Schriften des Vereins für Socialpolitik

Band 227

Beiträge zur Standortforschung

Von

Horst Behnke, Johannes Bröcker, Hans Joachim Schalk, Klaus Schöler, Horst Todt, Reiner Wolff

Herausgegeben von Horst Todt

Duncker & Humblot · Berlin

Schriften des Vereins für Socialpolitik

Gesellschaft für Wirtschafts- und Sozialwissenschaften

Neue Folge Band 227

SCHRIFTEN DES VEREINS FÜR SOCIALPOLITIK
Gesellschaft für Wirtschafts- und Sozialwissenschaften
Neue Folge Band 227

Beiträge zur Standortforschung

Duncker & Humblot · Berlin

Beiträge zur Standortforschung

Von

Horst Behnke, Johannes Bröcker, Hans Joachim Schalk, Klaus Schöler, Horst Todt, Reiner Wolff

Herausgegeben von Horst Todt

Duncker & Humblot · Berlin

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Beiträge zur Standortforschung / von Horst Behnke ... Hrsg. von Horst Todt. – Berlin : Duncker und Humblot, 1994
(Schriften des Vereins für Socialpolitik, Gesellschaft für Wirtschafts- und Sozialwissenschaften ; N. F., Bd. 227)
ISBN 3-428-07881-0
NE: Behnke, Horst; Todt, Horst [Hrsg.]; Gesellschaft für Wirtschafts- und Sozialwissenschaften: Schriften des Vereins ...

Druck: Berliner Buchdruckerei Union GmbH, Berlin
Printed in Germany

ISSN 0505-2777
ISBN 3-428-07881-0

Vorwort

Der vorliegende Band gibt einige Vorträge wieder, die anläßlich der Sitzung des Regionalwissenschaftlichen Ausschusses im Herbst 1992 gehalten wurden. Die Zusammenkunft stand nicht unter einem Motto; es gab kein Generalthema und es gab auch kein Raster von Themen, auf das sich der Ausschuß eingeengt hätte. Vielmehr haben die Mitglieder des Gremiums und einschlägig interessierte Gäste aus ihrer Arbeit vorgetragen. Diese Konzeption hat Vor- und Nachteile. Es darf einerseits nicht erwartet werden, daß alle wichtigen oder aktuellen Felder der Regionalwissenschaft vertreten sind, noch kann die Auswahl unter irgendeinem relevanten Aspekt repräsentativ sein; auch engere Teilgebiete sind so nicht systematisch zu erfassen. Andererseits darf man hoffen, daß die Fachvertreter genau die Themen behandeln, die sie zur Zeit für wichtig halten; ein Optimist mag hoffen, daß dies auch das tatsächlich Wichtige sei. Eine solche Hoffnung gab auch den Ausschlag für den Weg des Ausschusses.

Bei dieser Sachlage mag es überraschen, daß die Referate doch ganz gut zusammenpassen. Vor allem gehören alle Vorträge in den theoretischen Bereich.

Zwei der Arbeiten sind theoretisch fundierte empirische Untersuchungen, die trotz großer methodischer Differenz engen inhaltlichen Bezug haben und aktuelle Probleme im wiedervereinigten Deutschland betreffen. Es sind dies die Beiträge von Herrn J. Bröcker (Kapitalakkumulation und Migration im vereinigten Deutschland – Ein dynamisches Zwei-Regionen-Modell), der direkt die Probleme der Vereinigung angeht, und Herrn H. J. Schalk (Technische Effizienz und Kapitalmobilität in der Verarbeitenden Industrie: Ein interregionaler Vergleich für die Bundesrepublik Deutschland), dessen auf die alten Länder beschränkte Daten hohe Relevanz auch für Gesamtdeutschland besitzen dürften.

Die übrigen Arbeiten, die in diesem Band zur Diskussion gestellt werden, gehören der reinen Theorie an. Herrn R. Wolffs Beitrag (Strategien der Investitionspolitik einer Region: Der Fall des Wachstums mit konstanter Sektorstruktur) schließt thematisch an die vorgenannten Untersuchungen an, während die beiden folgenden Aufsätze von H. Behnke (Ein Erklärungsansatz für unterschiedliche Veränderungen in der Standortstruktur von Geschäften innerhalb verschiedener Städte bei gleichen Veränderungen in den Rahmenbedingungen) und H. Todt (Wettbewerb durch Standortwahl in der Fläche – Die Cournot- und die Stackelberg-Lösung) Probleme der reinen Standorttheo-

rie behandeln. Eine weitere wettbewerbstheoretische Untersuchung wurde von K. Schöler (Zur Anwendung des Konzepts konsistenter konjekturaler Reaktionen auf die Theorie der räumlichen Preisbildung) beigesteuert.

Wenn der Ausschuß hier theoretische Fragen erörtert, so bedeutet dies keine Abkehr von wirtschaftspolitisch – institutionellen Problemen. Vielmehr wird die Absicht postuliert, auch die Theorie nicht zu vernachlässigen.

Horst Todt

Inhaltsverzeichnis

Kapitalakkumulation und Migration im vereinigten Deutschland

Ein dynamisches Zwei-Regionen-Modell

Von *Johannes Bröcker*, Dresden

I. Einleitung

Zwei Jahre nach der deutschen Vereinigung befinden wir uns noch mitten in einem stürmischen Übergangsprozeß, in dem überlebte Strukturen verschwinden und neue entstehen, ohne daß wir in unseren Beobachtungen den einen von dem anderen Prozeß klar zu trennen vermögen. Bei der Produktion lokaler Güter ist der Wachstumsprozeß bereits voll im Gange, während in industriellen Basisbereichen noch der schmerzliche Tod des Todgeweihten das Bild bestimmt. Trotz zunehmender Skepsis herrscht weitgehende Einigkeit darüber, daß es auf mittlere Sicht auf breiter Front zu einem Aufholprozeß in Ostdeutschland kommen wird. Aber welche Zeit er in Anspruch nehmen wird, mit welchen Wanderungsbewegungen und welchen finanziellen Lasten für heutige und zukünftige Generationen er verbunden sein wird, darüber läßt sich nur spekulieren.

Ich stelle in diesem Papier eine Modellstudie vor, mit der ich versuche, mich einer Antwort auf folgende vier Fragen zu nähern:

1. Wie lange dauert es unter günstigen Voraussetzungen, bis Ostdeutschland annähernd den Westen aufgeholt hat?
2. Wie wird der Anpassungspfad durch die Festlegung von Tariflöhnen und Subventionierung von Faktorkosten beeinflußt?
3. Welche Bevölkerungsverschiebungen[1] sind im Laufe der Anpassung zu erwarten?
4. Welche finanziellen Lasten sind mit unterschiedlichen Übergangsstrategien verbunden?

Die Antworten sollen aus der numerischen Auswertung eines empirisch gestützten dynamischen Gleichgewichtsmodells gewonnen werden. Die grundlegenden Charakteristika dieses Modells sind folgende:

---

[1] *Raffelhüschen* (9) hat eine statische Abschätzung der nach der Vereinigung zu erwartenden Wanderungen vorgenommen. Seiner Arbeit verdanke ich wesentliche Anregungen.

1. Das Modell ist ein intertemporales neoklassisches Konkurrenzmodell. Als Akteure treten Firmen, Haushalte und ein Staat auf. Firmen sind profitmaximierende, Haushalte sind nutzenmaximierende Mengenanpasser. Firmen und Haushalte haben vollständige Voraussicht. Flexible Faktorpreise sorgen für die Räumung der Faktormärkte – mit einer Ausnahme: In den Ergebnissen werde ich u.a. ein Interventionsszenario präsentieren, in welchem die ostdeutschen Löhne durch Tarifvereinbarungen auf ein Mindestniveau fixiert werden, das höher als der markträumende Lohn ist.
2. Im Modell gibt es nur ein einziges produziertes, investiertes, konsumiertes und international gehandeltes Gut sowie zwei Faktoren, Arbeit und Kapital.
3. Das Modell ist ein reines Realmodell. Faktorpreise sind reale Austauschverhältnisse. Insbesondere ist der Zins das reale Austauschverhältnis zwischen Gegenwarts- und Zukunftsgut.
4. Im Modell wird eine kleine offene Volkswirtschaft mit zwei Regionen, West- und Ostdeutschland dargestellt. „Kleine offene Volkswirtschaft" heißt, daß zum herrschenden Produktpreis in beliebiger Höhe exportiert oder importiert werden kann und daß zum herrschenden Zins auf dem Weltkapitalmarkt beliebig geliehen oder verliehen werden kann.
5. Die Haushaltsseite wird als ein Kontinuum überlappender Generationen (OLG-Ansatz) modelliert. Im Prinzip kennen wir in der neoklassischen dynamischen Analyse zwei Ansätze, den Haushaltssektor darzustellen, nämlich einerseits *Ramseys* Ansatz des ewig lebenden repräsentativen Haushaltes (10), und andererseits den OLG-Ansatz, der zuerst von *Allais* (2) vorgeschlagen wurde. Dieser ist hier der geeignetere.

II. Das Modell

Ich beginne mit dem *Produktionssektor*. Er wird durch zwei repräsentative Firmen dargestellt, eine im Osten und eine im Westen. Beide Firmen produzieren dasselbe Gut mit den Faktoren Arbeit und Kapital mit Hilfe einer neoklassischen Technologie. Diese wird, der Tradition folgend, durch eine linear homogene CD-Funktion dargestellt:

$$Y = K^{\varepsilon} L^{1-\varepsilon} e^{\gamma t}.$$

Hier bezeichnen Y den Output, K und L die eingesetzten Mengen an Kapital und Arbeit, ε die partielle Produktionselastizität des Kapitals und γ die Rate des exogenen technischen Fortschritts. Alle Variablen bis auf den Zins sind Funktionen sowohl der Zeit als auch der Region[2]. Die Parameter dage-

[2] Der Übersichtlichkeit halber wird soweit möglich auf Regions- und Zeitindices verzichtet.

gen sind in der Zeit konstant und für beide Regionen identisch. Damit ist implizit unterstellt, daß man für jede neu investierte Mark im Osten auf dieselbe Technologie wie im Westen Zugriff hat. Der bereits vorhandene Kapitalstock im Osten ist jedoch größtenteils obsolet, und die Produktivität ist entsprechend gering. In der CD-Funktion kann man das einfach auf die Weise darstellen, daß für den Osten zum Vereinigungszeitpunkt eine geringere Kapitalausstattung pro Kopf als für den Westen angenommen wird. Obgleich es in Wirklichkeit noch vielerlei andere Gründe geben mag, weswegen der Osten hinter dem Westen in seiner Produktivität zurückbleibt, wird hier das Problem alleine auf das Fehlen eines modernen Kapitalbestandes reduziert.

Die Beschreibung der Technologie ist damit noch nicht beendet. Würde man es, was die Beschreibung der Technologie betrifft, bei dem bisher gesagten belassen, würde sich ein vollkommen abwegiges Resultat einstellen: Im Moment der Vereinigung wäre das Grenzprodukt des Kapitals im Osten weit höher als der Zins, und es gäbe keine Veranlassung, mit der Anpassung der östlichen Kapitalintensität an das Westniveau, bei welchem Zins und Grenzprodukt gleich sind, auch nur einen Moment zu zögern. Es gäbe einen einmaligen Kapital*schub,* keinen über die Zeitachse verteilten Investitions*strom.* Es gäbe nach dieser Spezifikation mit anderen Worten keine Investitionsfunktion im eigentlichen Sinne. Ich folge deswegen der mikroökonomisch fundierten Version der Tobinschen q-Theorie der Investition (7). In dieser Theorie unterstellt man, daß durch die Investition Anpassungskosten entstehen. Sie sind um so höher, je schneller der Kapitalbestand wächst. Die Anpassungskosten sind in einer Installationsfunktion ausgedrückt, welche angibt, wie groß der Flow an Gütern und Diensten ist, der erforderlich ist, um einen gegebenen Kapitalbestand mit einer gegebenen Rate wachsen zu lassen. Wir wählen hier eine in der Literatur übliche quadratische Form:

$$I = \left(1 + \eta \hat{K}\right) \dot{K}.$$

I ist der Aufwand an investierten Gütern, $\dot{K}$ der Zuwachs des Kapitalbestandes pro Zeiteinheit und $\hat{K} := \dot{K} / K$ die Wachstumsrate des Kapitalbestandes. η ist der sogenannte Anpassungskosten-Parameter, der die Systemdynamik wesentlich bestimmt. Nach dieser Spezifikation gilt das übliche $I = \dot{K}$ nur für den Grenzfall $\hat{K} = 0$. Je schneller das Kapital wächst, um so mehr übersteigt I den Zuwachs des Kapitalbestandes $\dot{K}$, weil Anpassungskosten auftreten. Die Anpassungskosten bewirken, daß es vorteilhaft ist, den gewünschten Zuwachs des Kapitalbestandes nicht auf einmal, sondern verteilt über die Zeitachse zu realisieren.

Die Firma wählt einen Zeitpfad ihrer Investition und der Beschäftigung derart, daß der Gegenwartswert ihrer Nettoerträge maximal wird. Ihr sind ein bestimmter Ausgangskapitalbestand K^0 sowie die Installationsfunktion als Restriktionen vorgegeben. Da ich in einem der später präsentierten Szenarien

auch die Möglichkeit der Investitionsförderung vorsehe, ist im Firmenkalkül auch der eventuelle Fördersatz zu berücksichtigen. Die Firma löst also das Problem

$$\max_{L,\, I} \int_0^\infty [K^\varepsilon L^{1-\varepsilon} e^{\gamma t} - wL - (1-h)I]e^{-it}\mathrm{d}t$$

unter den Nebenbedingungen

$$I = (1 + \eta\hat{K})\,\dot{K},$$
$$K(0) = K^0.$$

w ist der Lohnsatz und i der Zins. Das eine produzierte Gut ist der Numeraire des Systems. Sein Preis ist zu jeder Zeit = 1. Wegen der Annahme „kleine offene Volkswirtschaft" ist der Zins über die Zeit konstant, und wegen der Annahme eines integrierten Kapitalmarktes ist er in beiden Regionen und im Rest der Welt identisch, wenngleich – wohlgemerkt – das partielle Grenzprodukt des Kapitals im kapitalknappen Osten höher als im Westen ist. h ist der Fördersatz für Investitionen. Er wird nur in Ostdeutschland gewährt, und zwar nur solange, bis das ostdeutsche sich dem westdeutschen Niveau angepaßt hat.

Wegen der Konkavität des Problems sind die folgenden aus dem Maximumprinzip hergeleiteten Bedingungen notwendig und hinreichend für das Optimum. (f_L und f_K bezeichnen die partiellen Grenzprodukte von Arbeit und Kapital.)

$$w = f_L, \tag{1}$$

$$\hat{K} = \frac{q - 1 + h}{2\eta(1-h)}, \tag{2}$$

$$\dot{q} = iq - f_K - (1-h)\eta\hat{K}^2, \tag{3}$$

$$K(0) = K^0, \tag{4}$$

$$0 = \lim_{t \to \infty} K(t)q(t)e^{-it}. \tag{5}$$

(1) ist die vertraute Bedingung „Lohnsatz = Grenzprodukt der Arbeit". In den weiteren Bedingungen tritt das Tobinsche q auf, das sich aus der Cozustandsvariable herleitet, die zur Zustandsvariable K gehört. q ist als laufender Marktpreis oder Kurswert des installierten Kapitals zu interpretieren, d.h. im Gleichgewicht muß für das Anteilseigentum an einer Einheit des installierten Kapitals ein Betrag q – gemessen in Einheiten des Numeraire-Gutes – gezahlt werden. (2) ist Tobins Investitionsfunktion. Danach ist die Wachstumsrate des Kapitals eine wachsende Funktion von q und h. Bedingung (3) ist die Non-Arbitrage-Bedingung. Sie sagt, daß man beim Erwerb bestehenden Real-

kapitals mit diesem Kapital unter Berücksichtigung des Kursverlustes gerade den herrschenden Marktzins realisiert. Integriert man (3) unter der Restriktion (5), dann erkennt man, daß q nichts anderes als der Barwert der zukünftigen Grenzerträge des Kapitals ist (siehe [4, S. 62]). (4) ist die Anfangswertbedingung von K, (5) ist die Transversalitätsbedingung. Sie stellt sicher, daß auf dem Markt für bestehendes Kapital keine Spekulationsblasen aufgepustet werden.

Bedingung (1) determiniert normalerweise den Lohnsatz. Lediglich für den Fall, daß der tarifvertraglich fixierte Lohn den markträumenden Lohn übersteigt, determiniert (1) die Beschäftigung. Die vier Gleichungen (2) bis (5) determinieren die Zeitpfade des Kapitals K und seines Preises q. Die Besonderheit bei der Bestimmung des Zeitpfades für q liegt darin, daß für q kein Anfangswert bekannt ist. An seine Stelle tritt die Transversalitätsbedingung. Jedem Anfangswert von q ist ein Zeitpfad von K und q zugeordnet. Aber alle diese Pfade divergieren, bis auf den einen, der die Transversalitätsbedingung erfüllt. Das ist die bekannte Sattelpfadeigenschaft der Gleichgewichtstrajektorie.

Ich komme zu den *Haushalten*. Beide Regionen sind mit einem Kontinuum überlappender Generationen bevölkert, die sich zu Anfang gleichmäßig über das Altersintervall 0 bis 50 Jahre verteilen. Das Alter 0 ist hier der Beginn des Erwerbslebens, nach 40 Jahren endet die Erwerbsperiode und nach 50 Jahren tritt der sichere – und vorhergesehene – Tod ein. Jede Generation gebiert eine Nachfolgegeneration, die 25 Jahre später als sie selbst ins Erwerbsleben tritt. Jede Generation wiederum umfaßt ein Kontinuum von Haushalten. Jeder Haushalt entscheidet nach einem Nutzenmaximierungskalkül über den Konsumpfad für den gesamten Rest seines Lebens sowie über den Wohn- und Arbeitsort. Eine Trennung von Wohn- und Arbeitsort ist – obwohl inzwischen praktisch von erheblicher Bedeutung – nicht zugelassen.

Bei der Wahl der Funktionsform der Nutzenfunktion folge ich wieder der Tradition. Ich wähle eine zeitseparable, zeitneutrale und isoelastische Form:

$$U(t) = \int_{t}^{t^{\dagger}} u[c(s)]\,\mathrm{d}s \text{ mit } u' = c^{-1/\rho}.$$

$U(t)$ ist der zum Planungszeitpunkt t bis zum Lebensende $t^{\dagger} > t$ erzielbare Nutzen. s ist die Zeit, über die der Nutzenfluß u von t bis $t^{\dagger}$ integriert wird. u ist eine wachsende konkave Funktion des Konsumflusses c, weil der Grenznutzen u' eine positive und sinkende Funktion desselben ist. Der Parameter $\rho > 0$ ist die intertemporale Substitutionselastizität.

Die Spezifikation der Nutzenfunktion ist im Modell nur von zweitrangiger Bedeutung. Sie beeinflußt lediglich die Leistungsbilanzreaktion. Das liegt

daran, daß in der unterstellten offenen Ökonomie der Zins im Rest der Welt fixiert wird. Im Modell einer geschlossenen Ökonomie wäre dies anders; denn hier würde die Form der Präferenzen die Zinsreaktion beeinflussen und damit auch auf den Pfad der Kapitalakkumulation wirken.

Der Haushalt maximiert den Nutzen für den Rest seines Lebens unter der Budgetrestriktion

$$V(t) = \int_t^{t^1} c(s) e^{i(t-s)} \, \mathrm{d}s.$$

$V(t)$ ist das Vermögen zum Zeitpunkt t. Es setzt sich zusammen aus dem bis t angesparten Vermögen und dem „Restlebensvermögen". Das „Restlebensvermögen" ist der Barwert der zukünftigen Lohneinkommen bzw. Arbeitsloseneinkommen.

Wie gesagt wählen die Haushalte außer ihrem Konsumpfad auch ihren Standort. Sie wandern von Ost nach West, wenn der Nutzengewinn aus dem erwarteten höheren Einkommen größer ist als der durch den Verlust der Heimat entstehende Nutzenverlust. Da verschiedene Haushalte diesen Verlust unterschiedlich empfinden, treffen natürlich nicht alle Haushalte dieselbe Entscheidung. Einige Haushalte werden bereits durch einen geringen Einkommensgewinn, andere erst durch einen höheren Einkommensgewinn zum Wandern veranlaßt. Dies kann durch den folgenden einfachen Ansatz abgebildet werden:

$$-\dot{P}_o(l, t) = \dot{P}_w(l, t) = \beta v(l, t) P_o(l, t)$$

mit

$$v(l, t) = \frac{W_w(l, t) - W_o(l, t)}{W_o(0, t)}.$$

$P_o(l, t)$ und $P_w(l, t)$ sind die Populationsdichten der l-jährigen Haushalte zum Zeitpunkt t im Osten und im Westen. $-\dot{P}_o(l, t) = \dot{P}_w(l, t)$ ist deswegen die Ost-West-Wanderung der l-jährigen Haushalte pro Zeiteinheit. Die gesamte Ost-West-Wanderung pro Zeiteinheit ergibt sich durch Integration über das Erwerbsalter l. Das Wanderungsmodell unterstellt, die relative Abwanderungsrate $-\dot{P}_o(l, t)/P_o(l, t)$ der l-jährigen ostdeutschen Haushalte sei proportional zum relativen Zuwachs des Restlebensvermögens $v(l, t)$, den ein l-jähriger Haushalt beim Wechsel von Ost nach West noch erzielen kann. $W_o(l, t)$ bzw. $W_w(l, t)$ bezeichnet das Restlebensvermögen eines l-jährigen Haushaltes zum Zeitpunkt t im Osten bzw. im Westen. Da in West und Ost dieselben Preise und Zinsen herrschen, kann das Restlebensvermögen als indirekte Nutzenfunktion verwendet werden, die den Konsumnutzen

für den Rest des Lebens bei Wahl des jeweiligen Standortes in Einheiten des Numeraires ausdrückt. Aus der Spezifikation von v ergibt sich, daß die Haushalte bei ihrer Wanderungsentscheidung den Zuwachs an Konsumnutzen, den sie durch den Wechsel von Ost nach West erzielen könnten, mit dem Konsumnutzen des ganzen Lebens vergleichen, welcher durch das Lebensvermögen eines Osthaushaltes gemessen wird, der zum Zeitpunkt t neu in das Erwerbsleben eintritt.

Die Proportionalitätskonstante β ist der einzige Parameter des Wanderungsmodells. Implizit mißt β, wie hoch die Haushalte die Kosten des Heimatverlustes veranschlagen. Die gewählte Spezifikation impliziert, daß die Abwanderungsrate mit zunehmendem Alter sinkt, wie wir es ja in der Realität auch tatsächlich beobachten.

Soweit zu den Haushalten. Als dritten Akteur im Bunde gibt es neben Firmen und Haushalten schließlich noch den Staat, der gewisse Interventionen vornehmen kann, die ich gleich noch erkläre.

Soweit die formale Modellspezifikation. Obwohl das Modell – gemessen an der komplexen Realität – stark vergröbert, ist es doch bereits zu komplex, als daß man auf analytischem Wege aus ihm noch allzuviel schließen könnte. Was wir analytisch sagen können ist lediglich dieses: Beginnen wir mit einer Steady-State Situation der westdeutschen Ökonomie und wird diese schlagartig und unerwartet durch die Vereinigung mit Ostdeutschland gestört, dann passiert folgendes: Tobins q im Osten steigt sprunghaft an, was dort die Investition in Gang setzt. Diese wird durch Kapitalimport finanziert. Gleichzeitig wandern Arbeitskräfte von Ost nach West, so lange es Haushalte im Osten gibt, deren Vermögensgewinn beim Standortwechsel die subjektiv empfundenen Wanderungskosten mindestens aufwiegt. Langfristig kommt es zum Ausgleich der Kapitalintensität, der Löhne, des Pro-Kopf-Einkommens usw. bei einer gegenüber dem Startpunkt geringeren Bevölkerung im Osten bzw. höheren Bevölkerung im Westen.

Über diese qualitative Prognose möchte ich jedoch hinauskommen und auch quantitative Aussagen zu treffen versuchen. Ich bediene mich deswegen der Simulationstechnik. Um konkrete Zeitpfade auszurechnen, muß ich die freien Parameter des Systems mit Werten belegen. Da ich kein strukturstabiles System über einen langen Zeitraum beobachte, kann ich die Parameter allerdings nicht nach den Regeln der ökonometrischen Kunst schätzen. Deswegen verlasse ich mich teils auf die Literatur, teils auf das Kalibrieren, d.h. auf das Schätzen mit Null Freiheitsgraden.

III. Quantifizierung der Parameter

Im Produktionssektor sind drei Parameter zu bestimmen, ε, γ und η.

- Die partielle Produktionselastizität des Kapitals ϵ kann man bekanntlich im Konkurrenzgleichgewicht an der Einkommensverteilung ablesen. Für die Bundesrepublik findet man in der amtlichen Statistik einen Wert von knapp 0,3.
- Die Fortschrittsrate γ ist hier von minderer Relevanz. Gestützt auf Schätzungen von *Madison* (8) nehme ich eine Fortschrittsrate von 1,4% p.a. an. Bei konstanter Beschäftigung entspricht das einer Steady-State-Wachstumsrate von 2% p.a.
- Den Investitionskostenparameter η kann man aus der Parameterschätzung für Tobins Investitionsfunktion (siehe Gleichung (2)) ablesen. Die einschlägigen Schätzungen sind allerdings mit beachtlichen Unsicherheiten behaftet, da überhaupt die empirische Evidenz zugunsten der neoklassischen Investitionshypothese nicht gerade überwältigend ist (1, 6). Nach einem Survey von *Auerbach* und *Hines* (3) liefern amerikanische Schätzungen einen Wert von 5 Jahren oder mehr. Aus einer Schätzung von *Funke* (5) für Deutschland und Großbritannien kann ein Wert von 3 Jahren abgelesen werden. Je größer η gewählt wird, desto höher sind die unterstellten Anpassungskosten und desto träger reagiert die Investition. Im Sinne einer optimistischen Parametrisierung wähle ich daher mit $\eta = 3$ Jahre den kleinsten mir aus der Literatur bekannten Wert.

Im Konsummodell gibt es nur einen einzigen Parameter, die intertemporale Substitutionselastizität ρ. Sie wird am beobachteten Kapitalkoeffizienten geeicht, d.h. ρ wird so gewählt, daß im Steady-State-Gleichgewicht ein Kapitalkoeffizient von 4,5 Jahren realisiert wird. Dies ist der Wert, den man in der Bundesrepublik in der Vergangenheit nach amtlichen Quellen, die ich an dieser Stelle nicht im Detail diskutieren möchte, beobachten konnte. Das liefert für ρ einen Wert von knapp 0,6, der auch mit ökonometrischen Schätzungen verträglich ist.

Offen bleibt dann noch der Mobilitätsparameter β sowie die ostdeutsche Kapitalausstattung zum Zeitpunkt der Vereinigung. Wie diese beiden Größen kalibriert werden, wird im nächsten Abschnitt bei der Präsentation der Ergebnisse erklärt.

IV. Ergebnisse

Ich spare mir hier Ausführungen über die Lösungstechniken, die nicht gerade einfach sind. Die folgenden Bilder zeigen die Bewegung verschiedener Variablen in der Zeit für den Westen und den Osten. Ich simuliere zwei Szenarien:

- Ein Laissez-Faire-Szenario, in dem die Faktorpreise von Anfang an allein durch Marktkräfte determiniert werden. Es herrscht ohne Unterbrechung Vollbeschäftigung (natürlich bei einem im Osten sehr niedrigen Lohnniveau), und es werden keinerlei Förderinstrumente eingesetzt. Dieses Szenario entspricht etwa dem, was *Sinn* und *Sinn* (11) mit „organischer Systemtransformation" meinen.
- Ein Interventionsszenario, mit dem ich versuche, mich den wirklichen Verhältnissen zu nähern. Es unterscheidet sich in den folgenden drei Punkten vom Laissez-Faire-Szenario:
  - Die Löhne werden im Osten auf einem höheren als dem markträumenden Niveau fixiert, indem sie mit einer schrittweise ansteigenden Relation an den Westlohn gebunden werden. Die Relation ist im ersten Jahr 45% und steigt in Schritten von jährlich 10%-Punkten bis auf 85% an. Tatsächlich ist ja bereits eine vollständige Angleichung der Tariflöhne vereinbart. Effektiv werden die Löhne aber auch in Zukunft nicht vollkommen mit denen im Westen gleichziehen. Dies rechtfertigt die getroffene Annahme.
  - Die Erwerbspersonen, die dadurch arbeitslos werden, erhalten vom Staat 60% des Lohnes eines Beschäftigten.
  - Investitionen werden mit 30% der Investitionssumme staatlich gefördert. Zwar gibt es in den neuen Ländern Fördermöglichkeiten bis zu 50% der Investitionssumme. Aber sie werden nur unter bestimmten Bedingungen gewährt, und die Beschaffung der Fördermittel ist mit erheblichen Informations- und Unsicherheitskosten verbunden, so daß eine allgemeine 30%-Förderung als Repräsentant aller Investitionsförderinstrumente wohl eher hoch angesetzt ist. Auf den Versuch einer seriösen Quantifizierung des repräsentativen Fördersatzes h will ich mich hier jedoch nicht einlassen.

  Sinn und *Sinn* nennen dieses zweite Szenario das „Hochlohn-High-Tech-Szenario".

Im Interventionsszenario fallen Staatsausgaben für die Finanzierung der Arbeitslosen und für die Investitionssubvention an. Ich nehme hier an, diese würden in vollem Umfang durch Neuverschuldung finanziert, deren Rückzahlung so weit in die Zukunft geschoben wird, daß lebende Generationen davon nicht betroffen sind. Ich meine nicht, daß dies eine realistische Annahme ist. Ich benutze lediglich diese Annahme als Kunstgriff, um Finanzierungsfragen hier ganz herauszuhalten. Diese möchte ich in einer Folgeuntersuchung studieren.

Ich komme zu den Resultaten. Im Laissez-Faire-Szenario stellt sich im Osten ein Gleichgewichtslohn und ein Pro-Kopf-Einkommen von knapp 40% des westdeutschen Lohnes bzw. Pro-Kopf-Einkommens ein. Von diesem Ausgangspunkt steigt das Niveau gleichmäßig an und nähert sich asym-

ptotisch dem Westniveau (Abb. 1)[3]. Alle 10 Jahre halbiert sich der Abstand. Nach 40 Jahren ist er auf knapp 3% zusammengeschmolzen.

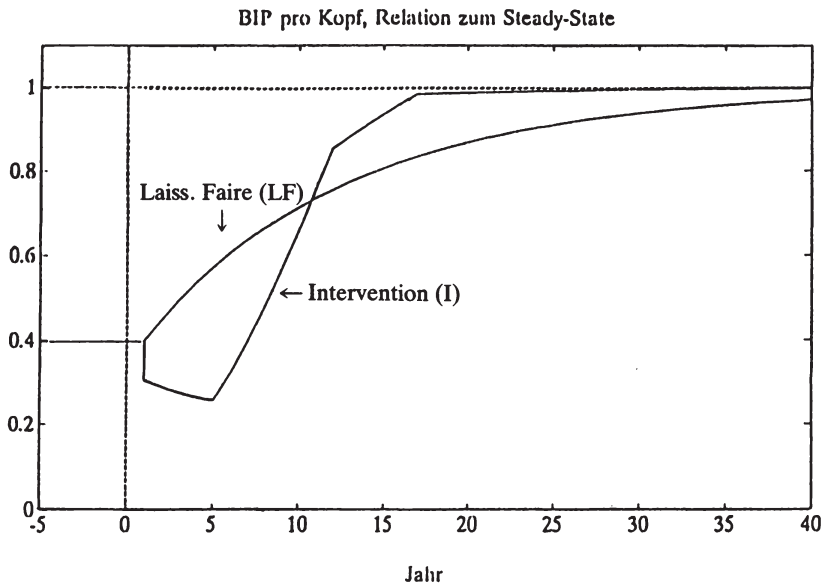


Abbildung 1

Im Interventionsszenario wird der Anfangslohnsatz im Osten auf 45% des Westlohnes fixiert und steigt dann in 4 Jahren linear auf 85%. Entsprechend kommt es zu einer Arbeitslosenquote von anfangs über 30%, die in den 4 Jahren auf über 70% hochschnellt (Abb. 2). Die Kapitalakkumulation ist nicht schnell genug, um diesen Lohnanstieg aufzufangen. Das Pro-Kopf-Einkommen sinkt von dem Anfangswert von 30% des westdeutschen Pro-Kopf-Einkommens in den 4 Jahren noch weiter ab. Vom fünften Jahr an verharrt die fixierte Lohnrelation annahmegemäß auf 85%. Output und Beschäftigung nehmen von da an schnell zu. Nach 12 Jahren geht die Arbeitslosigkeit auf Null zurück. Nach erreichter Vollbeschäftigung verlangsamt sich das Wachstum. Nach insgesamt 16 Jahren ist das Westniveau fast erreicht und die Investitionsförderung wird eingestellt.

[3] In den Abbildungen 1, 3 und 4 zeigen die durchgezogenen Linien die Entwicklung im Osten, die unterbrochenen die im Westen. Die westdeutsche Entwicklung ist vom Steady-State kaum zu unterscheiden.

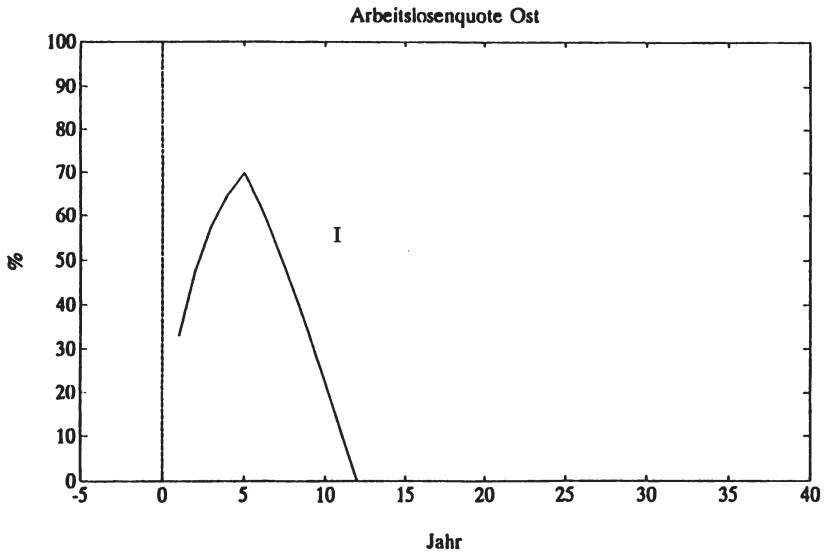


Abbildung 2

Interessant ist, daß trotz erheblicher Wanderungen, zu denen ich gleich noch komme, die westdeutsche Ökonomie durch die Vereinigung nur wenig berührt wird, abgesehen von den fiskalischen Belastungen, die hier aber noch nicht im Blickfeld sind.

An dieser Stelle komme ich noch einmal auf die Schätzung des ostdeutschen Kapitalbestandes zurück. Ich habe ihn derart quantifiziert, daß zum Zeitpunkt der Vereinigung das ostdeutsche Pro-Kopf-Einkommen im Interventionsszenario 30% des westdeutschen beträgt. Das ist die Relation, die sich nach der amtlichen Statistik 1991 ergeben hat.

Wie sieht die Entwicklung auf dem Kapitalmarkt aus? Wegen der günstigen Investitionsmöglichkeiten ist der Kurswert des ostdeutschen Kapitalbestandes im Verhältnis zu seiner Effizienz nach der Vereinigung hoch und sinkt dann im Laufe der Zeit (Abb. 3). Im Interventionsszenario ist dieser Kurswert, von dem nach *Tobins* Investitionsfunktion ja die Investitionen abhängen, zwar kleiner als im Laissez-Faire-Szenario. Aber die Investitionen sind im Interventionsszenario solange höher als im Laissez-Faire-Szenario, wie die 30prozentige Investitionsförderung gezahlt wird (Abb. 4).

Wie sehen die Bevölkerungsbewegungen aus? Die Wanderung von Ost nach West beträgt – hochgerechnet auf die Gesamtbevölkerung – am Anfang 270 Tsd. Personen p.a. im Laissez-Faire-Szenario und 200 Tsd. Personen p.a.

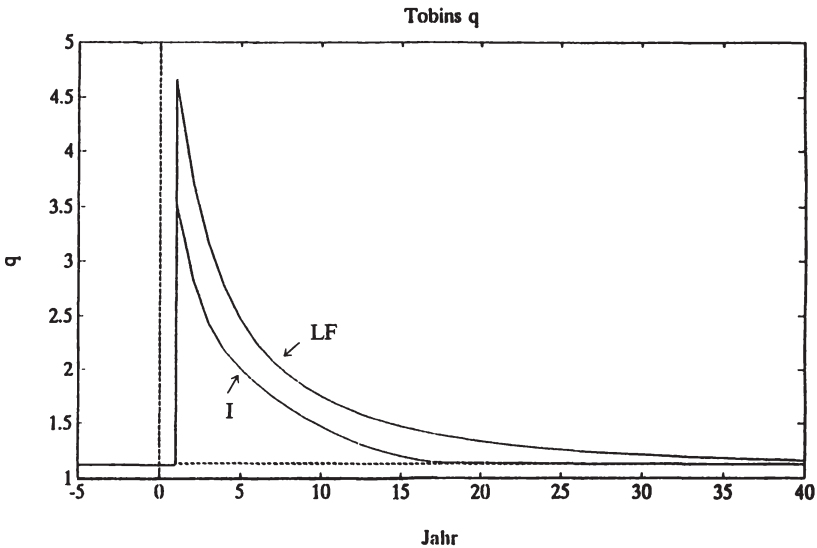


Abbildung 3

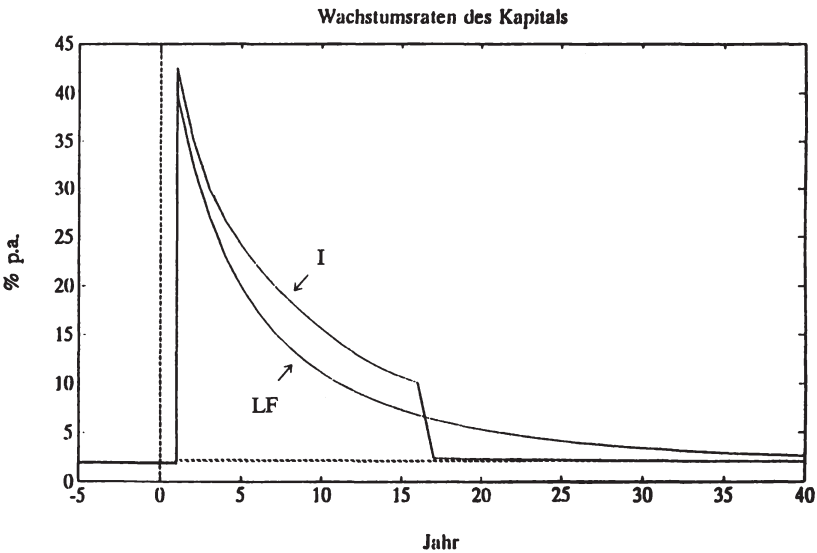


Abbildung 4

im Interventionsszenario (Abb. 5). Letzteres ist der beobachtete Wert für die Zeit vom 1.7.90 bis 30.6.91. Daß diese Beobachtung im Modell genau reproduziert wird, liegt am Kalibrierverfahren. Der Wanderungsparameter β, dessen Quantifizierung ich oben noch offengelassen hatte, wird nämlich

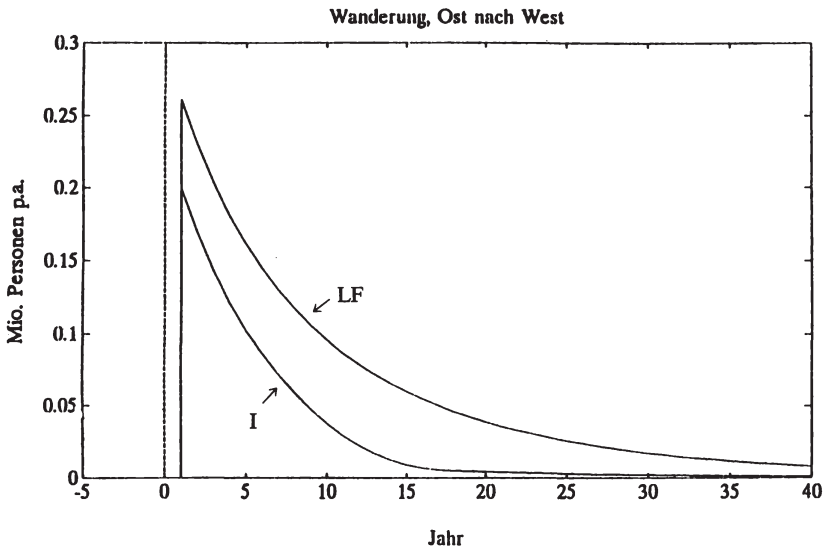


Abbildung 5

gerade derart fixiert, daß im Modell auf dem Gleichgewichtspfad die Nettowanderung im ersten Jahr nach der Vereinigung der genannten Beobachtung entspricht. In Folge der Wanderungen geht die Bevölkerung im Osten im Interventionsszenario um ca. 1,5 Millionen und im Laissez-Faire-Szenario um ca. 3,5 Millionen Personen zurück (Abb. 6).

Was tut sich auf dem Gütermarkt? Annahmegemäß treten keine Preis- und Zinsreaktionen auf. Die zusätzliche Nachfrage nach der Vereinigung wird mit zusätzlichen Importen befriedigt. Der Außenbeitrag, der hier im Ausgangs-Steady-State Null ist, wird schlagartig negativ (Abb. 7). Im Laissez-Faire-Szenario geht das Defizit aber schnell zurück und weicht einem Überschuß; denn die Inländer wollen ihre Schulden zurückzahlen und auf lange Sicht den Marktwert des inländischen Kapitals als Vermögen besitzen. Im Interventionsszenario kommt es zu einem exorbitanten Defizit, das mit dem Ende der Investitionsförderung im 17. Jahr in einen Überschuß umschlägt.

Insgesamt scheint das Interventionsszenario ganz ermutigend zu sein, aber es hat seinen Preis. So, wie ich das Szenario formuliert habe, bleibt die Rech-

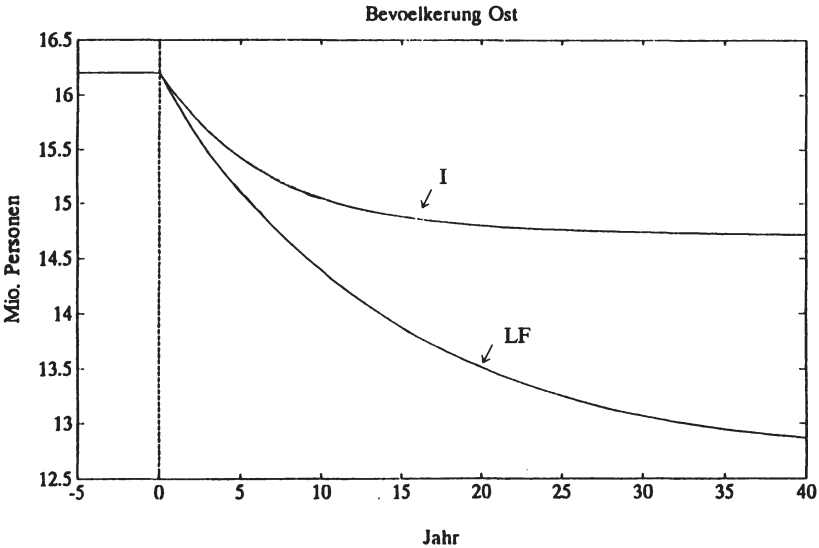


Abbildung 6

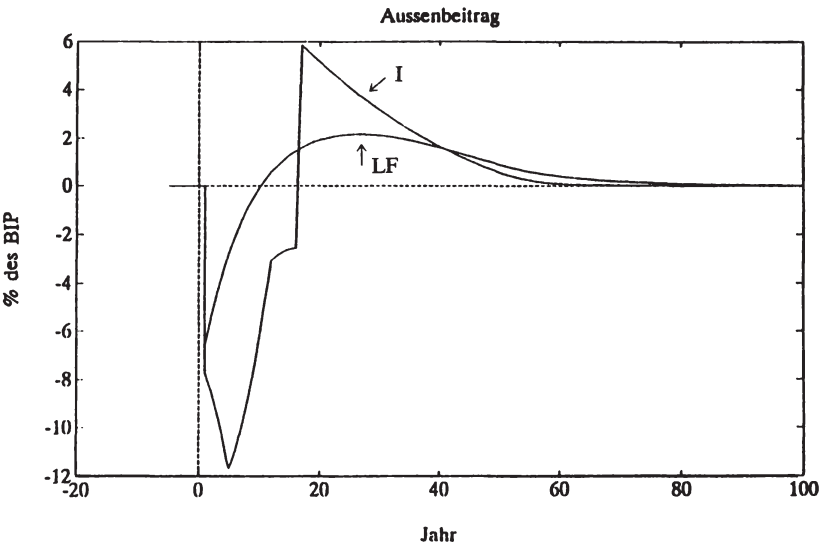


Abbildung 7

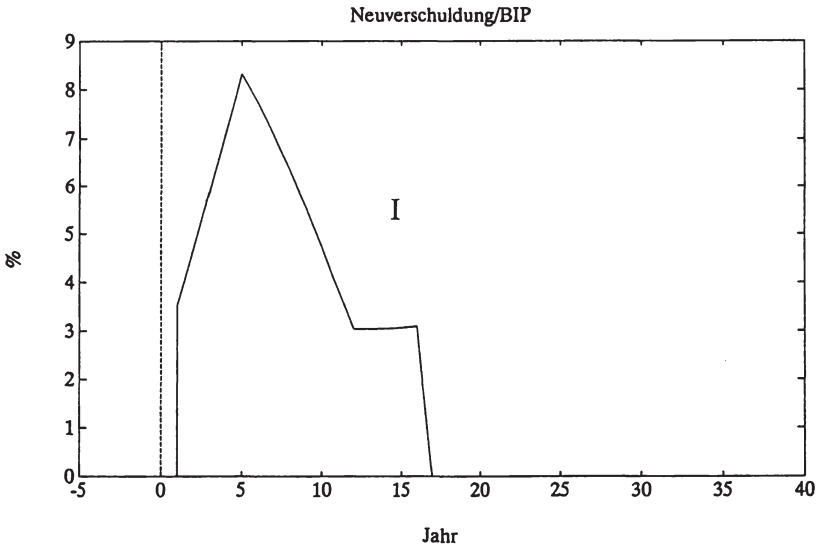


Abbildung 8

nung unbezahlt, weil keine Finanzierung der Arbeitslosenunterstützung und der Investitionsförderung vorgesehen ist. Implizit werden sie also durch Staatsverschuldung finanziert. Anfangs machen diese Ausgaben 3,5% des Sozialproduktes aus, schnellen dann mit dem Zuwachs der Arbeitslosigkeit auf 8,5% hoch, fallen dann mit dem Abbau der Arbeitslosigkeit wieder auf gut 3% und gehen erst nach dem Ende der Investitionsförderung auf Null zurück (Abb. 8). Finanzierte man diese Ausgaben tatsächlich voll durch Staatsverschuldung, so baute sich eine Staatsschuld auf, die in diesen 16 Jahren einschließlich Verzinsung die Höhe eines Sozialproduktes erreicht (Abb. 9)[4].

V. Ausblick

Bevor ich meine Schlußfolgerungen ziehe, sind wohl einige Anmerkungen zu der eingeschränkten Aussagekraft dieser Modellstudie angebracht. Selbst-

[4] Es wird unterstellt, die Staatsschuld sei vor der Vereinigung null. Nach der Vereinigung tritt dann anfangs eine negative Staatsschuld (d.h. ein positives Staatsvermögen) auf, weil ich annehme, daß der Staat den ostdeutschen Kapitalbestand in sein Vermögen übernimmt. Einen Teil dieses Vermögens, der dem Wert der ostdeutschen Sparguthaben entspricht, zahlt er nach der Vereinigung an die Ostdeutschen wieder aus.

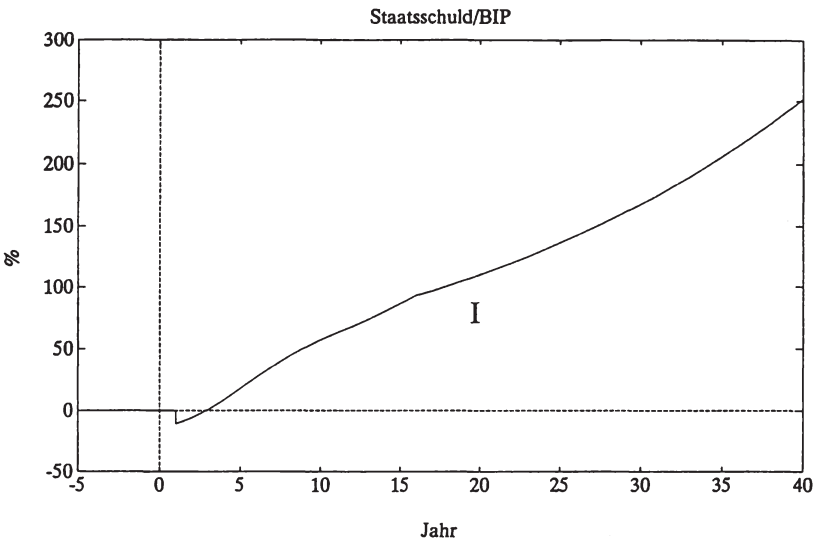


Abbildung 9

verständlich kann die wirkliche Anpassung nicht so harmonisch ablaufen, wie in diesem Modell geschildert. Die Reibungsverluste sind ja bekannt:

- Die Eigentumsverhältnisse sind in Ostdeutschland teils ungeklärt, auf den Wohnungsmärkten sind niedrigere als die markträumenden Preise fixiert, die Verwaltung ist noch mangelhaft entwickelt, und die Qualifikationen sind gewiß noch nicht angeglichen. Ferner ist offensichtlich, daß in Wirklichkeit auch das ostdeutsche Kapital nicht, wie es in diesem Modell unterstellt wurde, vom ersten Tag der Vereinigung an zu seinem fundamentalen Wert gehandelt werden konnte; dafür sind die Märkte zu dünn und z.T. fehlen Informationen über die wirklichen Vermögensbestände.
- Abgesehen davon, daß ich diese speziellen ostdeutschen Hemmnisse vernachlässige, sind selbstredend auch gegen diesen Typ neoklassischer Modelle ganz allgemein Einwände zu erheben. In der wirklichen Welt sind Preise nicht völlig flexibel, es treten Mengenrestriktionen auf, die das System in ein Gleichgewicht treiben können, in dem nicht alle Märkte geräumt sind. Die wirkliche Welt ist voller Risiken, die sich nicht vollkommen diversifizieren lassen. In der wirklichen Welt sind die Märkte mehr oder weniger monopolistisch. Eingesessene Firmen haben einen Anreiz, Marktzutrittsschranken aufzubauen, die ein Aufholen des Ostens erschweren könnten.
- Schließlich kann man natürlich auch gegen meine konkrete Modellspezifikation Einsprüche erheben; ein 5-Parameter-Modell muß ein allzu grobes

Werkzeug sein für ein so vielschichtiges Problem. Mir selbst scheint in dieser Hinsicht der vernachlässigte Zinseffekt der schwerwiegendste Einwand zu sein. Um ihn einzubauen, muß man zwischen Tradeables und Non-Tradeables unterscheiden. Das macht die Dinge allerdings wesentlich komplizierter und wäre ein nächster Schritt.

Trotz dieser Einschränkungen gibt uns die Analyse Hinweise auf die makroökonomischen Restriktionen, die jede Diskussion über ökonomische Perspektiven und Strategien in Deutschland respektieren muß. Alle Annahmen liegen nämlich auf der optimistischen Seite. Es gibt wenige Gründe, weswegen die Entwicklung schneller gehen könnte oder billiger sein könnte, als in den Szenarien gezeigt.

Die Studie macht erneut deutlich, daß man drei Dinge

- eine Lohnangleichung in 4 bis 5 Jahren,
- eine schnelle Angleichung des ökonomischen Niveaus
- und solide Staatsfinanzen ohne ernsthafte Steuererhöhungen oder Ausgabenreduzierungen

nicht gleichzeitig realisieren kann.

Literatur

1. *Abel,* Andrew/*Blanchard,* Olivier J. (1986): The present value of profits and cyclical movement in investment. Econometrica, 54: 249 - 273.
2. *Allais,* Maurice (1947): Economie et interet. Imprimerie Nationale, Paris.
3. *Auerbach,* Alan J./*Hines,* James R. (1986): Tax Reform, Investment, and the Value of the Firm. NBER Working Paper Nr. 1803, Cambridge.
4. *Blanchard,* Olivier J./*Fischer,* Stanley (1989): Lectures on Macroeconomics. MIT-Press, Cambridge, Massachusetts.
5. *Funke,* Michael/*Wadewitz,* Sabine/*Willenbockel,* Dirk (1989): Tobin's q and sectoral investment in West Germany and Great Britain: a pooled cross-section and time-series study. Zeitschrift für Wirtschafts- und Sozialwissenschaften, 109: 399 - 420.
6. *Galeotti,* Marzio (1988): Tobin's marginal „q" and tests of the firm's dynamic equilibrium. Journal of Applied Econometrics, 3: 267 - 277.
7. *Hayashi,* Fumio (1982): Tobin's marginal *q* and average *q*: a neoclassical interpretation. Econometrica, 50: 213 - 224.
8. *Maddison,* Angus (1987): Growth and slowdown in advanced capitalist economies: techniques of quantitative assessment. Journal of Economic Literature, 25: 649 - 698.
9. *Raffelhüschen,* Bernd (1992): Labor migration in Europe: Experiences from Germany after unification. European Economic Review, 7:1453 - 1471.

10. *Ramsey,* Frank (1928): A mathematical theory of saving. Economic Journal, 38: 543 - 559.

11. *Sinn,* Gerlinde/*Sinn,* Hans-Werner (1991): Kaltstart: Volkswirtschaftliche Aspekte der deutschen Vereinigung. Mohr/Siebeck, Tübingen.

Technische Effizienz und Kapitalmobilität in der Verarbeitenden Industrie: Ein interregionaler Vergleich für die Bundesrepublik Deutschland

Von *Hans Joachim Schalk*, Münster*

I. Einführung

Nach der neoklassischen Theorie weichen die Einkommen pro Kopf zweier Länder oder Regionen deshalb voneinander ab, weil der Faktor Arbeit unterschiedlich mit Kapital ausgestattet ist. In der „reichen" Region mit dem höheren Pro-Kopf-Einkommen (der höheren Produktivität) wird die Kapitalintensität größer sein als in der „armen". Gleichzeitig impliziert dies, daß die physischen Grenzproduktivitäten des Kapitals in der armen Region höher und in der reichen Region niedriger sind. Bei völliger Mobilität von Kapital (vollkommenem Wettbewerb, freien Faktormärkten) dürften aufgrund dieses Gefälles in den Grenzproduktivitäten neue Investitionen praktisch nur in den armen Regionen stattfinden, Kapital müßte von den reichen in die armen Regionen fließen, bis sich die Kapitalintensitäten und damit auch die Einkommen pro Kopf angeglichen haben (vgl. *Lucas* 1990).

Die in der Realität der Bundesrepublik Deutschland zu beobachtende Entwicklung läßt sich allerdings mit den Vorhersagen dieses Modells nicht vereinbaren. So lag beispielsweise die Bruttowertschöpfung je Einwohner im Jahre 1988 in der „reichsten" Arbeitsmarktregion München zweieinhalbmal so hoch wie in der „ärmsten" Region Freyung (vgl. *Ziegler* 1992). Ähnlich groß ist die Kluft zwischen Ost- und Westdeutschland. Das Bruttoinlandsprodukt je Erwerbstätigen betrug 1992 im Westen das 2,6fache des Ostens (vgl. *Jahreswirtschaftsbericht* 1993). Infolge dieses Einkommens- bzw. Produktivitätsgefälles müßten sich nach dem neoklassischen Wachstumsmodell rapide Kapitalströme aus den reichen in die armen Regionen ergeben, die aber weder zwischen den Regionen Westdeutschlands noch zwischen Ost und West gegenwärtig zu beobachten sind.

Offenbar sind die Annahmen, die diesem Modell zugrunde liegen, völlig unrealistisch (vgl. *Herdzina* 1993). In diesem Beitrag soll untersucht werden,

* Ich danke J. Lüschow für wertvolle Hilfe bei der Durchführung der Berechnungen. Hilfreiche Kommentare und kritische Anmerkungen erhielt ich von I. Deitmer, W. Franz, M. Wolf und den Teilnehmern an der Sitzung des Ausschusses für Regionaltheorie und -politik, Oktober 1992 in Borken.

welche Konsequenz sich für die Konvergenzaussage ergibt, wenn eine der Annahmen des Modells, nämlich die völlig identischer Produktionstechnologien in allen Regionen, nicht zutrifft. Hierzu werden in einer empirischen Analyse die regionalen Unterschiede in den „technischen Effizienzgraden" der Verarbeitenden Industrie zu ermitteln versucht. Die Messung der technischen Effizienz basiert auf der Schätzung von „frontier production functions". Das theoretische Konzept dieser Funktionen wird im nächsten Abschnitt entwickelt. Die methodische Vorgehensweise zur Schätzung der Produktionsfunktion und technischen Effizienzgrade ist Gegenstand des dritten Abschnitts. Im vierten Abschnitt werden empirische Ergebnisse dargestellt und im fünften Abschnitt schließlich versucht, auf der Grundlage des empirischen Modells eine Antwort auf die Frage zu geben, warum physisches Kapital nicht von den reichen in die armen Regionen fließt[1].

Eine Warnung sei gleich zu Beginn und damit an exponierter Stelle ausgesprochen (dafür aber nicht mehr wiederholt). Jeder Empiriker weiß selbst sehr genau, „that there is no such thing as perfect data" (*Petersen* 1989, 203). Dies gilt für regionale Daten noch viel mehr als für gesamtwirtschaftliche und im besonderen für die in dieser Arbeit verwendeten Daten, die teilweise, wie z.B. die regionalen Kapitalbestände, auf eigenen Berechnungen beruhen. Die ausgewiesenen quantitativen Resultate können deshalb nur ungefähre Vorstellungen über Größenordnungen vermitteln und dürfen nicht als bis zum letzten Dezimalpunkt genau mißinterpretiert werden. Dennoch sei die Auffassung vertreten, daß die ausgewiesenen quantitativen Ergebnisse nützliche Informationen für die Beantwortung der in diesem Beitrag interessierenden Frage liefern können, warum privates Kapital nicht (ohne weiteres) aus den reichen Regionen der Bundesrepublik in ihre ärmeren Gebiete strömt.

II. Technische Effizienz und die „frontier production function"

Eine an der regionalen Wirtschaftspolitik häufig geäußerte Kritik geht dahin, daß ihr System der Investitionsförderung in strukturschwachen Räumen die Faktorpreise verzerre und somit zu einer fehlerhaften interregionalen Allokation der Faktoren Arbeit und Kapital führe. Die Regionalförderung verursache demnach allokative Ineffizienz oder Preis-Ineffizienz, die in einem Verlust an Wachstum resultiere, da die Produktionsfaktoren nicht ihrer wachstumsoptimalen Verwendung zugeführt werden. Ein anderer Aspekt von ökonomischer Ineffizienz, nämlich technische Ineffizienz, fand dagegen in der regionalwissenschaftlichen Diskussion in der Bundesrepublik m.W. bisher weniger Beachtung. Technische Ineffizienz liegt dann vor, wenn ein

[1] Die Anregung, das empirische Modell für die gestellte Frage auszuwerten, erhielt der Verf. durch den Artikel von *Lucas*, R. J. (1990) (Why Doesn't Capital Flow from Rich to Poor Countries?). *Lucas* geht allerdings methodisch einen etwas anderen Weg.

Betrieb, eine Branche oder Region mit einem bestimmten Einsatz an Inputs nicht den größtmöglichen Output erzielt, die Kombination der Produktionsfaktoren somit nicht in der effizientesten Weise erfolgt. Die Produktionsverluste aufgrund von technischen Ineffizienzen können nicht minder bedeutsam sein als die infolge allokativer Ineffizienz[2]. Wenn es deshalb regionale Unterschiede in der technischen Effizienz geben sollte, wäre es wichtig zu wissen, wie groß diese sind, denn die Effizienz der Kombination von Produktionsfaktoren stellt einen wichtigen Faktor dar, der die relative Attraktivität einer Region als Standort für Investitionen beeinflußt (*Krakowski/Lav/Lux* 1992, 531).

Die gewöhnlich zur ökonometrischen Schätzung der Produktionsfunktion angewandten Verfahren verstoßen gegen das Konzept der Produktionsfunktion, die technisch effiziente Produktion voraussetzt und nur Punkte maximalen Outputs für gegebene Inputkombinationen beschreiben soll (vgl. *Schumann* 1992, 139)[3]. Beispielsweise weist die mit der Methode der kleinsten Quadrate geschätzte Produktionsfunktion sowohl positive als auch negative Residuen auf. Dies kommt daher, weil mit der Regressionsanalyse eine „durchschnittliche" Produktionsfunktion geschätzt wird, die für gegebene Inputmengen den „im Durchschnitt" erzielbaren und nicht den „maximalen" Output angibt. Das produktionstheoretische Konzept verlangt aber ein Schätzverfahren, das Abweichungen von der Produktionsfunktion nur in Richtung auf niedrigere Outputwerte zuläßt. Eine mit solchen Verfahren geschätzte Funktion wird als „frontier (production) function" oder „Rand-(Produktions-) Funktion" bezeichnet. Durch das Wort „Rand" soll hervorgehoben werden, daß es sich um eine Funktion handelt, die nur den maximalen Output beschreibt, der mit alternativen Inputkombinationen erzielbar ist. Die Randfunktion repräsentiert somit die „best-practice"- im Gegensatz zur „average-practice"-Technologie, die man mit der Methode der kleinsten Quadrate erhält (vgl. *Beeson/Husted* 1989, 15).

Worauf ist es zurückzuführen, daß in bestimmten Regionen unterhalb der Randfunktion produziert wird? Der Grund hierfür liegt in der unterschiedlichen Ausstattung der Regionen mit „Faktorinputs", die in empirischen Produktionsfunktionen oft nicht erfaßt werden (können). Dazu zählen beispielsweise regionale Unterschiede in der qualitativen Zusammensetzung der in

[2] So stellt *Leibensein* (1966, 394f.), dessen Begriff X-Effizienz mit technischer Effizienz verwandt aber nicht identisch ist (vgl. *Frantz* [1990], 150f.), in einer allerdings schon etwas älteren Studie für die USA fest, daß Wohlfahrtsgewinne, die mit einer Steigerung der allokativen Effizienz erzielt werden können, relativ gering sind; vgl. dazu auch *Bohnet/Beck* (1986) und *Böbel* (1984) für neuere Untersuchungen, die zu höheren Ergebnissen gelangen.

[3] Vgl. hierzu ausführlicher *Schalk* (1976), 72ff. In der Zwischenzeit ist eine umfangreiche Literatur entstanden, die sich mit der Schätzung des maximal möglichen Outputs als Funktion von Inputmengen beschäftigt. Für eine Übersicht vgl. z.B. *Schmidt* (1985/86).

Produktionsfunktionen meistens nur berücksichtigten klassischen Faktoren Arbeit und Kapital, in der Innovationsleistung ansässiger Unternehmen, im Agglomerationsgrad der Region und in der quantitativen und qualitativen Ausstattung mit Infrastrukturkapital[4,5]. Die unterschiedliche branchenmäßige Zusammensetzung des hier betrachteten Aggregats Verarbeitende Industrie kann sich ebenfalls in regionalen Effizienzunterschieden niederschlagen, da sie unterschiedliche regionale Produktionsfunktionen implizieren. Diese Unterschiede kommen dann in regional differierenden Meßwerten für die Effizienzgrade zum Ausdruck. Durch die Schätzung einer Rand-Produktionsfunktion kann nun der Outputeffekt dieser Faktoren ermittelt werden, ohne daß die Faktoren selbst gemessen werden müssen.

III. Methodische Vorgehensweise

Das produktionstheoretische Konzept verlangt ein Schätzverfahren, das Outputwerte nur genau auf oder unter dem Rand der Funktion zuläßt. Diese Forderung läßt sich als ein Problem auffassen, das mit den Methoden der linearen oder quadratischen Programmierung gelöst werden kann. *Aigner* und *Chu* (1968) waren die ersten, die dieses Instrumentarium zur Bestimmung einer Rand-Produktionsfunktion einsetzten[6]. Für die Bundesrepublik wurde es von *Brockhoff* (1970) und *Schalk* (1975 und 1976) zur Schätzung einer Produktionsfunktion für ein Unternehmen der chemischen Industrie bzw. regionaler Randfunktionen verwendet[7]. Auf Regionen übertragen kann das (linearisierte) Produktionsmodell von *Aigner* und *Chu* wie folgt geschrieben werden:

$$(1) \qquad Y_r = \alpha + X_r\beta - u_r,\ u_r \geq 0,\ r = 1, \ldots, R$$

Darin bezeichnen Y und X den Output bzw. einen Vektor mit K Inputs (in Logarithmen), α und β den Effizienzparameter bzw. die partiellen Produk-

[4] Zu einem großen Teil handelt es sich bei diesen in Produktionsfunktionen üblicherweise nicht berücksichtigten Faktorinputs um solche Faktoren, die Untersuchungsgegenstand der innovationsorientierten Regionalpolitik sind; vgl. *Ewers/Brenck* (1992) für eine Übersicht.

[5] Man kann deshalb den technischen Effizienzgrad auch als quantitativen Ausdruck der Wettbewerbsfähigkeit einer Region interpretieren, die wiederum durch die aufgezählten Faktoren beeinflußt wird; vgl. *Werner* (1992), 583.

[6] Die Arbeit von *Aigner* und *Chu* kann als der Beginn der parametrischen Ansätze zur Schätzung der Randfunktion betrachtet werden. Ein nicht-parametrisches Verfahren, bei dem keine explizite Form für die Produktionsfunktion spezifiziert wird, war erstmals von *Farrell* (1957) zur Messung allokativer und technischer Ineffizienzen verwendet worden; vgl. für Übersichten zu diesen Arbeiten *Førsund/Lovell/Schmidt* (1980) und *Schmidt* (1985/86).

[7] Weitere Beispiele für Schätzungen mit Programmierungsmethoden sind *Timmer* (1971) und *Carlsson* (1972).

tionselastizitäten und r den Regionsindex. Der Term u_r repräsentiert regionale Ineffizienz und ist entweder Null oder positiv. Er bewirkt, daß die regionalen Produktionswerte entweder auf oder unterhalb der Randfunktion liegen:

$$Y_r \leq \alpha + X_r\beta \tag{2}$$

Die Parameter der Randfunktion können mit dem linearen Programmierungsverfahren „geschätzt" werden, wozu die Zielfunktion

$$\sum_{r=1}^{R} (\alpha + X_r\beta - Y_r) \tag{3}$$

unter den R Nebenbedingungen (2) zu minimieren ist[8]. Die technischen Ineffizienzen für jede Region werden schließlich aus den Abweichungen der tatsächlichen Outputwerte von den bei den gegebenen Inputkombinationen mit der geschätzten Randfunktion ermittelten Werten abgeleitet.

Das Verfahren hat einige schwerwiegende Nachteile (vgl. auch *Førsund/Lovell/Schmidt* 1980, 10). Es stehen keine statistischen Prüfmaße (Standardfehler, t-Werte etc.) zur Beurteilung der Schätzung zur Verfügung, da keine Annahmen über die Verteilung der Residuen gemacht werden. Das Verfahren besitzt die Eigenschaft, daß die Lage der Funktion und damit deren Parameterschätzwerte gerade durch die extremen Beobachtungswerte am äußersten Rand der Punktwolke bestimmt werden. Das Modell reagiert deshalb auf „Ausreißer" extrem empfindlich. Da die Menge der Produktionswerte, die maximal genau auf der Randfunktion liegen, gleich sein muß der Anzahl der zu schätzenden Parameter, wird die Zahl der Regionen, die technisch (100prozentig) effizient sein können, durch die Methode und die gewählte Produktionsfunktion (willkürlich) limitiert.

Diese Nachteile sog. deterministischer Modelle, deren Fehlerterm nur eine einseitige negative Komponente enthält, können vermieden werden, wenn sog. Paneldaten verfügbar sind und dann stochastische Schätzverfahren angewendet werden[9]. Für diese Untersuchung liegt ein solcher Panel-Datensatz vor, d.h. regionale Daten mit Beobachtungen für jede Region über mehrere Zeitperioden. Das Produktionsmodell (1) wird jetzt wie folgt formuliert:

$$Y_{rt} = \alpha + X_{rt}\beta + v_{rt} - u_r,\ u_r \geq 0,\ r = 1, \ldots, R,\ t = 1, \ldots, T \tag{4}$$

---

[8] Wird anstelle von (3) die Summe der quadrierten Abweichungen minimiert, erhält man ein mit der quadratischen Programmierung zu lösendes Problem; vgl. *Aigner/Chu* (1968), 827.

[9] Ansatz (1) hätte auch als stochastisches Modell formuliert und geschätzt werden können. Mit diesen Modellen treten dann aber bisher noch nicht gelöste Probleme insbesondere bei der Ermittlung der technischen Effizienzgrade auf; vgl. dazu *Schmidt/Sickles* (1984) und *Beeson/Husted* (1989).

Der Index t steht für die Zeitperiode. u reflektiert wieder (regionale) technische Effizienz, während v alle zufälligen Einflüsse auf den Produktionsprozeß repräsentiert.

Das Modell (4) kann auf verschiedene Weise geschätzt werden[10]. Geht man davon aus, daß v normal, u einseitig normal verteilt und v und u unabhängig voneinander und von den Regressoren X sind, läßt sich das Maximum-Likelihood(ML)-Verfahren zur Schätzung einer Randfunktion anwenden. Aus der geschätzten Funktion können dann nach einem Verfahren von *Battese/Coelli* (1988) für jede Region die technischen Effizienzgrade ermittelt werden[11].

IV. Empirische Ergebnisse

Das ML-Verfahren wurde zur Schätzung von Randfunktionen für die Verarbeitende Industrie mit kombinierten regionalen Querschnittsdaten auf Kreisbasis (328 Kreise) und Zeitreihendaten für die Jahre 1978 bis 1989 angewendet. Insgesamt stand somit ein gepoolter Datensatz von 3936 Beobachtungen zur Verfügung. Der Output wird durch die Bruttowertschöpfung gemessen. Als Faktorinputs werden „Arbeit" und „Kapital" berücksichtigt. Für den Faktor Arbeit werden die Beschäftigten aus den statistischen Berichten der Statistischen Landesämter verwendet. Der Faktor Kapital wurde mit Investitionszahlen aus den gleichen Datenquellen mit Hilfe der perpetual-inventory Technik berechnet[12]. Output und Kapital werden in Mill. DM zu Preisen von 1980 gemessen, die Beschäftigten in Tsd.

Den Schätzungen wurden verschiedene Funktionstypen zugrundegelegt. Von allen haben sich für die Cobb-Douglas-Produktionsfunktion die nach den Kriterien der Statistik beurteilten besten Resultate ergeben, auf die deshalb im folgenden ausschließlich eingegangen werden soll[13]. Die Schätzwerte der ML-Methode werden dem Ordinary-Least-Squares(OLS)-Resultat gegenübergestellt.

---

[10] Vgl. zu den verschiedenen Verfahren *Lüschow/Schalk/Untiedt* (1993).

[11] Vgl. zur Vorgehensweise ausführlicher *Aigner/Lovell/Schmidt* (1977), *Jondrow/Lovell/Materov/Schmidt* (1982) und *Lüschow/Schalk/Untiedt* (1993).

[12] Für Einzelheiten zur Datenerstellung vgl. *Asmacher/Schalk/Thoss* (1987). *Deitmer* (1993) hat die Daten neu berechnet und bis 1989 verlängert.

[13] Außer der Cobb-Douglas-Funktion wurden Produktionsfunktionen vom CES (Constant Elasticity of Substitution)-Typ, eine von *Vinod* (1973) gerade für regionale Zwecke vorgeschlagene und von *Beeson* (1987) in einer neueren regionalen Untersuchung verwendete VES(Variable Elasticity of Substitution)-Funktion sowie die von allen flexibelste Form einer Translog-Produktionsfunktion (vgl. *Beeson/Husted* 1989) getestet. Die Schätzergebnisse für die Parameter dieser Produktionsfunktionen unterscheiden sich allerdings nicht signifikant vom Cobb-Douglas-Fall. Die mit den entsprechenden Ansätzen auftretende hohe Multikollinearität läßt außerdem die Interpretation der Schätzresultate problematisch erscheinen.

Die in logarithmisch-linearer Form geschätzte Cobb-Douglas-Funktion lautet:

$$(5) \qquad \ln Y_{rt} = \alpha + \beta_1 \ln A_{rt} + \beta_2 \ln K_{rt} + \beta_3 t + v_{rt} - u_r$$

A und K bezeichnen darin den Arbeits- bzw. Kapitalinput und mit der Zeitvariable t soll der neutrale technische Fortschritt erfaßt werden. Tabelle 1 enthält die Ergebnisse, die für diesen Ansatz mit dem OLS- bzw. ML-Verfahren erhalten wurden. Die OLS-Schätzung liefert für die partiellen Produktionselastizitäten ziemlich ähnliche Resultate wie das ML-Verfahren für die Randproduktionsfunktion. Das bedeutet, daß die „best-practice"-Technologie weitgehend eine neutrale Transformation der „average-practice"-Technologie darstellt. Dies muß auch der Fall sein, wenn das Produktionsmodell richtig spezifiziert sein soll (vgl. *Schmidt* 1985/86, 321f.). Beide Verfahren liefern geringfügig zunehmende Skalenerträge von 1.01. Die technische Fortschrittsrate liegt bei rd. 0,75 Prozent pro Jahr. Diese Ergebnisse stimmen sehr gut mit denen einer früheren Untersuchung für die Industrie in Nordrhein-Westfalen überein (vgl. *Schalk* 1976). Im Falle der ML-Methode handelt es sich bei σ und σ_u/σ_v um Parameter der zu schätzenden Likelihood-Funktion. Sie sind hoch gegen die Nullhypothese gesichert. Dies zeigt an, daß sich die Regionen signifikant im technischen Effizienzgrad voneinander unterscheiden. Im (ungewichteten) Durchschnitt liegt er bei 71 Prozent, mit einem niedrigsten Effizienzgrad von lediglich 44 Prozent.

Obwohl das Ergebnis der ML-Schätzung auf erhebliche Unterschiede im technischen Effizienzgrad hinweist, haben sich auch mit dem OLS-Verfahren, bei dem ein für alle Regionen identischer Effizienzgrad unterstellt wird, „gesicherte" Resultate ergeben. Bei 3936 Beobachtungen kann dies auch erwartet werden, da bei einer entsprechend hohen Anzahl von Freiheitsgraden die Standardfehler sehr klein werden.

Die Häufigkeitsverteilung der Effizienzgrade ist in Abbildung 1 dargestellt. Die meisten Kreise (209) weisen einen mittleren Effizienzgrad um 70 Prozent aus. 51 Kreise sind 80 Prozent oder mehr effizient. Der Rest liegt unterhalb einer Effizienz von 60 Prozent, davon 10 unter 50 Prozent. Ineffizienz scheint somit in einigen Regionen ein nicht unerhebliches Problem darzustellen.

Tabelle 1

Parameterschätzwerte für die Produktionsfunktion (endogene Variable = ln Y; siehe Gleichung 5)[a)]

Unabhängige Variable	OLS-Modell	„Frontier"-Modell (ML-Schätzung)
Konstante	2.530 (48.0)	2.704 (37.8)
ln A	0.683 (56.0)	0.636 (34.9)
ln K	0.330 (29.2)	0.372 (23.5)
t	0.00662 (6.0)	0.00813 (4.7)
$\bar{R}^2$	0.94	
σ	0.230	0.551 (214.4)
σ_u / σ_v		1.864 (35.3)
∅ Effizienz		0.70
minimale Effizienz		0.44

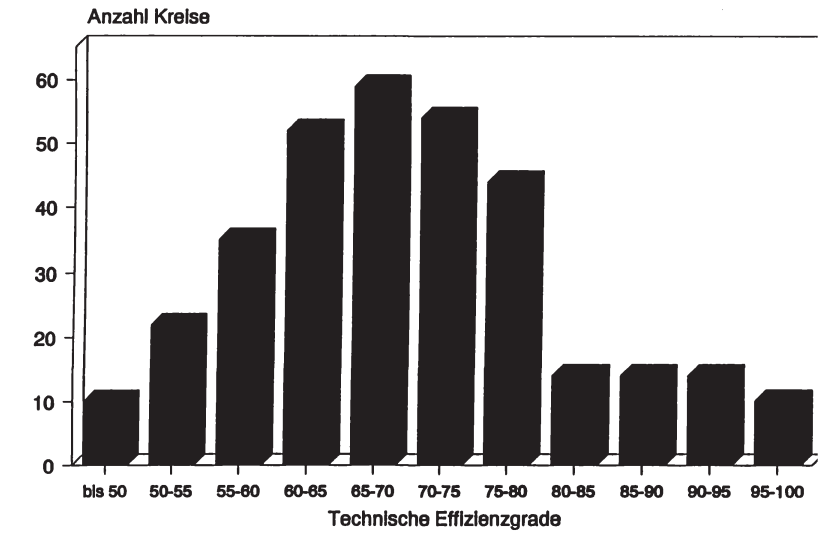
a) t-Werte in Klammern unter den Regressionskoeffizienten; vgl. Text für weitere Erläuterungen.

V. Technische Effizienz und Kapitalmobilität

Ob, wie schnell und unter welchen Voraussetzungen Länder bzw. Regionen unterschiedlichen Entwicklungsstandes konvergieren, sind nicht nur für die Bundesrepublik Deutschland nach der Vereinigung beider deutscher Staaten Fragen von großem nationalen Interesse. Sie werden gegenwärtig verstärkt auch im internationalen Kontext im Rahmen des europäischen Binnenmarktprojekts und weltweit diskutiert und wissenschaftlich untersucht[14]. Nach dem neoklassischen Modell des Wirtschaftswachstums gibt es einen automatischen Mechanismus, der zu Konvergenz in den Einkommens- und Produktionsniveaus pro Kopf führt. Die Grenzproduktivität des Kapitals ist in der armen Region mit dem niedrigen Pro-Kopf-Einkommen höher als in der reichen. Auf einem freien Kapitalmarkt müßte dieses Produktionsgefälle dazu führen, daß Kapital aus der reichen in die arme Region fließt. Im folgenden

[14] Vgl. *Berthold* (1992) für eine Übersicht zur Konvergenzdiskussion und zu einigen empirischen Arbeiten *Barro/Sala-i-Martin* (1991, 1992), *Barro* (1991), *Tamura* (1991) und *Blanchard/Katz* (1992).

soll geprüft werden, inwieweit die Bedingungen für einen solchen Kapitalstrom in der Bundesrepublik erfüllt sind.



Quelle: Eigene Berechnung

Abbildung 1: Häufigkeitsverteilung der technischen Effizienzgrade

Hierzu sei zunächst angenommen, daß die Produktion in allen Regionen (Kreisen) durch die Produktionsfunktion von Cobb-Douglas-Typ mit identischer Technologie, die in der ökonometrischen Analyse mit dem OLS-Verfahren erhalten wurde, beschrieben werden kann. Unter der mit der Empirie gut verträglichen Annahme konstanter Skalenerträge ist der Output pro Arbeiter $Y/A = a\,(K/A)^{1-b}$ und die Grenzproduktivität des Kapitals durch $GPK = a\,(1-b)\,(K/A)^{-b}$ gegeben oder in Abhängigkeit von der Arbeitsproduktivität[15]:

$$GPK = (1-b)\,a^{\frac{1}{1-b}} \left(\frac{Y}{A}\right)^{-\frac{b}{1-b}} \tag{6}$$

In Tabelle 2 sind die nach dieser Formel berechneten Grenzproduktivitäten für Regionen mit hohem, mittlerem und niedrigem Einkommen dargestellt

[15] Um die Darstellung und Argumentation im folgenden zu vereinfachen, wird von konstanten Skalenerträgen ausgegangen, obwohl die ökonometrische Schätzung (geringfügige) zunehmende Skalenerträge ergeben hat. Die dargestellten empirischen Ergebnisse wurden aber mit der ökonometrisch geschätzten Funktion erhalten.

(Spalte (4)). Der Wert 0,12 für die reiche Region entspricht ziemlich genau den Opportunitätskosten („user cost of capital") in den nicht von der regionalen Strukturpolitik geförderten einkommensstarken Gebieten (Spalte (6))[16]. Für die arme Region erhält man eine 2,8mal höhere Grenzproduktivität des Kapitals. Der Grenzertrag des Kapitals (Spalte (4)/Spalte (6))[17] ist in der armen Region infolge der dort subventionierten Kosten für die Kapitalnutzung sogar 3,4mal höher als in der reichen Region. Die traditionelle Theorie besagt, daß ein solches Produktivitäts- bzw. Ertragsgefälle des Faktors Kapital auf Dauer nicht bestehen bleiben kann. Es müßte durch einen starken Zufluß von physischem Kapital aus der reichen in die arme Region eliminiert werden[18]. Dies war allerdings in den letzten Jahren nicht zu beobachten: Weder die Kapitalintensitäten noch die Pro-Kopf-Einkommen haben sich zwischen der armen und reichen Region in dem Maße angenähert, wie durch das Modell vorhergesagt.

Noch viel drastischer fällt ein entsprechender Vergleich zwischen den alten und neuen Bundesländern aus. Gemäß *Görzig* (1992) war der Output pro Arbeiter in der Industrie Westdeutschlands im Jahre 1990 fast 3,9mal höher als in Ostdeutschland. Daraus wird nach (6) eine rd. 17mal höhere Grenzproduktivität des Kapitals für Ostdeutschland errechnet. Hinzu kommt, daß die „user cost of capital" infolge verschiedener Maßnahmen der regionalen Strukturpolitik in Ostdeutschland weniger als 7 Prozent betragen (vgl. *Schalk,* 1992). Bezogen auf die Kapitalnutzungskosten erhöht sich damit das Ertragsgefälle gegenüber dem Westen (Region mittleren Einkommens) auf sogar das 22fache. Dieses Gefälle sollte eigentlich ausreichen, um einen erheblich stärkeren Kapitalstrom von West nach Ost zu begründen als den tatsächlich beobachteten.

Zu einem völlig anderen und mit den empirischen Fakten verträglicheren Ergebnis gelangt man, wenn die regionalen Unterschiede in den technischen Effizienzgraden bei der Ermittlung der Grenzproduktivitäten des Kapitals berücksichtigt werden (vgl. Spalte (5) in Tabelle 2). Das Grenzproduktivitätsverhältnis zwischen reicher und armer Region wird 1, d.h. der Unterschied im technischen Effizienzgrad (Spalte (3)) eliminiert die Differenz in den Grenzproduktivitäten völlig. Kapital fließt nicht aus der reichen in die arme Region, und zwar nicht wegen irgendwelchen Marktunvollkommenheiten,

---

[16] Vgl. zu Berechnungen der user cost of capital, differenziert nach Art der Investition und Fördergebietstypen für alte und neue Bundesländer, *Schalk* (1992), Tabelle 2.

[17] Bei dieser im folgenden als Grenzertrag des Kapitals bezeichneten Relation handelt es sich genauer gesagt um den „Grenzertrag des Geldes, wenn dieses im Faktor (Kapital) angelegt wird" (*Schumann* 1992, 158).

[18] Ein umgekehrter Strom wäre beim Faktor Arbeit zu erwarten, dessen hier nicht ausgewiesenes Grenzproduktivitätsgefälle ungefähr der relativen Lohndifferenz (Spalte (1)) zwischen reicher und armer Region von über 70 Prozent entspricht; vgl. auch *Luger/Evans* (1988), 421f.

Tabelle 2

Grenzproduktivität des Kapitals in „reichen" und „armen" Regionen der Bundesrepublik Deutschland (alte Bundesländer) für die Verarbeitende Industrie[a)]

– Durchschnitt 1985 - 1989 –

Region[b)]	Bruttolöhne pro Arbeitnehmer in Tsd. DM	Y/A in Tsd. DM	TE in Prozent	GPK (OLS) in Prozent	GPK (ML) in Prozent	„user cost of capital" (1992) in Prozent	K/A in Tsd. DM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hohes Einkommen	55	88	81	12	23	12	154
Mittleres Einkommen	40	70	70	20	22	9	127
Niedriges Einkommen	31	51	59	34	23	8	97

a) Geldgrößen in Preisen von 1980; Quelle der Daten: *Deitmer* (1993) und eigene Berechnungen; Klassifikation von „reich" bis „arm" nach Bruttolöhnen pro Arbeitnehmer (Spalte (1)); vgl. Text für weitere Erläuterungen.

b) Region mit hohem Einkommen („reiche" Region): Durchschnitt der 15 Land-Kreise bzw. kreisfreien Städte Hamburg, Wolfsburg, Düsseldorf, Köln, Leverkusen, Erftkreis, Frankfurt, Ludwigshafen, Stuttgart, Böblingen, Mannheim, Bodenseekreis, München, München (LK), Erlangen. Regionen mit niedrigem Einkommen („arme" Regionen): Durchschnitt der 15 Land-Kreise bzw. kreisfreien Städte Wittmund, Pirmasens, Freyung-Grafenau, Passau (LK), Regen, Rottal-Inn, Cham, Tirschenreuth, Hof, Hof (LK), Kronach, Lichtenfels, Ansbach (LK), Neustadt a.d.A.-Bad Windsheim, Schweinfurt (LK). Regionen mit mittlerem Einkommen: Übrige Regionen der Bundesrepublik. Werte für diese Region entsprechen gleichzeitig dem Bundesdurchschnitt.

welche die Mobilität von Kapital behindern, sondern weil kein wirtschaftlicher Anreiz dazu besteht. Aufgrund der größeren Effizienz der reichen Region ist Kapital dort trotz höherer Kapitalausstattung pro Arbeiter genauso ergiebig wie in der armen Region. Oder m.a.W.: Die arme Region ist nicht nur arm, weil der Faktor Arbeit dort schlechter mit Kapital ausgestattet ist, sondern auch da Arbeit und Kapital um rd. 27 Prozent weniger effizient sind als in der reichen Region. Aufgrund der niedrigeren Kapitalnutzungskosten in der armen Region besteht aber immer noch ein geringes (Grenz-)Ertragsgefälle des Kapitals gegenüber der reichen Region von 1,5.

Die Attraktivität der armen Region für die Anlage von Kapital aus der reichen Region ist somit erheblich geringer als es bei Nicht-Berücksichtigung der regional unterschiedlichen technischen Effizienzen erscheint. Dies mag

mit ein Grund für den in empirischen Untersuchungen festgestellten geringen Umverteilungseffekt der Regionalförderung von privatem Kapital aus den nicht-geförderten in die geförderten Regionen sein (vgl. z.B. *Deitmer* 1993). Die Beseitigung der Effizienzunterschiede könnte diesen Effekt erhöhen und damit die Wirksamkeit der regionalen Investitionsförderung in der armen Region steigern.

Für einen Vergleich zwischen Ost- und Westdeutschland ist zunächst das Verhältnis der technischen Effizienzgrade dieser beiden Landesteile zu ermitteln. Hierzu werden die Pro-Arbeiter-Produktionsfunktionen $Y/A = a\,(K/A)^{1-b}\,TE$ für Ost und West nach TE aufgelöst und zueinander in Beziehung gesetzt. In Ermangelung statistischer Informationen über die Kapitalausstattung in Ostdeutschland sei angenommen, das (wirtschaftlich nutzbare) private Kapital pro Arbeiter sei im Osten genauso hoch wie die Kapitalintensität der armen Region Westdeutschlands (vgl. Spalte (7) der Tabelle 2). In diesem Fall erhält man für das Verhältnis des Effizienzgrades zwischen Ost- und Westdeutschland bezogen auf den westlichen Durchschnitt einen Wert von 28 Prozent. Dies impliziert einen technischen Effizienzgrad für Ostdeutschland von 20 Prozent. Unter Berücksichtigung dieses Effizienzgefälles reduziert sich die Relation für die Grenzproduktivitäten des Kapitals von 17 auf einen Wert von 0,33. Eine Investition im Osten könnte dennoch einen höheren Ertrag als im Westen erbringen, da die Kapitalnutzungskosten durch die Regionalförderung in Ostdeutschland niedriger liegen (7 Prozent im Vergleich zu 9 Prozent im westlichen Durchschnitt). Diese Differenz in den Kapitalkosten vermag aber die relative Ertragsrate lediglich auf 0,42 anzuheben, m.a.W., im Westen beträgt der Grenzertrag des Kapitals immer noch das 2,4fache im Osten. Die Kapitalnutzungskosten müßten im Osten auf 3 Prozent heruntersubventioniert werden oder auf 4 Prozent bei gleichzeitiger völliger Abschaffung der Regionalförderung im Westen, um das Ertragsgefälle wenigstens einzuebnen. Dazu müßte aber die gegenwärtige Investitionsförderung im Osten um mindestens die Hälfte aufgestockt werden.

Über die Gründe für die Effizienzunterschiede kann hier nur spekuliert werden[19]. Es ist aber zu vermuten, daß Infrastrukturmängel, geringeres Humankapital[20], administrative Hemmnisse, organisatorische Mängel etc. (vgl. *Weichselberger/Jäckel* 1991) neben dem veralteten Kapitalstock zu Ineffizienzen in Ostdeutschland beitragen. Durch die Beseitigung dieser

[19] Eine empirische Ursachenanalyse der regionalen Effizienzunterschiede wird in der bereits zitierten Arbeit von *Lüschow/Schalk/Untiedt* (1993) durchgeführt.

[20] Zwar lassen die formalen Abschlüsse auf eine qualifizierte Arbeiterschaft auch im Osten schließen, „allerdings ist der Hinweis auf die ‚formalen' Abschlüsse nicht unwichtig, wenn es um einen Vergleich mit westdeutschen Qualifizierungsniveaus geht" (*Franz* 1992, 260). Hinzu kommt, daß die beruflichen Kenntnisse mit einem im Vergleich zu Westdeutschland veralteten Kapitalstock erworben wurden (vgl. *Krakowski/Lav/Lux* 1992, 534).

Mängel und Hemmnisse könnte der geringe Effizienzgrad der ostdeutschen Industrie angehoben und der Grenzertrag des Kapitals wahrscheinlich stärker gesteigert werden als mittels einer noch höheren Subventionierung der Kapitalnutzungskosten durch die Regionalförderung.

VI. Zusammenfassung und Ausblick

Nach neoklassischen Modellvorstellungen differieren die Einkommen pro Kopf zwischen Ländern bzw. Regionen aufgrund unterschiedlicher Ausstattung des Faktors Arbeit mit Kapital. Bei einem freien Kapitalmarkt können diese Unterschiede allerdings nicht auf Dauer bestehen bleiben, da mobiles Kapital den interregionalen Ausgleich der Grenzproduktivitäten erzwingen müßte. Offensichtlich stehen diese Modellaussagen in einem eklatanten Widerspruch zur Realität: Kapital fließt nicht in dem von dem Modell prognostizierten Umfang aus den reichen Regionen mit hohem Pro-Kopf-Einkommen in die armen Regionen mit niedrigem Output pro Kopf. Einer der möglichen Gründe hierfür ist, daß eine fundamentale Annahme des Modells, nämlich die regional invarianter Produktionstechnologien, in der Realität nicht erfüllt ist. Arme Regionen können nicht nur wegen niedriger Kapitalausstattung arm sein, sondern auch wegen geringer technischer Effizienz ihrer Produktionsfaktoren. Aufgrund technischer Ineffizienzen besteht auch kein Anreiz, privates Kapital in den armen Regionen zu lokalisieren. Zur Überwindung dieser Barriere für die Kapitalmobilität müßten die Kosten für die Nutzung von Kapital in den armen Regionen erheblich billiger gemacht werden als es teilweise schon der Fall ist. Die Alternative dazu wäre, für einen Abbau der technischen Ineffizienzen zu sorgen. Die für eine solche Strategie benötigten Informationen sind noch zu erarbeiten. Geographische Unterschiede in der Produktionstechnologie können jedenfalls weitreichende regionalpolitische Konsequenzen haben und sollten deshalb, stärker als dies bisher geschehen ist, in empirischen Untersuchungen beachtet werden.

Literatur

Aigner, D. J./*Chu*, S. F. (1968): On Estimating the Industry Production Function, The American Economic Review 58, 826 - 839.

Aigner, D./*Lovell*, C. A. K./*Schmidt*, P. (1977): Formulation and Estimation of Stochastic Frontier Production Function Models, Journal of Econometrics 6, 21 - 37.

Asmacher, Chr./*Schalk*, H. J./*Thoss*, R. (1987): Analyse der Wirkungen regionalpolitischer Instrumente, Beiträge zum Siedlungs- und Wohnungswesen und zur Raumplanung 120, Münster.

Barro, R. J. (1991): Economic Growth in a Cross Section of Countries, The Quarterly Journal of Economics 56, 107 - 443.

Barro, R. J./*Sala-i-Martin,* X. (1991): Convergence across States and Regions, Brookings Papers on Economic Activity, 107 - 182.

Barro, R. J./*Sala-i-Martin,* X. (1992): Convergence, Journal of Political Economy 100, 223 - 251.

Battese, G. E./*Coelli,* T. J. (1988): Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data, Journal of Econometrics 38, 387 - 399.

Beeson, P. E. (1987): Total Factor Productivity Growth and Agglomeration Economics in Manufacturing, 1959 - 73, Journal of Regional Science 27, 183 - 199.

Beeson, P. E./*Husted,* S. (1989): Patterns and Determinants of Productive Efficiency in State Manufacturing, Journal of Regional Science 29, 15 - 28.

Berthold, N. (1992): Das Binnenmarktprojekt 1992. Konvergenz oder Divergenz in Europa?, WiSt H. 12, 598 - 604.

Böbel, I. (1984): Wettbewerb und Industriestruktur, Industrial Organization – Forschung im Überblick, Berlin etc.

Bohnet, A./*Beck,* M. (1986): X-Effizienz-Theorie: Darstellung, Kritik, Möglichkeiten der Weiterentwicklung, Zeitschrift für Wirtschaftspolitik 35, 211 - 233.

Blanchard, O. J./*Katz,* L. F. (1992): Geographic Differences in Production Technology, Regional Science and Urban Economics 18, 399 - 424.

Brockhoff, K. (1970): Zur Quantifizierung der Produktivität industrieller Forschung durch die Schätzung einer einzelwirtschaftlichen Produktionsfunktion – Erste Ergebnisse, Jahrbücher für Nationalökonomie und Statistik 184, 248 - 276.

Carlsson, B. (1972): The Measurement of Efficiency in Production: An Application to Swedish Manufacturing Industries 1968, The Swedish Journal of Economics 74, 468 - 485.

Deitmer, I. (1993): Effekte der regionalen Strukturpolitik auf Investitionen, Beschäftigung und Wachstum, Dissertation, Münster.

Ewers, H.-J./*Brenck,* A. (1992): Innovationsorientierte Regionalpolitik. Zwischenfazit eines Forschungsprogramms, in: Birg, H./Schalk, H. J. (Hrsg.), Regionale und sektorale Strukturpolitik, Münster.

Farrell, M. J. (1957): The Measurement of Productive Efficiency, Journal of the Royal Statistical Society 120, 253 - 290.

Førsund, F. R./*Lovell,* C. A. K./*Schmidt,* P. (1980): A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement, Journal of Econometrics 13, 5 - 25.

Frantz, R. S. (1988): X-Efficiency: Theory, Evidence and Applications, 2nd printing 1990, Boston.

Franz, W. (1992): Im Jahr danach – Bestandsaufnahme und Analyse der Arbeitsmarktentwicklung in Ostdeutschland, in: Gahlen, B./Hesse, H./Ramser, H. J. (Hrsg.), Von der Plan- zur Marktwirtschaft, Tübingen, 245 - 274.

Görzig, B. (1992): Produktion und Produktionsfaktoren in Ostdeutschland, DIW-Beiträge zur Strukturforschung, Heft 135, Berlin.

Herdzina, K. (1993): Regionale Disparitäten, ländliche Räume und Ansatzpunkte einer integrierten Regionalpolitik, Europäischer Forschungsschwerpunkt Ländlicher Raum, Diskussionsbeiträge Nr. 1/1993, Universität Hohenheim.

Jondrow, J./*Lovell,* C. A. K./*Materov,* J. S./*Schmidt,* P. (1982): On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model, Journal of Econometrics 19, 233 - 238.

Krakowski, M./*Lau,* D./*Lux,* A. (1992): Die Standortqualität Ostdeutschlands, Wirtschaftsdienst X, 530 - 536.

Leibenstein, H. (1966): Allocative Efficiency vs. X-Efficiency, The American Economic Review 56 (3), 392 - 415.

Lucas, R. E. (1990): Why Doesn't Capital Flow from Rich to Poor Countries?, American Economic Review 80, Papers and Proceedings, 92 - 96.

Lüschow, J./*Schalk,* H. J./*Untiedt,* G. (1993): Messung regionaler Effizienzunterschiede in der Verarbeitenden Industrie, in Vorbereitung.

Luger, M. I./*Evans,* W. N. (1988): Geographic Differences in Production Technology, Regional Science and Urban Economics 18, 399 - 424.

Peterson, W. (1989): Rates of Return on Capital: An International Comparison, Kyklos 42, 203 - 217.

Schalk, H. J. (1975): Die Bestimmung der regionalen Produktivität durch die Schätzung von Produktionsfunktionen, Seminarberichte 11 der Gesellschaft für Regionalforschung, 181 - 196.

— (1976): Die Bestimmung regionaler und sektoraler Produktivitätsunterschiede durch die Schätzung von Produktionsfunktionen, Münster.

— (1992): Effects of Regional Policy on Investment and Employment in the Federal Republic of Germany, Volkswirtschaftliche Diskussionsbeiträge der Universität Münster Nr. 155.

Schmidt, P. (1985/86): Frontier Production Functions, Econometric Reviews 4 (2), 289 - 328.

Schmidt, P./*Sickles,* R. C. (1984): Production Frontiers and Panel Data, Journal of Business and Economic Statistics 2, 367 - 374.

Schumann, J. (1992): Grundzüge der mikroökonomischen Theorie, 6. Auflage, Berlin u.a.

Tamura, R. (1991): Income Convergence in an Endogenous Growth Model, Journal of Political Economy 99, 522 - 540.

Timmer, C. P. (1971): Using a Probabilistic Frontier Production Function to Measure Technical Efficiency, Journal of Political Economy 79, 776 - 794.

Vinod, H. D. (1973): Interregional Comparison of Production Structures, Journal of Regional Sciences 13, 261 - 267.

Weichselberger, A./*Jäckel,* P. (1991): Investitionsaktivitäten westdeutscher Unternehmen in der ehemaligen DDR, Ifo-Schnelldienst 12, 6 - 10.

Werner, G. (1992): Regionalpolitische Konsequenzen der europäischen Integration, Wirtschaftsdienst X, 581 - 589.

Ziegler, A. (1992): Angleichung der Lebensbedingungen zwischen ost- und westdeutschen Regionen als Herausforderung der Strukturpolitik, WSI-Mitteilungen 11/1992, 729 - 735.

Strategien der Investitionspolitik einer Region: Der Fall des Wachstums mit konstanter Sektorstruktur

Von *Reiner Wolff*, Siegen

Wir diskutieren das Problem des effizienten Wachstums heterogener Kapitalstöcke in einer regionalen Ökonomie mit substitutionaler Produktionstechnologie und variablen Skalenerträgen. Wir zeigen, daß die Ökonomie einen von Neumannschen Expansionspfad mit Turnpike-Charakter besitzt, wenn ihre Technologie homothetisch und (additiv) separabel hinsichtlich der Input- und Outputvariablen ist. Darüber hinaus weisen wir nach, daß diese Voraussetzungen bei Hicks-neutralem technischen Fortschritt lokal notwendige Bedingungen für die Existenz und die Turnpike-Eigenschaft des von Neumann-Pfades sind.

I. Einführung

Die formal-analytische Diskussion von Fragen des regionalen Wirtschaftswachstums unter dem Gesichtspunkt der Effizienz der intertemporalen Faktorallokation reicht bis in die 60er Jahre zurück. Seit dieser Zeit haben sich die kontrolltheoretischen Modelle vom sogenannten Rahman-Typ als eine Art Standard für die Ermittlung von Strategien der Investitionspolitik mehrerer Regionen mit dem Ziel der Maximierung zukünftiger Gesamteinkommen herausgebildet (*Rahman* [21, 22], *Intriligator* [15], *Takayama* [25, 26]). Die optimalen Investitionspläne der sektoral aggregierten Rahman-Modelle besitzen ‚bang-bang'-Charakter und sind somit dadurch gekennzeichnet, daß das gemeinsame Investitionsvolumen der einbezogenen Regionen in Zeitintervallen auf jeweils eine Region konzentriert wird. Dabei hängen die Zeitpunkte der Umlenkung der Investitionsströme von einer Region in eine andere von der Höhe der regionalen Investitionsquoten und Kapitalkoeffizienten ab, aber auch vom Stellenwert der Transportkosten im Produktionsprozeß (*Domazlicky* [6]) und der Art der Modellierung des öffentlichen Sektors (*Sakashita* [23], *Ohtsuki* [20], *Bagchi/Olsder/Strijbos* [2]; vgl. dazu auch *Helpman/Pines* [14]). Von der Einbeziehung weiterer Wachstumsziele neben der Zielsetzung der Maximierung zukünftiger Einkommen geht ebenfalls ein wesentlicher Einfluß auf die Abfolge aus, in denen die Investitionen den Regionen zugewiesen werden (*Michel/Pestieau/Thisse* [19], *Friesz/Luque* [11]). Das gleiche gilt bei einer Flexibilisierung der Kapitalkoeffizienten in den Regionen im Wege der Einführung neoklassischer Technologien (*Datta-Chaudhuri* [5], *Fujita* [13, Kap. 6]).

Fujita [12] hat gezeigt, daß im Fall der Produktion mehrerer Güter in den Regionen ein von Neumann-Pfad des gleichschrittigen Wachstums der regionalen Kapitalstöcke mit maximaler Rate existiert, dem die effizienten Wachstumspfade während des größten Teils des Optimierungszeitraumes folgen. Aussagen dieser Form werden allgemein in der Theorie mehrsektoraler Wachstumsmodelle als Turnpike-Theoreme bezeichnet, da dem von Neumannschen Expansionspfad in diesen Theoremen die Rolle einer Schnellstraße des Wachstums zufällt. (Einen Literaturüberblick gibt *McKenzie* [18].) Diese Ergebnisse beruhen aus der Sicht der mikroökonomischen Produktionstheorie zentral auf der Annahme, daß die Menge der Produktionsmöglichkeiten einer Volkswirtschaft einen konvexen Kegel bildet. Damit ist aber insbesondere die Voraussetzung konstanter Skalenerträge verbunden, die vor allem unter regionalwirtschaftlichen Aspekten theoretisch und empirisch nicht befriedigt.

Daher wollen wir in der vorliegenden Arbeit untersuchen, ob das Konzept des ausgewogenen Wachstums überhaupt eine Richtlinie für die Investitionspolitik schon einer *einzelnen* regionalen Ökonomie (*Dorfman* [7]) sein kann, wenn diese Ökonomie ihre Outputs unter den Bedingungen nicht-konstanter Skalenerträge herstellt, die als (positive oder negative) Folgen der räumlichen Konzentration der Produktionsaktivitäten auftreten können. Diese Folgen wollen wir im weiteren kurz als Agglomerationseffekte bezeichnen. Da wir uns auf die Analyse der produktionstheoretischen Voraussetzungen des Turnpike-Theorems in einer räumlichen Volkswirtschaft beschränken wollen, werden wir alle übrigen raumspezifischen Faktoren vernachlässigen.

Im einzelnen ist diese Arbeit wie folgt aufgebaut. Wir werden das Modell einer Region mit substitutionaler Produktionstechnologie vorstellen, die ihre Bestände an Realkapital in einer zukünftigen Periode maximiert (Abschnitt II). Anhand dieses Modells werden wir sodann den Fall des gleichschrittigen Wachstums der Kapitalstöcke allgemein charakterisieren (Abschnitt III). Anschließend zeigen wir, daß es einen Turnpike-Expansionspfad der betrachteten Region gibt, wenn ihre Technologie sowohl im Hinblick auf die Inputs und Outputs (additiv) separabel als auch in bezug auf diese Variablen homothetisch ist (Abschnitt IV). Zuletzt weisen wir nach, daß die soeben genannten Bedingungen bei Hicks-neutralem technischen Fortschritt gleichzeitig lokal notwendige Voraussetzungen für die Anwendung des Turnpike-Theorems auf eine regionale Wirtschaft sind (Abschnitt V). Mit einer kritischen Würdigung der gewonnenen Resultate beschließen wir die Abhandlung (Abschnitt VI).

Wir legen die folgende Notation zugrunde: Elemente des $\mathbb{R}^n$, $n > 1$, werden wir als Spaltenvektoren oder einfach als Vektoren bezeichnen und zu ihrer Veranschaulichung Kleinbuchstaben in fetter Schrift verwenden. Für die Komponente i des Vektors $\mathbf{x}$ schreiben wir einfach x_i. Entsprechend geben

fett gesetzte Großbuchstaben wie zum Beispiel **A** Matrizen mit den Elementen a_{ij} an. Sei ferner $f: S \to T$ eine über $S \in \mathbb{R}^n$ definierte differenzierbare Funktion mit dem Wertebereich $T \in \mathbb{R}$. In diesem Fall benutzen wir f_{x_i} und $f_{x_i x_j}$ als Abkürzungen für die partiellen und gemischt-partiellen Ableitungen $\partial f(\mathbf{x}) / \partial x_i$ und $\partial^2 f(\mathbf{x}) / \partial x_i \partial x_j$ und weiter $f_{\mathbf{x}}$ als Abkürzung für den Gradienten $\nabla f(\mathbf{x})$. Totale Differentiation nach der ‚Zeit' machen wir durch einen Punkt ‚·' kenntlich, und das Superskript ‚T' zeigt die Transponierung eines Vektors oder einer Matrix an. Für den Nullvektor mit jeweils geeigneter Länge reservieren wir das Symbol **o**.

II. Das neoklassische Grundmodell

Wir betrachten zunächst das Modell einer regionalen Volkswirtschaft, die ihren Output mit Hilfe von $n\,(> 1)$ heterogenen Kapitalstöcken $\mathbf{k}^T := (k_1 \ldots k_n)$ erzeugt. Dieser Output wird vollständig für die Kapitalbildung herangezogen und besteht aus Einheiten n verschiedener Netto-Investitionen $\dot{\mathbf{k}}^T := (\dot{k}_1 \ldots \dot{k}_n)$.[1] Die Produktionstechnologie der Volkswirtschaft läßt sich in jeder Periode t durch eine Transformations-Randfunktion der Form

$$T(\mathbf{k}, \dot{\mathbf{k}}, t) = 0 \tag{1}$$

beschreiben, in der die Zeitvariable t den exogenen technischen Fortschritt repräsentiert. Die Funktion $T(\cdot)$ soll die folgenden Eigenschaften haben:

(P1) Sie ist über ihrem Definitionsbereich $\mathbb{R}^n_+ \times \mathbb{R}^n \times \mathbb{R}_+$ stetig differenzierbar.

(P2) Sie besitzt nicht verschwindende partielle Ableitungen und Kreuzableitungen zweiter Ordnung in bezug auf die Komponenten von **k**. Eine entsprechende Annahme gilt für $\dot{\mathbf{k}}$.

(P3) Sie ist steigend in **k** und fallend in $\dot{\mathbf{k}}$.

(P4) Sie ist jeweils streng quasi-konkav im Hinblick auf **k** und $\dot{\mathbf{k}}$.

(P5) Alle Kapitalstöcke genügen den Inada-Regularitätsbedingungen $\lim_{k_i \to 0} T_{k_i} = \infty$ und $\lim_{k_i \to \infty} T_{k_i} = 0$.

Dadurch wird eine neoklassische Produktionstechnologie definiert, die die üblichen Merkmale fallender Grenzraten der Substitution zwischen den Kapitalstöcken und steigender Grenzraten der Transformation zwischen den produzierten Investitionsgütern aufweist. Allerdings beinhalten die genannten

[1] Zur Vereinfachung der Darlegungen werden wir nachfolgend nicht zwischen Brutto- und Netto-Investitionen unterscheiden. Erstere können modelliert werden, indem man anstelle von $\dot{\mathbf{k}}$ den Ausdruck $\dot{\mathbf{k}} + \text{diag}\,(\mathbf{d})\,\mathbf{k}$ ansetzt, worin **d** den Vektor der (konstanten) Abschreibungsraten angibt.

Eigenschaften keine Annahme über die Entwicklung der Skalenerträge im Produktionsprozeß. Somit ist insbesondere denkbar, daß die Skalenerträge im Zeitablauf aufgrund von Agglomerationseffekten variieren. Daher beschreibt (1) den Rand einer in ihren Input- und Outputzweigen jeweils konvexen, aber nicht notwendig global konvexen Technologie, für die *Lau* [17] den Begriff der bikonvexen Technologie in die Literatur eingeführt hat. Wir werden die Funktion $T(\mathbf{k}, \dot{\mathbf{k}}, t)$ im weiteren einfach als die *Transformationsfunktion* der betrachteten Volkswirtschaft bezeichnen.

Wir wollen nun annehmen, daß die Volkswirtschaft das Ziel verfolgt, in einem endlichen und abgeschlossenen Zeitintervall $[t_0, t_1]$ eine positive Linearkombination der Endniveaus der Kapitalstöcke $\mathbf{k}(t_1)$ bei gegebener Anfangsausstattung $\mathbf{k}_0$ zu maximieren:

$$(2) \qquad \max_{\mathbf{k}(t),\, t\in[t_0,t_1]} \mathbf{p}^T \mathbf{k}(t_1)$$

$$\text{mit } T(\mathbf{k}, \dot{\mathbf{k}}, t) = 0 \text{ und } \mathbf{k}(t_0) = \mathbf{k}_0.$$

Darin kann $\mathbf{p}$ ökonomisch als ein Vektor (diskontierter) Kapitalgüterpreise interpretiert werden, der für die Periode t_1 erwartet wird. In diesem Fall liegt ein klassisches Investitionsproblem vor: Die Volkswirtschaft ist bestrebt, den erwarteten Gegenwartswert ihrer Kapitalstöcke mit Hilfe einer gegebenen Technologie und gegebener Anfangsbestände zu maximieren.

Analytisch handelt es sich bei der formulierten dynamischen Optimierungsaufgabe um ein Mayersches Variationsproblem. Wir unterstellen, daß dieses Problem eine innere Lösung in der Klasse der nicht-negativen und zweifach differenzierbaren Funktionen $\mathbf{k}(t)$ für $t \in [t_0, t_1]$ besitzt und daß für mindestens einen Kapitalstock i die Ungleichung $k_i(t_1) > k_i(t_0)$ zutrifft. Da ferner die Funktion $T(\cdot)$ nach Voraussetzung quasi-konkav in $\mathbf{k}$ ist, wird die Legendre-Bedingung für ein Maximum von $\mathbf{p}^T \mathbf{k}(t_1)$ global erfüllt. Somit ist die Eindeutigkeit der angenommenen Lösung gewährleistet (vgl. *Kamien/Schwartz* [16]).

III. Eine Charakterisierung stationärer Kapitalstrukturen

Aus der Literatur ist bekannt, daß die optimalen Wachstumspfade $\mathbf{k}(t)$ die sogenannten Eigenzinssatzgleichungen der Theorie des optimalen Wachstums befriedigen, die wiederum äquivalent zu den Eulerschen Gleichungen des Optimierungsproblems (2) sind:

PROPOSITION 1: *Ist* $\mathbf{k}(t)$ *eine Lösung der Optimierungsaufgabe (2), so gilt*

$$\frac{T_{k_i}}{T_{\dot{k}_i}} = \frac{T_{k_n}}{T_{\dot{k}_n}} + \frac{d}{dt}\log\left(\frac{T_{k_i}}{T_{\dot{k}_n}}\right) \quad \text{für alle } i \neq n. \tag{3}$$

BEWEIS: Vgl. *Wan*, Jr. [27, S. 277 - 278].

Die Gestalt der optimalen Wachstumspfade $\mathbf{k}(t)$ wird durch die somit vorliegenden $n-1$ Differentialgleichungen (3) unter Berücksichtigung der Randbedingungen und der Technologie-Restriktion (1) vollständig determiniert. Wir werden im weiteren $\mathbf{k} > \mathbf{o}$ (komponentenweise) für alle $t \in [t_0, t_1]$ annehmen. Ferner sei $\mathbf{s}^T := (k_1/k_n \ldots k_{n-1}/k_n)$. Der uns interessierende Fall des effizienten gleichgewichtigen (im Sinne eines gleichschrittigen) Wachstums der Kapitalstöcke läßt sich dann in der Form $\mathbf{k}(t) = k_n(t)\,(\bar{\mathbf{s}}^{*T}, 1)^T$ angeben, worin $\bar{\mathbf{s}}^*$ für einen Vektor von $n-1$ positiven Konstanten steht. Eine solche Lösung kann dann wie folgt charakterisiert werden:

PROPOSITION 2: *Jeder Strahl $\bar{\mathbf{s}}^*$ des effizienten gleichgewichtigen Wachstums der Kapitalstöcke mit den Raten $w(t)$ (> 0) genügt den Bedingungen*

$$-\frac{T_{k_1}}{T_{\dot{k}_1}} = \ldots = -\frac{T_{k_n}}{T_{\dot{k}_n}} = w + \frac{\dot{w}}{w} - \Theta(t), \tag{4}$$

in denen $\Theta(t)$ die relative Änderung der gemeinsamen Wachstumsrate w durch technischen Fortschritt wiedergibt.

BEWEIS: Wir beachten, daß $\dot{\mathbf{k}} = w\mathbf{k}$ entlang $\bar{\mathbf{s}}^*$ und folglich $T(\mathbf{k}, w\mathbf{k}, t) = 0$. Da $T(\cdot)$ fallend in $\dot{\mathbf{k}}$ ist, können wir nach w als Funktion von $\mathbf{k}$ und t auflösen: $w = f(\mathbf{k}, t)$. Somit erhalten wir wegen $\mathbf{k} = k_n\,(\bar{\mathbf{s}}^{*T}, 1)^T$ den Ausdruck $w = f(k_n\bar{\mathbf{s}}^*, k_n, t)$. Wir beachten weiterhin, daß ein gegebener Expansionspfad $\bar{\mathbf{s}}$ nur dann effizient sein kann, wenn es sich zu keiner Zeit im Verlaufe des Wachstumsprozesses lohnt, den Pfad zu verlassen. Mit anderen Worten: Jeder effiziente Strahl $\bar{\mathbf{s}}^*$ weist stets in die Richtung des maximalen Anstiegs der Expansionsrate w (beziehungsweise in die Richtung ihrer minimal möglichen Reduktion). Daher können wir entlang $\bar{\mathbf{s}}^*$ den Umhüllenden-Satz anwenden:

$$\begin{aligned} -\frac{T_{k_n}}{T_{\dot{k}_n}} &= \frac{\partial \dot{k}_n}{\partial k_n} = \frac{\partial (w k_n)}{\partial k_n} = \sum_{i=1}^{n-1} f_{k_i}\bar{s}_i^* k_n + f_{k_n} k_n + w \\ &= \sum_{i=1}^{n} f_{k_i} k_i + w. \end{aligned} \tag{5}$$

Analoge Relationen müssen für die übrigen Kapitalstöcke gelten, da der Index n jedem beliebigen Kapitalstock zugeordnet werden kann. Zuletzt ermittelt man durch totale Differentiation der Funktion $f(\cdot)$ nach t:

$$\dot{w} = \sum_{i=1}^{n} f_{k_i} \dot{k}_i + \frac{\partial f}{\partial t} = w \left(\sum_{i=1}^{n} f_{k_i} k_i + \frac{1}{w} \frac{\partial f}{\partial t} \right). \tag{6}$$

Wir definieren nun $\Theta(t) := \frac{1}{w} \frac{\partial f}{\partial t}$ und lösen nach dem Summenterm im rechten Klammerausdruck von (6) auf. Einsetzen in (5) bestätigt dann die in Proposition 2 aufgestellte Behauptung. □

Proposition 2 impliziert insbesondere, daß die Grenzrate der Substitution zwischen zwei beliebigen Kapitalstöcken entlang $\bar{\mathbf{s}}^*$ mit der Grenzrate der Transformation zwischen den zugehörigen Investitionen übereinstimmt. (Man beachte die ersten $n-1$ Gleichheitszeichen in (4) und forme um.) Die Transformationsraten wiederum stellen sich als Konstanten ein, da der zweite Ausdruck auf der rechten Seite von (3) aufgrund der Gleichungen (4) zu Null wird. Dadurch wird eine invariante Beziehung zwischen $\bar{\mathbf{s}}^*$ und den volkswirtschaftlichen Grenzraten der Substitution und der Transformation begründet. Somit erhalten wir

KOROLLAR 1: *Entlang eines Strahls $\bar{\mathbf{s}}^*$ des effizienten gleichschrittigen Wachstums der Kapitalstöcke stimmen die volkswirtschaftlichen marginalen Substitutionsraten und zugehörigen Transformationsraten überein:*

$$\frac{T_{k_i}}{T_{k_n}} = \frac{T_{\dot{k}_i}}{T_{\dot{k}_n}} \quad \textit{und} \quad \frac{d}{dt}\left(\frac{T_{k_i}}{T_{k_n}}\right) = \frac{d}{dt}\left(\frac{T_{\dot{k}_i}}{T_{\dot{k}_n}}\right) = 0 \quad \textit{für alle } i \neq n. \tag{7}$$

Wir wollen nun für einen Moment annehmen, daß die Transformationsfunktion $T(\cdot)$ sowohl zeitunabhängig als auch homogen vom Grade Eins in bezug auf $\mathbf{k}$ und $\dot{\mathbf{k}}$ ist. Das ist der Fall einer stationären Technologie mit konstanten Skalenerträgen, den zuerst *Dorfman/Samuelson/Solow* [8, Kap. 12] untersucht haben. In diesem Fall ist $\Theta(t) \equiv \dot{w}(t) \equiv 0$ und wir können schließen, daß die Grenzraten der Substitution und Transformation zwischen Inputs und Outputs entlang $\bar{\mathbf{s}}^*$ konstant bleiben. Die Bedingungen (7) konstituieren dann als $\bar{\mathbf{s}}^*$ den von Neumann-Pfad des gleichschrittigen exponentiellen Wachstums der betrachteten Volkswirtschaft mit maximaler Rate. Von diesem Pfad ist bekannt, daß er die Eigenschaften eines Sattelpunktes besitzt und daher ein Turnpike für solche Wachstumspfade $\mathbf{k}(t)$ ist, die nicht entlang dem Strahl $\bar{\mathbf{s}}^*$ starten oder enden.

Allerdings werden in anderen Fällen von Technologien die Quotienten T_{k_i}/T_{k_n} und $T_{\dot{k}_i}/T_{\dot{k}_n}$ $(i = 1, \ldots, n-1)$ in der Regel sowohl von der ‚Zeitvariablen' t abhängen (infolge technischen Fortschritts) als auch vom ‚Niveau' k_n der ökonomischen Aktivitäten (infolge variabler Skalenerträge). Daher werden sie selbst dann im Zeitablauf steigen oder fallen, wenn die Struktur des volkswirtschaftlichen Kapitalstocks unverändert bleibt. In diesen Fällen wird es keinen Pfad des gleichschrittigen Wachstums geben, der als

effizient im Sinne der Bedingungen (3) gelten kann. Mit anderen Worten: Das Turnpike-Theorem gilt nicht ohne weiteres für eine regionale Volkswirtschaft mit zeitabhängiger Technologie und variablen Skalenerträgen.

Das bedeutet, daß in einer regionalen Wirtschaft die Existenz eines Turnpike-Pfades des gleichschrittigen Wachstums der Kapitalstöcke mit variabler Rate zusätzliche Annahmen über die Produktionstechnologie erfordert. Damit stellt sich insbesondere die Frage nach funktionalen Formen der Transformationsfunktion $T(\cdot)$, die mit (7) vereinbar sind.

IV. Separable Produktionstechnologien

Wir haben im vorangegangenen Abschnitt darauf hingewiesen, daß in einer regionalen Volkswirtschaft ein Turnpike-Wachstumspfad nicht ohne weiteres existiert. Dafür sind neben dem technischen Fortschritt vor allem Agglomerationseffekte verantwortlich, die sich im Produktionsprozeß in variablen Skalenerträgen niederschlagen und auf diesem Wege die Grenzraten der Faktorsubstitution und Outputtransformation beeinflussen. Daher fordern wir nun, daß die Technologie der betrachteten Volkswirtschaft weitergehenden Voraussetzungen genügt. Die erste Voraussetzung betrifft den technischen Fortschritt:

(P6) Der technische Fortschritt ist Hicks-neutral.

Ferner wollen wir annehmen, daß die Technologie unabhängige Input- und Outputzweige besitzt. Insbesondere soll gelten:

(P7) $T(\cdot)$ ist sowohl (additiv) separabel als auch homothetisch mit Bezug zu $\mathbf{k}$ auf der einen Seite und $\dot{\mathbf{k}}$ auf der anderen Seite.

(Man beachte, daß (P6) durch (P7) impliziert wird.) Diese Annahmen sind hinreichend für die Konstanz der Substitutions- und Transformationsraten T_{k_i}/T_{k_n} und $T_{\dot{k}_i}/T_{\dot{k}_n}$ $(i = 1, \ldots, n-1)$ entlang beliebiger Expansionspfade $\bar{\mathbf{s}}$. Damit können wir nun die folgende theoretische Aussage formulieren:

THEOREM 1: *Wenn die Transformationsfunktion (1) die Gestalt*

$$T(\mathbf{k}, \dot{\mathbf{k}}, t) = J(G(\mathbf{k}), t) - H(\dot{\mathbf{k}}) \tag{8}$$

mit linear-homogenen Funktionen $G(\cdot)$ und $H(\cdot)$ annimmt, dann gibt es einen eindeutigen Vektor von positiven Konstanten $\bar{\mathbf{s}}^*$ mit der folgenden Eigenschaft: Er konstituiert

1. *einen Expansionspfad mit globalem Turnpike-Charakter für $n = 2$, und zwar im Hinblick auf alle optimalen Investitionsstrategien, die im gesamten Intervall $[t_0, t_1]$ der Ungleichung $\dot{k}_2/k_2 > 0$ genügen, so daß mindestens ein (geeignet indizierter) Kapitalstock kontinuierlich steigt.*

2. *einen Expansionspfad mit lokalem Turnpike-Charakter für beliebiges* n (≥ 2).

BEWEIS: Die Existenz und Eindeutigkeit eines effizienten Expansionspfades $\bar{\mathbf{s}}^*$ gehen unmittelbar aus unserer Diskussion im Zusammenhang mit Korollar 1 und den Annahmen (P1) - (P4) hervor. Darüber hinaus ist aufgrund der Annahme (P5) gewährleistet, daß alle Komponenten des Vektors $\bar{\mathbf{s}}^*$ streng positiv sind. Damit bleibt lediglich zu zeigen, daß dieser Vektor die Merkmale eines Sattelpunktes besitzt.

1. Teil: $n = 2$, $\dot{k}_2/k_2 > 0$. In diesem Fall reduziert sich der Vektor $\mathbf{s}$ auf ein Skalar s. Wir definieren nun $u := T_{k_1}/T_{k_2} = G_{k_1}/G_{k_2} =: a(s)$ sowie $v := T_{\dot{k}_1}/T_{\dot{k}_2} = H_{\dot{k}_1}/H_{\dot{k}_2} =: b(\dot{k}_1/\dot{k}_2)$. Weiter sei $\bar{u}^* := a(\bar{s}^*)$ und $\bar{v}^* := b(\bar{s}^*)$. Da aufgrund der angenommenen Quasi-Konkavität der Transformationsfunktion in $\mathbf{k}$ und $\dot{\mathbf{k}}$ (P4) eine umkehrbar eindeutige Beziehung zwischen der stationären Kapitalstruktur $\bar{s}^*$ und den zugehörigen Substitutions- und Transformationsraten $\bar{u}^*$ und $\bar{v}^*$ besteht, können wir den Nachweis der (globalen) Turnpike-Eigenschaft von $\bar{s}^*$ indirekt führen, indem wir stattdessen bestätigen, daß der Punkt $(\bar{u}^*, \bar{v}^*)$ ein (globaler) Sattelpunkt ist.

Zu diesem Zweck ermitteln wir in einem ersten Schritt zunächst aus $u = a(s)$ und $s = k_1/k_2$ im Wege der Differentiation:

$$\dot{u} = a'\dot{s} = a' \frac{\dot{k}_1 k_2 - k_1 \dot{k}_2}{k_2^2} = w_2(t) a' \left(\frac{\dot{k}_1}{\dot{k}_2} - s \right). \tag{9}$$

Darin ist $w_2(t) := \dot{k}_2/k_2$. Ferner dürfen wir wiederum aufgrund der Modellannahme (P4) $a', b' \neq 0$ voraussetzen und deshalb die Größen $\dot{k}_1/\dot{k}_2$ und s über die Umkehrfunktionen $b^{-1}(v)$ und $a^{-1}(u)$ ausdrücken:

$$\dot{u} = w_2(t) a' [b^{-1}(v) - a^{-1}(u)]. \tag{10}$$

Im zweiten Schritt bringen wir die Variablen u und v in die Optimumbedingung (3) ein und beachten die Differentiationsregel $d \log (\cdot)/dt = (\cdot)^{-1} d(\cdot)/dt$. Einfache Umformung ergibt dann:

$$\dot{v} = \lambda(t)(u - v) \tag{11}$$

mit $\lambda(t) = T_{k_2}(\mathbf{k}(t), \dot{\mathbf{k}}(t), t)/T_{\dot{k}_2}(\mathbf{k}(t), \dot{\mathbf{k}}(t), t) < 0$. Die somit gewonnenen Relationen (10) und (11) bilden ein System zweier nichtlinearer und explizit zeitabhängiger Differentialgleichungen erster Ordnung in den Variablen u und v. Der Punkt $(\bar{u}^*, \bar{v}^*)$ mit $a^{-1}(\bar{u}^*) = b^{-1}(\bar{v}^*)$ und $\bar{u}^* = \bar{v}^*$ stellt eine stationäre Lösung dieser Differentialgleichungen dar.

An dieser Stelle ist nun zu beachten, daß aufgrund der Annahmen (P3) und (P4) $a(\cdot)$ eine fallende Funktion von s und $b(\cdot)$ eine steigende Funktion von

$\dot{k}_1/\dot{k}_2$ ist. Folglich ist $a' < 0$. Ferner sind die Umkehrfunktionen $a^{-1}(\cdot)$ und $b^{-1}(\cdot)$ fallend in u beziehungsweise steigend in v. Somit stellt der Ausdruck $b^{-1}(v) - a^{-1}(u) = 0$ eine inverse Beziehung zwischen den Variablen u und v her. Da nun aber nach Voraussetzung $w_2(t) > 0$ gilt, erhalten wir

$$(12) \qquad \dot{u} \begin{cases} > 0 \text{ für } b^{-1}(v) < a^{-1}(u) \\ = 0 \text{ für } b^{-1}(v) = a^{-1}(u) \\ < 0 \text{ für } b^{-1}(v) > a^{-1}(u) \end{cases}, \quad \dot{v} \begin{cases} > 0 \text{ für } u < v \\ = 0 \text{ für } u = v \\ < 0 \text{ für } u > v \end{cases}.$$

Diesen Beziehungen entspricht im Phasendiagramm das Strömungsbild eines Sattelpunktes. Damit ist der Beweis des ersten Teiles von Theorem 1 abgeschlossen. □

Bevor wir den noch ausstehenden Teil unseres Theorems beweisen, wollen wir kurz auf eine interessante Implikation der bisherigen Diskussion für die Konstellation $n > 2$ hinweisen. Diese Implikation betrifft den Spezialfall, daß $G(\cdot)$ und $H(\cdot)$ additive Funktionen vom ACMS-Typ mit konstanter Substitutions- und Transformationselastizität sind. Die Quotienten T_{k_i}/T_{k_n} und $T_{\dot{k}_i}/T_{\dot{k}_n}$ hängen dann nur von den Größen s_i und $\dot{k}_i/\dot{k}_n$ ab, so daß $n-1$ voneinander unabhängige Gleichungssysteme vom Typ (10) - (11) entstehen. Damit können wir festhalten:

KOROLLAR 2: *Wenn die Funktionen $G(\cdot)$ und $H(\cdot)$ vom ACMS-Typ sind, ist der Expansionspfad $\hat{\mathbf{s}}^*$ unabhängig von n ein globaler Turnpike-Pfad für alle optimalen Investitionsstrategien, die im gesamten Intervall $[t_0, t_1]$ der Bedingung $\dot{k}_n/k_n > 0$ genügen, also gewährleisten, daß mindestens ein (mit n geeignet indizierter) Kapitalstock stetig wächst.*

2. Teil: $n (\geq 2)$ beliebig. An die Stelle der Skalare u und v treten nun Vektoren $\mathbf{u}$ und $\mathbf{v}$ der Länge $n-1$. Entsprechend ersetzen wir $a(s)$ und $b(\dot{k}_1/\dot{k}_2)$ durch vektorwertige Funktionen $\mathbf{a}(\mathbf{s})$ und $\mathbf{b}(\dot{k}_1/\dot{k}_n, \ldots, \dot{k}_{n-1}/\dot{k}_n)$. Das Gleichungssystem (10) - (11) geht dann in

$$(13) \qquad \dot{\mathbf{u}} = w_n(t)\,\mathbf{A}\,[\mathbf{b}^{-1}(\mathbf{v}) - \mathbf{a}^{-1}(\mathbf{u})], \quad \dot{\mathbf{v}} = \lambda(t)\,(\mathbf{u} - \mathbf{v})$$

mit entsprechend definierten $w_n(t)$ und $\lambda(t)$ und der (regulären) Jacobi-Matrix $\mathbf{A}$ der Funktionen $\mathbf{a}(\cdot)$ über. Wir können jetzt im allgemeinen keine globalen Resultate mehr erwarten und beschränken uns deshalb auf eine lokale Analyse der Sattelpunkteigenschaft der stationären Lösung $(\bar{\mathbf{u}}^*, \bar{\mathbf{v}}^*)$. Dies geschieht, indem wir die Gleichungen (13) durch eine Taylor-Entwicklung zweiter Ordnung in der Umgebung dieses Punktes approximieren. Dabei beachten wir die Beziehungen $\dot{\bar{\mathbf{u}}}^* \equiv \dot{\bar{\mathbf{v}}}^* \equiv \mathbf{o}$ und $\mathbf{b}^{-1}(\bar{\mathbf{v}}^*) = \mathbf{a}^{-1}(\bar{\mathbf{u}}^*)$ und nähern die Wachstumsrate $w_n(t)$ durch $w(t)$ an. Darüber hinaus berücksichtigen wir den Tatbestand, daß das Produkt der Jacobi-Matrizen von $\mathbf{a}(\cdot)$ und $\mathbf{a}^{-1}(\cdot)$ die Einheitsmatrix ergibt. Wir gelangen dann zu einem System von

2 $(n-1)$ linearen Differentialgleichungen erster Ordnung mit zeitabhängigen Koeffizienten:

$$\dot{\mathbf{x}} = w(t)\,(\mathbf{A}\mathbf{B}^{-1}\mathbf{y} - \mathbf{x}), \quad \dot{\mathbf{y}} = \lambda(t)\,(\mathbf{x} - \mathbf{y}). \tag{14}$$

Darin stehen $\mathbf{x}$ und $\mathbf{y}$ für die Abweichungen $\mathbf{u} - \bar{\mathbf{u}}^*$ und $\mathbf{v} - \bar{\mathbf{v}}^*$, und die Matrix $\mathbf{B}$ ist die (reguläre) Jacobi-Matrix von $\mathbf{b}(\cdot)$. Der stationären Lösung $(\bar{\mathbf{u}}^*, \bar{\mathbf{v}}^*)$ von (13) entspricht jetzt der stationäre Punkt $(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*) = (\mathbf{o}, \mathbf{o})$ des linearisierten Systems (14).

An dieser Stelle ist der Hinweis wichtig, daß $(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*) = (\mathbf{o}, \mathbf{o})$ auch ein stationärer Punkt des autonomen Systems

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{B}^{-1}\mathbf{y} - \mathbf{x}, \quad \dot{\mathbf{y}} = \mathbf{y} - \mathbf{x}, \tag{15}$$

ist. Dieser Tatbestand erweist sich insbesondere aus dem folgenden Grund als hilfreich:

LEMMA 1: *Wenn* $(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*) = (\mathbf{o}, \mathbf{o})$ *ein Sattelpunkt der autonomen Gleichungen (15) ist, dann ist er auch ein Sattelpunkt des zeitabhängigen Systems (14).*

BEWEIS: Siehe Appendix.

Das zu (15) gehörende charakteristische Polynom lautet:

$$|\mathbf{A}\mathbf{B}^{-1} - (1 - \eta^2)\mathbf{I}| = 0. \tag{16}$$

Für dessen Wurzeln gilt nun wegen (P3) - (P4):

LEMMA 2: *Wenn* η *eine Wurzel von (16) ist, dann ist auch* $-\eta$ *eine Wurzel von (16). Beide Wurzeln sind reell und ungleich Null.*

BEWEIS: Siehe Appendix.

Damit haben wir gezeigt, daß die Wurzeln des charakteristischen Polynoms (16) reell und von Null verschieden sind und ferner in Paaren betragsmäßig gleicher Wurzeln mit unterschiedlichen Vorzeichen auftreten. Daher ist $(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*) = (\mathbf{o}, \mathbf{o})$ ein symmetrischer Sattelpunkt von (15) und somit wegen Lemma 1 auch ein Sattelpunkt von (14). Das bedeutet aber wiederum, daß $(\bar{\mathbf{u}}^*, \bar{\mathbf{v}}^*)$ die Eigenschaft eines lokalen Sattelpunktes für das Ausgangssystem (13) besitzt. Aufgrund der umkehrbar eindeutigen Beziehung zwischen $(\bar{\mathbf{u}}^*, \bar{\mathbf{v}}^*)$ und $\bar{\mathbf{s}}^*$ können wir zuletzt auch auf den (lokalen) Turnpike-Charakter des Expansionspfades $\bar{\mathbf{s}}^*$ schließen und die Beweisführung beenden. □

V. Lokale Abbildungen regionaler Turnpike-Technologien

Im letzten Abschnitt haben wir *hinreichende* Bedingungen für die Gültigkeit des Turnpike-Theorems in einer regionalen Volkswirtschaft angegeben. Den ökonomischen Kern dieser Bedingungen bildet die gewählte Beschreibung der Produktionstechnologie in der Form (8). Damit stellt sich sofort die Frage, ob es noch andere funktionale Formen der Transformationsfunktion gibt, die bei variablen Skalenerträgen zu vergleichbaren Resultaten führen. Noch weiter gedacht: Können wir *notwendige* Anforderungen an die Transformationsfunktion einer regionalen Ökonomie formulieren? Dieser Frage wollen wir uns nun unter zwei einschränkenden Voraussetzungen zuwenden. Zunächst einmal wollen wir uns nur mit solchen Anforderungen befassen, die in der Umgebung eines Expansionsstrahls $\bar{\mathbf{s}}^*$ erfüllt sein müssen. Diese Einschränkung liegt insofern nahe, als das Turnpike-Theorem in der Regel ebenfalls nur eine lokale Stabilitätsaussage beinhaltet. Daneben wollen wir zur Vereinfachung a priori unterstellen, daß der technische Fortschritt neutral sowohl im Hinblick auf die Faktorsubstitution als auch hinsichtlich der Outputtransformation ist. Wir gelangen dann zu folgendem Ergebnis:

THEOREM 2: *Wenn ein Turnpike-Pfad* $\bar{\mathbf{s}}^*$ *existiert und entlang diesem Pfad alle Substitutions- und Transformationsraten zwischen den Inputs beziehungsweise Outputs von der Zeitvariablen t unabhängig sind, dann kann man die Produktionstechnologie lokal durch eine Randfunktion der Form (8) abbilden.*

BEWEIS: Wir definieren zunächst $\tilde{T}(k_n, \dot{\mathbf{k}}, t) := T(k_n \bar{\mathbf{s}}^*, k_n, \dot{\mathbf{k}}, t) = 0$. Da $\tilde{T}(\cdot)$ eine steigende Funktion von k_n ist, läßt sie sich nach dieser Variablen in Abhängigkeit von $\dot{\mathbf{k}}$ und t auflösen: $k_n = \tilde{F}(\dot{\mathbf{k}}, t)$. Nach Voraussetzung hängt nun aber keine der Transformationsraten zwischen den Outputs von t ab. Daher folgt wegen (7) und der Quasi-Konkavität von $T(\cdot)$ in bezug auf $\dot{\mathbf{k}}$, daß $\tilde{F}(\cdot)$ eine homothetische und quasi-konvexe Funktion von $\dot{\mathbf{k}}$ ist. (Vgl. *Färe* [9] und [10, S. 49 - 61]; siehe auch *Shephard* [24].) Somit gilt $k_n = F(H(\dot{\mathbf{k}}), t) = F(wH(\mathbf{k}), t)$, worin $H(\cdot)$ ohne Beschränkung der Allgemeinheit als linear-homogen angenommen werden darf. Mithin erhält man durch Multiplikation beider Seiten der in Abschnitt 3 eingeführten Funktion $w = f(\mathbf{k}, t)$ mit $H(\mathbf{k})$ den Ausdruck

$$wH(\mathbf{k}) = H(\dot{\mathbf{k}}) = f(\mathbf{k}, t)\, H(\mathbf{k}) =: \tilde{J}(\mathbf{k}, t). \tag{17}$$

Er separiert $T(\cdot)$ in einen homothetischen und quasi-konvexen Outputzweig $H(\cdot)$ und einen verbleibenden Inputzweig $\tilde{J}(\cdot)$. An dieser Stelle beachten wir, daß die Gradienten der Funktionen $H_{\dot{\mathbf{k}}}$ und $\tilde{J}_{\mathbf{k}}$ wegen (7) stets in die gleiche Richtung weisen. Weiter stellen wir in Rechnung, daß der technische Fortschritt neutral in bezug auf die Faktorsubstitution sein soll und daß $T(\cdot)$

quasi-konkav in $\mathbf{k}$ ist. Demnach muß $\tilde{J}(\cdot)$ sowohl homothetisch als auch quasi-konkav in $\mathbf{k}$ sein: $\tilde{J}(\mathbf{k}, t) = J(G(\mathbf{k}), t)$. Folglich erhalten wir aus (17) eine hinsichtlich der Vektoren $\mathbf{k}$ und $\dot{\mathbf{k}}$ quasi-konkave Abbildungsvorschrift der Form

$$T(\mathbf{k}, \dot{\mathbf{k}}, t) = J(G(\mathbf{k}), t) - H(\dot{\mathbf{k}}) \tag{18}$$

mit linear-homogenen Funktionen $H(\cdot)$ und $G(\cdot)$. Dieses Resultat entspricht der aufgestellten Behauptung. □

VI. Kritische Würdigung aus regionalökonomischer Sicht

Unsere Analyse hat zwei wesentliche Ergebnisse hervorgebracht. Wir haben gezeigt, daß das Turnpike-Theorem auch für den Fall einer regionalen Akkumulationswirtschaft mit einer zeitabhängigen bikonvexen Produktionstechnologie und variablen Skalenerträgen gilt, wenn diese Technologie (additiv) trennbar und homothetisch in bezug auf die Kapitalstöcke und Investitionen ist. Ferner haben wir gesehen, daß diese Eigenschaften gleichzeitig lokal notwendige Voraussetzungen für die Gültigkeit der Turnpike-Aussage ist, sobald der technische Fortschritt den Kriterien der Hicks-Neutralität genügt. Im Hinblick auf eine ökonomische Bewertung dieser Resultate rücken zwei Aspekte in den Vordergrund.

Zunächst einmal sollten wir darauf hinweisen, daß wir die Zeitvariable t in der Transformationsfunktion der betrachteten Ökonomie inhaltlich weitergehender interpretieren dürfen als bisher geschehen. Diese Variable kann nämlich neben dem exogenen technischen Fortschritt auch andere im Zeitablauf gegebene ökonomische Größen wie zum Beispiel Arbeitsinputs und Konsumoutputs repräsentieren. Davon bleiben unsere Theoreme solange qualitativ unberührt, wie solche Größen sich in einer entsprechend erweiterten Transformationsfunktion von den Kapitalstöcken und Investitionen separieren lassen. Dabei ist völlig unerheblich, aus welchem Entscheidungsprozeß die zeitlichen Entwicklungspfade der zusätzlichen ökonomischen Variablen hervorgegangen sind. Insbesondere wäre denkbar und zulässig, daß diese Pfade optimale Pfade im Sinne eines umfassenderen intertemporalen Nutzenkriteriums sind.

Dieser Tatbestand macht einmal mehr deutlich, daß die Annahme der Separabilität im allgemeinen starke Implikationen für die Struktur der Lösung eines ökonomischen Modells nach sich zieht. Unter anderem deshalb hat diese Annahme in der angewandten Ökonomik seit jeher einen großen Stellenwert. Allerdings haben wir uns in der vorliegenden Arbeit nicht mit der schwierigen Frage nach der Existenz von Aggregatorfunktionen wie $G(\cdot)$ und $H(\cdot)$ befassen können. (Wir verweisen den interessierten Leser auf

Blackorby/Schworm [3].) Jedenfalls scheint es sich aus dem Blickwinkel der mikroökonomischen Aggregationstheorie bei der durch (8) repräsentierten Produktionstechnologie um eine recht spezielle Technologie zu handeln. Insoweit ist die vorliegende Arbeit eher ernüchternd: Das Konzept der gleichschrittigen Ausweitung der ökonomischen Aktivitäten einer arbeitsteiligen regionalen Volkswirtschaft mit signifikanten Agglomerationseffekten erscheint unter dem Gesichtspunkt der intertemporalen Allokationseffizienz als eine eher suboptimale Strategie des wirtschaftlichen Wachstums.

Appendix: Beweise der Lemmata 1 - 2

Beweis von Lemma 1: Wir betrachten im $(\mathbf{x}, \mathbf{y})$-Raum die durch (14) beziehungsweise (15) implizit definierten Phasentrajektorien. Wegen $w(t) > 0$ und $\lambda(t) < 0$ stimmt die Menge $S = \{(\mathbf{x}, \mathbf{y}) \mid \dot{\mathbf{x}} = \mathbf{o} \text{ oder } \dot{\mathbf{y}} = \mathbf{o}\}$ in beiden dynamischen Systemen überein. Aus dem gleichen Grund gilt für alle anderen Punkte $(\mathbf{x}, \mathbf{y}) \notin S$, daß die Vorzeichen von $\dot{\mathbf{x}}$ und $\dot{\mathbf{y}}$ sich nicht ändern, wenn wir von (15) zu (14) wechseln. Somit bleibt die topologische Struktur des (mehrdimensionalen) Strömungsdiagramms erhalten. □

Beweis von Lemma 2: Die erste Aussage dieses Satzes resultiert unmittelbar aus dem Tatbestand, daß η quadratisch in (16) auftritt. Für die Bestätigung der zweiten Aussage sind weitergehende Überlegungen erforderlich. Deshalb halten wir zunächst fest, daß die Matrix $\mathbf{A}$, deren Komponenten wir in der Form $a_{ij} = \partial a_i / \partial s_j = k_n \, \partial (G_{k_i} / G_{k_n}) / \partial k_j = -k_n \, (\partial^2 k_n / \partial k_i \, \partial k_j)$ schreiben können, der Hesse-Matrix der wegen (P3) - (P4) streng konvexen Isoquantenfunktion für den Kapitalstock k_n in bezug auf $k_1, \ldots, k_{n-1}$ negativ proportional und somit streng negativ definit ist. Entsprechend können wir argumentieren, daß die Matrix $\mathbf{B}$ und mithin auch $\mathbf{B}^{-1}$ streng positiv definit ist.

Der angestrebte Beweis ließe sich nun im Prinzip indirekt mit Hilfe eines Satzes von *Arrow* [1, S. 200] über die Eigenwerte des Produktes einer positiven quasi-definiten Matrix mit einer symmetrischen Matrix führen (*Das-Gupta* [4, S. 315]). Stattdessen wollen wir hier einen direkten Beweis angeben, der sich unmittelbar die Definitheit der Matrizen $\mathbf{A}$ und $\mathbf{B}$ zunutze macht. Insbesondere bestätigen wir, daß die Wurzeln des charakteristischen Polynoms (16) (also die Eigenwerte des Produktes $\mathbf{AB}^{-1}$) mit den Eigenwerten von $\mathbf{A}$ übereinstimmen.

In der linearen Algebra wird gezeigt, daß eine reguläre Matrix $\mathbf{R}$ existiert, die für die positiv definite Matrix $\mathbf{B}^{-1}$ die Gleichung $\mathbf{B}^{-1} = \mathbf{R}\mathbf{R}^T$ erfüllt. Ferner wird gezeigt, daß für jede reguläre $(n-1, n-1)$-Matrix $\mathbf{X}$ der Ausdruck $\mathbf{X}^{-1} \, \mathbf{AB}^{-1} \, \mathbf{X}$ eine Ähnlichkeitstransformation des Produktes $\mathbf{AB}^{-1}$ darstellt, die dessen Eigenwerte unberührt läßt. (Beweise dieser Sätze finden sich in jedem Lehrbuch zur linearen Algebra.) Wir wählen nun $\mathbf{X} = (\mathbf{R}^T)^{-1}$. Dann ist

$$\begin{aligned}\mathbf{X}^{-1}\,\mathbf{A}\,\mathbf{B}^{-1}\mathbf{X} &= \mathbf{X}^{-1}\,\mathbf{A}\mathbf{R}\mathbf{R}^T\,\mathbf{X} \\ &= \mathbf{R}^T\,\mathbf{A}\mathbf{R}\mathbf{R}^T\,(\mathbf{R}^T)^{-1} \\ &= \mathbf{R}^T\,\mathbf{A}\mathbf{R}.\end{aligned}$$

Nun sei $\mathbf{z} \in I\!R^{n-1}$ ein beliebiger Vektor mit mindestens einer von Null verschiedenen Komponente. Ferner sei $\mathbf{r} := \mathbf{R}\mathbf{z}$. Man beachte, daß $\mathbf{r} \neq \mathbf{o}$ gilt, da $\mathbf{R}$ eine reguläre Matrix ist. Somit beinhaltet die negative Definitheit von $\mathbf{A}$:

$$\mathbf{z}^T\,\mathbf{R}^T\,\mathbf{A}\mathbf{R}\mathbf{z} = \mathbf{r}^T\,\mathbf{A}\mathbf{r} < 0.$$

Daher können wir schließen, daß $\mathbf{R}^T\,\mathbf{A}\mathbf{R} = \mathbf{X}^{-1}\,\mathbf{A}\mathbf{B}^{-1}\,\mathbf{X}$ ebenfalls eine negativ definite Matrix darstellt und demnach reelle, negative Eigenwerte besitzt. Mithin müssen auch alle Eigenwerte des Matrixproduktes $\mathbf{A}\mathbf{B}^{-1}$ reell und negativ sein. (Allerdings ist $\mathbf{A}\mathbf{B}^{-1}$ normalerweise nicht symmetrisch und folglich auch nicht negativ definit.) Damit haben wir Lemma 2 bewiesen. □

Literatur

1. *Arrow,* K. J.: Stability Independent of Adjustment Speed, in: Trade, Stability, and Macroeconomics (Horwich, G./Samuelson, P. A., Eds.), 181 - 202, Academic Press, New York, 1974.
2. *Bagchi,* A./*Olsder,* G. J./*Strijbos,* R. C. W.: Regional Allocation of Investment as a Hierarchical Optimization Problem, Regional Science and Urban Economics 11 (1981), 205 - 213.
3. *Blackorby,* C./*Schworm,* W.: The Existence of Input and Output Aggregates in Aggregate Production Functions, Econometrica 56 (1988), 613 - 643.
4. *DasGupta,* S.: A Local Analysis of Stability and Regularity of Stationary States in Discrete Symmetric Optimal Capital Accumulation Models, Journal of Economic Theory 36 (1985), 302 - 318.
5. *Datta-Chaudhuri,* M.: Optimum Allocation of Investments and Transportation in a Two-Region Economy, in: Essays on the Theory of Optimal Economic Growth (Shell, K., Ed.), 129 - 140, MIT Press, Massachusetts, 1967.
6. *Domazlicky,* B.: A Note on the Inclusion of Transportation in Models of the Regional Allocation of Investment, Journal of Regional Science 17 (1977), 235 - 241.
7. *Dorfman,* R.: Regional Allocation of Investment: Comment, Quarterly Journal of Economics 77 (1963), 162 - 165.
8. *Dorfman,* R./*Samuelson,* P. A./*Solow,* R. M.: Linear Programming and Economic Analysis, McGraw-Hill, New York, 1958.
9. *Färe,* R.: On Linear Expansion Paths and Homothetic Production Functions, in: Production Theory (Eichhorn, W./Henn, R./Opitz, R., Eds.), 53 - 64, Springer, Berlin, 1974.

10. *Färe,* R.: Fundamentals of Production Theory, Springer, Berlin, 1988.

11. *Friesz,* T. L./*Luque,* J.: Optimal Regional Growth Models: Multiple Objectives, Singular Controls, and Sufficiency Conditions, Journal of Regional Science 27 (1987), 201 - 224.

12. *Fujita,* M.: Optimum Growth in Two-Region, Two-Good Space Systems: The Final State Problem, Journal of Regional Science 13 (1973), 385 - 407.

13. *Fujita,* M.: Spatial Development Planning, A Dynamic Convex Programming Approach, North-Holland, Amsterdam, 1978.

14. *Helpman,* E./*Pines,* D.: Optimal Public Investment and Dispersion Policy in a System of Open Cities, American Economic Review 70 (1980), 507 - 514.

15. *Intriligator,* M. D.: Regional Allocation of Investment: Comment, Quarterly Journal of Economics 78 (1964), 659 - 662.

16. *Kamien,* M. I./*Schwartz,* N. L.: Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management, North-Holland, Amsterdam, 1981.

17. *Lau,* L. J.: A Characterization of the Normalized Restricted Profit Function, Journal of Economic Theory 12 (1976), 131 - 163. Reprinted as Essay VI in: The Hamiltonian Approach to Dynamic Economics (Cass, D./Shell, K., Eds.), 131 - 163, Academic Press, New York, 1976.

18. *McKenzie,* L. W.: Optimal Economic Growth, Turnpike Theorems and Comparative Dynamics, in: Handbook of Mathematical Economics (Arrow, K. J./Intriligator, M. D., Eds.), Vol. III, 1281 - 1355, North-Holland, Amsterdam, 1986.

19. *Michel,* P./*Pestieau,* P./*Thisse,* J.-F.: Regional Allocation of Investment with Distributive Objectives, Journal of Regional Science 23 (1983), 199 - 209.

20. *Ohtsuki,* Y.: Regional Allocation of Public Investment in an *n*-Region Economy, Journal of Regional Science 11 (1971), 225 - 233.

21. *Rahman,* A.: Regional Allocation of Investment, An Aggregative Study in the Theory of Development Planning, Quarterly Journal of Economics 77 (1963), 26 - 39.

22. *Rahman,* A.: Regional Allocation of Investment: The Continuous Version, Quarterly Journal of Economics 80 (1966), 159 - 160.

23. *Sakashita,* N.: Regional Allocation of Public Investment, Papers of the Regional Science Association 19 (1967), 161 - 182.

24. *Shephard,* R. W.: Theory of Cost and Production Functions, Princeton University Press, Princeton, 1970.

25. *Takayama,* A.: Regional Allocation of Investment: A Further Analysis, Quarterly Journal of Economics 81 (1967), 330 - 337.

26. *Takayama,* A. : Regional Allocation of Investment: Corrigendum, Quarterly Journal of Economics 82 (1968), 526 - 527.

27. *Wan,* H. Y. Jr.: Economic Growth, Harcourt Brace Jovanovich, New York, 1971.

Ein Erklärungsansatz für unterschiedliche Veränderungen in der Standortstruktur von Geschäften innerhalb verschiedener Städte bei gleichen Veränderungen in den Rahmenbedingungen

Von *Horst Behnke,* Hamburg

I. Einführung

Dieser Beitrag stellt einen theoretischen Erklärungsansatz für unterschiedlich starke Fluktuationen von Geschäften in verschiedenen Städten vor.

Die Erklärung für die Beobachtung unterschiedlicher Reaktionen der räumlichen Struktur des städtischen Marktes auf Veränderungen in den Marktbedingungen kann bei Abstraktion von anderen Aspekten[1] vollkommen durch die Theorie des reinen Wettbewerbs durch Standortwahl erfolgen[2].

Um auf eine idealtypische Struktur zu reduzieren, wird angenommen, daß sämtliche Städte monozentrisch sind und ein zentrumsorientiertes Verkehrsnetz haben. Da sich das Gewünschte an Hand einfachster Strukturen zeigen läßt, werden von diesem Verkehrsnetz auch nur die Haupteinfallstraßen betrachtet, die in gleichmäßiger räumlicher Verteilung vom Rand bzw. Umland einer jeden Stadt direkt in das Zentrum führen. Diesen Haupteinfallstraßen wird die Bevölkerung jeder Stadt wie folgt zugeordnet: Jeder Haushalt, bzw. die von ihm entfaltete Nachfrage, wird dem zu ihm nächstgelegenen Punkt einer Haupteinfallstraße zugeordnet. Für die weitere Argumentation wird sowohl die Dichte der Nachfrage entlang aller Haupteinfallstraßen aller Städte konstant[3] gesetzt, als auch die Länge aller dieser Haupteinfallstraßen für eine Stadt als gleich angesehen. Verschiedene Städte unterscheiden

---

[1] Wie z.B. unvollständige Informationen der Marktteilnehmer, lokal unterschiedliche Mieten oder Steuern.

[2] Der reine Wettbewerb durch Standortwahl wurde durch das Modell des Wettbewerbs zweier Anbieter auf einer Linie von *Hotelling,* H.: Stability in Competition, in: Economic Journal Vol. 39 (1929), S. 41 - 57, begründet.

[3] Natürlich sind auch andere Nachfragedichten denkbar. Eine Alternative wäre eine vom Stadtrand entlang der Haupteinfallstraßen steigende Nachfragedichte, um eine zum Zentrum hin zunehmende Bevölkerungsdichte zu betonen. Dabei ist allerdings zu bedenken, daß dieser Umstand auch schon durch eine konstante Nachfragedichte impliziert ist, da einem jedem Punkte einer jeden Haupteinfallstraßen eine um so größere Fläche zugeordnet ist, je weiter er vom Zentrum entfernt ist.

sich dann nur noch durch die unterschiedliche Länge und Anzahl ihrer Haupteinfallstraßen. Solche Straßennetze werden in den weiteren Abschnitten als s-Sterne bezeichnet, wobei $s > 3$ gelten muß.

Um die so lokalisierte Nachfrage der Haushalte besteht ein Wettbewerb zwischen den örtlichen Händlern. Zur Vereinfachung wird der Wettbewerb um den Absatz eines einzigen, mit Ausnahme seiner räumlichen Angebotsorte, homogenen Gutes bzw. Güterbündels betrachtet. Daraus ergibt sich für alle Händler die gleiche Kostensituation. Diese sei durch einen linearen Verlauf der variablen Kosten und zunächst durch Fixkosten in Höhe von Null gekennzeichnet. Die überall für eine Entfernungseinheit gleichen Transportkosten müssen die Nachfrager selber tragen. Die Händler besitzen keine Information über die räumliche Herkunft der Nachfrager, so daß ihnen allen nur die Möglichkeit der Preispolitik eines einheitlichen Preises ab Händler verbleibt.

Ferner seien die Transportkosten im wesentlichen nicht in Geld meßbar, sondern lediglich als eine Minderung des Nutzens der Nachfrager anzusehen, welche durch die Überwindung von Entfernungen verursacht wird.

Über jeden Nachfrager wird daher zusätzlich angenommen, daß seine Nachfragemenge nicht von diesen bei der Überwindung der Entfernung zu einem Händler anfallenden Nutzeneinbußen abhängt, sondern lediglich vom dort verlangten Preis ab Händler. Somit läßt sich einerseits in Unabhängigkeit von räumlichen Aspekten ein einheitlicher Marktpreis wie im Standardmodell der vollständigen Konkurrenz bestimmen und andererseits werden die Nachfrager stets den ihnen nächst gelegenen Händler zum Kauf des Gutes aufsuchen. Folglich entstehen zusammenhängende Marktgebiete entlang der Haupteinfallstraßen mit Marktgrenzen, die durch die gleiche Entfernung zu (mindestens[4]) zwei verschiedenen Standorten von Händlern gekennzeichnet sind. Jeder Händler strebt seinerseits danach, durch die Wahl seines Standortes einen möglichst großen Teil an Nachfrage zu erhalten. Da ferner angenommen wird, daß alle Händler ihren Standort unabhängig voneinander wählen müssen[5], ist der seitens der Händler entstehende Wettbewerb durch Standortwahl im Sinne der Spieltheorie als ein n Personen Konstantsummenspiel anzusehen. Die möglichen Standorte, hier alle Punkte des durch die

[4] Die Gleichgewichtsbedingungen im folgenden Abschnitt zeigen, daß in gleichgewichtigen Konstellationen von Händlern das Zentrum stets besetzt sein muß, so daß Marktgrenzen immer genau die Marktgebiete zweier Händler trennen.

[5] Die Möglichkeit zur Bildung von Kooperationen zwischen Händlern ist natürlich denkbar, soll hier aber nicht weiter verfolgt werden. Erfolgen diese in Form von Gebietskartellen, so handelt es sich um das gleiche Modell wie beim Wettbewerb durch Standortwahl für Filialketten. Diese wird in *Behnke,* H.: Spatial Competition among Firms with more than one shop, (Bisher unveröffentlichtes Manuskript 1992), für den linearen Markt vollständig abgehandelt und für Straßennetze finden sich dort wesentliche Zwischenergebnisse.

Haupteinfallstraßen und das Zentrum gebildeten Sternes, bilden die Menge der reinen Strategien eines jeden Händlers.

Zunächst stellt sich nun die Frage, ob in diesem statischen Modell Konstellationen von Standorten existieren, die Nash-Gleichgewichte[6] in reinen Strategien darstellen[7]. Eine solche Konstellation liegt dann vor, wenn keiner der Händler eine Veranlassung sieht seinen Standort zu ändern; d.h. er sein Marktgebiet demnach nicht ohne Hilfe anderer Händler vergrößern kann[8].

In den folgenden Abschnitten werden zuerst Bedingungen für solche Gleichgewichte bei gegebener Anzahl von Anbietern vorgestellt und allgemeine Existenzaussagen getroffen. Anschließend werden positive Fixkosten in das Modell eingeführt, womit dann die Modellierung eines freien Marktzu- und -abganges möglich wird und die Zahl der Anbieter im Gleichgewicht endogen vom Modell bestimmt wird. Darauf folgt ein Beispiel zur unterschiedlichen Sensitivität gegebener Gleichgewichte in verschiedenen Städten bei einer Veränderung der für alle Anbieter in allen Städten gleichen Fixkosten in allen Städten. Den Abschluß bildet ein kurzer Ausblick auf mögliche Erweiterungen des Modells und den damit verbundenen weiteren Erklärungswert durch Modelle des reinen Wettbewerbs durch Standortwahl.

II. Gleichgewichtsbedingungen und Existenzaussagen bei gegebener Anzahl von Händlern

In diesem Abschnitt werden zuerst die notwendigen und hinreichenden Bedingungen für Nash-Gleichgewichte in reinen Strategien der Standortwahl

---

[6] Das Nash-Gleichgewicht wurde von *Nash,* J. F.: Non-Cooperative Games, in: Annals of Mathematics Vol. 54 (1951), S. 286 - 295, als spieltheoretische Verallgemeinerung der Gleichgewichtslösung im Cournotschen Oligopol entwickelt.

[7] Da jeder Punkt eines *s*-Sternes als Standort für jeden Händler in Frage kommt, sind die Mengen ihrer reinen Strategien gleich und haben unendlich viele Elemente. Für solche *n* Personen Konstantsummenspiele gibt es keine allgemeingültigen Aussagen über Nash-Gleichgewichte, weder über deren Existenz in reinen oder gemischten Strategien, noch über die Anzahl solcher Gleichgewichte.

[8] Dabei sieht jeder Händler die Standorte der anderen Händler als gegeben an. Diese Verhaltensweise wird in der einschlägigen Literatur als ‚ZCV-Strategy (Zero Conjectural Variation)' bezeichnet. Andere Möglichkeiten des Verhaltens der Händler führen zu anderen Aussagen über die Existenz von Nash-Gleichgewichten und selbstverständlich auch zu anderen Mengen solcher Gleichgewichte. Beispiele sind *Eaton,* B. C./*Lipsey,* R. G.: The Principle of Minimum Differentiation Reconsidered, in: Review of Economic Studies Vol. 42 (1975), S. 27 - 39, die einen Vergleich mit der Minimax-Strategie durchführen und *Hannesson,* R.: Defensive Foresight rather than Minimax, in: Review of Economic Studies Vol. 49 (1982), S. 653 - 657, wo die Händler nur dann die Verlegung ihres Standortes durchführen, wenn auch nach evtl. auftretenden Folgereaktionen von anderen Händlern auf diesen Standortwechsel noch eine Verbesserung ihres Marktanteils verbleibt. Es stellt sich hier allerdings die Frage, warum nicht auch weitergehende Reaktionsketten berücksichtigt werden.

auf s-Sternen angegeben. Anschließend wird diskutiert, bei welchen Anzahlen von Anbietern und Haupteinfallstraßen überhaupt Gleichgewichte existieren.

Satz II.1. (Bedingungen für das Vorliegen eines Nash-Gleichgewichtes mit n Händlern auf s-Sternen):

Für eine gegebene Anzahl n von Anbietern auf jedem s-Stern sind die folgenden Bedingungen notwendig und hinreichend für die Existenz eines Nash-Gleichgewichtes in reinen Strategien[9]. (Dabei wird mit Teilmarktgebiet eines Standortes jeder gesamte Teil seines Marktgebietes bezeichnet, der zwischen ihm und einer seiner Marktgrenzen liegt[10].)

G1: Das Zentrum jedes s-Sterns ist mindestens von einem und höchstens von s Händlern besetzt.

G2: Alle Standorte auf den Haupteinfallstraßen sind höchstens von zwei Händlern besetzt.

G3: Auf allen Haupteinfallstraßen ist die Randposition zur Stadtgrenze entweder doppelt oder die Haupteinfallstraße ist gar nicht von Händlern besetzt.

G4: Alle Teilmarktgebiete aller doppelt besetzten Standorte und jedes mit s Händlern besetzten Zentrums haben die gleiche Größe a.

G5: Händler auf einfachbesetzten Standorten besitzen ein Marktgebiet zwischen a und $2a$, und jedes Zentrum hat, falls es mit $k < s$ Händlern besetzt ist, ein Marktgebiet zwischen ka und $(k + 1)a$. (Mit a aus G4.)

G6: Der Abstand zwischen zwei beliebigen, benachbarten Standorten beträgt höchstens $2a$. (Mit a aus G4.)

Auf Grundlage dieser notwendigen und hinreichenden Bedingungen ergibt sich folgender Satz über die Existenz und die Anzahl von Nash-Gleichgewichten in Abhängigkeit von der vorgegebenen Anzahl n von Händlern[11]:

Satz II.2. (Satz über die Existenz von Nash-Gleichgewichten auf s-Sternen mit gegebener Zahl von Händlern):

---

[9] Diese Gleichgewichtsbedingungen behalten ihre Gültigkeit auch, wenn die Haupteinfallstraßen unterschiedliche Längen haben.

[10] Auf den Beweis für die Notwendigkeit jeder einzelnen Bedingung und die Tatsache, daß alle Bedingungen zusammen hinreichend für ein Nash-Gleichgewicht in reinen Strategien sind, wird hier verzichtet, da sich in Beweistechnik und -idee keine Unterschiede zum Beweis der entsprechenden Bedingungen für den Wettbewerb durch Standortwahl auf einer Linie mit konstanter Dichte der Nachfrage ergeben. Dieser Beweis befindet sich in: *Selten,* R.: Anwendungen der Spieltheorie auf die politische Wissenschaft, in: *Maier,* H./*Ritter,* K./*Matz,* U. (Hrsg.): Politik und Wissenschaft, München 1971, S. 287 - 320; hier: S. 297.

[11] Der Beweis und eine Herleitung an Hand von Beispielen befindet sich in *Behnke,* H.: Gleichgewicht im Wettbewerb durch Standortwahl auf Straßennetzen, Dissertation Hamburg 1990, S. 70 - 81.

Für jeden s-Stern mit einer beliebigen festen Zahl $n > 1$ von Händlern gilt:

A) Für $n \leq s$ existieren eindeutige,

B) für $s < n < 3s-1$ existieren keine

C) und für $3s - 1 \leq n$ existieren (genau für die Fälle $n = 3s - 1$ und $n = 3s$ eindeutige) Nash-Gleichgewichte in reinen Strategien.

Hierzu ist anzumerken, daß sich in jedem eindeutigen Gleichgewicht in A) stets alle Händler im Zentrum befinden. Für die beiden eindeutigen Fälle in C) gilt Folgendes: Es sind auf jeder Haupteinfallstraße stets zwei Händler positioniert, die sich zusammen auf dem Punkt in 2/3 der Entfernung von Zentrum zum Stadtrand befinden. Die übrigen $s - 1$ bzw. s Händler haben ihren Standort im Zentrum.

Im nächsten Abschnitt dient dieser Satz als Grundlage für den Existenzsatz III.2. im um freien Marktzugang und Mindestmarktanteile erweiterten Modell.

III. Gleichgewichtsbedingungen und Existenzaussagen bei Vorliegen von Fixkosten und permanent möglichem freien Marktzugang

Als Erweiterung des Modells werden nun für jeden Händler in gleicher Höhe entstehende Fixkosten zum Unterhalt seines Geschäftes angenommen. Um diese decken zu können, benötigt dann jeder Händler eine bestimmte Mindestmenge an Nachfrage, die sich wegen der konstanten Nachfragedichte in einem Mindestmarktanteil a_{fix} ausdrückt. Wird zusätzlich angenommen, daß sich einerseits jederzeit neue Händler in jeder Stadt niederlassen können, sofern sie einen Standort finden auf dem sie mehr als a_{fix} erhalten können[12], und andererseits jeder Händler, der an seinem Standort nicht a_{fix} erhält, sein Geschäft aufgeben muß, so ändern sich die Bedingungen G1-G6 aus dem vorigen Abschnitt dergestalt, daß in G4, G5 und G6 a_{fix} anstelle des endogen durch eine betrachtete Konstellation erzeugten a eingesetzt werden muß. Die dann entstandenen Bedingungen lauten:

---

[12] Um nicht in eine Diskussion um infinitesimal kleine Marktanteile zu geraten, wird hier angenommen, daß potentielle Neulinge am Markt immer die Chance auf einen Gewinn sehen müssen, während am Markt befindliche Anbieter stets solange auf diesem beharren, bis sie Verluste machen. Dieses ist durchaus angemessen, da in der Bestimmung von a_{fix} natürlich auch sämtliche Opportunitätskosten für alternative Kapitalanlagen enthalten sind. Ein Händler wird dann nur umdisponieren, wenn er sich dadurch auch wirklich verbessert.

Da diesem Modell keine Dynamik innewohnt, kann eine Gleichgewichtskonstellation dann evtl. durch Herausnahme von Händlern, die lediglich a_{fix} erwirtschaften, wieder in ein neues Gleichgewicht überführt werden, wenn sich dadurch keine Gewinnchancen für potentielle neue Händler ergeben.

Satz III.1. (Bedingungen für das Vorliegen eines Nash-Gleichgewichtes bei freiem Marktzugang auf s-Sternen):

Für ein gegebenen Mindestanteil an Nachfrage von a_{fix} beim Wettbewerb durch Standortwahl und freien Marktzugang auf einem s-Stern[13] sind die folgenden Bedingungen notwendig und hinreichend für die Existenz eines Nash-Gleichgewichtes in reinen Strategien[14].

F1: Das Zentrum jedes s-Sterns ist mindestens von einem und höchstens von s Händlern besetzt.

F2: Alle Standorte auf den Haupteinfallstraßen sind höchstens von zwei Händlern besetzt.

F3: Auf jeder Haupteinfallstraße sind die Randpositionen zur Stadtgrenze entweder doppelt oder gar nicht besetzt.

F4: Alle Teilmarktgebiete aller doppelt besetzten Standorte und jedes mit s Händlern besetzte Zentrum haben die gleiche Größe a_{fix}.

F5: Händler auf einfachbesetzten Standorten besitzen ein Marktgebiet von mindestens a_{fix} und höchstens $2a_{fix}$, und jedes Zentrum hat, falls es mit $k < s$ Händlern besetzt ist, ein Marktgebiet von mindestens ak und höchstens a_{fix} $(k + 1)$.

F6: Der Abstand zwischen zwei beliebigen, benachbarten Standorten beträgt höchstens $2a_{fix}$.

Jedes im Modell erzeugte Nash-Gleichgewicht in reinen Strategien bestimmt nun nicht mehr den Mindestanteil an Nachfrage bei gegebener Zahl von Händlern, sondern die Zahl der Händler bei durch Fixkosten induziertem a_{fix}.

Für den Existenzsatz ergeben sich durch diese Erweiterung des Modells folgende Konsequenzen:

– Eine Konstellation von Händlern, die im Grundmodell des vorherigen Abschnittes ein Nash-Gleichgewicht in reinen Strategien war, kann in der Erweiterung um freien Marktzugang und Mindestanteil an Nachfrage nur dann noch ein Gleichgewicht sein, wenn das induzierte a auch gleich a_{fix} ist.

  Ist a_{fix} größer, so gibt es immer eine Möglichkeit für neue Anbieter auf dem Markt zu bestehen[15]. Ist dagegen das induzierte a größer, müßten solange Anbieter, welche weniger Nachfrage als a_{fix} besitzen, die Stadt

[13] Diese Gleichgewichtsbedingungen behalten ihre Gültigkeit auch hier, wenn die Haupteinfallstraßen unterschiedlich lang sind.

[14] Auch hier erübrigt sich die Angabe des exakt formulierten Beweises, da durch die weiteren Ausführungen zu diesem Satz seine Richtigkeit als Folge von Satz II.1. unmittelbar einsichtig ist.

[15] Besteht das Gleichgewicht in einer Konstellation mit allen Händlern im Zentrum, so ist, aufgrund von F1, auf jeder Haupteinfallstraße eine größere Nachfrage als a_{fix}.

verlassen, bis alle verbleibenden Händler mindestens ihre Fixkosten decken können.

– Da für jede Konstellation von Händlern, die $a = a_{fix}$ erfüllt, die Bedingungen F1-F6 mit G1-G4 aus dem vorherigen Abschnitt zusammenfallen, können im erweiterten Modell keine anderen gleichgewichtigen Konstellationen von Händlern vorliegen als jene durch Satz II.2. charakterisierten.

Folglich schränkt die in diesem Abschnitt vorgenommene Erweiterung des Modellrahmens die Menge der möglichen Nash-Gleichgewichte ein. Der Existenzsatz II.2. bildet nun die Grundlage für den nachstehenden Existenzsatz des erweiterten Modells.

Satz III.2. (Satz über die Existenz von Nash-Gleichgewichten auf s-Sternen mit freiem Marktzugang):

Für jeden s-Stern mit einer beliebigen festen $a_{fix} > 0$ gilt (l_s bezeichne dabei die Länge der Haupteinfallstraßen)[16]:

1. Für $a_{fix} > l_s s$ existiert ein Gleichgewicht ohne Händler.
2. Für $l_s \leq a_{fix} \leq l_s s$ existiert genau ein Gleichgewicht mit int $[l_s s/a_{fix}]$ Anbietern.
3. Für $l_s/3 < a_{fix} < l_s$ existiert kein Gleichgewicht.
4. Für $a_{fix} = l_s/3$ existieren genau zwei Gleichgewichte mit $n = 3s-1$ oder $n = 3s$ Händlern.
5. Für $l_s/(3 + 2/s) < a_{fix} < l_s/3$ existieren keine Gleichgewichte.
6. Für $l_s/5 < a_{fix} \leq l_s/(3 + 2/s)$ existiert ein Gleichgewicht mit $n = 3s + \text{int}\,[s\,(l_s-3a_{fix})/2a_{fix}]$ Händlern.
   Ist int $[s\,(l_s-3a_{fix})/2a_{fix}] = [s\,(l_s-3a_{fix})/2a_{fix}]$, so existiert zusätzlich stets ein Gleichgewicht mit $s\,(l_s-3a_{fix})/2a_{fix}]-1$ Händlern.
7. Für $l_s/5 = a_{fix}$ existieren mehrere mögliche Gleichgewichte mit Anbieterzahlen zwischen $n = 4s-1$ und $n = 5s$ Händlern.
8. Für $a_{fix} < l_s/5$ existieren stets mehrere Gleichgewichtskonstellationen. Die mögliche Zahl der Händler ist jede ganze Zahl

$$n \in [s\,(\text{int}[l_s/2a_{fix}] + 2) + \max\{1,\ \text{int}[s\,(l_s/a_{fix} - \text{int}[l_s/a_{fix}])/2a_{fix}]\};$$

$$s\,(\text{int}[l_s/a_{fix}]) + \text{int}[s\,(l_s/a_{fix} - \text{int}[l_s/a_{fix}])/2a_{fix}]]$$

Ein neuer Anbieter kann sich dann direkt neben das Zentrum auf eine der Haupteinfallstraßen begeben. Liegt ein Gleichgewicht mit Händlern auf den Haupteinfallstraßen vor, so besteht stets folgende Möglichkeit zum Markteintritt: Aufgrund von F3 sind beide Teilmarktgebiete jedes zum Stadtrand nächsten Standortes größer als a_{fix}, so daß sich ein neuer Händler auf einer der beiden Seiten direkt neben diesen Standorten niederlassen wird.

[16] Der Beweis von Satz III.2. befindet sich im Anhang.

Durch die vielen in Satz III.2. gekennzeichneten Bereiche wird klar, daß bei unterschiedlichen s-Sternen[17] für die alle das gleiche a_{fix} gilt, auch unterschiedliche Situationen bzgl. der Existenz und Art von Nash-Gleichgewichten fast zwangsläufig auftreten müssen. Auch die Sensitivität bestehender Gleich- oder Ungleichgewichte auf Änderungen von a_{fix} kann dann sehr unterschiedlich sein. Der nächste Abschnitt verdeutlicht diesen Sachverhalt an einem Beispiel mit drei Städten.

IV. Ein exemplarisches Beispiel für unterschiedliche Reaktionen der Standortstruktur auf eine Veränderung der Fixkosten

Im folgenden Beispiel mit drei Städten, die jeweils auf einen s-Stern mit $s = 3$ reduziert sind, liegt bei einem gegebenem a_{fix} für alle ein Nash-Gleichgewicht in reinen Strategien vor. Bei einer, nur aus Gründen der Überschaubarkeit, deutlichen Erhöhung von a_{fix} auf den doppelten Wert entstehen in zwei Städten neue Gleichgewichte mit nun verringerter Zahl von Händlern, wobei in der ersten Stadt deren Gewinne gleich bleiben und in der zweiten Stadt steigen. In der dritten Stadt entsteht dagegen eine Situation des Ungleichgewichtes.

Die Ausgangssituation in den drei Städten vor Verdopplung von a_{fix} sei gegeben durch:

Stadt 1 befindet sich in einem gewinnlosen Nash-Gleichgewicht:

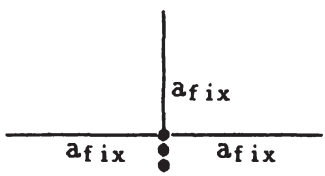


Abb. 1a: Stadt 1 vor Erhöhung der Fixkosten.

Auch Stadt 2 befindet sich in einem Nash-Gleichgewicht ohne Gewinne für die Händler:

[17] Die verschiedenen Sterne, und die durch sie repräsentierten Städte, können sich sowohl durch eine unterschiedliche Anzahl von Haupteinfallstraßen (s), als auch durch deren unterschiedliche Länge (l_s) ergeben.

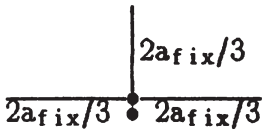


Abb. 2a: Stadt 2 vor Erhöhung der Fixkosten.

Stadt 3, mit $b = a_{fix}/3$, befindet sich in einem Nash-Gleichgewicht, in dem die Händler teilweise Gewinne erzielen und teilweise nicht.

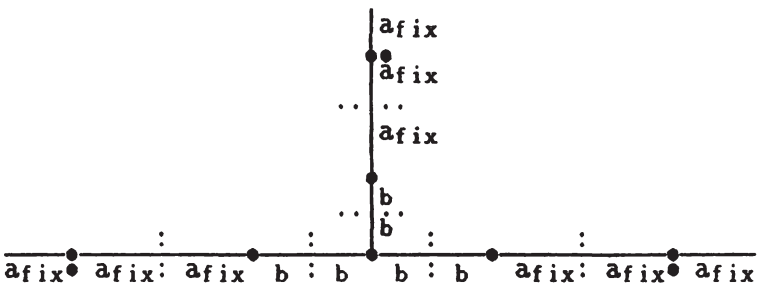


Abb. 3a: Stadt 3 vor Erhöhung der Fixkosten.

Nach einer Erhöhung der Fixkosten, die sich in einer Veränderung auf $a^*_{fix} = 2a_{fix}$ niederschlägt, verändert sich die Struktur der Händlerstandorte in den einzelnen Städten wie folgt:

In Stadt 1 befindet sich nur noch ein Händler, der nun einen Gewinn in Höhe der neuen Fixkosten macht:

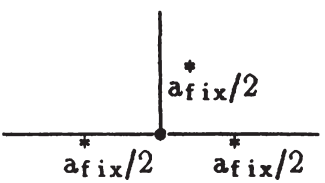


Abb. 1b: Stadt 1 nach Erhöhung der Fixkosten.

Auch Stadt 2 befindet sich weiterhin in einem Nash-Gleichgewicht mit nur noch einem Händler. Dieser erzielt aber, wie vor der Erhöhung der Fixkosten, keinen Gewinn.

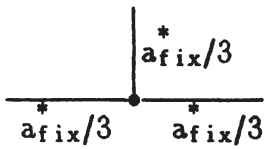


Abb. 2b: Stadt 2 nach Erhöhung der Fixkosten.

In Stadt 3 herrscht überhaupt kein Gleichgewicht mehr. Von der in ihr vorhandenen Nachfrage her, würden zwar bis zu fünf Händler ihre Fixkosten decken können und sogar einen Gewinn erzielen, aber die Struktur ist so ungünstig, daß durch Umzüge ständig Verbesserungen für einen oder mehrere Anbieter möglich sind.

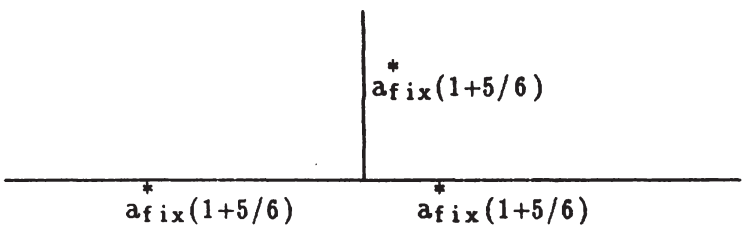


Abb. 3b: Situation in Stadt 3 nach Erhöhung der Fixkosten.

V. Abschließende Bemerkungen

Die vorangegangenen Abschnitte zeigten, daß schon bei einer Reduktion von städtischen Strukturen zu s-Sternen unterschiedliche Anzahlen, Arten und Sensitivitäten von Nash-Gleichgewichten in reinen Strategien vorliegen können. Wird bedacht, daß Städte selbst bei grober Betrachtung nicht immer ein im wesentlichen zentrumsorientiertes Verkehrsnetz haben müssen[18], und auch geographische Gegebenheiten eine Regelmäßigkeit der städtischen Struktur aufbrechen können, so wird klar, daß bei Beibehaltung der konstanten Nachfragedichte[19] und genügend kleinem a_{fix}, Gleichgewichte beim Ver-

[18] Denkbar sind vor allem Straßennetze mit Subzentren, Ringstraßen oder auch solche in der sogenannten Manhattan-Norm.

[19] Für beliebige Straßennetze mit konstanter Nachfragedichte kann auch allgemein gezeigt werden, daß bei genügend großen Anzahlen von Anbietern (bzw. genügend kleinem a_{fix}), stets Nash-Gleichgewichte in reinen Strategien existieren. Wird auf andere kontinuierliche Nachfragedichten mit Straßenabschnitten auf denen eine strikt monotone Nachfragedichte herrscht übergegangen, so kann aber gezeigt werden, daß es im Gegensatz hierzu Anzahlen von Anbietern (bzw. Werte für a_{fix}) gibt, für die

gleich von Städten gleicher Größenordnung zwar sehr wahrscheinlich sind, aber deren Sensitivität auf Veränderungen der Rahmenbedingungen bzgl. Änderungen der räumlichen Struktur der Standorte der Händler aber höchst unterschiedlich ausfallen wird. Werden die Überlegungen zum Wettbewerb durch Standortwahl auch auf den Fall verschiedener Güter mit unterschiedlichen Mindestanteilen an deren Nachfrage und unterschiedlichen räumlichen Nachfragepotentialen ausgedehnt, so wird auch die Möglichkeit sichtbar die Christallersche Theorie der zentralen Orte[20] durch den reinen Wettbewerb durch Standortwahl von der Angebotsseite her zu begründen[21].

Mit dieser kurzen Schlußbemerkung, die theoretisch interessante Möglichkeiten zu Erweiterungen anriß, endet dieser Beitrag.

Anhang

Beweis und Erläuterungen zu Satz III.2.:

Der Beweis erfolgt mit Hilfe der notwendigen und hinreichenden Bedingungen für Nash-Gleichgewichte in reinen Strategien aus Satz III.1. und durch Verwendung des Existenzsatzes II.2.

Ferner ist unmittelbar einsichtig, daß sich bei sinkendem a_{fix} die mögliche(n) Zahl(en) der Anbieter im Gleichgewicht nicht verringern können, so daß beginnend mit 1. schrittweise bis 8. vorgegangen werden kann.

zu 1.:

Wenn $a_{\text{fix}} > l_s s$ (, welches das gesamte Marktgebiet des s-Sterns ist), dann erhält nicht einmal ein Händler seine Fixkosten, so daß sich dieser städtische Markt zwar in einem Gleichgewicht befindet, aber dieses ohne Händler ist.

zu 2.:

Ist $l_s < a_{\text{fix}} \leq l_s s$, so ist a_{fix} größer als (bzw. gleich mit) jede(r) einzelne(n) Haupteinfallstraße. Aufgrund von F1 aus Satz III.1. ist mindestens

---

und alle höheren Anzahlen (niedrigeren Werte) keine Gleichgewichte existieren können; s. hierzu: *Behnke,* H.: Gleichgewicht im Wettbewerb durch Standortwahl auf Straßennetzen, Dissertation Hamburg 1990; hier: S. 82 - 91.

[20] Die Theorie von *Christaller,* W.: Die zentralen Orte in Süddeutschland, Jena 1933, begründet eine räumliche Hierarchie von Städten von der Nachfrageseite her. Durch maximale Reichweiten von Gütern zu Haushalten wird eine vom Gut mit der niedrigsten dieser Reichweiten bestimmte, von ihrer Kompaktheit optimale Struktur bestimmt, die dann aus geometrischen Überlegungen heraus in ein Sechseckmuster für jede Hierarchiestufe mündet.

[21] Zumindest kann die Christallersche Hierarchie der Städte mitsamt ihrer räumlichen Anordnung als ein mögliches Nash-Gleichgewicht in reinen Strategien konstruiert werden.

ein Händler im Zentrum positioniert, so daß sowohl keine Händler auf den Haupteinfallstraßen sind als auch neue Händler nur im Zentrum erfolgreich sein können. Hier finden genau soviele Händler Platz, wie der gesamte städtische Markt verkraften kann. Das sind $n = \text{int}\,[l_s s / a_{\text{fix}}]$. Die Zahl der Händler liegt folglich zwischen 1 und s; analog zu Teil A aus Satz II.2. und wie von Bedingungen F1 und F3 aus Satz III.1. gefordert.

zu 3.:

Gilt $a_{\text{fix}} < l_s$, so ist jede Haupteinfallstraße in der Lage mindestens einem Händler die Deckung seiner Fixkosten zu ermöglichen. Da F4 fordert, daß im Gleichgewicht kein Teilmarktgebiet größer als a_{fix} ist, müssen sich nun im Gleichgewicht auch Händler auf den Haupteinfallstraßen befinden. Wegen F3 muß jede der s Haupteinfallstraßen aber mindestens zwei Händlern das Überleben auf dem Markt sichern können, da sich schließlich aufgrund von F1 immer noch mindestens ein Anbieter im Zentrum befinden muß.

F1, F3 und F4 sind demnach solange nicht zu erfüllen bis nicht $a_{\text{fix}} \leq l_s/3$ ist.

zu 4.:

Für $a_{\text{fix}} = l_s/3$ ergibt sich genau der Teil C des Satzes II.2., so daß sich jeweils genau zwei Händler zusammen auf dem Punkt in Entfernung $2l_s/3$ vom Zentrum auf jeder der s Haupteinfallstraßen befinden. Da dann für das Zentrum $a_{\text{fix}} s$ verbleibt, können sich dort s oder $s-1$ Händler befinden ohne F1 oder F5 zu verletzen. Sind s Händler im Zentrum, so erhalten alle Händler nur a_{fix}.

zu 5.:

Ist $a_{\text{fix}} < l_s/3$, so können, immer unter Beachtung von F1, auf jeder Haupteinfallstraße in der zu 4. beschriebenen Konstellation neue Händler auf jeder der s Haupteinfallstraßen eindringen.

F1 bis F6 können dann nur zusammen erfüllt werden, wenn erstens die Randpositionen weiterhin doppelt besetzt bleiben und sich genau in der Mitte eines $2a_{\text{fix}}$ großen Marktgebietes befinden, zweitens ein dritter Anbieter auf jeder Haupteinfallstraße ebenfalls mindestens a_{fix} erhält und drittens mindestens ein Händler im Zentrum zu a_{fix} gelangt.

Das bedeutet für l_s in bezug auf a_{fix}, daß $l_s \geq 3a_{\text{fix}} + 2a_{\text{fix}}/s$ werden muß. Nach a_{fix} umgeformt ergibt sich, daß für weitere Gleichgewichte $a_{\text{fix}} \leq l_s (3 + 2/s)$ werden muß.

zu 6.:

Ist nun $a_{\text{fix}} \leq l_s/(3 + 2/s)$, so sind F1 - F6 stets erfüllt, wenn auf jeder der s Haupteinfallstraßen drei Händler folgendermaßen positioniert sind: Zwei befinden sich zusammen in Entfernung a_{fix} vom Stadtrand und der Dritte in

Entfernung $2a_{fix}$ von diesen beiden. Bis zum Zentrum verbleibt dann ein Rest von $b = l_s - 3a_{fix}$. Solange $b < 2a_{fix}$ ist, können keine weiteren Händler auf den Haupteinfallstraßen Platz finden. (D.h. solange $a_{fix} > l_s/5$ ist.) Zu den so auf jeden Fall vorhandenen $3s$ Händlern, kommen im Zentrum noch soviele hinzu wie dort ihre Fixkosten decken können. Dies sind $\text{int}[sb/2a_{fix}]$ weitere Händler.

Ist $sb/2a_{fix}$ ein ganzes Vielfaches von a_{fix}, so kann sich im Zentrum, wie in 4., auch ein Händler weniger befinden, ohne F1 - F6 zu verletzen.

zu 7.:

Ist nun $a_{fix} = l_s/5$, so bleibt die Konstellation von 6. die einzig mögliche Gleichgewichtskonstellation, allerdings kann die mit nur einem Händler besetzte Position auf jeder Haupteinfallstraße nun auch mit einem zweiten Händler besetzt sein. Auf allen Haupteinfallstraßen können sich daher zwischen $3s$ und $4s$ Händler befinden. Zu diesen kommen dann $s-1$ oder s im Zentrum hinzu. Im letzten Fall erhalten dann alle Händler gerade a_{fix}.

zu 8.:

Wenn $a_{fix} < l_s/5$, so ist es, ausgehend von den Gleichgewichtskonstellationen unter 7., stets möglich beliebig viele Gleichgewichte zu konstruieren. Um den Bereich der möglichen Zahl der Händler im Gleichgewicht zu ermitteln, ist auch hier folgendes zu beachten:

Zunächst müssen natürlich, aufgrund von F3, die Randpositionen der Haupteinfallstraßen doppelt besetzt sein und sich auf dem Mittelpunkt eines $2a_{fix}$ großen Gebietes befinden. Gleichzeitig muß, wegen F1, im Zentrum mindestens ein Händler angesiedelt sein.

Die möglichen Anzahlen von Händlern im Gleichgewicht bei $a_{fix} < l_s/5$ lassen sich nun wie folgt eingrenzen:

Minimale Zahl der Händler im Gleichgewicht:

Auf den s Haupteinfallstraßen wird die minimale Zahl von Händlern erreicht, wenn zwischen der Randposition und der dem Zentrum benachbarten Position nur $2a_{fix}$ große Marktgebiete mit jeweils einem Händler auf ihrem Mittelpunkt existieren. Das Marktgebiet der ebenfalls einfach besetzten Nachbarposition des Zentrums ist dann entweder kleiner oder genauso groß.

Auf diese Weise ergibt sich für die Zahl der Händler auf den Haupteinfallstraßen:

a) Wenn $\text{int}[l_s/a_{fix}]$ gerade ist, sind es $s\,(\text{int}[l_s/2a_{fix}]) + 1$ Händler, und

b) wenn $\text{int}[l_s/a_{fix}]$ ungerade ist, sind es $s\,(\text{int}[l_s/2a_{fix}]) + 2$ Händler bzw. $s\,(\text{int}[l_s/2a_{fix}]) + 1$ Händler, falls $l_s/a_{fix} = \text{int}[l_s/a_{fix}]$.

Im Zentrum ergibt sich:

a) Wenn $\text{int}[l_s/a_{\text{fix}}]$ gerade ist, sind es $\text{int}[s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}} - 1])/2a_{\text{fix}}]$ Händler.

b) wenn $\text{int}[l_s/a_{\text{fix}}]$ ungerade ist, sind es $\max\{1, \text{int}[s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}}])/2a_{\text{fix}}]\}$ Händler.

(Ist $\text{int}[s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}}])/2a_{\text{fix}}] = 0$, so muß folgendermaßen umgestellt werden, damit F1 erfüllt ist:

Auf genau einer Haupteinfallstraße rückt die Nachbarposition des Zentrums vom Zentrum weg bis zu dem Punkt, der $2a_{\text{fix}}$ vom Zentrum entfernt ist. Nun hat diese Position zu ihrer anderen Nachbarposition den gleichen Abstand, wie vorher zum Zentrum. Genau ein weiterer Händler kann dann seinen Standort im Zentrum finden.)

Zusammengefaßt ergibt sich als minimale Zahl der Händler:

$$s\ (\text{int}[l_s/2a_{\text{fix}}] + 2) + \max\{1, \text{int}[s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}}])/2a_{\text{fix}}]\}$$

Maximale Zahl der Händler im Gleichgewicht:

Die größtmögliche Zahl von Händlern wird erreicht, wenn die für die Konstruktion der minimalen Anzahl von Händlern angegebenen Positionen zwischen Randposition und Nachbarposition des Zentrums alle doppelt besetzt sind. Die Nachbarposition des Zentrums bleibt einfachbesetzt. (Es sei denn, l_s ist ein ungerades Vielfaches von a_{fix}, dann wird auch diese Position doppelt besetzt.)

Auf diese Weise ergibt sich für die s Haupteinfallstraßen:

a) Wenn $\text{int}[l_s/a_{\text{fix}}]$ gerade ist, sind es $s\ (\text{int}[l_s/a_{\text{fix}}]) - 1$ Händler.

b) Wenn $\text{int}[l_s/a_{\text{fix}}]$ ungerade ist, sind es $s\ (\text{int}[l_s/a_{\text{fix}}])$ Händler.

(Ist außerdem $l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}}] = 0$, so werden hier schon die s im Zentrum befindlichen Händler mitgezählt.)

Im Zentrum werden dann soviele Händler wie möglich aufgestellt.

a) Wenn $\text{int}[l_s/a_{\text{fix}}]$ gerade ist, sind es $\text{int}[s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}} - 1])/2a_{\text{fix}}]$ Händler, und

b) wenn $\text{int}[l_s/a_{\text{fix}}]$ ungerade ist, sind es $\text{int}[s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}}])/2a_{\text{fix}}]$ Händler.

(Ist dieser Ausdruck Null, aber $s\ (l_s/a_{\text{fix}} - \text{int}[l_s/a_{\text{fix}}]$ positiv, so wird, um F1 zu erfüllen, mit gleichem Ergebnis folgendermaßen umgestellt:

Um ein genügend großes Marktgebiet für einen Händler im Zentrum zu schaffen, wird auf genau einer Haupteinfallstraße die doppelt besetzte Position neben der Nachbarposition des Zentrums in eine einfach besetzte Position umgewandelt. Anschließend entfernt sich die Nachbarposition des

Zentrums um $2a_{fix}$ vom Zentrum. Der entfernte Händler bekommt seinen neuen Standort im Zentrum.

Ist schon $s\ (l_s/a_{fix} - \text{int}[l_s/a_{fix}]) = 0$, so wurde die maximale Anzahl s der im Zentrum bereits bei den Haupteinfallstraßen mitgezählt.)

Durch Zusammenfassen der Fälle ergibt sich dann als maximale Zahl der Händler:

$s\ (\text{int}[l_s/a_{fix}]) + \text{int}[s\ (l_s/a_{fix} - \text{int}[l_s/a_{fix}])/2a_{fix}]$

Ist l_s ein ungerades Vielfaches von a_{fix}, so ist ein Gleichgewicht möglich indem alle Händler nur a_{fix} erhalten.

Literatur

Behnke, H. (1990): Gleichgewicht im Wettbewerb durch Standortwahl auf Straßennetzen, Dissertation an der Universität Hamburg.

Behnke, H. (1992): Spatial Competition among Firms with more than one shop befindet sich in: Karmann, A./Moster, K./Schader, M./Uebe, U: (Hrsg.) Operations Research '92, Heidelberg 1993, S. 399 - 401.

Christaller, W. (1933): Die zentralen Orte in Süddeutschland, Jena.

Eaton, B. C./*Lipsey,* R. G. (1975): The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition, in: Review of Economic Studies Vol. 42, S. 27 - 39.

Hannesson, R. (1982): Defensive Foresight rather than Minimax: A Comment on Eaton and Lipsey's Model of Spatial Competition, in: Review of Economic Studies Vol. 49, S. 653 - 657.

Hotelling, H. (1929): Stability in Competition, in: Economic Journal Vol. 39, S. 41 - 57.

Nash, J. F. (1951): Non-Cooperative Games, in: Annals of Mathematics Vol. 54, S. 286 - 295.

Selten, R. (1971): Anwendungen der Spieltheorie auf die politische Wissenschaft, in: Politik und Wissenschaft (Hrsg.: Maier, H./Ritter, K./Matz, U.), München, S. 229 - 243.

Wettbewerb durch Standortwahl in der Fläche

Die Cournot- und die Stackelberg-Lösung

Von *Horst Todt*, Hamburg *

I. Einführung

Seit jeher steht bei der Diskussion von Wettbewerbsphänomenen der Preiswettbewerb im Vordergrund. Andere Formen des Wettbewerbes wurden typischerweise ignoriert. Dies geschieht, soweit die Wettbewerbssituation nicht von vornherein jenseits von Preiserwägungen liegt, indem der *Preis* als differenzierendes Element durch *Fixierung auf einen einheitlichen Wert* ausgeschaltet wird.

Ansatzpunkt dieser Untersuchung und Beispiel für derartige Wettbewerbssituationen ist das Hotelling-Modell: Auf einer geraden Strecke ist die fest vorgegebene Nachfrage mit konstanter Dichte verteilt, die sich demjenigen Konkurrenten zuwendet, der am nächsten ist. Aktionsparameter der Konkurrenten ist allein der Standort. Eine Standortkonfiguration gilt hier als stabil, wenn keiner der Konkurrenten seinen Marktanteil vergrößern kann, indem er seinen Standort ändert. Für zwei Anbieter ist der Mittelpunkt der Strecke eine stabile Konfiguration (*Hotelling*), bei drei Anbietern existiert keine stabile Anordnung von Standorten (*Chamberlin*), wohl aber bei jeder höheren Anzahl von Anbietern (*Selten*). Bekannter als die standorttheoretische Einkleidung (Badestrand mit Eisverkäufern) ist die Interpretation der Strecke als links-rechts-Skala im politischen Meinungsspektrum, wobei die Parteien durch Standortwahl um Stimmen konkurrieren (*Downs*).

Auf der Basis eines eindimensionalen Wirtschaftsraumes wurde in der Literatur eine Reihe von Modellen entwickelt, die verschiedene ökonomische Aspekte berücksichtigen (Standortänderungskosten, allgemeine Kosten, Preise, Skalenerträge). Demgegenüber soll hier zum *zwei-dimensionalen Wirtschaftsraum* übergegangen werden, wobei jedoch der Wettbewerb um Marktanteile durch Standortwahl der einzige Gesichtspunkt bleibt. Dieser Ansatz erlaubt nicht nur bessere Einblicke in die Natur des räumlichen Wett-

* Für Anregungen und z.T. auch mathematische Unterstützung danke ich Herrn Prof. Dr. F. Bolle. Frau Dipl.-Mathematikerin Angelika Holdorf hat eine kritische Durchsicht des Manuskripts besorgt.

bewerbs, sondern bietet auch die Möglichkeit, die Standortwahl und den Parteienkampf um die Wählergunst allgemeiner zu diskutieren und weitere Wettbewerbssituationen unter das Modell zu subsumieren.

II. Annahmen und Problemstellung

1. Das *totale Marktgebiet* $M \subset \mathbb{R}^2$ sei beschränkt und konvex. Ein beliebiger Punkt $Q \in M$ wird durch den euklidischen Abstand $\Theta\ (Q, P)$ zu einem fest gewählten Referenzpunkt $P \in M$ und einer Richtung α (Winkel zu einer vorgegebenen Geraden durch P) dargestellt. Die Kurzbezeichnung Θ oder Θ_P wird alternativ eingesetzt, wenn daraus keine Unklarheit resultiert.

Über M ist die auf 1 normierte (nicht beeinflußbare) Gesamtnachfrage gemäß einer beliebigen Verteilungsfunktion n verteilt. M ist auch die Menge aller potentiellen Standorte von Anbietern.

Die Annahme eines konvexen totalen Marktgebietes M bedarf eines Kommentars. Durch die Wahl eines *konvexen* Marktgebietes soll sichergestellt sein, daß alle relevanten Standorte auch zulässig sind und nicht durch Randbedingungen ausgeschlossen werden. Die Bedingung kann als *freie Standortwahl* interpretiert werden.

2. Zum Marktgebiet M_P des Standortes P ($M_P \subset M$) gehören alle Punkte $R \in M$, die zu keinem konkurrierenden Standort weniger weit entfernt sind als zu P: Sei $\Theta\ (R, R')$ die euklidische Entfernung zwischen $R, R' \in M$ und $\tilde{S} \subset M$ die endliche Menge aller Standorte (hier im wesentlichen auf zwei konkurrierende Standorte beschränkt), dann ist $M_P := \{R \in M \mid \Theta\ (R, P)\ \delta \leq \Theta\ (R, Q), \forall\ Q \in \tilde{S} \setminus \{P\}\}$. Die Menge $G_{PQ} := M_P \cap M_Q$ heißt Grenze der Marktgebiete von Standort Q, P (kurz: Grenze von P bzw. Q oder Grenze der Marktgebiete P und Q). Wenn die Grenze G_{PQ} nicht leer ist, bildet sie das Segment einer Geraden (Mittelsenkrechte zur geraden Verbindung zweier Standorte). Ein Punkt R kann mehreren Grenzen angehören. Ein Punkt, der nicht zu einer Grenze gehört, liegt im Inneren des Marktgebietes eines Standortes. Die Nachfrage im Inneren eines Marktgebietes fällt dem zugehörigen Standort voll zu. Die Nachfrage eines Grenzpunktes R fällt zu gleichen Teilen auf alle Marktgebiete, denen er angehört. Haben zwei oder mehrere Anbieter den gleichen Standort, so fällt auf sie die Nachfrage des gemeinsamen Standortes zu gleichen Teilen. Die Zuordnung des Randes von M_p erweist sich als nicht wesentlich für das Problem.

3. *Eine stabile Standortkonfiguration* entspricht dem *spieltheoretischen Gleichgewichtspunkt in reinen Strategien.* Eine reine Strategie eines Anbieters i besteht in der Wahl eines Standortes $S_i \in M$. Liegen alle Standorte fest, so sind offensichtlich auch die Marktanteile $m_i\ (\tilde{S})$ determiniert. Sei $\tilde{S}/S_i$ eine Standortkonfiguration, bei der alle Standorte bis auf S_i festgelegt sind, so ist ein Gleichgewichtspunkt in reinen Strategien $\tilde{S}^*$ gegeben falls gilt:

$$m_i(\tilde{S}^*) = \max_{S_i \in M} m_i(\tilde{S}^*/S_i) \quad \forall_i.$$

Von jetzt ab wird *Gleichgewicht* synonym mit *Gleichgewichtspunkt in reinen Strategien verwendet.*

Hier soll allein der Fall von zwei Anbietern, also das Duopol behandelt werden. Weil sich die Marktanteile zu 1 addieren, genügt es, den Anteil eines Anbieters, sagen wir $m_1 = m$ zu betrachten.

Anbieter 1 will m maximieren, Anbieter 2 minimieren.

Wenn Standorte $S_1^*, S_2^* \in M$ von Anbieter 1 bzw. 2 existieren, so daß gilt:

$$m(S^1, S_2^*) \leq m(S_1^*, S_2^*) \leq m(S_1^*, S_2),$$

dann liegt eine stabile Standortkonfiguration vor, also ein ‚Nash-Gleichgewicht in reinen Strategien' oder spezieller ein ‚Sattelpunkt in reinen Strategien'.

Angenommen es existiere kein (S_1^*, S_2^*), das diese Bedingung erfüllt, dann ist – zumindest bei einer standorttheoretischen Interpretation – ein eventuell existierendes Gleichgewicht in gemischten Strategien keine sinnvolle Lösungskonzeption; denn die realisierte, konkrete Lösung wird stets mindestens einem Anbieter Veranlassung geben, seinen Standort zu ändern. Die Standortkonfiguration ist nicht stabil. Es wird hier darum schlechthin gesagt, daß eine Cournot-Lösung nicht existiere.

4. Das Modell beschreibt eine Situation angemessen, in der zwei Anbieter in den Markt einsteigen wollen, von einander wissen, aber entscheiden müssen, ohne die Wahl des jeweils anderen zu kennen. Eine solche Informationsstruktur entspricht dem Cournotschen Duopol. (S_1^*, S_2^*) – sofern existent – bildet demnach den Cournot-Punkt beim Standortwettbewerb.

Es ist zu fragen, inwieweit die Cournot-Lösung angemessen ist. Beim Mengen- oder Preisduopol bindet eine Entscheidung die Konkurrenten nur für eine mehr oder weniger kurze Periode. Auf lange Sicht ist ein (stillschweigendes) Zusammenwirken (Collusion) eher realistisch. Wird eine kurze Periode zugrunde gelegt, so handelt es sich in der Regel um weniger schwerwiegende Entscheidungen. In solchen Situationen ist es häufig sinnvoll anzunehmen, daß die Akteure ihre Entscheidung unabhängig („gleichzeitig") treffen.

Beim Standortwettbewerb bindet die Entscheidung meist lange. Eine ‚einmalige' Entscheidung ohne offene oder stillschweigende Absprachen ist plausibler. Andererseits fällt es schwer zu glauben, daß die Entscheidungsprozesse gleichzeitig ablaufen. Typischerweise wird sich *ein* Konkurrent zuerst niederlassen, durchaus in dem Bewußtsein, daß ein weiterer folgen wird; der

nachfolgende wird sich in Kenntnis des gegnerischen Standortes entscheiden. Ein Problem besteht primär für den ersten Anbieter, der seinen Standort so wählen muß, daß der vom zweiten erreichbare Marktanteil möglichst klein wird. Diese Informationsstruktur entspricht der Stackelberg-Lösung.

Auch die Stackelberg-Lösung kann in ähnlicher Weise wie die Cournot-Lösung charakterisiert werden:

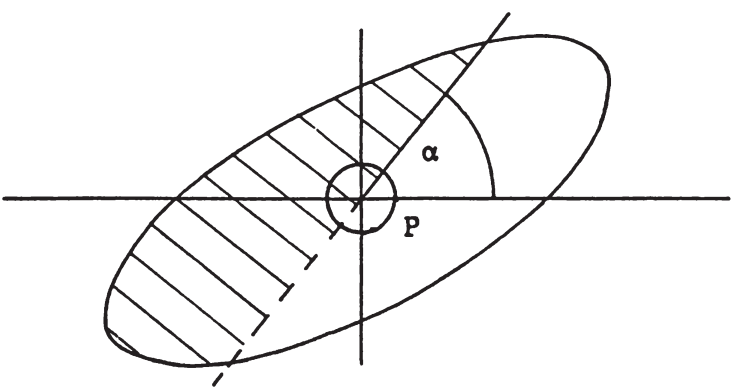
$$m(S_1, S_2^0(S_1)) \leq m(S_1^0, S_2^0(S_1^0)) \leq m(S_1^0, S_2(S_1^0)),$$

wobei $(S_1^0, S_2^0(S_1^0))$ die Stackelberg-Lösung darstellt. Anbieter 1 tritt zuerst in den Markt ein; Anbieter 2 setzt seinen Standort in Abhängigkeit vom Standort des Anbieters 1 fest und rechnet damit, daß der andere optimal auf diese Wahl reagiert $(S_2^0 = S_2^0(S_1))$. Für Anbieter 1 ist dann S_1^0 die beste Entscheidung. Anbieter 2 kann bei einer anderen nicht-optimalen Reaktion z.B. $S_2(S_1^0)$ günstigstenfalls gleich gut fahren.

III. Der Cournot-Punkt in der Ebene

Als Halbebene des Marktgebietes M zum Punkt P mit Richtung $\alpha \in [0; 2\pi]$ sei definiert:

$$M_{P\alpha} := \{Q_P(\Theta, \tau) \in M \mid 0 < \Theta, \alpha < \tau \leq \tilde{\alpha}\}$$



Hierbei ist $\tilde{\alpha} = \alpha + \pi$ die Gegenrichtung zu α.

Es gilt folgender Satz:

Satz 1[1]:

(1) Eine stabile Standortkonfiguration für ein beschränktes und konvexes Marktgebiet $M \subset \mathbb{R}^2$ und eine beliebige Nachfrageverteilung n über M mit $n(M) = 1$ existiert genau dann, wenn es einen Punkt $P \in M$ gibt, so daß gilt:

$$| n(M_{P\alpha}) - n(M_{P\bar{\alpha}}) | \leq n(P) \qquad \forall \alpha$$

(2) P bildet einen Gleichgewichtsstandort für beide Anbieter, falls die unter (1) genannte Bedingung erfüllt ist.

Anmerkung: Es wird nicht ausgeschlossen, daß für beide Anbieter zwei verschiedene Standorte P_1, P_2 eine stabile Konfiguration bilden.

Trivialbeispiel:

$^1/_2$	$^1/_2$	Nachfrage
*	*	
P_1	P_2	

Wenn aber (P_1, P_2) mit $P_1 \neq P_2$ ein Gleichgewicht ist, dann existiert auch ein gleichgewichtiger Doppelstandort, wie im Beispiel sofort zu sehen ist (P_1 oder P_2 oder jeder Punkt auf der geraden Verbindungslinie von P_1 und P_2). Es ist bemerkenswert, daß der Gleichgewichtspunkt P seine Lage nicht ändert, wenn die Entfernung der Nachfrage bzw. der Nachfragedichte in einer *fixierten Richtung* α beliebig variiert wird. Es ist also möglich, die gesamte Nachfrage – mit Ausnahme von $n(P)$ – richtungstreu auf den Einheitskreis um P zu projizieren, ohne daß sich das Problem in relevanter Weise ändert. Genau der gleiche Freiheitsgrad wird bei dem klassischen Standortproblem von *Launhardt* und *Weber* beobachtet. Hier ist für eine Nachfrageverteilung n über M der Produktionsstandort zu finden, der die Belieferung der Nachfrage entlang der Luftlinie bei minimalen Wegen erlaubt.

Genauer:

$$\int_{Q \in M} \| P - Q \|_2 \, dn \rightarrow \min_{P \in M}$$

Die Punkte P und Q sind hier, der üblichen Darstellung beim *Launhardt-Weber*-Problem folgend, in kartesischen Koordinaten erfaßt. $\| . \|_2$ ist die euklidische Norm. Der optimale Punkt P ist eine Verallgemeinerung des Zentralwertes auf den $\mathbb{R}^2$. Das Minimum existiert immer und ist eindeutig, weil das Integral eine strikt konvexe Funktion von P ist, sofern nicht die gesamte Nachfrage (bis auf eine Nullmenge) auf einer Geraden liegt. P steht hinfort

[1] Beweis siehe Anhang

für den Zentralwert und ist der Referenzpunkt (bzw. Ursprung), sofern nicht etwas anderes eigens vermerkt wird.

Angenommen P sei gefunden und das Integral sei zerlegt in

$$\int\limits_{\substack{Q \in M \\ Q \notin P}} \| P - Q \|_2 \, dn + n(P).$$

Bei Darstellung in kartesischen Koordinaten, $P = (x_P, y_P)$ und $Q = (x_Q, y_Q)$ ist

$$\| P - Q \|_2 = \sqrt{(x_P - x_Q)^2 + (y_P - y_Q)^2}.$$

Die partiellen Ableitungen lauten für $Q \neq P$:

$$\frac{\partial \sqrt{\ }}{\partial (x_P, y_P)} = \left[\frac{x_P - x_Q}{\sqrt{\ }}, \frac{y_P - y_Q}{\sqrt{\ }} \right] = (\cos \alpha, \sin \alpha),$$

wobei α der Winkel zur Abszisse ist. Das Integral

$$\| \int\limits_{\substack{Q \in M \\ Q \notin P}} (\cos \alpha, \sin \alpha) \, dn \|_2 = \| \int \cos \alpha \, dn, \int \sin \alpha \, dn \|_2$$

ist die Resultierende der *nachfragegewichteten Richtungen*. Für $n(P) = 0$ muß es den Wert Null annehmen. Allgemein gilt das *Kriterium von Kuhn* für den Zentralwert:

$$\| \int\limits_{\substack{Q \in M \\ Q \notin P}} (\cos \alpha, \sin \alpha) \, dn \|_2 \leq n(P)$$

das notwendig und hinreichend dafür ist, daß P ein Minimum ist.

Es ist nun interessant, das Kriterium für eine stabile Standortkonfiguration mit dem *Kuhnschen* Kriterium für den Zentralwert zu vergleichen. Für die Beziehung zwischen beiden Kriterien gilt folgender Satz:

Satz 2:

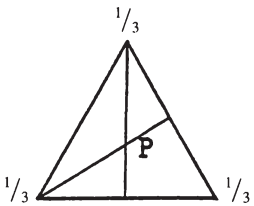
$$| n(M_{P\alpha}) - n(M_{P\bar{\alpha}}) | \leq n(P) \quad \forall \alpha \implies$$

$$\| \int\limits_{\substack{Q \in M \\ Q \notin P}} (\cos \alpha, \sin \alpha) \, dn \|_2 \leq n(P)$$

Der Gleichgewichtsstandort liegt also stets im Zentralwert. Allerdings gilt das Umgekehrte nicht allgemein; denn der Zentralwert existiert immer, die Gleichgewichtslösung dagegen nicht.

Beispiel:

Sei M ein gleichseitiges Standortdreieck mit gleicher Nachfrage an den Ecken. Der Zentralwert ergibt sich in diesem Fall als Schnittpunkt der Seitenhalbierenden:



Der Zentralwert P ist offensichtlich kein Gleichgewichtsstandort.

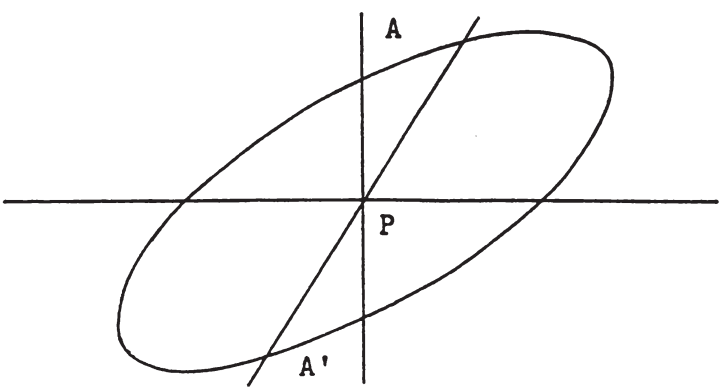
Es treffen für den Gleichgewichtsstandort jedoch die allgemeinen Aussagen über den Zentralwert zu. Insbesondere ist das Gleichgewicht eindeutig, wenn die Nachfrage nicht auf eine Gerade konzentriert ist. Der Fall mehrerer Gleichgewichtsstandorte ist mithin durch folgende Beschreibung charakterisiert:

1. Die Gesamtnachfrage (bis auf eine Nullmenge) liegt auf einer Geraden.
2. Es gibt auf der Geraden eine Lücke mit einer Null-Nachfrage, so daß auf beiden Seiten genau die Hälfte der Gesamtnachfrage liegt.

Es handelt sich bei dieser Situation um einen sehr speziellen Fall.

Schließlich gibt es für den Zentralwert (zumindest für die diskrete Nachfrageverteilung) einen leistungsfähigen Algorithmus, der auch die Gleichgewichtslösung liefert, sofern eine solche überhaupt existiert.

Eine Interpretation des Resultates: Die Bedingung für das Gleichgewicht bedeutet für $n(P) = 0$ totale Symmetrie:



Jede Gerade durch P halbiert das Marktgebiet in bezug auf die Nachfrage. *Gegenüberliegende* Segmente (A und A') tragen stets die gleiche Nachfrage. Für $n(P) > 0$ impliziert die Bedingung, daß die *Nachfragemasse* in $n(P)$ ausreichen muß, um für alle gegenüberliegenden Segmente mit unterschiedlicher Nachfrage einen Ausgleich zu schaffen.

Diese auch in dieser abgeschwächten Form noch sehr starke Forderung wird in Anwendungen nur selten erfüllt sein, so daß in der Regel *kein Gleichgewicht* existiert.

Eine allgemeine Behandlung des Wettbewerbs für drei und mehr Anbieter wurde m.W. bisher noch nicht durchgeführt. *Shaked* [7] zeigt für stetige Verteilungen (nicht geradlinig begrenztes Marktgebiet mit positiver Dichte), daß bei drei Anbietern kein Gleichgewicht existiert. Bei vier Anbietern scheinen nur sehr spezielle Situationen ein Gleichgewicht zu ermöglichen. Allgemein dürfte ein Gleichgewicht eher die Ausnahme als die Regel sein.

IV. Der Stackelberg-Punkt

1. Nunmehr soll ein Vergleich der Cournot-Lösung mit der Stackelberg-Lösung durchgeführt werden. Wenn ein Cournot-Gleichgewicht existiert, gilt nach Satz 1:

$$| n(M_{P\alpha}) - n(M_{P\tilde{\alpha}}) | \leq n(P) \quad \forall \alpha$$

Angenommen der zuerst in den Markt tretende Anbieter siedele am Zentralwert, dann erhält der später hinzutretende Anbieter, wenn er ebenfalls am Zentralwert siedelt, genau wie der erste einen Marktanteil von $\frac{1}{2}$; an keinem anderen Standort kann er mehr bekommen. Also ist der Cournotpunkt auch der optimale Standort für beide Stackelberg-Anbieter[2].

2. Es soll weiterhin geprüft werden unter welchen Voraussetzungen eine Stackelberg-Lösung auch ohne Cournot-Lösung existiert. Die jeweils ‚zusammengehörenden' Halbebenen $M_{R\alpha}$ und $M_{R\tilde{\alpha}}$ bestimmen zu $R \in M$ alternative Richtungen α. Wenn sich der erste Anbieter in $R \in M$ niederläßt, wird der zweite diejenige Richtung α bestimmen für die $| n(M_{R\alpha}) - n(M_{R\tilde{\alpha}}) |$ maximal wird. Der erste wiederum wird von vornherein R so festlegen, daß gilt:

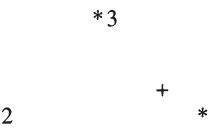
$$D_{S\alpha^*} := | n(M_{S\alpha^*}) - n(M_{S\tilde{\alpha}^*}) | = \min_R \max_\alpha | n(M_{R\alpha}) - n(M_{R\tilde{\alpha}}) |.$$

Der erste Anbieter wird sich – sofern er rational handelt – in S niederlassen, der zweite ‚in der Nähe' von S (d.h. in einem Abstand, der mit einem ver-

[2] Ein Algorithmus für den diskreten Fall wurde von *Drezner* a.a.O. entwickelt.

nachlässigbaren Verlust verbunden ist) u.z. auf dem Lot in S zur Richtung α^* zur ‚größeren' Marktseite. Exakt in S darf der zweite Anbieter seinen Standort nicht wählen; denn dann erhielte er den Marktanteil $\frac{1}{2}$ und würde sich seines strategischen Vorteils begeben. Weil ‚in der Nähe' ein ungenauer Ausdruck ist, bleibt eine kleine Unbestimmtheit bei der Stackelberg-Lösung. S ist die kennzeichnende Größe und soll deshalb als Stackelberg-Punkt bezeichnet werden.

3. Ein einfaches Beispiel bietet das klassische Standortdreieck vom Typ ‚Launhardt-Weber'.



Drei Punkte (mit der zugehörigen Nachfrage als Zahl) sollen von zwei Stackelberg-Anbietern bedient werden. Bei + liegt der Zentralwert. Was immer der erste Anbieter tut, der zweite kann sich zwei Nachfragepunkte sichern. Also wird der erste sich den Punkt mit größter Nachfrage sichern wollen; d.h. die Ecke mit dem Gewicht 4 ist Stackelberg-Punkt. Dies kann er nur, indem er dort siedelt. Dem zweiten Anbieter stehen viele äquivalente optimale Standorte zur Verfügung, nicht unbedingt in der Nähe von S.

4. Weniger durchsichtig ist ein zweites Beispiel. Es sei ebenfalls ein Standortdreieck (= Marktgebiet M) gegeben; die Nachfrage soll allerdings mit *konstanter Dichte* über das gesamte Marktgebiet verteilt sein.

Zunächst sei festgestellt, daß keine Cournot-Lösung existiert. Weil es keine diskrete Nachfrage im Beispiel gibt, ist $n(R) = 0 \; \forall \; R \in M$. Die Bedingung für ein Cournot-Gleichgewicht impliziert dann $| n(M_{P\alpha}) - n(M_{P\tilde{\alpha}}) | = 0 \; \forall \; \alpha$. Dies bedeutet, daß alle das Marktgebiet halbierenden Geraden sich in einem Punkt, dem Zentralwert, schneiden. Wie aus dem folgenden Satz ersichtlich ist, gibt es einen derartigen Punkt nicht.

Satz 3:

(1) Wenn das Marktgebiet M ein beliebiges Dreieck mit konstanter Nachfragedichte ist, dann liegt der Stackelberg-Punkt S im Schwerpunkt (Schnittpunkt der drei Seitenhalbierenden).

(2) Seien α_1, α_2, α_3 die Steigungen der drei Seiten des Marktdreiecks, dann gilt:

$$| n(M_{S\alpha_i}) - n(M_{S\tilde{\alpha}_i}) | = \min_R \max_\alpha | n(M_{R\alpha}) - n(M_{R\tilde{\alpha}}) | = D_{S\alpha^*} > 0$$

für i = 1, 2, 3

Anmerkung: Parallel zu den drei durch (S, α_i) charakterisierten Geraden verlaufen das Marktgebiet halbierende Geraden. Sie bilden ein dem Marktgebiet ähnliches Dreieck, dessen Schwerpunkt ebenfalls S ist. $D_{S\alpha^*}$ beträgt $\frac{1}{9}$ des Marktgebietes, unabhängig von der Gestalt des Marktdreiecks.

Der zweite Anbieter siedelt dicht bei S und kann zwischen drei Richtungen wählen. Er erhält (fast) $\frac{5}{9}$ des Marktes.

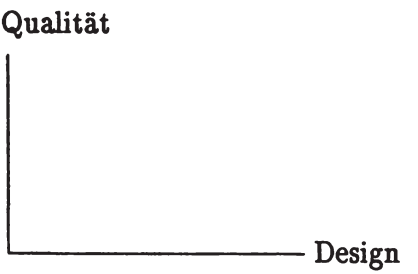
V. Resultate und Anwendungen

Wettbewerb durch Standortwahl sichert kein räumliches Gleichgewicht. Die Konzeption scheint allerdings auch zu einfach zu sein, um einen Vergleich mit der Realität zu rechtfertigen. Die regionalwissenschaftliche Diskussion hat eine Reihe von Argumenten geliefert, die für eine gewisse Stabilität räumlicher Konfigurationen spricht. Einmal bedingt der hohe Wert von Immobilien auch hohe Änderungskosten für den Standort. Nur langfristig ist ein Wandel wahrscheinlich, so daß zumindest temporär Stabilität vorherrscht. Zum anderen ist die Qualität eines Standortes von der Nachbarschaft anderer Standorte, nicht zuletzt von der Konkurrenz, geprägt. Viele konkurrierender Anbieter an einem Platz bieten der Kundschaft reichliche Möglichkeiten, zwischen den verschiedenen Alternativen und den Varianten im Wettbewerb zu wählen. Dies erhöht die Nachfrage an einem Ort intensiver Konkurrenz. Derartige Vorteile binden Standorte vor allem des tertiären Sektors aneinander. Sie sind sogar von überragender Bedeutung für die wirtschaftliche Entwicklung und die letzte Ursache der Stadtbildung. Eine kleinräumige Unruhe im Standortgefüge wird hierdurch jedoch nicht unbedingt ausgeschlossen. Mag sich die Standortstruktur somit auch nur langfristig und in kleinen Schritten wandeln, es gibt keinen Hinweis auf stationäres Verhalten. Solange nicht eine geschlossene Theorie der Standortfaktoren Gegenteiliges begründet muß der ständige Wandel durch Wettbewerb als eine Möglichkeit akzeptiert werden.

2. Im politischen Raum ist dieser ständige Wandel Realität. Erklärt das simple Hotelling-Modell trotz einer starken Vereinfachung, weshalb 2-Parteien-Systeme stabil sind, weshalb Führungen linker Parteien rechts von der Parteimitte und Führungen rechter Parteien links von der Parteimitte agieren, so bleibt doch ein wichtiger Aspekt bei eindimensionaler Betrachtung unberücksichtigt. Liegen zwei große Parteien auf einer Hauptdimension *in der Mitte* (mit kleinem Anstandsabstand) dicht beieinander, so wird Neigung bestehen, den Wettbewerb entlang einer weiteren Dimension, einer Nebendimension, zu betreiben. Diese Nebendimension (z.B. liberal - doktrinär) kann dann entscheidend für die Mehrheit sein. Gleichzeitig wird die postulierte Stabilität des 2-Parteien-Systems relativiert. Ständige Verlagerung des politischen Standortes der Parteien und sei es auch nur in einem mittleren Bereich, werden erklärlich.

3. *Chamberlin* verdanken wir das folgende, weithin akzeptierte Argument: An oligopolistischen Märkten ist jeder Preiswettbewerb mit erheblichen Risiken verknüpft, weshalb Preisänderungen nur ungern vorgenommen werden; der eigentliche Wettbewerb verlagert sich deshalb auf ein anderes Gebiet, nämlich die Produktvariation.

Hier sei eine alternative Erklärung des Qualitätswettbewerbs postuliert. Verschiedene Produkte, die als Substitute miteinander in Wettbewerb stehen, unterscheiden sich regelmäßig in einer großen Zahl von Merkmalen. Der Mensch ist nicht in der Lage, in einem solchen vieldimensionalen Feld sinnvolle Vergleiche anzustellen. Er aggregiert darum die relevanten Einzelmerkmale entlang von – sagen wir – zwei Skalen der Bewertung, *Qualität* und *Design*. Er akzeptiert, daß bessere Qualität und fortschrittliches Design einen höheren Preis rechtfertigen. Bei gleichem Preis haben (im Urteil der Kunden) schlechtere und weniger moderne Artikel keine Chancen. Der Preisunterschied stellt die Chancengleichheit der Produkte wieder her. Bei festem Preisgefüge verteilt sich dann die Nachfrage über Qualitäts-Design-Feld:



Eine Marke bildet einen Standort in diesem Feld. Nachfrager wählen die ihren Bedürfnissen nächste Marke.

Das grundlegende Problem besteht darin, eine sachgerechte Entfernung zwischen jeweils zwei Marken in diesem Diagramm zu definieren. Nun ist die euklidische Entfernung auch in übertragenen Anwendungen der elementaren Anschauung am nächsten. Es ist darum naheliegend, diese Entfernungskonzeption zur approximativen Beschreibung eines Sachverhaltes heranzuziehen. Auch das im Qualität-Design-Raum entscheidende Individuum wird größere oder kleinere Ähnlichkeit zwischen Marken feststellen. Wenn es kein Gleichgewicht bei einer derartigen Wettbewerbskonstellation gibt, ist der ständige Positionswechsel von Marken bzw. der Untergang oder Neueintritt von Marken zu erklären. Alternative Entfernungsmaße wie z.B. die Absolutnorm können allerdings in vielen Fällen der Euklidischen Entfernung vorgezogen werden. Ein solcher Ansatz soll hier nicht verfolgt werden.

4. Schließlich bietet sich der Standortwettbewerb generell für eine realitätsnähere Beschreibung des menschlichen Entscheidungsverhaltens (in einem eingeschränkt rationalem Sinne) an als sie die herkömmliche Nutzentheorie bietet. Es sei angenommen, die Menschen tragen ein ‚Bild' in sich von dem was erstrebenswert ist (z.B. ein Vorbild, ein ‚Ideal', einen ‚guten' Erfahrungsrichtwert, eine Faustregel etc.) und versuchen diesem Bild möglichst nahe zu kommen. Ferner soll dieses Bild als Punkt in einem Kriterienraum darstellbar sein. Dann könnte der tatsächliche Erfolg durch kreisförmige ‚Indifferenzkurven' dargestellt werden.

Der Abstand von einem erstrebenswerten Standard, Ideal, etc. wäre ein ‚Nutzenmaß'. Es wird hierbei das Axiom der Unersättlichkeit geopfert. Die üblichen sonstigen Axiome könnten aufrecht erhalten werden.

Anhang

Beweis zu Satz 1:

Es ist zu zeigen, daß $| n(M_{P\alpha}) - n(M_{P\tilde{\alpha}}) | \leq n(P)\ \delta,\ \forall\ \alpha$ $(\tilde{\alpha} = \alpha + \pi)$ notwendig und hinreichend für ein Gleichgewicht ist und daß P dann ein Gleichgewichtsstandort beider Anbieter ist.

1. Die Bedingung ist hinreichend.

Es existiert ein $P \in M$ mit $| n(M_{P\alpha}) - n(M_{P\tilde{\alpha}}) | \leq n(P)\ \delta$

O.B.d.A. gilt dann:

$$0 \leq n(M_{P\alpha}) - n(M_{P\tilde{\alpha}}) \leq n(P) \quad \forall\ \alpha$$

(ansonsten vertausche man α und $\tilde{\alpha}$).

Außerdem gilt nach Voraussetzung:

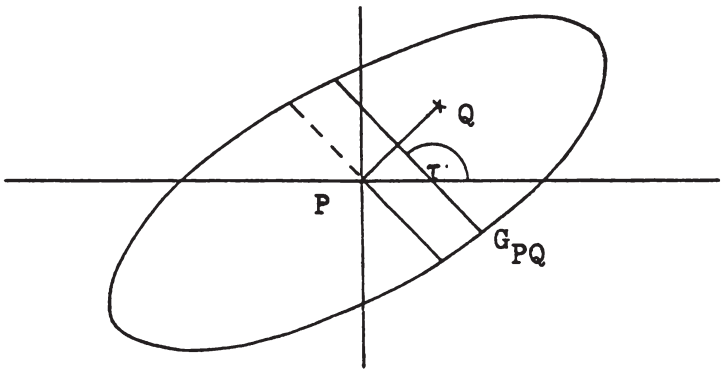
$$2n(M_{P\alpha}) \leq n(M_{P\alpha}) + n(M_{P\tilde{\alpha}}) + n(P) = n(M) = 1$$

Angenommen, die beiden Anbieter seien in P und $Q \neq P$, dann gibt es eine Marktgrenze G_{PQ} mit Steigungswinkel τ, so daß $M_Q \subset M_{P\tau}$.

Es ist dann

$$n(M_Q) \leq n(M_{P\tau}) \leq \frac{1}{2}$$

Der Standort Q bringt keine Verbesserung gegenüber P. Dies gilt für jedes zulässige Q, also ist P Gleichgewichtsstandort für beide Anbieter.



2. Die Bedingung ist notwendig.

a) Sei P Gleichgewichtsstandort für beide Anbieter und es gelte

$$| n(M_{P\alpha}) - n(M_{P\tilde{\alpha}}) | > n(P),$$
$$\text{dann gibt es } n(M_{P\alpha}) > \frac{1}{2}.$$

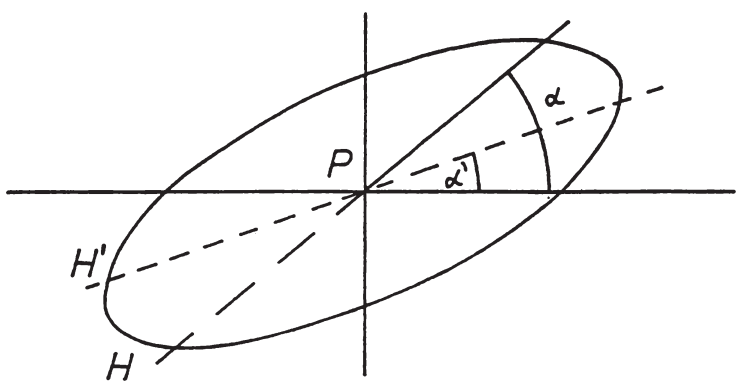
Eine offene Halbebene $M^0_{P\alpha}$ des Marktgebietes M zum Punkt P mit Richtung α sei wie folgt definiert:

$$M^0_{P\alpha} := \{Q_P(\theta, \tau) \in M, \mid 0 < \theta, \alpha < \tau < \tilde{\alpha}\}$$

Eine offene Halbebene $M^0_{P\alpha}$, mit $\alpha - \alpha' > 0$ kann dann so gewählt werden, daß die Nachfrage n' über dem Segment $HH'P$ die Bedingung

$$0 \leq n' < n(M_{P\alpha}) - \frac{1}{2}$$

erfüllt.



Nun ist $n(M^0_{P\alpha'}) > \frac{1}{2}$. Weil $M^0_{P\alpha'}$ offen ist, gibt es eine gerade Strecke $G_{PQ} \subset M^0_{P\alpha'}$ mit Steigung α', welche die Marktgrenze zu einem Standort $Q \neq P$ bildet, mit $n(M_Q) > \frac{1}{2}$. Jeder Standort Q' auf der Verbindungslinie QP bringt demjenigen Anbieter in P Vorteile, der nach Q wechselt, entgegen der Voraussetzung, daß P ein stabiles Gleichgewicht sei.

b) Seien P und Q zwei verschiedene, stabile Gleichgewichtsstandorte, so daß in jedem Standort ein Anbieter sitzt. Dann hat jeder den Marktanteil $m_1(P) = m_2(Q) = \frac{1}{2}$. Ansonsten könnte sich derjenige Anbieter mit geringerem Marktanteil an den Standort desjenigen mit größerem Marktanteil begeben und gewinnen, entgegen der Annahme, daß Gleichgewicht herrsche. Bewegt sich nun der Anbieter in Q auf der Strecke $\overline{PQ}$ auf P zu und vergrößert dabei seinen Marktanteil, so herrscht ebenfalls – entgegen der Voraussetzung – kein Gleichgewicht. Also vergrößert sich der Marktanteil nicht, und jeder Punkt Q' auf $\overline{PQ}$ ist ebenfalls möglicher Gleichgewichtsstandort. Das analoge Argument gilt für P. Darum ist jedes Punktepaar aus $\overline{PQ}$ einschließlich der Grenzen gleichgewichtig. Das heißt also, daß jeder Punkt Q' ein möglicher Doppelstandort ist.

Beweis zu Satz 2:

Es ist zu zeigen daß

$$\text{(i)} \qquad I = \| \int_{\substack{Q \in M \\ Q \notin P}} [\cos \alpha(Q), \sin \alpha(Q)]\, dn \|_2 \leq n(P)$$

immer erfüllt ist, wenn

$$\text{(ii)} \qquad | n(M_{P\alpha}) - (M_{P\tilde{\alpha}}) | \leq n(P)$$

gilt.

Hier ist $\alpha(Q)$ der Winkel des Vektors $Q \in M \subset \mathbb{R}^2$ in bezug auf den Ursprung P. Wie aus (i) ersichtlich, hat nur die Richtung α des Punktes Q, nicht aber die Entfernung zwischen Q und P Einfluß auf das Integral. Nun werde die Nachfrage in allen $Q \subset M$, $Q \neq P$ winkeltreu auf den Einheitskreis um P projiziert. Die entsprechende Verteilung der Nachfrage ist wie folgt definiert:

$$\bar{n}(\alpha) := n(\{Q_P(\theta, \tau) \mid 0 < \theta, 0 \leq \tau < \alpha\})$$

Bei geeigneter Wahl des Koordinatensystems wird $\int_0^{2\pi} \cos \alpha\, d\bar{n} = 0$; I läßt sich dann darstellen als

$$\text{(iii)} \qquad I = \int_0^{2\pi} \sin \alpha\, d\bar{n}(\alpha).$$

Das Integral kann als nicht-negativ angenommen werden (ansonsten drehe man das Koordinatensystem um 180°). Wegen $\sin(\alpha + \pi) = \sin \alpha$ folgt aus (iii)

$$I = \int_0^{\pi} \sin \alpha \, d(\bar{n}(\alpha) - \bar{n}(\alpha + \pi))$$

Als Grenzwert ausgedrückt ergibt dies

(iv)
$$I = \int_0^{\pi} \sin \alpha \, d(\bar{n}(\alpha) - \bar{n}(\alpha + \pi)) = \lim_{s \to \infty} \sum_{i=1}^{s} \sin \frac{\pi}{s} i$$
$$\left[\bar{n}\left(\frac{\pi}{s} i\right) - \bar{n}\left(\frac{\pi}{s}(i-1)\right) - \bar{n}\left(\frac{\pi}{s} i + \pi\right) + \bar{n}\left(\frac{\pi}{s}(i-1) + \pi\right)\right]$$

Der Inhalt der großen Klammer ist gleich der Differenz der Nachfrage auf den beiden Halbebenen $M_{Pi\pi/s}$ und $M_{P\,(i-1)\pi/s}$. Aus (ii) folgt:

(v)
$$\frac{1}{2} - n(P) \leq n(M_{P\alpha}) \leq \frac{1}{2}.$$

Angenommen, es gelte entgegen (v) $n(M_{P\alpha}) > \frac{1}{2}$, so ist $n(M_{P\alpha}) > n(M_{P\bar{\alpha}})$ und es folgt aus (ii)

$$\frac{1}{2} < n(M_{P\alpha}) \leq n(P) + n(M_{P\bar{\alpha}}) \qquad \text{oder}$$
$$1 < 2n(M_{P\alpha}) \leq n(P) + n(M_{P\bar{\alpha}}) + n(M_{P\alpha}) = 1.$$

Somit gilt (v).

Sei $X_P(\alpha) := n(M_{P\alpha}) - \left(\frac{1}{2} - n(P)\right) \geq 0.$

Es gilt wegen (v)

(vi)
$$X_P(\alpha) \leq n(P)$$

Für die Differenz der Nachfrage auf zwei Halbebenen ergibt sich:

$$n(M_{P\bar{\alpha}}) - n(M_{P\alpha}) = X_P(\bar{\alpha}) - X_P(\alpha)$$

Die rechte Seite von (iv) kann wie folgt umgeformt werden:

$$
\begin{aligned}
I &= \lim_{s \to \infty} \sum_{i=1}^{s} \sin \frac{\pi}{s} i \, (n(M_{Pi\pi/s}) - n(M_{P(i-1)\pi/s})) \\
&= \lim_{s \to \infty} \sum_{i=1}^{s} \sin \frac{\pi}{s} i \left[X_P\left(\frac{\pi}{s} i\right) - X_P\left(\frac{\pi}{s}(i-1)\right)\right] \\
&= \lim_{s \to \infty} \sum_{i=1}^{s-1} X_P\left(\frac{\pi}{s} i\right) \cdot \left(\sin \frac{\pi}{s}(i+1) - \sin \frac{\pi}{s} i\right) - X_P(\pi) \sin \pi + X_P(0) \sin \frac{\pi}{s}
\end{aligned}
$$

Wegen $\sin \pi = \lim\limits_{s \to \infty} \frac{\pi}{s} = 0$ ergibt sich

$$
I = \lim_{s \to \infty} \sum_{i=1}^{s-1} X_P\left(\frac{\pi}{s} i\right) \cdot \left(\sin \frac{\pi}{s}(i+1) - \sin \frac{\pi}{s} i\right)
$$

Sei $\left[\frac{s \pm 1}{2}\right]$ der auf die nächste ganze Zahl abgerundete Wert des Ausdrucks in der Klammer.

$$
\begin{aligned}
&\text{Für } i > \left[\frac{s+1}{2}\right] \text{ ist } \sin \frac{\pi}{s}(i+1) - \sin \frac{\pi}{s} i \leq 0; \\
&\text{für } \left[\frac{s-1}{2}\right] \geq i \text{ gilt } 0 \leq \sin \frac{\pi}{s}(i+1) - \sin \frac{\pi}{s} i.
\end{aligned}
$$

Wegen (v) ist dann

$$
\begin{aligned}
I &= \lim_{s \to \infty} \sum_{i=1}^{s-1} X_P\left(\frac{\pi}{s} i\right) \left(\sin \frac{\pi}{s}(i+1) - \sin \frac{\pi}{s} i\right) \\
&\leq \lim_{s \to \infty} \sum_{i=1}^{\left[\frac{s-1}{2}\right]} X_P\left(\frac{\pi}{s} i\right) \left(\sin \frac{\pi}{s}(i+1) - \sin \frac{\pi}{s} i\right) \\
&\leq n(P) \lim_{s \to \infty} \left(\sin \frac{\pi}{s}\left[\frac{s-1}{2}\right] - \sin \frac{\pi}{s}\right) = n(P).
\end{aligned}
$$

Beweis zu Satz 3:

Zuerst soll Teil (2) des Satzes bewiesen werden.

Die Ecken des Dreiecks seien mit $X_i = (x_i, y_i)$, $i = 1, 2, 3$, bezeichnet.

O. B. d. A. kann $X_3 = (0, 0)$ gesetzt werden.

Der Schwerpunkt ist dann $S = (x_s, y_s) = (\frac{1}{3}X_1 + \frac{1}{3}X_2)$

$$\text{Es sei } f\colon \left[\frac{1}{2}, 1\right] \to \left[\frac{1}{2}, 1\right]$$
$$\text{mit } f(\rho) = \frac{\rho}{3\rho - 1}.$$

Mit $\overline{X_i X_j}$ wird die Strecke zwischen den Punkten X_i und X_j bezeichnet.

Lemma 1: $\overline{f(\rho)X_1\ \rho X_2}$ und $\overline{\rho X_1\, f(\rho)X_2}$ schneiden sich in S für alle zulässigen ρ.

Beweis: Es ist zu zeigen, daß sich S als Konvexkombination $\lambda \cdot f(\rho)\, X_1 + (1 - \lambda) \cdot \rho\, X_2$ (bzw. $\lambda \cdot \rho\, X_1 + (1 - \lambda) \cdot f(\rho)\, X_2$) mit $\lambda \in [0, 1]$ darstellen läßt.

$\lambda = \dfrac{3\rho - 1}{3\rho}$ für zulässige Werte von ρ ergibt mit der ersten Gleichung S.

Die zweite Gleichung geht in die erste über, wenn man ρ durch $f(\rho)$ substituiert. Dies ist zulässig, weil $f(\rho)$ für das gleiche Intervall definiert ist wie ρ. Q.e.d.

Lemma 2: $g(\rho) := \dfrac{\rho^2}{3\rho - 1}$ gibt den Flächenanteil des Dreiecks $(\rho X_1, f(\rho) X_2, X_3)$ am Gesamtdreieck (X_1, X_2, X_3) an.

Beweis:

Die Fläche des Gesamtdreiecks beträgt $\Delta := {}^1/_2 \,\|X_1\|_2 \cdot \|X_2\|_2 \cdot \sin\gamma$ (γ ist der Winkel bei X_3). Der entsprechende Wert für das Teildreieck ist

$\Delta_\rho = {}^1/_2 \rho \cdot \|X_1\|_2 \cdot f(\rho) \cdot \|X_2\|_2 \cdot \sin\gamma$, so daß sich für den Anteil $\frac{\Delta_\rho}{\Delta} = \frac{\rho^2}{3\rho - 1}$ ergibt. Q.e.d.

Der Anteil $g(\rho)$ wird ein Minimum wenn:

$$\frac{dg}{d\rho} = \frac{3\rho^2 - 2\rho}{(3\rho - 1)^2} = 0 \quad \text{und} \quad \frac{d^2 g}{d\rho^2} > 0$$

Für den zulässigen Bereich ist dies allein der Wert $\rho = \frac{2}{3}$. Für diesen Wert von ρ ist auch $f(\rho) = \frac{2}{3}$. Also ist die größte Differenz zwischen Δ_ρ und Δ erreicht, wenn beide Dreiecke einander ähnlich sind, d.h. die gleichen Winkel aufweisen. Aufgrund der formalen Analogie gilt dieses Argument auch, wenn es für die beiden anderen Ecken durchgeführt wird. Anders ausgedrückt:

$$|n(M_{S\alpha_i}) - n(M_{S\tilde{\alpha}_i})| = \max_{\alpha} |n(M_{S\alpha}) - n(M_{S\tilde{\alpha}})|$$
$$\text{für } \alpha_i = \text{Steigung einer Seite.}$$

Um Teil (1) des Satzes einzusehen, beachte man zunächst, daß die das Marktgebiet M halbierende Gerade parallel zu einer Seite durch ein Dreieck

$\Delta_{\frac{1}{2}} = \frac{1}{2} \cdot \frac{1}{\sqrt{2}} \|X_1\|_2 \cdot \frac{1}{\sqrt{2}} \|X_2\|_2 \sin \gamma$ repräsentiert wird. Auf alle Seiten angewandt bedeutet dies, daß die drei halbierenden Geraden parallel zu einer Seite von M jeweils alle drei Seiten im Verhältnis $1 : \frac{1}{\sqrt{2}}$ verkürzen. Das entstehende M ähnliche Dreieck (zentrales Dreieck) hat S als Schwerpunkt.

Es ist nicht möglich sich im zentralen Dreieck von S zu entfernen ohne auch den Abstand zu einer der halbierenden Geraden zu vergrößern. Dadurch wird die Chance des zweiten Anbieters verbessert. Also ist S das Optimum für Anbieter 1.

Literatur

1. *Chamberlin,* E.: The Theory of Monopolistic Competition (1933), 8th edition.
2. *Downs,* A.: An Economic Theory of Political Action in a Democracy, in: Journ. Pol. Econ. 65 (1957), S. 135 - 150.
3. *Drezner,* Z.: Competitive Location Strategies for two Factilities, in: Regional Science and Urban Economics 12 (1982) S. 485 - 493.
4. *Hotelling,* H.: Stability in Competition, in: Economic Journal 39 (1929), S. 41 - 57.
5. *Kuhn,* H. W.: On a Pair of Dual Nonlinear Programs, in: Abadie, J. (Ed.), Nonlinear Programming, Amsterdam: North Holland Publishing Comp., 1967.
6. *Selten,* R.: Anwendungen der Spieltheorie auf die politische Wissenschaft, München: Verlag C. H. Beck, 1971, 287 - 320.
7. *Shaked,* A.: Non-existence of Equilibrium for the Two-dimensional Three-firms Location Problem, in: Rev. Econ. Stud. 42 (1975), S. 51 - 56.

Zur Anwendung des Konzepts konsistenter konjekturaler Reaktionen auf die Theorie der räumlichen Preisbildung

Von *Klaus Schöler*, Siegen

I. Einführung

Ebenso wie in der Makroökonomie wurde zu Beginn der Achtziger Jahre auch in der Mikroökonomie versucht, Erwartungen der Wirtschaftssubjekte zu endogenisieren. In der Makroökonomie wurde der Gedanke der „Rationalen Erwartungen" in stochastischen Modellen aufgegriffen; in die Oligopoltheorie wurden die konsistenten Erwartungen eines Unternehmens über die Reaktionen der Konkurrenten auf eigene Marktaktivitäten eingeführt. Der Begriff „konsistent" bedeutet in diesem Zusammenhang, daß die konjekturalen Reaktionen mit den tatsächlichen Reaktionen übereinstimmen müssen, anders gesagt, die von einem Unternehmen 1 erwarteten Reaktionen des Unternehmens 2 müssen konsistent sein mit der Zielsetzung und der ökonomischen Situation des Unternehmens 2 (vgl. z.B. *Bresnahan* [1981], *Kamien/Schwartz* [1983], *Boyer/Moreaux* [1984]). Im Rahmen der Oligopoltheorie wird üblicherweise die Zielsetzung der Gewinnmaximierung angenommen und die ökonomische Situation des Unternehmens durch Nachfrage- und Kostenfunktionen beschrieben. Dies bedeutet aber auch, daß – läßt man einmal alternative Zielsetzungen außer acht – die konsistenten konjekturalen Reaktionen in ihrer Stärke von den angenommenen Nachfrage- und Kostenfunktionen abhängen.

Die Verwendung exogener konjekturaler Reaktionen führt – wie sich am Lehrbuchbeispiel des Cournotmengenoligopols leicht zeigen läßt – zu einer unerfreulichen Inkonsistenz. Jedes Unternehmen erwartet im Cournotmodell, daß die Produktionsmenge der Konkurrenten bei eigenen Änderungen des Outputs unverändert bleibt; m.a.W. die konjekturalen Reaktionen betragen genau Null. Nun ist jedem Unternehmen aber die Tatsache bekannt, daß die eigene gewinnmaximale Menge vom Output der anderen Firmen im Markt abhängt. Aus dieser Situation ergibt sich die Frage, welche Informationen die Anbieter zu der Überzeugung bringen, daß die eigene tatsächliche Reaktionsfunktion sich von den konjekturalen Reaktionsfunktionen der anderen Firmen fundamental unterscheidet. Jede Firma müßte über Informationen verfügen,

die sie zu der Auffassung veranlassen, die eigene Position im Markt sei unterschiedlich von der aller anderen Firmen. Diese Annahme kann selbstverständlich aus rein logischen Gründen nicht für alle Unternehmen zutreffen, da in diesem Fall die Marktsituation aller Firmen gleich wäre. Folglich müssen sich einige oder alle Firmen zu Beginn eines zum Gleichgewicht führenden Marktprozesses irren. Dieser Irrtum über die Reaktionen der Konkurrenten wird jedoch im Zuge des Wettbewerbsprozesses nicht korrigiert; im traditionellen Oligopolmodell findet keine Anpassung der konjekturalen Reaktionen an die sich abzeichnenden tatsächlichen Verhaltensweisen statt. Der Verzicht aber, aus den Erwartungsfehlern der Vergangenheit zu lernen, steht in Widerspruch zu der Gewinnmaximierungsannahme des Modells. Jeder Prognosefehler über das Verhalten der Konkurrenten läßt Opportunitätskosten in Höhe des entgangenen Gewinns entstehen. Eine Lösung dieser Probleme kann in der Frage nach genau jenem konjekturalen Reaktionskoeffizienten gesehen werden, der im Einklang mit der gewinnmaximierenden Verhaltensweise des Konkurrenten steht, kurz gesagt, der mit dem Modell selbst konsistent ist. Diese Überlegungen gelten uneingeschränkt auch für Preisoligopole.

In den letzten Jahren haben die Ergebnisse dieser Diskussion auch in der räumlichen Preistheorie Beachtung gefunden: *Capozza* und *Van Order* [1985, 1989] und *DeCanio* [1984] haben das Konzept der konsistenten konjekturalen Reaktionen auf preisunelastische räumliche Nachfragefunktionen übertragen und gezeigt, daß der konsistente Reaktionskoeffizient genau $\frac{1}{3}$ beträgt. Dieses Ergebnis behält seine Gültigkeit auch für die Bestimmung räumlicher Gleichgewichtspreise (vgl. *Schöler* [1991]). Das Konzept der konsistenten konjekturalen Variationen kann ferner auf den Fall heterogener Güter bezogen werden, wobei die konsistenten Reaktionskoeffizienten von der Intensität der Substitutionsbeziehungen bestimmen werden (vgl. *Schöler* [1990]). Schließlich ist das Konzept der konsistenten konjekturalen Reaktionen für langfristige Marktgleichgewichte (*Schöler/Schlemper* [1991] und für unterschiedliche eindimensionale Marktstrukturen (*Schöler* [1992]) untersucht worden.

Welche Konsequenzen hat nun die Einführung des Konzepts der konsistenten konjekturalen Reaktionen im Rahmen der räumlichen Preistheorie? Zunächst einmal reduziert sich die Menge der Modellösungen auf genau eine Lösung, nämlich die konsistente Lösung. Seit etwa zehn Jahren hat sich in der Literatur zur räumlichen Preistheorie die Unterscheidung von drei Modellvarianten mit exogenen konjekturalen Reaktionen etabliert:

- Lösch-Wettbewerb (L) mit konjekturalen Reaktionen von 1,
- Hotelling-Smithies-Wettbewerb (HS) mit konjekturalen Reaktionen von 0,
- Greenhut-Ohta-Wettbewerb (GO) mit konjekturalen Reaktionen von -1.

Das nachfolgende Modell zeigt, daß die konsistenten Reaktionen unter den üblichen Modellannahmen weder den Wert 1 noch den Wert -1 annehmen.

Damit sind zwei der Wettbewerbsmodelle nicht konsistent in dem beschriebenen Sinne. Allenfalls kann der Hotelling-Smithies-Wettbewerb mit einem Reaktionskoeffizient von Null für ein bestimmtes Intervall von Marktgebietsgrößen als hinreichende Approximation der konsistenten Lösung angesehen werden.

Ferner läßt sich ein weiteres, zwar plausibles, aber keineswegs triviales Resultat, mit Hilfe der konsistenten Lösungen ableiten: Je größer die Anzahl der direkten, das heißt in unmittelbarer räumlicher Nachbarschaft angesiedelten Konkurrent ist, um so geringer ist das Ausmaß der konsistenten Reaktion, die von einem dieser Konkurrenten auf eine Preisvariation des betrachteten Unternehmens erwartet wird. Dieses Ergebnis läßt sich leicht erklären: Je mehr direkte Konkurrenten existieren, um so unbedeutender ist die Preisveränderung eines Anbieters und um so schwächer ist die Reaktion auf diese Änderung.

Schließlich leistet die Einbeziehung konsistenter konjekturaler Reaktionen zwei – wie ich meine – wichtige Beiträge zu langjährig kontrovers diskutierten Fragen innerhalb der Preistheorie und insbesondere innerhalb der räumlichen Preistheorie.

(1) Die im deutschen Schrifttum der fünfziger Jahre geführte Diskussion „Marktform oder Verhaltensweise" oder – wie man auch sagen könnte – „Marktstruktur oder Verhaltensweise", erscheint in einem neuen Licht und wird entscheidbar. Wenn die konsistenten Reaktionen von Nachfrage- und Kostenfunktionen abhängen, dann folgen die nunmehr endogenisierten Verhaltensweisen der Marktstruktur.

(2) Seit den grundlegenden Arbeiten von *Lösch* [1944] und *Mills/Lav* [1964] werden in der Literatur die Marktergebnisse des räumlichen Wettbewerbs unter Verwendung zweidimensionaler Marktgebiete diskutiert. In einer Vielzahl von Aufsätzen steht dabei die optimale geometrische Form des Marktgebietes im Vordergrund des Interesses (vgl. *Hartwick* [1973], *Eaton/Lipsey* [1976]). Bis zu dem Beitrag von *Ishikawa* und *Toda* (*Ishikawa/Toda* [1990]) wurden die Wettbewerbsmodelle im zweidimensionalen Raum ohne die Marktreaktionen der Konkurrenten modelliert; ein seltsamer Mangel für regionale Oligopole. *Ishikawa* und *Toda* fanden heraus, daß für den Lösch-Wettbewerb und den Hotelling-Smithies-Wettbewerb aus wohlfahrtstheoretischer Sicht die Rangfolge der Marktfiguren lautet: Sechseck-Viereck-Dreieck, ein Ergebnis, das aus früheren Untersuchungen schon bekannt war. Für den Greenhut-Ohta-Wettbewerb zeigt sich jedoch, daß bei kleinen Marktgebieten das Sechseck, bei großen Marktgebieten das Viereck aus wohlfahrtstheoretischer Sicht vorzuziehen ist. Unter Verwendung konsistenter konjekturaler Reaktionen ergibt sich jedoch wieder die eindeutige Rangfolge: Sechseck-Viereck-Dreieck. Ich bin der Auffassung, daß die Auswahl einiger kurz vorgestellter Ergebnisse zeigt, daß die Mühe der Verwendung konsistenter

konjekturaler Reaktionen in der räumlichen Preistheorie durchaus lohnenswert ist.

In der vorliegenden Untersuchung sollen zunächst die Annahmen vorgestellt werden, unter denen ein Modell mit zweidimensionalen Marktgebieten diskutiert wird (Abschnitt II). Die Abschnitte III bis V umfassen das Grundmodell, die konsistenten konjekturalen Reaktionen und die Marktergebnisse sowie Wohlfahrtseffekte des Modells. In Abschnitt VI werden die Ergebnisse kurz zusammengefaßt.

II. Modellannahmen

Für die nachfolgenden Überlegungen ist es sinnvoll, einige vereinfachende Annahmen einzuführen und eine Reihe von Variablen auf 1 zu standardisieren, ohne daß dadurch die Allgemeinheit der Modellaussagen eingeschränkt wird.

A1: Die Nachfrager sind mit einer konstanten Dichte von $B = 1$ über die zweidimensionale Gesamtmarktfläche kontinuierlich verteilt.

A2: Die Nachfrage q ist für alle Konsumenten gleich und ist eine lineare Funktion des Ortspreises p am Standort des Nachfragers:

$$q(r) = 1 - p \quad \text{mit} \quad p = m + r \tag{1}$$

Der Ortspreis setzt sich zusammen aus dem Ab-Werk-Preis m des Anbieters und den Transportkosten zwischen Produktions- und Konsumtionsstandort, wobei r die Entfernung zwischen beiden Orten sei und die Transportkosten je Mengen- und Entfernungseinheit mit $f = 1$ angenommen werden.

A3: Die Entfernung zwischen den Standorten der Firmen sei jeweils L; die Marktgebiete der Unternehmen haben eine Ausdehnung zwischen Standort und Wettbewerbsgrenze von R. Die Standorte liegen jeweils in der Mitte des gesamten firmenindividuellen Marktgebietes.

A4: Es wird angenommen, daß zwischen den zweidimensionalen Marktgebieten keine unversorgten Restgebiete verbleiben, so daß als regelmäßige Formen der Marktgebiete regelmäßige Dreiecke ($n = 3$), Vierecke ($n = 4$) oder Sechsecke ($n = 6$) in Betracht kommen.

A5: Das gehandelte Gut wird von allen Firmen unter den gleichen technologischen Bedingungen produziert; die variablen Kosten der Produktion mögen 0 sein:

$$K = K_f \qquad K_f = \text{Fixkosten} \tag{2}$$

A6: Die Firmen verfolgen das Ziel der Gewinnmaximierung; die Nachfrager verfolgen das Ziel der Konsumentenrentenmaximierung und kaufen das Gut jeweils bei dem Unternehmen, das das Gut zum niedrigsten Ortspreis anbietet. Daraus folgt, daß an der Marktgebietsgrenze zwischen zwei Anbietern i und j die Ortspreise identisch sind ($p_i = p_j$) und die Entfernung zwischen Standort und Marktgebietsgrenze aus

$$m_i + R = m_j + (L - R) \tag{3}$$

mit

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

bezeichnet werden kann.

III. Das Modell

Die auf ein Unternehmen i entfallende Marktnachfrage S_i beträgt unter Verwendung der Annahmen A1 bis A4 im Fall regelmäßiger zweidimensionaler Marktgebiete:

$$S_i(m_{i,n}, R) = 2n \int_0^{\pi/n} \int_0^{R/\cos\theta_n} r(1 - m_{i,n} - r)\, dr\, d\theta_n \tag{4}$$

wobei n die Anzahl der Ecken (oder Seitenlinien) des regelmäßigen firmenindividuellen Marktgebietes repräsentiert, ferner sind R der zugehörige Innenkreisradius des Marktgebietes, π in der oberen Integrationsgrenze die Zahl 3,1416... und θ_n der Winkel, der am Standort des Anbieters zwischen einer zu einem Eckpunkt verlaufenden Linie und der nächsten Seitenhalbierenden entsteht (vgl. z.B. *Mills/Lav* [1964]). Der Gewinn der Firma i kann unter Berücksichtigung der Annahme A4 mit

$$\Pi_i(m_{i,n}, R) = m_{i,n} S_i(m_{i,n}, R) - K_{fi} \tag{5}$$

bezeichnet werden und lautet nach Lösung des Doppelintegrals in Gleichung (4):

$$\Pi_i(m_{i,n}, R) = 2nR^2 m_{i,n} [\alpha_n (1 - m_{i,n}) - \beta_n R] - K_{fi} \tag{6}$$

mit:

$$\alpha_3 = \sqrt{\frac{3}{2}}, \quad \alpha_4 = \frac{1}{2}, \quad \alpha_6 = \frac{1}{(2\sqrt{3})},$$
$$\beta_3 = 0{,}79684, \quad \beta_4 = 0{,}38259, \quad \beta_6 = 0{,}20266$$

Die Bedingung 1. Ordnung für ein Gewinnmaximum lautet:

$$\partial \Pi_i / \partial m^*_{i,n} = 2n\,[(2\alpha_n R\,(1 - m^*_{i,n}) - 3\beta_n R^2)\,R'_n m^*_{i,n} + R^2\,(\alpha_n - \beta_n R - 2\alpha_n m^*_{i,n})] = 0 \tag{7}$$

mit $R'_n = dR/dm_{i,n}$. Es kann davon ausgegangen werden, daß die konjekturalen Reaktionsannahmen nicht nur hinsichtlich eines Nachbarkonkurrenten j, sondern auch bezüglich der Nachbarkonkurrenten k_1 bis k_{n-1} identisch sind:

$$(dm_{j,n}/dm_{i,n})^c = (dm_{k_1,n}/dm_{i,n})^c = \ldots = (dm_{k_{n-1},n}/dm_{i,n})^c = \phi_{ji,n} = \phi_{k_1 i,n} = \ldots = \phi_{k_{n-1} i,n} = \phi_n \tag{8}$$

Da wiederum die Kostenfunktionen aller am Markt befindlichen Unternehmen und die auf sie unter gleichen Bedingungen entfallenden Nachfragemengen gleich sind, gibt es keinen rationalen Grund für Unternehmen i, unterschiedliche konjekturale Verhaltensweisen den Konkurrenten j und k_1 bis k_{n-1} zuzuordnen. Somit kann R' auch als

$$R'_n = (\phi_n - 1)/2\omega_n \tag{9}$$

geschrieben werden, wobei $\omega_3 = 2$, $\omega_4 = \sqrt{2}$, $\omega_6 = \frac{2}{\sqrt{3}}$ ist. Für die Bedingung 2. Ordnung muß ferner gelten:

$$2nR'_n\,[2R\alpha_n\,(2 - 4m^*_{i,n} - R/R'_n) - 6\beta_n R^2 + (2\alpha_n\,(1 - m^*_{i,n}) - 6\beta_n R)\,R'_n m^*_{i,n}] < 0 \tag{10}$$

Diese Bedingung ist für alle $\phi_n \in [-1, 1)$ erfüllt. Aus der Gleichung (7) kann der gewinnoptimale Ab-Werk-Preis errechnet werden:

(11)

$$m^*_{i,n}(R, \phi_n) = \left[\frac{1}{2} - \frac{R}{2}\left[\frac{2\omega_n}{\phi_n - 1} + \frac{3\beta_n}{2\alpha_n}\right]\right] - \left[0{,}25\left[1 - R\left[\frac{2\omega_n}{\phi_n - 1} + \frac{3\beta_n}{2\alpha_n}\right]\right]^2 + \frac{\omega_n R}{\phi_n - 1} - \frac{\omega_n \beta_n R^2}{\alpha_n\,(\phi_n - 1)}\right]^{0{,}5}$$

Ein Vergleich der Marktergebnisse bei exogenen und endogenen Reaktionskoeffizienten setzt voraus, daß zunächst die konsistenten konjekturalen Reaktionskoeffizienten ermittelt werden.

IV. Konsistente konjekturale Reaktionen

Konsistente konjekturale Reaktionen liegen vor, wenn die von Firma i erwarteten Reaktionen des Konkurrenten j im Gleichgewicht mit den tatsächlichen Reaktionen der Firma j übereinstimmen. Diese Konsistenzforderung impliziert, daß das Unternehmen i die Preisreaktionsfunktion des Nachbarunternehmens j (oder einer Firma k) kennt:

$$2n\left[(2\alpha_n R(1 - m^*_{j,n}) - 3\beta_n R^2) R'_n m^*_{j,n} + R^2(\alpha_n - \beta_n R - 2\alpha_n m^*_{j,n})\right] = 0 \tag{12}$$

Der tatsächliche Reaktionskoeffizient $\hat{\phi}_n$ wird durch die Steigung der Reaktionsfunktion j (oder k) ausgedrückt, die durch die Differentiation der impliziten Funktion (12) nach $m^*_{i,n}$ errechnet wird. Dabei ist zwischen der reaktionsauslösenden Firma i und den passiv beteiligten und die Firma j umgebenden Firmen k_1 bis k_{n-1} zu unterscheiden. Da diese Firmen in jeder Hinsicht identisch sind, können die Terme für eine von ihnen mit $n-1$ multipliziert werden:

$$\begin{aligned} &2\left[(2R_i\alpha_n(1 - m^*_{j,n}) - 3R_i^2\beta_n) R'_{i,n} m^*_{j,n}\right. \\ &\quad \left. + R_i^2(\alpha_n - \beta_n R_i - 2\alpha_n m^*_{j,n})\right] \\ &\quad + 2(n-1)\left[(2R_k\alpha_n(1 - m^*_{j,n}) - 3R_k^2\beta_n) R'_{k,n} m^*_{j,n}\right. \\ &\quad \left. + R_k^2(\alpha_n - \beta_n R_k - 2\alpha_n m^*_{j,n})\right] = 0 \end{aligned} \tag{12'}$$

Die Größe R_i repräsentiert die kürzeste Entfernung zwischen dem Standort des Unternehmens j und der mit Firma i gemeinsamen Marktgebietsgrenze, die Größe R_k hat die gleiche Bedeutung für die gemeinsamen Grenzen mit den $n-1$ k-Unternehmen. Der tatsächliche Reaktionskoeffizient $\hat{\phi}$ lautet schließlich:

$$\hat{\phi}_n = \frac{dm^*_{j,n}}{dm^*_{i,n}} = -\frac{\xi_{i,n}}{\xi_{j,n}} \tag{13}$$

mit:

$$\xi_{i,n} = m^*_{j,n}\left[(\phi_n - 1)/(2\omega_n)\right]\left[\alpha_n(1 - m^*_{j,n})/\omega_n - 3\beta_n R_i/\omega_n\right] + R_i(\alpha_n - \beta_n R_i - 2\alpha_n m^*_{j,n} - \beta_n R_i/2)/\omega_n$$

und

$$
\begin{aligned}
\xi_{j,n} = {} & [2\alpha_n R_i (1 - m^*_{j,n}) - 3\beta_n R_i^2]\,[(\phi_n - 1)/(2\omega_n)] \\
& + m^*_{j,n}[(\phi_n - 1)/(2\omega_n)]\,[-\alpha_n (1 - m^*_{j,n})/\omega_n - 2\alpha_n R_i + 3\beta_n R_i/\omega_n] \\
& + (-\alpha_n R_i/\omega_n + 3\beta_n R_i^2/2\omega_n - 2\alpha_n R_i^2 + 2\alpha_n m^*_{j,n} R_i/\omega_n) \\
& + (n-1)\,[(\phi_n - 1)/(2\omega_n)]\,[2\alpha_n R_k (1 - m^*_{j,n}) - 3\beta_n R_k^2] + m^*_{j,n} \\
& \quad (n-1)\,[(\phi_n - 1)/(2\omega_n)]\,[2\alpha_n (\phi_n - 1)/2\omega_n - 2\alpha_n R_k \\
& - (\phi_n - 1)/\alpha_n m^*_{j,n}/\omega_n - 3\beta_n (\phi_n - 1) R_k/\omega_n] \\
& + (n-1)\,[\alpha_n (\phi_n - 1) R_k/\omega_n - 3\beta_n (\phi_n - 1) R_k^2/2\omega_n \\
& - 2\alpha_n R_k^2 - 2\alpha_n m^*_{j,n} R_k (\phi_n - 1)/\omega_n]
\end{aligned}
$$

Da im Fall konsistenter konjekturaler Reaktionen ϕ^*_n gilt: $\hat{\phi}_n = \phi_n$ und ferner $R_i = R_k = R$ angenommen werden kann, läßt sich Gleichung (13) umformen zu

$$
\begin{aligned}
& - [\alpha_n (m^*_{j,n})^2 (\phi^*_n - 1)^2 (n\phi^*_n - \phi^*_n - 1) + m^*_{j,n}(\phi^*_n - 1) \\
& \quad (\alpha_n (n\phi^*_n (8\omega_n R - \phi^*_n + 1) - (\phi^*_n + 1)(4\omega_n R - \phi^*_n + 1)) \\
& + 3\beta_n R (\phi^*_n - 1)(n\phi^*_n - \phi^*_n - 1)) \\
& + \omega_n R (2\alpha_n (2n\phi^*_n (\omega_n R - \phi^*_n + 1) + \phi^{*2}_n - 1) \\
& + 3\beta_n R (\phi^*_n - 1)(2\phi^*_n (n-1) + \phi^*_n - 1))]/2\omega_n^2 = 0
\end{aligned} \tag{14}
$$

Durch Einsetzen des gewinnoptimalen Ab-Werk-Preises der Firma j

(11')

$$
\begin{aligned}
m^*_{j,n}(R, \phi_n) = {} & \left[\frac{1}{2} - \frac{R}{2}\left[\frac{2\omega_n}{\phi^*_n - 1} + \frac{3\beta_n}{2\alpha_n}\right]\right] \\
& - \left[0{,}25\left[1 - R\left[\frac{2\omega_n}{\phi^*_n - 1} + \frac{3\beta_n}{2\alpha_n}\right]\right]^2 + \frac{\omega_n R}{\phi^*_n - 1} - \frac{\omega_n \beta_n R^2}{\alpha_n (\phi^*_n - 1)}\right]^{0{,}5}
\end{aligned}
$$

in Gleichung (14) können nunmehr die endogenen, d.h. konsistenten Reaktionskoeffizienten ϕ^*_n in Abhängigkeit vom Innenkreisradius R, der die Marktgebietsausdehnung repräsentiert, und von der geometrischen Form des Marktgebietes $n = 3, 4, 6$ iterativ ermittelt werden[1]. In allen Fällen erfüllen die gefundenen Koeffizienten die Bedingung 2. Ordnung (10) für ein Gewinnmaximum bezüglich der Preise m^*. Die Ergebnisse finden sich in Abbildung 1, zu der zwei Anmerkungen zu machen sind, die auch für alle weiteren graphischen Darstellungen Gültigkeit haben. Zum einen sind die gewinnmaximalen Ausdehnungen der Innenkreisradien als jeweilige End-

[1] Die Bestimmung der Nullstellen der Funktion (14) wurde bis zu einer Genauigkeit von kleiner/gleich $1 * 10^{-5}$ durchgeführt.

punkte der Kurven zu berücksichtigen. Unter Verwendung der Nichtnegativitätsbedingung für die Ortsnachfrage $1 - m - r \geq 0$ erhält man $R_3^{\max} = \frac{3}{8}$ $R_4^{\max} = \frac{3}{(4\sqrt{2}}$ und $R_6^{\max} = \frac{3}{\sqrt{\frac{3}{8}}}$ (vgl. *Greenhut/Ohta* [1973]).

Zum anderen wurden die Innenkreisradien für alle drei Figuren auf flächengleiche Radien standardisiert. Beträgt beispielsweise $R_4 = 0{,}1$, so lauten die zugehörigen flächengleichen Radien $R_3 = 0{,}087738$ und $R_6 = 0{,}107465$.

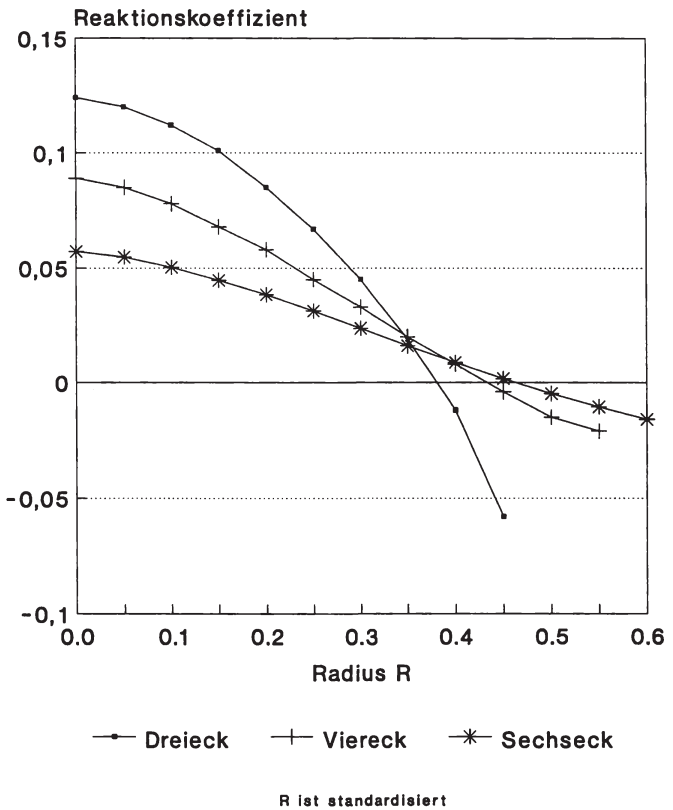


Abbildung 1: Konsistente konjekturale Reaktionen in zweidimensionalen Marktgebieten

Die Abbildung 1 zeigt, daß die konsistenten konjekturalen Reaktionskoeffizienten die Wertebereiche $0{,}116 > \phi_3 > -0{,}011$ für das Dreieck, $0{,}088 > \phi_4 > -0{,}010$ für das Viereck und $0{,}057 > \phi_6 > -0{,}013$ für das Sechseck aufweisen. Grundsätzlich kann festgehalten werden, daß die endo-

genen Reaktionskoeffizienten mit R und mit n variieren und nicht – wie in dem Beitrag von *Ishikawa* und *Toda* [1990] – als konstant anzunehmen und exogen einzuführen sind. Es kann aber auch nicht übersehen werden, daß die numerischen Werte dicht bei Null liegen, und dies um so deutlicher, je höher die Anzahl der Ecken der Marktfigur ist. Der Hotelling/Smithies-Wettbewerb mit $\phi_{HS} = 0$ stellt somit eine geeignete Approximation der konsistenten Lösung dar, was aber auch gleichzeitig bedeutet, daß sowohl der Lösch-Wettbewerb mit $\phi_L = 1$ als auch der Greenhut/Ohta-Wettbewerb mit $\phi_{GO} = -1$ von ungeeigneten Annahmen ausgehen. Der Übergang von exogenen zu endogenen Reaktionskoeffizienten läßt somit zwei wichtige Neuerungen erkennen: (1) Da jeder der möglichen exogenen Reaktionskoeffizienten $-1 \leq \phi < 1$, der die Gewinnmaximierungsbedingung 2. Ordnung erfüllt, ein eigenes Modell begründet, reduziert sich unter sonst gleichen Bedingungen bei Verwendung der endogenen, d.h. konsistenten konjekturalen Reaktionskoeffizienten die Vielzahl der möglichen Ansätze auf genau *ein* Modell. (2) Die konsistenten Reaktionskoeffizienten werden bestimmt durch die Struktur des Modells, insbesondere durch die Kosten- und Nachfragefunktion, durch die Standortverteilung und durch die Figur des Marktgebietes. Damit ist wiederum die Möglichkeit zu einer Auffächerung der Modelle gegeben.

V. Marktergebnisse und Wohlfahrtseffekte

In diesem Abschnitt sollen einige Marktergebnisse des konsistenten Modells diskutiert werden. Bis zum Erscheinen des Beitrags von *Ishikawa* und *Toda* [1990] bestand in der Literatur Konsens darüber, daß das Sechseck – jedenfalls bei Lösch-Wettbewerb – sowohl die gewinnmaximale als auch wohlfahrtsmaximale Figur für zweidimensionale Marktgebiete darstelle. Die beiden Autoren haben gezeigt, daß die Verwendung alternativer exogener Verhaltensannahmen ($\phi_L = 1$, $\phi_{HS} = 0$ und $\phi_{GO} = -1$) die bislang gültigen Rangfolgen der Marktfiguren hinsichtlich der Gewinne ($\Pi_6 > \Pi_4 > \Pi_3$) und Wohlfahrtseffekte ($\Omega_6 > \Omega_4 > \Omega_3$) nicht unverändert läßt[2]: (1) Die Annahme des GO-Wettbewerbs führt zu einer Umkehr der Rangfolge bei Gewinnen, und die Annahme des HS-Wettbewerbs zeigt das gleiche Ergebnis zwischen $R = 0$ und einem Schnittpunkt der Gewinnkurven. (2) Unter Verwendung des GO-Wettbewerbs weist das Sechseck bei geringen Marktgebietsausdehnungen die niedrigsten Wohlfahrtseffekte und bei großen Radien die höchsten Wohlfahrtseffekte auf. Es ist nunmehr zu prüfen, ob diese neuen Resultate auch unter konsistenten Reaktionskoeffizienten gültig bleiben.

[2] *Ishikawa* und *Toda* [1990] untersuchen nicht die Entwicklung der kurzfristigen Gewinne in Abhängigkeit von der Marktgebietsausdehnung, sondern betrachten in der langfristigen Null-Gewinn-Situation den Zusammenhang zwischen der Höhe der Fixkosten und der Anzahl der Anbieter. Aus diesen Resultaten kann allerdings auf die Rangfolge der kurzfristigen Gewinne bei gegebenen Fixkosten geschlossen werden.

In Abbildung 2 sind zunächst die Ab-Werk-Preise in Abhängigkeit von der Marktgebietsausdehnung dargestellt. Wie in Kenntnis der konsistenten Reaktionskoeffizienten nicht anders erwartet werden kann, ist die Differenz zwischen HS-Preisen und modellkonsistenten Preisen nicht sehr groß. Die Maxima der konsistenten Preislinien liegen bei $R_3^{m\,\max} = 0{,}335$, $R_4^{m\,\max} = 0{,}433$ und $R_6^{m\,\max} = 0{,}497$. In Abbildung 3 sind die Gewinnfunktionen bei konsistenten Reaktionen und alternativen Marktfiguren zusammengefaßt, wobei die gezeichneten Kurven die hinsichtlich ihrer numerischen Größe nicht näher bestimmten Fixkosten enthalten ($\hat{\Pi}_n = \Pi_n + K_f$). Für die Marktausdehnung zwischen $R = 0$ und $R = 0{,}323$ zeigt sich eine Rangfolge der Gewinne von $\hat{\Pi}_3 > \hat{\Pi}_4 > \hat{\Pi}_6$, jenseits von $R = 0{,}323$ kehrt sich diese Rangfolge um.

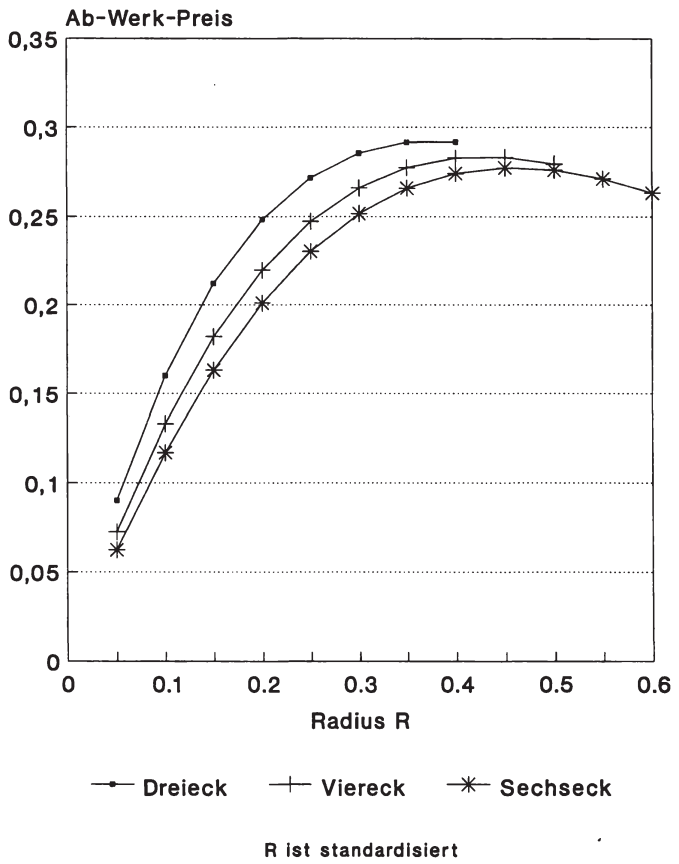


Abbildung 2: Konsistente Ab-Werk-Preise

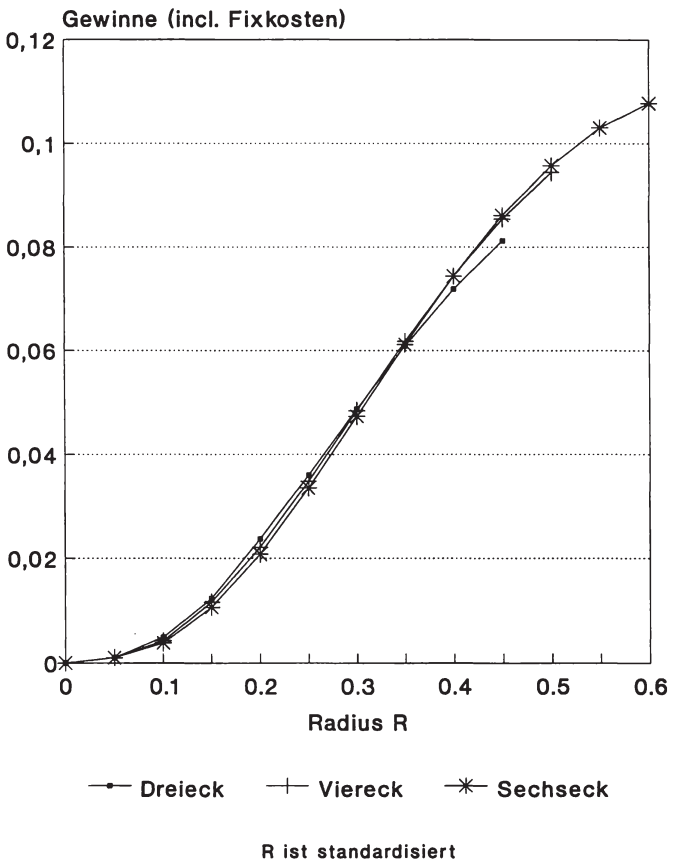


Abbildung 3: Gewinne und Marktfiguren

Neben den firmenindividuellen Marktergebnissen sollen die sozialökonomischen Wirkungen alternativer Marktgebietsformen – und damit Standortverteilungen – untersucht werden. Die Wohlfahrtseffekte werden als Summe aus Produzentenrenten und Konsumentenrenten definiert. Die über ein Firmenmarktgebiet aggregierte Konsumentenrente lautet unter Verwendung der konsistenten Reaktionskoeffizienten und der sich daraus ergebenden Ab-Werk-Preise:

$$C_i(\phi_n^*, R) = 2n \int_0^{\pi/n} \int_0^{R/\cos_n \theta_n} 0{,}5r\,(1 - m_{i,n}^* - r)\,dr\,d\theta_n \tag{15}$$

mit π und θ_n wie in Gleichung (4) definiert. Aus Gleichung (15) ergibt sich schließlich:

$$(16) \qquad C_i(\phi_n^*, R) = n[\alpha_n(1 - m_{i,n}^*)^2 R^2 - 2\beta_n(1 - m_{i,n}^*)R^3 + \gamma_n R^4]$$

mit α_n und β_n, wie in Gleichung (6) definiert, und mit $\gamma_3 = 0{,}866025$, $\gamma_4 = \frac{1}{3}$ und $\gamma_6 = 0{,}160375$. Die Ergebnisse für Konsumentenrente und Wohlfahrtseffekte sind in den Abbildungen 4 und 5 festgehalten. Die Figur des regelmäßigen Sechsecks läßt die höchsten Konsumentenrenten bei jeder zulässigen Ausdehnung des Marktgebietes im zweidimensionalen Raum entstehen. Dieses Resultat bleibt uneingeschränkt auch für die Summe aus Gewinnen und Konsumentenrenten erhalten; das Sechseck läßt die höchsten Wohlfahrtseffekte bei allen Marktausdehnungen entstehen. (In Abbildung 5 enthalten die

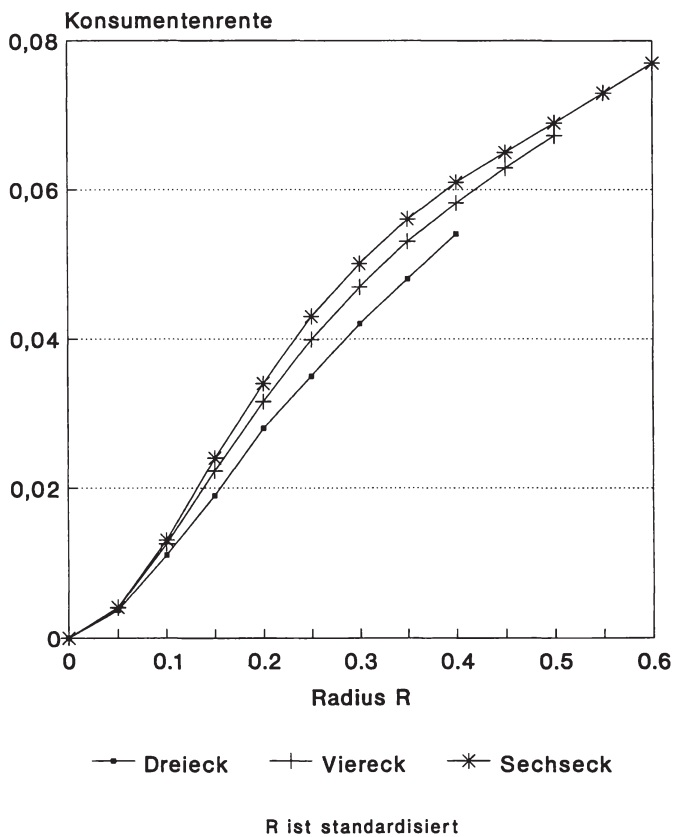


Abbildung 4: Konsumentenrente und Marktfiguren

Kurven einen numerisch nicht näher quantifizierten Fixkostenblock.) Eine sozialökonomische Inferiorität der Marktgebietsfigur des Sechsecks ist in einem Modell mit konsistenten Verhaltensannahmen nicht nachweisbar.

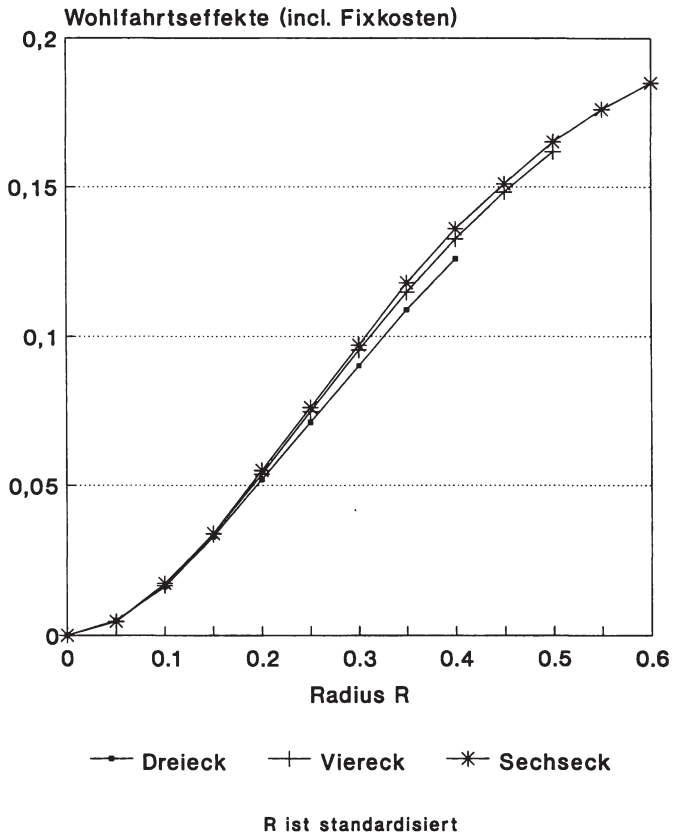


Abbildung 5: Wohlfahrtseffekte und Marktfiguren

VI. Zusammenfassung

Das vorgestellte Modell des räumlichen Wettbewerbs weist gegenüber den traditionellen Ansätzen eine Reihe neuer Elemente auf. Zunächst werden die üblicherweise exogen eingeführten konjekturalen Reaktionen durch endogene, d.h. konsistente Reaktionen ersetzt. Damit reduziert sich die Vielfalt der Lösungen auf genau *eine* Lösung. Die wichtigsten Ergebnisse können wie folgt zusammengefaßt werden: (1) An die Stelle der drei Wettbewerbsmo-

delle von *Ishikawa* und *Toda* [1990], die von traditionellen exogenen Verhaltensannahmen ausgehen, tritt ein Modell, daß eine eindeutige wohlfahrtstheoretische Beurteilung alternativer Marktfiguren erlaubt; das Sechseck führt zu den höchsten Wohlfahrtseffekten bei allen Ausdehnungen des Marktgebietes, die theoretisch zulässig sind. (2) Zwar kann im zweidimensionalen Raum – insbesondere bei hexagonalen Marktgebieten – der HS-Wettbewerb als eine hinreichende Approximation des konsistenten Modells und seiner Ergebnisse angesehen werden, jedoch gibt es – nach meiner Auffassung – keinen Anlaß für die Verwendung von Näherungslösungen, da die konsistenten Resultate ermittelt werden können und vorliegen.

Literatur

Boyer, M./*Moreaux,* M. (1984): Equilibres de duopole et variations conjecturales rationelles, in: Canadian Journal of Economics, Bd. 17, 111 - 125.

Bresnahan, T. (1981): Duopoly Models with Consistent Conjectural Variations, in: American Economic Review, Bd. 71, 934 - 945.

Capozza, D. R./*Van Order,* R. (1985): Spatial Competition with Consistent Conjectures, Discussion Paper, University of British Columbia, Vancouver.

Capozza, D. R./*Van Order,* R. (1989): Spatial Competition with Consistent Conjectures, in: Journal of Regional Science, Bd. 29, 1 - 13.

DeCanio, S. J. (1984): Delivered Pricing and Multiple Basing Point Equilibria: A Reevaluation, in: Quarterly Journal of Economics, Bd. 99, 329 - 349.

Eaton, B. C./*Lipsey,* R. (1976): The Non-Uniqueness of Equilibrium in the Löschian Location Model, in: American Economic Review, Bd. 66, 77 - 93.

Greenhut, M. L./*Ohta,* H. (1973): Spatial Configurations and Competitive Equilibrium, in: Weltwirtschaftliches Archiv, Bd. 109, 87 - 104.

Hartwick, J. M. (1973): Lösch's Theorem on Hexagonal Market Areas, in: Journal of Regional Science, Bd. 13, 213 - 222.

Holahan, W. L. (1975): The Welfare Effects of Spatial Price Discrimination, in: American Economic Review, Bd. 65, 498 - 503.

Ishikawa, T./*Toda,* M. (1990): Spatial Configurations, Competition and Welfare, in: Annals of Regional Science, Bd. 24, 1 - 12.

Kamien, M. I./*Schwartz,* N. L. (1983): Conjectural Variations, in: Canadian Journal of Economics, Bd. 16, 191 - 211.

Lösch, A. (1944): Die räumliche Ordnung der Wirtschaft, 2. Aufl., Jena.

Mills, E. S./*Lav,* M. R. (1964): A Model of Market Areas with Free Entry, in: Journal of Political Economy, Bd. 72, 278 - 288.

Schöler, K. (1990): Heterogene Güter und alternative Preistechniken in einem räumlichen Oligopol, in: Jahrbücher für Nationalökonomie und Statistik, Bd. 207, 143 - 156.

— (1991): A Note on „Price Variation in Spatial Markets: The Case of Perfectly Inelastic Demand", in: Annals of Regional Science, Bd. 25, S. 55 - 62.

— (1992): Konsistente konjekturale Reaktionen und Marktstrukturen im räumlichen Oligopol, in: Jahrbuch für Sozialwissenschaft, Bd. 43, S. 402 - 415.

Schöler, K./*Schlemper,* M. (1991): Räumlicher Wettbewerb bei konsistenten konjekturalen Reaktionen, in: Jahrbücher für Nationalökonomie und Statistik, Bd. 208, S. 449 - 458.

Sachverzeichnis