), *Arbeitsplatzentwicklung in Industriebetrieben*, Berlin (de Gruyter).
- (1990b), Zur Bedeutung des kleinbetrieblichen Sektors für die Regionalpolitik, in: Berger, J./Domeyer, V./Funder, M. (Hrsg.), *Kleinbetriebe im wirtschaftlichen Wandel*, Frankfurt a.M./New York (Campus), S. 241 - 268.
- (1992), Regional Differences in New Firm Formation: Evidence from West Germany, in: *Regional Studies*, Vol. 25, S. 233 - 241.
- (1994), Determinants of New Firm Formation in West German Regions 1986 - 1989: An Empirical Analysis, in: Witteloostuijn, A. v. (Ed.), *Market Evolution – Competition and Cooperation across Markets and over Time*, Dordrecht/Boston (Kluwer Academic Publishers), S. 283 - 298.
- Fritsch, M./König, A./Weißhuhn, G.* (1994), Die Beschäftigtenstatistik als Betriebspanel – Ansatz, Probleme und Analysepotentiale, in: Hochmuth, U./Wagner, J. (Hrsg.), *Firmenpanelstudien in Deutschland – Konzeptionelle Überlegungen und empirische Analysen*, Tübingen (Tübinger Volkswirtschaftliche Schriften, Francke), S. 65 - 82.
- Gerlach, K./Wagner, J.* (1993), Gross and Net Employment Flows in Manufacturing Industries, in: *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 113. Jg, S. 17 - 28.
- (1994), Entries, Exits, and the Dynamics of Employment: Lower Saxony, 1978 - 1990, in: Buttler, F./Schwarze, J./Wagner, G. (Eds.), *Labor Market Dynamics in Present Germany*, Frankfurt a.M./New York (Campus), S. 240 - 256.
- Gudgin, G.* (1978), *Industrial Location Processes and Regional Employment Growth*, Westmead (Saxon House).
- König, A.* (1994), *Betriebliche Beschäftigungsdynamik und personeller Strukturwandel – Eine Longitudinalanalyse*, Frankfurt a.M./New York (Campus).
- O'Farrell, P.* (1986), *Entrepreneurs and Industrial Change*, Dublin (Irish Management Institute).

- Reynolds, P. D./Maki, W. R. (1990), Business Volatility and Economic Growth (Final Project Report). Minneapolis (mimeo).*
- Reynolds, P. D. (1994), Autonomous Firm Dynamics and Economic Growth in the United States: 1986 - 90, in: Regional Studies, Vol. 27.*
- Storey, D. J./Johnson, S. (1987), Job Generation and Labour Market Change, London (MacMillan).*

Zusammenfassung der Diskussion

Die Diskussion konzentrierte sich auf Datenprobleme und wirtschaftspolitische Implikationen. Für die Wirtschaftspolitik läßt sich aus den Ergebnissen der Schluß ziehen, daß eine Förderung nach dem „Gießkannenprinzip“ sicherlich nicht als optimal angesehen werden kann. Erfolgt die Förderung dagegen nach einer „Pick the Winner“ – Strategie, und versucht relativ erfolgreiche Neugründungen zu fördern, so stellt sich das Problem, solche Fälle vorab identifizieren zu müssen. Zu fragen wäre daher, welche Potentiale für eine „Make more Winners“ – Strategie bestehen, mit welcher die Anzahl der erfolgreichen Neugründungen erhöht werden soll.

Schwächen der verwendeten Datenbasis können unter Umständen darin gesehen werden, daß Betriebe als Erhebungseinheit verwendet wurden, und eine Zurechnung zu Unternehmen nicht möglich ist. Weitere Analysepotentiale werden in einer stärkeren Disaggregation der Betrachtung (z. B. auf Brancheebene) sowie in zeitlich längeren Beobachtungen gesehen. Betont wurde auch, daß es sich bei der Analyseperiode um eine Phase relativer Hochkonjunktur handelte und sich daher der Zusammenhang zwischen Arbeitslosenquoten und Gründungswahrscheinlichkeit bei rezessiver gesamtwirtschaftlicher Lage auch anders darstellen könnte.

Die interessante Frage, ob hier evtl. der Schumpetersche innovative Unternehmer identifizierbar wäre, muß unter den vorliegenden, aus den Daten gewonnenen Erkenntnissen eher verneint werden.