



Universität Potsdam

Sascha Frohwerk

Asymmetrien in der Neuen Ökonomischen Geographie

Modelle, Simulationsmethoden und
wirtschaftspolitische Diskussion

Potsdamer Schriften zur Raumwirtschaft | 3
Prof. Dr. Klaus Schöler (Hrsg.)

Potsdamer Schriften zur Raumwirtschaft
Prof. Dr. Klaus Schöler (Hrsg.)

Potsdamer Schriften zur Raumwirtschaft | 3
Prof. Dr. Klaus Schöler (Hrsg.)

Sascha Frohwerk

Asymmetrien in der Neuen Ökonomischen Geographie

Modelle, Simulationsmethoden und
wirtschaftspolitische Diskussion

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Universitätsverlag Potsdam 2011

<http://info.ub.uni-potsdam.de/verlag.htm>

Am Neuen Palais 10, 14469 Potsdam

Tel.: +49 (0)331 977 4623 / Fax: 3474

E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Potsdamer Schriften zur Raumwirtschaft** wird herausgegeben von Prof. Dr. Klaus Schöler.

ISSN (print) 2190-8702

ISSN (online) 2190-8710

Zugl.: Potsdam, Univ., Diss., 2010

Das Manuskript ist urheberrechtlich geschützt.

Online veröffentlicht auf dem Publikationsserver der
Universität Potsdam:

URL <http://pub.ub.uni-potsdam.de/volltexte/2011/4915/>

URN <urn:nbn:de:kobv:517-opus-49158>

<http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-49158>

Zugleich gedruckt erschienen im Universitätsverlag Potsdam:

ISBN 978-3-86956-089-2

Vorwort

Die vorliegende Arbeit wurde im Oktober 2010 von der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität Potsdam als Dissertation angenommen. Sie entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Volkswirtschaftslehre, insbesondere Wirtschaftstheorie.

Danken möchte ich zunächst meinem Doktorvater Prof. Dr. Klaus Schöler, der durch seine Forschung auf dem Gebiet der Neuen Ökonomischen Geographie einen entscheidenden Anstoß zu dieser Arbeit gegeben hat. Das Zweitgutachten hat Prof. Dr. Malcolm Dunn übernommen, dem ich für die Bereitschaft, sich kritisch mit dieser Thematik zu beschäftigen, danken möchte.

Für viele hilfreiche Hinweise, Korrekturen und Verbesserungsvorschläge möchte ich denen danken, die jeweils einen Teil des Manuskripts kritisch gelesen haben: Alexander Baumgärtner, Arne Beck, Ursula Frohwerk, Christiane Nordus, Sven Paasch sowie Florian Wetzig.

Wesentliche Vorarbeiten konnte ich im Rahmen der Winterseminare der Gesellschaft für Regionalforschung vorstellen und diskutieren. Diese Seminare haben wichtige Impulse für mein Fortkommen geliefert. Allen Teilnehmern gilt hierfür mein Dank. Abschließend möchte ich noch zwei Personen erwähnen, die mit dieser Arbeit nicht direkt in Verbindung stehen, die mich aber während und nach meinem Studium darin bestärkt haben, mich weiter mit wissenschaftlichen Fragen zu beschäftigen. Dies sind Prof. Dr. Gerhard Prosi sowie Prof. Toker Doganoglu. Ohne sie hätte ich vermutlich nie begonnen, eine Dissertation anzufertigen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Fragestellung	1
1.2	Stand der Forschung	3
1.3	Aufbau	4
2	Grundlagen der Neuen Ökonomischen Geographie	7
2.1	Grundidee und Modellaufbau	7
2.2	Regionen	10
2.3	Agglomeration, Konzentration und Spezialisierung	11
2.4	Skalenerträge	13
2.5	Transportkosten	15
2.6	Inter- und intraindustrieller Handel	18
2.7	Faktormobilität	20
2.8	Monopolistische Konkurrenz	23
2.9	Das Core-Periphery-Modell	32
2.10	Kritikpunkte	45
3	Modelle für interregionale Agglomeration	53
3.1	Grundmodell	54
3.2	Symmetrische Regionen mit lokalen Präferenzen	56
3.3	Asymmetrische Regionen ohne lokale Präferenzen	65
3.4	Asymmetrische Regionen mit lokalen Präferenzen	70
3.5	Zusammenfassung	74
4	Modelle für internationale Agglomeration	75
4.1	Modell mit Zwischenprodukten	77
4.2	Zwischenprodukte bei unterschiedlich großen Ländern	90
4.3	Konzentration und Spezialisierung von Ländern	93

4.4	Spezialisierung bei sektorspezifischen Transportkosten	106
4.5	Spezialisierung bei unterschiedlich großen Ländern	111
4.6	Spezialisierung bei unterschiedlich großen Sektoren	115
4.7	Spezialisierung bei verschiedenen Asymmetrien	117
4.8	Gleichgewichte und Löhne	118
4.9	Zusammenfassung	119
5	Simulationsmethoden für die Neue Ökonomische Geographie	121
5.1	Einleitung	121
5.2	Computersimulation	123
5.3	Lösung von NEG-Modellen	131
5.3.1	Erster Schritt: Gleichgewicht bei konstanten Transportkosten	132
5.3.2	Zweiter Schritt: Gleichgewichte bei Änderung der Transportkosten	135
5.4	Zusammenfassung	158
6	Wirtschaftspolitische Implikationen	159
6.1	Neue Ökonomische Geographie und Regionalpolitik	160
6.2	Regionale Entwicklung in Deutschland nach 1989	161
6.3	Das empirische Bild	163
6.4	Modell	168
6.5	Wirtschaftspolitische Konsequenzen	174
7	Fazit	181
	Literaturverzeichnis	189
	Anhänge	197

Abbildungsverzeichnis

2.1 Ursachen für Agglomeration (in Anlehnung an Pflüger (2008)) . . .	8
2.2 Agglomeration, Konzentration und Spezialisierung	12
2.3 Entwicklung der Kosten für Luft- und Seefracht (World Bank (2009), S. 179)	17
2.4 Handel mit Zwischenprodukten 1962 - 2006 (World Bank (2009), S. 171)	21
2.5 Tangentenlösung bei Monopolistischer Konkurrenz	25
2.6 Schematischer Aufbau des CP-Modells (eigene Darstellung in Anlehnung an Brakman et al. (2001), S.63)	33
2.7 Reallohndifferenz bei $T = 2,1$	40
2.8 Reallohndifferenz bei $T = 1,5$	41
2.9 Reallohndifferenz bei $T = 1,7$	42
2.10 Gleichgewichte in Abhängigkeit vom Transportkostensatz	43
3.1 Diagramme des CP-Modells	55
3.2 Verteilung der Präferenzen für Region 1 bei unterschiedlichen Werten für die Standardabweichung	59
3.3 Arbeitsangebot bei unterschiedlichen Werten für die Standardabweichung	60
3.4 Gleichgewichte im LW-Modell bei $T = 1,7; s = 0,001$	62
3.5 Gleichgewichte bei $T = 1,5$	63
3.6 Gleichgewichte bei $T = 2,1$	64
3.7 Gleichgewichte bei $T = 1,7$	65
3.8 Gleichgewichte bei $T = 1,45$ und $s = 0,01$	66
3.9 Bifurkationsdiagramm bei $s = 0,01$ und $s = 0,005$	67
3.10 Gleichgewichte bei $\varepsilon = 1,05$ und $s = 0$	68
3.11 CP-Modell mit $\varepsilon = 1,05$	69
3.12 LW-Modell bei $s = 0,01$ und $\varepsilon = 1,05$	72

3.13 Lohnverhältnis bei $s = 0,01$, $\varepsilon = 1,05$ und $T = 1,5$	73
4.1 Inlandslinie im IM-Modell bei $w_h = 1$ und $\mu = 0,4$; $\alpha = 0,5$; $\sigma = 5$	84
4.2 IM-Modell bei extrem hohen Transportkosten ($T = 1.000$)	85
4.3 IM-Modell bei hohen Transportkosten ($T = 3,0$)	86
4.4 IM-Modell bei niedrigen Transportkosten ($T = 1,5$)	87
4.5 IM-Modell bei mittleren Transportkosten ($T = 2,15$)	88
4.6 Bifurkationsdiagramm mit Zwischenprodukten	89
4.7 In- und Auslandsleichgewichte bei unterschiedlich großen Ländern ($\varepsilon_h = 1,05$)	91
4.8 IM-Modell bei hohen Transportkosten ($T = 3$) und unterschiedlich großen Ländern (hier für das größere Land dargestellt)	93
4.9 Schematischer Aufbau des Clustermodells (ein Land)	94
4.10 Kurzfristiges Gleichgewicht im In- und Ausland	101
4.11 Hohe Transportkosten ($\sigma = 3$; $\alpha = 0,4$; $\nu = 0,1$; $T = 3,0$)	102
4.12 Niedrige Transportkosten ($\sigma = 3$; $\alpha = 0,4$; $\nu = 0,1$; $T = 2,0$)	104
4.13 Mittlere Transportkosten ($\sigma = 3$; $\alpha = 0,4$; $\nu = 0,1$; $T = 2,35$)	105
4.14 Bifurkationsdiagramm des symmetrischen Cluster-Modells	106
4.15 Gleichgewichte bei verschiedenen Transportkostensätzen	108
4.16 Gleichgewichte und Transportkosten	111
4.17 Gleichgewichtskurven für unterschiedlich große Länder	113
4.18 Asym. Länder, mittlere Transportkosten ($\varepsilon = 0,025$, $T = 2,4$)	114
4.19 Großes und kleines Land im Cluster-Modell	115
4.20 Das Cluster-Modell bei unterschiedlich großen Sektoren	116
4.21 Unterschiedlich große Länder und Sektoren ($\varepsilon = 0,025$; $\mu = 0,52$)	118
4.22 Gleichgewichte mit Spezialisierung bei $L_f = 1$	119
5.1 Kurzfristiges Gleichgewicht des Standard-Modells	134
5.2 Gleichgewichte in Abhängigkeit vom Transportkostensatz	135
5.3 Nicht-Stetigkeit	137
5.4 Approximierte und tatsächliche Werte der Simulation	141
5.5 Fehler im LW-Modell	142
5.6 Plot des CP-Modells für $T = (1,6; 1,7; 1,8; 1,9)$	143
5.7 Gleichgewichte bei verschiedenen Transportkostensätzen	145
5.8 Beispiel zur Approximierung mittels Durchschnitt	146
5.9 Invertieren der Wertepaare mit Approximation	146

5.10	Bifurkationsdiagramm des LW-Modells bei Verwendung von Splines	154
5.11	Konstruktion des Bifurkationsdiagramms für zwei Länder	155
5.12	Mögliche Szenarien bei $L_f = 1$	156
6.1	Bevölkerungsentwicklung alte und neue Länder (aggregiert, ohne Berlin)	164
6.2	BIP absolut und BIP pro Kopf 2007 für Bundesländer (Daten siehe Anhang)	165
6.3	Durchschnittliche Bruttomonatsverdienste im produzierenden Gewerbe und im Dienstleistungsbereich ohne Sonderzahlungen (Quelle: Statistisches Bundesamt)	166
6.4	Bruttolohn und Wanderung (Daten siehe Anhang)	167
6.5	Bifurkationsdiagramm bei $s = 0,01$ und $s = 0,005$	169
6.6	CP-Modell bei $\varepsilon = 1,5$	170
6.7	LW-Modell bei $s = 0,01$ und $\varepsilon = 1,05$	171
6.8	LW-Modell bei $T = 1,5$ und $\varepsilon = 1,05$	172
6.9	Dauerhafte Lohnsatzdifferenz bei $s = 0,01$, $\varepsilon = 1,05$ und $T = 1,5$.	173
6.10	Unterschiede in der Attraktivität der Regionen ($\Omega = 0,501$, $T = 1,5$ und $\varepsilon = 1,05$	175
6.11	Unterschiede in der Attraktivität der Regionen ($\Omega = 0,499$, $T = 1,5$ und $\varepsilon = 1,05$	177
7.1	Schema der Simulation kurzfristiger Gleichgewichte	181

Symbolverzeichnis

A	Agrarsektor (als Referenz)
E	Ausgaben
L	ges. Arbeitskräfte im Industriesektor (oder i)
M	Aggregat der industriellen Produkte
Q	Preisindex
R	Anzahl der Regionen
T	Transportkosten
Y	Einkommen eines Landes
f	Index für Auslandsvariablen
$g(\cdot)$	Dichtefunktion der lokalen Präferenzen
h	Index für Inlandsvariablen
l	Arbeitsnachfrage pro Unternehmen
m	Nachfragemenge einer Variante
p	Preis
r	Index für die Region
s	Standardabweichung der lokalen Präferenzen
w	Nominallohn
x	Angebotsmenge einer Variante
α	Parameter der Produktionsfunktion für den eigenen Sektor (nur CM + IM)
β	Parameter der Produktionsfunktion für Arbeit
ε	Größenunterschied für Länder oder Regionen
λ	Aufteilung der Arbeitskräfte
μ	Exponent der Nutzenfunktion für Industriegüter
ν	Parameter der Produktionsfunktion für den anderen Sektor (nur CM)
π	Gewinn der Unternehmen
ρ	Parameter für die Homogenität von Gütern
σ	Substitutionselastizität

Symbolverzeichnis

$\varphi_i, \bar{\varphi}$	Gewichtungsfaktor eines Arbeiters / des Grenzarbeiter im LW-Modell
$\phi(\cdot)$	Dichte der Standard-Normalverteilung
$\Phi(\cdot)$	kumulative Verteilungsfunktion der Standard-Normalverteilung
Ω	Mittelwert der lokalen Präferenzen
ω_i	Reallohn im Land i
ω	Reallohnverhältnis ω_1/ω_2

1 Einleitung

1.1 Fragestellung

Die *Neue Ökonomische Geographie*¹ erklärt anhand von verschiedenen Modellen, unter welchen Bedingungen sich Agglomerationen bilden. Dabei wird auf exogene Unterschiede bewusst verzichtet und die Agglomerationsbildung aus einem mikroökonomischen Totalmodell abgeleitet. Die Literatur zur Neuen Ökonomischen Geographie befasst sich dabei fast ausschließlich mit symmetrischen Modellen. Darunter ist zu verstehen, dass beide Länder oder Regionen gleich groß sind, die Ausgabenanteile für beide Sektoren (sofern zwei Sektoren betrachtet werden) gleich sind und für alle industriellen Produkte die gleichen Transportkosten anfallen.² Dies ist eine Vereinfachung, wie sie in der Wirtschaftstheorie üblich ist. Der Zweck dieser Vereinfachung ist es, die Modelle auf das Wesentliche zu reduzieren um möglichst allgemeine Ergebnisse zu erhalten.

Aus verschiedenen Gründen ist es jedoch interessant zu untersuchen, welche Erkenntnisse sich gewinnen lassen, wenn man diese Symmetrieanahmen aufgibt. Zum einen stellt sich die Frage, inwieweit die Modelle dann überhaupt noch lösbar sind. Die Theorien der NEG sind sehr komplex und werden erst durch Vereinfachungen und Standardisierungen in eine Form gebracht, die eine relativ einfache Lösbarkeit durch Computersimulationen erlaubt. Daher führen Veränderungen an ihnen schnell dazu, dass Gleichgewichte nicht mehr bestimmt werden können. Zum anderen ist zu untersuchen, ob die Ergebnisse der Basis-Modelle erhalten bleiben, wenn man einzelne Symmetrieanahmen aufgibt und ob sich daraus zusätzliche Erkenntnisse gewinnen lassen. Beispielsweise lassen fast alle Modelle die Frage offen, in welchem Land oder in welcher Region sich eine Agglomeration bildet, wenn dies geschieht. Es ist offensichtlich, dass dies eine Folge davon

¹Im englischen *New Economic Geography* und daher in dieser Arbeit auch als NEG abgekürzt.

²In der Regel wird von zwei Ländern ausgegangen.

ist, dass hinsichtlich der Regionen und der Sektoren stets angenommen wird, dass diese identisch sind. Daher liegt die Vermutung nahe, dass sich konkretere Ergebnisse erzielen lassen, wenn man diese Annahmen ändert.

Aus wirtschaftspolitischer Perspektive ist diese Fragestellung interessant, weil sich die vereinfachenden Annahmen in der Regel empirisch nicht nachweisen lassen. Außerdem ist die Erkenntnis, dass man zwar erklären kann, dass sich eine Agglomeration herausbildet, aber nicht wo dies geschieht, aus politischer Sicht unbefriedigend. Ohne diese Prognose ist es oft schwierig, die richtigen regionalpolitischen Maßnahmen rechtzeitig zu planen.

Aus diesen Gründen soll in der vorliegenden Arbeit der Frage nachgegangen werden, wie sich die Ergebnisse der jeweiligen Basis-Modelle ändern, wenn Asymmetrien angenommen werden.³

Grundsätzlich können zwei Typen von Modellen der NEG unterschieden werden: Solche, bei denen die Agglomeration durch Wanderung von Produktionsfaktoren begründet wird, und jene, in denen die Faktoren immobil sind und Zwischenprodukte diese Funktion übernehmen.⁴ Der erste Typ von Modellen wird verwendet, wenn sich die Fragestellung auf Regionen eines Landes bezieht und daher von relativ hoher Mobilität ausgegangen werden kann. Der zweite Typ von Modellen geht von Staaten aus. Hier wird die Mobilität als gering eingeschätzt.

Aufgrund der in den letzten Jahren stark gestiegenen Anzahl von Modelle beider Typen, die aber dennoch ähnliche Strukturen aufweisen, sollen im folgenden die theoretisch einflussreichsten Modelle verwendet werden, um die Auswirkungen von verschiedenen Asymmetrien zu untersuchen.⁵ Dazu wird jeweils das Grundmodell dargestellt und im Anschluss die Asymmetrie modelliert. Unter Verwendung der in Kapitel 5 dargestellten Methoden werden die Ergebnisse abgeleitet, die später in Kapitel 6 auf wirtschaftspolitische Fragen angewandt werden.

In der Entwicklung geeigneter Simulationsmethoden liegt ein weiterer Schwerpunkt dieser Arbeit. Denn obwohl die meisten Modelle der NEG nicht analytisch lösbar sind, fehlen technische Aspekte der Simulation nahezu vollständig in der Literatur.

³Unter *Basis-Modell* wird in dieser Arbeit jeweils das bereits in der Literatur vorhandene symmetrische Modell verstanden.

⁴Vgl. auch Knaap (2004), S. 28.

⁵Einen Hinweis darauf, welche Artikel den größten Einfluss auf die Literatur hatten, gibt der *Social Citation Index*, der zählt, wie oft ein Beitrag zitiert wurde.

“[...] an important but often neglected aspect of the geographical economic approach: computer simulations” [Brakman et al. (2001), S. xix]

In Kapitel 5 werden daher verschiedene Aspekte der Simulation unter Verwendung der Software *Scilab* diskutiert. Auf dieser Basis kann für jedes Modell eine entsprechende Simulation programmiert werden. Dieses Verfahren ist transparent und kann von jedem nachvollzogen werden (keine “Black Box”). Außerdem lassen sich Veränderungen in den Annahmen - auch solche, die über reine Parameteränderungen hinausgehen - in der Simulation berücksichtigen.

1.2 Stand der Forschung

Die in dieser Arbeit entwickelten Modellvarianten basieren alle auf den grundlegenden Modellen der NEG. Für regionale Agglomeration ist dies das Core-Periphery-Modell von Krugman (1991a) mit der Erweiterung von Ludema und Wootton (1997). Betrachtet man internationale Agglomeration, werden die Modelle von Krugman und Venables (1995) und Krugman und Venables (1996) als Basis verwendet. Auf analytisch lösbare Modelle wurde bewusst verzichtet, da diese Lösbarkeit durch Vereinfachung erreicht wird. Damit sind die Möglichkeiten der Veränderung dieser Modelle stark eingeschränkt, möchte man auch die Varianten analytisch lösen. Hier erlaubt die Simulation deutlich mehr Freiheiten.

In den grundlegenden Modellen der NEG wird fast ausschließlich Symmetrie unterstellt. Entsprechend übersichtlich ist die Literatur zu Annahmen der Asymmetrie. Baldwin et al. (2003) nennen zumindest die Möglichkeit, dass das CP-Modell auch für unterschiedlich große Länder modelliert werden könnte. In den angrenzenden Disziplinen, die teilweise auch auf der Marktform der monopolistischen Konkurrenz aufbauen, werden Asymmetrien in verschiedenen Fällen behandelt. In der klassischen Außenhandelstheorie etwa, die man im Zusammenhang mit den in Kapitel 4 dargestellten Modellen sehen kann, sind Unterschiede in der Landesgröße enthalten.⁶ Allerdings ist der Fokus dieser Modelle ein anderer. Statt Agglomeration erklären sie vorrangig das Entstehen von Handelsströmen. Cabrales und Motta (2001) untersuchen die Folgen von Handelsintegration bei asymmetrischen Ländern, verwenden dabei aber ein Oligopolmodell. Daher liegt der

⁶Vgl. z.B. Siebert (1997), Kap. 5.

1 Einleitung

Schwerpunkt dieser Untersuchung auf strategischem Verhalten. Hervorzuheben ist allerdings, dass Unternehmen im großen Land in diesem Modell ebenfalls einen Vorteil durch den großen Binnenmarkt haben. Diese Ergebnisse passen zu denen in Abschnitt 4.2 dieser Arbeit, obwohl der Modellrahmen sich deutlich unterscheidet. Tharakan und Thisse (2002) kommen in einem sehr ähnlichen Modell mit Cournot-Wettbewerb zu genau entgegengesetzten Aussagen.

Im Rahmen der Stadtökonomik, deren Modelle den hier vorgestellten relativ ähnlich sind, untersuchen Takatsuka und Zeng (2005) wie sich Agglomerationen bei sektoral unterschiedlichen Transportkosten entwickeln. Deren Modell basiert auf Ottaviano et al. (2002) und hat daher ähnliche Eigenschaften, so z.B. Pendelkosten und entfernungsabhängige, quadratische Transportkosten. Obwohl hier ein anderer Modellrahmen verwendet wird, kommt auch dieses Modell zu dem Ergebnis, dass der Sektor mit den höheren Transportkosten im größeren Land einen kleineren Anteil der Produktion hat als der Sektor mit den niedrigeren Transportkosten. Allerdings ist die unvollständige Spezialisierung, die in dieser Arbeit nur durch die angenommene Asymmetrie entsteht, bereits bei Ottaviano et al. (2002) enthalten. Das Modell von Takatsuka und Zeng (2005) bestimmt daher nur die Aufteilung der Industrie. Unvollständige Spezialisierung wird schließlich auch von Tabuchi und Thisse (2006) erklärt. Bei diesem Modell werden allerdings handelbare und nicht handelbare Güter unterstellt. Außerdem beinhaltet es Bodenpreise und kommt im Ergebnis zu einem Städtesystem.

1.3 Aufbau

Kapitel 2 dieser Arbeit gibt einen Überblick über empirische und theoretische Grundlagen der Neuen Ökonomischen Geographie. Damit werden zwei Ziele verfolgt: Zum einen sollen die verschiedenen Komponenten dieser Modellwelt (z.B. Transportkosten, Skalenerträge sowie Faktor- und Gütermobilität) definiert werden. Es wird jeweils auch ein kleiner Einblick in die empirische Forschung gegeben um zu verdeutlichen, welches Ausmaß die verschiedenen Größen in der Realität haben. Zum anderen finden sich hier die wichtigsten theoretischen Grundlagen. Dabei wird in Abschnitt 2.8 die Marktform der monopolistischen Konkurrenz diskutiert und die Nachfragefunktion sowie der Preisindex hergeleitet, da diese beiden Formeln in allen Modellen verwendet werden. Abschnitt 2.9 stellt

darauf aufbauend das Grundmodell von Krugman vor und diskutiert Kritikpunkte an dieser Theorie.

Die Kapitel 3 und 4 bilden den theoretischen Kern dieser Untersuchung. Für die oben genannten Modelle werden hier die verschiedenen symmetrischen und asymmetrischen Varianten formal dargestellt und durch Simulationen gelöst. In Kapitel 5 werden ausführlich verschiedene Aspekte der Computersimulation vorgestellt. Zunächst werden kurz die wichtigsten Elemente der Syntax erklärt um dann auf die Umsetzung der Simulationen verschiedener NEG-Modelle einzugehen. Wie bereits erwähnt, fehlt eine solche Darstellung in der Literatur. Daher ist dieses Kapitel die Grundlage für alle Simulationen dieser Arbeit.

In Kapitel 6 wird zunächst diskutiert, welche Konsequenzen die Neue Ökonomische Geographie für die Wirtschaftspolitik haben kann, bevor die Ergebnisse aus Kap. 3 beispielhaft auf die Entwicklung der deutschen Wirtschaft nach der Wiedervereinigung angewendet werden. Hier folgt nach einem kurzen empirischen Teil eine Anpassung des Modells aus Abschnitt 3.4 auf die konkrete Fragestellung. Abschließend werden das Ergebnis und mögliche Kritikpunkte diskutiert. Das letzte Kapitel zieht ein Fazit und beantwortet zusammenfassend die eingangs aufgeworfenen Fragen.

2 Grundlagen der Neuen Ökonomischen Geographie

2.1 Grundidee und Modellaufbau

Die Modelle der Neuen Ökonomischen Geographie⁷ lassen sich von klassischen Standortmodellen und anderen Agglomerationstheorien⁸ vor allem dadurch abgrenzen, dass sie die Ansiedlung von Unternehmen und Haushalten nicht durch externe Faktoren, sondern aus sich selbst heraus erklären. Sie gehören damit zu den so genannten *Second Nature* Ansätzen⁹. Unter *First Nature* versteht man dabei Standortfaktoren wie Bodenschätze, Klima, Zugang zu Flüssen oder Seehäfen, fruchtbarer Boden, usw.. *Second Nature* bedeutet, dass die Anhäufung von Unternehmen oder Nachfragern selbst einen Einfluss auf die Standortwahl weiterer Unternehmen oder Haushalte hat. Dadurch entstehen Zuzugsprozesse, die sich dann wieder selbst begründen und verstärken. Um solche Prozesse zu untersuchen, ist es sinnvoll, zunächst von exogenen Unterschieden abzusehen, auch wenn klar ist, dass "first Nature" die initiale Begründung für die Ansiedlung von Städten gegeben hat. Dies wird unmittelbar deutlich, wenn man historische Plätze für Siedlungen betrachtet. Andererseits erklären solche Unterschiede, z.B. der Zugang zur See, zwar kleinräumig den Ort von Ansiedlungen aber nicht unbedingt das unterschiedliche Wachstum von verschiedenen Standorten mit ähnlichen Voraussetzungen. Die gleiche Überlegung gilt auch großräumig: Klimaunterschiede mögen erklären, warum z.B. im Mittelmeerraum eine größere Bevölkerungsdichte

⁷Krugman selbst bezeichnet seine Theorie als *New Economic Geography*, während andere Autoren, z.B. Brakman et al. (2001), die Theorie als *Geographical Economics* bezeichnen. Dies kommt der Natur der Modelle meiner Ansicht nach näher, da es sich offensichtlich um ökonomische Modelle handelt. Zu dieser Diskussion siehe auch Kapitel 2.10.

⁸Vgl. Roos (2002).

⁹Vgl. z.B. Pflüger (2008).

herrscht als in Sibirien, die Entwicklungsunterschiede zwischen den verschiedenen mediterranen Ländern können aber nur unzureichend erklärt werden. Allerdings muss man wohl unterschiedliche politische und institutionelle Verhältnisse auch zur *First Nature* zählen, auch wenn dies in der Literatur in der Regel nicht getan wird. Dadurch kann schon ein Teil der Entwicklungsunterschiede von ansonsten ähnlichen Ländern (z.B. Italien und Kroatien) erklärt werden.

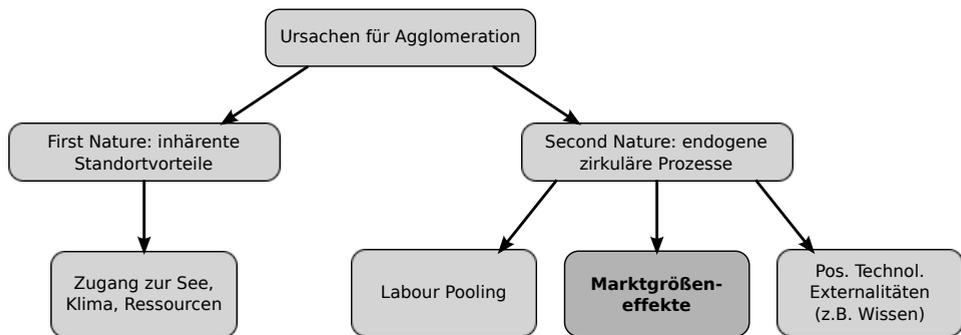


Abbildung 2.1: Ursachen für Agglomeration (in Anlehnung an Pflüger (2008))

Neben den hier thematisierten Ansätzen gehören vor allem Labour-Pooling Ansätze und Knowledge Spillovers zu den Modellen der *Second Nature* zur Erklärung von Agglomerationen. Beim Labour-Pooling geht man davon aus, dass in einem Gebiet, in dem sich viele Unternehmen einer Branche befinden, auch viele spezialisierte Arbeitskräfte angesiedelt sind. Eine Annahme, die nicht unrealistisch ist. Dies zieht weitere Unternehmen der Branche an, die wiederum Arbeitskräfte ausbilden. Diese Idee spiegelt sich z.T. auch in einigen Weiterentwicklungen der NEG wieder. Der Ansatz positiver externer Effekte ist dem des Labour Pooling ähnlich. Hier geht man davon aus, dass Unternehmen der gleichen Branche Wissen austauschen und daraus ein Standortfaktor erwächst. Dieser, gerade unter Geographen sehr beliebte Ansatz muss aus ökonomischer Sicht kritisch gesehen werden. Wissen ist ein Wettbewerbsfaktor, den man nicht verschenkt. Daher ist anzuzweifeln, dass Unternehmen wirklich bereit sind, Wissen zu teilen. Dies gilt insbesondere deshalb, weil Qualität und Quantität von Wissen schlecht messbar ist und daher der Austausch schwierig vertraglich zu regeln ist. Daher ist der einzige offensichtliche Austausch durch den Wechsel von den Trägern des Wissens, den Arbeitskräften, möglich.

Beide Ansätze, Labour-Pooling und Knowledge-Spillovers, sind im Unterschied

zur NEG oftmals nicht formalisiert und wenn überhaupt, dann nur in partialanalytischen Modellen. Genau hier liegt auch die Besonderheit der NEG: Sie erklärt Agglomerationen aus einem mikroökonomischen Totalmodell heraus. Neben den üblichen Verhaltensannahmen der Gewinn- und Nutzenmaximierung sind lediglich Annahmen bezüglich der Beschaffenheit der Güter (Substitutionselastizität) und der Produktionsbedingungen sowie der Transportkosten notwendig. Dabei gibt es drei Komponenten, die in allen Modellen anzutreffen sind, wenn auch in unterschiedlicher Ausprägung: Steigende Skalenerträge, Mobilität von Faktoren oder Gütern und Transportkosten. Diese Grundbausteine sollen in den folgenden Abschnitten ausführlicher diskutiert werden.

Auf die Einordnung der NEG bei den Second-Nature Ansätzen werden wir später noch zurückkommen: Wenn die NEG Agglomerationen bei Abwesenheit von inhärenten Standortunterschieden erklärt und daher von homogenen Regionen ausgeht, so können die Einsichten wirtschaftspolitisch auch nur angewendet werden, wenn die betrachteten Regionen sich nicht zu stark unterscheiden. So ist es z.B. ein Unterschied, ob man zwei Industrieländer mit hochqualifizierten Arbeitskräften aber relativ arm an Bodenschätzen miteinander vergleicht, oder etwa ein Industrieland und ein Entwicklungsland mit wenig Humankapital aber vielen Bodenschätzen.

Als Folge der genannten Annahmen entstehen Kräfte in dem Modell, die in Richtung einer Agglomeration des Kerngebietes und einer entsprechenden Entleerung der Peripherie wirken, und Kräfte, die diese Agglomerationstendenz begrenzen. Für eine Agglomeration wirken die so genannten *forward linkages* und *backward linkages*. Als *forward linkages* bezeichnet man den Zusammenhang, dass eine Ansiedlung von neuen Unternehmen an einem Standort den Reallohn steigen lässt und daher den Zuzug neuer Arbeitskräfte hervorruft (Vorteil auf dem Beschaffungsmarkt). Diese neuen Arbeitskräfte erhöhen die Nachfrage nach Gütern in der Region, was wiederum Unternehmen anzieht. Dies wird als *backward linkage* bezeichnet (Vorteil auf dem Absatzmarkt). Beide Kräfte verstärken sich gegenseitig und wirken so wie eine Spirale. Als Gegengewicht zu diesen Kopplungseffekten gibt es einen regional immobilen Faktor oder nicht transportierbare Güter.

2.2 Regionen

In den Modellen der Neuen Ökonomischen Geographie werden mehrere Regionen angenommen (in der Regel 2), da eben die Verteilung ökonomischer Aktivität Gegenstand dieser Theorien ist. Dabei wird die Region in Abhängigkeit von der im jeweiligen Modell behandelten Fragestellung unterschiedlich definiert. Es wird keine bestimmte Regionsgröße unterstellt. Grundsätzlich lassen sich Modelle unterteilen in solche, die die Verteilung ökonomischer Aktivität auf *Regionen* untersuchen, und solche, die die Verteilung auf *Länder* zum Gegenstand haben. Die Abgrenzung kann prinzipiell auf politischen Grenzen aufbauen, aber auch an sprachlichen, kulturellen oder handelsrechtlichen Kriterien orientiert sein¹⁰. Die Definition von Regionen an politischen Grenzen hat einen Vorteil bei der empirischen Forschung, da die Datenbeschaffung deutlich einfacher wird.

Im ökonomischen Kontext ist es sinnvoll, die Regionen an der Faktormobilität abzugrenzen. Dabei kann festgestellt werden, dass Kapital international sehr mobil ist. Diese Mobilität wird allenfalls durch instabile politische Systeme und schwankende Wechselkurse sowie das damit verbundene Inflationsrisiko begrenzt. Der Faktor Arbeit hingegen ist, vor allem in Europa, eher immobil. Da der Arbeitsmobilität in den Modellen der NEG eine Schlüsselrolle zukommt, erscheint es sinnvoll, Länder von Regionen dadurch abzugrenzen, ob der Faktor Arbeit über die Grenze mobil ist oder eher nicht.¹¹ Wir verwenden daher den Begriff Regionen, wenn Arbeitskräfte aufgrund von Lohnunterschieden den Standort wechseln und den Begriff Länder, wenn sie dies nicht tun. In Europa wäre es z.B. sinnvoll, Nationalstaaten als Länder einzuordnen und Bundesländer als Regionen.

Darüber hinaus ist es bei manchen Fragestellungen auch sinnvoll, supranationale Regionen zu definieren, Dabei könnte man z.B. die Länder Osteuropas zu einer Region zusammenfassen.

¹⁰Zur Definition von Regionen vgl. Maier und Tödtling (2002), S. 15ff.

¹¹Dabei ist es eine Vereinfachung, davon auszugehen, dass Arbeiter gar nicht das Land wechseln, denn in der Realität gibt es immer Wanderungsbewegungen. Diese sind jedoch zwischen den Ländern Europas sehr gering und auf bestimmte Gruppen von Arbeitnehmern konzentriert.

2.3 Agglomeration, Konzentration und Spezialisierung

Die Neue Ökonomische Geographie beschäftigt sich mit der Anhäufung von Unternehmen (oder Haushalten) in einer Region oder in einem Land, abhängig vom Maßstab der Betrachtung. Da in einigen Modellen mehrere Sektoren betrachtet werden, müssen die Begriffe Agglomeration, Konzentration und Spezialisierung genau unterschieden werden.¹²

Agglomeration: Von Agglomeration wird gesprochen, wenn es eine Anhäufung von Unternehmen in einer Region gibt, unabhängig von Sektoren, denen diese Unternehmen zugerechnet werden. Das Gegenteil von Agglomeration ist Peripherie oder Hinterland.

Konzentration: Wenn ein Großteil einer Branche in einer Region angesiedelt ist, ist diese Branche hier konzentriert. Dieser Begriff ist aus Sicht der Unternehmen definiert. Eine Branche ist konzentriert, nicht eine Region.

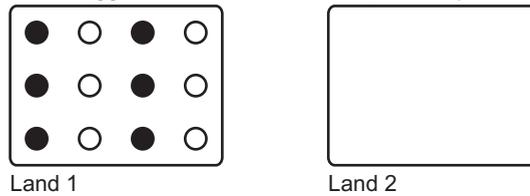
Spezialisierung: Eine Spezialisierung einer Region liegt vor, wenn viele Unternehmen in einer Region einer bestimmten Branche angehören. Dieser Begriff ist aus Sicht der Region definiert.

Diese Begriffe schließen sich weder gegenseitig aus noch ein. Abbildung 2.2 verdeutlicht einige Fälle schematisch. Schwarze und weiße Kreise stellen Unternehmen zweier Branchen dar. Fall a beschreibt die Situation, in der sich die gesamte Industrie in Land 1 angesiedelt hat. Hier liegt eine Agglomeration in Land 1 vor. Beide Branchen sind außerdem in Land 1 konzentriert. Eine Spezialisierung einer Region ist hier nicht zu beobachten. Im Fall b hingegen spezialisieren sich beide Regionen auf eine Branche und beide Branchen konzentrieren sich jeweils auf ein Land. Hier liegt aber keine Agglomeration vor. Fall c schließlich zeigt ein Bild mit geringer Agglomeration in Land 1. Eine Branche (schwarz) ist in Land 1 konzentriert. Die andere Branche ist nicht konzentriert, auch wenn Land 1 auf diese Branche (weiß) spezialisiert ist.

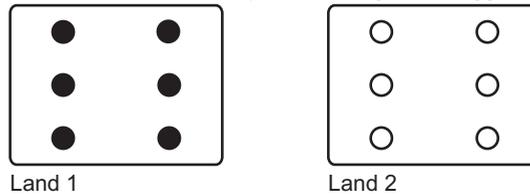
Der Grad der Konzentration für eine Branche zu einem Zeitpunkt über mehrere Regionen wird durch den regionalen GINI-Koeffizienten gemessen. Dieser wird

¹²Vgl. Brakman et al. (2001), S. 130f.

Fall a: Agglomeration, Konzentration, keine Spezialisierung



Fall b: Konzentration, Spezialisierung, keine Agglomeration



Fall c: Konzentration einer Branche (schwarz), Spezialisierung in Land 2, keine (vollständige) Agglomeration

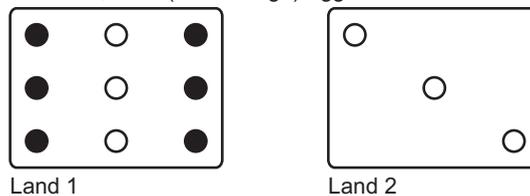


Abbildung 2.2: Agglomeration, Konzentration und Spezialisierung

als

$$GINI_{it} = \frac{N}{N-1} \cdot \left[\frac{2}{N} \cdot \frac{\sum_{r=1}^N (r \cdot s_{irt})}{\sum_{r=1}^N s_{irt}} - \frac{N+1}{N} \right] \quad \text{mit} \quad s_{irt} = \frac{x_{irt}}{\bar{x}_{it}} \quad (2.1)$$

berechnet¹³. Dabei ist N die Anzahl der Regionen, x_{it} die Anzahl der Beschäftigten in Region r zum Zeitpunkt t in Sektor i und \bar{x}_{it} die entsprechende Anzahl der Beschäftigten in allen Regionen. Zur Berechnung sind die s_{irt} aufsteigend zu sortieren. Der regionale GINI-Koeffizient nimmt den Wert 0 an, wenn die Branche in allen Regionen gleich stark vertreten ist und den Wert 1, wenn sie nur in einer Region konzentriert ist. Der Nachteil an diesem Konzentrationsmaß ist, dass die Struktur der Unternehmensgrößen nicht erfasst wird. Ob eine hohe Anzahl von Arbeitnehmern in einer Region durch ein einziges dort ansässiges Unternehmen

¹³Vgl. Haas und Südekum (2005) und Südekum (2006).

verursacht wird, oder ob es sich eine Agglomeration von kleinen und mittleren Unternehmen handelt, bleibt im Dunkeln.

Der Grad der Spezialisierung einer Region wird in der Literatur üblicherweise mit dem *Krugman Specialization Index KSI* gemessen¹⁴. Dieser ergibt sich aus den aufsummierten Abweichungen der regionalen Beschäftigungsanteile aller Branchen von den durchschnittlichen bundesweiten Anteilen. Der KSI liegt zwischen 0 und 2. Er nimmt den Wert 0 an, wenn die Industriestruktur der Region genau der im Gesamtgebiet entspricht und 2, wenn sie sich von dieser maximal unterscheidet.

$$KSI_{rt} = \sum_{i=1}^N |x_{irt} - \bar{x}_{it}| \quad (2.2)$$

Für Agglomeration im Allgemeinen gibt es kein spezielles Maß. Hier wird üblicherweise das regional BIP als Maß für wirtschaftliche Aktivität verwendet. Die hier genannten Kennzahlen sind keineswegs die einzig möglichen, sondern die am weitesten verbreitetsten. Eine detaillierte Diskussion zu Kennzahlen und Messmethoden findet sich z.B. in Bickenbach und Bode (2006) und Mare (2005).

2.4 Skalenerträge

Steigende Skalenerträge liegen vor, wenn die Durchschnittskosten mit steigender Outputmenge sinken.¹⁵ Man kann dabei zwischen internen und externen sowie zwischen technologischen und pekuniären Skalenerträgen unterscheiden.¹⁶ Als intern werden Skalenerträge bezeichnet, wenn die Durchschnittskosten mit steigendem Output des Unternehmens sinken. Im Fall externer Skalenerträge sinken sie mit steigender produzierter Menge der Branche, auch wenn die Menge des betrachteten Unternehmens konstant bleibt. Damit es zu solchen externen Economies of Scale kommen kann, müssen externe Effekte vorliegen (Spill-Overs)¹⁷.

¹⁴Vgl. Krugman (1991a).

¹⁵Vgl. Mas-Collel et al. (1995), S. 132 oder Shy (2000), S. 45f. Wobei im eigentlichen Sinne die sinkenden Durchschnittskosten dadurch verursacht werden, dass die Outputmenge bei einer Verdoppelung aller Inputs um mehr als das Doppelte steigt. Hier werden sie jedoch durch sehr hohe Fixkosten verursacht, was einer langfristigen Betrachtungsweise widerspricht.

¹⁶In der gängigen mikroökonomischen Literatur wird der Begriff meist im engeren Sinne nur für technologische interne Skalenerträge verwendet.

¹⁷Vgl. Tibor (1954).

Diese sind auch Grundlage einer anderen Klasse von Agglomerationsmodellen (vgl. Abb. 2.1). Technologische Skalenerträge liegen in der Produktionsfunktion der Unternehmen gegründet, die pekuniären hingegen laufen über Preise auf Märkten ab. Typisch sind hier Marktgrößeneffekte. In der Neuen Ökonomischen Geographie werden sowohl interne als auch externe Skaleneffekte verwendet. Die internen sind im Marktmodell der unvollkommenen Konkurrenz enthalten und stellen sicher, dass ein Unternehmen immer nur an einem Standort produziert.¹⁸ Pekuniäre externe Effekte werden in den Marktgrößenvorteilen auf dem Arbeitsmarkt (Standard-Modell) oder dem Markt für Zwischenprodukte (IM- und CM-Modell) deutlich. Sie begründen die Vorteile für ein Unternehmen, sich in der Agglomeration anzusiedeln.

Aus empirischen Studien ist zu erkennen, dass ca. 30% aller Unternehmen der Güterproduktion unter steigenden Skalenerträgen arbeiten.¹⁹ Dabei wurde unter Verwendung von Daten amerikanischer und chilenischer Unternehmen ein durchschnittlicher Homogenitätsgrad zwischen 1,2 und 1,4 ermittelt. Branchen mit hohen Skalenerträgen sind Maschinenbau, chemische Industrie, Raffinerien, Fahrzeugbau, Tabakwaren und Druck. Konstante oder nur leicht steigende Skalenerträge wurden für Holz, Lebensmittelproduktion, Lederwaren, Textilien und Schuhe festgestellt.²⁰ An dieser Aufstellung kann abgelesen werden, dass Skalenerträge in stark industrialisierten Branchen stärker sind als in Wirtschaftszweigen mit eher traditioneller Produktion oder Nähe zur Landwirtschaft. Eine Annahme, die auch in der Neuen Ökonomischen Geographie oft zu finden ist.

Aus theoriegeschichtlicher Sicht sind steigende Skalenerträge in ökonomischen Modellen seit den 70er Jahren des letzten Jahrhunderts zu finden. In der Industrieökonomik formalisiert von Spence (1976) und Dixit und Stiglitz (1977) wurden sie zu einem wichtigen Bestandteil der Stadtökonomik, der Außenwirtschaftstheorie²¹, der endogenen Wachstumstheorie²² und der Neuen Ökonomischen Geographie.

¹⁸Steigende Skalenerträge sind mit vollkommener Konkurrenz nicht vereinbar.

¹⁹Vgl. World Bank (2009), S. 129ff, sowie die dort angegebene Literatur.

²⁰Vgl. World Bank (2009), Tab. 4.2.

²¹Z.B. Ethier (1982), Krugman (1980), Krugman (1981).

²²Z.B. Romer (1986), Grossman und Helpman (1991) und Aghion und Howitt (2005).

2.5 Transportkosten

Als Transportkosten verstehen wir alle Kosten der Raumüberwindung. Die NEG geht davon aus, dass innerhalb einer Region keine Transportkosten anfallen, beim Überschreiten der Grenze zwischen zwei Regionen fällt der jeweilige Transportkostensatz in voller Höhe an. Dies ist eine vereinfachende Annahme und es muss an dieser Stelle berücksichtigt werden, wie in dem jeweiligen Modell die Region definiert ist. Zwischen subnationalen Regionen sind Transportkosten in der Regel nur die tatsächlichen Kosten des physischen Transports von Gütern. Zwischen Ländern werden Transportkosten eher wie Handelskosten aufgefasst und umfassen auch tarifäre und nicht-tarifäre Handelshemmnisse, wie z.B. Zölle und auch Transaktionskosten. Anderson und van Wincoop (2004) definieren Handelskosten als die Differenz aller Kosten, die für ein Produkt am Ort des Konsums entstanden sind und den tatsächlichen Produktionskosten. Die Handelskosten teilen sich auf in echte Transportkosten (die auch die Kapitalbindung der in Transport befindlichen Güter einschließen), tarifäre und nicht-tarifäre Handelshemmnisse, Informationskosten, Transaktionskosten (Kontaktkosten, Vertragskosten, Kosten des Zahlungsverkehrs einschließlich Wechselkursrisiko), sowie Kosten der lokalen Logistik am Verbrauchsort.²³

Der zuletzt genannte Ansatz hat den Vorteil, dass die Handelskosten relativ gut messbar sind. Denn während es recht einfach ist, die Kosten des physischen Transports zu messen, sind die anderen Bestandteile der Handelskosten oft nur schwer zu quantifizieren. Handelskosten werden in empirischen Studien gemessen als Steuer-äquivalenter Aufschlag auf den Wert der gehandelten Waren. Die durchschnittlichen 170% Handelskosten, die Anderson und van Wincoop (2004) grob für den Handel zwischen Industrienationen schätzen, teilen sich nach Untersuchungen der Autoren auf in 55% Kosten der lokalen Logistik, 21% Kosten des Transports zwischen den Ländern (einschließlich ca. 9% Kapitalbindungskosten) und 44% Kosten der Grenzüberschreitung²⁴. Diese Aufstellung macht zwei Dinge deutlich: Zum einen sind Transportkosten relevant und verschwinden keinesfalls durch neuere Technologien und preiswerteren Gütertransport, wie gern behauptet

²³Vgl. Anderson und van Wincoop (2004), S. 691.

²⁴ $1,55 \cdot 1,44 \cdot 1,21 - 1 = 1,7$. Der Wert von 170% wirkt relativ hoch, beinhaltet aber z.B. auch Zölle, Versicherungen, Wechselkursverluste sowie die Kosten der Geschäftsanbahnung. Zur detaillierten Diskussion der einzelnen Komponenten und verschiedener Schätzverfahren sei auf Anderson und van Wincoop (2004) verwiesen.

tet wird. Zum zweiten entsteht ein nicht unerheblicher Anteil (44%) tatsächlich bei der Überschreitung der Grenze. Dies lässt die oft kritisierte Annahme der Neuen Ökonomischen Geographie, dass Transportkosten nur zwischen den Ländern oder Regionen aber nicht darin anfallen, zumindest etwas realistischer erscheinen. Dies gilt zumindest für Modelle, die sich auf Nationalstaaten beziehen.

Natürlich sind die genannten 170% nur ein grober Durchschnitt, wobei allerdings anzumerken ist, dass dieser Wert für Industrieländer gilt und daher eher eine Untergrenze darstellen dürfte. Dennoch schwanken die Transportkosten (die sich ja auf den Warenwert und nicht auf deren Gewicht beziehen) nach Warengruppe und betrachteten Ländern enorm und haben sich in den letzten Jahrzehnten dramatisch verändert.

Siebert (1997) weist darauf hin, dass die Kosten für Transport und Kommunikation in dem Zeitraum von 1920 bis 1990 massiv gefallen sind.²⁵ Auch für den Zeitraum von 1974 bis 2004 können wir dies aus Abbildung 2.3 erkennen. Hier sehen wir auch, welche Dimension die Frachtkosten im Bezug auf den Wert der zu transportierenden Ware haben.²⁶ Allerdings geht dieser Rückgang der Transportkosten weniger auf einen Rückgang der tatsächlichen Frachtkosten pro Gewichtseinheit zurück, als vielmehr auf ein günstigeres Preis / Gewichts-Verhältnis der Waren, eine Veränderung des Modal Split sowie eine Änderung der Handelspartner zurück.²⁷ Da das Potential solcher Veränderungen begrenzt ist und in Zukunft mit steigenden Ölpreisen zu rechnen ist, werden die Transportkosten langfristig wohl nicht weiter sinken. So sind die realen Kosten pro Mengeneinheit im Luft- und Seeverkehr seit 1980 kaum gesunken, da der technologische Fortschritt durch steigende Ölpreise fast vollständig kompensiert wurde.²⁸

Die Handelskosten hängen in erheblichem Maße von der Art der gehandelten Güter und den am Handel beteiligten Ländern ab. Folgt man der oben dargestellten Unterteilung von Handelskosten in reine Frachtkosten und Kosten der Grenzüberschreitung, wie z.B. Zölle, so zeigt sich, dass es in beiden Bereichen große Unterschiede gibt. So werden in vielen Fällen Fertig- und Halbfertigprodukte anders verzollt als Rohstoffe. Auch die reinen Frachtkosten unterscheiden sich.

²⁵Vgl. Siebert (1997), S. 12 und auch World Bank (2009), S. 173ff.

²⁶Bei den Kurven mit der Bezeichnung "fitted" wurden die Daten um Änderungen in der Zusammensetzung der transportierten Waren und den handelnden Ländern bereinigt.

²⁷Vgl. World Bank (2009), S. 179.

²⁸Vgl. World Bank (2009), S. 179, sowie Hummels (2007).

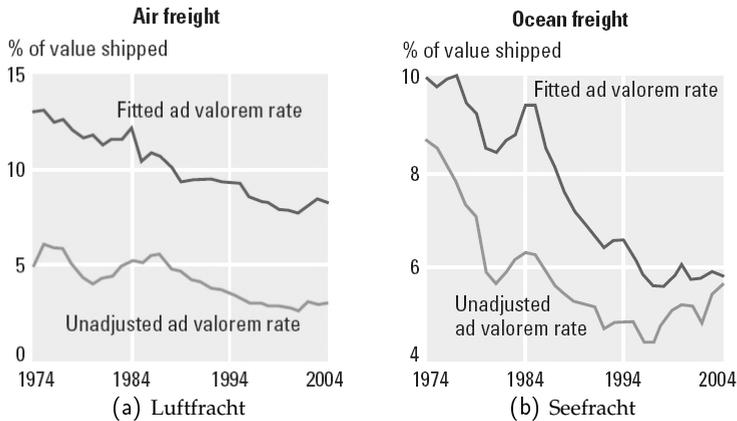


Abbildung 2.3: Entwicklung der Kosten für Luft- und Seefracht (World Bank (2009), S. 179)

Hier kommen verschiedene Faktoren zum Tragen. Zunächst sind Frachtkosten in der Regel vom Gewicht der transportierten Güter abhängig. Ein unterschiedliches Verhältnis von Preis zu Gewicht führt daher zu anderen “ad valorem” Transportkosten. Insbesondere haben industrielle Fertigprodukte (z.B. Elektronik) ein günstigeres Preis/Gewicht-Verhältnis als viele Rohstoffe (z.B. Steinkohle). Daneben haben auch weitere Produkteigenschaften einen Einfluss, z.B. die Notwendigkeit der Kühlung, die Möglichkeit des Transports in Tanks. Als letztes beeinflussen auch die beteiligten Länder durch Hafengebühren und andere Kosten vor Ort sowie natürlich die Distanz und Erreichbarkeit die Gesamtkosten des Transports. Beispielsweise kostete 1996 der Transport einer Tonne Kaffee zwischen Brasilien und Europa \$ 51 (2,6% des Wertes), zwischen Kolumbien und Europa aber \$ 95 (4,9% des Wertes).²⁹ Vergleicht man das Verhältnis der Frachtkosten zum Wert der importierten Güter in die USA für verschiedene Gütergruppen, so zeigt sich, dass z.B. bei Maschinen und Transportausrüstung durchschnittlich 5,7% Frachtkosten anfallen, während bei Rohmaterialien 15,1% Frachtkosten entstehen.³⁰ Diese Werte enthalten nur die reinen Frachtkosten und sind grobe Durchschnitte auf Basis der SITC-Klassifikation. Die Schwankungen bei feiner Unterteilung dürften noch deutlich höher ausfallen. Diese empirischen Befunde sind auch die Motivation, Transportkostenunterschiede in den Sektoren zu modellieren, wie dies in Abschnitt 4.4 dieser Arbeit geschieht.

²⁹Vgl. Hummels (1999a), Tabelle 1.

³⁰Vgl. Hummels (1999b), Tabelle 1.

In der NEG werden Transportkosten im Sinne von Eisberg-Transportkosten aufgefasst³¹. Diese gehen auf Samuelson (1952) zurück und machen die Modellierung eines eigenständigen Transportsektors überflüssig. Die Idee ist, dass die Kosten des Transports in Gütereinheiten des transportierten Gutes gemessen werden und, bildlich gesprochen, während des Transports wegschmelzen. Wenn keine Transportkosten entstehen, ist der Transportkostensatz in diesem Sinne $T = 1$. Niedrigere Sätze sind nicht sinnvoll. Nehmen wir z.B. an, dass $T = 1,1$ sei und Güter von Region 1 nach Region 2 transportiert werden. Damit in Region 2 z.B. 100 Einheiten des Gutes ankommen, müssen in Region 1 110 Einheiten abgeschickt werden.³² Der Wert der restlichen 10 Einheiten deckt die Transportkosten. Diese Art der Modellierung hat allerdings auch Nachteile. Die Faktoren, die für den Transport aufgewendet werden, stehen eigentlich anderen Sektoren nicht mehr zur Verfügung. Da es sich bei der NEG durchweg um Totalmodelle handelt, ist dies bedenklich. Außerdem wird unterstellt, dass die Transportkosten und daher die Transporttechnologie statisch ist, sich also weder in der Zeit noch mit der transportierten Menge ändert. Dies ist in der Realität nicht der Fall. So weist z.B. der World Development Report 2009 darauf hin, dass gerade beim Transport Skaleneffekte auftreten, die z.B. für die geringen Seefrachtkosten in China verantwortlich sind.³³

Abschließend können wir zusammenfassen: Transportkosten sollen hier im Sinne von Handelskosten verstanden werden. Diese sind zwar gesunken, haben aber immer noch eine beträchtliche Höhe und daher einen großen Einfluss auf die Bildung von Agglomerationen und Handelsströmen. Transportkosten unterscheiden sich deutlich nach Sektoren und den am Handel beteiligten Ländern

2.6 Inter- und intraindustrieller Handel

Das Ausmaß des internationalen Handels steht in enger Beziehung zur Höhe der Transportkosten oder der Handelskosten. Aber auch kulturelle und sprachliche Unterschiede, die nur schwer monetär messbar sind, beeinflussen den Handel zwischen zwei Ländern.³⁴ Die Zunahme des internationalen Handels in Folge des

³¹Vgl. Schöler (2005), S. 183 und Brakman et al. (2001), S. 80ff.

³² $\frac{110}{100} = 1,1$

³³Vgl. World Bank (2009), S. 172.

³⁴Vgl. Rossi-Hansberg (2005).

langfristigen Rückgangs der Transportkosten wird in der Literatur auch als Globalisierung bezeichnet. Diese kann in zwei Etappen unterteilt werden, wobei die erste bis etwa zur Mitte des letzten Jahrhunderts reicht und vorwiegend durch interindustriellen Handel gekennzeichnet ist, während die zweite, daran anschließende Etappe eher durch intraindustriellen Handel geprägt ist.³⁵ Die klassische Außenhandelstheorie erklärt Handel durch komparative Vorteile (Ricardo) oder unterschiedliche Faktorausstattungen (Heckscher-Ohlin). Letzteres führt dazu, dass arbeitsreiche Länder Produkte exportieren, bei deren Produktion der Faktor Arbeit relativ intensiv eingesetzt wird. Entsprechend exportieren kapitalreiche Länder kapitalintensiv produzierte Güter. Dadurch wird interindustrieller Handel begründet.

Steigende Skalenerträge und der Wunsch nach Produktvielfalt der Konsumenten führen dagegen zu intraindustriellem Handel. Dabei handeln zwei Länder miteinander, indem beide die Produkte des gleichen Sektors im- und exportieren.³⁶ Ein empirisches Maß für intraindustriellen Handel stellt der Grubel-Lloyd-Index dar.³⁷ Er wird in berechnet als

$$GL_{cd,i} = 1 - \frac{|Exporte_{cd,i} - Importe_{cd,i}|}{Exporte_{cd,i} + Importe_{cd,i}}$$

und misst den Anteil der Im- und Exporte des Sektors i zwischen den Ländern c und d , bei denen das Unternehmen im exportierenden Land dem gleichen Sektor angehört, wie das im importierenden Land, gemessen an der Summe aus Im- und Exporten dieses Sektors. Dieser Index nimmt einen Wert zwischen 0 und 1 an und ist umso höher, je größer der Anteil des intraindustriellen Handels in dieser Branche ist. Er lässt sich über mehrere Jahre, Länder oder auch Branchen aggregieren.³⁸ Einen entscheidenden Einfluss auf den Wert dieses Index hat die Art und Weise, nach der Sektoren abgegrenzt werden. Zum einen gibt es mehrere Klassifikationen, zum anderen haben diese verschiedene Ebenen. Bei der Wahl der Klassifikation ist darauf zu achten, dass möglichst lange Zeitreihen vorhanden sind. Dies ist aber zwangsläufig damit verbunden, dass technologische Veränderungen über die Zeit nicht berücksichtigt werden können. Für die häufig verwendete *Standard*

³⁵Vgl. World Bank (2009), S.170.

³⁶Vgl. Krugman und Obstfeld (2009), S. 195ff.

³⁷Vgl. Grubel und Lloyd (1975).

³⁸Vgl. auch Brühlhart (2009), S. 204f.

International Trade Classification (SITC) liegen Daten seit 1962 vor. Sie hat eine Tiefe von 5 Ebenen, wobei die meisten Publikationen 3-stellige Codes verwenden³⁹. Bei deren Verwendung ergibt sich eine Steigerung des weltweiten intraindustriellen Handels von ca. 25% 1962 auf über 50% 2006.⁴⁰ Generell ist der GL-Index bei hochentwickelten Ländern höher als bei Entwicklungsländern.⁴¹

Der größte Teil des intraindustriellen Handels ist der Handel mit Zwischenprodukten. Bei Verwendung von 5-stelligen SITC-Codes erreichte der GL-Index 2006 einen Wert von 0,355 für Zwischenprodukte (gegenüber 0,272 für Fertigwaren und 0,061 für Rohstoffe), bei Verwendung von 3-stelligen Codes sogar über 0,5.⁴² In Abbildung 2.4 ist die Entwicklung des GL-Index von 1962 bis 2006 für Zwischenprodukte, Fertigwaren und Rohstoffe dargestellt. Hier wird der hohe Anteil von Zwischenprodukten deutlich sichtbar. Dies ist ein Beleg dafür, dass die zentrale Annahme des Handels mit Zwischenprodukten der Modelle in Abschnitt 4 auch empirisch begründet werden kann.

2.7 Faktormobilität

Das Grundmodell der Neuen Ökonomischen Geographie sowie eine Reihe darauf aufbauender Modelle basieren auf der Mobilität von Faktoren, meist des Faktors Arbeit. Ebenso wie in den vorangegangenen Unterabschnitten dieses Kapitels sollen auch hier einige empirische Daten zu dieser Annahme dargestellt werden, um deren Plausibilität beurteilen zu können.

Üblicherweise wird in ökonomischen Modellen von den Faktoren Kapital und Arbeit ausgegangen. Oft, und insbesondere auch in der Neuen Ökonomischen Geographie, wird Arbeit noch in hoch- und geringqualifizierte Arbeit unterschieden. Bei Modellen, in denen die Mobilität der Arbeit im Mittelpunkt der Betrachtung steht, wird der Faktor Kapital oft weggelassen. Dies kann damit gerechtfertigt werden, dass Kapital eine hohe Mobilität aufweist und daher ubiquitär verfügbar ist.

³⁹Z.B. ergibt sich bei Verwendung von 5-stelligen Codes gegenüber 3-stelligen im ungewichteten Durchschnitt von 216 Ländern 0,073 statt 0,138, für die USA 0,317 statt 0,503 und für Deutschland 0,419 statt 0,570. Vgl. Brülhart (2009), S. 410ff.

⁴⁰Vgl. Brülhart (2009).

⁴¹Vgl. Brülhart (2009), Abbildung 2.

⁴²Vgl. Brülhart (2009), S. 418 und 252 sowie World Bank (2009), S. 171.

Evolution of global intraindustry trade, by 3-digit product group, 1962–2006

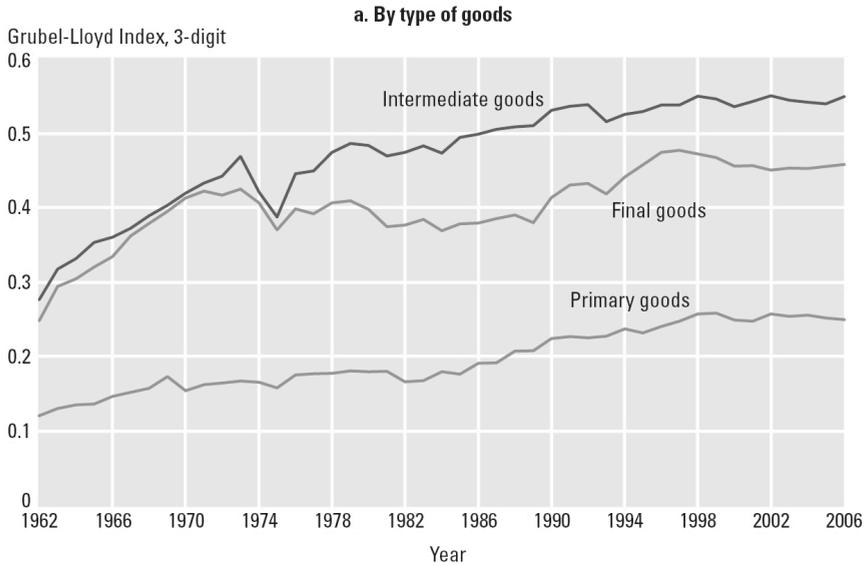


Abbildung 2.4: Handel mit Zwischenprodukten 1962 - 2006 (World Bank (2009), S. 171)

Empirisch lässt sich das jedoch nur teilweise belegen. Verglichen mit dem Faktor Arbeit ist Kapital deutlich mobiler.⁴³ Dennoch hat die Kapitalmobilität nicht kontinuierlich zugenommen, sondern war im Jahr 2000 etwa auf dem Niveau von 1914.⁴⁴ Das liegt daran, dass vor 1914 die Goldparität Währungen abgesichert hat. Mit den beiden Weltkriegen und der Weltwirtschaftskrise ist der internationale Finanzmarkt zusammengebrochen und wurde dann stetig wieder aufgebaut. Daher hat die Kapitalmobilität in den letzten Jahrzehnten stetig zugenommen. Die oft geäußerte Ansicht, sie habe in den letzten Jahren extreme Werte angenommen und sei daher maßgeblich an weltwirtschaftlichen Tendenzen beteiligt, kann aber aus historischer Sicht nicht bestätigt werden.

Eine ganz andere Situation ist beim Faktor Arbeit zu beobachten. Nach der *International Organization of Immigration* leben ca. 214 Mio. Menschen (ca. 3,1% der Weltbevölkerung) außerhalb des Landes, in dem sie geboren sind.⁴⁵ Auch wenn dies eine hohe Zahl ist, so muss die Arbeitsmobilität dennoch als relativ gering

⁴³Vgl. World Bank (2009), S. 149.

⁴⁴Vgl. Obstfeld und Taylor (2003).

⁴⁵Vgl. International Organization for Migration (2010).

eingeschätzt werden. Dies gilt insbesondere dann, wenn man sich die Lohnsatzdifferenzen zwischen Entwicklungs- und Industrieländern vor Augen hält. Die These des Standardmodells der NEG, dass Arbeiter bereits bei geringen Unterschieden im Lohnsatz die Region wechseln, kann daher für Nationalstaaten nicht bestätigt werden. Dies ist die empirische Grundlage dafür, dass in Modellen, die sich auf die Agglomeration auf Länderebene beziehen, keine Arbeitsmobilität unterstellt wird, sondern industrielle Verflechtungen (vgl. auch Abschnitt 2.6) zur Erklärung von Agglomerationsbildung herangezogen werden.

Trotz dieser Tatsache ist ein Trend zu verzeichnen, nachdem der Teil der Migration, der auf (extreme) Lohnunterschiede zurückgeht, ansteigt. Während im 19. Jhd. noch der größte Teil der Emigranten aus Industrieländern stammten, betrug der Anteil der Auswanderer aus Entwicklungs- und Schwellenländern in Industrieländer 2000 39% der gesamten Wanderungen in diesem Jahr. Weitere 16% waren Wanderungen zwischen Industrieländern und 24% Wanderungen zwischen Entwicklungs- und Schwellenländern.⁴⁶

Die Mobilität innerhalb von Ländern muss als deutlich höher angesehen werden. Allerdings gibt es über Migration innerhalb von Staaten (sogenannte interne Migration) keine einheitlichen Daten. Es sind aber verschiedene Untersuchungen auf Landesebene aus den Jahren 1992 bis 2006 verfügbar.⁴⁷ Interne Wanderungsbewegungen in Deutschland werden auch in Abschnitt 6.2 dieser Arbeit untersucht.

Im Fall der internen Migration folgen die Wanderungen oft einem Lohngefälle. Menschen wandern von Regionen mit geringem durchschnittlichen Einkommen oder hoher Arbeitslosigkeit in wirtschaftlich bessere Regionen. In den meisten Ländern ist dies mit Wanderungen aus ländlichen Gebieten in die Städte verbunden. Allerdings gehen die größten Wanderungsbewegungen nicht auf ökonomische Motive, sondern auf politische Instabilität und gewaltsame Auseinandersetzungen zurück.⁴⁸ Im Fall von ökonomisch motivierter Migration kann ein positiver Zusammenhang zwischen Bildungsniveau und Wahrscheinlichkeit der internen Wanderung festgestellt werden. Dies belegt eine häufig getroffene Annahme der Neuen Ökonomischen Geographie, dass nämlich die Industriearbeiter

⁴⁶Vgl. World Bank (2009), S. 150.

⁴⁷In World Bank (2009), S. 153ff sind die Ergebnisse dieser Untersuchungen zusammengetragen. Die in diesem Abschnitt genannten Trends stammen aus dieser Quelle.

⁴⁸So fällt auf, dass in Tabelle 5.3 des World Development Reports 2009 die Staaten mit der höchsten internen Wanderung in den letzten Jahren vor der jeweiligen Erhebung instabile politische Verhältnisse hatten.

wandern und die Landarbeiter (eher allgemein als wenig qualifizierte Arbeiter zu verstehen) regional immobil sind.

2.8 Monopolistische Konkurrenz

Die Marktform der Monopolistischen Konkurrenz spielt in der Neuen Ökonomischen Geographie eine zentrale Rolle. Sie begründet die steigenden Skalenerträge, die dazu führen, dass jedes Unternehmen nur einen Standort hat und daher die Standortwahl der Firmen das Güterangebot in einer Region festlegt. Alle hier behandelten Modelle der NEG beinhalten mindestens einen Sektor mit dieser Marktform.⁴⁹ Im Sinne einer Strukturierung der komplexen Modelle der NEG und um bei Annahmen und Kritik besser unterscheiden zu können, ob sie sich auf das Dixit-Stiglitz-Modell oder dessen räumliche Erweiterungen beziehen, soll hier zunächst die Monopolistische Konkurrenz dargestellt und diskutiert werden.

Intuition und Abbildung Die Marktform der Monopolistischen Konkurrenz beschreibt einen Markt, auf dem keine homogenen, sondern heterogene Güter gehandelt werden. Das bedeutet, die Güter sind nahe Substitute, aber nicht identisch. Beispielsweise gibt es viele Produkte auf dem Markt für Erfrischungsgetränke, die alle das gleiche Bedürfnis befriedigen, dennoch gibt es nur einen Anbieter von Coca Cola.⁵⁰ Infolge dessen ist jeder Hersteller Monopolist für seine eigene Produktvariante, im Gegensatz zum Monopol können die Nachfrager aber auf andere Produktvarianten ausweichen. Diese Marktform kann als recht realistisch angesehen werden, da diese Beschreibung auf viele Märkte zutrifft. Dabei muss es sich nicht um eine Unterschiedlichkeit der Produkte in physischen Eigenschaften handeln. Es kann auch sein, dass eine Variante schneller zu beziehen ist, in einem Geschäft in der Nähe des Kunden verkauft wird oder in einer als passender empfundenen Packungsgröße angeboten wird.⁵¹ Mit dieser Marktform werden in der Neuen Ökonomischen Geographie Industriesektoren beschrieben. Jedes Unternehmen hat hier einen Entscheidungsspielraum bei Preis und angebotener Menge seines Gutes. Dennoch fließen die Entscheidungen der anderen

⁴⁹Dem Autor ist auch kein Modell dieser Gattung bekannt, bei dem das nicht der Fall ist.

⁵⁰Vgl. Varian (2007), S. 545ff.

⁵¹Vgl. Tirole (1994), S. 277.

über die Nachfrage nach jeder einzelnen Variante in den Gewinn ein. Dieser Einfluss wird umso größer sein, je ähnlicher sich die Produktvarianten sind. Durch die Unterschiedlichkeit der Güter ist es den Unternehmen also möglich, den Preis zu setzen und daher prinzipiell auch Gewinne zu realisieren (im Gegensatz zur vollkommenen Konkurrenz). Im Gegensatz zum Oligopol reagieren die Anbieter aber nicht strategisch aufeinander, sondern die Reaktion findet indirekt über Preise und Mengen auf dem Markt statt.

Weiterhin wird bei der monopolistischen Konkurrenz freier Marktzutritt angenommen. Da nun Unternehmen einen positiven Gewinn erzielen, treten neue Unternehmen in den Markt ein. Dadurch verschiebt sich für jeden Anbieter die Kurve der auf ihn entfallenden Nachfrage (seine Preis-Absatz-Funktion) nach unten, worauf das Unternehmen mit einer Preisänderung entsprechend der Preissetzungsregel im Monopol reagiert. Dies geschieht solange, bis der Gewinn aller Anbieter auf Null geschrumpft ist.⁵²

Durch diese Überlegungen kommt man zu der bekannten Tangentiallösung: Da die realisierte Preis-Mengen-Kombination sowohl auf der Nachfragekurve als auch auf der Durchschnittskostenkurve liegen muss (Gewinn = 0) und die Nachfragekurve definitionsgemäß einen fallenden Verlauf hat, muss sich das Gleichgewicht in einem Tangentialpunkt von Nachfragekurve und Durchschnittskostenkurve einstellen, der links vom Durchschnittskostenminimum liegt (Abb. 2.5). Daraus kann unmittelbar erkannt werden, dass dieser Punkt nicht effizient ist. Außerdem kann aus der Graphik auch eine weitere Eigenschaft der Monopolistischen Konkurrenz abgelesen werden:

In principle imperfect competition drives a wedge between prices and marginal costs. (Brakman und Garretsen (2003))

Mathematische Formulierung Die in der Neuen Ökonomischen Geographie verwendete Marktform der Monopolistischen Konkurrenz geht (in dieser Form) auf Dixit und Stiglitz (1977) zurück und wird daher auch als Dixit-Stiglitz-Modell bezeichnet. Da Angebot und Nachfrage in dieser Marktform an vielen Stellen der Arbeit eine Rolle spielen und deren Herleitung nicht in jedem Modell wiederholt

⁵²Im Prinzip ist ein solcher Anpassungsprozess auch ohne Markteintritte möglich. Vgl. Chamberlin (1933).

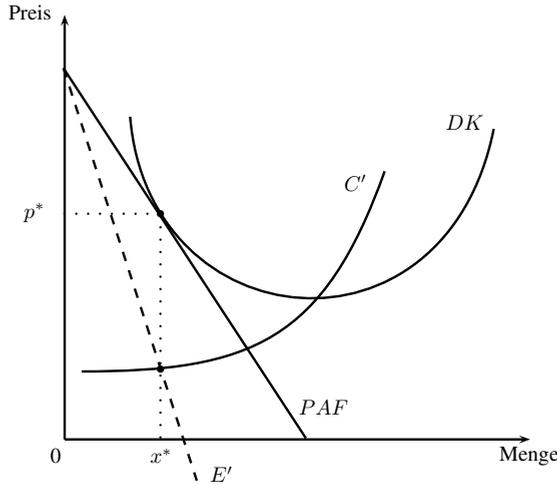


Abbildung 2.5: Tangentiallösung bei Monopolistischer Konkurrenz

werden sollen, folgt an dieser Stelle eine detaillierte Darstellung, auf die dann später an den entsprechenden Stellen verwiesen wird. Dabei ist die hier beschriebene Formulierung etwas spezialisierter als die von Dixit und Stiglitz (1977) in dem Sinne, als dass hier eine Nutzenfunktion auf oberer Ebene vom Cobb-Douglas Typ und auf unterer Ebene vom CES Typ verwendet wird, während das Grundmodell nur fordert, dass die Nutzenfunktion der oberen Ebene separabel ist und konvexe Indifferenzkurven besitzt und die der zweiten Ebene symmetrisch ist.⁵³ Das bedeutet, dass auch andere Formulierungen als die im Standard-Modell von Krugman, wie z.B. die quasi-lineare oder die quadratische Nutzenfunktion kompatibel mit dem Modell sind.⁵⁴

Nachfrage Wir nehmen auf oberer Ebene eine Nutzenfunktion vom Cobb-Douglas Typ an. In diese gehen zwei Sektoren (A und M) ein. Sektor M hat die Marktform der monopolistischen Konkurrenz, über den zweiten soll hier keine Annahme getroffen werden.⁵⁵

⁵³Vgl. Dixit und Stiglitz (1977), S. 298. In Abschnitt III wird dort auch der Fall der Asymmetrie betrachtet. Die hier beschriebene Version ist eine Kombination aus Abschnitten I und II des Beitrags von Dixit und Stiglitz.

⁵⁴Eine quasi-lineare Nutzenfunktion mit einer quadratischen Sub-Nutzenfunktion wird z.B. in Ottaviano et al. (2002) verwendet.

⁵⁵Während im CP-Modell für den zweiten Sektor die Marktform der vollkommenen Konkurrenz angenommen wird, verwendet z.B. das CM ebenfalls monopolistische Konkurrenz (Vgl. Kap. 4.3).

$$U = M^\mu \cdot A^{1-\mu} \quad \text{mit} \quad M = \left[\int_0^n m_k^\rho dk \right]^{\frac{1}{\rho}}$$

Nun folgt ein zweistufiger Optimierungsprozess, bei dem zunächst der Nutzen aus den Aggregaten unter der Nebenbedingung der Budgetrestriktion maximiert wird. Daraus erhalten wir den Nutzen, den die beiden Aggregate ergeben. In einem zweiten Schritt werden dann die Ausgaben pro Gütergruppe minimiert, so dass der Wert der jeweiligen Sub-Nutzenfunktion genau den im ersten Schritt ermittelten Aggregaten entspricht. Die Preisindizes der Aggregate bezeichnen wir mit Q_A und Q_M . Y ist das Einkommen. Der Parameter μ spiegelt die Gewichtung der beiden Gütergruppen in der linear homogenen Nutzenfunktion des repräsentativen Nachfragers wieder. Für eine Cobb-Douglas-Nutzenfunktion erhalten wir im ersten Schritt

$$\max U = M^\mu \cdot A^{1-\mu} \quad \text{s.t.} \quad A \cdot Q_A + M \cdot Q_M = Y \quad (2.3)$$

mit der Lagrangefunktion

$$\mathcal{L} = M^\mu \cdot A^{1-\mu} - \lambda (A \cdot Q_A + M \cdot Q_M - Y) \quad (2.4)$$

Als Bedingungen erster Ordnung ergeben sich

$$\frac{\partial \mathcal{L}}{\partial M} = \mu M^{\mu-1} A^{1-\mu} - \lambda Q_M = 0 \quad (2.5)$$

$$\frac{\partial \mathcal{L}}{\partial A} = (1-\mu) A^{-\mu} M^\mu - \lambda Q_A = 0 \quad (2.6)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = A \cdot Q_A + M \cdot Q_M - Y = 0 \quad (2.7)$$

Nun wird Gleichung (2.5) durch (2.6) geteilt und entsprechend gekürzt

$$\frac{\mu}{1-\mu} \cdot \frac{A}{M} = \frac{Q_M}{Q_A}$$

und nach einer Menge (hier A) aufgelöst.

$$A = \frac{Q_M}{Q_A} \cdot \frac{1-\mu}{\mu} \cdot M \quad (2.8)$$

Dies kann man nun in Gleichung (2.7) einsetzen und umformen zu

$$M \cdot Q_M + \frac{Q_M}{Q_A} \cdot \frac{1-\mu}{\mu} \cdot M \cdot Q_A = Y$$

Aufgelöst nach M ergibt sich

$$M = \frac{\mu Y}{Q_M} \quad (2.9)$$

und durch Einsetzen in (2.8)

$$A = \frac{(1-\mu) Y}{Q_A} \quad (2.10)$$

Dies ist das wohl bekannte Ergebnis, dass im Falle einer Cobb-Douglas-Nutzenfunktion der Ausgabenanteil jedes Gutes gleich dem entsprechenden Exponenten der Nutzenfunktion ist. Um nun die Nachfrage nach einer einzelnen Variante von M zu ermitteln, minimieren wir die Ausgaben für differenzierte Güter unter der Nebenbedingung, dass der Nutzen daraus M entspricht.

$$\min. \int_0^n p_k m_k dk \quad \text{s.t.} \quad \left[\int_0^n m_k^\rho dk \right]^{\frac{1}{\rho}} = M \quad (2.11)$$

Die linke Seite der Nebenbedingung ist die Sub-Nutzenfunktion vom CES Typ (Constant Elasticity of Substitution), wobei der Parameter ρ eben diese Substitutionselastizität misst und damit ein Maß der Vorliebe für Vielfalt der Konsumenten darstellt. Die Menge jeder einzelnen Variante ist mit m_k bezeichnet, p_k ist deren Preis. Für dieses Optimierungsproblem lautet die Lagrange-Funktion

$$\mathcal{L} = \int_0^n p_k m_k dk - \lambda \left(\left[\int_0^n m_k^\rho dk \right]^{\frac{1}{\rho}} - M \right)$$

mit den Bedingungen erster Ordnung für zwei beliebige Güter i und j aus dem

Aggregat.

$$p_i - \lambda \cdot \frac{1}{\rho} \cdot \left[\int_0^n m_k^\rho dk \right]^{\frac{1}{\rho}-1} \cdot \rho \cdot m_i^{\rho-1} = 0 \quad (2.12)$$

$$p_j - \lambda \cdot \frac{1}{\rho} \cdot \left[\int_0^n m_k^\rho dk \right]^{\frac{1}{\rho}-1} \cdot \rho \cdot m_j^{\rho-1} = 0 \quad (2.13)$$

$$\left[\int_0^n m_k^\rho dk \right]^{\frac{1}{\rho}} - M = 0 \quad (2.14)$$

Teilt man die ersten beiden Ableitungen durcheinander und kürzt, ergibt sich

$$\frac{m_i^{\rho-1}}{m_j^{\rho-1}} = \frac{p_i}{p_j}$$

Löst man dies nach m_i auf, erhält man

$$m_i = m_j \cdot \left(\frac{p_j}{p_i} \right)^{\frac{1}{1-\rho}}$$

Dies kann nun in Gleichung (2.14) eingesetzt werden. Nach Umformung erhält man

$$m_j = \frac{p_j^{\frac{1}{\rho-1}} \cdot M}{\left[\int_0^n p_i^{\frac{\rho}{\rho-1}} di \right]^{\frac{1}{\rho}}} \quad (2.15)$$

Die Ausgaben für eine Gütervariante können auch als $p_j m_j$ berechnet werden. Das Minimum der Ausgaben über alle Gütervarianten (2.11) erhält man dann als Integral über alle n Varianten

$$\int_0^n p_j m_j dj = \left[\int_0^n p_i^{\frac{\rho}{\rho-1}} di \right]^{\frac{\rho}{\rho-1}} \cdot M \quad (2.16)$$

Da die linke Seite von (2.16) die gesamten Ausgaben für Industriegüter darstellt und M das Aggregat dieser Güter repräsentiert, ist der Ausdruck davor der Preisindex.⁵⁶ Es gilt: $\sigma = \frac{1-\rho}{\rho}$. Dieser Wert ist positiv, da ρ zwischen 0 und 1 liegt und

⁵⁶Auf den Index M wird an dieser Stelle im Sinne einer übersichtlicheren Darstellung verzichtet.

beschreibt die Elastizität der Nachfrage.

$$Q = \left[\int_0^n p_i^{\frac{\rho}{\rho-1}} di \right]^{\frac{\rho-1}{\rho}} = \left[\int_0^n p_i^{1-\sigma} di \right]^{\frac{1}{1-\sigma}} \quad (2.17)$$

Q sind dann die minimalen Kosten, die nötig sind, um eine Einheit des Zusammengesetzten Gutes M zu kaufen. Setzt man (2.17) in (2.15) ein, so kann man für die Nachfrage nach einer Einheit der Variante j schreiben⁵⁷

$$m_j = M \cdot \left[\frac{p_j}{Q} \right]^{\frac{1}{\rho-1}} = M \cdot \left[\frac{p_j}{Q} \right]^{-\sigma}$$

oder unter Verwendung von (2.9):

$$m_j = \mu Y \frac{p_j^{-\sigma}}{Q^{1-\sigma}} \quad \forall j \in [0, n] \quad (2.18)$$

Dies ist die einzelwirtschaftliche Nachfrage nach der Produktvariante j . Sie hängt neben dem Parameter μ vom Einkommen, dem Preis der Variante und dem Preisniveau ab.

Angebot Da jedes Unternehmen Alleinanbieter seiner Produktvariante ist, handelt es entsprechend als Monopolist. Aus der Maximierung der Gewinnfunktion

$$\pi = E(x) - C(x)$$

ergibt sich daher die bekannte Bedingung "Grenzerlös gleich Grenzkosten",

$$E'(x) = C'(x)$$

wobei der Grenzerlös in zweifacher Weise von der Menge abhängt:

$$E' = \frac{\partial E}{\partial x} = \frac{\partial x}{\partial x} \cdot p(x) + \frac{\partial p}{\partial x} \cdot x \quad (2.19)$$

⁵⁷Dies entspricht Gleichung (7) in Dixit und Stiglitz (1977).

Dabei gilt $\frac{\partial x}{\partial p} = 1$. Erweitert man den letzten Term von (2.19) mit p , so ergibt sich

$$E' = p + x \cdot \frac{\partial p \cdot p}{\partial x \cdot p} = p + p \cdot \frac{\partial p}{\partial x} \cdot \frac{x}{p} \quad (2.20)$$

$\frac{\partial p}{\partial x} \cdot \frac{x}{p}$ ist aber genau der Kehrwert der Elastizität. Wir können daher (2.20) durch Ausklammern von p auch schreiben als

$$E' = p \left[1 - \frac{1}{\varepsilon} \right] \quad (2.21)$$

Dieser Ausdruck ist in der Literatur als Amoroso-Robinson-Relation bekannt, wobei in dieser Formulierung die Elastizität bereits negativ definiert ist.⁵⁸ Berücksichtigt man nun, dass $E' = C'$ gelten soll, so folgt aus (2.21) durch Auflösen nach p

$$p = \frac{C'}{1 - \frac{1}{\varepsilon}}$$

als Regel für die Preissetzung. Da die Elastizität der CES-Nachfragefunktion σ entspricht, kann der Preis angegeben werden als⁵⁹

$$p = C' \left(\frac{\sigma}{\sigma - 1} \right) \quad \text{bzw.} \quad p = \frac{C'}{\rho} \quad (2.22)$$

Da freier Zu- und Abgang von Unternehmen am Markt unterstellt wird, muss der Gewinn im Gleichgewicht null betragen. Voraussetzung hierfür ist, dass n hinreichend groß ist, da sonst der Gewinn des letzten hinzukommenden Unternehmens, das noch keine Verluste macht, größer als null wäre. Der Gewinn ist genau dann null, wenn die Deckungsbeiträge den fixen Kosten entsprechen.⁶⁰

$$F = (p - C') \cdot x$$

Setzt man hier für den Preis die Gleichung (2.22) ein, so ergibt sich

$$F = \left(\frac{C'\sigma}{\sigma - 1} - C' \right) x$$

⁵⁸Genau genommen müsste hier der Betrag der Elastizität stehen.

⁵⁹Dies ist ein Aufschlag auf die Grenzkosten, was auch als *Markup-Pricing* bezeichnet wird.

⁶⁰Hier wird unterstellt, dass die variablen Kosten den Grenzkosten entsprechen.

Löst man dies nach der Menge auf, erhält man den gewinnmaximalen Output pro Unternehmen⁶¹

$$x^* = (\sigma - 1) \frac{F}{C'} \quad (2.23)$$

Mit dieser Formulierung von Angebot und Nachfrage ist es nun möglich, die optimale Menge des Produktionsfaktors sowie die Anzahl der Unternehmen zu bestimmen. Da sich die Annahmen über die verwendeten Faktoren zwischen den verschiedenen vorgestellten Modellen unterscheiden, werden diese Ergebnisse in den entsprechenden Abschnitten präsentiert. Weiterhin kann unter Verwendung von mehreren Sektoren und mehreren Regionen auf Basis dieses Modells ein räumliches Gleichgewicht hergeleitet werden.

Kritik am Dixit-Stiglitz-Modell Grundsätzlich basiert das Dixit-Stiglitz-Modell auf den Grundannahmen der neoklassischen Theorie, die hier jedoch nicht diskutiert werden soll. Darüber hinaus kann die Annahme einer über den gesamten Bereich fallenden Durchschnittskostenkurve kritisiert werden, die dazu führt, dass Unternehmen nur einen Standort haben. Diese fallenden Durchschnittskosten entstehen durch Fixkosten bei der Produktion. Außerdem wird angenommen, dass der gleiche Produktionsfaktor als fester und variabler Faktor in die Produktion einfließt, was bezweifelt werden kann.⁶² Obwohl für die Produktion Fixkosten angenommen werden, können Unternehmen frei den Standort wechseln. Das bedeutet, dass diese fixen Kosten keine versunkenen Kosten sind.

Einige Autoren⁶³ führen darüber hinaus noch das Fehlen von strategischem Verhalten der Unternehmen sowie die Tatsache, dass es sich um Ein-Produkt-Unternehmen handelt, als Kritikpunkte an. Dies sind jedoch Eigenschaften des Modells, die man durchaus als sinnvoll erachten kann. Anderenfalls müsste man ein Oligopolmodell verwenden, das jedoch einen ganz anderen Erklärungsschwerpunkt hätte.

Das Dixit-Stiglitz-Modell hat seit seinem Erscheinen großen Einfluss auf verschiedene Teile der Volkswirtschaftslehre, z.B. der Außenhandelstheorie, der Industrieökonomik, der Wachstumstheorie und eben auch der Raumwirtschaftstheorie. Es

⁶¹Das gleiche Ergebnis erhält man, wenn man (2.22) in die Gewinngleichung einsetzt und diese gleich null setzt.

⁶²Vgl. Woll (2003), S. 251ff.

⁶³Z.B. Eckey und Kosfeld (2004). Für eine Gegenargumentation vgl. Brakman et al. (2001), S. 69.

ermöglichte eine konkrete Formulierung von vielen Aussagen, die bei Chamberlin (1933) nur verbal und oft vage getroffen werden konnten.⁶⁴

2.9 Das Core-Periphery-Modell

Das Core-Periphery-Modell (im folgenden CP-Modell) von Krugman (1991b) ist das Grundmodell der Neuen Ökonomischen Geographie. Es wurde in diversen Veröffentlichungen ausführlich dargestellt. Zu den umfassendsten Darstellungen zählen Fujita et al. (1999) und Brakman et al. (2001) sowie in deutscher Sprache Klüver (2000) und Roos (2002). Eine bekannte Kritik findet sich in Neary (2001). Schöler (2005) bindet das Modell in die anderen Theorien der Raumwirtschaftstheorie ein. Da das CP-Modell einerseits in der Literatur bereits umfangreich behandelt wurde, andererseits aber die Grundlage aller in dieser Arbeit behandelten Modelle bildet, soll es hier kurz dargestellt werden.

In Abbildung 2.6 ist der Grundaufbau des Modells vereinfacht dargestellt. Das Modell besteht aus zwei Ländern mit jeweils zwei Sektoren, Landwirtschaft und Industrie. Es handelt sich um ein Totalmodell. Dabei ist die Bezeichnung "Landwirtschaft" nicht allzu wörtlich zu nehmen. Gemeint ist ein Referenzsektor, in dem ein homogenes Gut unter vollkommener Konkurrenz mit konstanten Skalenerträgen produziert wird. Im Industriesektor wird ein differenziertes Produkt unter steigenden Skalenerträgen produziert und in der Marktform der Monopolistischen Konkurrenz nach Dixit und Stiglitz (1977) gehandelt.⁶⁵ Beide Gütergruppen werden konsumiert und fließen daher in eine Cobb-Douglas-Nutzenfunktion der Konsumenten ein. Dabei geben diese einen festen Anteil μ für Industrieprodukte aus und entsprechend $1 - \mu$ für Agrarprodukte. Die Güter können gehandelt werden, wobei im Falle der Industrieprodukte Transportkosten in Höhe von T anfallen.

Damit es zur Herausbildung von Agglomerationen kommen kann, müssen die industriellen Arbeitskräfte ihren Standort wechseln können. Sie wandern jeweils in die Region, in der ihr Reallohn am höchsten ist. Die Landarbeiter (auch dies ist nicht wörtlich zu nehmen) sind immobil und sorgen somit dafür, dass sich eine

⁶⁴Ein ähnliches Modell wurde von Spence (1976) vorgestellt. Dieses hatte aber bedeutend weniger Einfluss.

⁶⁵Zum Dixit-Stiglitz Modell siehe auch Abschnitt 2.8.

Region nicht völlig entleert, denn auch sie konsumieren sowohl Agrar- als auch Industrieprodukte. Durch Preiseffekte auf dem Absatz- und dem Beschaffungsmarkt kann es nun zur Bildung von Agglomerationen kommen. Ziel des Modells ist es, diese Agglomerationsbildung in Abhängigkeit von verschiedenen Parametern, insbesondere den Transportkosten, zu erklären. Dabei wird von völlig identischen Regionen ausgegangen. Genau diese Annahme, die zunächst sinnvoll ist, um ein möglichst allgemeingültiges Modell zu erhalten, wird später ein Ansatzpunkt zur Modifikation sein.

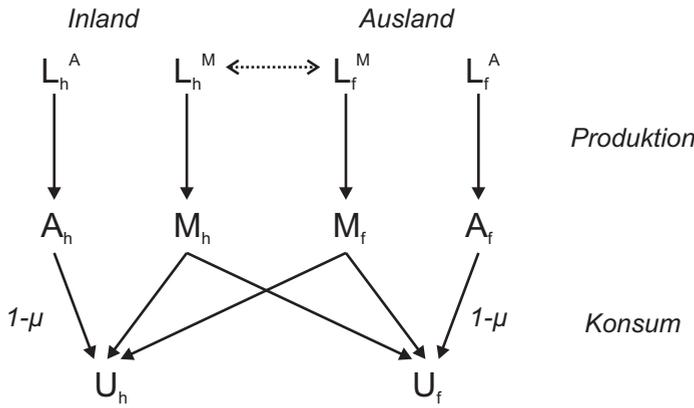


Abbildung 2.6: Schematischer Aufbau des CP-Modells (eigene Darstellung in Anlehnung an Brakman et al. (2001), S.63)

Nachfrage Die Konsumenten haben eine Nutzenfunktion vom Cobb-Douglas-Typ, die von einem homogenen Gut A und einer großen Anzahl differenzierter, aber ähnlicher Produkte M abhängt.

$$U = M^\mu \cdot A^{1-\mu} \quad (2.24)$$

wobei μ zwischen 0 und 1 liegt. Das Aggregat M ist mit einer CES-Subnutzenfunktion beschrieben.

$$M = \left[\int_0^n m_i^\rho di \right]^{\frac{1}{\rho}} \quad \rho \in (0, 1) \quad (2.25)$$

Der Nutzen hängt nun positiv von dem Konsum des homogenen Gutes A und jeder Produktvariante des differenzierten Gutes m_i sowie von der Anzahl ver-

fügbarer Varianten n ab. Dies wird als *Vorliebe für Vielfalt* bezeichnet und leitet sich aus der strikten Konvexität der Präferenzen ab. Eine Mischung von mehreren Produktvarianten spendet immer einen höheren Nutzen als der Konsum von nur einer Variante.⁶⁶ Der Parameter ρ ($0 < \rho < 1$) ist definiert als $\rho = \frac{\sigma-1}{\sigma}$ und gibt das Maß dieser Vorliebe an⁶⁷. Ein ρ nahe 1 bedeutet, dass die einzelnen Varianten nahezu vollkommene Substitute sind (die Vorliebe für Vielfalt ist dann gering), bei ρ nahe 0 werden die Varianten als sehr unterschiedlich empfunden.

Ein einzelner Haushalt sieht sich unter diesen Bedingungen folgender Budgetrestriktion gegenüber:

$$p^A \cdot A + \int_0^n p_i m_i di = Y \quad (2.26)$$

Um nun die Nachfragemengen nach A und m_i herzuleiten, verwendet man, wie in Abschnitt 2.8 beschrieben, ein zweistufiges Optimierungsverfahren. Zunächst wird (2.24) unter der Nebenbedingung (2.26) maximiert, um die gesamte nachgefragte Menge M und A zu erhalten. Danach werden die auf die differenzierten Güter entfallenden Ausgaben so minimiert, dass das im ersten Schritt berechnete Aggregat M konsumiert wird. Daraus erhält man die Nachfragemengen der einzelnen Produktvarianten m_i . Dieses Vorgehen ist möglich, da die Präferenzen für A und M separierbar sind und die M homothetisch in m_i ist.⁶⁸ Im Falle von Cobb-Douglas-Nutzenfunktionen, bei denen sich die Exponenten zu 1 addieren, entsprechen die Ausgabenanteile den Exponenten. Wir erhalten daher

$$A = \frac{(1 - \mu) Y}{p^A}$$

und

$$M = \frac{\mu Y}{Q}$$

Nach dem in Abschnitt 2.8 beschriebenen Verfahren wird die Nachfrage nach einer Produktvariante (2.18) errechnet. Geht man nun von R Regionen aus, in denen das Gut nachgefragt wird, und berücksichtigt die Transportkosten T , so kann die

⁶⁶Vgl. Mas-Collel et al. (1995), S. 44, Dixit und Stiglitz (1977), S. 297.

⁶⁷Für σ gilt dementsprechend $\sigma = \frac{1}{1-\rho}$.

⁶⁸Vgl. Klüver (2000), S. 55.

Nachfrage am Ort der Produktion einer Variante geschrieben werden als

$$m_r = \mu \sum_{r=1}^R Y_r \frac{p_r^{-\sigma} T^{-\sigma}}{Q^{1-\sigma}} \cdot T$$

Dabei entstehen die multiplikativ verknüpften Transportkosten am Ende der Gleichung, weil T Einheiten am Produktionsort verschickt werden müssen, damit eine Einheit am Konsumort ankommt. Die Transportkosten im Zähler hingegen stellen den Effekt auf den Preis dar, der die einzelwirtschaftliche Nachfrage beeinflusst.⁶⁹ Diese Gleichung kann nach dem Preis aufgelöst werden, um die inverse Nachfragefunktion zu erhalten, mit der später der Nominallohn ermittelt wird

$$p_r = \left[\frac{\mu}{m_r} \sum_{r=1}^R Y_r (T)^{1-\sigma} Q_r^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (2.27)$$

Angebot Das Agrargut A wird annahmegemäß mit konstanten Skalenerträgen unter Einsatz von Arbeit als einzigem Produktionsfaktor hergestellt. Die Einheiten werden so gewählt, dass genau eine Einheit Arbeit zur Produktion von einer Einheit A benötigt wird. Damit ist $A = l_A$. Das differenzierte Gut wird ebenfalls nur mit Arbeit hergestellt. Die benötigte Menge an Arbeit setzt sich dabei zusammen aus einem fixen Einsatz von F pro Standort und einem variablen Einsatz von c pro Outputeinheit

$$l_m = F + c \cdot q \quad (2.28)$$

Die Tatsache, dass F Arbeitseinheiten pro Standort benötigt werden, ist verantwortlich dafür, dass Unternehmen immer nur an einem Standort produzieren. Dadurch kommen die steigenden Skalenerträge zum Ausdruck. Die Gewinnfunktion eines Unternehmens lautet

$$\pi = p \cdot q - w (F + c \cdot q) \quad (2.29)$$

Das Unternehmen maximiert nun seinen Gewinn im Hinblick auf p , wobei der Preisindex Q als gegeben und unabhängig von p angesehen wird. Begründet wird

⁶⁹Hier wird vereinfachend davon ausgegangen, dass die Transportkosten zwischen allen Ländern gleich sind. Das ist insofern nicht problematisch, als dass alle hier beschriebenen Modelle zwei Länder betrachten. Die Vereinfachung lässt sich aber auch aufheben, ohne die Lösbarkeit des Modells zu gefährden.

dies mit der großen Anzahl von Unternehmen.⁷⁰ Entsprechend der im letzten Abschnitt beschriebenen Vorgehensweise erhält man den optimalen Preis aus Gleichung (2.22) als

$$p = cw \left(\frac{\sigma}{\sigma - 1} \right) \quad (2.30)$$

Dabei entspricht die Elastizität σ und die Grenzkosten $c \cdot w$. Dieser Preis ist für alle Güter gleich. Der gewinnmaximale Output pro Unternehmen (2.23) ist dann

$$q^* = (\sigma - 1) \frac{F}{c} \quad (2.31)$$

Setzt man nun (2.31) in (2.28) ein, so erhält man die Menge an Arbeit, die pro Unternehmen zur Erzeugung dieses optimalen Outputs notwendig ist.

$$l^* = F + cq^* = F\sigma \quad (2.32)$$

Daraus kann die Anzahl der Firmen n bestimmt werden, indem man die verfügbare industrielle Arbeitskraft L durch den Arbeitsbedarf pro Unternehmen teilt.

$$n = \frac{L}{l^*} = \frac{L}{F\sigma} \quad (2.33)$$

Aus den Gleichungen (2.30) und (2.31) wird insbesondere deutlich, dass eine Veränderung der Marktgröße keinen Einfluss auf den Output eines einzelnen Unternehmens und den Preis einer Produktvariante hat. Eine externe Steigerung der Nachfrage würde zu einer Erhöhung der Anzahl (identischer) Unternehmen n und somit Produktvarianten führen. Der Grund hierfür liegt in der konstanten Preiselastizität der Nachfrage, die aus der CES-Nutzenfunktion resultiert.

Die Lohngleichung Im Gleichgewicht muss der Preis aus der inversen Nachfragefunktion (2.27) dem optimalen Preis der Anbieter (2.30) entsprechen. Löst man dies nach w auf und berücksichtigt die optimale Menge q^* , erhält man den Lohn-

⁷⁰Vgl. Dixit und Stiglitz (1977), S. 299.

satz im Gleichgewicht.

$$\begin{aligned}
 cw \left(\frac{\sigma}{\sigma-1} \right) &= \left[\frac{\mu}{q^*} \sum_{r=1}^R Y_r (T)^{1-\sigma} Q_r^{\sigma-1} \right]^{\frac{1}{\sigma}} \\
 w &= \left(\frac{\sigma-1}{c\sigma} \right) \left[\frac{\mu}{q^*} \sum_{r=1}^R Y_r (T)^{1-\sigma} Q_r^{\sigma-1} \right]^{\frac{1}{\sigma}}
 \end{aligned} \tag{2.34}$$

Reallohn Um den Reallohn ω zu erhalten, wird der Nominallohn durch das Preisniveau geteilt. Dabei ist zu berücksichtigen, dass das Preisniveau der Agrarprodukte 1 ist und die Arbeiter ihr Einkommen zu einem Anteil von μ für industrielle Produkte ausgeben.

$$\omega = wQ^{-\mu} \tag{2.35}$$

Standardisierungen Um die Anzahl freier Variablen des Modells zu reduzieren und so seine Lösung zu vereinfachen, werden die Einheiten der Ausbringungsmenge sowie des fixen Inputs F sinnvoll gewählt. Dabei werden die Grenzkosten (gemessen in Arbeitseinheiten) festgelegt als

$$c \equiv \frac{\sigma-1}{\sigma} = \rho \tag{2.36}$$

Eine zusätzliche Einheit Output kann also mit ρ zusätzlichen Einheiten Arbeit produziert werden. (2.30) wird dann unter Verwendung von (2.36) zu

$$p = w$$

und aus (2.31) und (2.32) ergibt sich

$$q^* = l^*$$

Der fixe Faktoreinsatz F ist exogen und kann daher in beliebigen Mengeneinheiten gemessen werden. Setzt man

$$F \equiv \frac{\mu}{\sigma}$$

wird (2.33) zu

$$n = \frac{L}{\mu}$$

und aus (2.31) wird

$$q^* = l^* = \mu$$

Mit diesen Standardisierungen und unter Berücksichtigung der Transportkosten kann nun der Preisindex (2.17) geschrieben werden als

$$Q = \left[\sum_{r=1}^R n_r (p_r T_r)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = \left[\frac{1}{\mu} \sum_{r=1}^R L_r (w_r T_r)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (2.37)$$

und die Lohngleichung zu

$$w = \left(\frac{\sigma - 1}{\sigma c} \right) \cdot \left[\frac{1}{q^*} \sum_{r=1}^R Y_r T_r^{1-\sigma} Q_r^{\sigma-1} \right]^{\frac{1}{\sigma}} = \left[\sum_{r=1}^R Y_r T_r^{1-\sigma} Q_r^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (2.38)$$

Diese Standardisierungen sind zulässig und schränken nicht die Allgemeingültigkeit des Modells ein. Es ist aber wichtig, sie bei der Interpretation einzelner Variablen zu berücksichtigen.⁷¹

Kurzfristiges Gleichgewicht Um ein Gleichgewicht bestimmen zu können, muss zunächst eine Annahme über den Anpassungsmechanismus getroffen werden. Es wird angenommen, dass die Arbeiter in die Region mit dem höheren Reallohn wandern und dass diese Wanderung nicht sofort, sondern mit einer zeitlichen Verzögerung eintritt.

$$\dot{\lambda}_r = \gamma (\omega_r - \bar{\omega}) \lambda_r \quad \text{mit} \quad \bar{\omega} = \sum_{r=1}^R \lambda_r \omega_r$$

Es wird angenommen, dass die (immobilen) landwirtschaftlichen Arbeitskräfte, die auch nicht den Sektor wechseln können, gleichmäßig auf die Regionen verteilt sind. Wählt man nun die Einheiten für landwirtschaftliche und industrielle Arbeitskräfte in sinnvoller Weise, indem $L^A = 1 - \mu$ und $L^M = \mu$ gesetzt wird, so

⁷¹Vgl. Klüver (2000) S. 76.

ergibt sich folgende Gleichung für das Budget der Haushalte⁷²

$$Y_r = L_r^M w_r \lambda_r + \frac{1 \cdot L^A}{R} = \mu \lambda_r w_r + \frac{(1 - \mu)}{2} \quad (2.39)$$

Dabei nehmen wir zwei Regionen an. Wir können nun die Gleichungen (2.35), (2.37), (2.38) und (2.39) jeweils für beide Regionen formulieren und erhalten folgendes Gleichungssystem:

$$Y_1 = \mu \lambda w_1 + \frac{1 - \mu}{2} \quad (2.40)$$

$$Y_2 = \mu(1 - \lambda)w_2 + \frac{1 - \mu}{2} \quad (2.41)$$

$$Q_1 = \left[\lambda w_1^{1-\sigma} + (1 - \lambda)(w_2 T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (2.42)$$

$$Q_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1 - \lambda)w_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (2.43)$$

$$w_1 = \left[Y_1 Q_1^{\sigma-1} + Y_2 T^{1-\sigma} Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (2.44)$$

$$w_2 = \left[Y_1 Q_1^{\sigma-1} T^{1-\sigma} + Y_2 Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (2.45)$$

$$\omega_1 = w_1 Q_1^{-\mu} \quad (2.46)$$

$$\omega_2 = w_2 Q_2^{-\mu} \quad (2.47)$$

Dieses Gleichungssystem kann nicht algebraisch gelöst werden. Die Lösung ist jedoch durch eine Simulation möglich, hat dann aber natürlich keinen Anspruch auf Allgemeingültigkeit. Nach Brakman et al. (2001) werden zunächst Werte für die externen Parameter festgelegt. Danach werden die folgenden Schritte so lange durchlaufen, bis ein Abbruchkriterium erfüllt ist. Damit sind dann die Werte für die endogenen Variablen Y_i , Q_i und w_i festgelegt. Jeder Durchlauf besteht aus folgenden Schritten:

1. Festlegen von Startwerten für den Lohnsatz für beide Regionen (z.B. $w_1 = w_2 = 1$).
2. Diese werden in (2.40) bis (2.47) eingesetzt, um für beide Regionen jeweils Y

⁷²Diese Standardisierung ist nicht unbedingt notwendig, so kommen z.B. Baldwin et al. (2003) in ihrer Darstellung des CP-Modells ohne aus.

und Q zu bestimmen.

3. Mit diesen Werten können nun die Lohnsätze neu berechnet werden.
4. Entsprechen die so ermittelten Lohnsätze einem Abbruchkriterium, sind die endgültigen Werte festgelegt; entsprechen sie einem solchen Kriterium nicht, werden die Schritte 2-4 erneut durchlaufen.

Es stellt sich die Frage, wann dieses Abbruchkriterium erreicht ist. Die Werte, die in jeder Iteration errechnet werden, verändern sich immer weniger. Es ist daher sinnvoll, die Simulation abzubrechen, wenn die Differenz zwischen den in zwei Iterationen bestimmten Werten einen festzulegenden Schwellenwert ϵ unterschreitet.

$$\left| \frac{w_{r,i} - w_{r,i-1}}{w_{r,i-1}} \right| < \epsilon$$

wobei der Wert für ϵ in der Literatur üblicherweise auf 0,0001 festgesetzt wird. Verwendet man $\sigma = 5$, $\mu = 0,4$ und Transportkosten $T = 2,1$, $T = 1,7$ oder $T = 1,5$, erhält man Abbildung 2.7.⁷³

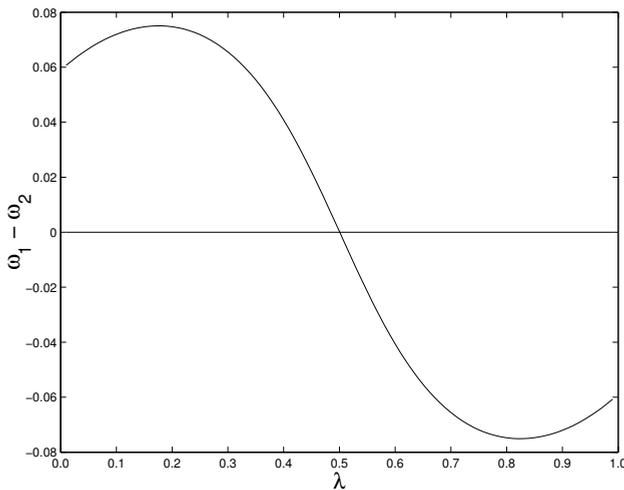


Abbildung 2.7: Reallohndifferenz bei $T = 2,1$

Liegt der Transportkostensatz bei 2,1, ist für $0 < \lambda < 0,5$ die Reallohndifferenz positiv ($\omega_1 > \omega_2$). Daher wandern Arbeiter aus Region 2 in Region 1 ab und lassen λ steigen. Für $0,5 < \lambda < 1$ ist die Argumentation genau umgekehrt. Hier

⁷³Der Quellcode für die Simulationen findet sich in Kapitel 5.3.1.

ist $\omega_2 > \omega_1$, so dass Arbeiter aus Region 1 in Region 2 wandern und λ sinkt. Das (stabile) Gleichgewicht liegt bei $\lambda = 0,5$. Es kommt nicht zu Spezialisierung.

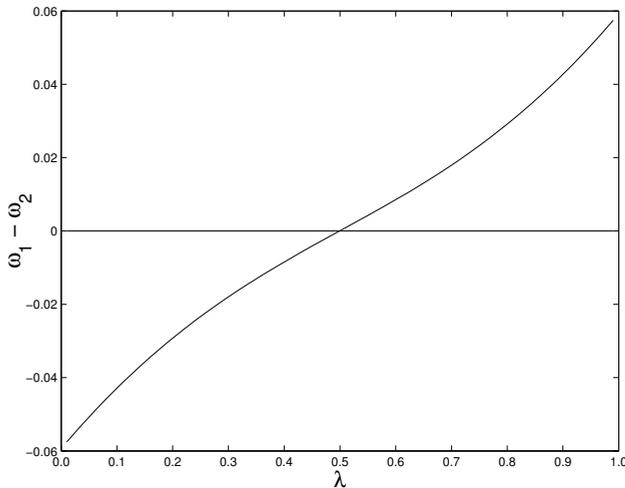


Abbildung 2.8: Reallohndifferenz bei $T = 1,5$

Im Falle von $T = 1,5$ (geringe Transportkosten) ist für $0 < \lambda < 0,5$ die Reallohndifferenz negativ, also $\omega_2 > \omega_1$. Dadurch wandern aus Region 1, die ohnehin schon einen geringeren Anteil der industriellen Produktion hat, weitere Arbeitnehmer in Region 2 ab. Dadurch sinkt λ weiter, bis schließlich alle industriellen Arbeitskräfte in Region 2 sind ($\lambda = 0$). Für $0,5 < \lambda < 1$ gilt der entgegengesetzte Effekt. Bei $\lambda = 0,5$ liegt zwar ein Gleichgewicht vor, dies ist aber instabil, da bereits eine geringe Abweichung zu den beschriebenen Wanderungstendenzen führt und sich eines der stabilen Gleichgewichte einstellt. Für geringe Transportkosten kommt es also zu einer Konzentration der Industrie in einer Region. In der anderen Region findet dann nur noch Landwirtschaft statt.

Liegen die Transportkosten in einem Bereich dazwischen, z.B. bei 1,7, so ergeben sich insgesamt fünf Gleichgewichte: Das symmetrische, die beiden extremen sowie zwei dazwischen; dies zeigt Abbildung 2.9. Aufgrund der bereits oben beschriebenen Kräfte sind in diesem Fall sowohl das symmetrische als auch die beiden Gleichgewichte bei $\lambda = 0$ und $\lambda = 1$ stabil, die beiden dazwischen liegenden instabil. Mit steigenden Transportkosten bewegen sich die instabilen Gleichgewichte immer mehr nach außen und verschwinden dann.

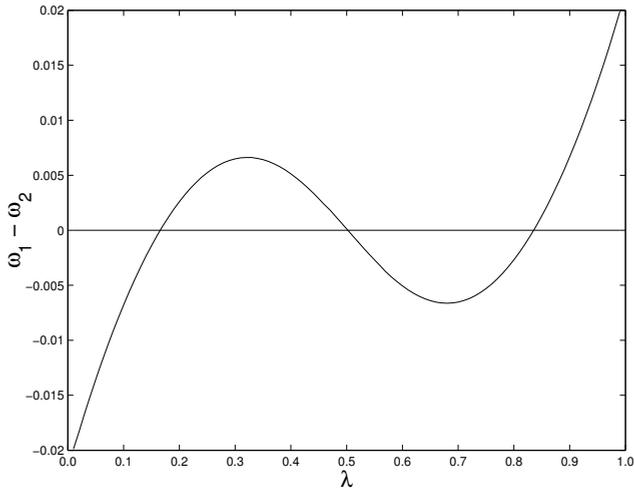


Abbildung 2.9: Reallohndifferenz bei $T = 1,7$

Aus diesen drei betrachteten Fällen ist zu ersehen, dass es für relativ geringe Transportkosten zwei stabile Gleichgewichte (Konzentration in einer der beiden Regionen) und ein instabiles Gleichgewicht bei symmetrischer Aufteilung der Industrie gibt. Im Falle von relativ hohen Transportkosten ist das einzige Gleichgewicht stabil und liegt bei einer hälftigen Aufteilung der industriellen Produktion auf beide Regionen. In einem Bereich mittlerer Transportkosten gibt es fünf mögliche Gleichgewichte. Dabei stellt sowohl eine vollkommene Konzentration der Industrie auf eine der beiden Regionen als auch eine gleichmäßige Aufteilung ein stabiles Gleichgewicht dar. Dazwischen gibt es noch zwei instabile Gleichgewichte, bei denen sich jeweils ein größerer Teil der Industrie in einer und ein kleinerer in der anderen Region befindet.

Diese Gleichgewichte können auch graphisch dargestellt werden, indem man ausgehend von 1 die Transportkosten T erhöht und die daraus resultierende Aufteilung λ abträgt. Dabei stellen in Abb. 2.10 durchgezogene Linien stabile Gleichgewichte und gestrichelte Linien instabile Gleichgewichte dar. Diese Abbildung wird auch als Bifurkationsdiagramm bezeichnet. Der Wert der Transportkosten \underline{T} , bis zu dem ein symmetrisches Gleichgewicht bei abnehmenden Transportkosten stabil ist, wird als *Break-Point* bezeichnet. Nimmt man andersherum steigende Transportkosten an, so sind Gleichgewichte bei vollständiger Agglomeration bis zu einem bestimmten Wert der Transportkosten \bar{T} stabil. Dieser wird *Sustain-Point*

genannt. Aus der Abbildung wird ebenfalls deutlich, dass bei Überschreitung dieser kritischen Werte aus der jeweiligen Richtung ein sprunghafter Wechsel des Gleichgewichts auftritt. Eine Tatsache, die oft für Kritik am CP-Modell sorgt (vgl. auch Abschnitt 2.10) und die durch verschiedene Modifikationen vermieden werden kann, wie in Abschnitt 3.2 gezeigt wird.

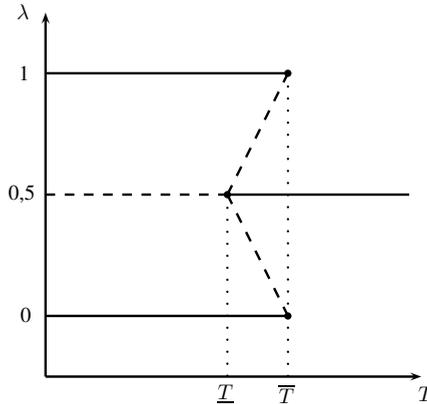


Abbildung 2.10: Gleichgewichte in Abhängigkeit vom Transportkostensatz

Price-Index- und Home-Market-Effect In diesem Abschnitt sollen die Kräfte, die zur Agglomeration führen, näher untersucht und auch eine Bedingung für die Begrenzung der Bildung von Agglomerationen abgeleitet werden. Hierfür betrachten wir die Preisgleichungen (2.42) und (2.43) sowie die Lohngleichungen (2.44) und (2.45). Aufgrund der Symmetrie des Modells erhalten wir für $L_1 = L_2$ und $Y_1 = Y_2$ für die Preisindizes und die Löhne $Q_1 = Q_2$ und $w_1 = w_2$. Im Gleichgewicht ist dann folgende Bedingung erfüllt:⁷⁴

$$1 + T^{1-\sigma} = \frac{\mu}{L} \cdot \left(\frac{Q}{w}\right)^{1-\sigma} = \frac{w}{Y} \cdot \left(\frac{Q}{w}\right)^{1-\sigma} \quad (2.48)$$

wobei auf den Index für die Region verzichtet wurde, da die Werte hier ja für beide Regionen gleich sind. Wir nehmen nun vereinfachend an, dass in einer ausreichend kleinen Umgebung des Gleichgewichtspunktes diese Gleichungen linear sind. Dann ist $dQ = dQ_1 = -dQ_2$ und $dw = dw_1 = -dw_2$. Differenziert man bei-

⁷⁴Zur Berechnung dieser Effekte und auch der *No Black Hole Bedingung* vgl. Fujita et al. (1999), S. 55f.

de Gleichungen und bildet die Wachstumsraten erhält man

$$(1 - \sigma) \frac{dQ}{Q} = \frac{L}{\mu} \cdot \left(\frac{Q}{w}\right)^{\sigma-1} (1 - T^{1-\sigma}) \left[\frac{dL}{L} + (1 - \sigma) \frac{dw}{w}\right] \quad (2.49)$$

und

$$\sigma \frac{dw}{w} = \frac{Y}{w} \cdot \left(\frac{Q}{w}\right)^{\sigma-1} (1 - T^{1-\sigma}) \left[\frac{dY}{Y} + (\sigma - 1) \frac{dQ}{Q}\right] \quad (2.50)$$

Wir nehmen nun an, dass das Angebot an Arbeitskräften vollkommen elastisch ist ($dw = 0$). Da die Transportkosten positiv sind ($T > 0$) und $1 - \sigma < 0$ ist, hat eine Änderung der Anzahl von Industriearbeitern $\frac{dL}{L}$ einen negativen Effekt auf den Preisindex $\frac{dQ}{Q}$. Das bedeutet, je größer der industrielle Sektor ist, desto geringer ist der Preisindex für industrielle Güter. Der Grund hierfür ist einfach: Wenn mehr Industriegüter in der Region hergestellt werden, müssen weniger importiert werden und es fallen daher weniger Transportkosten an. Diesen Zusammenhang bezeichnet man als *Price-Index-Effect*.

Fujita et al. (1999) folgend wird nun ein Index Z definiert, der bei vollkommen kostenfreiem Handel den Wert 0 annimmt und bei unendlich hohen Transportkosten den Wert 1.

$$Z \equiv \frac{1 - T^{1-\sigma}}{1 + T^{1-\sigma}} \quad (2.51)$$

Löst man nun (2.49) und (2.50) nach $\frac{\partial Q}{Q}$ auf und setzt diese gleich, so ergibt sich unter Berücksichtigung von (2.51)

$$\left[\frac{\sigma}{Z} + Z(1 - \sigma)\right] \frac{dw}{w} + Z \frac{dL}{L} = \frac{dY}{Y} \quad (2.52)$$

Aus dieser Gleichung kann zum einen abgelesen werden, dass ein Nachfragezuwachs einen überproportionalen Zuwachs an Beschäftigung und Produktion in Höhe von $\frac{1}{Z} > 1$ nach sich zieht. Das ist der *Home-Market-Effect*. Zum anderen sind die Unternehmen mit wachsendem Industriesektor in der Lage, höhere Löhne zu zahlen. Im CP-Modell mit seinem vollkommen elastischen Arbeitsangebot tritt dieser Effekt offensichtlich nicht ein, aber in den in den Abschnitten 3.2 und 3.4 dargestellten Modellen spielt er eine Rolle.

Die "No black Hole"-Bedingung Diese Effekte werden normalerweise durch gegenläufige Kräfte begrenzt. So hat der Zuzug von industriellen Arbeitern einen negativen Effekt auf den Preis, was den Reallohn erhöht, aber einen negativen Effekt auf den Nominallohn. Dies hat zur Folge, dass ein weiterer Zuzug irgendwann nicht mehr sinnvoll ist und das System ein stabiles Gleichgewicht erreicht hat. Es gibt allerdings Parameterwerte, bei denen der Einfluss der Preisniveaus immer stärker ist als der des Nominallohns, so dass ein Zuzug den Reallohn beständig erhöht. In dem Fall würden die steigenden Skalenerträge dazu führen, dass sich die gesamte Produktion nur an einem Ort oder in einer Region ansiedelt. Fujita et al. (1999) machen dies am Beispiel einer geschlossenen Volkswirtschaft deutlich. Bildet man das totale Differenzial des Nominallohns

$$\frac{d\omega}{\omega} = \frac{dw}{w} - \mu \frac{dQ}{Q}$$

und verwendet (2.49) und (2.50), so erhält man

$$\frac{d\omega}{\omega} = (1 - \mu) \frac{dY}{Y} + \left[\frac{\mu\sigma}{\sigma - 1} - 1 \right] \frac{dL}{L}$$

Damit der Reallohn nun nicht kontinuierlich steigt, muss

$$\frac{\sigma - 1}{\sigma} > \mu$$

gelten. Dies wird als *No Black Hole Bedingung* bezeichnet, weil sie sicherstellt, dass es Kräfte gibt, welche die Agglomerationsbildung begrenzen. Sie gilt in dieser Form für die Modelle der regionalen Agglomeration. Für die in Abschnitt 4 beschriebenen Modelle für internationale Agglomeration ist die Formulierung etwas anders. Die Argumentation ist aber die gleiche.

2.10 Kritikpunkte

Die Neue Ökonomische Geographie hat seit den ersten Beiträgen 1991 eine relativ kontroverse Diskussion sowohl unter Ökonomen als auch unter Geographen ausgelöst. Da diese Arbeit die Modelle dieser Theorie erweitert, erscheint es notwendig, die teilweise recht scharfe Kritik an der NEG kurz zu diskutieren.

Hierbei ist zu unterscheiden, ob es bei der Kritik um die Theorie selbst geht oder um die Erklärungsweite, welche einige Autoren in Anspruch nehmen und die damit zusammenhängenden Konsequenzen auf andere Gebiete der Wissenschaft. Neben einigen, zum Teil wenig sachlichen Diskussionsbeiträgen, die entweder für oder gegen die Neue Ökonomische Geographie argumentieren, sind die bekanntesten Beiträge dazu Neary (2001) und Martin (1999). Allerdings können diese Beiträge die neuen Entwicklungen nicht berücksichtigen und beziehen sich daher vorwiegend auf das Core-Periphery-Modell.

Begriff und Reichweite des Modells Zunächst einmal ist festzustellen, dass die Bezeichnung "New Economic Geography", die von einigen der wichtigsten Vertreter selbst "erfunden" wurde, nicht ganz richtig ist. Martin (1999) stellt dazu fest: "*it is not that new, and it most certainly is not geography*".⁷⁵ Dies stellt zunächst einmal keine Kritik an der Modellwelt selbst dar, sondern am Begriff "New Economic Geography", der nämlich auch so verstanden werden kann, dass alle anderen Theorien dann "Old Economic Geography" sind.⁷⁶ Dies hat einigen Autoren, insbesondere aus der tatsächlichen Wirtschaftsgeographie, wenig gefallen. Auch der Anspruch, den insbesondere Krugman für diese Theorie anmeldet, die verschiedensten Phänomene zu erklären, hat sicher etwas Überhebliches:

"This book shows, in particular, how a common approach - one that emphasizes the three-way interaction among increasing returns, transportation costs, and the movement of productive factors - can be applied to a wide variety of issues in regional, urban, and international economics".⁷⁷

Aus diesem Grund verwenden einige Autoren, wie z.B. Brakman et al. (2001), Brakman und Garretsen (2003) und Brakman et al. (2005), allesamt Vertreter dieser Forschungsrichtung, den Begriff *Geographical Economics*. Dieser verdeutlicht, worum es bei diesen Modellen geht: Um die Einbeziehung des Raumes in Modelle der Mainstream-Ökonomik.⁷⁸ Auch der Begriff *Economic Geography* findet häufig Anwendung, z.B. in Baldwin et al. (2003), Knaap (2004) oder Forslid und Wooton

⁷⁵Martin (1999), S. 67.

⁷⁶Zu Unterschieden und Gemeinsamkeiten der geographisch geprägten und der ökonomisch geprägten Forschungsrichtung siehe auch Overman (2004).

⁷⁷Fujita et al. (1999), S. vi.

⁷⁸Vgl. Brakman et al. (2001), S. xx.

(2003). Dieser hat auch den Vorteil, dass er den jeweiligen Beitrag eindeutig einer Forschungsrichtung zuordnet, ohne dabei durch den Zusatz "new" einen überzogenen Anspruch zu erheben. In dieser Arbeit wird der Begriff "Neue Ökonomische Geographie" nur deshalb verwendet, weil er inzwischen sehr weit verbreitet ist. Vom Gegenstand der Untersuchung her wäre "Räumliche Wirtschaftstheorie" als Übersetzung von "Geographical Economics" sicher sinnvoller, dabei wäre aber die Gefahr einer uneindeutigen Zuordnung wegen der Ähnlichkeit zum allgemeineren Begriff der Raumwirtschaftstheorie zu groß.

Damit wäre der zweite Teil des oben genannten Zitats von Ron Martin "[...] *and it most certainly is not geography*" besprochen. Bleibt die Frage, was an der Theorie neu ist. Martin argumentiert hier, dass bereits lange vor Krugman räumliche Strukturen und auch Agglomerationen mit Modellen erklärt wurden. Außerdem, so stellt er und auch andere Autoren fest, besteht die Grundidee der NEG darin, das Dixit-Stiglitz-Modell und Transportkosten in ein räumliches Wirtschaftsmodell zu integrieren, wie dies auch in ähnlicher Form in der Außenhandelstheorie und der Wachstumstheorie geschehen ist. Die Idee und die einzelnen Bestandteile sind also alle nicht neu. Auch hier muss man den Kritikern in Grunde Recht geben. Die entscheidende Neuerung ist aber, dass dieses Modell Agglomerationen *ohne* exogene Unterschiede und *ohne* Spillover-Effekte erklärt. Und dies mit einem soliden mikroökonomischen Unterbau. Das ist in der Tat neu und durchaus beeindruckend. Dennoch kann, das wurde aber bereits am Ende von Abschnitt 2.8 getan, an dem verwendeten Marktmodell von Dixit und Stiglitz (1977) selbst Kritik geübt werden, hier insbesondere an der Abwesenheit von strategischem Verhalten der Unternehmen.

Modelltheoretische Kritik Möchte man die NEG modelltheoretisch kritisieren, muss überprüft werden, ob die Theorie mit den empirischen Erkenntnissen konsistent ist. Dabei soll zunächst vom Grundmodell ausgegangen werden, um dann bei einigen Kritikpunkten auf neuere Arbeiten zu verweisen, die eben diese Punkte ausräumen. Zunächst fehlt ein Mechanismus, der vollständige Agglomeration in einer Region und damit verbunden die vollständige Entleerung der anderen Region verhindert. Dies kann aber durch eine geringfügige Veränderung am CP-Modell behoben werden, wie es Ludema und Wooton (1997) vorschlagen. Diese führen lokale Präferenzen der Wirtschaftssubjekte ein, die einer Normalverteilung folgen. Dadurch verlassen nicht alle Arbeiter sofort die Region mit dem geringe-

ren Lohn, sondern einige erst bei stärkeren Lohnsatzunterschieden. Dieses Modell ist auch Grundlage der Überlegungen in Kapitel 3 dieser Arbeit.

Es gibt aber noch eine Reihe weiterer Modelle, die nicht zum sogenannten Katastrophenergebnis führen. So berücksichtigen Baldwin et al. (2003) in ihrem *Localized-Spillover-Modell* technisches Wissen, das an einen Ort gebunden ist, und erhalten ein Bifurkationsdiagramm mit einem langsamen Übergang von Gleichverteilung zu Agglomeration.⁷⁹

In den Modellen der internationalen Agglomeration (Kapitel 4) führen Asymmetrien, die ja Gegenstand dieser Arbeit sind, in bestimmten Bereichen zu stabilen Gleichgewichten, die nicht bei einer Gleichverteilung oder bei vollständiger Agglomeration liegen.

Eine weitere Annahme, die wenig sinnvoll erscheint ist die, dass Transportkosten nur für industrielle Güter, nicht aber für landwirtschaftliche anfallen. Unabhängig davon, wie man nun Agrargüter interpretiert, kann das empirisch nicht belegt werden. Es wurde bereits in Fujita et al. (1999) untersucht, welche Konsequenzen Transportkosten auf Agrargüter haben.⁸⁰ Dabei hängt das Ergebnis entscheidend von der Höhe der Transportkosten auf Agrarprodukte ab. Sind diese ausreichend hoch, kommt es immer zu einer Gleichverteilung der industriellen Produktion. Dies ist nicht sehr befriedigend.

Eckey und Kosfeld (2004) haben ebenfalls ein Modell mit Transportkosten im Agrarsektor aufgestellt, kommen aber zu deutlich positiveren Ergebnissen. Nach diesem Modell hängt die Agglomerationsbildung von der Höhe beider Transportkostensätze ab. Qualitativ führen hohe Transportkosten in beiden Sektoren zur Gleichverteilung und geringe zu Agglomeration. Dies ist kompatibel mit dem CP-Modell.

Ein häufig angebrachter Kritikpunkt ist die Notwendigkeit, zu numerischen Methoden zu greifen, da das CP-Modell (wie viele andere Modelle der NEG auch) nicht analytisch lösbar ist. Zunächst einmal muss man dieser Kritik nicht folgen. Fackler und Miranda (2002) führen z.B. an, dass es sinnvoller sei, ein realistisches Modell numerisch zu lösen als es soweit zu vereinfachen, dass es analytisch lösbar ist, dann aber nicht mehr viel mit dem realen Problem zu tun hat. Es gibt

⁷⁹Vgl. Baldwin et al. (2003), S. 173ff.

⁸⁰Vgl. Fujita et al. (1999), S. 97ff.

jedoch eine ganze Reihe von Modellen, die zu ähnlichen Aussagen wie das CP-Modell kommen und empirisch lösbar sind. Am häufigsten genannt werden hier Puga (1999) sowie das *Footloose Capital Model* von Martin und Rogers (1995) und das *Footloose Entrepreneur Model* von Forslid und Ottaviano (2003).⁸¹ Es gibt dabei grundsätzlich zwei Ansätze, um die Lösbarkeit zu erreichen: Es wird eine zirkuläre Verbindung unterbrochen oder es wird eine andere Form der Nutzenfunktion unterstellt, z.B. eine quasi-lineare.

Die bisher angesprochenen Kritikpunkte konnten also durch Modifikationen und Erweiterungen ausgeräumt werden. Problematisch ist aber sicher die Formulierung der Transportkosten. Diese fallen nur zwischen Regionen an (und dann in voller Höhe), nicht aber innerhalb von Regionen. Insbesondere haben die Transportkosten keinen Bezug zur geographischen Entfernung, die in den Modellen ja gar nicht enthalten ist.⁸² Diese Kritik ist sicher berechtigt. Wie relevant sie ist, hängt von der untersuchten Fragestellung ab. Betrachtet man zwei Regionen eines Landes (z.B. die östlichen und die westlichen deutschen Bundesländer in Abschnitt 6.2), so ist die Annahme sicher unrealistisch. Dass beim Transport eines Gutes von Kiel nach München keine Transportkosten entstehen, beim Transport des gleichen Gutes von Kassel nach Erfurt aber schon, ist nicht sinnvoll. Betrachtet das Modell internationale Agglomeration, so ist die Annahme zwar immer noch nicht realistisch, aber weniger kritisch. Wie in Abschnitt 2.5 dargestellt, hängt ein nicht unerheblicher Teil der Handelskosten nicht von der physischen Entfernung sondern tatsächlich von der überschrittenen Grenze ab. Durch diese empirische Erkenntnis erscheint die Realitätsferne hier weniger drastisch.

Empirische Forschung und wirtschaftspolitische Empfehlungen Soll die NEG nicht nur von rein theoretischem Interesse sein, sondern Hinweise für die Wirtschaftspolitik geben, so ist zunächst zu prüfen, ob ihre Aussagen auch empirischen Tests standhalten. Im Falle der grundlegenden Modelle tritt hier das erste Problem auf, die Alles-oder-Nichts-Lösung. Daher müssen für konkrete empirische Untersuchungen eher fortgeschrittene Modelle, z.B. mit Berücksichtigung von lokalen Präferenzen, herangezogen werden.

⁸¹Zu Weiterentwicklungen des *Footloose Entrepreneur Modells* vgl. auch Pflüger und Südekum (2008).

⁸²Ein Modell mit entfernungsabhängigen Transportkosten wurde von Ottaviano et al. (2002) vorgestellt. Hier entwickeln sich die Transportkosten allerdings quadratisch zur Entfernung, was auch nicht sehr plausibel ist.

Bei empirischen Arbeiten zur NEG gibt es grundsätzlich zwei Richtungen. Zum einen kann man versuchen, ein Modell komplett zu schätzen und zu kalibrieren und dann auf der Grundlage der so ermittelten Parameter Prognosen zu erstellen. Auf den ersten Blick bietet sich das bei der NEG an, weil die Modelle sehr komplex und damit vergleichsweise nahe an der Realität sind. Dabei entstehen aber zwei Probleme. Zum einen handelt es sich um Totalmodelle. Aus dieser Sicht sind die Vereinfachungen trotz allem noch relativ stark. Viele Parameter können nur sehr ungenau geschätzt werden (z.B. eine durchschnittliche Substitutionselastizität oder die Ausgabenanteile). Head und Mayer (2004) schätzen diese Parameter für verschiedene Sektoren auf Grundlage einer modifizierten Version des Modells von Puga (1999). Für den Ausgabenanteil errechnen Sie Werte zwischen 6,78% (Lebensmittel) und 0,1% (Mineralstoffe), für die Substitutionselastizität σ innerhalb der Erzeugnisse einer Branche Werte von 11,02 (Computer) und 2,32 (Metalle) und für den Anteil an Zwischenprodukten aus dem eigenen Sektor β 49% (Fahrzeuge) bis 0,12% (Schiffe). Diese Werte zeigen aber, dass die Parameterwerte, die üblicherweise angenommen werden, zumindest nicht völlig unrealistisch sind.

Des Weiteren treten ökonometrische Probleme bei der Schätzung von nicht linearen, nicht monotonen Zusammenhängen auf.⁸³ Möglicherweise wäre ein Lösungsansatz mit *Computable General Equilibrium* (CGE) Modellen denkbar, wie sie z.B. in der Verkehrswissenschaft eingesetzt werden.⁸⁴ Ein generelles Problem ist jedoch die Nicht-Verfügbarkeit von regionalen Preisindizes, wenn sich die Untersuchung auf Regionen und nicht auf Staaten bezieht.

Ein anderer Ansatz ist der, Teile des Modells zu schätzen oder einzelne Wirkungszusammenhänge empirisch zu unterlegen. Dieser Ansatz führt natürlich eher zu Ergebnissen, erfüllt aber nicht den Anspruch, das ganze Modell empirisch zu untermauern. Dieses Vorgehen wird auch von Head und Mayer (2004) vorgeschlagen, die aus den theoretischen Modellen fünf Aussagen ableiten und die empirische Literatur dazu jeweils zusammentragen. Darunter ist auch die Hypothese des Home-Market-Effects, die als eine der ersten getestet wurde, so z.B. von Davis und Weinstein (1998) oder Hanson (2005).

Insgesamt führen diese Untersuchungen zu unterschiedlichen Ergebnissen, was bei der Vielzahl verschiedener Parameter nicht verwundert. Auch besteht hier

⁸³Vgl. Neary (2001), S. 556.

⁸⁴Vgl. Bröcker (2004).

noch ein erheblicher Forschungsbedarf, denn im Gegensatz zur Theorie, die seit der grundlegenden Veröffentlichung von Krugman (1991b) bereits erhebliche Fortschritte gemacht hat und in ersten Büchern zusammenhängend dargestellt wurde,⁸⁵ befindet sich die empirische Forschung dazu noch in einem recht frühen Stadium. Allerdings wird hier zurzeit viel getan, so dass sich das bald ändern dürfte.

Die wirtschaftspolitischen Empfehlungen müssen sich aufgrund des Stands der empirischen Forschung an einzelnen Zusammenhängen orientieren. So wurden bereits die Wirkungen von verschiedenen Maßnahmen, vor allem Steuern, theoretisch untersucht. Baldwin et al. (2003) geben hier einen Überblick, der allerdings nicht mehr auf dem neusten Stand ist. Anwendungen, wie sie in dieser Arbeit in Kapitel 6 dargestellt sind, finden sich z.B. auch in Geppert und Gornig (2005), Südekum (2006) oder Schöler (2007). Auch die Tatsache, dass die Weltbank ihren World Development Report 2009 auf diesem Theoriegebäude aufbaut, belegt die Tragfähigkeit der NEG für wirtschaftspolitische Schlussfolgerungen.

⁸⁵Z.B. Fujita et al. (1999), Brakman et al. (2001) oder Baldwin et al. (2003).

3 Modelle für interregionale Agglomeration

Dieses Kapitel untersucht eine Asymmetrie der Regionsgröße in zwei verschiedenen Modellen zur Erklärung von interregionaler Agglomeration. Dazu wird zunächst das Grundmodell aus Abschnitt 2.9 noch einmal betrachtet. Abschnitt 3.2 stellt dann eine Erweiterung des Modells um beschränkte Mobilität nach Ludema und Wooton (1997) dar. Diese ist sowohl in den Annahmen als auch im Ergebnis realistischer als das CP-Modell selbst, weil Arbeiter nicht mehr bei geringsten Lohnunterschieden den Standort wechseln und im Ergebnis auch langfristig Gleichgewichte möglich sind, bei denen sich zwar Agglomerationen bilden, aber die Peripherie nicht vollständiger Abwanderung unterliegt. Die Gleichgewichte werden mit denen des CP-Modells verglichen, indem für beide Modelle Simulationen mit den gleichen Transportkostensätzen durchgeführt werden. Dabei werden Fälle mit unterschiedlich starken lokalen Präferenzen untersucht.

In den Abschnitten 3.3 und 3.4 wird darauf aufbauend für beide Modelle geprüft, welche zusätzlichen Erkenntnisse sich gewinnen lassen, wenn die Regionen nicht mehr gleich groß sind. Dazu wird der Größenunterschied zunächst in das jeweilige Gleichungssystem integriert. Durch Simulation unter Verwendung der in Kapitel 5 vorgestellten Methoden werden die neuen Gleichgewichte mit denen der bisherigen Modelle verglichen und gezeigt, dass bei sinkenden Transportkosten, die in der Realität meist zu beobachten sind, nun bessere Prognosen darüber möglich sind, wo sich die Agglomeration bildet. Das asymmetrische Modell mit lokalen Präferenzen (Abschnitt 3.4) ist auch Grundlage der wirtschaftspolitischen Anwendung in Kapitel 6.

3.1 Grundmodell

Das grundlegende Modell zur Erklärung von Agglomerationsbildung zwischen Regionen, also bei Mobilität der Arbeiter, ist das Core-Periphery-Modell, das bereits in Abschnitt 2.9 dargestellt wurde. Zum Vergleich mit den weiter unten folgenden Modellvarianten sollen hier nur die beiden wichtigsten Abbildungen dargestellt und kurz kommentiert werden. Dabei werden die im letzten Kapitel beschriebenen Simulationsmethoden eingesetzt. Das Modell beinhaltet insbesondere zwei Annahmen, die in den späteren Unterabschnitten fallen gelassen werden: Die unbegrenzte Mobilität der Arbeitskräfte und die vollkommene Gleichheit der beiden Regionen.

Ersteres bedeutet, dass Arbeiter bereits bei kleinsten Reallohndifferenzen die Region wechseln. Diese Annahme ist aufgrund von in der Realität vorhandenen Umzugskosten unrealistisch. Dabei sind neben den einmaligen monetären Kosten für den Ortswechsel (Umzugskosten im engeren Sinne) auch dauerhafte und zum Teil nicht monetäre Kosten gemeint. Dies können der Verlust an Lebensqualität sein, der entsteht, wenn ein Arbeitnehmer sein gewohntes Umfeld verlässt oder aber Kosten, die durch lange Fahrzeiten zu Freunden und Verwandten entstehen.⁸⁶ Die Annahme gleicher Regionen bedeutet neben einer homogenen Fläche und gleicher Ausstattung mit Produktionsfaktoren vor allem auch, dass beide Regionen die gleiche Größe und Bevölkerung haben.

Die Kernaussage des CP-Modells lässt sich am anschaulichsten durch das Bifurkationsdiagramm (Abbildung 3.1b) darstellen, das alle langfristigen Gleichgewichte (stabile und instabile) in einem Diagramm darstellt, bei dem die Transportkosten auf der horizontalen Achse und die regionale Aufteilung der Arbeiter auf der vertikalen Achse abgetragen werden. "Langfristig" meint dabei, dass die Wanderungen der Arbeiter aufgrund von Reallohndifferenzen abgeschlossen sind. Für jeden Transportkostensatz T kann die Aufteilung λ dadurch bestimmt werden, dass das Gleichungssystem (2.40) bis (2.47) für diesen T -Wert simuliert wird. Anders als in der Darstellung in Abschnitt 2.9 wurde hier auf der vertikalen Achse des Diagramms der kurzfristigen Gleichgewichte (Abb. 3.1a) nicht die Reallohndifferenz ($\omega_1 - \omega_2$), sondern deren Verhältnis (ω_1/ω_2) abgetragen. Der Grund dafür ist, dass

⁸⁶Hier sind natürlich auch negative Kosten, also Nutzenzuwächse, durch einen Umzug denkbar. Solche können z.B. entstehen, wenn die neue Region landschaftlich oder kulturell interessanter ist oder weil ein Ortswechsel generell als angenehm empfunden wird.

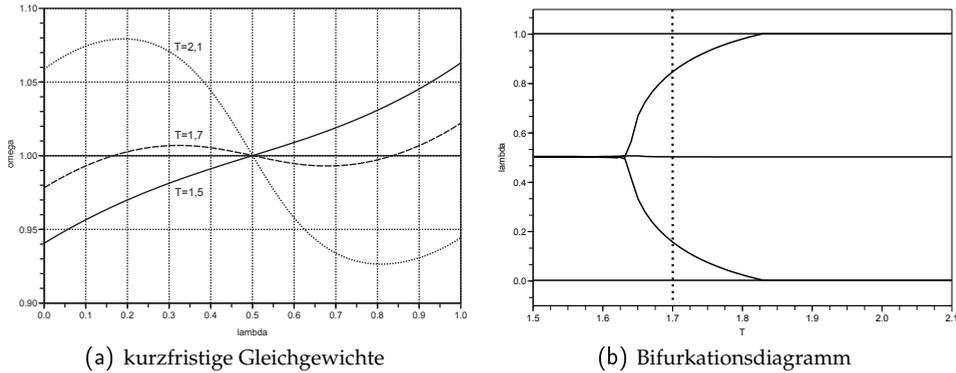


Abbildung 3.1: Diagramme des CP-Modells

bei den späteren Erweiterungen nur das Verhältnis errechnet werden kann und so eine bessere Vergleichbarkeit der Ergebnisse erreicht werden soll. Die grundlegende Form der Kurven bleibt aber auch bei dieser Darstellung erhalten.⁸⁷

Abbildung 3.1 zeigt beispielhaft die drei in der Literatur häufig verwendeten Sätze von $T = 1,5$ für geringe, $T = 1,7$ für mittlere und $T = 2,1$ für hohe Transportkosten. Das Bifurkationsdiagramm wurde direkt mit dem in Kapitel 5 beschriebenen Verfahren erstellt. Daher ist hier nicht, wie in der gezeichneten Variante in Abbildung 2.10, zwischen stabilen und instabilen Gleichgewichten unterschieden. Betrachtet man beide Abbildungen (3.1a und 3.1b) zusammen, kann man dies aber erkennen. Bei $T = 1,5$ ist im linken Bild die gepunktete Linie relevant. Das innere Gleichgewicht bei $\lambda = 0,5$ ist nicht stabil. Es liegen Randlösungen bei $\lambda = 0$ oder $\lambda = 1$ vor. Im rechten Bild befinden wir uns ganz am linken Rand. $1,5$ ist der geringste Wert für T , der hier dargestellt ist. Daher sind die dargestellten Gleichgewichte bei 0 und 1 stabil, das Gleichgewicht bei 0,5 instabil. Der höchste Wert der Transportkosten im Bifurkationsdiagramm von $T = 2,1$ ist in der linken Abbildung durch eine durchgezogene Linie dargestellt. Das symmetrische Gleichgewicht ist hier stabil. Für mittlere Transportkosten wurde hier beispielhaft $T = 1,7$ angenommen. Dieser Wert ist in Abbildung 3.1b durch eine gepunktete Linie gekennzeichnet. Man sieht, dass neben den stabilen Gleichgewichten an den Rändern und in der Mitte der Skala für λ auch noch je ein instabiles dazwischen

⁸⁷Die gebogene Form der Kurven, welche die instabilen Gleichgewichte darstellen, ergibt sich durch die Simulation. In vielen Beiträgen, und auch in Abbildung 2.10 in dieser Arbeit, wird die Abbildung vereinfacht mit geraden Linien dargestellt.

existiert. In Abb. 3.1a kann dies aus der gestrichelten Kurve abgelesen werden.

Auf ein Ergebnis soll hier noch einmal besonders hingewiesen werden, da es in den nächsten Abschnitten aufgehoben wird: Je nachdem, ob steigende oder sinkende Transportkosten untersucht werden, kommt es immer zu Sprüngen an Beak- oder Sustain-Point. Bei diesen ändert sich schlagartig die Verteilung der Arbeitskräfte auf die Regionen. Ein Verhalten, das nur wenig realistisch erscheint und daher häufig Gegenstand von Kritik ist.

3.2 Symmetrische Regionen mit lokalen Präferenzen

Das Modell von Ludema und Wooton (1997) ist eine Weiterentwicklung des Standardmodells, die durch Berücksichtigung von regionalen Präferenzen die Mobilität der Arbeitskräfte einschränkt und dadurch stabile Gleichgewichte zulässt, die zwischen der Gleichverteilung und der vollständigen Spezialisierung liegen. Da sich Modelle, die sich auf Staaten beziehen, von solchen, die sich auf Regionen innerhalb von Ländern beziehen, durch die Annahme zur Mobilität des Faktors Arbeit unterscheiden, ist es ersteren zuzuordnen. Im Gegensatz zu den in Kapitel 4 diskutierten Modellen (IM und CM) wird hier aber nur eine eingeschränkte Mobilität und keine totale Immobilität unterstellt. Daher können die Struktur und auch die Mechanismen des Grundmodells beibehalten werden und es wird keine zusätzliche Verflechtung - etwa über Zwischenprodukte - benötigt. Obwohl das Modell zeitlich nach den anderen Modellen einzuordnen ist, wird es bereits hier dargestellt, weil es zu großen Teilen dem Standard-Modell entspricht.

Wie bereits im CP-Modell wird auch hier von zwei Sektoren ausgegangen, von denen einer unter vollkommener Konkurrenz ein homogenes Gut produziert (Agrar-gut), das ohne Kosten transportiert werden kann. Der andere Sektor produziert ein differenziertes Gut unter der Marktform der monopolistischen Konkurrenz (vgl. Abschnitt 2.8). Bei der Produktion wird ausschließlich Arbeit eingesetzt, die in Landarbeiter (Arbeiter in Sektor *A*) und Industriearbeiter (Arbeiter in Sektor *M*) unterschieden wird. Die landwirtschaftlichen Beschäftigten sind immobil und deren Lohn durch die Annahmen des Modells auf 1 festgesetzt. Die Industriearbeiter hingegen sind mobil, haben aber - und hier liegt der Unterschied zum CP-Modell - mehr oder weniger starke Präferenzen für eine Region. Sie wandern nur dann, wenn der Reallohn der anderen Region auch diese regionalen Präferenzen

mit ausgleicht. Solche Präferenzen können z.B. in Sprache und Kultur aber auch in administrativen Voraussetzungen (Arbeitserlaubnis) begründet sein. Aufgrund der gleichen Struktur, sind auch die Gleichungen, die ein Gleichgewicht beschreiben denen im CP-Modell⁸⁸ gleich:

$$Y_1 = \mu\lambda w_1 + \frac{1-\mu}{2} \quad (3.1)$$

$$Y_2 = \mu(1-\lambda)w_2 + \frac{1-\mu}{2} \quad (3.2)$$

$$Q_1 = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (3.3)$$

$$Q_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (3.4)$$

$$w_1 = \left[Y_1 Q_1^{\sigma-1} + Y_2 T^{1-\sigma} Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (3.5)$$

$$w_2 = \left[Y_1 Q_1^{\sigma-1} T^{1-\sigma} + Y_2 Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (3.6)$$

$$\omega_1 = w_1 Q_1^{-\mu} \quad (3.7)$$

$$\omega_2 = w_2 Q_2^{-\mu} \quad (3.8)$$

Diese Gleichungen beschreiben bei Verwendung alternativer Werte für λ eine Linie, die als Arbeitsnachfrage interpretiert werden kann: Alle Punkte auf dieser Linie stellen ein Gleichgewicht in der Region dar, wobei die Größe des Industriesektors als gegeben angenommen wird. Ein Gleichgewicht des Systems kann aber nur dann erreicht werden, wenn die Industriearbeiter bei gegebenem Lohn auch bereit sind, die benötigte Menge an Arbeit anzubieten. Im Standard-Modell ist das der Fall, wenn $\omega_1 - \omega_1 = 0$ gilt. Daher kann die horizontale Linie bei 0 als Arbeitsangebotskurve interpretiert werden. Genau hier liegt nun der Unterschied zum Modell von Ludema und Wooton (1997). Die Arbeitskräfte haben lokale Präferenzen, die in ihre Entscheidung einfließen. Diese werden durch einen Faktor φ_i modelliert, mit dem der Lohn multipliziert wird. Ein Arbeiter ist genau dann indifferent, wenn

$$\varphi_i \omega_1 = (1 - \varphi_i) \omega_2 \quad (3.9)$$

⁸⁸Gleichungen (2.40) bis (2.47).

gilt. Dabei ist $\varphi_i \in (0, 1)$ der Gewichtungparameter, der die Präferenz des Arbeiters i für die Region 1 ausdrückt.

φ_i ist eine Zufallsvariable, die durch folgende Dichtefunktion beschrieben wird⁸⁹

$$g(\varphi_i, s) = \begin{cases} \frac{\frac{1}{s}\phi\left(\frac{\varphi-0,5}{s}\right)}{\Phi\left(\frac{1-0,5}{s}\right)-\Phi\left(\frac{0-0,5}{s}\right)}, & 0 < \varphi_i < 1 \\ 0, & \text{sonst,} \end{cases} \quad (3.10)$$

wobei $\Phi(\cdot)$ die kumulative Verteilungsfunktion der Normalverteilung mit einem Mittelwert von 0,5 und einer Standardabweichung von s ist. $\phi(\cdot)$ ist die Dichte der Standard-Normalverteilung.⁹⁰

$$\phi(x) = \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi}} \quad (3.11)$$

Durch den Mittelwert von 0,5 ist die Symmetrie sichergestellt: Im Mittel haben die Arbeiter keine Präferenz für eine der beiden Regionen. Je höher aber die Standardabweichung s ist, je größer ist die Präferenz für jeweils eine Region. Die Lohndifferenz muss dann entsprechend größer sein, um auch die Arbeiter mit einer starken Präferenz für die jeweils andere Region zum Wechsel zu bewegen. Bei $s = 0$ liegt der Fall des Standard-Modells vor: Die Arbeiter sind zwischen beiden Regionen indifferent und wandern bei kleinsten Lohnsatzunterschieden. In Abbildung 3.2 ist die Verteilung bei den Standardabweichungen $s = 0,05$ und $s = 0,005$ dargestellt.

Nun wird der sogenannte Grenzarbeiter definiert. Das ist der Arbeiter, dessen Präferenzen ihn bei gegebenem Lohnverhältnis genau indifferent erscheinen lassen. Seinen Gewichtungsfaktor bezeichnen wir mit $\bar{\varphi}$. Aus Gleichung (3.24) sehen wir, dass Indifferenz vorliegt, wenn

$$\frac{\omega_1}{\omega_2} = \frac{1 - \bar{\varphi}}{\bar{\varphi}} \quad (3.12)$$

⁸⁹Diese entspricht weder Gleichung (17) aus dem Artikel von Ludema und Wooton, noch aus der Arbeitspapierversion des gleichen Beitrages, die sich untereinander auch unterscheiden. Mit beiden konnte die Abbildung 2 aus Ludema und Wooton (1997) nicht reproduziert werden. Diese Gleichung entspricht der "Truncated Normal Distribution" aus der Literatur zur Stochastik. Diese liefert die richtigen Ergebnisse, die auch in Abb. 3.2 dargestellt sind.

⁹⁰Der Code zu diesen Funktionen befindet sich im Anhang.

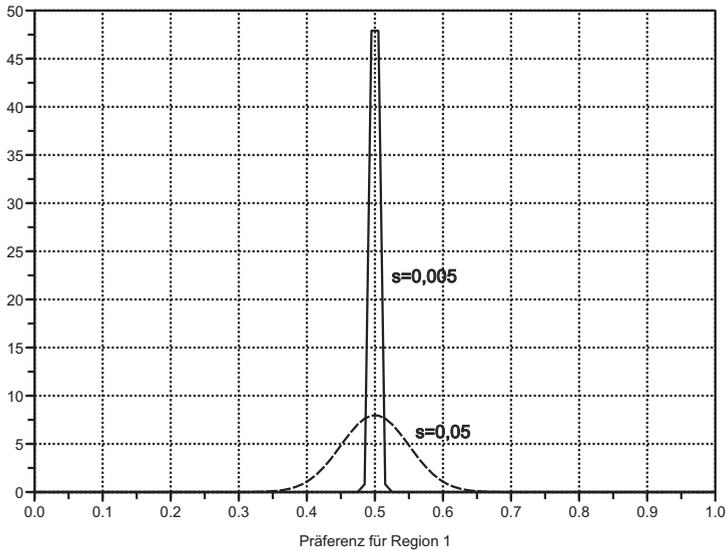


Abbildung 3.2: Verteilung der Präferenzen für Region 1 bei unterschiedlichen Werten für die Standardabweichung

gilt, wobei das Verhältnis $\frac{\omega_1}{\omega_2}$ vereinfachend als ω bezeichnet werden soll. Der Gewichtungsfaktor des Grenzarbeiters lässt sich nun bestimmen als

$$\bar{\varphi} = \frac{1}{\omega - 1} \tag{3.13}$$

Alle Arbeiter, die einen Gewichtungsfaktor von $\varphi_i \geq \bar{\varphi}$ haben, werden entsprechend in Region 1 arbeiten wollen.⁹¹ Dies entspricht $1 - \Phi(\bar{\varphi})$, da $\Phi(\bar{\varphi})$ die kumulierte Wahrscheinlichkeit dafür ist, dass $\varphi_i < \bar{\varphi}$ gilt.

Es wurde damit ein Zusammenhang zwischen dem Lohnverhältnis ω und der Verteilung des Arbeitsangebotes λ bei gegebener Standardabweichung der Verteilung der lokalen Präferenzen hergeleitet, der in Abbildung 3.3 für die bereits oben verwendeten Werte dargestellt ist. Dabei ist auf der vertikalen Achse der relative Lohnsatz abgetragen und auf der horizontalen der Anteil der Arbeiter, der

⁹¹Es ist in dem Modell nicht möglich, in einer Region zu leben und in der anderen zu arbeiten, wie es z.B. in Modellen der Stadtökonomik oft angenommen wird, z.B. in Ottaviano et al. (2002).

in Region 1 lebt (λ).⁹²

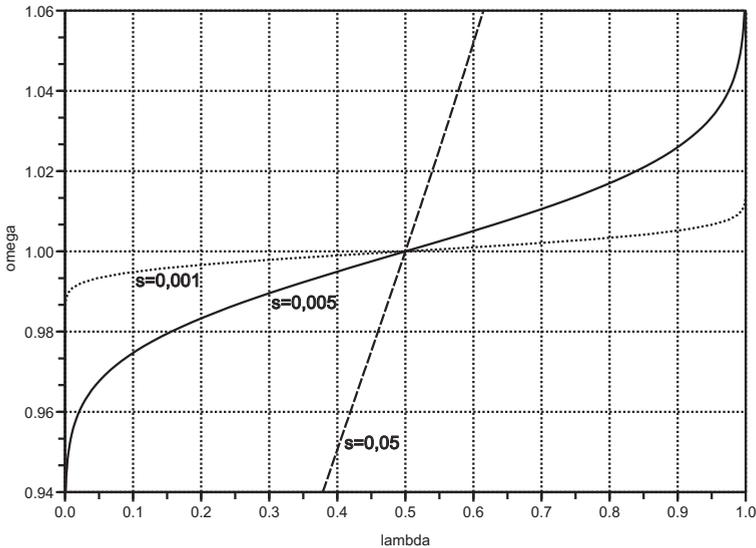


Abbildung 3.3: Arbeitsangebot bei unterschiedlichen Werten für die Standardabweichung

Aus dieser Abbildung lässt sich erkennen, dass im Fall relativ starker lokaler Präferenzen ($s = 0,05$) das Lohnsatzdifferenzial relativ groß sein muss, um mehr als die Hälfte der Arbeiter dazu zu bewegen, in eine Region zu ziehen. Je geringer s wird, desto geringer muss auch diese Prämie ausfallen. Bei $s = 0$ wechseln alle Arbeiter sofort den Standort wenn die Löhne sich unterscheiden. Dies ist der im CP-Modell betrachtete Fall (vgl. Abschnitt 2.9).

Wir haben nun sowohl das Verhalten der Unternehmen als auch das der Arbeiter beschrieben. Dabei bieten die Arbeiter ihre Arbeitsleistung an und die Unternehmen fragen diese zu einem Lohnsatz nach. In welcher Region die Arbeitskräfte ihre Arbeitskraft anbieten, entscheiden diese anhand von Gleichung (3.24), in die der Relativlohn und ihre lokalen Präferenzen eingehen. Nun kann Arbeitsangebot und Arbeitsnachfrage in einem Diagramm dargestellt werden (Abbildung 3.4). Dazu werden die Gleichgewichte aus dem Gleichungssystem (3.1) bis (3.8)

⁹²Die Angebotskurve bei einer Standardabweichung von 0 wäre eine horizontale Linie. Bei der Programmierung des Algorithmus mit Werten von nahezu 0 treten aber Probleme auf, so dass diese Linie hier nicht eingezeichnet ist.

so dargestellt, dass auf der vertikalen Achse $\omega = \omega_1/\omega_2$ abgetragen wird. Diese Veränderung zur Darstellung des Standard-Modells (vgl. Abb. 2.7 bis 2.9) ist notwendig, da das Arbeitsangebot aus Gleichung (3.13) nur in Abhängigkeit vom Relativlohnverhältnis nicht aber von der Lohndifferenz berechnet werden kann.

In Abhängigkeit von den Werten für die Transportkosten und der Standardabweichung der lokalen Präferenzen lassen sich nun unterschiedliche Gleichgewichte bestimmen. Zur Veranschaulichung ist in Abbildung 3.4 der Fall mit $T = 1,7$ und $s = 0,001$ dargestellt. Die Bestimmung der Gleichgewichte erfolgt nun analog zum CP-Modell. Bei einem Relativlohn oberhalb der Arbeitsangebotskurve (gestrichelt) werden mehr Arbeiter in Region 1 wechseln, da der mit ihrem Präferenzfaktor gewichtete relative Lohnsatz in Region 1 relativ hoch ist. Entsprechend wechseln Arbeiter von Region 1 in Region 2, wenn der Lohnsatz unter der Arbeitsangebotsfunktion liegt. Der Relativlohn, den die Unternehmen bei einer gegebenen Verteilung der Arbeitskräfte bieten ist durch die Arbeitsnachfragekurve dargestellt. Diese Kurven sind identisch zu denen im CP-Modell. Liegt die Arbeitsnachfragekurve nun über der Angebotskurve wechseln Arbeiter in Region 1 (λ steigt), im umgekehrten Fall wechseln Arbeiter in Region 2 (λ sinkt). Daraus ergeben sich nun 1, 3 oder 5 Gleichgewichte, abhängig vom Verlauf der Kurven. Bei den hier gewählten Werten kommt es zu fünf Gleichgewichten, von denen drei stabil sind (S) und zwei instabil (U). Die Stabilität der Gleichgewichtspunkte kann aus den mit Pfeilen dargestellten Wanderungen der Arbeiter unmittelbar erkannt werden. Die Randlösungen sind stabil, wenn bei $\lambda = 0$ die Angebotskurve über der Nachfragekurve und bei $\lambda = 1$ die Angebotskurve unter der Nachfragekurve liegt.

Betrachtet man nun die bereits in Abschnitt 2.9 verwendeten Werte für niedrige, mittlere und hohe Transportkosten ($T = 1,5$; $T = 1,7$; $T = 2,1$) zusammen mit den hier genannten Werten für schwache, mittlere und relativ starke lokale Präferenzen ($s = 0,001$; $s = 0,005$; $s = 0,05$), ergibt sich folgendes Bild: Im Fall von geringen Transportkosten (Abbildung 3.5) kommt es ohne lokale Präferenzen (Standard-Modell) und auch bei geringen oder mittleren lokalen Präferenzen zu vollständiger Konzentration der Industrie in einer der beiden Regionen ($\lambda = 0; 1$). Sind die Vorlieben für eine Region aber ausreichend stark (z.B. $s = 0,05$), so kommt es zu einer Gleichverteilung der Produktion. Das Gleichgewicht bei $\lambda = 0,5$ wird dann stabil, die extremen Gleichgewichte entfallen. Dies ist ein erster substanzieller Unterschied zum Standard-Modell, der wohl als realis-

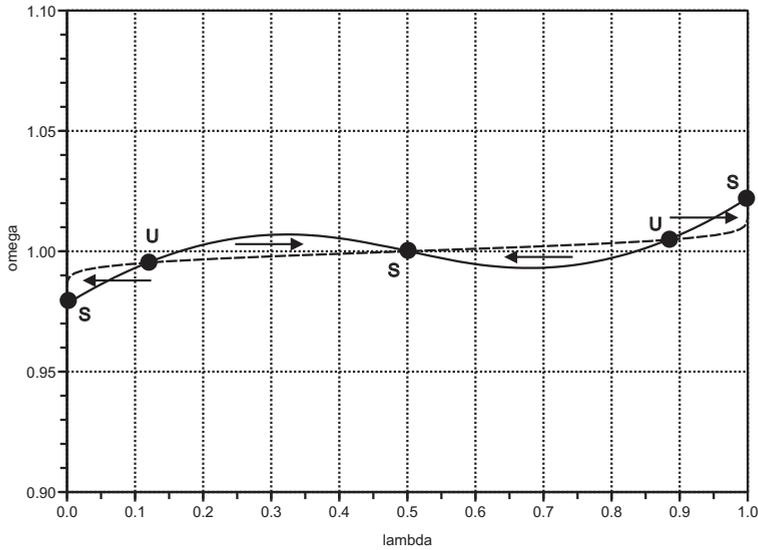
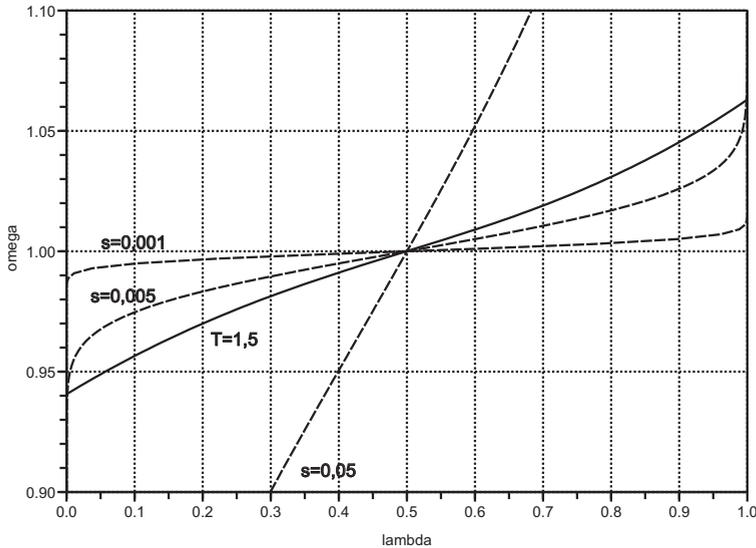


Abbildung 3.4: Gleichgewichte im LW-Modell bei $T = 1,7$; $s = 0,001$

tisch anzusehen ist: Wenn die lokalen Präferenzen der Arbeiter hoch sind, ist der Reallohn, der sie trotzdem zum Wandern bewegen würde, höher als der, den die Unternehmen aufgrund der Skalenvorteile bereit sind zu zahlen. Folglich kommt es nicht zur Herausbildung einer Kern-Peripherie-Struktur.

Untersucht man als nächstes den Fall hoher Transportkosten (Abb. 3.6), erkennt man, dass sich keine Änderungen zum Standardmodell ergeben. Das einzige Gleichgewicht, das stabil ist, liegt bei $\lambda = 0,5$. Die Berücksichtigung lokaler Präferenzen ändert an diesem Ergebnis nichts. Der Grund ist leicht einzusehen: Haben Arbeiter eine Vorliebe für eine Region, bremst dies die Bildung einer Agglomeration. Hier bildet sich diese auf Grund der hohen Transportkosten aber sowieso nicht.

Im Fall mittlerer Transportkosten und ohne lokale Präferenzen (Abb. 3.7) entstehen fünf Gleichgewichte, von denen das mit einer Gleichverteilung verbundene und die beiden extremen Gleichgewichte stabil sind, die beiden jeweils dazwischen liegenden sind instabil. Steigen die lokalen Präferenzen nun an (z.B. bei $s = 0,001$), so bewegen sich die instabilen Gleichgewichtspunkte auf die Ränder zu, bis die Randlösungen schließlich instabil werden und als einzige stabile

Abbildung 3.5: Gleichgewichte bei $T = 1,5$

Lösung die Gleichverteilung übrig bleibt.

Betrachtet man diese Ergebnisse, so zeigt sich, dass es auch möglich ist, zu stabilen Gleichgewichten zu gelangen, die nicht bei den Werten von $\lambda = [0; 0,5; 1]$ liegen. Dies geschieht, wenn bei $\lambda = 0$ die Arbeitsnachfrage über dem Arbeitsangebot liegt und beide Kurven einen Schnittpunkt zwischen $\lambda = 0$ und $\lambda = 0,5$ aufweisen (analog für den Bereich zwischen $0,5$ und 1). Dies ist z.B. bei $T = 1,45$ und $s = 0,01$ der Fall, wie in Abbildung 3.8 dargestellt. Durch die Berücksichtigung von Vorlieben der Arbeiter für eine Region lässt sich also ein stabiles Agglomerationsgleichgewicht darstellen, bei dem die Konzentration zwar stark aber nicht total ist. Dies räumt einen zentralen Kritikpunkt am CP-Modell mit einer relativ einfachen und empirisch plausiblen Erweiterung aus.

Abbildung 3.9 zeigt das bekannte Bifurkationsdiagramm für das LW-Modell mit Werten für die Standardabweichung der lokalen Präferenzen von $s = 0,01$ und $s = 0,005$. Dabei ist mit der gestrichelten Linie der Fall aus Abbildung 3.8 gekennzeichnet. Ein Arbeitsangebot mit $s = 0,05$ würde nur Gleichgewichte auf der Geraden bei $\lambda = 0,5$ hervorbringen. Man erkennt, dass die Abbildung eine

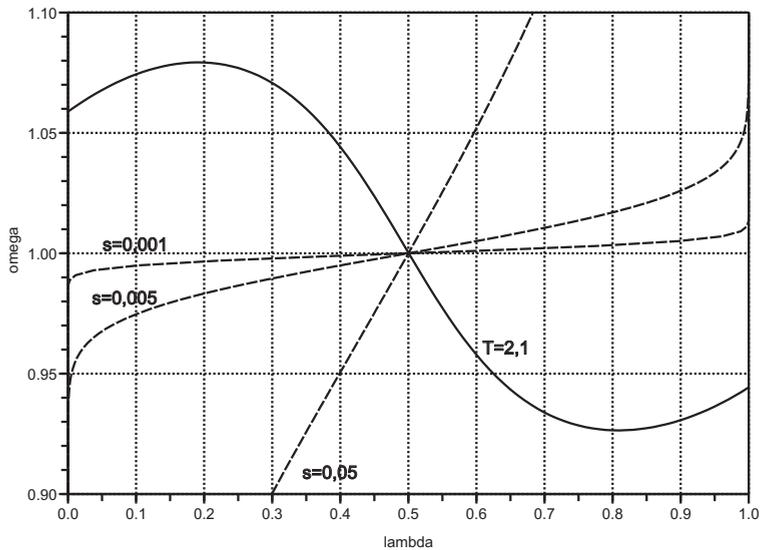


Abbildung 3.6: Gleichgewichte bei $T = 2,1$

ähnliche Form hat, wie im CP-Modell, dass die äußeren stabilen Gleichgewichte nun aber nicht mehr ganz am Rand liegen müssen. Es lässt sich außerdem feststellen, dass eine Konzentrationsbildung mit steigenden Präferenzen der Arbeiter für eine Region unwahrscheinlicher wird. Dieser Effekt tritt bereits hier sehr deutlich zu Tage, obwohl die Werte der Standardabweichung und die Annahme einer Normalverteilung sehr vorsichtig gewählt sind. So gibt es bei diesen Annahmen keine Individuen, die unter keinen Umständen den Standort wechseln würden. Dies könnte man z.B. mit einer T-Verteilung darstellen. Es wird aber bereits hier klar, dass dadurch Agglomerationsbildung leicht ausgeschlossen werden kann.

Zusammenfassend kann man feststellen, dass die Berücksichtigung von lokalen Präferenzen die Bildung von Agglomerationen tendenziell behindert. Starke lokale Präferenzen führen daher immer zur Gleichverteilung der ökonomischen Aktivität. Dies ist unmittelbar ersichtlich, da im LW-Modell genau wie im CP-Modell Agglomerationen durch die Wanderung von Arbeitskräften entstehen. Wollen diese nicht wandern (hohe Standardabweichung der lokalen Präferenzen) bilden sich daher auch keine Agglomerationen. Dies wird bereits in Abbildung

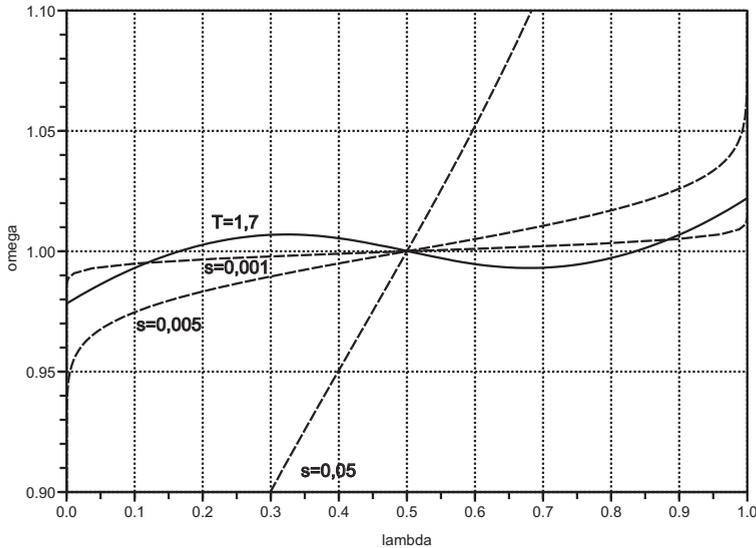


Abbildung 3.7: Gleichgewichte bei $T = 1,7$

3.9 deutlich. Hier liegt der Break-Point unter der Annahme $s = 0,01$ bereits bei deutlich niedrigeren Transportkosten als wenn geringere regionale Präferenzen mit $s = 0,005$ angenommen werden. Außerdem, und hier liegt die Leistung des LW-Modells, sind nun auch stabile Gleichgewichte erklärbar, die nicht bei Gleichverteilung oder extremer Konzentration in einer Region liegen. Dies ist z.B. bei $T = 1,45$ und $s = 0,01$ der Fall. Ebenfalls ändert sich bei steigenden lokalen Präferenzen die Anzahl der stabilen und instabilen Gleichgewichte. So erkennt man in Abbildung 3.9 dass nun auch drei Gleichgewichte möglich werden, von denen das mittlere bei $\lambda = 0,5$ instabil ist und die beiden äußeren, die aber keine Extreme darstellen, nun stabil sind. Dies ist aber bei sehr geringen lokalen Präferenzen (hier $s = 0,005$) nicht mehr der Fall.

3.3 Asymmetrische Regionen ohne lokale Präferenzen

Das Core-Periphery-Modell geht, wie die meisten Modelle der NEG auch, von symmetrischen Regionen aus. Dies geschieht, um das Modell möglichst allge-

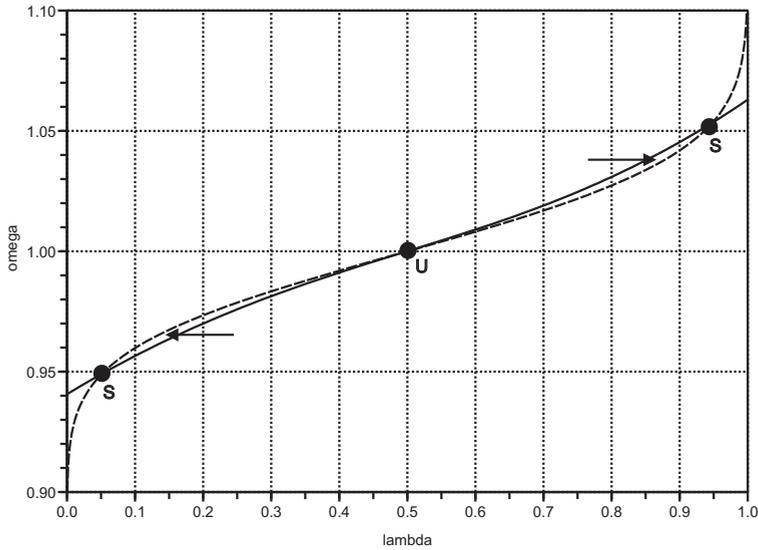


Abbildung 3.8: Gleichgewichte bei $T = 1,45$ und $s = 0,01$

meingütig zu halten und um explizit von externen Unterschieden (*first nature*) abzusehen. Als Ergebnis ergibt sich entweder die Gleichverteilung ökonomischer Aktivität oder die vollständige Konzentration der Industrie in einer Region. Dabei macht das Modell keine Aussage darüber, in welcher Region sich nun die Industrie ansiedelt. Um dies zu klären, ist es nun jedoch sinnvoll, exogene Unterschiede zu berücksichtigen. Daher soll hier ein Modell mit zwei unterschiedlich großen Ländern betrachtet werden um zu untersuchen, wie sich die Gleichgewichte aus Abschnitt 3.1 verändern.

In den hier betrachteten Modellen ist die Größe einer Region durch deren Angebot an Arbeit und der daraus resultierenden Kaufkraft definiert. Ein Maß für andere Größenangaben, z.B. der räumlichen Größe existiert nicht. Da die industrielle Arbeitskraft mobil ist, eignet sie sich aber wenig, um die Größe einer Region zu charakterisieren. Daher werden Größenunterschiede durch eine Differenz des immobilien Faktors, der landwirtschaftlichen Arbeit, modelliert. Dadurch verändert sich das Einkommen der Region, was wiederum in die Nachfrage nach Gütern beider Sektoren eingeht.

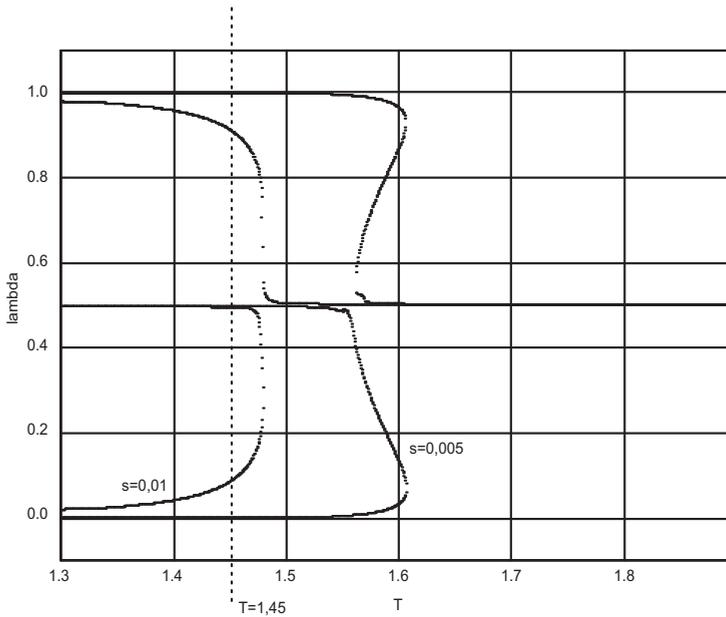


Abbildung 3.9: Bifurkationsdiagramm bei $s = 0,01$ und $s = 0,005$

Wenn wir annehmen, dass Region 1 größer ist als Region 2, so verändern sich Gleichungen (3.1) und (3.2) aus dem Standard-CP-Modell zu

$$Y_1 = \mu \lambda w_1 + \varepsilon \cdot \frac{1 - \mu}{2} \quad (3.14)$$

$$Y_2 = \mu (1 - \lambda) w_2 + \frac{1 - \mu}{2} \quad (3.15)$$

mit $\varepsilon \geq 1$, wobei $\varepsilon = 1$ den Fall symmetrischer Länder darstellt. An dieser Stelle sei daran erinnert, dass die Mengeneinheiten der Arbeitskräfte so festgesetzt wurden, dass $L^M = \mu$ und $L^A = 1 - \mu$ ist. Indirekt werden dadurch auch die anderen Gleichungen des Systems verändert, in die das Einkommen eingeht. Auf das Arbeitsangebot hingegen hat diese Veränderung keinen Einfluss, denn hier wird nur erklärt, wie die mobilen Arbeitskräfte auf Lohnsatzunterschiede reagieren. Alternativ könnte man auch die "2" im Nenner ändern. Die Verwendung eines Faktors (hier ε) hat den Vorteil, dass dieser als Prozentwert interpretiert werden kann. Bei $\varepsilon = 1,1$ beispielsweise ist Region 1 10% größer als Region 2. Führt man die Simulationen bei einem kleinen Unterschied in der Größe von 5% durch, so ergibt sich

folgendes Bild (Abbildung 3.10)⁹³:

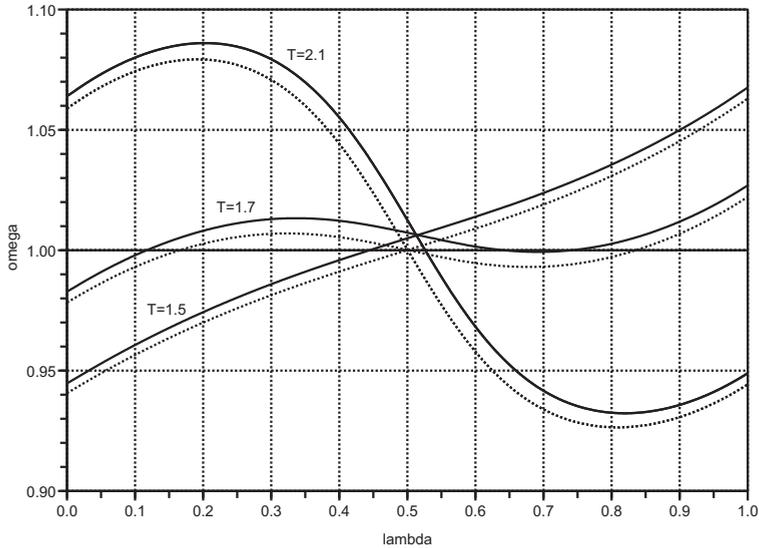


Abbildung 3.10: Gleichgewichte bei $\varepsilon = 1.05$ und $s = 0$

Im Fall geringer Transportkosten ($T = 1,5$) liegen die beiden stabilen Gleichgewichte bei vollständiger Konzentration der Industrie in einer der beiden Regionen. Dies unterscheidet sich nicht vom Standardmodell. Allerdings ist der Unterschied der Lohnsätze für $\lambda = 0$ geringer und für $\lambda = 1$ höher als im Standardmodell, was daran liegt, dass das Einkommen in Region 1 höher ist, da dort mehr landwirtschaftliche Arbeitskräfte leben. Dies erhöht über die Lohngleichung auch den Reallohn der Industriearbeiter, so dass $\omega = \frac{\omega_1}{\omega_2}$ steigt. Der gleiche Effekt führt auch dazu, dass das instabile Gleichgewicht nun nicht mehr bei $\lambda = 0,5$ liegt, sondern darunter. Der Ausgleich der Löhne ($\omega_1 = \omega_2$) wird nun bereits bei einer geringeren Wanderung der Arbeiter erreicht.

Im Fall hoher Transportkosten findet, genau wie bei symmetrischen Ländern, in beiden Regionen Produktion statt. Da nun in Region 1 mehr Arbeitnehmer leben und daher auch die Nachfrage nach Industrieprodukten höher ist, sorgt das Lohn-

⁹³Die gepunkteten Linien sind die des symmetrischen Falls, die hier zum Vergleich mit eingezeichnet wurden.

satzdifferential dafür, dass das stabile Gleichgewicht bei $\lambda > 0,5$ liegt. Gedanklich wird dies klar, wenn man extrem hohe Transportkosten annimmt, so dass in jeder Region die dort konsumierten Produkte gefertigt werden müssen.

Nimmt man nun Transportkosten im mittleren Bereich (hier $T = 1,7$) an, so zeigt sich die Asymmetrie besonders deutlich. Zum einen liegen die instabilen Gleichgewichte in dem Bereich, in dem es insgesamt zu 5 Gleichgewichten kommt, nicht mehr symmetrisch von $\lambda = 0,5$ entfernt, zum anderen kann es jetzt auch zu Situationen mit drei Gleichgewichten kommen, von denen zwei stabil sind. Dies liegt daran, dass sich (bei steigenden Transportkosten) Agglomerationen in der größeren Region durch die höhere Nachfrage länger halten können als in der kleineren Region. Dieser Effekt wird in Abbildung 3.11 noch deutlicher. Hier ist auf der horizontalen Achse der Transportkostensatz und auf der vertikalen Achse die Verteilung der industriellen Arbeiter λ abgetragen.

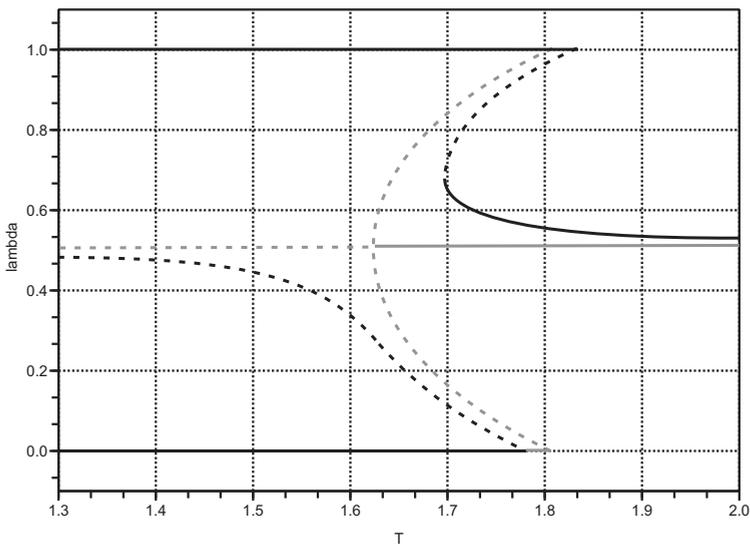


Abbildung 3.11: CP-Modell mit $\varepsilon = 1,05$

Geht man zunächst von geringen Transportkosten aus, so ist zu sehen, dass die gestrichelte Linie, die das instabile Gleichgewicht darstellt, weiter unten liegt als im symmetrischen Fall (grau dargestellt). Da die Fläche über dieser Linie dadurch

größer wird, steigt die Anzahl von Ausgangsverteilungen, die zu einem stabilen Gleichgewicht in der größeren Region führen. Steigen die Transportkosten an, wird dieser Effekt noch größer. Während im Standardfall das Gleichgewicht bei 0,5 immer besteht und am Break-Point vom instabilen zum stabilen Gleichgewicht wird, gilt dieser Zusammenhang hier nicht mehr.

Befindet sich das System bei geringen Transportkosten in einem stabilen Gleichgewicht, bei dem die gesamte industrielle Produktion in einer Region angesammelt ist, so ist ein solches Gleichgewicht bei steigenden Transportkosten länger stabil, wenn sich die Agglomeration in der größeren Region befindet. Ein einheitlicher Sustain-Point, wie im Standardmodell, besteht hier nicht. Zu begründen ist das damit, dass der Heimatmarkt größer ist und die steigenden Transportkosten, die sich ja nur auf die Exporte niederschlagen, daher später ihre agglomerationshemmende Wirkung entfalten.

Das wichtigste Ergebnis ist jedoch, dass bei sinkenden Transportkosten das stabile Gleichgewicht bei einer gleichverteilten Industrie immer in eine Agglomeration in der größeren Region übergeht. Aus der Grafik kann man dies erkennen, da die durchgezogene Linie, die das stabile Gleichgewicht kennzeichnet bei sinkenden Transportkosten in der Nähe des Break-Points nach oben gebogen ist und sich oberhalb der unteren gestrichelten Linie befindet. An dieser Stelle springt das System also vom stabilen Pfad bei ca. $\lambda = 0,65$ zur Agglomeration bei $\lambda = 1$. Ökonomisch kann das damit begründet werden, dass die Industrie in Region 1 hier bereits größer ist und in immer stärkerem Maße Skaleneffekte realisieren kann. Sinken die Transportkosten unter den Wert, der eine Produktion in beiden Ländern notwendig macht, wird nur in einem produziert und von dort aus die andere Region beliefert. Dies geschieht aufgrund des größeren Heimatmarktes.

Hieraus wird deutlich, dass die größere Region wegen des größeren Heimatmarktes im Vorteil ist. Bei den hier gewählten Werten, sind diese Effekte bereits bei einer Größendifferenz von 5% relativ stark.

3.4 Asymmetrische Regionen mit lokalen Präferenzen

In diesem Abschnitt sollen die beiden Modellvarianten aus 3.2 und 3.3 integriert werden. Das resultierende Modell enthält also sowohl Größenunterschiede zwi-

schen den Regionen als auch eingeschränkte Mobilität der Arbeitskräfte. Als Gleichungssystem für die Arbeitsnachfrage erhalten wir

$$Y_1 = \mu\lambda w_1 + \varepsilon \cdot \frac{1-\mu}{2} \quad (3.16)$$

$$Y_2 = \mu(1-\lambda)w_2 + \frac{1-\mu}{2} \quad (3.17)$$

$$Q_1 = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (3.18)$$

$$Q_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (3.19)$$

$$w_1 = \left[Y_1 Q_1^{\sigma-1} + Y_2 T^{1-\sigma} Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (3.20)$$

$$w_2 = \left[Y_1 Q_1^{\sigma-1} T^{1-\sigma} + Y_2 Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (3.21)$$

$$\omega_1 = w_1 Q_1^{-\mu} \quad (3.22)$$

$$\omega_2 = w_2 Q_2^{-\mu} \quad (3.23)$$

Ein langfristiges Gleichgewicht ist erreicht, wenn alle durch Reallohnunterschiede hervorgerufenen Wanderungen abgeschlossen sind. Dies ist der Fall, wenn

$$\varphi_i \omega_1 = (1 - \varphi_i) \omega_2 \quad (3.24)$$

erfüllt ist, wobei φ_i wie in Abschnitt 3.2 definiert ist. Zum Verständnis der Zusammenhänge ist es an dieser Stelle hilfreich, erst das Bifurkationsdiagramm und dann die dazugehörigen Liniendiagramme zu betrachten. Bei der Simulation wird, analog zu den anderen Abschnitten, natürlich umgekehrt verfahren.

Das Bifurkationsdiagramm in Abbildung 3.12, bei dem die Transportkosten auf der horizontalen und die Aufteilung der Arbeiter auf die Regionen auf der vertikalen Achse abgetragen werden, zeigt alle stabilen und instabilen Gleichgewichte. Vergleicht man es mit denen der vorangegangenen beiden Abschnitte, so ist zu erkennen, dass es die Ergebnisse beider Modellvarianten vereint. Einerseits sind, bedingt durch die lokalen Präferenzen, stabile Gleichgewichte möglich, bei denen ein großer Teil, aber eben nicht die gesamte Industrieproduktion in einer Region angesiedelt ist. Andererseits wird aber auch deutlich, dass sich das System bei

fallenden Transportkosten auf einem stabilen Gleichgewichtspfad von einer symmetrischen Aufteilung hin zur Agglomeration in der größeren Region bewegt. Dies geschieht, und das ist ein deutlich realistischeres Ergebnis als der Verlauf in Abschnitt 3.3, ohne Sprungstellen. In dem Modell führen die lokalen Präferenzen zu dem glatten Verlauf der stabilen Gleichgewichte und der Größenunterschied dazu, dass sich ohne wirtschaftspolitische Eingriffe die Agglomeration immer in der größeren Region bildet.

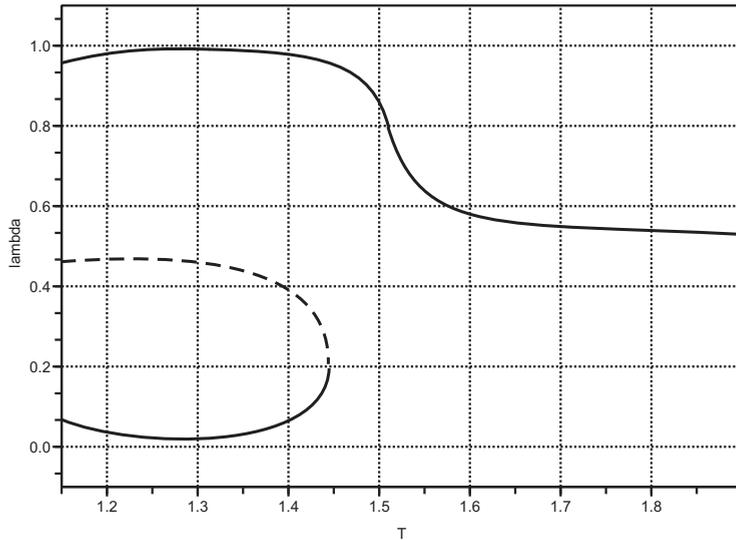


Abbildung 3.12: LW-Modell bei $s = 0,01$ und $\varepsilon = 1,05$

Betrachtet man nun an einer beliebigen Stelle (hier z.B. bei $T = 1,5$) das dazugehörige Diagramm mit den kurzfristigen Gleichgewichten (Abb. 3.13), werden zwei Dinge deutlich: Zum einen ergibt sich ein, im Vergleich zu den bisherigen Abbildungen, sehr asymmetrisches Bild. Während die Arbeitsangebotskurve (gestrichelte Linie) weiterhin bei $\omega = 1$ einen Wert von $\lambda = 0,5$ erreicht und einen zwar gekrümmten aber symmetrischen Verlauf aufweist, ist die Arbeitsnachfrage (durchgezogene Linie) deutlich asymmetrisch. Beides führt dazu, dass prinzipiell (abhängig von den jeweiligen Parameterwerten) sehr viele unterschiedliche Gleichgewichte denkbar sind und man hier nicht mehr mit irgendeiner Art von Symmetrie argumentieren kann. Zum anderen sieht man, wenn man die Anpassungskräfte im Arbeitsmarkt betrachtet, dass der hier gezeigte Schnittpunkt stabil ist. Rechts davon liegt der geforderte Lohn in Region 1 über dem gebotenen

und durch eine Abwanderung sinkt λ , links davon liegt der gebotene Lohn über dem geforderten und daher steigt λ . Wenn man bei diesem Gleichgewicht aber das Lohnsatzverhältnis ω betrachtet, so erkennt man, dass der Lohn in der größeren Region 1 über dem in Region 2 liegt ($\omega > 1$). Es bleibt also auch im stabilen Gleichgewicht ein Lohnsatzgefälle bestehen.

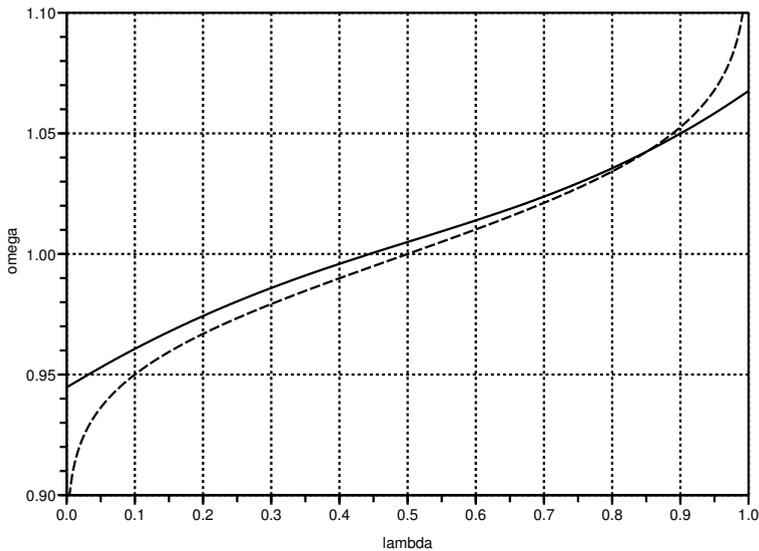


Abbildung 3.13: Lohnverhältnis bei $s = 0,01$, $\varepsilon = 1,05$ und $T = 1,5$

Es ist nun möglich, durch verschiedene Parameterwerte sowohl für die Mobilität der Arbeiter (s) als auch für die Mobilität der Güter (T) und zusätzlich für den Größenunterschied der Regionen (ε) verschiedenste Situationen zu untersuchen. Beispielhaft dafür wird in Abschnitt 6.2 dieses Modell verwendet, um die Entwicklung in Deutschland nach 1989 zu erklären. Dabei wird, anders als in der Literatur zur Neuen Ökonomischen Geographie üblich, nicht nur die Mobilität der Güter, sondern auch die der Arbeitskräfte variiert.

3.5 Zusammenfassung

In diesem Kapitel wurden zwei Modelle der Neuen Ökonomischen Geographie vorgestellt, die sich auf die Agglomerationsbildung auf regionaler Ebene beziehen und um Asymmetrien in den Regionsgrößen erweitert. Eines dieser Modelle ist das Core-Periphery-Modell, das einerseits das Standardmodell der gesamten NEG ist und andererseits auch das wichtigste Modell mit regionalem Fokus darstellt. Das andere ist eine Abwandlung des CP-Modells, bei dem berücksichtigt wird, dass einige Arbeitnehmer eine Präferenz für die Region haben, in der sie leben, und daher für den Nutzenverlust im Falle eines Umzugs durch einen Lohnaufschlag entschädigt werden müssen. Diese Erweiterung ist einerseits realitätsnah, andererseits bietet sie eine sehr einfache Möglichkeit, das "Katastrophenergebnis" des CP-Modells, das häufig kritisiert wurde, zu vermeiden.

Beide Modelle wurden in diesem Kapitel um die Annahme unterschiedlich großer Regionen erweitert. Das Ergebnis ist, dass in beiden Fällen sicher vorausgesagt werden kann, dass sich im Falle sinkender Transportkosten die Agglomeration in der größeren Region bildet. Im Modell mit regionalen Präferenzen ist dieser Anpassungsprozess bei bestimmten Parameterkonstellationen sogar ohne Sprünge möglich. Dies erlaubt auch das dauerhafte Bestehen von Lohnsatzdifferenzen. Insbesondere für wirtschaftspolitische Überlegungen, wie in Kapitel 6, bietet dies eine gute Grundlage.

4 Modelle für internationale Agglomeration

Sowohl das Core-Periphery-Modell als auch seine Erweiterung setzen die Mobilität von Arbeitskräften voraus. Diese ist jedoch in den Ländern Europas, und insbesondere im weniger qualifizierten Bereich als sehr gering zu bezeichnen. Das Modell kann daher kaum eine befriedigende Erklärung für Konzentrationsprozesse auf internationaler Ebene geben. In diesem Kapitel werden die beiden wichtigsten Modelle vorgestellt, die ohne Faktormobilität auskommen. Auch an diesen werden verschiedene Asymmetrieannahmen untersucht.

Das *Intermediate-Goods-Model (IM)* geht auf Krugman und Venables (1995) zurück. Es wird ausgegangen von zwei Ländern, zwei Sektoren (einem Agrarsektor und einem industriellen Sektor) und Arbeitskräften, die sektoral mobil aber regional immobil sind. Im Gegensatz zum Grundmodell werden hier also zwei Länder untersucht, im Sinne von Nationalstaaten. Durch diese Annahme fällt die Mobilität der Arbeitskräfte als Ausgleichsmechanismus weg. Dieser wird durch die Nachfrage der Industrie nach Zwischenprodukten ersetzt. Auch dadurch hat eine Agglomeration einen Vorteil, weil hier eine größere Menge produziert wird.⁹⁴ Es sind dann mehr Zwischenprodukte verfügbar und diese haben einen geringeren Preis, weil für Zwischenprodukte aus dem eigenen Land keine Transportkosten anfallen. Wegen geringerer Produktionskosten und einem großen Absatzmarkt wird dieser Standort dann attraktiver.

Das hier betrachtete Modell von Krugman und Venables (1995) stellt den wichtigsten Ansatz dar, die Ergebnisse des CP-Modells auch ohne Faktormobilität zu erklären. Aus diesem Grund wurde er auch hier als Basis für die Untersuchung

⁹⁴Im Modell wird nicht in Zwischen- und Endprodukte unterschieden, wie dies z.B. bei Venables (1996) der Fall ist. Das gleiche Aggregat, das konsumiert wird, fließt auch als Vorprodukt wieder in die Produktion zurück.

gewählt. Fujita und Hamaguchi (2001) stellen ebenfalls ein Modell vor, das Zwischenprodukte verwendet, diese werden jedoch von einem speziellen Sektor hergestellt, so dass der Erklärungsansatz eher dem Cluster-Modell (s.u.) entspricht. Gleiches gilt auch für das in Fujita und Thiesse (2002), Kap.11 dargestellte Modell. Puga (1999) hingegen vergleicht in einem ähnlichen Modell, welche Änderungen sich ergeben, wenn man alternativ von mobilen oder immobilen Arbeitskräften ausgeht.

Das zweite hier verwendete Modell ist das Cluster-Modell (CM) von Krugman und Venables (1996). Es kommt ebenfalls ohne Faktormobilität aus. Industriekonzentrationen werden durch vertikale Verflechtungen zwischen Unternehmen erklärt. So entstehen auch hier Forward- und Backward-Linkages. Im Gegensatz zum Modell mit Zwischenprodukten gibt es hier aber keinen Sektor mit vollkommener Konkurrenz mehr. Dieser wird durch einen zweiten Industriesektor ersetzt. Dadurch entsteht in dem Modell auch keine Kern-Periferie Struktur mehr. Vielmehr wird die Verteilung von zwei Sektoren auf zwei Länder erklärt.

Das Modell ist eine Weiterentwicklung des Modells von Krugman und Venables (1995), das wiederum auf Krugman (1991b) aufbaut. Venables (1999) stellt eine Modifikation des CM dar, in der statt zwei nun n Sektoren angenommen werden. Dafür verwendet jeder Sektor nur die eigenen Zwischenprodukte, ein intersektoraler Handel findet also nicht statt. Fujita und Hamaguchi (2001) modellieren ebenfalls Zwischenprodukte im Rahmen der NEG. Sie tun dies jedoch in einem stadtökonomischen Kontext mit eindimensionalem Raum und verwenden neben einem Agrar- und einem Industriesektor einen eigenen Sektor mit Zwischenprodukten. Die Ergebnisse sind daher nicht vergleichbar. Ein weiteres Modell dieser Klasse mit Zwischenprodukten wurde von Venables (1996) entwickelt. Er verwendet dabei drei Sektoren, einen Agrarsektor und zwei Industriesektoren, wobei einer davon Vorprodukte für den anderen herstellt, der dann ausschließlich Konsumprodukte auf den Markt bringt. Genau wie im vorliegenden Beitrag, kommt es auch in diesem Modell zu stabilen Gleichgewichten bei Spezialisierung, die nahe dem Sustain-Point knapp oberhalb von 0 bzw. knapp unterhalb von 1 liegen. Wegen der doppelten Symmetrie in Venables (1996) kommen diese Abweichungen auf beiden Seiten gleichzeitig vor, wie dies z.B. bei den hier dargestellten Überlegungen zu unterschiedlich großen Ländern (Abschnitt 4.5) der Fall ist. Allen Modellen mit Vorprodukten ist die Annahme gemeinsam, dass das CES-Aggregat, das in die Produktion eingeht, demjenigen im Konsum entspricht. Eine

Ausnahme bildet hier nur Gallo (2006) der allerdings die Adaption von technischem Fortschritt in einem ähnlichen Modell zum Inhalt hat.

Aus empirischer Sicht ist der Erklärungsansatz des Modells mit Zwischenprodukten und des Cluster-Modells, zumindest im Bereich der Industrieländer, besser zu begründen als der des Core-Periphery-Modells. Denn während die Mobilität des Faktors Arbeit international als eher gering einzuschätzen ist, hat der Handel mit Zwischenprodukten stetig zugenommen. Dies gilt insbesondere für den intraindustriellen Handel, was die Annahme des Clustermodells stützt, die besagt, dass ein Sektor einen relativ hohen Anteil der Vorprodukte aus der eigenen Produktion bezieht.⁹⁵

Nachdem im vorangegangenen Kapitel untersucht wurde, wie sich die Annahme unterschiedlich großer Länder auf die Agglomerationsbildung bei mobilen Arbeitskräften auswirkt, soll diese nun auch bei immobilen Arbeitskräften, also im Rahmen von Modellen der internationalen Agglomerationsbildung betrachtet werden. Dazu wird in den Abschnitten 4.2 und 4.5 sowohl das Modell mit Zwischenprodukten als auch das Cluster-Modell durch diese Asymmetrie erweitert. Da das Cluster-Modell zwei Industriesektoren enthält, lassen sich hier noch zwei weitere Asymmetrien untersuchen: Sektorspezifische Transportkosten (Abschnitt 4.4) und Unterschiede in der Größe der Sektoren (Abschnitt 4.6), die durch den Ausgabenanteil bestimmt wird. Diese Annahmen können auch miteinander kombiniert betrachtet werden, was Gegenstand von Abschnitt 4.7 ist.

4.1 Modell mit Zwischenprodukten

Vor der analytischen Darstellung des Modells sollen die zentralen Annahmen genannt werden: Wir betrachten zwei Länder, die zunächst als identisch angesehen werden. In beiden Ländern gibt es zwei Sektoren: Einen Agrarsektor und einen Industriesektor. Die Sektoren sind strukturell genau so modelliert, wie im CP-Modell. Das Agrargut wird mit konstanten Skalenerträgen bei vollkommener Konkurrenz hergestellt. Einziger Produktionsfaktor ist Arbeit, die ausschließlich als variabler Faktor in die Produktion einfließt. Das Industriegut wird aus Arbeit und Vorprodukten hergestellt. Dabei fließt ein Teil der industriellen Güter

⁹⁵Vgl. dazu Abschnitt 2.6 und 2.7.

als Input-Faktor in den Produktionsprozess ein. Die Zusammensetzung wird als identisch mit der Zusammensetzung des konsumierten Industriegutes angesehen. Die Präferenzen der Haushalte werden durch eine Cobb-Douglas-Nutzenfunktion dargestellt, wobei das industrielle Gut mit einer Sub-Nutzenfunktion vom CES-Typ einfließt. Arbeit ist räumlich immobil, kann aber frei zwischen den Sektoren wechseln. Es wird angenommen, dass Arbeitnehmer in den Sektor mit dem höheren Lohn wechseln.

Nachfrage Da die Struktur der Nachfrage identisch mit dem Grundmodell ist, wird auch hier die gleiche Nutzenfunktion verwendet

$$U = M^\mu \cdot A^{1-\mu} \quad (4.1)$$

Dabei ist M das differenzierte Produkt und A das einheitliche Produkt, auch als Agrarprodukt bezeichnet, das unter den Annahmen der vollkommenen Konkurrenz mit konstanten Skalenerträge produziert wird. Es wird angenommen, dass zur Produktion von A nur Arbeit eingesetzt wird und das Gut ohne Transportkosten im- und exportiert werden kann. Dieses Gut dient als Numeraire. Der Parameter μ gibt den Ausgabenanteil der Haushalte für das Industriegut an.

M wird durch eine Subnutzenfunktion vom CES-Typ dargestellt:

$$M = \left[\int_0^{n_h} x_{hi}^{\frac{\sigma-1}{\sigma}} di + \int_0^{n_f} x_{fi}^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}} \quad (4.2)$$

Das Aggregat besteht sowohl aus den im Inland hergestellten Mengen x_{hi} , als auch aus den im Ausland hergestellten Mengen x_{fi} , n_h und n_f sind die Anzahl der Produktvarianten oder Unternehmen in beiden Ländern. Nun werden der Preisindex und die individuelle Nachfrage bei monopolistischer Konkurrenz berechnet. Zur Herleitung der folgenden Gleichungen siehe Abschnitt 2.8. Es ergeben sich für den Preisindex die Gleichung

$$Q_h = \left[n_h p_h^{1-\sigma} + n_f (p_f T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.3)$$

und für die Nachfrage nach heimischen Gütern

$$m_h = \mu Y_h \frac{p_h^{-\sigma}}{Q_h^{1-\sigma}} + \mu Y_f \frac{p_f^{-\sigma}}{Q_f^{1-\sigma}} T^{1-\sigma} \quad (4.4)$$

Angebot Im Gegensatz zum Grundmodell wird in der Produktion neben Arbeitskraft auch ein Vorprodukt als Produktionsfaktor verwendet. Die Technologie wird dabei in der Gesamtkostenfunktion

$$TC = w^\beta Q^\alpha \cdot (F + cq^*), \quad \beta + \alpha = 1 \quad (4.5)$$

zusammengefaßt. Dabei sind $w^\beta \cdot Q^\alpha$ die Kosten für eine Einheit eines aus Arbeit und Vorprodukten zusammengesetzten Produktionsfaktors. Es fallen fixe Kosten F und variable Kosten c an. Diese Kostenfunktion impliziert drei Annahmen:

1. Das Aggregat M , das in der Produktion verwendet wird, ist das gleiche, das auch konsumiert wird. Mit anderen Worten: Die Produktion verwendet als Vorprodukt eine Anzahl differenzierter Produkte, die in ihrer quantitativen Zusammensetzung auch genau so konsumiert werden. α und β können dann als (konstante) Produktionselastizitäten angesehen werden.
2. Der fixe und der variable Produktionsfaktor sind gleich aus Arbeit und Vorprodukten zusammengesetzt.
3. Der zusammengesetzte Produktionsfaktor wird mit konstanten Skalenerträgen produziert. Dies ist notwendig, um zu den konstanten Stückkosten zu gelangen. Denn nur dann lässt sich diese Kostenfunktion (leicht vereinfacht) aus einer Cobb-Douglas Produktionsfunktion ableiten.⁹⁶

Leitet man (4.5) nach der Menge ab, erhält man die Grenzkosten

$$C' = w^\beta \cdot Q^\alpha \cdot c$$

⁹⁶Vgl. Jehle und Reny (2001), S. 131. Die Produktionsfunktion selbst taucht, wie in den meisten anderen Modellen der Neuen Ökonomischen Geographie, nicht auf. Stattdessen wird mit der dualen Kostenfunktion gearbeitet, wobei allerdings der konstant Term $\alpha^\alpha \cdot (1 - \alpha)^{1-\alpha}$ zur Vereinfachung weggelassen wird. Daher muss auch $\alpha + \beta = 1$ gelten, da die Homogenität der Produktionsfunktion eins geteilt durch die Summe der Produktionselastizitäten entspricht.

und daraus nach der Preissetzungsregel im Monopol⁹⁷ den Preis

$$p = w^\beta Q^\alpha \cdot c \cdot \frac{\sigma}{\sigma - 1} \quad (4.6)$$

In der Marktform der monopolistischen Konkurrenz erwirtschaften die Unternehmen einen Gewinn von 0. Um die produzierte Menge zu bestimmen, wird daher zunächst die Gewinnfunktion aufgestellt, die null ergeben muss.

$$\pi_i = p_i q_i - w^\beta \cdot Q^\alpha \cdot (F - c q_i) = 0 \quad (4.7)$$

In diese wird der Preis (4.6) eingesetzt

$$\pi_i = w^\beta \cdot Q^\alpha \cdot c \cdot \frac{\sigma}{\sigma - 1} q_i - w^\beta \cdot Q^\alpha \cdot F - w^\beta \cdot Q^\alpha \cdot c q_i = 0 \quad (4.8)$$

und diese Gleichung vereinfacht

$$\pi = w^\beta Q^\alpha \left[c q_i \left(\frac{\sigma}{\sigma - 1} - 1 \right) - F \right] = 0$$

Wenn man nun $1 = \frac{\sigma-1}{\sigma-1}$ ersetzt und die Gleichung durch den Kostenterm teilt, bekommt man die optimale Menge bei einem Gewinn von 0

$$q_i^* = (\sigma - 1) \frac{F}{c} \quad (4.9)$$

Zu beachten ist, dass die optimale Menge auch in diesem Modell konstant ist und insbesondere nicht von Variablen der Nachfrageseite abhängt. Daraus kann bereits hier der Schluss gezogen werden, dass eine Ausweitung der gesamten angebotenen Menge, z.B. wegen eines Anstiegs der Nachfrage, nur durch eine Erhöhung der Anbieterzahl realisiert werden kann. Der einzelne Produzent bietet aber immer die gleiche Menge an.⁹⁸

⁹⁷Die umgeformte Amoroso-Robinson-Relation lautet: $p = C' \cdot \frac{\sigma}{\sigma-1}$.

⁹⁸Vgl. hierzu auch Gleichung (4.13).

Standardisierung Nun wird die übliche Standardisierung durchgeführt mit $c = \frac{\sigma-1}{\sigma}$, wonach sich (4.6) vereinfacht zu

$$p = w^\beta \cdot Q^\alpha \quad (4.10)$$

und (4.9) zu

$$q^* = F\sigma \quad (4.11)$$

Weiterhin werden die Einheiten so gewählt, dass $F = \frac{1}{\beta\sigma}$ gilt. Dadurch vereinfacht sich (4.11) weiter zu

$$q^* = \frac{1}{\beta} \quad (4.12)$$

Preisindex Die Summe der Löhne der Industriearbeiter in einem Land ist $L \cdot w$.⁹⁹ Der Wert der gesamten industriellen Produktion ist $n \cdot p \cdot q^*$. Da die Unternehmen keinen Gewinn machen, entspricht die Lohnsumme dem Anteil β des Wertes der produzierten Menge. Daher gilt

$$wL = \beta n p q^*$$

Setzt man hier die optimale Menge (4.12) ein und löst nach n auf, so kann die Menge der Unternehmen im Optimum bestimmt werden als

$$n^* = \frac{wL}{p} \quad (4.13)$$

Diese optimalen Werte für den Preis und die Anzahl der Unternehmen sollen nun verwendet werden, um die Preisindices zu bestimmen. Dazu werden (4.12) und (4.13) in (4.3) eingesetzt.

$$Q_h = \left[\frac{w_h L_h}{w_h^\beta Q_h^\alpha} \cdot \left(w_h^\beta Q_h^\alpha \right)^{1-\sigma} + \frac{w_f L_f}{w_f^\beta Q_f^\alpha} \cdot \left(w_f^\beta Q_f^\alpha T \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.14)$$

⁹⁹Die Anzahl der Arbeitnehmer pro Region wurde hier auf 1 standardisiert. Dann kann L als Anteil interpretiert werden und die Lohnsumme ist $L \cdot 1 \cdot w$.

Dies lässt sich vereinfachen zu

$$Q_h = \left[L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} + L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.15)$$

Ausgabengleichung Um die gesamten Ausgaben des Inlandes für Industrieprodukte zu bestimmen, wird zunächst die Gleichung für das gesamte Einkommen aufgestellt

$$Y = w \cdot L + 1 \cdot (1 - L) \quad (4.16)$$

Die Ausgaben setzen sich zusammen aus den Ausgaben der privaten Haushalte, die den Anteil μ ihres Einkommens ausgeben, und denen der Unternehmen, die das Industriegut als Vorprodukt einsetzen.

$$E = \mu Y + \alpha n p q \quad (4.17)$$

Dabei ergibt sich der Term für die Ausgaben der Unternehmen für Vorprodukte, weil diese bei q^* einen Gewinn von 0 erzielen. Dann entsprechen die Ausgaben der Unternehmen dem Anteil α am Gesamtwert der Produktion. Diese errechnet sich aus der Anzahl der Unternehmen, dem Preis und der Menge, die jedes Unternehmen produziert: $n^* \cdot q^* \cdot p$. Setzt man nun die optimalen Werte für die Anzahl der Unternehmen (4.13) und die Produktionsmenge (4.12) sowie die Formel für das Einkommen (4.16) der privaten Haushalte in diese Gleichung ein, so erhält man die Ausgabengleichung, hier für das Inland formuliert

$$E_h = \mu [w_h L_h + 1 - L_h] + \frac{\alpha w_h L_h}{\beta} \quad (4.18)$$

Lohngleichung Da das zusammengesetzte Industriegut, das als Vorprodukt in die Produktionsfunktion eingeht, annahmegemäß das gleiche ist, das auch konsumiert wird, können die in (4.4) hergeleiteten Nachfragefunktionen auch als Nachfragen der Unternehmen verstanden werden. Setzt man hier statt den privaten Ausgaben als Budget E_h ein, gelangt man zu

$$m_h = E_h \frac{p_h^{-\sigma}}{Q_h^{1-\sigma}} + E_f \frac{p_f^{-\sigma}}{Q_f^{1-\sigma}} T^{1-\sigma} \quad (4.19)$$

Dies muss im Optimum gleich der Menge aus (4.12) sein. Außerdem kann der optimale Preis aus (4.10) verwendet werden. Daraus erhalten wir die Lohngleichung.

$$\frac{\left[w_h^\beta Q_h^\alpha\right]^\sigma}{\beta} = E_h Q_h^{\sigma-1} + E_f Q_f^{\sigma-1} T^{1-\sigma} \quad (4.20)$$

die nach w_h aufgelöst werden kann

$$w_h = \left[\frac{\beta \left(E_h Q_h^\sigma Q_f + E_f Q_f^\sigma T^{1-\sigma} Q_h \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_h^{-\frac{\alpha}{\beta}} \quad (4.21)$$

Eine analoge Gleichung lässt sich für den Industrieloohn im Ausland aufstellen. Die Löhne in den beiden landwirtschaftlichen Sektoren sind auf 1 normiert.

Gleichgewicht Das Gleichungssystem, das den Gleichgewichtszustand beschreibt, besteht aus dem Preisindex (4.14), der Ausgabengleichung (4.18) und der Lohngleichung (4.21), jeweils formuliert für das Inland und das Ausland¹⁰⁰

$$Q_h = \left[L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} + L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.22)$$

$$Q_f = \left[L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} + L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.23)$$

$$E_h = \mu [w_h L_h + 1 - L_h] + \frac{\alpha w_h L_h}{\beta} \quad (4.24)$$

$$E_f = \mu [w_f L_f + 1 - L_f] + \frac{\alpha w_f L_f}{\beta} \quad (4.25)$$

$$w_h = \left[\frac{\beta \left(E_h Q_h^\sigma Q_f + E_f Q_f^\sigma T^{1-\sigma} Q_h \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_h^{-\frac{\alpha}{\beta}} \quad (4.26)$$

$$w_f = \left[\frac{\beta \left(E_f Q_f^\sigma Q_h + E_h Q_h^\sigma T^{1-\sigma} Q_f \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_f^{-\frac{\alpha}{\beta}} \quad (4.27)$$

¹⁰⁰Im Vergleich zum Cluster-Modell werden hier keine Gleichungen für den zweiten Sektor benötigt, da dieser wie im CP-Modell als Referenzsektor modelliert ist. Das Gleichungssystem besteht daher nur aus sechs statt aus 12 Gleichungen.

Auch dieses Gleichungssystem kann nicht analytisch gelöst werden, so dass auch hier Simulationen durchgeführt werden müssen, um zu einer Lösung zu gelangen. Der Anpassungsmechanismus in diesem Modell ist die Mobilität der Arbeitskräfte zwischen den Sektoren, nicht zwischen den Ländern.¹⁰¹ Da im Agrarsektor immer ein Lohnsatz von eins herrscht, wechseln Arbeiter vom Industriesektor in den Agrarsektor, wenn der Lohn im Industriesektor w unter 1 liegt und in die andere Richtung, wenn er über 1 liegt. Dementsprechend finden bei $w = 1$ keine Bewegungen statt. Die Kurve in Abbildung 4.1 ist der Ort aller Kombinationen von L_f und L_h , bei der $w_h = 1$ gilt. Dabei wird der Wert für L_f festgesetzt und das o.g. Gleichungssystem solange iterativ durchlaufen, bis $w_h = 1$ gilt. Links von der Kurve ist $w_h > 1$. Dadurch wechseln Arbeiter in den Industriesektor und L_h erhöht sich, was grafisch zu einer Bewegung nach rechts führt. Entsprechend gilt rechts von der Kurve $w_h < 1$, was zu einer Verringerung von L_h führt.

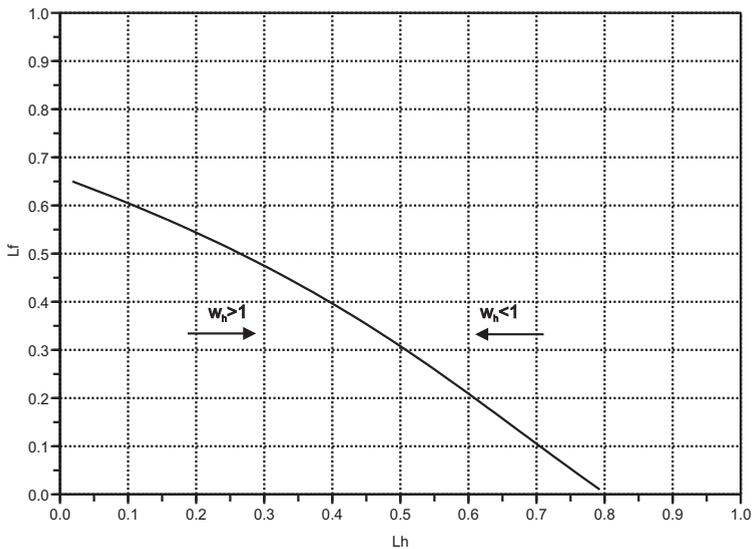


Abbildung 4.1: Inlandslinie im IM-Modell bei $w_h = 1$ und $\mu = 0,4$; $\alpha = 0,5$; $\sigma = 5$

¹⁰¹Daher wird in den folgenden Graphiken die Verteilung der Arbeitskräfte L_h und L_f verwendet. Außerdem kann das Modell dann besser mit den anderen verglichen werden. Sehr ähnliche Abbildungen lassen sich auch erzeugen, wenn stattdessen die Anzahl der Firmen n_h und n_f verwendet wird. Vgl. hierzu Venables (1995).

Eine entsprechende Kurve kann auch für das Ausland erzeugt werden. In diesem Fall ändern sich, da alle Parameter als gleich unterstellt werden, nur Abhängige und Unabhängige Variable. Die Auslandsline gibt ein L_f bei gegebenem L_h an, bei dem $w_f = 1$ gilt. Die Anpassungen finden grafisch gesehen dann nicht in horizontaler, sondern in vertikaler Richtung statt. Die Gleichgewichtskurve des Auslands ist in den folgenden Abbildungen gepunktet dargestellt. Wie auch in den anderen Modellen der NEG, lassen sich Fälle mit hohen, mittleren und niedrigen Transportkosten unterscheiden.

Im Fall $T = \infty$, als Extremfall hoher Transportkosten, ist die Inlandskurve eine vertikale Linie bei $L_h = \mu$ und die Auslandslinie entsprechend eine horizontale Linie bei $L_f = \mu$ (Abb. 4.2).

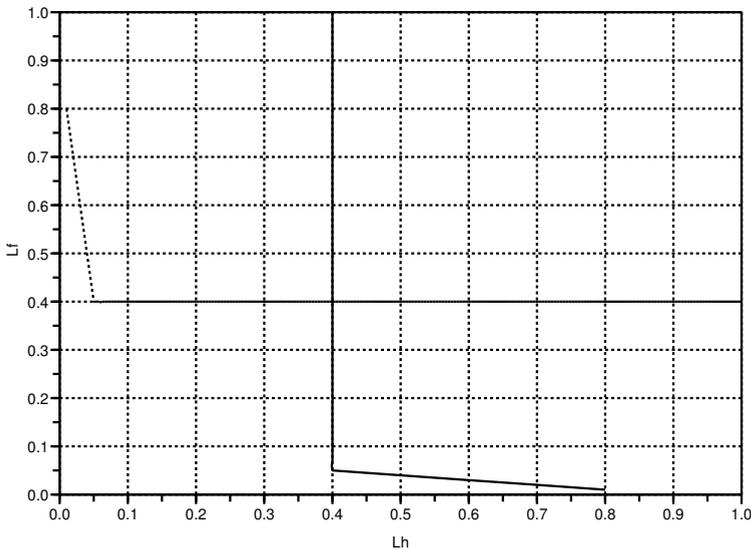


Abbildung 4.2: IM-Modell bei extrem hohen Transportkosten ($T = 1.000$)

Die beiden Länder sind vollständig autark und produzieren je einen Anteil von μ an industriellen Gütern und $1 - \mu$ an landwirtschaftlichen Gütern. Einzige Ausnahme sind die Stellen, an denen das jeweils andere Land gar keine Produktion des Industriegutes hat. In diesem Fall (z.B. $L_f = 0$ aus Sicht des Inlandes) muss das Land die gesamte Produktion übernehmen. Hier wäre dann $L_h = 2 \cdot \mu$. Da

dies jedoch kein Gleichgewicht ist, sind diese Abschnitte nicht relevant.

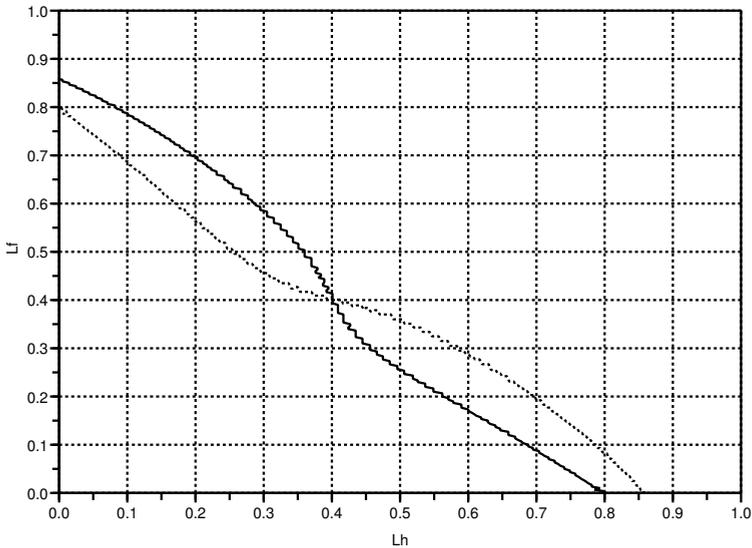


Abbildung 4.3: IM-Modell bei hohen Transportkosten ($T = 3,0$)

Auch im Fall von hohen, aber nicht utopisch hohen Transportkosten (z.B. $T = 3$) entsteht ein einziges stabiles Gleichgewicht bei $L_h = L_f = \mu$ (Abb. 4.3). Stabil ist dieses Gleichgewicht, weil links oberhalb des Schnittpunktes die Auslandskurve unter der Inlandskurve liegt. Aufgrund der oben beschriebenen Anpassungsmechanismen ergibt sich in diesem Bereich eine Bewegung nach rechts unten, in Richtung Schnittpunkt. Eine analoge Überlegung lässt sich für den Bereich rechts unten anstellen. Das System bewegt sich bei jeder denkbaren Anfangsaufteilung auf diesen Schnittpunkt zu, der daher ein stabiles Gleichgewicht darstellt.

Im hier unterstellten Fall von symmetrischen Ländern bestimmt die Ausgabenverteilung, die sich aus der Nutzenfunktion ergibt (Parameter μ), die Aufteilung der Arbeitskräfte auf die Sektoren. Es gibt keine Netto-Überschüsse im Handel zwischen beiden Ländern, ebenso gibt es keine Spezialisierung.

Bei niedrigen Transportkosten (hier $T = 1,5$; Abb. 4.4) verlaufen beide Linien mit negativer Steigung, allerdings liegt links vom Schnittpunkt die Auslandslinie über der des Inlands. Daraus folgt, dass das symmetrische Gleichgewicht nun instabil

ist.¹⁰² Dafür ergeben sich aber zwei weitere Gleichgewichte mit vollkommener Spezialisierung an den Achsen. Diese Gleichgewichte sind stabil und haben jeweils die Werte $L_h = 0; L_f = 2\mu$ oder $L_h = 2\mu; L_f = 0$. In der Abbildung 4.4 kann man dies erkennen, da bei $\mu = 0,4$ die Achsenabschnitte bei 0,8 liegen.

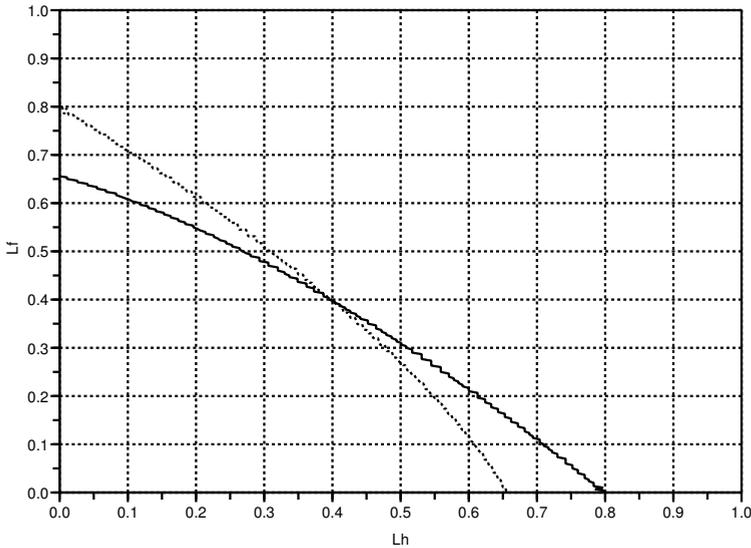


Abbildung 4.4: IM-Modell bei niedrigen Transportkosten ($T = 1,5$)

Ökonomisch kann dies dadurch begründet werden, dass die Vorteile der Agglomeration bei geringen Transportkosten überwiegen. Aufgrund der Spezialisierung einer Region auf Industrieproduktion können die Vorprodukte hier nun ohne Transportkosten verwendet werden. Die fertigen Produkte werden zur Hälfte in das andere Land exportiert, das nun nur noch Agrarprodukte herstellt.

Bei mittleren Transportkosten (hier $T = 2,15$) ergibt sich ein Bild mit unterschiedlichen möglichen Lösungen (Abbildung 4.5), da sich die Linien, für die $w_i = 1$ gilt, dreimal schneiden. Dadurch entsteht ein stabiles Gleichgewicht bei $L_h = L_f = \mu$. Außerdem sind die beiden Randlösungen mit vollständiger Spezialisierung ebenfalls stabile Gleichgewichte. Dazwischen entsteht noch jeweils ein weiteres, insta-

¹⁰²Zur Verdeutlichung kann man oben beschriebene Überlegungen für diesen Fall anstellen. Dann gelangt man zu Anpassungskräften, die nach links oben gerichtet sind.

biles Gleichgewicht. In diesem Fall entscheidet die Ausgangssituation darüber, welches stabile Gleichgewicht am Ende erreicht wird.

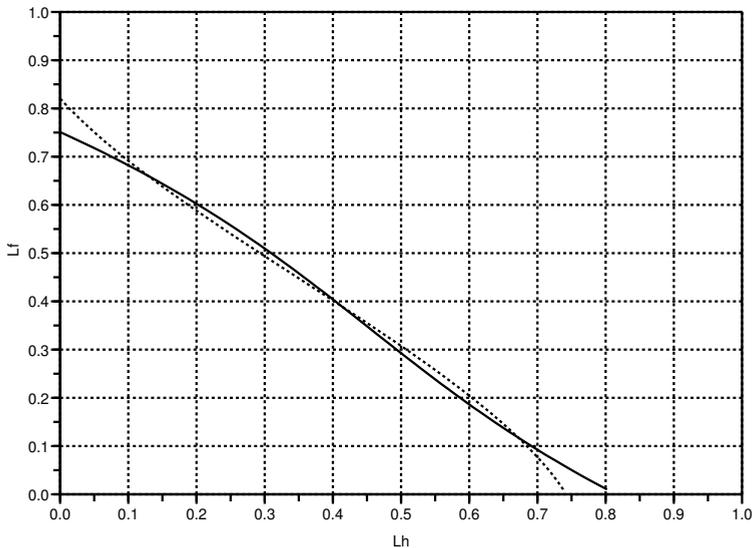


Abbildung 4.5: IM-Modell bei mittleren Transportkosten ($T = 2, 15$)

Agglomerative und Deglomerative Kräfte Auch in diesem Modell gibt es zwei Mechanismen, die zu Agglomeration führen (Forward- und Backward-Linkages) sowie einen Effekt, der die Bildung von Agglomerationen hemmt. Der Forward-Linkage bezeichnet den Vorteil einer großen Anzahl von Zwischenprodukten, die auf dem heimischen Markt produziert wird. Weil dafür keine Transportkosten anfallen, reduziert sich der Preisindex (4.15), der in die Kostenfunktion der Unternehmen (4.5) eingeht.¹⁰³ Der Gewinn steigt (zumindest vorübergehend) und es siedeln sich mehr Unternehmen an, die wiederum das Angebot an Zwischenprodukten erhöhen.

Begrenzt wird die Agglomerationsbildung durch den Wettbewerb auf dem Arbeitsmarkt. Wechseln immer mehr Arbeitnehmer in den Industriesektor, so fällt

¹⁰³Das ist anders als im CP-Modell. Dort läuft dieser Effekt über das Lohnniveau ab, denn der Preisindex tritt nicht in der Kostenfunktion auf.

dort der Lohn. Liegt dieser jedoch unter 1, ist ein Wechsel nicht mehr attraktiv. Damit dieser Mechanismus wirksam ist, muss μ jedoch kleiner als 0,5 sein. Ist das nicht der Fall, geben die Konsumenten mehr als die Hälfte ihres Einkommens für Produkte des Industriesektors aus. Aufgrund der steigenden Skalenerträge kann es dann dauerhaft zu Löhnen über eins kommen. Außerdem muss $\alpha > 0$ sein, damit es überhaupt zu Agglomerationen kommen kann.¹⁰⁴

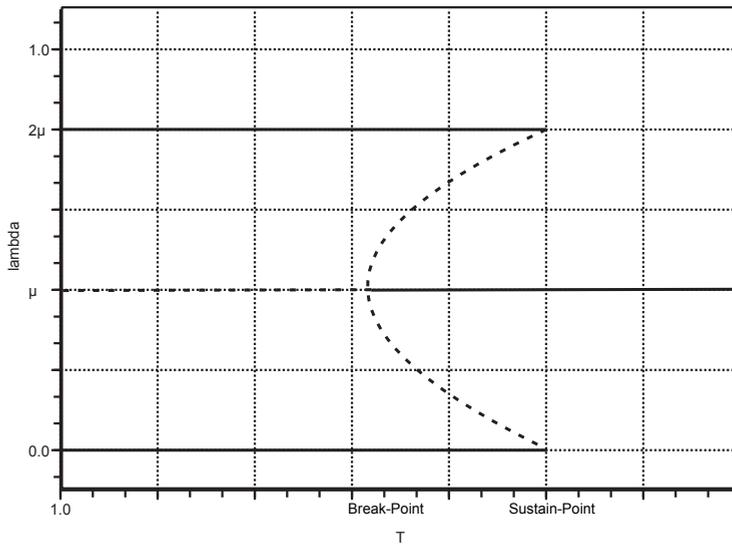


Abbildung 4.6: Bifurkationsdiagramm mit Zwischenprodukten

Auch dieses Modell kann in Form eines Bifurkationsdiagramms dargestellt werden. Das Ergebnis ist dabei das gleiche, wie im Core-Periphery-Modell, und genau das ist ja Ziel des Modells. Es ist also auch ohne die im CP-Modell unterstellte Mobilität der Arbeitskräfte möglich, Agglomerationen zu begründen. Bei ansonsten gleicher Modellstruktur wurden hier lediglich Vorprodukte eingeführt und damit die gleichen Kernaussagen generiert.

Bei den hier verwendeten Parametern, insbesondere $\mu = 0,4$, ergibt sich das in Abbildung 4.6 dargestellte Bild, wobei der Break-Point bei ca. 1,85 und der Sustain-Point bei ca. 2,4 liegt. Dabei kommt es bei geringen Transportkosten zu Spezialisierung, die allerdings nicht vollständig ist, da der Ausgabenanteil für Industrieprodukte zu gering ist, um die gesamte in einem Land vorhandene Ar-

¹⁰⁴Vgl. Klüver (2000), S. 132f.

beitskraft zu beschäftigen. Daher liegt der Anteil der im Inland im Industriesektor beschäftigten Arbeitskräfte bei 2μ . Das Land bedient also die Nachfrage beider Länder nach Industrieprodukten. Das ist anders als im CP-Modell. Da dort die gesamte Arbeitskraft in einem Land variabel ist und sich durch Wanderungen ändert, kommt es bei sehr geringen Transportkosten immer zur vollständigen Agglomeration, bei der ein Land nur landwirtschaftliche und das andere nur industrielle Güter herstellt. Im Modell mit Zwischenprodukten dagegen verbleibt auch ein Teil der landwirtschaftlichen Produktion in dem Land, das die gesamte Industrieproduktion beinhaltet, wenn μ kleiner als 0,5 ist. Der Fall, in dem $\mu > 0,5$ ist, wurde per Definition ausgeschlossen (siehe oben).

Unterstellt man hingegen hohe Transportkosten, so produziert jedes Land seinen Konsum an industriellen Gütern selbst, so dass $L_i = \mu$ gilt. Da der Break-Point unterhalb des Sustain-Points liegt, gibt es dazwischen einen Bereich mit fünf möglichen Gleichgewichten, genau wie im Core-Periphery-Modell.

4.2 Zwischenprodukte bei unterschiedlich großen Ländern

Nun soll, analog zu den Abschnitten 3.3 und 3.4, auch in diesem Modell der Fall unterschiedlich großer Länder untersucht werden. Dabei wird die Größe eines Landes an deren gesamter Menge an Arbeitskräften gemessen. Das ist in diesem Fall möglich, da der Faktor Arbeit regional immobil ist. Das Gleichungssystem, welches das kurzfristige Gleichgewicht beschreibt, hat dann die folgende Form

$$Q_h = \left[\varepsilon_h L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} + \varepsilon_f L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.28)$$

$$Q_f = \left[\varepsilon_f L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} + \varepsilon_h L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.29)$$

$$E_h = \mu [w_h \varepsilon_h L_h + \varepsilon_h (1 - L_h)] + \frac{\alpha w_h \varepsilon_h L_h}{\beta} \quad (4.30)$$

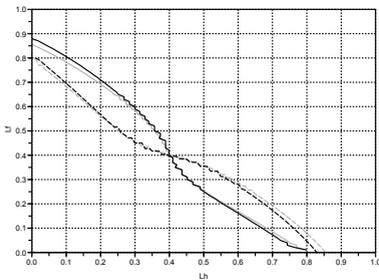
$$E_f = \mu [w_f \varepsilon_f L_f + \varepsilon_f (1 - L_f)] + \frac{\alpha w_f \varepsilon_f L_f}{\beta} \quad (4.31)$$

4.2 Zwischenprodukte bei unterschiedlich großen Ländern

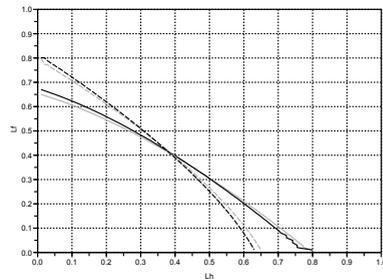
$$w_h = \left[\frac{\beta \left(E_h Q_h^\sigma Q_f + E_f Q_f^\sigma T^{1-\sigma} Q_h \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_h^{-\frac{\alpha}{\beta}} \quad (4.32)$$

$$w_f = \left[\frac{\beta \left(E_f Q_f^\sigma Q_h + E_h Q_h^\sigma T^{1-\sigma} Q_f \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_f^{-\frac{\alpha}{\beta}} \quad (4.33)$$

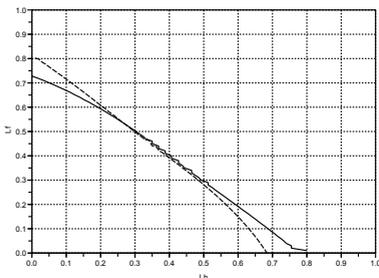
Dabei stehen die Faktoren ε_h und ε_f für die Größen der Länder. Möchte man also dieses Modell für einen Größenunterschied von 5% simulieren (das Inland ist 5% größer als das Ausland) so gelten diese Gleichungen mit $\varepsilon_h = 1,05$ und $\varepsilon_f = 1$ für das Inland und mit $\varepsilon_h = 1$ und $\varepsilon_f = 1,05$ für das Ausland, da das größere Land aus dessen Sicht das Ausland darstellt.



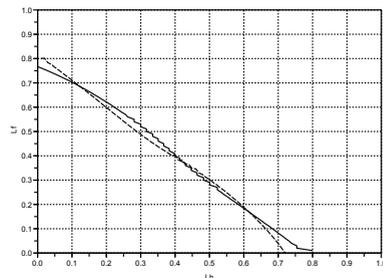
(a) hohe Transportkosten ($T = 3,0$)



(b) geringe Transportkosten ($T = 1,5$)



(c) mittlere Transportkosten - nur asym. Fall ($T = 1,9$)



(d) mittlere Transportkosten - nur asym. Fall ($T = 2,15$)

Abbildung 4.7: In- und Auslandsleichgewichte bei unterschiedlich großen Ländern ($\varepsilon_h = 1,05$)

Die oberen beiden Diagramme in Abbildung 4.7 vergleichen den Fall unterschiedlich großer Länder (schwarze Linien) mit dem gleich großer Länder (graue Linien) bei hohen (Diagramm a) und geringen (Diagramm b) Transportkosten. Im Fall hoher Transportkosten ändern sich zwar die Achsenabschnitte der Gleichgewichtslinien. Diese sind aber nicht relevant, da die Mechanismen des Modells bei hohen Transportkosten immer zu einem stabilen Gleichgewicht bei $L_h = L_f = \mu$ führen. Im Fall geringer Transportkosten findet die Produktion aufgrund der Agglomerationskräfte immer nur in einem Land statt. Daher sind hier die beiden Randlösungen stabile Gleichgewichte. Im Fall gleich großer Länder liegen diese bei $L_h = 2\mu$ für $L_f = 0$ und umgekehrt (siehe oben). Sind die Länder unterschiedlich groß, treten zwei weitere Effekte auf. Zum einen ist die Bevölkerung insgesamt nun größer und daher auch die Nachfrage nach industriellen Produkten. Im Inland, von dem hier angenommen wird, dass es 5% größer ist als das Ausland, stehen aber bei vollkommener Spezialisierung auf industrielle Produkte mehr Arbeitskräfte zur Verfügung, als wenn diese Spezialisierung im kleineren Ausland stattgefunden hätte. Deshalb liegen die relevanten Achsenabschnitte jetzt bei $L_h < 2\mu$ und $L_f > 2\mu$.

Außerdem ist das größere Land eher in der Lage, Skaleneffekte zu realisieren. Dadurch entsteht im Modell eine Asymmetrie, die in den Diagrammen (c) und (d) von Abb. 4.7 besonders deutlich wird. Hier ist für eine bessere Übersicht nur der Fall asymmetrischer Länder dargestellt. Im dazugehörigen Bifurkationsdiagramm (Abb. 4.8), hier für das Inland dargestellt, sind die in Abb. 4.7 verwendeten Transportkostensätze mit roten Linien markiert. Bei beiden Transportkostensätzen ($T = 1,9$ und $T = 2,15$), stellen die Randlösungen stabile Gleichgewichte dar. Dies kann aus Abbildung 4.7 (c) und (d) abgelesen werden, da bei $L_h = 0$ die Auslandslinie über der Inlandslinie und bei $L_f = 0$ die Inlandslinie rechts von der Auslandslinie liegt. Entsprechend stellen die jeweils äußeren Schnittpunkte mit der Achse stabile Gleichgewichte dar, bei denen aber eine Lohnsatzdifferenz bestehen bleibt. Bei $T = 1,9$ besteht, genau wie bei $T = 1,5$ ein weiteres instabiles Gleichgewicht, das nun aber nicht mehr bei $L_h = \mu$ liegt, sondern darunter. Bei $T = 2,15$ befinden wir uns im Bereich mittlerer Transportkosten. Vergleicht man diese Grafik mit Abb. 4.5, die den symmetrischen Fall darstellt, so wird auch hier die Asymmetrie deutlich, die bereits bei einem Größenunterschied von 5% auftritt: Das mittlere stabile Gleichgewicht liegt nun über 0,4 und auch die instabilen Gleichgewichte dazwischen sind nicht mehr symmetrisch.

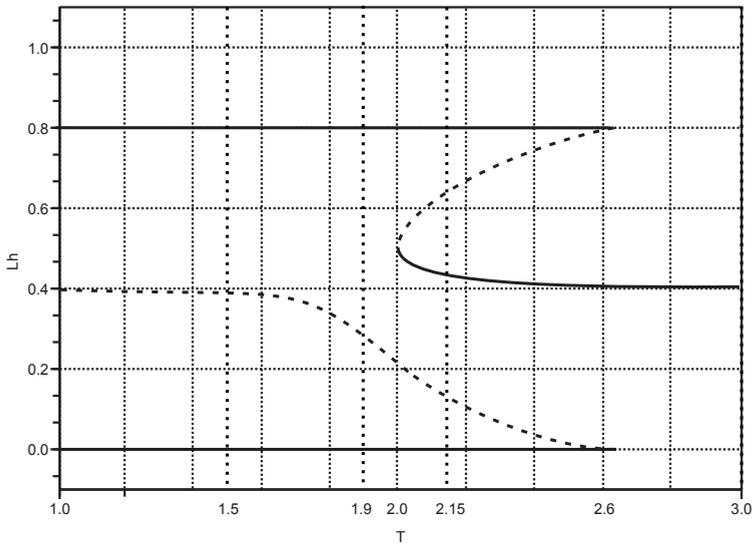


Abbildung 4.8: IM-Modell bei hohen Transportkosten ($T = 3$) und unterschiedlich großen Ländern (hier für das größere Land dargestellt)

Insgesamt kann festgestellt werden, dass bei sinkenden Transportkosten das System von einer gleichmäßigen Verteilung der Industrie immer zu einer Agglomeration im größeren Land springt. Anders als im symmetrischen Modell mit Zwischenprodukten (Abschnitt 4.1) ist der Ort der Agglomerationsbildung nun durch die Größe der Länder bestimmt. Dieses Ergebnis ist qualitativ gleich dem des asymmetrischen CP-Modells in Abschnitt 3.3. Obwohl die Mechanismen, die zur Bildung von Kern und Peripherie führen in diesem Modell andere sind als im Grundmodell von Krugman, kann durch die Annahme eines Größenunterschieds also auch hier die Richtung der Entwicklung bei sinkenden Transportkosten sicher vorausgesagt werden.

4.3 Konzentration und Spezialisierung von Ländern

Das Cluster-Modell (CM) von Krugman und Venables (1996) ist ein Ansatz der Neuen Ökonomischen Geographie, der die Konzentration von Industriesektoren

auf verschiedene Länder auch bei regional immobilen Arbeitskräften erklärt. Die resultierenden Gleichgewichte sind, wie bei allen hier dargestellten Modellen, abhängig von der Höhe der Transportkosten.

Das Modell basiert auf folgenden Grundannahmen: Es gibt zwei Sektoren, in denen die Marktform der monopolistischen Konkurrenz herrscht. Im Gegensatz zum Modell mit Zwischenprodukten aus Abschnitt 4.1 gibt es hier keinen Sektor mit vollkommener Konkurrenz. Ein Teil der produzierten Güter fließt als Vorprodukte wieder in die Produktion ein. Produktionsfaktoren sind daher Arbeit und Vorprodukte sowohl des eigenen als auch des anderen Sektors. Für die Bildung von Agglomerationen ist es allerdings nötig, dass mehr eigene als fremde Vorprodukte eingesetzt werden. Arbeit ist regional immobil, kann aber in beiden Sektoren eingesetzt werden.

Agglomerationen entstehen in diesem Modell, weil ein zusätzliches Unternehmen des Sektors 1, das sich z.B. in Region 1 niederlässt, in dieser Region das Angebot an Vorprodukten des Sektors 1 erhöht. Davon profitieren Unternehmen des Sektors 1 mehr als Unternehmen des Sektors 2. Andererseits verstärkt sich aber der Wettbewerb auf dem Arbeits- und Absatzmarkt. Dieser Effekt wirkt sich auf beide Sektoren aus, so dass es zu gegenläufigen Effekten kommt.

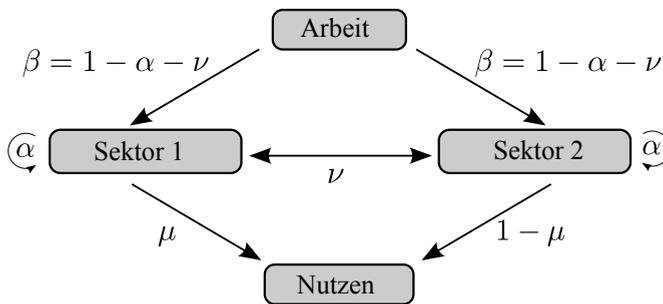


Abbildung 4.9: Schematischer Aufbau des Clustermodells (ein Land)

Nachfrage Die Nutzenfunktion der Haushalte lautet

$$U = M_1^\mu \cdot M_2^{1-\mu} \quad (4.34)$$

wobei die Aggregate M_1 und M_2 jeweils durch Subnutzenfunktionen vom CES-Typ dargestellt werden

$$M_j = \left[\int_0^{n_h} x_{hj}^{\frac{\sigma-1}{\sigma}} dj + \int_0^{n_f} x_{fj}^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}} \quad (4.35)$$

In die Aggregate fließen sowohl im Inland als auch im Ausland hergestellte Produktvarianten ein. x_{hi} und x_{fi} bezeichnen die im Inland und im Ausland produzierte Menge des Gutes i . n_h und n_f sind die Anzahl der Produktvarianten oder Unternehmen in beiden Ländern.

Bei einer Cobb-Douglas-Nutzenfunktion, bei der sich die Exponenten zu eins addieren, entsprechen die Exponenten den Ausgabenanteilen für die einzelnen Güter (bzw. hier den einzelnen Güteraggregaten) M_1 und M_2 .¹⁰⁵ Die Aufteilung dieser Aggregate wird nun durch eine Ausgabenminimierung bestimmt, wobei die Summe der Nachfragen nach den einzelnen Varianten eben genau diesen Aggregaten entsprechen muss.

Nach dem in Abschnitt 2.8 beschriebenen Verfahren werden Nachfragefunktion und Preisindex bestimmt. Unter Berücksichtigung der Nachfrage von ausländischen Konsumenten ergeben sich für die Nachfrage nach Produkten des Sektors 1

$$m_1 = \mu Y_h \frac{p_1^{-\sigma}}{Q^{1-\sigma}} + \mu Y_f \frac{p_1^{-\sigma}}{(Q/T)^{1-\sigma}} \quad (4.36)$$

und des Sektors 2

$$m_2 = (1 - \mu) Y_h \frac{p_2^{-\sigma}}{Q^{1-\sigma}} + (1 - \mu) Y_f \frac{p_2^{-\sigma}}{(Q/T)^{1-\sigma}} \quad (4.37)$$

sowie für den Preisindex (hier für Sektor j des Inlands)

$$Q_{hj} = \left[n_{hj} p_{hj}^{1-\sigma} + n_{fj} (p_{fj} T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.38)$$

Für den zweiten Sektor sowie das Ausland gilt Entsprechendes. Es gibt also in diesem Modell insgesamt vier Preisindizes.

¹⁰⁵Vgl. Mas-Collel et al. (1995) S. 56.

Angebot Es wird ein Produktionsprozess angenommen, der gedanklich in zwei Stufen zerlegt werden kann. Zunächst werden aus den Preisen der Produktionsfaktoren (Arbeit und (Vor-)Produkte beider Sektoren) die Kosten einer zusammengesetzten Inputeinheit berechnet. Dann werden diese in fixe und variable Kosten aufgeteilt. Durch diese Vorgehensweise wird die Rechnung vereinfacht. Es wird aber implizit unterstellt, dass die Zusammensetzung der fixen und variablen Faktoren identisch ist. Ferner wird angenommen, dass die verwendeten Aggregate der Zwischenprodukte denen im Konsum entsprechen. Wäre dies nicht so, könnten die oben bestimmten Preisindizes nicht verwendet werden.

Die Produktion findet mit der Gesamtkostenfunktion

$$TC_i = C_i \cdot (F + cq_i) \quad (4.39)$$

statt. F sind die Fixkosten und c die variablen Kosten, jeweils ausgedrückt in Kosteneinheiten C_i des zusammengesetzten Inputfaktors.

$$C_i = w_i^\beta \cdot Q_i^\alpha \cdot Q_j^\nu \quad (4.40)$$

Dabei bezeichnet i den eigenen und j den anderen Sektor. Die Produktion des Einheitsfaktors unterliegt konstanten Skalenerträgen.¹⁰⁶ Daher gilt $\beta + \alpha + \nu = 1$. Außerdem wird, wie schon erwähnt, $\alpha > \nu$ unterstellt. Der gewinnmaximale Preis wird nach der Amoroso-Robinson-Relation bestimmt als

$$p_i = \underbrace{w_i^\beta Q_i^\alpha Q_j^\nu}_{C_i} \cdot c \cdot \frac{\sigma}{\sigma - 1} \quad (4.41)$$

Um nun die Menge zu berechnen, bei der sich ein Gewinn von 0 ergibt, wird zunächst die Gewinnfunktion aufgestellt und der in (4.41) bestimmte Preis eingesetzt.

$$\pi_i = p_i q_i - C_i \cdot (F + cq_i) = \frac{c \cdot C_i}{\sigma - 1} \cdot \left[q_i - \frac{F(\sigma - 1)}{c} \right] = 0 \quad (4.42)$$

Auflösen nach q_i ergibt die für das Unternehmen optimale Menge, bei der sich

¹⁰⁶Eine solche Kostenfunktion lässt sich leicht vereinfacht aus einer Cobb-Douglas Produktionsfunktion ableiten. Damit die Stückkosten konstant sind, muss die Produktionsfunktion konstante Skalenerträge aufweisen (vgl. Jehle und Reny (2001), S. 131).

aufgrund der Marktform ein Gewinn von 0 ergibt

$$q_i^* = (\sigma - 1) \frac{F}{c} \quad (4.43)$$

Diese optimale Menge ist für alle Unternehmen gleich und unabhängig von der Nachfrageseite. Eine Zunahme der Nachfrage führt daher nicht zu einer größeren Angebotsmenge des einzelnen Unternehmens sondern zu einer steigenden Zahl von Unternehmen und somit Produktvarianten.

Standardisierungen Wie in anderen Modellen der NEG auch, werden nun zwei Standardisierungen durchgeführt, indem die Mengeneinheiten für die variablen Kosten c und die fixen Kosten F geeignet festgelegt werden.

$$c \equiv \frac{\sigma - 1}{\sigma} \quad (4.44)$$

setzt man dies in (4.41) ein, so vereinfacht sich der optimale Preis zu

$$p_i = w_i^\beta \cdot Q_i^\alpha \cdot Q_j^\nu \quad (4.45)$$

Aus (4.43) wird dann

$$q^* = F\sigma \quad (4.46)$$

Die Fixkosten werden standardisiert auf

$$F = \frac{1}{\beta\sigma} \quad (4.47)$$

Dann wird aus (4.46)

$$q^* = \frac{1}{\beta} \quad (4.48)$$

Preisindex Die Anzahl der Unternehmen kann nun bestimmt werden, indem die Lohnsumme dem Anteil des Faktors Arbeit an dem Wert der gesamten Produktion gleichgesetzt wird

$$w_i L_i = \beta n_i p_i q^* \quad (4.49)$$

setzt man hier (4.48) ein, erhält man

$$w_i L_i = n_i p_i \quad (4.50)$$

oder

$$n_i = \frac{w_i L_i}{p_i} \quad (4.51)$$

Setzt man nun die Anzahl der Unternehmen (4.51) und den Preis (4.45) in (4.38) ein, bekommt man

$$Q_{hi} = \left[\frac{w_{hi} L_{hi}}{w_{hi}^\beta Q_{hi}^\alpha Q_{hj}^\nu} \cdot \left(w_{hi}^\beta Q_{hi}^\alpha Q_{hj}^\nu \right)^{1-\sigma} + \frac{w_{fi} L_{fi}}{w_{fi}^\beta Q_{fi}^\alpha Q_{fj}^\nu} \cdot \left(w_{fi}^\beta Q_{fi}^\alpha Q_{fj}^\nu T \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.52)$$

Durch Ausmultiplizieren und Kürzen ergibt sich

$$Q_{hi} = \left[L_{hi} w_{hi}^{1-\beta\sigma} Q_{hi}^{-\alpha\sigma} Q_{hj}^{-\nu\sigma} + L_{fi} w_{fi}^{1-\beta\sigma} Q_{fi}^{-\alpha\sigma} Q_{fj}^{-\nu\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.53)$$

und Entsprechendes für das Ausland und jeweils für beide Sektoren. Insgesamt ergeben sich so vier Gleichungen.

Ausgabengleichung Das gesamte Einkommen in einem Land ist

$$Y = w_1 L_1 + w_2 L_2 \quad (4.54)$$

wobei $L_1 + L_2 = 1$ gilt¹⁰⁷. Die Ausgaben für die Produkte eines Sektors E_i setzen sich aus den Ausgaben der privaten Haushalte und denen der Unternehmen für die Zwischenprodukte dieses Sektors zusammen. Die Haushalte geben aufgrund ihrer Nutzenfunktion (4.34) den Anteil μ ihres Einkommens für Produkte des einen Sektors und den Anteil $1 - \mu$ für die des anderen Sektors aus. Da die Unternehmen keine Gewinne machen, entsprechen ihre Ausgaben für Vorprodukte des eigenen und des anderen Sektors dem Anteil α bzw. ν ihrer Erträge.

$$E_{h1} = \mu \cdot (w_{h1} L_{h1} + w_{h2} L_{h2}) + \frac{\alpha w_{h1} L_{h1} + \nu w_{h2} L_{h2}}{\beta} \quad (4.55)$$

¹⁰⁷Die gesamte Menge an Arbeitskräften in einem Land wird auf 1 normiert. Daher können die Werte L_1 und L_2 als Anteile interpretiert werden. Von Arbeitslosigkeit wird abgesehen.

$$E_{h2} = (1 - \mu) \cdot (w_{h1}L_{h1} + w_{h2}L_{h2}) + \frac{\alpha w_{h2}L_{h2} + \nu w_{h1}L_{h1}}{\beta} \quad (4.56)$$

Auch diese Gleichung kann für beide Sektoren und beide Länder analog aufgestellt werden.

Lohngleichung Wir nehmen an, dass in der Produktion das gleiche zusammengesetzte Gut M_i verwendet wird, welches auch konsumiert wird. Damit ist die Subnutzenfunktion (4.35) auch als Subproduktionsfunktion zu verstehen und die Unternehmen werden im Rahmen ihrer Kostenminimierung zu den gleichen Nachfragefunktionen kommen, wie die Haushalte bei deren Ausgabenminimierung. Daher kann in (4.36) und (4.37) als Budget auch die Ausgabenfunktion (4.56) verwendet werden um zur Nachfrage nach einer einzelnen Produktvariante zu gelangen. Setzt man daher für μY in (4.36) E_i ein, erhält man

$$q_i = p_i^{-\sigma} \cdot \left[E_i Q_i^{\sigma-1} + E_{fi} \left(\frac{Q_{fi}}{T} \right)^{\sigma-1} \right] \quad (4.57)$$

Dies muss im Gleichgewicht dem optimalen Angebot bei einem Gewinn von 0 aus Gleichung (4.48) entsprechen. Setzt man also hier (4.45) ein und setzt (4.57) gleich (4.48), so erhält man die Lohngleichung

$$\frac{\left[w_{hi}^\beta Q_{hi}^\alpha Q_{hj}^\nu \right]^\sigma}{\beta} = Q_{hi}^{\sigma-1} E_{hi} + Q_{fi}^{\sigma-1} E_{fi} T^{1-\sigma} \quad (4.58)$$

die sich nach dem Lohnsatz auflösen lässt

$$w_{hi} = \left[\frac{\beta \left(Q_{hi}^\sigma \cdot E_{hi} \cdot Q_{fi} + Q_{fi}^\sigma \cdot E_{fi} \cdot T^{1-\sigma} \cdot Q_{hi} \right)}{Q_{hi} \cdot Q_{fi}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{hi}^{-\frac{\alpha}{\beta}} \cdot Q_{hj}^{-\frac{\nu}{\beta}} \quad (4.59)$$

Nur die Preisindizes lassen sich nicht auflösen und müssen durch geschachtelte Iterationen bestimmt werden.

Gleichgewicht Formuliert man die drei Gleichungen (4.53), (4.56) und (4.59) für zwei Sektoren (1 und 2) und für Inland h und Ausland f , so ergibt sich ein Gleichungssystem aus 12 Gleichungen.

4 Modelle für internationale Agglomeration

$$Q_{h1} = \left[L_{h1} w_{h1}^{1-\beta\sigma} Q_{h1}^{-\alpha\sigma} Q_{h2}^{-\nu\sigma} + L_{f1} w_{f1}^{1-\beta\sigma} Q_{f1}^{-\alpha\sigma} Q_{f2}^{-\nu\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.60)$$

$$Q_{h2} = \left[L_{h2} w_{h2}^{1-\beta\sigma} Q_{h2}^{-\alpha\sigma} Q_{h1}^{-\nu\sigma} + L_{f2} w_{f2}^{1-\beta\sigma} Q_{f2}^{-\alpha\sigma} Q_{f1}^{-\nu\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.61)$$

$$Q_{f1} = \left[L_{f1} w_{f1}^{1-\beta\sigma} Q_{f1}^{-\alpha\sigma} Q_{f2}^{-\nu\sigma} + L_{h1} w_{h1}^{1-\beta\sigma} Q_{h1}^{-\alpha\sigma} Q_{h2}^{-\nu\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.62)$$

$$Q_{f2} = \left[L_{f2} w_{f2}^{1-\beta\sigma} Q_{f2}^{-\alpha\sigma} Q_{f1}^{-\nu\sigma} + L_{h2} w_{h2}^{1-\beta\sigma} Q_{h2}^{-\alpha\sigma} Q_{h1}^{-\nu\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.63)$$

$$E_{h1} = \mu \cdot (w_{h1} L_{h1} + w_{h2} L_{h2}) + \frac{\alpha w_{h1} L_{h1} + \nu w_{h2} L_{h2}}{\beta} \quad (4.64)$$

$$E_{h2} = (1 - \mu) \cdot (w_{h1} L_{h1} + w_{h2} L_{h2}) + \frac{\nu w_{h1} L_{h1} + \alpha w_{h2} L_{h2}}{\beta} \quad (4.65)$$

$$E_{f1} = \mu \cdot (w_{f1} L_{f1} + w_{f2} L_{f2}) + \frac{\alpha w_{f1} L_{f1} + \nu w_{f2} L_{f2}}{\beta} \quad (4.66)$$

$$E_{f2} = (1 - \mu) \cdot (w_{f1} L_{f1} + w_{f2} L_{f2}) + \frac{\nu w_{f1} L_{f1} + \alpha w_{f2} L_{f2}}{\beta} \quad (4.67)$$

$$w_{h1} = \left[\frac{\beta \left(Q_{h1}^\sigma \cdot E_{h1} \cdot Q_{f1} + Q_{f1}^\sigma \cdot E_{f1} \cdot T^{1-\sigma} \cdot Q_{h1} \right)}{Q_{h1} \cdot Q_{f1}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{h1}^{-\frac{\alpha}{\beta}} \cdot Q_{h2}^{-\frac{\nu}{\beta}} \quad (4.68)$$

$$w_{h2} = \left[\frac{\beta \left(Q_{h2}^\sigma \cdot E_{h2} \cdot Q_{f2} + Q_{f2}^\sigma \cdot E_{f2} \cdot T^{1-\sigma} \cdot Q_{h2} \right)}{Q_{h2} \cdot Q_{f2}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{h2}^{-\frac{\alpha}{\beta}} \cdot Q_{h1}^{-\frac{\nu}{\beta}} \quad (4.69)$$

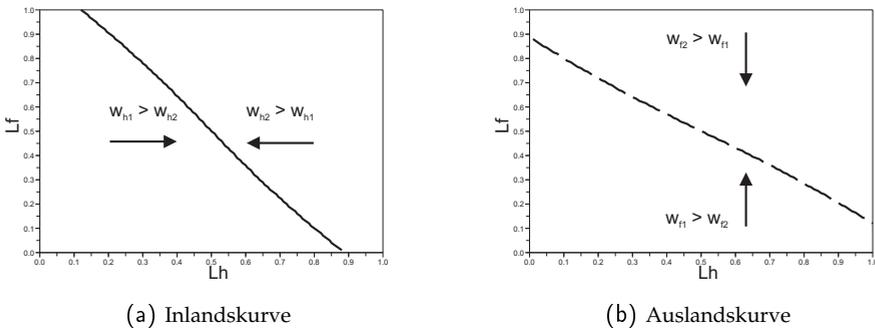
$$w_{f1} = \left[\frac{\beta \left(Q_{f1}^\sigma \cdot E_{f1} \cdot Q_{h1} + Q_{h1}^\sigma \cdot E_{h1} \cdot T^{1-\sigma} \cdot Q_{f1} \right)}{Q_{f1} \cdot Q_{h1}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{f1}^{-\frac{\alpha}{\beta}} \cdot Q_{f2}^{-\frac{\nu}{\beta}} \quad (4.70)$$

$$w_{f2} = \left[\frac{\beta \left(Q_{f2}^\sigma \cdot E_{f2} \cdot Q_{h2} + Q_{h2}^\sigma \cdot E_{h2} \cdot T^{1-\sigma} \cdot Q_{f2} \right)}{Q_{f2} \cdot Q_{h2}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{f2}^{-\frac{\alpha}{\beta}} \cdot Q_{f1}^{-\frac{\nu}{\beta}} \quad (4.71)$$

Kurzfristig beschreiben die Gleichungen (4.60) bis (4.71) das Gleichgewicht. Die Verteilung der Arbeiter auf die Sektoren eines Landes ist konstant. Die Löhne in

den Sektoren können sich unterscheiden. Liegt der Lohn in einem Sektor über dem des anderen, werden die Arbeiter in den Sektor mit dem höheren Lohnsatz wechseln bis die Lohnsätze zum Ausgleich kommen. Da das Gleichungssystem nicht analytisch lösbar ist, wird im Folgenden die numerische Simulation verwendet, um zu Ergebnissen zu gelangen.

Durch Simulation wird der Beschäftigungsanteil (L) des Sektors 1 im In- und Ausland bei gegebenem Anteil des Sektors 1 im jeweils anderen Land errechnet, bei dem die Löhne im betrachteten Land in beiden Sektoren gleich sind. Die Ausgabenteile μ wurden, wie in der Literatur üblich, auf den Wert 0,5 festgelegt. Man erhält $L_h (L_f)$ für das Inland und $L_f (L_h)$ für das Ausland. Da die beiden Länder identisch sind, handelt es sich bei der Gleichgewichtslinie für das Ausland um die Inverse der entsprechenden Linie des Inlandes.¹⁰⁸



(a) Inlandskurve

(b) Auslandskurve

Abbildung 4.10: Kurzfristiges Gleichgewicht im In- und Ausland

Die Ergebnisse der Simulationen werden in einem Diagramm dargestellt, in dem auf der horizontalen Achse die Aufteilung der inländischen Arbeitskräfte auf die Sektoren und auf der vertikalen Achse die der ausländischen Arbeitskräfte auf die Sektoren dargestellt ist. Ein Punkt in der linken unteren Ecke stellt eine Situation dar, bei der keine Arbeitskräfte im inländischen und keine im ausländischen Sektor 1 arbeiten, sondern in beiden Ländern alle Arbeiter in Sektor 2 beschäftigt sind. Bewegt man sich nach rechts, nimmt der Anteil der im Inland in Sektor 1 beschäftigten Arbeiter zu und folglich der in Sektor 2 beschäftigten ab. Eine Bewegung nach oben bedeutet analog eine Wanderung von Arbeitern von Sektor 2

¹⁰⁸Die Simulationsmethode wird in Kapitel 5 ausführlich dargestellt.

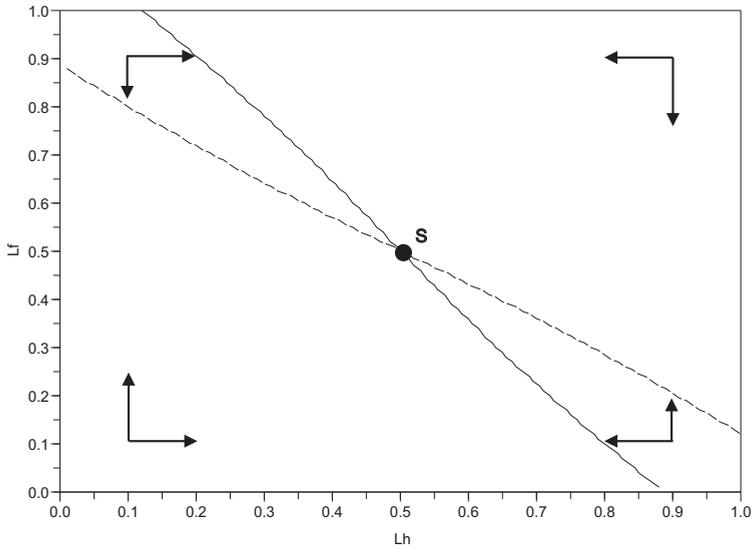


Abbildung 4.11: Hohe Transportkosten ($\sigma = 3$; $\alpha = 0,4$; $\nu = 0,1$; $T = 3,0$)

nach Sektor 1 im Ausland. Die durchgezogene Linie verbindet solche Werte für L_h bei denen bei gegebenem L_f der inländische Lohn für beide Sektoren gleich ist. Passen die Werte von L_h und L_f so zusammen, dass bei jeweils gegebenem anderen Wert in beiden Ländern Lohnausgleich herrscht¹⁰⁹, so handelt es sich um ein langfristiges Gleichgewicht. Dies ist den Schnittpunkten beider Kurven erreicht. Aufgrund von Lohnunterschieden wechseln Arbeitnehmer den Sektor. Für das Inland bedeutet das graphisch eine Verlagerung der Verteilung in *horizontaler* Richtung (rechts oder links). Eine analoge Bewegung für das Ausland bedeutet dann eine *vertikale* Verlagerung (vgl. Abb. 4.10). Diese Verlagerungen geschehen zu der jeweiligen Gleichgewichtskurve hin. Daraus folgt, dass es stabile und instabile Gleichgewichte geben kann.

Stabile Gleichgewichte sind solche, bei denen sich das System bei einer kleinen Abweichung vom Gleichgewichtspunkt von selbst wieder zu diesem zurück bewegt. In Abbildung 4.11 stellt der Punkt S ein solches stabiles Gleichgewicht dar.

¹⁰⁹Es muss also gelten $w_{h1} = w_{h2}$ und $w_{f1} = w_{f2}$. w_{h1} und w_{f1} sowie w_{h2} und w_{f2} dürfen sich aber voneinander unterscheiden, da die Arbeiter regional immobil sind.

Stellt man sich ausgehend von Punkt S eine geringe Verschiebung nach unten vor, würde der ausländische Lohnsatz in Sektor 1 den in Sektor 2 übersteigen ($w_{f1} > w_{f2}$) und ein Teil der Arbeitskräfte würde zu Sektor 1 wechseln. Im Inland wäre $w_{h1} > w_{h2}$. Auch hier würden Arbeiter von Sektor 2 in Sektor 1 wechseln. Ohne Annahmen über die Reihenfolge und die Stärke der Reaktionen können wir den Anpassungsprozess nicht im Detail beschreiben. Wir wissen aber, dass sich am Ende ein Gleichgewicht in Punkt S einstellt. Gehen wir alternativ von Punkt U in Abbildung 4.12 aus und nehmen wieder eine geringe Verlagerung nach unten an, so erreichen wir diesen Punkt durch einen Anpassungsprozess nicht, sondern den Punkt S in der unteren rechten Ecke der Abbildung. Der Grund dafür ist, dass in dem Bereich zwischen den beiden Kurven rechts unterhalb von U keine Tendenz zur Mitte, sondern zum Eckpunkt vorliegt. Das liegt daran, dass hier $w_{h1} > w_{h2}$ und $w_{f2} > w_{f1}$ gilt. Daher ist der Punkt U instabil und die mit S bezeichneten Punkte stabil.

Für die nun folgenden Simulationen wurden die Parameter $\sigma = 3$, $\mu = 0,5$, $\alpha = 0,4$ und $\nu = 0,1$ gesetzt und nur die Transportkosten variiert. Aufgrund von Lohndifferenzen steigt im langfristigen Gleichgewicht bei jedem Punkt unterhalb der L_f -Kurve L_f an und sinkt bei jedem Punkt oberhalb dieser Kurve. Für das Inland steigt L_h für jeden Punkt links von der L_h -Kurve und sinkt für jeden Punkt rechts von dieser Kurve. Dies wird durch die Pfeile in den Abbildungen verdeutlicht.

Bei hohen Transportkosten (Abb. 4.11) existiert nur ein Gleichgewicht, bei dem jedes Land jeweils die Hälfte der Produktion eines jeden Sektors besitzt ($L_h = L_f = 0,5$). Die agglomerativen Kräfte sind hier nicht stark genug, um die hohen Transportkosten zu kompensieren. Jedes Land produziert die im Land verbrauchten Güter selbst.

Im Fall niedriger Transportkosten (Abb. 4.12) gibt es drei Gleichgewichte: Das symmetrische sowie die Agglomeration von jeweils einem Sektor in einer Region. Das symmetrische Gleichgewicht ist nun allerdings nicht mehr stabil. Von jedem anderen Punkt des Diagramms gelangt man aufgrund der beschriebenen Kräfte zu den beiden Extrempunkten, in denen sich jedes Land auf die Produktion eines Gutes vollständig spezialisiert. Im jeweils anderen Land findet dann keine Produktion dieses Gutes mehr statt (obere linke und untere rechte Ecke).

Liegen die Transportkosten im mittleren Bereich, so entstehen fünf Gleichgewich-

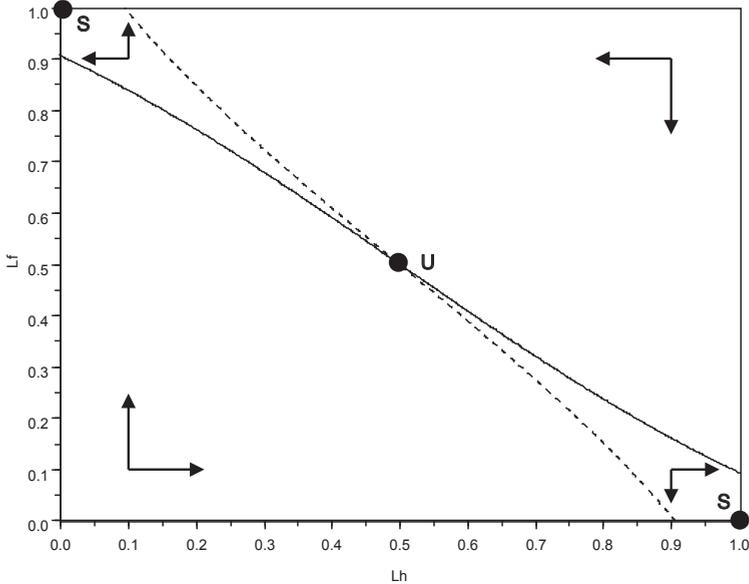


Abbildung 4.12: Niedrige Transportkosten ($\sigma = 3$; $\alpha = 0,4$; $\nu = 0,1$; $T = 2,0$)

te: Ein stabiles in der Mitte und in den beiden Ecken sowie je ein instabiles dazwischen (Abb. 4.13). Ist die Produktion in der Ausgangslage relativ gleich verteilt führt ein Absenken der Transportkosten auf einen mittleren Wert zu einem symmetrischen Gleichgewicht. Ist die Verteilung in der Ausgangslage relativ ungleich, wird dagegen diese Ungleichheit weiter verstärkt und es kommt zur vollständigen Spezialisierung der beiden Regionen.

Es stellt sich allerdings die Frage, was unter "mittleren Transportkosten" zu verstehen ist. Der obere Grenzwert \bar{T} wird als Sustain-Point bezeichnet. Ab dieser Höhe der Transportkosten kommt es immer zur Bildung von Agglomerationen. Der untere Grenzwert \underline{T} heißt Break-Point. Unterhalb dieser Schwelle entstehen in keinem Fall Agglomerationen.

Der Sustain-Point lässt sich bestimmen, indem $L_h = 1$ und $L_f = 0$ gesetzt wird (oder andersherum). Für diese Werte muss $w_{h1} \geq w_{h2}$ und $w_{f2} \geq w_{f1}$ gelten, damit es sich um ein langfristiges Gleichgewicht handelt. Bei $w_{h1} = w_{h2}$ und $w_{f2} = w_{f1}$ muss dann für T die Bedingung

$$\bar{T}^{\sigma(\alpha-\nu)} = (1 + \alpha - \nu) \bar{T}^{1-\sigma} + (1 - \alpha + \nu) \bar{T}^{\sigma-1} \quad (4.72)$$

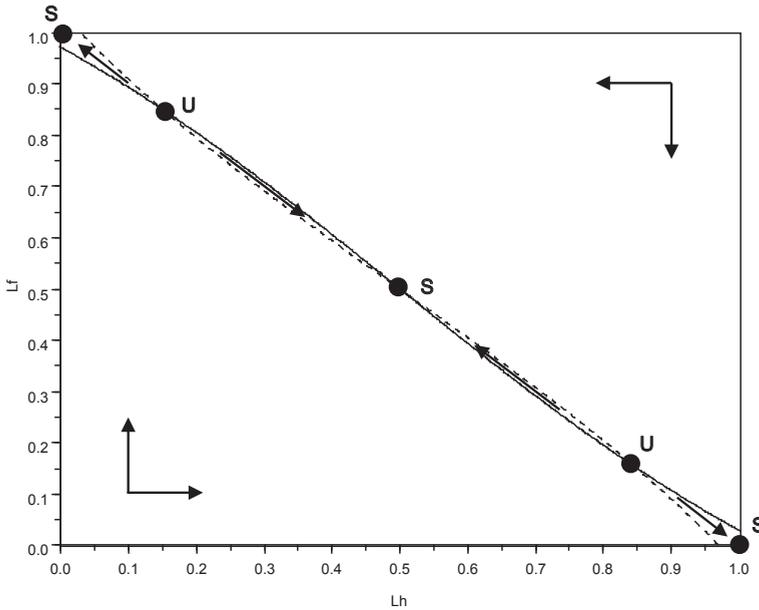


Abbildung 4.13: Mittlere Transportkosten ($\sigma = 3$; $\alpha = 0,4$; $\nu = 0,1$; $T = 2,35$)

erfüllt sein. Um den Break-Point zu bestimmen, werden zunächst alle Variablen im symmetrischen Gleichgewicht berechnet, wobei die doppelte Symmetrie des Systems (Regionen und Sektoren) die Rechnung deutlich vereinfacht. Danach wird das totale Differenzial der Lohngleichung gebildet und der Wert von T bestimmt, bei dem dieses Differenzial gerade 0 ist¹¹⁰.

$$\underline{T}^{\sigma-1} = \frac{[1 + \alpha - \nu] \cdot [\sigma(1 + \alpha - \nu) - 1]}{[1 - \alpha + \nu] \cdot [\sigma(1 - \alpha + \nu) - 1]} \quad (4.73)$$

Wenn die Bedingung $\alpha > \nu$ eingehalten wird, gilt immer $1 < \underline{T} < \bar{T}$. Dies ist auch sinnvoll, denn ein Transportkostensatz im Sinne der Eisberg-Theorie von 1 bedeutet ja Transportkosten von 0. Negative Transportkosten machen aber keinen Sinn. Bei den in diesem Kapitel durchgängig verwendeten Werten von $\sigma = 3$ und $\alpha - \nu = 0,3$ lauten die Werte $\underline{T} = 2,21$ und $\bar{T} = 2,49$ ¹¹¹. Stellt man die Verteilung der Arbeitskräfte im Inland auf die beiden Sektoren in Abhängigkeit von der Höhe der Transportkosten grafisch dar, ergibt sich Abbildung (4.14).

¹¹⁰Für eine genaue Herleitung der Gleichungen (4.72) und (4.73) vgl. Klüver (2000).

¹¹¹Vgl. Krugman und Venables (1996), Tabelle 2.

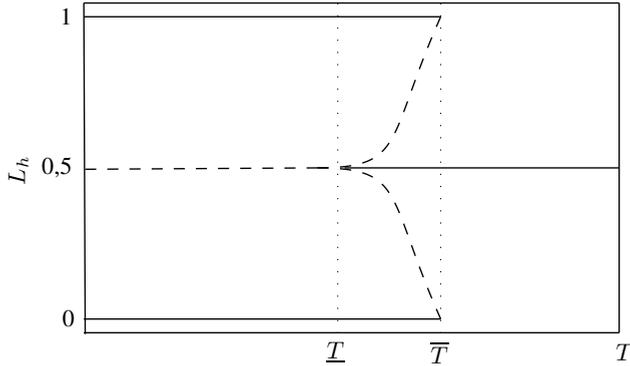


Abbildung 4.14: Bifurkationsdiagramm des symmetrischen Cluster-Modells

4.4 Spezialisierung bei sektorspezifischen Transportkosten¹¹²

Wie bereits in Kapitel 2.5 gezeigt wurde, unterscheiden sich in der Realität die Transportkosten pro transportierter Werteinheit für verschiedene Güter oft erheblich. Auch kommt es zu Veränderungen der Transportkosten, die sich nur auf bestimmte Güterarten auswirken. Diese können durch Weiterentwicklungen in der Transporttechnologie (z.B. Verbreitung des Internet für Informationsgüter) oder durch technischen Fortschritt in den Gütern selbst (z.B. Verringerung von Gewicht und Volumen bei technischen Geräten) - die dann indirekt die Transportkosten senken - begründet sein. Dieser Abschnitt erweitert das Modell von Krugman und Venables um Transportkostenunterschiede und zeigt die sich daraus ergebenden Gleichgewichte.

Dazu wird der einheitliche Satz in Höhe von T durch T_1 und T_2 ersetzt. Die Nachfragefunktionen (4.36) und (4.37) lauten nun

$$m_1 = \mu Y_h \frac{p_1^{-\sigma}}{Q^{1-\sigma}} + \mu Y_f \frac{p_1^{-\sigma}}{(Q/T_1)^{1-\sigma}} \quad (4.74)$$

und

¹¹²Dieser Abschnitt basiert teilweise auf Frohwerk (2008).

$$m_2 = (1 - \mu)Y_h \frac{p_2^{-\sigma}}{Q^{1-\sigma}} + (1 - \mu)Y_f \frac{p_2^{-\sigma}}{(Q/T_2)^{1-\sigma}} \quad (4.75)$$

und aus dem Preisindex (4.38) wird

$$Q_{hi} = \left[n_{hi} p_{hi}^{1-\sigma} + n_{fi} (p_{fi} T_i)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.76)$$

bzw. nach Einsetzen aus Gleichung (4.53)

$$Q_{hi} = \left[L_{hi} w_{hi}^{1-\beta\sigma} Q_{hi}^{-\alpha\sigma} Q_{hj}^{-v\sigma} + L_{fi} w_{fi}^{1-\beta\sigma} Q_{fi}^{-\alpha\sigma} Q_{fj}^{-v\sigma} T_i^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.77)$$

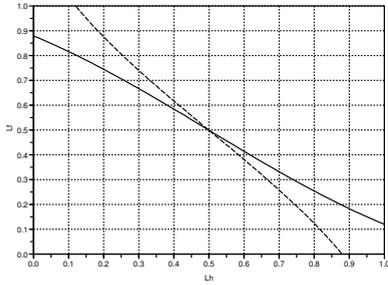
Die Lohngleichung (4.59) lautet nun

$$w_{hi} = \left[\frac{\beta \left(Q_{hi}^\sigma \cdot E_{hi} \cdot Q_{fi} + Q_{fi}^\sigma \cdot E_{fi} \cdot T_i^{1-\sigma} \cdot Q_{hi} \right)}{Q_{hi} \cdot Q_{fi}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{hi}^{-\frac{\alpha}{\beta}} \cdot Q_{hj}^{-\frac{v}{\beta}} \quad (4.78)$$

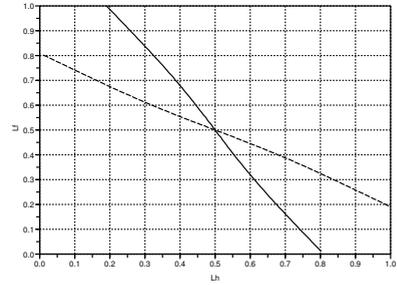
Die Ausgabengleichung bleibt unverändert.

Im Fall identischer Transportkosten kommt es zu fünf möglichen Gleichgewichten, wenn T zwischen \underline{T} und \bar{T} liegt. Geht man nun von unterschiedlichen Transportkostensätzen aus und führt die Simulation mit $T_1 < \underline{T}$ und $T_2 > \bar{T}$ durch, so entstehen auch hier fünf mögliche Gleichgewichte. Dies kann durch die folgende Überlegung nachvollzogen werden: Geht man von gleichen Transportkosten in beiden Sektoren aus und senkt nun die Transportkosten in Sektor 1, so reduziert sich zunächst das Preisniveau in Sektor 1. Da die Produkte dieses Sektors sowohl in der Produktion des eigenen als auch des anderen Sektors als Vorprodukte einfließen, senken sich auch die jeweiligen Produktionskosten. Aufgrund der Annahme, dass die Vorprodukte des eigenen Sektors jeweils einen größeren Anteil in der Produktion haben als die des jeweils anderen Sektors, wirkt diese Kostenreduktion in Sektor 1 stärker als in Sektor 2. Daher kommt es auch zu 5 Gleichgewichten, wenn ein Transportkostensatz am oberen (unteren) Rand des Bereichs mittlerer Transportkostensätze liegt, wenn der andere darunter (darüber) liegt.

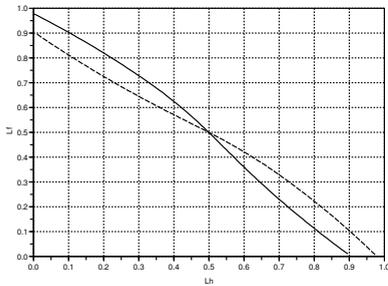
4 Modelle für internationale Agglomeration



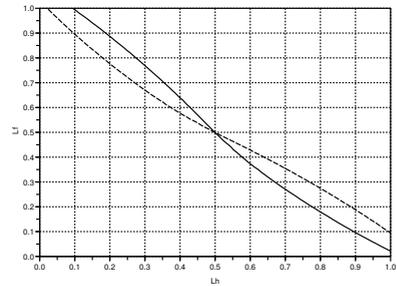
(a) $T_1 = T_2 = 1,8$



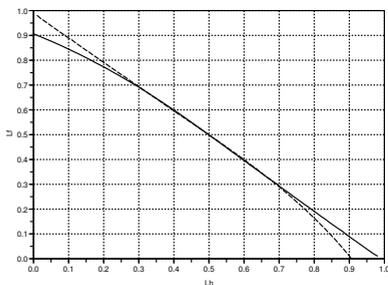
(b) $T_1 = T_2 = 3,5$



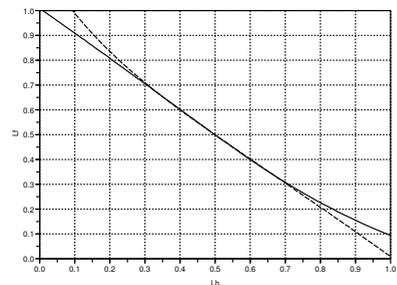
(c) $T_1 = 1,8; T_2 = 3,5$



(d) $T_1 = 3,5; T_2 = 1,8$



(e) $T_1 = 2,6; T_2 = 1,2$



(f) $T_1 = 1,2; T_2 = 2,6$

Abbildung 4.15: Gleichgewichte bei verschiedenen Transportkostensätzen

Die Argumentation kann an dem Gleichungssystem nachvollzogen werden, das der Simulation zu Grunde liegt. Im Folgenden untersuchen wir den Einfluss der in den jeweiligen Gleichungen unabhängigen Variablen auf die Abhängige in einer Iteration.¹¹³

Da die Struktur der Preisindizes gleich ist, betrachten wir beispielhaft Gleichung (4.77)

$$Q_{h1}(w_{h1}^+, w_{f1}^+, L_{h1}^-, L_{f1}^-, Q_{h1}^+, Q_{h2}^+, Q_{f1}^+, Q_{f2}^+, T_1) \quad (4.79)$$

Eine analoge Darstellung ist auch für die anderen drei Q s möglich. Die Preisindizes hängen positiv vom Lohnsatz des eigenen Sektors im In- und Ausland sowie negativ vom Anteil der Beschäftigung im eigenen Sektor im In- und Ausland ab. Eine direkte Abhängigkeit von den entsprechenden Variablen des anderen Sektors besteht nicht. Diese wirken jedoch indirekt über die jeweiligen Preisindizes, denn jeder Preisindex hängt in positiver Weise von allen anderen Indizes sowie von sich selbst ab. Ferner ist die Abhängigkeit jedes Preisindex von den Transportkosten des eigenen Sektors positiv.

Für die Ausgaben ergibt sich

$$E_{h1}(w_{h1}^+, w_{h2}^+, L_{h1}^+, L_{h2}^+) \quad (4.80)$$

Die Ausgaben hängen positiv von den Lohnsätzen und den Beschäftigungsanteilen beider Sektoren des jeweiligen Landes ab. Die gleiche Überlegung für die Lohngleichung führt zu

$$w_{h1}(Q_{h1}^?, Q_{h2}^-, Q_{f1}^?, Q_{f2}^+, E_{h1}^+, E_{f1}^+, T_1) \quad (4.81)$$

Hier lässt sich feststellen, dass der Lohnsatz sich bei einer Änderung der Ausgaben eines jeden Landes für die Güter des eigenen Sektors gleichgerichtet entwickelt, in Bezug auf die Transportkosten gegengerichtet. Ferner hängt der Lohnsatz positiv vom Preisniveau des anderen Sektors im Ausland sowie negativ vom Preisniveau des anderen Sektors im Inland ab. Die Abhängigkeit von den Preisniveaus des eigenen Sektors ist hingegen unbestimmt.

¹¹³Dies ist nicht gleichbedeutend mit der ersten Ableitung. Diese wäre zwar aussagekräftiger, aber leider kann insbesondere die Ableitung der Lohngleichung nach den Transportkosten nicht gebildet werden.

Ein weiterer Effekt ungleicher Transportkosten besteht darin, dass die Verteilung der Arbeitskräfte in beiden Ländern auf die Sektoren nicht mehr genau entgegengesetzt ist. Am deutlichsten ist dies bei vollkommener Spezialisierung eines Landes zu erkennen. Es ist nämlich möglich, dass auch in diesem Fall ein Teil der Arbeitskräfte des anderen Landes im gleichen Sektor beschäftigt ist (Abb. 4.15 e und f).

Der Grund hierfür liegt in der substitutionalen Produktionsfunktion, die der Kostenfunktion (4.40) zugrunde liegt. Diese Kostenfunktion erreicht man, indem man eine Kostenminimierung mit einer Produktionsfunktion vom Cobb-Douglas-Typ $q_i = L_i^\beta \cdot M_i^\alpha \cdot M_j^\nu$ durchführt und den dann in der Kostenfunktion entstandenen konstanten Faktor $\beta^{-\beta} \cdot \alpha^{-\alpha} \cdot \nu^{-\nu}$ weglässt. Dabei müssen sich die Exponenten zu eins addieren. Dies ist auch inhaltlich notwendig, da die gedankliche Zerlegung des Produktionsprozesses nur sinnvoll ist, wenn der zusammengesetzte Produktionsfaktor mit konstanten Skalenerträgen produziert wird. Dann ist auch die Produktionsfunktion homogen, denn die Elastizität der Kostenfunktion bezogen auf die Produktionsmenge (in diesem Fall die Menge des zusammengesetzten Faktors) entspricht der inversen Skalanelastizität der Produktionsfunktion¹¹⁴. Damit die Stückkosten konstant sind, muss daher $\beta + \alpha + \nu = 1$ gelten.

Der Faktor, dessen Transportkosten sinken, wird nun in der Produktion verstärkt eingesetzt. Daher wird von ihm insgesamt auch mehr produziert als von dem anderen Faktor, der dadurch relativ teurer geworden ist. Bildet man die Verteilung der Arbeitskräfte im Inland L_h in Abhängigkeit der Transportkosten des Sektors 1 bei konstanten Transportkosten in Sektor 2 ab, ergibt sich Abbildung (4.16b). Dabei kann aufgrund der Symmetrie der Länder und der Asymmetrie der Sektoren der Knick immer nur auf einer Seite auftreten. Inhaltlich bedeutet dieser Knick, dass sich ein Land vollständig auf einen Sektor spezialisiert und das andere trotzdem einen geringen Anteil dieses Sektors beinhaltet.

¹¹⁴Vgl. Jehle und Reny (2001), S. 131.

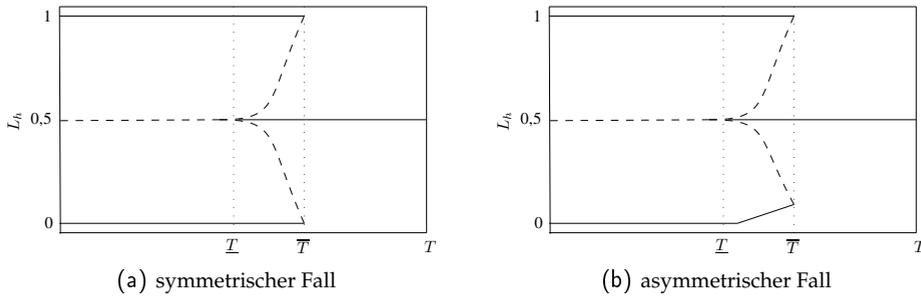


Abbildung 4.16: Gleichgewichte und Transportkosten

4.5 Spezialisierung bei unterschiedlich großen Ländern

In diesem Abschnitt soll die zweite Symmetrieannahme aufgehoben werden und analog zu den Abschnitten 3.3 und 3.4 der Fall unterschiedlich großer Länder auch im CM untersucht werden. Dabei wird ein Größenunterschied von 2,5% angenommen¹¹⁵. Im CM-Modell gibt es keinen Sektor, der seine Größe nicht endogen verändert. Aufgrund der räumlichen Immobilität der Arbeitskräfte addieren sich aber die Beschäftigtenzahlen beider Sektoren jedes Landes jeweils zu deren gesamtem Arbeitsangebot auf. Im Basis-Modell ist dies auf 1 normiert, damit die Werte für L_i als Anteile verstanden werden können. Dies dient aber nur einer besseren Interpretierbarkeit und ist für die Lösung des Modells nicht notwendig. Die Größe eines Landes wird daher in dieser Variante als Summe der Arbeitskräfte beider Sektoren einer Region modelliert, wobei die Annahme $L_{i1} + L_{i2} = 1$ beibehalten wird, diese Werte aber für das Inland in den Preis- und Ausgaben-gleichungen jeweils mit ε_h und für das Ausland mit ε_f multipliziert werden. Das Gleichungssystem wird dann einmal mit $\varepsilon_h = 1,025$ und $\varepsilon_f = 1$ und einmal mit $\varepsilon_h = 1$ und $\varepsilon_f = 1,025$ gerechnet. Daraus ergeben sich dann die Gleichgewichtslinien für das Inland und das Ausland. Berücksichtigt man ε_h und ε_f und $L_{h2} = 1 - L_{h1}$ sowie $L_{f2} = 1 - L_{f1}$ und bezeichnet vereinfachend L_{h1} mit L_h und L_{f1} mit L_f , so können Gleichungen (4.60) bis (4.67) umgeschrieben werden zu:¹¹⁶

¹¹⁵Aufgrund der anderen Modellstruktur wirkt der Größenunterschied hier stärker als dies in dem in Abschnitt 4.2 behandelten Modell der Fall ist. Das ist jedoch kein generelles Ergebnis, sondern liegt an der Wahl der Parameterwerte. Daher wurde hier ein geringerer Größenunterschied gewählt.

¹¹⁶Gleichungen (4.68) bis (4.71) bleiben unverändert.

$$Q_{h1} = \left[\varepsilon_h L_h w_{h1}^{1-\beta\sigma} Q_{h1}^{-\alpha\sigma} Q_{h2}^{-v\sigma} + \varepsilon_f L_f w_{f1}^{1-\beta\sigma} Q_{f1}^{-\alpha\sigma} Q_{f2}^{-v\sigma} T_1^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.82)$$

$$Q_{h2} = \left[\varepsilon_h (1 - L_h) w_{h2}^{1-\beta\sigma} Q_{h2}^{-\alpha\sigma} Q_{h1}^{-v\sigma} + \varepsilon_f (1 - L_f) w_{f2}^{1-\beta\sigma} Q_{f2}^{-\alpha\sigma} Q_{f1}^{-v\sigma} T_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.83)$$

$$Q_{f1} = \left[\varepsilon_f L_f w_{f1}^{1-\beta\sigma} Q_{f1}^{-\alpha\sigma} Q_{f2}^{-v\sigma} + \varepsilon_h L_h w_{h1}^{1-\beta\sigma} Q_{h1}^{-\alpha\sigma} Q_{h2}^{-v\sigma} T_1^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.84)$$

$$Q_{f2} = \left[\varepsilon_f (1 - L_f) w_{f2}^{1-\beta\sigma} Q_{f2}^{-\alpha\sigma} Q_{f1}^{-v\sigma} + \varepsilon_h (1 - L_h) w_{h2}^{1-\beta\sigma} Q_{h2}^{-\alpha\sigma} Q_{h1}^{-v\sigma} T_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (4.85)$$

$$E_{h1} = \mu \cdot (w_{h1} \varepsilon_h L_h + w_{h2} \varepsilon_h (1 - L_h)) + \frac{\alpha w_{h1} \varepsilon_h L_h + v w_{h2} \varepsilon_h (1 - L_h)}{\beta} \quad (4.86)$$

$$E_{h2} = (1 - \mu) \cdot (w_{h1} \varepsilon_h L_h + w_{h2} \varepsilon_h (1 - L_h)) + \frac{v w_{h1} \varepsilon_h L_h + \alpha w_{h2} \varepsilon_h (1 - L_h)}{\beta} \quad (4.87)$$

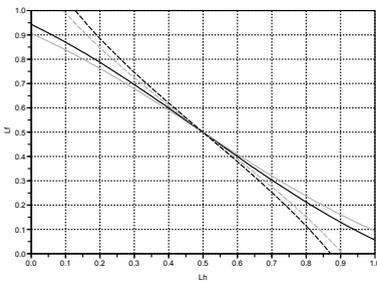
$$E_{f1} = \mu \cdot (w_{f1} \varepsilon_f L_f + w_{f2} \varepsilon_f (1 - L_f)) + \frac{\alpha w_{f1} \varepsilon_f L_f + v w_{f2} \varepsilon_f (1 - L_f)}{\beta} \quad (4.88)$$

$$E_{f2} = (1 - \mu) \cdot (w_{f1} \varepsilon_f L_f + w_{f2} \varepsilon_f (1 - L_f)) + \frac{v w_{f1} \varepsilon_f L_f + \alpha w_{f2} \varepsilon_f (1 - L_f)}{\beta} \quad (4.89)$$

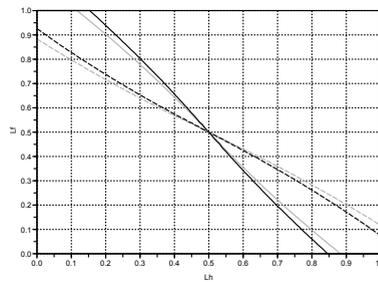
Abbildung 4.17 stellt das kurzfristige Gleichgewicht (L_h) im Inland bei gegebener Verteilung der Arbeitskräfte im Ausland (L_f) und umgekehrt dar. Die Ausgabenanteile wurden hier wiederum mit $\mu = 0,5$ angenommen. Dabei sind für den Fall einer Größendifferenz von 2,5% mit schwarzen Linien die Gleichgewichte des kleineren Auslands (gestrichelt) und die des größeren Inlands (durchgezogen) dargestellt. Die grauen Linien zeigen zum Vergleich die Situation mit gleich großen Ländern. Es ist zu erkennen, dass die Größe des Landes bei gleich großen Sektoren im Ausland ($L_f = 0,5$) keinen Einfluss auf deren Verteilung im Inland hat. Ist der Sektor 1 im Ausland kleiner als die Hälfte, so ist dieser im Inland umso kleiner, je größer das Land ist. Das kann daran abgelesen werden, dass in Abb. 4.17a bei $L_f < 0,5$ die schwarzen Linien weiter links liegt als die grauen. Für $L_f > 0,5$ gilt der entgegengesetzte Fall.

Ökonomisch lässt sich das wie folgt begründen: Je größer ein Land ist, desto weniger ist es vom Ausland abhängig. Daher verläuft seine Gleichgewichtslinie mit immer stärkerer Steigung und wäre bei extremen Größenunterschieden eine ver-

tikale Linie bei 0,5, wobei die hälftige Aufteilung in diesem Extremfall auf die Exponenten der Nutzenfunktion zurückgeht. Die Landesgröße nimmt dabei auf zwei Arten Einfluss auf das Gleichgewicht: Zum einen geht sie direkt in die Ausgabefunktion ein. Je mehr Menschen in einem Land leben, desto größer sind die Ausgaben für beide Gütergruppen (Nachfrageeffekt). Zum anderen geht die Landesgröße in den Preisindex ein, weil dieser von der Anzahl der Produktvarianten positiv abhängt, denn die Outputmenge je Unternehmen ist konstant. Mehr Einwohner bedeuten aber auch mehr Varianten und daher c.p. ein höheres Preisniveau (Preiseffekt). Diese beiden Effekte wirken gegenläufig, gleichen sich aber nicht aus. Da aber z.B. ein kleines Ausland selbst bei vollständiger Spezialisierung auf einen Sektor absolut gesehen nur wenig anbietet und von der anderen Gütergruppe auch nur wenig nachgefragt wird, sinkt der Einfluss des Auslands, je kleiner dieses im Vergleich zum Inland ist.



(a) Bei niedrigen Transportkosten ($T = 2,0$)



(b) Bei hohen Transportkosten ($T = 3,0$)

Abbildung 4.17: Gleichgewichtskurven für unterschiedlich große Länder

Ungeachtet dieser Effekte ergeben sich bei sehr geringen und sehr hohen Transportkosten die gleichen langfristigen Gleichgewichte, die aus dem symmetrischen Modell bekannt sind. In Abb. 4.17a liegen bei vollkommener Spezialisierung beider Länder Gleichgewichte vor. In Abb. 4.17b ist das einzige stabile Gleichgewicht durch eine hälftige Aufteilung beider Industrien gegeben. Ein Unterschied ergibt sich aber nun bei der Betrachtung der mittleren Transportkosten. In Abb. 4.18 wurden Transportkosten von 2,4 unterstellt. Es ist zu erkennen, dass in den beiden Gleichgewichten mit Spezialisierung das kleinere Ausland sich jeweils auf einen Sektor spezialisiert, im größeren Inland aber keine vollständige Spezialisierung auftritt. Hier erkennt man den oben beschriebenen Nachfrageeffekt.

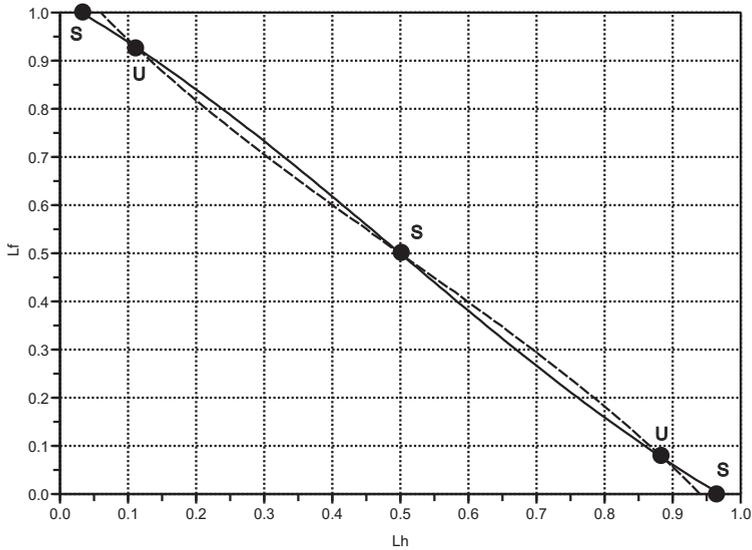
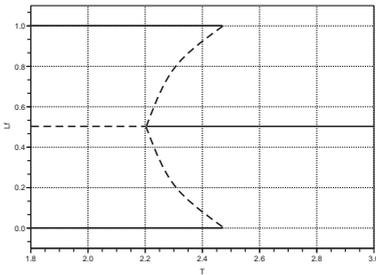


Abbildung 4.18: Asym. Länder, mittlere Transportkosten ($\varepsilon = 0,025$, $T = 2,4$)

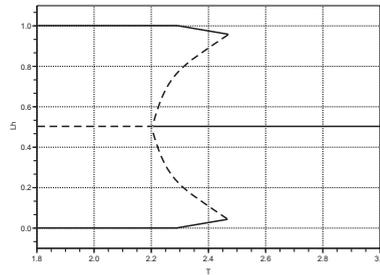
Auch diese Modellvariante kann nun in Form eines Bifurkationsdiagramms dargestellt werden, in dem der Transportkostensatz variiert. In diesem Diagramm werden auf der Abszisse die Transportkosten abgetragen und auf der Ordinate der Anteil der Beschäftigten, die in einem Land in Sektor 1 arbeiten. Ein solches Diagramm bezieht sich also immer auf ein Land. Daher ist für das kleine und das große Land hier je ein Diagramm notwendig.¹¹⁷

In Abbildung 4.19 kann die Lage der stabilen und instabilen Gleichgewichte im kleineren und im größeren Land in Abhängigkeit von der Höhe der Transportkosten abgelesen werden. Dabei lässt sich zunächst feststellen, dass Sustain- und Break-Point in beiden Ländern identisch sind. Bei den hier gewählten Werten der Parameter liegt der Break-Point \underline{T} bei ca. 2,2 und der Sustain-Point \bar{T} bei ca. 2,47. Allerdings tritt im Fall des kleineren Landes bei Transportkosten unter \bar{T} stets vollständige Spezialisierung auf ($L_f = 0$ oder $L_f = 1$), während beim größeren Land ab ca. $T = 2,29$ die Spezialisierung nicht mehr vollständig ist. Die öko-

¹¹⁷Im Basismodell und in der Variante mit unterschiedlichen Transportkosten war das nicht notwendig, da die Länder hier identisch waren.



(a) kleineres Land



(b) größeres Land

Abbildung 4.19: Großes und kleines Land im Cluster-Modell

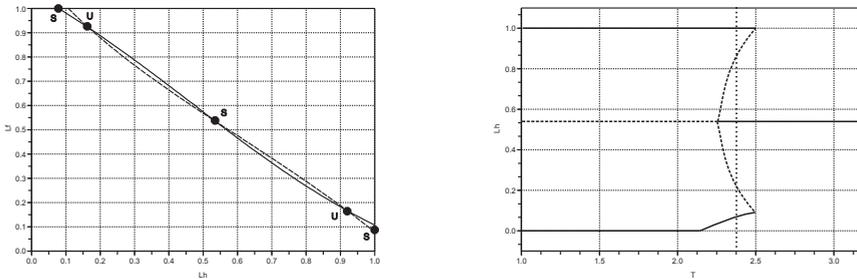
nomische Erklärung ist folgende: Angenommen, das kleine Land (Ausland) produziert ausschließlich Gut 2 und das große Land (Inland) ausschließlich Gut 1. Dann stünden insgesamt weniger Arbeitskräfte für die Produktion von Gut 2 zur Verfügung als für Gut 1. Es würde weniger Gut 2 produziert und die Produktion ist aufgrund der geringeren Skaleneffekte teurer. Auch die Nachfragemengen nach beiden Gütern sind nicht identisch, allerdings bleibt die Nachfrage der Konsumenten konstant, während sich nur die Nachfrage nach Zwischenprodukten von Gut 2 verringert (wegen $\mu > \nu$, vgl. auch Gleichungen (4.86) bis (4.89)). Da steigende Transportkosten die im Ausland produzierten Güter für Nachfrager im Inland weiter verteuern, lohnt es sich, auch im Inland eine geringe Menge von Gut 2 herzustellen.

4.6 Spezialisierung bei unterschiedlich großen Sektoren

In Krugman und Venables (1996) und der darauf aufbauenden Literatur wird der Parameter μ , der die Ausgabenanteile der Haushalte wiedergibt, auf 0,5 festgelegt. Die Darstellung des Modells in Abschnitt 4.3 ist an dieser Stelle allgemeiner gehalten, so dass es relativ einfach ist, zu untersuchen, welche Konsequenzen es hat, wenn die Haushalte ihr Einkommen in einem anderen Verhältnis auf die beiden Gütergruppen aufteilen. Eine formale Änderung des Modells ist nicht not-

wendig.

Für die Simulation wurde ein Wert von $\mu = 0,55$ angenommen. Das bedeutet, dass die Haushalte in beiden Ländern 55% ihres Einkommens für Produkte des Sektors 1 ausgeben und entsprechend 45% für Produkte des Sektors 2. Auf die detaillierte Darstellung der Fälle mit niedrigen, mittleren und hohen Transportkosten soll hier verzichtet werden. Abbildung 4.20b gibt das Bifurkationsdiagramm an. Für den mit einem roten Balken markierten Fall mittlerer Transportkosten (hier $T = 2,375$) ist in Abbildung 4.20a zusätzlich das Diagramm der kurzfristigen Gleichgewichte gegeben. Dabei wurden stabile Gleichgewichte mit S, instabile mit U gekennzeichnet, was im Bifurkationsdiagramm durch gestrichelte und durchgezogene Linien gekennzeichnet ist.



(a) Gleichgewichte im In- und Ausland ($\mu = 0,55$; $T = 2,375$) (b) Bifurkationsdiagramm bei asymmetrischen Sektoren ($\mu = 0,55$)

Abbildung 4.20: Das Cluster-Modell bei unterschiedlich großen Sektoren

Zunächst ist festzustellen, dass das symmetrische Gleichgewicht, das gilt sowohl für das instabile unterhalb des Break-Points als auch für das stabile darüber, etwas unterhalb von 0,55 liegt. Der Grund liegt darin, dass eine Ausgabenerhöhung der Haushalte für diesen Sektor sich in vollem Umfang auf die direkte Nachfrage auswirkt, der Multiplikator über die indirekte Nachfrage aber kleiner als 1 ist. Daher steigt der zweite Ausdruck in den Ausgabengleichung (4.64) und (4.66) unterproportional an. Außerdem wird die indirekte Nachfrage nach Gütern des Sektors 1 verringert, weil bei einer Erhöhung von μ automatisch die Ausgaben für den anderen Sektor sinken, in deren Produktion die Güter aus M_1 als Vorprodukte eingehen. Insgesamt ist dieser Effekt aber schwach.

Ein recht deutlicher Effekt der Zunahme von μ ist erkennbar, wenn man den Be-

reich mittlerer Transportkosten betrachtet. Spezialisiert sich ein Land hier auf Sektor 1, so kann es dazu kommen, dass auch das andere Land eine geringe Menge der Produktion dieses Sektors herstellt. Graphisch gesehen entfernen sich die Enden nahe dem Sustain-Point von der jeweiligen Achse. Da die Länder als symmetrisch angesehen werden, ist unbestimmt, in welchem Land dies geschieht.

Ökonomisch kann dieses Phänomen wie folgt erklärt werden: Eine Erhöhung des Ausgabenanteils μ lässt die Ausgaben des Sektors 1 (E_{1h} und E_{1f}) steigen und die des zweiten Sektors sinken. Dadurch verschiebt sich das Lohnverhältnis zugunsten von w_1 in beiden Ländern.¹¹⁸ Dies erklärt die Verlagerung beider Kurven nach rechts oben. Die Erklärung für unvollständige Spezialisierung ist die gleiche, wie im vorangegangenen Abschnitt, nur dass diese jetzt nur im größeren Sektor auftreten kann. Diese Betrachtung gilt allerdings nur, wenn die Größendifferenz nicht zu groß wird. Werden extreme Unterschiede unterstellt, kommt es nicht zu einer Konvergenz der Löhne, weil durch die Agglomerationskräfte (Skaleneffekte) die Löhne in einem Sektor dann immer weiter steigen würden.

4.7 Spezialisierung bei verschiedenen Asymmetrien

Bisher wurden drei Arten von Asymmetrien untersucht: Unterschiedliche Transportkosten je Sektor, unterschiedliche Größen der Länder und unterschiedliche Ausgabenanteile für die Sektoren. In allen Fällen kommt es bei bestimmten Parameterkonstellationen zu Gleichgewichten, in denen in mindestens einem Land unvollständige Spezialisierung auftritt. Darunter verstehen wir, dass dieses Land einen großen Teil eines Sektors beinhaltet aber auch ein geringer Teil anderer Güter dort produziert wird.

Kombiniert man diese Asymmetrien miteinander, so verstärken sich deren Effekte oder heben sich gegenseitig auf. Beispielhaft soll hier gezeigt werden, welche Ergebnisse bei unterschiedlich großen Ländern ($\varepsilon = 0,025$) und unterschiedlichen Ausgabenanteilen ($\mu = 0,52$) entstehen.

Betrachtet man die beschriebenen Asymmetrien einzeln, wie in den letzten Abschnitten geschehen, ergibt sich unvollständige Spezialisierung entweder im oberen und unteren Bereich in gleicher Ausprägung (im kleineren Land bei Betrachtung

¹¹⁸Vgl. hierzu auch Gleichung (4.81).

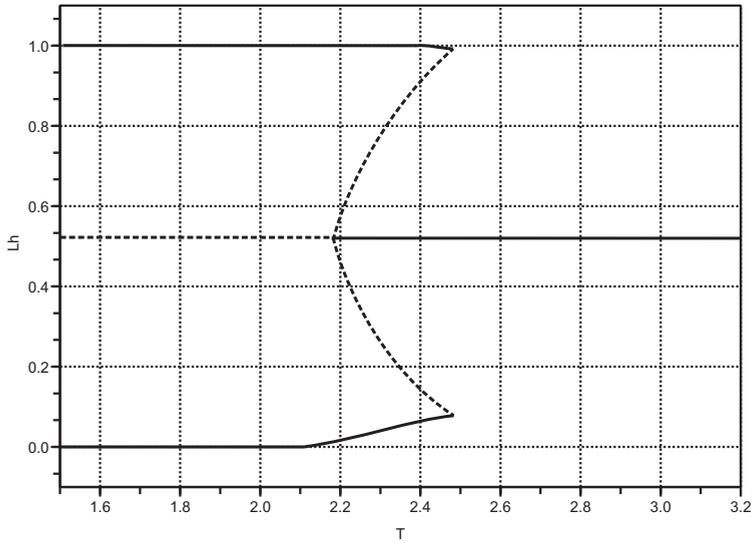


Abbildung 4.21: Unterschiedlich große Länder und Sektoren ($\varepsilon = 0,025$; $\mu = 0,52$)

tung von Größenunterschieden) oder aber ausschließlich in einem Bereich (z.B. bei ungleichen Ausgabenanteilen). In der Kombination ist es nun möglich, dass z.B. bei Konzentration auf Sektor 1 noch 10% von Sektor 2 im Land verbleiben; bei alternativer Konzentration auf Sektor 2 jedoch nur 5% von Sektor 1. Das entspricht in etwa dem in Abb. 4.21 dargestellten Fall. Dort ist das Bifurkationsdiagramm für das größere Land dargestellt. Im kleinen Land kommt es, außer bei extremen Werten von μ , immer zur vollständigen Spezialisierung.

4.8 Gleichgewichte und Löhne

An dieser Stelle soll noch einmal auf das Lohnsatzverhältnis in den verschiedenen möglichen Gleichgewichtssituationen des Cluster-Modells eingegangen werden. Diese Überlegungen gelten für alle hier betrachteten Varianten. Bei hohen Transportkosten bildet sich immer ein Gleichgewicht, dass durch einen Schnittpunkt der L_h^* - und der L_f^* -Kurve gekennzeichnet ist. In diesem Punkt gibt es in beiden

Ländern einen Lohnausgleich zwischen den Sektoren, die Löhne können sich aber zwischen den Ländern unterscheiden.

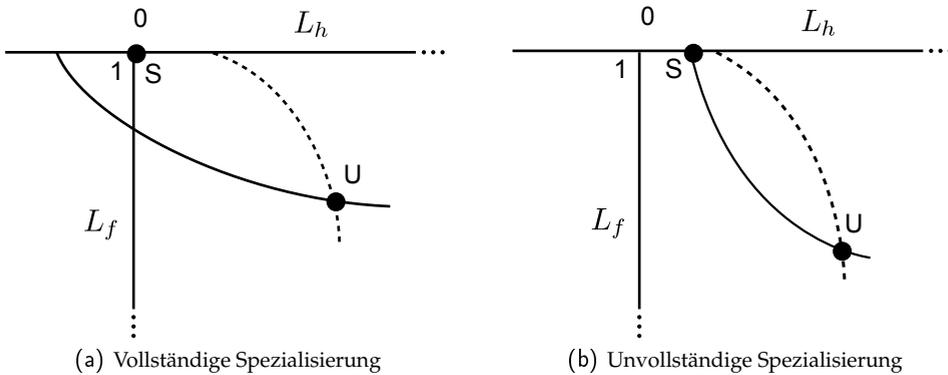


Abbildung 4.22: Gleichgewichte mit Spezialisierung bei $L_f = 1$

In den anderen Gleichgewichten, bei vollständiger oder unvollständiger Spezialisierung, ist das nicht der Fall. Bei vollständiger Spezialisierung (Abbildung 4.22a) herrscht in beiden Ländern ein Lohngefälle. Im Inland gilt $w_{h1} < w_{h2}$ und im Ausland $w_{f1} > w_{f2}$. Wenn jedoch unvollständige Spezialisierung vorliegt (in Abb. 4.22b z.B. im Inland), so gibt es in diesem Land einen Ausgleich der Löhne, im anderen jedoch nicht.

Mit dieser Überlegung kann nun erklärt werden, warum bei allen betrachteten Asymmetrien die unvollständige Agglomeration erst ab einem bestimmten Transportkostensatz auftritt. Die Transportkosten reduzieren, das wurde bereits oben beschrieben, die Löhne. Im symmetrischen Modell werden die Löhne in beiden Sektoren gleich stark reduziert; bei den verschiedenen Asymmetrien geschieht dies in unterschiedlicher Weise. Dennoch wird dieser Effekt erst wirksam, wenn die Löhne eines Landes ausgeglichen sind. Erst dann, wechseln nicht alle Arbeiter den Sektor und erst dann wird der Knick im Bifurkationsdiagramm sichtbar.

4.9 Zusammenfassung

In diesem Kapitel wurden die beiden bedeutendsten Modelle für internationale Agglomeration und Spezialisierung dargestellt und verschiedene Symmetriean-

nahmen aufgehoben. Obwohl beide Modelle strukturell sehr ähnlich sind, haben sie doch einen unterschiedlichen Erklärungsgehalt. Während das Modell mit Zwischenprodukten zeigt, dass die Ergebnisse des Core-Periphery-Modells auch ohne Faktormobilität möglich sind, und daher Agglomeration und Entleerung zum Gegenstand hat, erklärt das Cluster-Modell die Spezialisierung von Regionen, bzw. die räumliche Konzentration von Branchen. Aufgrund der höheren Komplexität bietet dieses auch mehr Möglichkeiten, Asymmetrien zu untersuchen.

Im Modell mit Zwischenprodukten wurde der Fall unterschiedlich großer Länder untersucht. Dabei ergibt sich ein sehr ähnliches Bild wie im Core-Periphery-Modell mit unterschiedlich großen Regionen. Das wichtigste Ergebnis ist hier, dass sich bei fallenden Transportkosten die Agglomeration immer im größeren Land bildet. Modelliert man den gleichen Sachverhalt im Cluster-Modell, tritt dieses Ergebnis nicht auf. Das ist dadurch zu erklären, dass die Agglomerationskräfte, die durch steigende Skalenerträge begründet sind, nun in beiden Sektoren auftreten. In dieser Modellvariante kann aber gezeigt werden, dass es im größeren Land zu stabilen Gleichgewichten bei unvollständiger Spezialisierung kommen kann, was im Grundmodell nicht möglich ist. Die Betrachtung unterschiedlich hoher Transportkosten und unterschiedlich großer Ausgabenanteile für die Sektoren, zeigen ein sehr ähnliches Resultat. In beiden Fällen kann es bei einem Land zu unvollständiger Spezialisierung kommen. Dies ist auch verständlich, denn in beiden Fällen werden die Agglomerationsmechanismen eines Sektors verstärkt oder geschwächt.

5 Simulationsmethoden für die Neue Ökonomische Geographie

5.1 Einleitung

Große Teile und wichtige Ergebnisse dieser Arbeit basieren auf Computersimulationen. Aus diesem Grund werden in diesem Kapitel eine Einführung in die Methoden und ein Überblick über die technische Umsetzung gegeben. Dies stellt weder ein Handbuch zu der Simulationssoftware dar, noch eine Einführung in numerische Lösungsverfahren allgemein, sondern ist eine Anwendung von beiden auf die Modelle der NEG.

Eine hervorragende Einführung in computergestützte Verfahren bietet Fackler und Miranda (2002), sowie auf einer eher theoretischen Ebene Judd (1998). Das verwendete Softwareprogramm Scilab wird detailliert beschrieben in Pincon (k.A.) oder Zogg (2007). Da die Syntax jedoch der von Matlab sehr ähnlich ist, können auch Bücher zu Matlab verwendet werden. Der einzige mir bekannte Text, der sich mit der Simulation von NEG-Modellen beschäftigt, ist Brakman et al. (2001) und auch das nur im Umfang von weniger als einer Seite. Diese sind jedoch wenig technisch und helfen daher bei der Implementierung von Simulationen nur bedingt weiter. Die vorliegende Arbeit soll diese Lücke schließen und betrachtet systematisch die verschiedenen Aspekte der Simulation von Modellen der Neuen Ökonomischen Geographie.

Numerische Verfahren sind immer dann ein Ausweg, wenn zu einem Modell keine analytische Lösung gefunden werden kann. In solchen Fällen gibt es zwei Möglichkeiten: Das Modell wird so weit vereinfacht, bis es analytisch lösbar ist oder die Lösung wird durch Simulation bestimmt. Eine solche Lösung hat nun den Nachteil, dass sie keinen Anspruch auf Allgemeingültigkeit erheben kann. Allerdings ist es oft möglich, durch ökonomische Überlegungen Rückschlüsse auf die

Allgemeingültigkeit bei realistischen Parameterwerten zu ziehen. Dennoch stehen numerische Verfahren bei einigen Ökonomen wegen dieses Problems in der Kritik.¹¹⁹ Doch auch wenn es auf dem Gebiet der NEG einige analytisch lösbarere Modelle gibt,¹²⁰ so muss für die meisten doch auf numerische Methoden zurückgegriffen werden. In dieser Arbeit soll der Meinung von Fackler und Miranda gefolgt werden:

“We believe that it is better to derive economic insights from a realistic numerical model of an economic system than to derive irrelevant results, however general, from an unrealistic but explicitly soluble model.” (Fackler und Miranda (2002), S. xvi)

Die hier vorgestellten Modelle lassen sich in Form eines nicht-linearen Gleichungssystems zusammenfassen. Je nach Modell kann dies aus unterschiedlich vielen Gleichungen bestehen. Auch kann es vorkommen, dass nicht alle endogenen Variablen isoliert werden können. Dies ist z.B. beim Cluster-Modell der Fall und macht eine weitere Simulationsstufe erforderlich.

Im einfachsten Fall sind zur Lösung eines nicht-linearen Gleichungssystems durch Iteration folgende Schritte notwendig:¹²¹

1. Festlegung der exogenen Variablen.
2. Schätzung sinnvoller Startwerte für die endogenen Variablen.
3. Berechnung “neuer” Werte für die Lösungsvariablen.
4. Dies wird so lange wiederholt, bis sich alte und neue Werte der Lösungsvariablen nicht mehr zu stark unterscheiden.

Dabei ist durch einen Toleranzwert festzulegen, was unter “nicht mehr zu stark” zu verstehen ist.

Ergebnis dieser Simulation ist ein kurzfristiges Gleichgewicht, wobei der Begriff “kurzfristig” zwar in der Literatur durchgehend verwendet wird, aber nicht ganz treffend ist. Hierunter wird eine Konstellation von Werten für die endogenen Variablen verstanden, bei dem alle Gleichungen des Systems erfüllt sind - aber vor dem Einsetzen des Ausgleichsmechanismus (z.B. Wandern der Arbeiter). Bezieht

¹¹⁹Speziell bei NEG Modellen vgl. Baldwin et al. (2003), S. 2.

¹²⁰Z.B. das Footloose Capital Model von Martin und Rogers (1995) oder das Footloose Entrepreneur Model von Forslid und Ottaviano (2003).

¹²¹Vgl. Brakman et al. (2001), S. 102 und Fackler und Miranda (2002), S. 16ff.

man diesen mit ein, spricht man von einem langfristigen Gleichgewicht. Das kurzfristige Gleichgewicht wird in der Regel graphisch dargestellt, so dass auf der entstandenen Abbildung ein langfristiges Gleichgewicht ablesbar ist. Möchte man aber dieses langfristige Gleichgewicht nicht nur ablesen, sondern genau ermitteln, so muss auch dies durch das Programm erfolgen.¹²² Dies ist auch notwendig, um z.B. das verbreitete Bifurkationsdiagramm zu erzeugen oder um die Ergebnisse bei unterschiedlichen Parameterwerten zu vergleichen. Dazu sind Programmteile notwendig, die die Ergebnisse zwischenspeichern, abrufen und weiterverarbeiten. Auch diese werden im Folgenden erklärt.

Auch wenn die verschiedenen Modelle in den Gleichungen und auch im Ausgleichsmechanismus Unterschiede aufweisen, so ist doch das grundsätzliche Vorgehen bei ihrer Lösung gleich. In jedem Fall werden zunächst beide Seiten des Ausgleichsmechanismus (z.B. Arbeitsangebot und -nachfrage oder inländische und ausländische Aufteilung der Sektoren) bei gegebenen Transportkosten simuliert. Danach müssen die Schnittpunkte gefunden werden. Dies wird für verschiedene Werte der Transportkosten (oder anderer exogener Parameter, z.B. den lokalen Präferenzen im LW-Modell) wiederholt, um die Abhängigkeit der langfristigen Gleichgewichte von den externen Werten darzustellen. Beide Schritte werden durch Abbildungen dargestellt. Dieses allgemeine Vorgehen wird in Abschnitt 5.3 unter Verwendung verschiedener Beispiele erläutert, nachdem im Abschnitt 5.2 zunächst eine Einführung in die verwendete Software und deren Syntax gegeben wurde. Der letzte Abschnitt fasst die Ergebnisse zusammen.

5.2 Computersimulation

Verwendete Software Für die Berechnungen wurde das Programm *Scilab* genutzt. Es wird von einem Konsortium unter der Leitung des *Institut National de Recherche en Informatique et en Automatique (INRIA)* entwickelt und ist frei auf der Homepage¹²³ für verschiedene Betriebssysteme erhältlich. Für diese Arbeit wurden die Versionen 4.1.2 und 5.0.3 verwendet. Auch wenn Scilab einen eigenen Editor mitbringt, so ist bei intensiver Arbeit mit dem Programm die Benutzung

¹²²Programm meint hier den Simulationscode. Diese Begriffe werden hier, wie in der Informatik üblich, synonym benutzt.

¹²³<http://www.scilab.org>.

eines komfortableren externen Editors empfehlenswert. Hier wurde *Notepad++*¹²⁴ benutzt. Aufgrund der großen Ähnlichkeit der Syntax zu Matlab, kann die Syntaxhervorhebung genutzt werden.

Im Editor werden für die einzelnen Aufgaben bei der Simulation Skripte geschrieben, die als Text-Datei gespeichert und dann mit Scilab ausgeführt werden. Es handelt sich also um eine interpretierte und nicht um eine kompilierte Programmiersprache mit den entsprechenden Konsequenzen für die Performance.¹²⁵ Die erzeugten Ergebnisse können als Text-Dateien (für Matrizen) oder als EPS-Graphiken gespeichert werden.

Arbeiten mit Werten und Matrizen Variablen können auf einfache Weise mit Werten belegt werden. Der Ausdruck `a=5` weist der Variablen `a` den Wert 5 zu, während `a==5` den Wert der Variablen `a` mit der Zahl 5 vergleicht und dann `true` oder `false` zurückgibt. Ungleichheit wird dabei mit `~` symbolisiert. "Größer als" und "kleiner als" entsprechen der mathematischen Formulierung. Solche Vergleiche werden häufig in `if`-Abfragen verwendet. Variablen müssen, im Gegensatz zu vielen anderen Programmiersprachen, nicht vorher deklariert werden.¹²⁶

Matrizen werden z.B. mit

```
m = [1.6 1.7 1.8 1.9; 1.0 2.3 1.5 3.0];
```

für die Matrix

$$m = \begin{bmatrix} 1,6 & 1,7 & 1,8 & 1,9 \\ 1,0 & 2,3 & 1,5 & 3,0 \end{bmatrix}$$

erzeugt und mit Werten versehen. Vektoren sind Matrizen mit nur einer Zeile oder Spalte und werden entsprechend genauso behandelt. Die Elemente einer Zeile sind mit Leerzeichen getrennt. Die einzelnen Zeilen werden mit einem Semikolon abgegrenzt. Möchte man auf einzelne Elemente aus einer Matrix zugreifen, so geschieht das mit

¹²⁴<http://notepad-plus.sourceforge.net>.

¹²⁵Kompilierte Sprachen werden vor der Ausführung in Binärcode übersetzt, der dann bedeutend schneller ausgeführt werden kann, z.B. C und C++. Beispiele für interpretierte Sprachen sind PHP, Ruby oder Java Script. Zur Performance siehe auch weiter unten im folgenden Unterabschnitt. Bei besonders zeitaufwendigen Berechnungen können aber auch kompilierte Funktionsbausteine aus C oder FORTRAN eingebunden werden.

¹²⁶Unter der Deklaration von Variablen versteht man, dass diese vor dem eigentlichen Programmcode benannt und mit einem Datentyp versehen wird, wie dies beispielsweise in Pascal der Fall ist.

```
a=m(i,j)
```

Diese Zeile weist den Wert der Matrix *m* in Zeile *i* und Spalte *j* der Variablen *a* zu. Dies wird häufig in Schleifen (s.u.) verwendet. Eine ganze Zeile oder Spalte kann mit

```
m(:,4)
```

ausgelesen werden. Diese Schreibweise ist z.B. beim Erstellen von Abbildungen hilfreich. Soll ein Vektor gleichmäßig mit Werten zwischen einem Start- und einem Zielwert gefüllt werden, kann dies mit

```
L = linspace(0,1,1001)';
```

geschehen. Hier wird *L* mit 1001 Werten zwischen 0 und 1 (Intervallgrenzen eingeschlossen) gefüllt. Mit Vektoren und Matrizen kann entsprechend den mathematischen Gesetzen gerechnet werden. So füllt z.B. die Anweisung

```
x = d-s;
```

den Vektor *x* mit den jeweiligen Differenzen zwischen den Elementen der Vektoren *d* und *s*. Entsprechend sind auch alle anderen zulässigen Operationen mit Matrizen möglich. Ist die Größe einer Matrix nicht vor dem Start des Programms bekannt, kann sie auch ermittelt werden. Z.B.:

```
m = [1.6 1.7 1.8 1.9];
t = size(m);
s = t(2);
```

Dabei ermittelt `size()` die Anzahl der Zeilen und die Anzahl der Spalten. Da hier nur die Anzahl der Spalten ausgelesen werden soll, muss aus dem Vektor *t* das zweite Element verwendet werden. In diesem Beispiel hätte *s* den Wert 4.

Kontrollstrukturen Auch Scilab unterstützt die üblichen Kontrollstrukturen, wie Verzweigungen und verschiedene Arten von Schleifen. Verzweigungen werden mit

```
if expr1 then statements
elseif expr2 then statements
....
else statements
end
```

realisiert. Dabei ist der `else`- und der `elseif`-Block optional. Eine solche Verzweigung ist z.B. in Zeile 20-24 des Programms in Abschnitt 5.3.1 zu finden. Verzweigungen können auch verschachtelt werden. Unterstützt wird auch die `case`-Anweisung, die hier aber nicht verwendet wird.

Sich wiederholende Programmteile können als `for`-Schleife oder als `while`-Schleife implementiert werden. Die aus anderen Programmiersprachen bekannte Variante mit `repeat until` gibt es nicht.

```
for i = 1:5
...
end;

while a < 5
a = a+1;
end;
```

In Simulationen sind inhaltlich `while`-Schleifen sinnvoll, da eine Schleife meist solange durchlaufen werden soll, bis ein bestimmtes Kriterium erfüllt ist, z.B. die Differenz zwischen altem und neuem Wert einer Variablen unter einen Schwellenwert fällt. Der Nachteil von `while`-Schleifen ist jedoch, dass a priori nicht sicher ist, ob das Abbruchkriterium überhaupt erfüllt wird. Wenn ein Gleichungssystem z.B. im Laufe der Simulation divergiert, läuft das Programm auf unbestimmte Zeit und muss manuell abgebrochen werden. Aus diesem Grund werden hier `for`-Schleifen mit einem sehr hohen Zielwert für die Iterationsvariable verwendet, die bei Erreichen des Abbruchkriteriums durch einen `break`-Befehl unterbrochen werden.

Führt ein Durchlauf einer Schleife möglicherweise zu Fehlern, gibt es die Möglichkeit, die Ausführung bei Auftreten des Fehlers abzubrechen und beim nächsten Durchlauf der Schleife fortzufahren. Dazu dient das `try-catch` Konstrukt.

```
for i = 1 : length(T)
    try
        // hier kann ein Fehler auftreten
    catch
        // wird nur bei Fehlern ausgeführt
    end
end;
```

Einen weiteren Block, der sowohl mit als auch ohne Fehler ausgeführt wird, wie dies in einigen Programmiersprachen üblich ist (`finally`-Block), gibt es bei Scilab

nicht.

Datenverwaltung Daten können zur späteren Weiterverwendung in Dateien geschrieben und daraus wieder eingelesen werden. Durch

```
fprintfMat(f,m);
```

wird eine Matrix m in die Datei f geschrieben. f ist dabei ein String (eine Zeichenkette), der den kompletten Pfad zur Datei enthält. Zum Einlesen wird

```
m = fscanfMat(f);
```

verwendet. Dabei ist es oft sinnvoll, die verwendeten Parameterwerte in den Dateinamen aufzunehmen. Um z.B. die Matrix mit den Simulationsergebnissen für einen bestimmten Wert der Transportkosten T zu speichern, kann der Dateiname f festgelegt werden mit

```
f = '\data\matrix-' + string(T*100) + '.txt';
```

Für $T = 2,1$ würde sich z.B. der Dateiname `matrix-210.txt` ergeben. Das ist sinnvoll, da die eigentliche Simulation sehr zeitaufwendig ist. Das Einlesen einer Datei aber schnell geht.

Modularisierung Wird ein Programm sehr groß und werden die gleichen Programmteile an verschiedenen Stellen verwendet, bietet es sich an, das Programm in Module zu unterteilen. So wurde zur Simulation des Cluster-Modells in Abschnitt 4.3 die Bibliothek `lib_cm.sci` angelegt, in die verschiedene Funktionen ausgelagert wurden. Das eigentliche Programm ruft dann nur noch diese Funktionen auf und wird dadurch erheblich übersichtlicher. Damit Funktionen aus einer anderen Datei verwendet werden können, muss diese eingebunden werden. Dies geschieht mit dem Befehl

```
chdir(Verzeichnis);  
getf(Datei);
```

Wobei der erste Befehl das aktuelle Arbeitsverzeichnis ändert und der zweite dann die Datei einbindet. Auf die erste Anweisung kann auch verzichtet werden. Dann muss bei den eingebundenen Dateien aber der Pfad mit angegeben werden. Die Verwendung von `chdir()` ist besonders dann zu empfehlen, wenn das

Programm auch Dateien schreibt. Werden mehrere Dateien aus dem gleichen Verzeichnis eingebunden, muss das Verzeichnis nur einmal geändert werden. `getf()` wird dann ohne absoluten Pfad aufgerufen, z.B.:

```
chdir('D:\CM-Transportkostendifferenz\scilab');  
getf("lib_cm.sci");
```

In der externen Datei werden Funktionen mit dem Schlüsselwort `function` deklariert:

```
function [xy] = myFunction(param)  
xy = ...  
endfunction
```

Dabei ist `xy` der Ausgabewert der Funktion. Hier kann auch eine ganze Matrix stehen. Die Funktion `myFunction` wird dann mit dem Parameter `param` aufgerufen, der auch leer gelassen werden kann. Durch `endfunction` wird die Deklaration beendet. Der Aufruf aus dem Hauptprogramm sieht wie folgt aus:

```
a = myFunction(b);
```

Grafik Die grafischen Ausgabemöglichkeiten in Scilab sind vielfältig. Die Ergebnisse von Simulationen werden jedoch meist als zweidimensionaler `x-y`-Plot dargestellt. Daher soll sich die Beschreibung auf diese Art von Abbildungen beschränken. Der allgemeine Aufruf zum Erstellen einer 2D-Abbildung lautet¹²⁷

```
plot2d([x],y,<opt_args>)
```

Dabei enthält der Vektor `x` die Werte der horizontalen Achse, `y` die der vertikalen. Wird `x` weggelassen, nimmt das System den Vektor `1..n` an. Da das für unsere Zwecke wenig sinnvoll ist, sind hier immer beide Vektoren angegeben. In den optionalen Argumenten können Angaben zum Layout der Grafik gemacht werden, insbesondere können mit `style` Linien zur besseren Unterscheidung farbig dargestellt werden. Die folgende Zeile erzeugt z.B. eine Abbildung, in der die Werte der dritten Spalte gegen die der ersten Spalte der Matrix `c` durch eine grüne Kurve geplottet werden:

```
plot2d(c(:,1),c(:,3), style=3);
```

¹²⁷Der Befehl `plot()` sollte ist nur noch aus Gründen der Kompatibilität vorhanden und sollte vermieden werden. Er hat den Nachteil, dass die meisten Formatierungsanweisungen mit ihm nicht verwendet werden können.

Der Doppelpunkt bedeutet dabei "alle Werte dieser Spalte oder Zeile". Erhält der Parameter `style` einen Wert kleiner als eins, werden die einzelnen Wertepaare mit verschiedenen Symbolen dargestellt, aber nicht zu einer Linie verbunden. `style=0` zeichnet einzelne Punkte. Ein weiteres wichtiges Argument ist `rect`. Dadurch kann angegeben werden, welcher Wertebereich an den Achsen dargestellt wird, z.B.

```
plot2d(c(:,1),c(:,3), rect=[1,-0.1,3,1.1]);
```

Eine vollständige Auflistung aller möglichen Argumente findet sich in der Hilfe. Nicht alle Formatierungen können so durchgeführt werden. Einige sind auch in separaten Anweisungen anzugeben, wie z.B. die Linienart (durchgezogen, gepunktet oder gestrichelt) mit

```
xset('line_style',2);
```

oder die Achsenbeschriftungen mit

```
xlabel("L");  
ylabel("w1_-_w2");
```

wie dies z.B. in den letzten beiden Zeilen des Programms in Abschnitt 5.3.1 zu sehen ist. Soll ein Gitternetz mit abgebildet werden, kann dies mit

```
xgrid();
```

geschehen. Häufig ist es auch erwünscht, die Strichstärke der Linien zu ändern um eine bessere Lesbarkeit im Druck zu gewährleisten. Dies geschieht mit dem folgenden Befehl.

```
xset('thickness',3);
```

Wie auch die anderen unabhängigen Formatierungsanweisungen (also nicht die, welche als Optionen von `plot2d()` angegeben werden), gilt diese Anweisung von der Stelle im Code, in der sie steht, bis die Formatierung wieder geändert wird. Dies kann in diesem Fall geschehen, indem die Strichstärke wieder auf 1 zurückgesetzt wird.

Neben den hier beschriebenen grundlegenden Funktionen gibt es noch eine große Anzahl verschiedener mitgelieferter Funktionen, die zum Teil später verwendet werden (z.B. `horner()`, `datafit()`, `sign()`, `poly()`, `roots()`). Da diese nicht

zum allgemeinen Verständnis der Programme notwendig sind, sondern vielmehr spezielle Probleme lösen, werden sie an den Stellen im Text erläutert, an denen sie eingesetzt werden.

Performance und Optimierung Ein durchaus wichtiges Thema bei der Programmierung von Simulationen ist die Performance, also die Geschwindigkeit, mit der eine Simulation berechnet werden kann. Diese wird einerseits durch die verwendete Hardware bestimmt, andererseits durch die Programmierung selbst. Dabei kann durch sinnvolle Programmstrukturen mehr eingespart werden, als durch die Verwendung eines schnelleren PCs, da sich bei geschachtelten Schleifen, wie sie hier verwendet werden, die überflüssigen Operationen potenzieren. Sind z.B. drei Schleifen mit je 100 Durchläufen geschachtelt und wird die äußere Schleife zehn Durchgänge zu oft durchlaufen, läuft die innere bereits $10 \cdot 100 \cdot 100 = 100.000$ mal zu häufig. Dabei sind vor allem zwei Dinge zu beachten:

1. Wahl geeigneter Startwerte: Wird, wie z.B. im CP-Modell einen Lohnsatz in Abhängigkeit vor einer gegebenen Verteilung der Arbeitskräfte im Ausland berechnet, und wird diese Verteilung in jeder Iteration ein wenig verändert, so ist zu erwarten, dass die Lösungsvariable sich auch nur gering ändert. In jedem Fall sollte als Startwert das Ergebnis des letzten Durchlaufs verwendet werden. In der ersten Runde sollte ein Wert genommen werden, der aufgrund der Kenntnis des Modells in etwa zu erwarten ist.
2. Schleifen sollten nur so oft durchlaufen werden, bis sich der Zielwert nicht mehr über den Schwellenwert hinaus ändert. Viele Systeme konvertieren sehr schnell. Daher liegt hier ein erhebliches Potential, Rechnerzeit einzusparen.

Die Performance kann einerseits durch Zeitmessungen, andererseits durch Zählen der Durchläufe der einzelnen Schleifenebenen erfolgen. Letzteres ist erstens einfacher, zweitens sinnvoller. Aufgrund der verwendeten Konstruktion der Schleifen, die Endlosdurchläufe verhindert, muss unbedingt kontrolliert werden, dass die Schleifen nicht die maximale Durchlaufzahl erreichen. Dann wäre nämlich nicht sichergestellt, dass die Werte tatsächlich konvergieren. Es ist auch sinnvoll, die minimale Anzahl von Durchläufen zu zählen. Dazu wird wie folgt vorgegangen:

1. Initialisierung der Variablen, wobei die Variable für den maximalen Wert eine niedrige Vorbelegung bekommt und die für den minimalen Wert eine

hohe.

2. Nach jedem kompletten Durchlauf (also nicht nach jeder Iteration) wird der Zähler mit diesen beiden Werten verglichen. Liegt der Wert unter dem Minimum, wird er in diese Variable übernommen, liegt er über dem Maximum, wird er entsprechend in die Variable übernommen.
3. Am Ende müssen die Werte nur noch ausgegeben werden.

Es ist sinnvoll, sich einen Block mit diesen Anweisungen zu definieren, der dann an die entsprechenden Stellen kopiert wird. Der Code könnte wie folgt aussehen:

```
i_min = 100;
i_max = 0;
...
if k < i_min then i_min = k
    end;
if k > i_max then i_max = k
    end;
...
mprintf("i_min:_%f,i_max:_%f_\n\r", i_min, i_max);
```

Die letzte Zeile gibt die Werte der Variablen auf dem Bildschirm aus.

5.3 Lösung von NEG-Modellen

Die Umsetzung einer Simulation in Scilab soll anhand des einfachsten Falls, des Grundmodells von Krugman, demonstriert werden, das später zur Veranschaulichung weiterer Techniken um lokale Präferenzen nach Ludema und Wooton (1997) erweitert wird. Das Verfahren wird, wie in Abschnitt 5.1 beschrieben, in zwei Schritte unterteilt: Die Bestimmung des kurzfristigen Gleichgewichts bildet Schritt 1. Danach wird das langfristige Gleichgewicht bei verschiedenen Werten der Transportkosten bestimmt und dargestellt. Der vollständige Simulationscode zu allen in dieser Arbeit vorgestellten Modellen befindet sich im Anhang.

5.3.1 Erster Schritt: Gleichgewicht bei konstanten Transportkosten

Das Grundmodell der NEG von Krugman kann in seinem kurzfristigen Gleichgewicht durch folgendes Gleichungssystem dargestellt werden.¹²⁸

$$Y_1 = \mu\lambda w_1 + \frac{1-\mu}{2} \quad (5.1)$$

$$Y_2 = \mu(1-\lambda)w_2 + \frac{1-\mu}{2} \quad (5.2)$$

$$Q_1 = \left[\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (5.3)$$

$$Q_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (5.4)$$

$$w_1 = \left[Y_1 Q_1^{\sigma-1} + Y_2 T^{1-\sigma} Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (5.5)$$

$$w_1 = \left[Y_1 Q_1^{\sigma-1} T^{1-\sigma} + Y_2 Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (5.6)$$

$$\omega_1 = w_1 Q_1^{-\mu} \quad (5.7)$$

$$\omega_2 = w_2 Q_2^{-\mu} \quad (5.8)$$

Wenn man diese Gleichungen simulieren will, ergibt sich folgender Code:

```

1 clear;
2 my = 0.4;
3 sigma = 5;
4 T = 1.7;
5 L = -0.01;
6 w1 = 1;
7 w2 = 1;
8
9 for j = 1:101
10     L = L+0.01;
11     for i = 1:1000
12         w11 = w1;
13         w21 = w2;
14         y1 = (my*L)*w1+(1-my)/2;

```

¹²⁸Vgl. Abschnitt 2.9.

```

15  y2 = (my*(1-L))*w2+(1-my)/2;
16  Q1 = (L*(w1^(1-sigma))+(1-L)*((w2*T)^(1-sigma)))^(1/(1-sigma));
17  Q2 = (L*((w1*T)^(1-sigma))+(1-L)*(w2^(1-sigma)))^(1/(1-sigma));
18  w1 = (y1*(Q1^(sigma-1))+y2*(Q2^(sigma-1)))*(T^(1-sigma))^(1/
      sigma);
19  w2 = ((y1*(Q1^(sigma-1)))*(T^(1-sigma))+y2*(Q2^(sigma-1)))^(1/
      sigma);
20  if w1<=0 then
21    break
22  elseif w2<=0 then
23    break
24  end;
25  if abs((w1-w11)/w11)<0.0001 then
26    if abs((w2-w21)/w21)<0.0001 then
27      break
28    end;
29  end;
30  temp1 = w1*(Q1^(-my));
31  temp2 = w2*(Q2^(-my));
32  end;
33  r1(j) = L;
34  r2(j) = temp1-temp2;
35  r3(j) = 0;
36  end;
37
38  plot2d(r1, r2, rect=[0,-0.02,1,0.02]);
39  plot2d(r1, r3);
40  xlabel("L");
41  ylabel("w1_/_w2");

```

Das Programm erzeugt Abbildung (5.1), die aus diversen Veröffentlichungen bekannt ist. Die geschwungene Kurve stellt die Arbeitsnachfrage aus den Gleichungen (5.1) bis (5.8) dar. Die Gerade bei 0 ist das Arbeitsangebot, das wegen der vollkommenen Mobilität unendlich elastisch ist. Die Schnittpunkte bestimmen dann später die langfristigen Gleichgewichte.¹²⁹

Im Folgenden soll der Aufbau des Programms kurz erläutert werden: In Zeile 1 wird der Speicher mit dem Befehl `clear` gelöscht, damit keine Variablen mit Werten eines vorherigen Durchlaufs belegt sind. Danach werden die Parameter mit Werten gefüllt (Zeile 2-5). Die zu bestimmenden endogenen Variablen `w1` und `w2` werden mit 1 initialisiert. Danach folgt eine doppelte Schleife. Die äußere Schleife

¹²⁹Zur inhaltlichen Diskussion der Abbildung und des Modells vgl. Abschnitt 2.9.

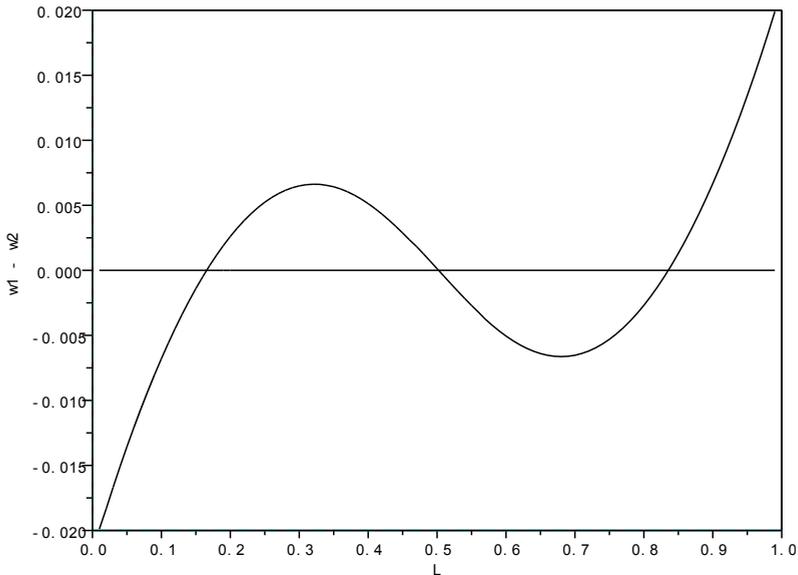


Abbildung 5.1: Kurzfristiges Gleichgewicht des Standard-Modells

von Zeile 9 bis 36 wird für jeden Wert für λ von 0 bis 1 durchlaufen. Die innere Schleife von Zeile 11 bis 32 enthält die eigentliche Simulation. Nach der Zuweisung neuer Werte für den Lohn werden zunächst die Gleichungen (5.1) bis (5.8) durchgerechnet (Zeile 14 - 19). Falls ein Lohnsatz unter 0 fällt, bricht das Programm ab (20 - 24). Danach wird geprüft, ob die neuen Werte für den Lohnsatz nicht mehr als die Toleranz von 0,0001 von den alten abweichen. Ist das der Fall, werden die Reallöhne ermittelt (30 und 31). Zeile 33 - 35 sind technischer Natur. Hier werden die Werte für λ für die Graphik festgehalten sowie eine 0 als Vergleichswert für die horizontale Linie gespeichert. Die letzten vier Zeilen erstellen die Graphik.

Bei der Simulation von anderen Modellen müssen entsprechend deren Gleichungen umgesetzt werden. Dabei ist zu beachten, dass es unter Umständen nicht möglich ist, die Gleichungen so aufzulösen, dass alle endogenen Variablen isoliert sind. In einem solchen Fall, z.B. im CM-Modell in Abschnitt 4.3, ist eine weitere geschachtelte Schleife in der Simulation notwendig. Der Anpassungsprozess, also z.B. die Wanderung von Arbeitskräften, ist nicht Teil dieser Simulation. Er kommt erst im Übergang zum langfristigen Gleichgewicht zum Tragen.

5.3.2 Zweiter Schritt: Gleichgewichte bei Änderung der Transportkosten

Die obigen Abbildungen können entweder 1 oder 3 Nullstellen haben, wobei eine immer bei 0,5 liegt. Bei $T = 1$ ist das Gleichgewicht bei 0,5 instabil, die extremen Gleichgewichte stabil. Das ist solange der Fall, bis bei steigendem T drei Nullstellen auftreten. Dann liegen stabile Gleichgewichte bei 0, 1 und 0,5 vor. Die Nullstellen dazwischen stellen instabile Gleichgewichte dar. Ab dem Punkt, an dem nur noch eine Nullstelle existiert, ist das einzige Gleichgewicht die Lösung bei 0,5. Es ist natürlich auch stabil. Möchte man die verschiedenen Gleichgewichte in Abhängigkeit von den Transportkosten darstellen, so muss die oben beschriebene Simulation ausgehend von $T = 1$ für verschiedene T -Werte durchgeführt werden.¹³⁰ Die daraus folgende "Tomahawk"-Abbildung wird in den meisten Veröffentlichungen konstruiert (vgl. Abb. 5.2).

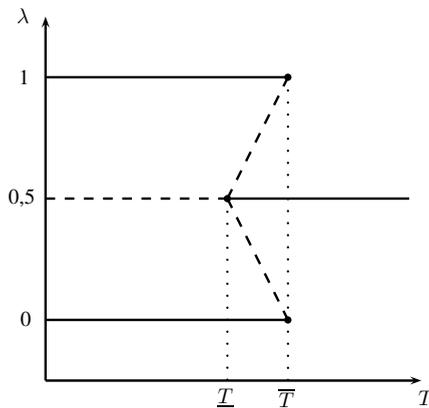


Abbildung 5.2: Gleichgewichte in Abhängigkeit vom Transportkostensatz

Wir wollen sie hier jedoch aus der Simulation berechnen, weil hier ein möglichst allgemein anwendbares Verfahren entwickelt werden soll, um ähnliche Modell auch ohne vorheriges Wissen über die Ergebnisse lösen zu können.

Nachdem das kurzfristige Gleichgewicht gefunden und in Form von zwei Kurven dargestellt wurde, kann das langfristige Gleichgewicht bestimmt werden. Dieses wird durch einen Anpassungsprozess, z.B. die Wanderung von Arbeitskräf-

¹³⁰Werte unter 1 sind nicht sinnvoll, da es sich um Eisberg-Transportkosten handelt. Daher werden tatsächliche Kosten von 0 mit dem Wert $T = 1$ dargestellt.

ten von einer Region in die andere oder durch den Wechsel des Arbeitsplatzes in einen anderen Sektor, erreicht. Grafisch ergibt sich das langfristige Gleichgewicht aus dem Schnittpunkt beider Kurven. Hier können aus methodischen Gründen nun folgende Fälle unterschieden werden:

Fall 1: Beide Marktseiten (Kurven) werden unabhängig voneinander simuliert, z.B. Arbeitsangebot und -nachfrage im LW-Modell.

Fall 2: Eine Kurve ist eine Inverse der anderen, z.B. Inlands- und Auslandsgleichgewicht im IM- oder CM-Modell.

Fall 3: Eine Kurve wird simuliert, die andere ist eine Konstante (z.B. der Nulllinie). Dieser Fall liegt z.B. beim CP-Modell vor.

Die letzten beiden Fälle sind Spezialisierungen des ersten. Dabei tritt folgendes technisches Problem auf: Die Abbildungen im ersten Schritt (bei gegebenen Parameterwerten) ergeben sich aus Simulationen bei denen ein Wert (Ordinate oder Abszisse) gegeben ist, und der andere aus dem Modell berechnet wird. Daraus entsteht eine Reihe von Punkten, die jeweils durch zwei Werte gekennzeichnet sind. Der vorgegebene Wert ist dabei eine regelmäßige Folge, über den berechneten Wert kann keine Aussage gemacht werden, er kann auch nicht definiert sein. Das Problem besteht nun darin, dass der Schnittpunkt von zwei Linien grafisch einfach bestimmt werden kann, die Berechnung aber deutlich schwieriger ist, da in vielen der betrachteten Modelle bei einer Kurve der Abszissenwert vorgegeben ist und der Wert der Ordinate berechnet wird, bei der andern Kurve aber genau andersherum. Beispielsweise ergibt sich das Arbeitsangebot im Modell vom Ludema und Wooton (1997), indem die Lohnsatzrelation vorgegeben wird und die Arbeiter entsprechend ihren Präferenzen den Standort bestimmen. Die Arbeitsnachfrage hingegen wird genau andersherum berechnet. Hier legen die Unternehmen den relativen Lohnsatz fest, den sie bei gegebener Verteilung der Arbeitskräfte bereit wären zu zahlen. Um nun Schnittpunkte zu bestimmen, benötigt man aber für beide Kurven den Funktionswert an der gleichen Stelle.

Zur Veranschaulichung sei auf Abbildung 5.3 verwiesen. Hier sind die beiden (nicht stetigen) Funktionen $f : a \rightarrow b$ und $g : b \rightarrow a$ dargestellt, wobei zur besseren Übersicht die Intervalle sehr groß gewählt sind. Möchte man nun zunächst das Intervall bestimmen, in dem sich der Schnittpunkt befindet, muss man zunächst eine Achse als Index festlegen, die Differenz für jeden Wert dieser Skala bilden und dann durch ein geeignetes Verfahren die Nullstellen bestimmen. Wählt man

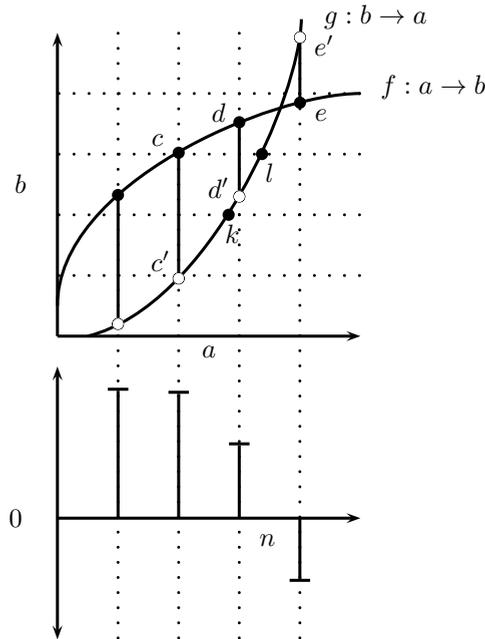


Abbildung 5.3: Nicht-Stetigkeit

z.B. a als Index, liegen die Werte für die Funktion f vor, aber nicht für die Funktion g . So liegt z.B. der Punkt d vor, der Punkt d' aber nicht. Für g liegen nur die Werte k und l vor, die nah an d' liegen. Selbst wenn das gelingt, bleibt ein weiteres Problem: Für den Schnittpunkt ist mit großer Wahrscheinlichkeit noch nicht einmal ein Punkt gegeben. Für diese Problematik gibt es grundsätzlich zwei Lösungsansätze:

Ansatz 1: Beide Kurven werden mittels eines Schätzverfahrens in eine stetige Funktion gebracht. Danach zieht man die beiden Funktionen von einander ab und berechnet die Nullstellen analytisch. Dieses Verfahren hat den Vorteil, dass es relativ schnell geht und einfach zu programmieren ist. Es kann allerdings nur angewendet werden, wenn die Kurve relativ gut geschätzt werden kann. Neben dem optischen Vergleich von geschätzter und tatsächlicher Kurve kann durch den Wert der *Sum of squared Errors (SSE)* die Güte der Schätzung beurteilt werden. An dieser Stelle muss aber beachtet werden, dass die Dimensionen in diesen Modellen relativ gering sind und daher ein SSE von z.B. 0,005 bereits ein unakzeptables Ergebnis sein kann. Im LW-Modell mit $T = 1,5$ und $s = 0,08$ beispielsweise ergibt

sich bei einem SSE von 0,005 für die Angebotskurve bereits ein Fehler von ca. 0,06 für λ , was einer Abweichung von 6% bezogen auf das zulässige Intervall für λ entspricht (vgl. Abb. 5.5).¹³¹

Ansatz 2: Für beide Kurven werden mittels Interpolation Funktionswerte für eine einheitliche Skala (z.B. $\lambda = 0; 0,001; 0,002; \dots; 1$) gebildet. Danach werden dann beide Kurven voneinander abgezogen und numerisch (z.B. durch die *Bisection Method*, *Newton* und *Quasi-Newton-Verfahren*) die Nullstellen der resultierenden Reihe gesucht. Das Verfahren ist aufwendiger und kann, je nach verwendeten Algorithmen, deutlich länger dauern. Dafür sind die Ergebnisse aber insbesondere bei komplizierten Kurvenverläufen genauer.

Fall 2 ist eine Variante von Fall 1, bei dem die zweite Kurve mit genau dem gleichen Algorithmus simuliert wird wie die erste, nur dass abhängige und unabhängige Variable vertauscht sind. Hier kann es sinnvoll sein um Zeit zu sparen, die Simulation nur einmal durchzuführen und dann numerisch die Inverse zu bilden. Generell können natürlich auch die gleichen Methoden angewandt werden wie im oben beschrieben, der Fall wird hier aber dennoch extra genannt, weil er in den Modellen der NEG, die sich mit mehreren Ländern beschäftigen, häufig anzutreffen ist. Aus diesem Grund wird ein Algorithmus für diesen Spezialfall angegeben.

Fall 3 ist eine Vereinfachung, da hier das Problem der unterschiedlichen abhängigen und unabhängigen Variablen nicht auftaucht. Nur die Problematik, dass der Schnittpunkt nicht als Datenpunkt vorliegt, bleibt bestehen. Dies ist ein Problem diskreter Werte allgemein. Es kann hier also ohne weiteres mit der numerischen Bestimmung der Nullstellen fortgefahren werden. Allerdings kann es in einigen Fällen die Performance und auch die Genauigkeit steigern, auch hier mit Interpolation zu arbeiten.

Abschließend sei noch darauf hingewiesen, dass durch eine Fallabfrage und mit Hilfe des theoretischen Wissens über das jeweils betrachtete Modell Nullstellen auszuschließen sind, die sich im negativen Bereich befinden. In diesem Fall kommt es zu Randlösungen, in denen zwar kein Gleichgewicht im Sinne eines Schnittpunkts vorliegt, die Anpassungskräfte des Systems aber dazu führen, dass der Schnittpunkt einer der beiden Kurven mit der Grenze des zulässigen Bereichs ei-

¹³¹Die Nachfragekurve hat einen SSE-Wert von 0,000002. Die Abweichung ist so gering, dass sie in der Abbildung nicht sichtbar wird.

ne stabile Situation bilden. Ist beispielsweise im CM oder IM-Modell der Lohn in Land 1 auch dann noch höher, wenn bereits alle Arbeitskräfte nach Region 1 gewandert sind, so stellt das eine stabile Situation dar, ohne ein Gleichgewicht im Sinne eines Lohnausgleichs zu sein.

Polynomschätzung

Approximierung eines Polynoms Bei diesem Verfahren wird zunächst anhand der Form der vermuteten Funktion ein Funktionstyp festgelegt. In den meisten hier betrachteten Fällen haben die Funktionen einen S-förmigen Verlauf, so dass wir ein Polynom dritten Grades annehmen. Danach werden die Parameter so gesetzt, dass die Summe der quadrierten Abweichungen zwischen den tatsächlichen Werten und den von der Funktion ermittelten minimiert werden. Dies kann auf drei verschiedenen Arten geschehen, die in Scilab ausgewählt werden können: Das Quasi-Newton Verfahren, das CG-Verfahren (*conjugate gradient*) und ND-Verfahren (*non-differentiable*). Als Standard ist das Quasi-Newton Verfahren vorgesehen. Die Wahl eines anderen Verfahrens bewirkt nur sehr geringe Änderungen der erreichten Güte oder der notwendigen Zeit. Die Qualität kann mit Hilfe der Variable SSE (*Sum of Squared Errors*) beurteilt werden, die Scilab mit ausgibt. In den folgenden Funktionen wird dieses Verfahren verwendet:

```
function [e]=errors(koeff,z)
e=z(2)-koeff(1)-koeff(2)*z(1)-koeff(3)*z(1)^2-koeff(4)*z(1)^3;
endfunction

function [result]=myRegress(m)
koeff0=[1;1;1;1];
X=m(:,1)'; Y1=m(:,2)';
measdata=[X;Y1];
[koeff,SSE]=datafit(errors,measdata,koeff0,algo='qn');
result = [koeff; SSE];
endfunction
```

Die erste Funktion `errors()` bestimmt dabei die Abweichung zwischen dem tatsächlichen Wert und dem Wert der geschätzten Funktion für jeden einzelnen Datenpunkt. In `koeff` werden dabei die Koeffizienten der Gleichung übergeben, `z` enthält die Wertepaare der Datenpunkte. Der Rückgabewert der Funktion ist die Differenz.

Die Funktion `myRegress()` wird mit der Matrix `m` aufgerufen, die ebenfalls die Wertepaare enthält und gibt die Koeffizienten des Polynoms zurück, welches eine beste Annäherung an die Punkte ist. Der letzte Wert in der Ergebnismatrix ist die Summe der quadrierten Abweichungen. Dabei werden zunächst Startwerte für die Parameter festgelegt (hier jeweils 1). Der Kern dieser Funktion ist der Befehl `datafit()`, der unter Verwendung der Fehlerfunktion die Koeffizienten ermittelt und auch die Summe der quadrierten Abweichungen mit ausgibt. Dabei wird über den (optionalen) Parameter `algo` der verwendete Algorithmus festgelegt.

Um nun simulierte und approximierte Werte zu plotten, muss eine zusätzliche Spalte der Matrix mit den Funktionswerten des Polynoms gefüllt werden. Dazu kann entweder

```
k(:,4) = koeff(1)+koeff(2)*k(:,1)+koeff(3)*k(:,1)^2+koeff(4)*k(:,1)^3;
```

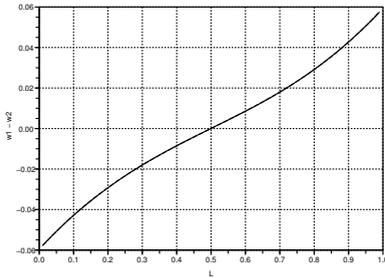
oder

```
t=poly(koeff,"x","coeff");  
k(:,4) = horner(t,k(:,1));
```

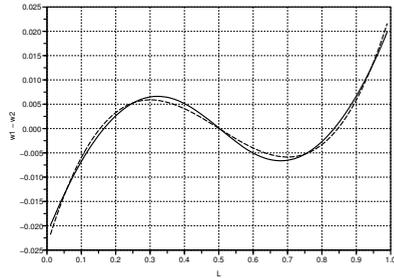
verwendet werden. Letzteres Verfahren ist etwas schneller und eleganter.

Güte der Schätzung Um zu prüfen, wie stark das approximierte Polynom von den simulierten Punkten abweicht, betrachten wir für die drei oft verwendeten Werte für T von 1,5, 1,7 und 2,1 beide Kurven und für den gesamten Bereich die Summe der quadrierten Abweichungen. Außerdem ist zu beachten, dass Abweichungen ja nur dann ins Gewicht fallen, wenn Sie bei $w_1 - w_2 = 0$ auftreten und auch dann nur für die beiden instabilen Gleichgewichte. Bei den folgenden Abbildungen kennzeichnet die durchgezogene Linie die simulierten Werte, die gepunktete gibt das Polynom wieder.

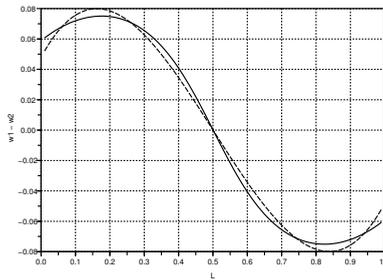
Bei diesen geringen Abweichungen scheint es durchaus sinnvoll, mit dem approximierten Polynom zu arbeiten. Nachdem für jeden Wert von T innerhalb des vorgegebenen Wertebereichs (hier (1,5 : 2,1) mit Schrittweite 0,01) die oben beschriebene Simulation durchgeführt wird, kann man die Parameter des Polynoms schätzen, das die resultierenden Punkte am besten beschreibt. Aus dem Wert für T (Spalte 1), diesen Parametern (Spalte 2-5) und der Summe der quadrierten Abweichungen als Gütemaß der Regression (Spalte 6), wird nun die Matrix `pk` gebildet,



(a) $T = 1,5$; $SSE = 0,000001$



(b) $T = 1,7$; $SSE = 0,000063$



(c) $T = 2,1$; $SSE = 0,001802$

Abbildung 5.4: Approximierte und tatsächliche Werte der Simulation

die Grundlage der weiteren Darstellungen ist:

$$pk = \begin{bmatrix} 1,50 & -0.059717 & 0.188860 & -0.207873 & 0.138231 & 0.000001 \\ 1,51 & -0.058086 & 0.189843 & -0.220598 & 0.146706 & 0.000001 \\ 1,52 & -0.056427 & 0.190946 & -0.233835 & 0.155519 & 0.000002 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2,10 & 0.047991 & 0.421414 & -1.551177 & 1.033567 & 0.001802 \end{bmatrix}$$

Plottet man nun die aus dem Polynom berechneten Linien für $T = (1,6; 1,7; 1,8; 1,9)$, so ergibt sich Abbildung 5.6.

Bei der Schätzung des Polynoms ist darauf zu achten, dass nur Werte innerhalb des zulässigen Intervalls verwendet werden. So gibt der Algorithmus für das CM-Modell z.B. auch Werte < 0 oder > 1 aus, wenn die Linie für das Ausland die

vertikalen Achsen schneidet. Das ist zunächst unproblematisch, denn durch Angabe von $\text{rect}=[]$, das den dargestellten Wertebereich einschränkt, werden diese Teile in der Graphik weggelassen. Verwendet man aber alle berechneten Werte zur Schätzung der Parameter des Polynoms, so verschlechtert sich die Qualität der Schätzung im relevanten Bereich, da nicht relevante Daten mit berücksichtigt werden.

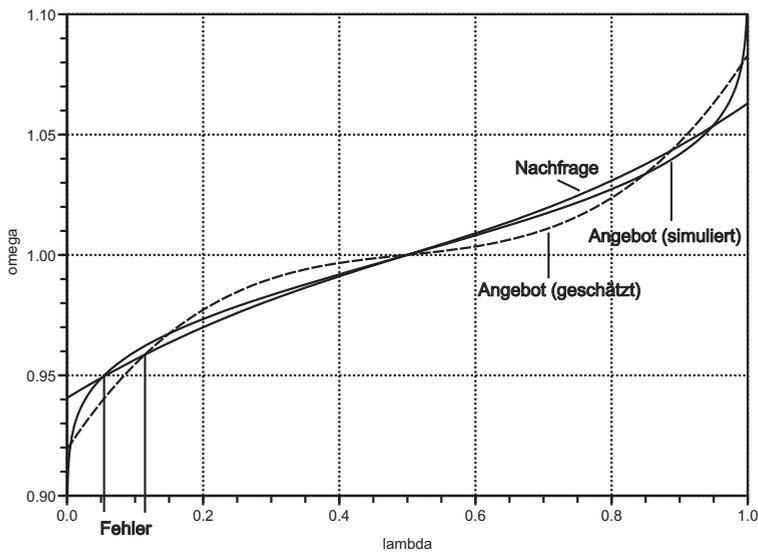


Abbildung 5.5: Fehler im LW-Modell

Auch bei korrekter Auswahl der Daten kann die Qualität der Schätzung nicht ausreichen sein. Eine Verbesserung ist zum einen möglich, indem der Grad des Polynoms erhöht wird. So ist es nach dem Weierstrass Theorem¹³² grundsätzlich möglich, jede Funktion durch ein Polynom mit beliebiger Genauigkeit zu approximieren, indem der Grad erhöht wird. Dies ist allerdings nur mäßig hilfreich, denn die Vorteile der Polynomschätzung schwinden bei zunehmendem Grad des Polynoms. Eine andere Möglichkeit, die Qualität zu erhöhen, liegt in der Verwendung der so genannten Chebychev Knoten. Dies basiert auf der Idee, die verwendeten Datenpunkte nicht gleich auf den gesamten relevanten Raum zu verteilen, son-

¹³²Dieses wurde 1885 von Karl Weierstrass formuliert.

dern an den Enden mehr Punkte zu verwenden als in der Mitte.

Wie die folgende Diskussion zeigen wird, ist es aber nicht sinnvoll, die Schätzung eines Polynoms für die gesamte Funktion zu verfeinern, da andere Verfahren zur Stückweisen Approximation zur Verfügung stehen, z.B. die Verwendung von Splines, die mit geringem Aufwand bessere Ergebnisse liefern.¹³³

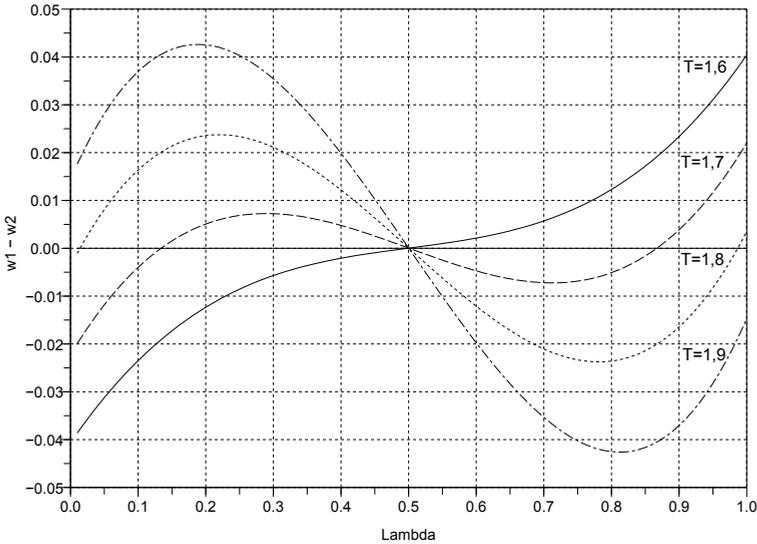


Abbildung 5.6: Plot des CP-Modells für $T = (1,6; 1,7; 1,8; 1,9)$

Berechnung der Nullstellen In einem nächsten Schritt werden für jedes T die Nullstellen des Polynoms bestimmt. An diesen Stellen gibt es keine Reallohn-Differenz und daher ein Gleichgewicht, wobei einige dieser Gleichgewichte allerdings instabil sein können. Dabei interessieren uns vor allem die inneren (instabilen) Gleichgewichte, weil wir deren Lage mit der Theorie alleine nicht exakt bestimmen können. Unter Verwendung des folgenden Programms entsteht Abbildung 5.7.¹³⁴

¹³³Zu Splines vgl. S. 147.

¹³⁴Das Programm bestimmt die Punkte des Bifurkationsdiagramms. Die Anweisungen für die grafische Ausgabe sind hier nicht dargestellt.

```
clear;
chdir("D:\Dokumente\Forschung\Dissertation\Simulation\CP");
pk = fscanfMat('Matrix-pk.txt');
t = size(pk); s = t(1);
for i = 1:s do
    koeff = [pk(i,2); pk(i,3); pk(i,4); pk(i,5)];
    p=poly(koeff,"x","coeff");
    temp = real(roots(p));
    for j = 1 : 3
        if temp(j) < 0 then temp(j) = 0; end;
        if temp(j) > 1 then temp(j) = 1; end;
    end;
    r(i,:) = [pk(i,1) temp' 0 1];
end;
fprintfMat('Matrix-roots.txt',r);
```

Dieses Programm erzeugt eine Matrix, die in der ersten Spalte den Transportkostensatz und in den darauf folgenden (drei) Spalten die Nullstellen der Funktion enthält. Die letzten beiden Spalten sind mit 0 und 1 belegt und dienen nur einer schöneren Grafik. Diese wird mit dadurch erzeugt, dass die Spalten 2-6 gegen Spalte 1 geplottet werden. Auf dieses Programm wird hier verzichtet.

Vergleicht man die konstruierte Abbildung 5.2 mit der aus der Simulation entstandenen Abbildung 5.7, so fällt folgendes auf:

In Abb. 5.2 sind instabile Gleichgewichte gestrichelt dargestellt. Dies ist in Abb. 5.7 nicht der Fall. Der Grund liegt darin, dass in der Simulation nicht zwischen stabilen und instabilen Gleichgewichten unterschieden wird. Dies wäre zwar möglich, der erhöhte Programmieraufwand steht aber einem sehr geringen Erkenntniszuwachs gegenüber. Schließlich sagt uns die Theorie zuverlässig, welche Gleichgewichte stabil sind und welche nicht. Die Linien für die extremen Gleichgewichte bei 0 und 1 sind in Abb. 5.7 über \bar{T} hinaus durchgezogen. Dies hat den technischen Grund, dass es dort durchaus Nullstellen gibt, diese liegen aber außerhalb des zulässigen Bereichs (unter 0 oder über 1). Auch hier gilt: Der Aufwand, dies in der Darstellung zu vermeiden, ist im Vergleich zum Erkenntniszuwachs gering.

Ein wesentlicher Unterschied ist aber zwischen den Abbildungen zu erkennen: Das instabile Gleichgewicht, deren Lage ja bei mittleren Transportkosten darüber entscheidet, ob das System zur Agglomeration oder zur Disagglomeration tendiert, verläuft nicht linear von 0,5 zu 0 bzw. zu 1. Damit ist der Bereich, der zur Gleichverteilung führt in der Simulation deutlich größer als man aufgrund von

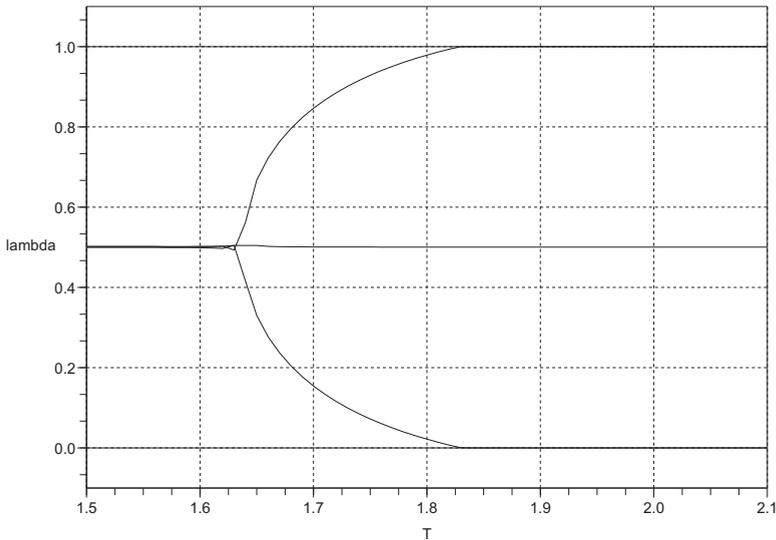


Abbildung 5.7: Gleichgewichte bei verschiedenen Transportkostensätzen

Abb. 5.2 annehmen könnte. Außerdem gleicht die Abbildung nun stärker der des Cluster-Modells (vgl. Abschnitt 4.3).

Interpolation

Bei diesem Verfahren wird nun, im Gegensatz zu dem eben vorgestellten, nicht die gesamte Funktion geschätzt, sondern nur die einzelnen Abschnitte zwischen den Datenpunkten. Dann werden die benötigten aber nicht vorhandenen Punkte durch Interpolation bestimmt und diese voneinander abgezogen. Die Nullstellen können daher nicht analytisch sondern nur numerisch bestimmt werden. Verwendet man ausreichend kleine Intervalle bei der Interpolation, können die Nullstellen aber mit ausreichender Genauigkeit gefunden werden.¹³⁵

Lineare Interpolation Ein mögliches Verfahren, um das beschriebene Problem der unterschiedlichen Skalen bei diskreten Werten zu lösen, liegt darin, die für

¹³⁵Zu Approximierungsmethoden allgemein vgl. Judd (1998), Kap. 6.

eine Skala fehlenden Werte durch einen Durchschnitt aus dem nächst geringeren und dem nächst höheren Wert zu approximieren.

	0,011	0,36	
$a \rightarrow$	0,022	0,55	
$b \rightarrow$	0,041	0,63	
	0,048	0,69	

Abbildung 5.8: Beispiel zur Approximierung mittels Durchschnitt

Liegen z.B. die Wertepaare aus Abb. 5.8 vor und möchte man diese auf eine Skala in Schritten von 0,01 bringen, muss z.B. der Wert für 0,03 geschätzt werden. Dieses Verfahren läuft die Indexwerte so lange nach oben durch, bis der letzte Wert vor dem Zielwert erreicht ist (hier a). Dann wird der erste Wert über dem Zielwert ermittelt (hier b) und der Durchschnitt der beiden Funktionswerte ermittelt.

$$F(0,03) \approx \frac{f(a) + f(b)}{2}$$

Diese Approximation wirkt auf den ersten Blick recht grob. Da aber die Punkte sehr nah beieinander liegen, ist der Fehler dennoch sehr gering. Das Verfahren kann auch verwendet werden, um eine Relation umzukehren (also von $a \rightarrow b$ zu $b \rightarrow a$) und dabei die Indexwerte auf eine Skala mit konstanten Schritten zu bringen:

	x	$f(x)$	
	0	0,011	
	0,01	0,022	$\leftarrow a$
	0,02	0,041	$\leftarrow b$
$z \rightarrow$	0,03	0,048	$f(z) = ?$
	0,04	0,051	

Abbildung 5.9: Invertieren der Wertepaare mit Approximation

Dies wird z.B. für die Simulation des Cluster-Modells von Krugman und Venables (1996) benötigt. Eine Implementierung in Scilab sieht dann wie folgt aus:

```
function result = myInvert_GGLine(m)
i = 1;
for p = 0:999
```

```

k = 1000-p;
ziel = k/1000;
while m(i,2) > ziel,
    if i==1000 then break end;
    i=i+1;
end;
t1 = i-1;
while ~(m(i,2) < ziel),
    if i==1000 then break end;
    i=i+1;
end;
t0 = i;
m(k,3) = (t0+t1)/2000;
if m(k,3) < 0.001 then m(k,3) = 0; end;
end;
result = m;
endfunction

```

Dabei wird zunächst für jeden gesuchten Wert der nächst höhere vorhandene gesucht. Danach wird der nächste vorhandene Wert bestimmt, der unter dem Zielwert liegt und schließlich wird daraus der Mittelwert gebildet. Auch hier gilt: Das Verfahren ist zwar grob, aufgrund des geringen Abstands zwischen den Werten aber dennoch recht genau. Zuletzt werden Schnittpunkte im negativen Bereich auf 0 gesetzt.

Interpolation durch kubische Splines Die Schätzung eines Polynoms für die gesamte Kurve hat den Vorteil, dass Nullstellen dann analytisch bestimmt werden können. Durch die Oszillation von Polynomen höheren Grades kann es aber zu Ungenauigkeiten kommen, wie z.B. in Abb. 5.5 zu erkennen ist. Bei linearer Approximation treten diese Fehler nicht auf, dafür ist das Verfahren aber sehr rechenintensiv und diese Intensität steigt, je genauer die Approximation wird. Ein Ausweg liegt in der Verwendung von Splines. Bei diesem Verfahren wird die Kurve nur zwischen jeweils zwei Datenpunkten durch ein Polynom geschätzt, so dass die Kurve möglichst "glatt" ist, also keine unnötigen Ausschläge hat, sich andererseits aber auch an den Knoten keine Brüche ergeben. Der Begriff *Spline* kommt ursprünglich aus dem Schiffsbau und bezeichnen das Verhalten von Latten, die sich beim Bau von Schiffen biegen, dabei aber möglichst wenig Spannung aufbauen. In der Mathematik wurden Splines erstmals von Schoenberg (1946) eingeführt. Nach der Schätzung des Splines können die Datenpunkte auf eine einheitliche

Skala interpoliert werden. Dieses Verfahren hat den Vorteil, dass es sehr genaue Ergebnisse liefert. Da Splines in der Physik weit verbreitet sind, liegen sehr effiziente Algorithmen zu ihrer Bestimmung vor, so dass die Berechnung relativ schnell ist. Der Nachteil ist, dass Splines keine handhabbare Funktionsgleichung liefern, wie die oben beschriebenen Polynome.

Die Güter einer Schätzung mit Splines ist schwer zu messen, da die gegebenen Daten definitionsgemäß auf dem Spline liegen. Aus diesem Grund ist auch ein optischer Vergleich der interpolierten Kurve mit den Datenpunkten aus der Simulation nicht sinnvoll. Aus den obigen Überlegungen kann man aber schlussfolgern, dass Splines gegenüber den anderen Verfahren die besten Ergebnisse liefern. Auch tritt das Problem nicht auf, dass Werte außerhalb des Definitionsbereichs den Spline beeinflussen, wie dies bei der Schätzung eines Polynoms der Fall ist.

Da, wie oben beschrieben, im CP-Modell diese Interpolation nicht unbedingt nötig ist, wird hier das Modell von Ludema und Wooton (1997) als Beispiel verwendet. Dieses erweitert das CP-Modell dadurch, dass das Arbeitsangebot nicht als vollkommen elastisch angesehen wird, sondern durch eine weitere Simulation als Kurve dargestellt wird.¹³⁶ Der in Abschnitt 5.3.1 dargestellte Code kann also weiterhin verwendet werden. Die beiden simulierten Kurven werden dann durch folgende Funktionen auf eine einheitliche Skala gebracht:

```
// Arbeitsangebot im LW-Modell
function zz = myLWS(s)
[y1, y2] = mySimulateLWS(s);
zz = interp(L, y1, y2, splin(y1,y2,"fast"));
endfunction

//Arbeitsnachfrage im LW-Modell
function zz = myLWD(sigma, my, T)
[k1, k2] = mySimulateLWD(sigma, my, T);
zz = interp(L, k1, k2, splin(k1,k2,"fast"));
endfunction
```

Dabei wird die eigentliche Simulation in den Funktionen `mySimulateLWS()` und `mySimulateLWD()` durchgeführt. Letztere enthält den Code von oben.¹³⁷ Die eingebaute Funktion `splin()` berechnet den Spline, wobei hier verschiedene Me-

¹³⁶Vgl. Abschnitt 3.2.

¹³⁷Der einzige Unterschied ist, dass hier der Quotient der beiden Reallöhne $\frac{\omega_1}{\omega_2}$ und nicht deren Differenz $\omega_1 - \omega_2$ verwendet wird. Das liegt daran, dass nur hierfür das Arbeitsangebot bestimmt werden kann.

thoden möglich sind. Durch `interp()` wird dann ein Vektor erzeugt, der die auf Grundlage dieses Splines interpolierten Werte auf den Vektor `L` enthält. Dadurch erhält man für beide Kurven Funktionswerte auf `L`, die man dann voneinander abziehen kann um die Schnittpunkte (Nullstellen) zu bestimmen. Zu beachten ist, dass die Funktion `splin()` eine Matrix entgegennimmt, deren erste Spalte strikt monoton steigend sein muss. Ist das nicht der Fall, muss diese vor dem Aufruf entsprechend geändert werden. Im LW- und IM-Modell ist dabei die Matrix für das Ausland umzukehren. Außerdem ist es in einigen Fällen nötig, die Anzahl der simulierten Punkte zu reduzieren, damit kein Wert doppelt auftritt, denn auch das verstößt gegen die strikte Monotonie.

Sind die simulierten Werte sehr unregelmäßig und ist daher bei nur einigen Werten die strikte Monotonie verletzt, können Simulationsschritte, bei denen dieser Fehler auftritt, auch durch eine try-catch-Konstruktion übersprungen werden. Insbesondere bei der Simulation der verschiedenen Varianten des Modells mit Zwischenprodukten in den Abschnitten 77 und 90 ist dies häufig der Fall. Bei der Konstruktion des Bifurkationsdiagramms des asymmetrischen IM-Modells wurde beispielsweise folgender Code verwendet:

```
L = linspace(0,1,1001)';
T = linspace(T0,T1,300)';
for i = 1 : length(T)
    try
        hh = myIMI(T(i));
        ff = myIMA(T(i));
        xx = hh-ff;
        r = myRoots(L,xx);
        for j = 1 : length(r)
            rts = [rts; T(i), r(j)];
        end;
    catch
        fprintf("Fehler_bei_T=%f_\n\r", T(i))
    end
end;
```

Dabei finden in den Funktionen `myIMI(T(i))` und `myIMA(T(i))` die Simulation und die Interpolation durch Splines statt, die gelegentlich zu Fehlern führen.

Methoden zur Bestimmung von Nullstellen Möchte man die Schnittpunkte zweier Funktionen bestimmen, ist es sinnvoll, die Differenz zu bilden und dann die

Nullstellen zu suchen. Das hat den Vorteil, dass die Bestimmung von Nullstellen methodisch einfacher ist. Wenn diese Funktion nun in stetiger Form vorliegt, können die Nullstellen analytisch gefunden werden.¹³⁸ Bei stetigen Werten braucht man einen Algorithmus um die Nullstellen zu finden. In der Literatur sind drei Verfahren, jeweils mit verschiedenen Varianten, bekannt: Das *Iterations-*, das *Bisection-* und das *Newton-Verfahren*. Im Folgenden sollen diese kurz vorgestellt werden und diskutiert werden, welches für die hier vorliegende Fragestellung geeignet ist.

Iteration Die Suche nach Nullstellen durch Iteration ist sehr einfach. Man vergleicht jeden Wert eines Vektors mit dem darauf folgenden. Haben beide unterschiedliche Vorzeichen, liegt dazwischen eine Nullstelle und der entsprechende Indexwert wird ausgegeben. Der Vorteil dieses Verfahrens ist dessen Einfachheit. Fehler können selbst bei stark schwankenden Vektorenwerten nicht auftreten. Außerdem muss vorher nicht bekannt sein wie viele Nullstellen im Intervall zu erwarten sind und ob der Startwert positiv oder negativ ist. Ferner werden, im Gegensatz zum Newton-Verfahren, keine numerischen Ableitungsverfahren benötigt. Der Nachteil ist allerdings die Ineffizienz des Verfahrens, da es immer $n - 1$ Schritte benötigt. Bei den hier verwendeten Dimensionen führt dies aber nicht zu langen Durchlaufzeiten. Die hier verwendete Funktion `myRoots(x,y)` ist im folgenden dargestellt.

```
function [result]=myRoots(x,y)
result = [];
for i = 1 : length(y)-1
    if sign(y(i)) ~= sign(y(i+1)) then
        result = [result ((x(i)+x(i+1))/2)];
    end;
end;
endfunction
```

Bisection-Methode Die Bisection-Methode basiert auf dem Zwischenwert-Theorem. Dies besagt folgendes: Wenn eine Funktion f stetig im Intervall $[a, b]$ ist und der Wert A liegt zwischen $f(a)$ und $f(b)$, dann gibt es mindestens ein x zwischen a und b , für das $f(x) = A$ gilt.¹³⁹ Bezogen auf das Problem der Nullstellen kann

¹³⁸Vgl. Abschnitt 5.3.2.

¹³⁹Vgl. Sydsaeter et al. (1999), S. 20.

es wie folgt umformuliert werden: Wenn eine Funktion f stetig ist und die Funktionswerte $f(a)$ und $f(b)$ unterschiedliche Vorzeichen haben, dann muss die Funktion in dem Intervall $[a, b]$ mindestens eine Nullstelle haben. Dazu wird zunächst das Intervall festgelegt und geprüft, ob die Funktionswerte an den beiden Enden unterschiedliche Vorzeichen haben. Ist das der Fall, teilt man das Intervall in zwei Hälften und prüft eines dieser Teile ebenfalls. Trifft die Bedingung zu, hat also auch dieses Teil-Intervall an den Enden jeweils einen positiven und einen negativen Funktionswert, muss die Nullstelle in diesem Teil liegen, sonst im anderen Teil. Der relevante Teil wird dann wieder geteilt und die Teile wiederum geprüft. Dies geschieht solange, bis ein hinreichend kleines Intervall bestimmt wurde, in dem sich die Nullstelle befinden muss. Ein Scilab-Programm zu diesem Algorithmus könnte so aussehen¹⁴⁰:

```

a = sign(f(a));
x = (a+b)/2;
d = (b-a)/2;
while d > tol
    d = d/2;
    if s == sign(f(x))
        x = x+d;
    else
        x = x-d;
    end;
end;

```

Die Bisection-Methode ist schneller als die Iterative Methode, aber langsamer als das Newton-Verfahren oder davon abgeleitete Algorithmen. Gegenüber letzteren hat es den Vorteil, dass man nicht die Jacobi-Matrix berechnen muss, was ohne Funktionsgleichung sehr aufwendig und fehleranfällig ist. Der Nachteil besteht jedoch darin, dass die Methode nur funktioniert, wenn sicher ist, dass die Funktion auf dem Intervall genau eine oder gar keine Nullstelle hat. Dies ist leicht einzusehen, denn bei zwei Nullstellen wären die Vorzeichen an beiden Enden gleich und der Algorithmus würde ohne Ergebnis abbrechen. In solchen Fällen ist das Iterative Verfahren im Vorteil, das auf jeden Fall alle Nullstellen findet, dafür aber auch $n - 1$ Durchläufe benötigt. Das Bisection-Verfahren liefert nach maximal $\frac{\log\left(\frac{b-a}{\text{tol}}\right)}{\log(2)}$ Durchläufen ein Ergebnis.

¹⁴⁰Vgl. Fackler und Miranda (2002), S. 30.

Newton-, Quasi-Newton- und Broyden-Methode Das Newton-Verfahren basiert auf folgendem Prinzip:¹⁴¹ An einer nicht-linearen Funktion, von der eine Nullstelle bestimmt werden soll, wird an einem festgelegten Startpunkt die erste Ableitung (Tangente) als Annäherung gebildet und deren Nullstelle bestimmt. Nun betrachtet man den Funktionswert der Funktion an dieser Stelle. Weicht er stark von 0 ab, wird an dieser Stelle wieder die Tangente berechnet und deren Nullstelle gesucht. Dies geschieht solange, bis eine vorher definierte Toleranz unterschritten wird. Formal kann man für jeden beliebigen Schritt dieser Iteration die Vorschrift

$$x_{n+1} = N_f(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$

angegeben werden, die als Newton-Iteration bekannt ist, wobei die Funktion N_f Newton-Operator heißt. Der Vorteil dieser Methode ist, dass sie unter bestimmten Bedingungen recht schnell konvergiert. Dem stehen verschiedene Nachteile gegenüber: Zum einen benötigt man eine funktionale Form um Ableitungen bilden zu können. Es gibt zwar Möglichkeiten numerisch Ableitungen zu bilden, diese sind aber so kompliziert, dass das Verfahren stark an Effizienz einbüßt. Außerdem findet das Newton-Verfahren ausgehend von einem Startpunkt immer nur eine Nullstelle. Man muss also vorher die Anzahl und die ungefähre Lage der Nullstellen kennen, wobei letzteres durch ein vorgeschaltetes Bisection-Verfahren möglich ist.

Das Problem der numerischen Ableitung und der damit verbundenen Ineffizienz kann durch Benutzung des Quasi-Newton-Verfahrens abgemildert werden. Hier wird die Ableitung nur an einem Punkt gebildet, dafür konvergiert das Verfahren weniger schnell. Eine ähnliche Vereinfachung liegt im Broydon-Verfahren vor, das die Sekante statt der Tangentialen nutzt.¹⁴²

Auch wenn das Newton-Verfahren oder Varianten davon häufig eingesetzt werden, so ist es in unserem Fall meist nicht sinnvoll. Einerseits ist der Aufwand vor der Anwendung (zusätzliche Bisection-Methode, Bestimmung der Startwerte, numerische Ableitungen) relativ hoch, andererseits ist es aufgrund der geringen Abstände der Wertepaare und der damit bereits vorhandenen Genauigkeit nur notwendig, die beiden Wertepaare zu finden, zwischen denen die Nullstelle liegt. Die Stärken des Newton-Verfahrens kommen bei dieser Fragestellung nicht zum

¹⁴¹Vgl. Judd (1998), S. 150ff.

¹⁴²Vgl. Fackler und Miranda (2002), S. 36ff.

Tragen.

Bifurkationsdiagramm Unter Verwendung von Splines und einfachen Iterationen zur Bestimmung der Nullstellen, kann nun das Bifurkationsdiagramm gezeichnet werden. Durch konsequente Modularisierung ist das Programm relativ einfach geworden:

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\LW");
getf("lib_lw.sci");
getf("../lib_ge.sci");
rts = [];
s = 0.005;
L = linspace(0,1,1001)';
T = linspace(1.1,1.9,300)';
for i = 1 : length(T)
    dd = myLWD(5, 0.4, T(i));
    ss = myLWS(s);
    xx = dd-ss;
    r = myRoots(L,xx);
    for j = 1 : length(r)
        rts = [rts; T(i), r(j)];
    end;
end;
fprintfMat('Matrix-roots.txt',rts);
plot2d(rts(:,1), rts(:,2), style=0);

```

Außerdem lässt es sich ohne größere Modifikationen auf andere Modelle anwenden. Es müssen die aufgerufenen Methoden angepasst werden. Gegebenenfalls müssen andere Parameter verwendet werden. Das oben dargestellte Programm erzeugt Abbildung 5.10.

Besonderheiten bei der Simulation für asymmetrische Länder In Modellen für internationale Agglomerationen wird das kurzfristige Gleichgewicht in einen Land ermittelt, indem die Verteilung der Arbeitskräfte im Inland gesucht wird, bei der das jeweilige Gleichgewichtskriterium bei gegebener Verteilung im Ausland erfüllt ist. An den Achsen steht daher λ_h und λ_f . Das Bifurkationsdiagramm wird nun dadurch ermittelt, dass der Schnittpunkt von Inlands- und Auslands-kurve bei jedem Transportkostensatz bestimmt wird. Dieser Schnittpunkt wird

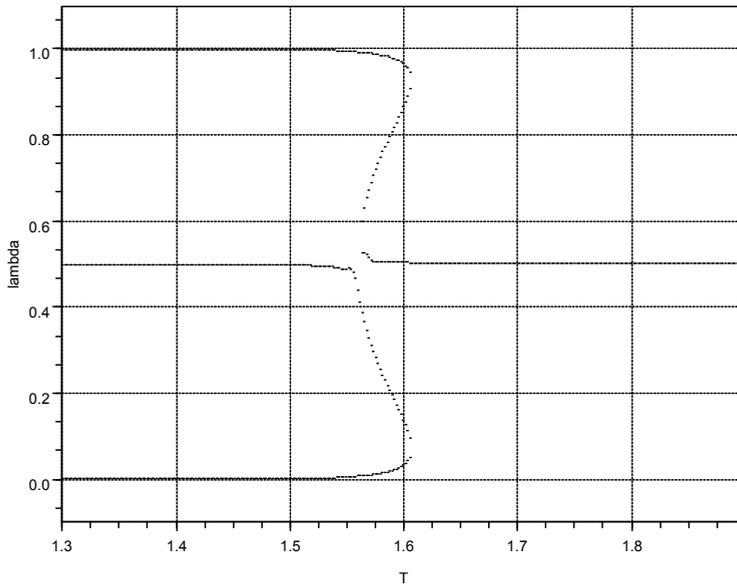


Abbildung 5.10: Bifurkationsdiagramm des LW-Modells bei Verwendung von Splines

dann zusammen mit dem verwendeten Wert von T in ein weiteres Diagramm eingetragen. Verbindet man die Punkte und bestimmt die Stabilität der Randlösungen, ergibt sich ein Bifurkationsdiagramm, wie beispielsweise in Abbildung 5.10.

Sind beide Länder symmetrisch, reicht es aus, entweder λ_h oder λ_f in Abhängigkeit von T darzustellen. Im Fall asymmetrischer Länder ist es jedoch notwendig, für jedes Land getrennt eine Graphik zu erstellen. Das Prinzip ist in Abbildung 5.11 gezeigt, das den Fall geringer Transportkosten ($T = 1,5$) im IM-Modell aus Abschnitt 4.1 darstellt. Um die Abbildung einfach zu halten sind hier tatsächlich symmetrische Länder verwendet worden. Das Prinzip ist dennoch erkennbar. Da der Parameter μ hier den Wert 0,4 annimmt, sind die (stabilen) Randlösungen mit Werten von $\lambda_f = 0$ und $\lambda_h = 2\mu$ oder umgekehrt verbunden. Auf diese Weise lassen sich für den Fall symmetrischer und asymmetrischer Länder je ein Bifurkationsdiagramm für das Inland und eines für das Ausland herstellen.

Im Gegensatz zu symmetrischen Modellvarianten kann es im Fall von Asymme-

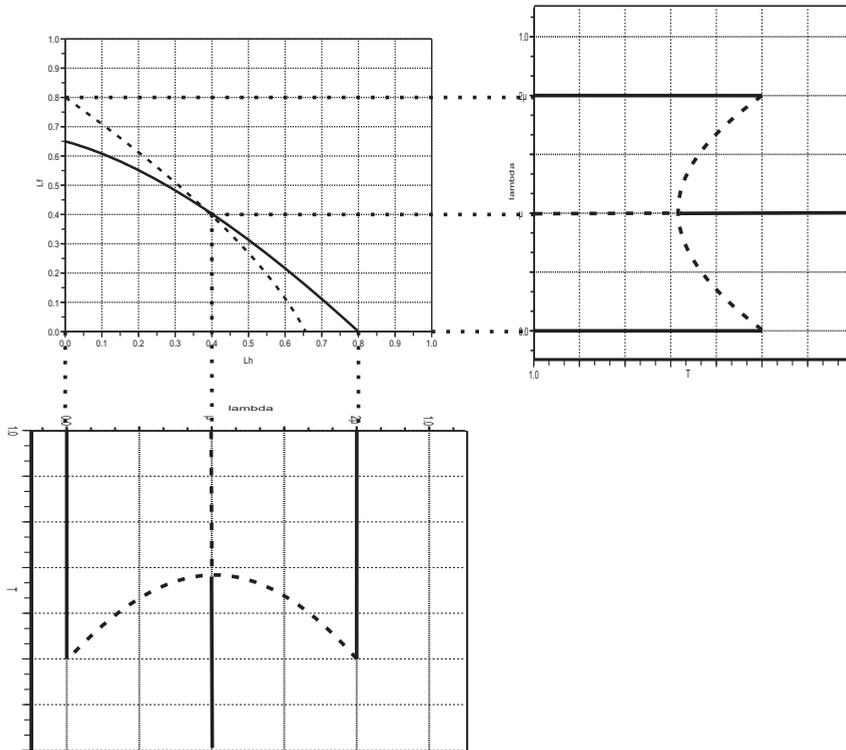


Abbildung 5.11: Konstruktion des Bifurkationsdiagramms für zwei Länder

trien im Clustermodell, dies wird Kapitel 4 gezeigt, zu unvollständiger Spezialisierung kommen. Darunter wird verstanden, dass sich zwar jeweils ein Land auf einen Sektor spezialisiert, in einem oder in beiden Ländern wird aber dennoch eine geringe Menge des anderen Sektors produziert. Dies wird im Bifurkationsdiagramm (so z.B. in Abbildung 4.20b) dadurch deutlich, dass sich das Agglomerationsgleichgewicht kurz vor dem Sustain-Point leicht von der Achse entfernt. Die bisher verwendeten Algorithmen haben diese Gleichgewichte aber gar nicht bestimmt. Das ist in den meisten Fällen auch nicht notwendig, denn wenn aus den kurzfristigen Gleichgewichten klar ist, dass die extremen sich nicht von null entfernen, ergeben sich diese aus dem geplotteten instabilen Gleichgewicht.

Ist aus den kurzfristigen Gleichgewichten zu erkennen, dass unvollständige Kon-

zentration vorliegt, so ist ein weiterer Schritt in der Simulation nötig, um dies auch im Bifurkationsdiagramm wiederzugeben. Dies wird hier am Beispiel des Gleichgewichts bei $L_h = 1$ gezeigt. In gleicher Weise lassen sich die Gleichgewichte in der Nähe von $L_h = 0$, $L_f = 1$ und $L_f = 0$ bestimmen.

Wie das Gleichgewicht bei $L_h = 1$ aussieht, hängt von der Lage der Inlands- und Auslandslinie in der oberen linken Ecke des Diagramms bei konstantem T ab. Die relevanten Möglichkeiten sind in Abbildung 5.12 dargestellt.

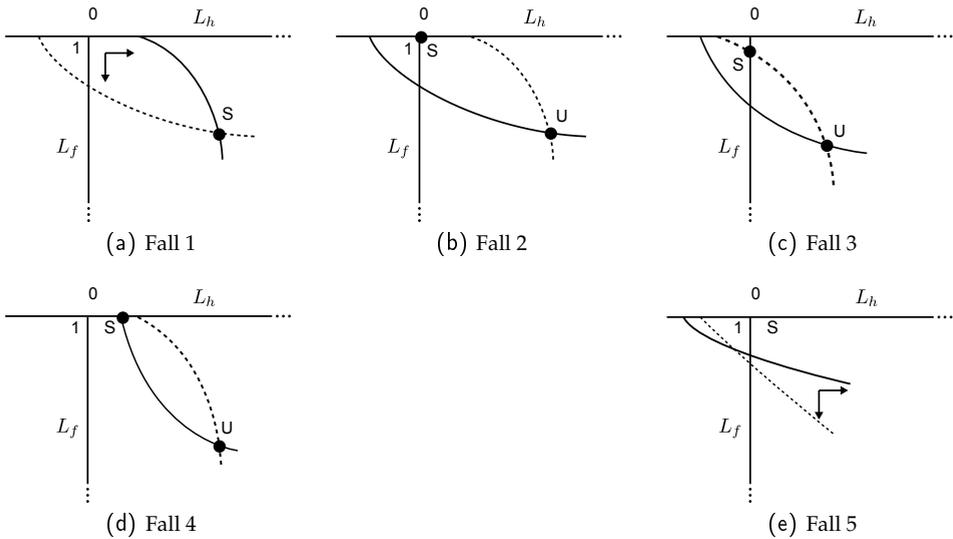


Abbildung 5.12: Mögliche Szenarien bei $L_f = 1$

Zunächst ist es sinnvoll, die Werte beider Linien bei $L_f = 1$ zu ermitteln, also am oberen Rand des Diagramms. Aufgrund der bereits beschriebenen Probleme der Nicht-Stetigkeit geschieht dies wieder mittels eines kubischen Splines. Da es sich hier um Gleichgewichtswerte handelt, bei denen die Lohnsatzdifferenz im jeweiligen Land null ist, werden sie hier mit L_h^* und L_f^* bezeichnet. Gilt $L_h^*(1) > L_f^*(1)$, so gibt es unabhängig von der konkreten Lage der Kurven kein Gleichgewicht (Fall 1) nahe des Extremwerts $L_h = 1$. Ebenso in Fall 5, der aber sehr unwahrscheinlich ist. In Fall 2 ist $L_h^*(1) < L_f^*(1)$ und außerdem $L_h^*(1) < 0$ und $L_f^*(1) > 0$. Hier gibt es ein stabiles Gleichgewicht bei $L_h = 1$, bei dem in beiden Ländern ein Lohnsatzgefälle herrscht.¹⁴³ Das gleiche Resultat ergibt sich auch im Fall 3, auch

¹⁴³Vgl. Abschnitt 4.8 für eine inhaltliche Diskussion.

wenn hier $L_f^*(1)$ positiv ist. Fall 4 ist die einzige Situation, bei der es (bezogen auf das Inland) zu unvollständiger Spezialisierung kommt. Hier gilt $L_h^*(1) < L_f^*(1)$ und $L_h^*(1) > 0$ und das stabile Gleichgewicht liegt dann bei $L_h = L_h^*(1)$.

In der Praxis ist es am wichtigsten, die Fälle 1 und 2 zu unterscheiden, da in allen anderen Konstellationen Fehler in der Fallunterscheidung sofort im Bifurkationsdiagramm durch negative Werte auffallen würden. Der folgende Code ermittelt die für das Bifurkationsdiagramm relevanten Punkte einschließlich Randlösungen (Fall 2 und 4).

```

...
for i = 1 : length(T)
    [k1, k2] = mySimulateCMI(T(i), T2, e, 0);
    m = size(k1);
    s = m(1);
    for j = 1 : s
        v1(j) = k1(s+1-j);
        v2(j) = k2(s+1-j);
    end;
    dd = interp(L, v2, v1, splin(v2,v1,"fast"));
    h0 = interp(0, k1, k2, splin(k1,k2,"fast"));
    h1 = interp(1, k1, k2, splin(k1,k2,"fast"));
    [k1, k2] = mySimulateCMI(T(i), T2, 0, e);
    ss = interp(L, k1, k2, splin(k1,k2,"fast"));
        for j = 1 : s
            v1(j) = k1(s+1-j);
            v2(j) = k2(s+1-j);
        end;
    f0 = interp(0, v2, v1, splin(v2,v1,"fast"));
    f1 = interp(1, v2, v1, splin(v2,v1,"fast"));
    x0 = 1;
    x1 = 0;
    if (h0 > f0) then x0 = min(h0, 1); end;
    if (h1 < f1) then x1 = max(h1, 0); end;
    rts = [rts; T(i), x0];
    rts = [rts; T(i), x1];
    xx = dd-ss;
    r = myRootsCM(L,dd, ss);
    for j = 1 : length(r) rts = [rts; T(i), r(j)]; end;
end;

```

Es ist allerdings sinnvoll, diese Schritte nur durchzuführen, wenn mit unvollständiger Spezialisierung zu rechnen ist, denn die Spline-Funktion führt, wie oben be-

schrieben, zu Fehlern bei nicht monoton steigenden Vektoren. Diese können aber bei den Werten für L_h^* und L_f^* leicht vorkommen, insbesondere wenn die Simulation sehr genau, also für viele Werte von T , durchgeführt wird.

5.4 Zusammenfassung

In diesem Abschnitt wurden die notwendigen Methoden vorgestellt, um Modelle der Neuen Ökonomische Geographie durch Computersimulationen zu lösen. Nach einer kurzen Einführung in die verwendete Software und deren Syntax erfolgte eine Diskussion der verschiedenen Methoden, mit denen man von kurzfristigen zum langfristigen Gleichgewicht gelangen kann. Bei der Darstellung wurde darauf geachtet, die Methoden möglichst allgemein anwendbar sind. Hier ist von Fall zu Fall zu entscheiden, welches Verfahren am sinnvollsten ist. Es kann aber festgehalten werden, dass die Verwendung von kubischen Splines zur Interpolation der kurzfristigen Kurven und das Bestimmen der Nullstellen mittels Iteration die höchste Genauigkeit bei hoher Robustheit bringen. Daher ist dies in den meisten Fällen das beste Verfahren.

Obwohl in diesem Abschnitt beispielhaft nur ein Ausschnitt der verwendeten Modelle gezeigt wurde, kann das dargestellte Verfahren auf alle anderen Modell angewandt werden. Es konnte also ein allgemeiner Ansatz zur Lösung von NEG-Modellen durch Computersimulation gefunden werden, der in weiten Teilen dieser Arbeit Anwendung findet und auch darüber hinaus verwendet werden kann.

6 Wirtschaftspolitische Implikationen

In dieser Arbeit wurden Aussagen darüber abgeleitet, wo sich Agglomerationen bilden, wenn Regionen oder Sektoren nicht identisch sind. Diese lassen sich in der wirtschaftspolitischen Anwendung sowohl für Erklärungen als auch für Prognosen zur Agglomerationsbildung nutzen. In der Literatur finden sich bereits diverse empirische Studien zur Agglomeration und zur Spezialisierung in Europa, die allerdings auf symmetrischen Modellen beruhen und Asymmetrien (wenn überhaupt) nur argumentativ in der Diskussion der Ergebnisse aufnehmen. Solche Studien, die meist auf den in Kapitel 4 genannten Modellen aufbauen, sind z.B. Klüver (2000), Midelfart-Knarvik et al. (2000) und Pflüger und Südekum (2005). Durch Anwendung der Ergebnisse aus Kapitel 4 auf die Industriestruktur Europas könnten die argumentativ dargelegten Resultate auch modelltheoretisch unterstützt werden.

In diesem Kapitel soll aber die wirtschaftspolitische Anwendung der theoretischen Ergebnisse dieser Arbeit beispielhaft an der Entwicklung Deutschlands nach der Wiedervereinigung dargestellt werden. Diese Anwendung wurde gewählt, weil es möglich ist, durch eine Veränderung der Argumentation von einer Variation der Transportkosten hin zu einer Variation der Faktormobilität eine theoretische Weiterentwicklung vorzustellen.

Zunächst erfolgt in Abschnitt 6.1 eine kurze Diskussion darüber, warum Regionalpolitik im Fall räumlicher Konzentration wirtschaftlicher Aktivität überhaupt notwendig sein könnte. Danach wird die räumliche Entwicklung in Deutschland von 1991 bis 2008 dargelegt und anhand statistischer Daten der Zusammenhang zur Theorie hergestellt, bevor im Abschnitt 6.4 das asymmetrische Modell mit räumlichen Präferenzen aus Abschnitt 3.4 leicht modifiziert und zur Erklärung herangezogen wird. Abschließend werden mögliche wirtschaftspolitische Maßnahmen diskutiert, einer nicht gewollten Entwicklung entgegenzuwirken.

6.1 Neue Ökonomische Geographie und Regionalpolitik

Die Aussagen der Neuen Ökonomischen Geographie sind positiver, erklärender Natur. Möchte man daraus wirtschaftspolitische Schlussfolgerungen ziehen bedarf es einer normativen Zielsetzung. Wie diese aussehen soll, hängt von dem betrachteten Fall ab. Es ist aber sinnvoll, Vor- und Nachteile von Agglomerationsbildung zu definieren um einzuschätzen, ob eine Entwicklung unerwünschte Folgen hat und regionalpolitische Eingriffe erforderlich macht.¹⁴⁴

In der Regel ist Regionalpolitik daran interessiert, Agglomerationsbildung und die damit verbundene Entleerung der Peripherie zu verhindern. Dafür kann man vor allem ökonomische und soziale Gründe nennen, wobei die sozialen überwiegen.¹⁴⁵

Aus ökonomischer Sicht kommt es vor allem zu einer Unterauslastung der immobilien Faktoren in der Peripherie und zu einer Überauslastung im Kern. Ein Totalmodell ohne vollkommene Konkurrenz, wie es der NEG zugrunde liegt, bedeutet eine Abkehr von der Aussage der klassischen Theorie, wonach die Märkte grundsätzlich zu einer optimalen Lösung kommen. Daher kann es hier dauerhafte nicht effiziente Allokationen geben.¹⁴⁶ Dies macht Regionalpolitik auch unabhängig von Gerechtigkeitsgedanken notwendig.

Man kann aber auch Argumente für eine genau entgegengesetzte Wirtschaftspolitik finden, welche die Bildung von Agglomerationen fördert. So beurteilen Pflüger und Südekum (2005) die Regionalpolitik der EU als effizienzmindernd, weil diese in starkem Maße versucht, Unterschiede in den Regionen auszugleichen.

Das Hauptargument für eine Begrenzung der Agglomerationstendenzen ist aber, dass viele Staaten, und auch die EU, ihren Bürgern annähernd gleiche Lebensverhältnisse in allen Regionen bieten wollen. Damit ist in der Regel neben Erwerbsmöglichkeiten auch eine Grundversorgung mit Infrastruktur, Bildungs- und Kultureinrichtungen und Wohnraum gemeint. Da viele dieser öffentlich bereitgestellt

¹⁴⁴Vgl. Pflüger und Südekum (2005).

¹⁴⁵Vgl. Maier und Tödting (2002), S. 178ff, Puga (2002). Maier und Tödting (2002) führen darüber hinaus noch ökologische Gründe an. Dies kann aber angezweifelt werden, da eine Agglomeration nicht zwangsläufig umweltschädlich ist und auch eine Gleichverteilung durch die Folgen der Zersiedelung der Landschaft Probleme mit sich bringen, wie die Autoren selbst anmerken.

¹⁴⁶Vgl. Ottaviano und Thiesse (2001), Maier und Tödting (2002), S. 118ff.

ten Güter (z.B. Schulen) eine Mindestauslastung benötigen, können daraus auch wieder ökonomische Argumente abgeleitet werden. Letztlich sind auch der soziale Frieden und der Zusammenhalt z.B. der Europäischen Union Ziele, für deren Erreichen zu starke Unterschiede in den Regionen zu vermeiden sind.

Wie man die Argumente für und gegen die Bildung von Agglomerationen nun einschätzt, hängt von der konkreten, zu beurteilenden Situation und auch von politischen Faktoren ab. Im unten diskutierten Fall von Deutschland nach der Wiedervereinigung kann man jedoch davon ausgehen, dass eine gewisse Gleichverteilung der industriellen Produktion erwünscht ist und daher einer zu einseitigen Entwicklung der Regionen gegengesteuert werden soll.

6.2 Regionale Entwicklung in Deutschland nach 1989

Als 1989 die Grenze zwischen der DDR und der Bundesrepublik Deutschland durchlässig wurde und die Vereinigung beider Staaten folgte, erwarteten viele eine Angleichung der Lebensverhältnisse und einen wirtschaftlichen Aufschwung im Osten. Zwanzig Jahre später sieht die Realität anders aus. Trotz erheblicher Subventionen und staatlicher Ausgaben liegt das wirtschaftliche Niveau der neuen Bundesländer erheblich unter dem der alten. Die Arbeitslosigkeit ist im Durchschnitt höher, die Löhne niedriger. Hinzu kommt eine z.T. umfangreiche Wanderung von Ost nach West.¹⁴⁷ Diese Wanderung hat zwei Konsequenzen: Zum einen bricht dadurch ein Teil der lokalen Nachfrage in Ostdeutschland weg, zum anderen fehlen dort qualifizierte Arbeitskräfte.

Bei der Analyse der Agglomerationsprozesse in Folge der Wiedervereinigung bietet sich die Neue Ökonomische Geographie an, weil sie die oben genannten Wanderungsbewegungen als einen wesentlichen Faktor zur Erklärung heranzieht. Daher verwenden einige Autoren¹⁴⁸ das Grundmodell von Krugman um die regionale Entwicklung nach 1989 zu erklären. Mit diesem lässt sich zwar schlüssig begründen, warum es in Folge von Handelserleichterungen zur Bildung von Agglomerationen und zur Herausbildung von Gebieten mit geringer ökonomischer Aktivität kommt. Dieser Theorie folgend müsste sich der Osten aber vollständig entleeren. Nur dann kann es dauerhaft Lohnsatzdifferenzen geben. Außerdem

¹⁴⁷Vgl. Abschnitt 6.3.

¹⁴⁸Z.B. Schempp (2000) und Schöler (2007).

wird nicht erklärt, warum sich die Agglomeration im Westen und nicht im Osten bildet. Natürlich kann man für beides leicht Erklärungen finden, diese sind aber nicht Bestandteil des Modells. Diese Lücke soll hier geschlossen werden. Dazu wird das modifizierte CP-Modell aus Abschnitt 3.4 herangezogen. Dieses enthält lokale Präferenzen nach Ludema und Wooton (1997) und berücksichtigt Größenunterschiede zwischen den beiden Regionen. Das ist zweifellos realistisch für Ost- und Westdeutschland.

Schempp (2000) verwendet das Modell von Krugman und reduziert es auf zwei zentrale Thesen um damit die oben nachgezeichnete Entwicklung zu erklären. Er bleibt dabei nahe an der Realität, jedoch ohne auf die Empirik oder die Theorie genauer einzugehen. Demgegenüber stellt Schöler (2007) das verwendete Modell detailliert dar und nennt Nominallohnverhältnisse zwischen Ost und West. Mögliche Abweichungen von den Modellprognosen, wie z.B. einzelne erfolgreiche Industrieansiedlungen im Osten, werden diskutiert.

Neben diesen beiden theorieorientierten Beiträgen gibt es einige empirische Untersuchungen, von denen die meisten sich jedoch nicht auf eine spezielle Modellwelt beziehen. Haas und Südekum (2005) und Südekum (2006) untersuchen für Deutschland die aus der NEG stammenden Thesen der Spezialisierung und Konzentration einzelner Branchen und Regionen. Diese Thesen stammen aber aus dem IM- und dem CM-Modell¹⁴⁹. Sie finden aber für Deutschland keine empirische Evidenz für diese Aussagen. Auf einem theoretischen Niveau untersucht Ross (2001) den Einfluss von verschiedenen Arten von Transfers auf die Bildung von Agglomerationen und leitet daraus sowohl Erklärungen für die gegenwärtige Lage der neuen Länder als auch wirtschaftspolitische Empfehlungen ab. Reichelt (2008) erklärt ebenfalls die wirtschaftliche Entwicklung der neuen Bundesländer auf Grundlage der NEG und bezieht dabei auch Modelle mit ein, die wirtschaftspolitische Eingriffe beinhalten. Seine Überlegungen bauen aber ausschließlich auf bekannten empirischen und theoretischen Arbeiten auf.

Eine empirische Studie, in der analog zu dem Verfahren von Hanson (2005)¹⁵⁰ die regionale Lohnstruktur in Deutschland untersucht wird, findet sich in Brakman et al. (2000b) und Brakman et al. (2000a)¹⁵¹. Brakman und Garretsen (1993) haben bereits kurz nach der Wiedervereinigung die Möglichkeit einer divergenten

¹⁴⁹Vgl. Abschnitte 4.1 und 4.3.

¹⁵⁰Der Beitrag von Hanson ist bereits 1998 als NBER Working Paper erschienen.

¹⁵¹Eine Zusammenfassung davon ist auch in Brakman et al. (2001) Kap. 5.6 zu finden.

Wirtschaftsentwicklung hervorgehoben und dies mit dem Core-Periphery-Modell und dem Verhältnis zwischen Lohn und Arbeitsproduktivität begründet. Danach ist eine Konvergenz nur möglich, wenn die Transportkosten sehr hoch bleiben oder wenn die Arbeitsproduktivität in den neuen Ländern deutlich über der in den alten Ländern liegt. Diese Aussagen sind mit den Ergebnissen dieses Kapitels vereinbar.

6.3 Das empirische Bild

Um die genannten Sachverhalte zu belegen, werden hier einige Statistiken dargestellt. Diese basieren auf Daten der Regionaldatenbank der Statistischen Ämter des Bundes und der Länder.¹⁵² Verlässliche Daten liegen hier leider erst ab 1991 flächendeckend vor, da ein einheitliches Erfassungssystem erst etabliert werden musste. Für viele Daten wäre es aber auch gar nicht sinnvoll, frühere Werte zu verwenden. Beispielsweise kann vermutet werden, dass Wanderungsbewegungen direkt nach der Grenzöffnung (1989 und 1990) nicht nur durch ökonomische, sondern auch durch persönliche oder ideelle Motive verursacht wurden. Ferner hat es einige Zeit gedauert, bis sich Löhne und Gehälter etabliert haben, die sich mit denen der alten Bundesländer vergleichen lassen.

Betrachtet man zunächst die Bevölkerungsentwicklung von 1991 bis 2008, so fällt zunächst auf, dass alle östlichen Bundesländer an Bevölkerung verloren und die meisten westlichen dazu gewonnen haben.¹⁵³ Dabei, wie bei allen weiteren Untersuchungen, muss Berlin gesondert betrachtet werden.¹⁵⁴ Die einzigen westlichen Länder mit rückläufiger Bevölkerung sind Bremen und das Saarland. Die höchsten Zuwächse hatten danach Baden-Württemberg, Bayern und Schleswig-Holstein mit jeweils über 7%. Der stärkste Rückgang ist in Sachsen-Anhalt mit 15,64% zu verzeichnen, gefolgt von Mecklenburg-Vorpommern, Thüringen und Sachsen, die jeweils Rückgänge von über 10% aufweisen. Die relativ konstante Bevölkerung in Brandenburg ist vermutlich auf einen Zuwachs in dem an Berlin

¹⁵²<https://www.regionalstatistik.de/>.

¹⁵³Die Daten finden sich in Tabelle A.4 im Anhang.

¹⁵⁴Die Daten für Berlin sind aus verschiedenen Gründen problematisch. Zu einen wird seit 1992 nicht mehr zwischen Ost- und Westteil unterschieden und auch die Zuordnung des Westteils zu den alten Bundesländern ist problematisch. Zum anderen kann das Umland aufgrund der Anordnung der Landkreise in Bandenburg nicht statistisch erfasst werden, so dass Suburbanisierungsprozesse nicht identifiziert werden können.

angrenzenden Gebiet zurückzuführen. Empirisch ist das jedoch schwer zu belegen, da die betreffenden Landkreise jeweils ein Stück dieses Gebietes einschließen und Daten unterhalb der Kreisebene nicht zugänglich sind. Die Vermutung liegt aber auch aufgrund der relativ konstanten Bevölkerung Berlins nahe.

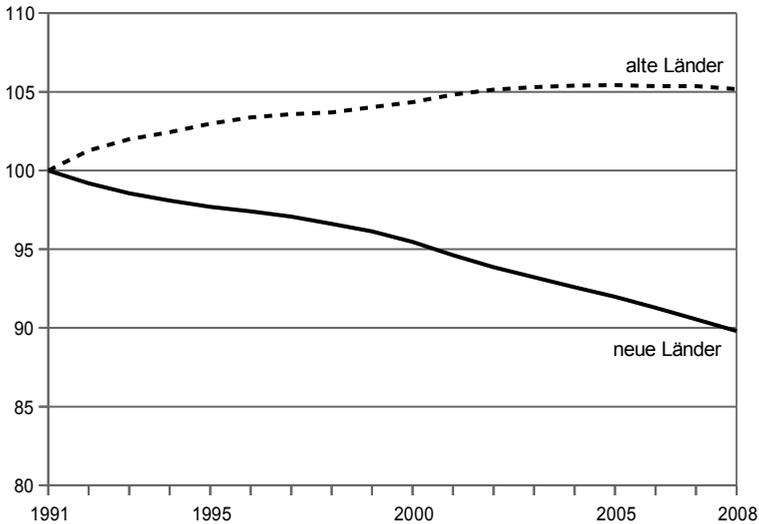


Abbildung 6.1: Bevölkerungsentwicklung alte und neue Länder (aggregiert, ohne Berlin)

Aggregiert man diese Daten über die alten und neuen Länder (ohne Berlin), wird dieser Trend noch deutlicher sichtbar. Dies ist in Abbildung 6.1 dargestellt. Das alte Bundesgebiet hat insgesamt einen Zuwachs von 5,17%, die neuen Länder einen Rückgang von 10,2%. Diese Daten enthalten allerdings alle Komponenten der Bevölkerungsentwicklung, also auch die so genannte natürliche Entwicklung, die durch Geburten und Sterbefälle entsteht. Es ist für den Zeitraum von 2000 - 2006 auch möglich, die Wanderung allein zu betrachten. Diese zeigt dieselben Tendenzen wie die Bevölkerungsentwicklung. Diese Daten wurden hier aber aus zwei Gründen nicht verwendet: Zum einen lassen sie sich zwar auf Ebene der Bundesländer aggregieren, enthalten aber auch Wanderungen zwischen Kreisen eines Bundeslandes. Zum anderen ist die Zeitreihe, wie bereits erwähnt, deutlich kürzer und enthält den interessanten Zeitraum von 1991 - 2000 gar nicht. Aus diesen Gründen wird hier die Bevölkerungsentwicklung statt der Wanderung genutzt.

Die Agglomeration von Unternehmen kann bei so großen Gebieten und ohne Fokussierung auf einzelne Branchen nur sehr allgemein gemessen werden. Nimmt

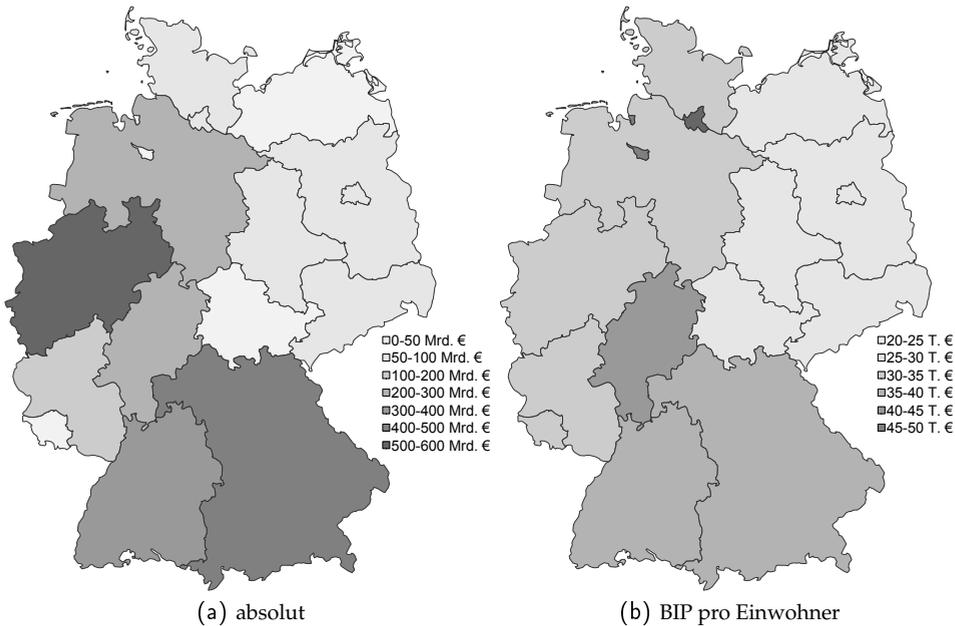


Abbildung 6.2: BIP absolut und BIP pro Kopf 2007 für Bundesländer (Daten siehe Anhang)

man an, dass die Industrie eine höhere Wertschöpfung hat als die im "Agrarsektor" zusammengefasste Produktion, so kann das regionale Bruttoinlandsprodukt (absolut oder pro Einwohner) als Näherungswert verwendet werden.¹⁵⁵ Dabei muss man beide Größen, absolutes BIP und BIP pro Einwohner betrachten, denn einerseits ändert sich die Einwohnerzahl auch in NEG Modellen, wenn die Arbeiter räumlich mobil sind, was für die Verwendung der absoluten Größe spricht, andererseits ist die Anzahl landwirtschaftlicher Arbeiter zumindest bei den meisten Modellen gleich und unveränderlich. Aus dieser Sicht wäre das BIP pro Kopf die aussagefähigere Größe.

Allerdings unterscheiden sich beide Werte qualitativ auch nur bei Bundesländern mit geringer Bevölkerung (insbesondere bei den Städten Hamburg und Bremen) und nicht zwischen Ost und West. Nur weist das absolute BIP ein deutlich größeres Intervall zwischen dem niedrigsten und dem höchsten Wert auf, als das relative.

Im Falle des absoluten BIP (Abbildung 6.2a) liegen die Bundesländer NRW, Bay-

¹⁵⁵Vgl. Brakman et al. (2001), S. 156.

ern und Baden-Württemberg mit Abstand vorn, während die hinteren Plätze durch Länder mit geringer Einwohnerzahl, wie Bremen und das Saarland besetzt sind.¹⁵⁶ Dennoch liegen auch in diesem Vergleich die neuen Länder weit hinten. Summiert man die Werte des absoluten BIP pro Bundesland für Ost und West auf, so ergibt sich ein Verhältnis von 15% in den neuen Ländern und 85% in den alten.¹⁵⁷

Die Größeneffekte drehen sich um, wenn das BIP pro Einwohner (in tausend €) betrachtet wird.¹⁵⁸ Hier liegen die Stadtstaaten Hamburg und Bremen deutlich vorn, was sicher an dem hohen Anteil der Industrie an der Gesamtproduktion liegt. Auch wird hier sehr deutlich, dass alle neuen Länder Werte zwischen 20 und 22 aufweisen, während die alten Länder zwischen 25 und 48 liegen. Selbst wenn man Hamburg und Bremen ausschließt erreichen diese Werte bis 35. Interessant ist an dieser Stelle auch, dass Berlin, als dritter Stadtstaat, nur ein BIP / Kopf von 24,81 erreicht.

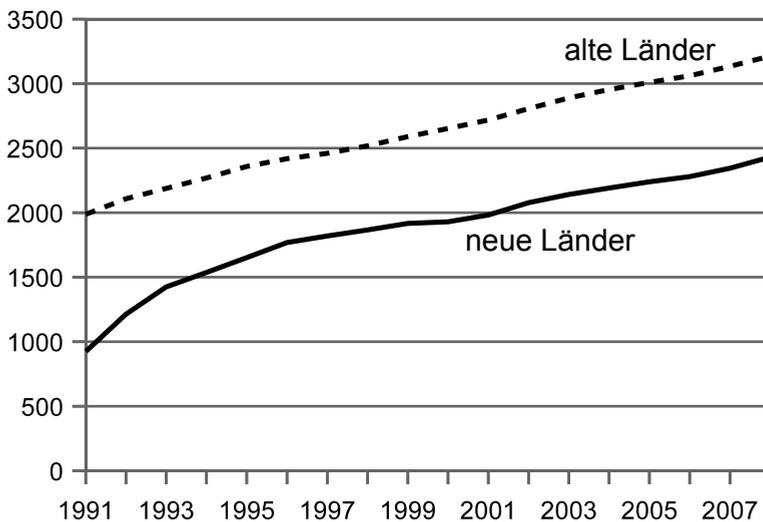


Abbildung 6.3: Durchschnittliche Bruttomonatsverdienste im produzierenden Gewerbe und im Dienstleistungsbereich ohne Sonderzahlungen (Quelle: Statistisches Bundesamt)

In den in dieser Arbeit betrachteten Modellen spielt der Lohnsatz als Auslöser für Wanderungsbewegungen eine zentrale Rolle. Da Daten über das regionale Preis-

¹⁵⁶Die Daten zur Abbildung finden sich in Tabelle A.2 im Anhang.

¹⁵⁷Summe Ost = 280.873.521; Summe West = 1.532.548.872; ohne Berlin.

¹⁵⁸Vgl. Abbildung 6.2b.

niveau in Deutschland fehlen, können annäherungsweise nur die Bruttolöhne¹⁵⁹ betrachtet werden.¹⁶⁰ Bei der Zeitreihe von 1991 bis 2008 (Abbildung 6.3) fällt auf, dass sich die Löhne von 1991 bis 1996 angenähert haben, sich seit dem aber fast parallel entwickeln. Bildet man das im Modell (s.u.) relevante Relativlohnverhältnis, so liegt dies 2008 bei 1,32.

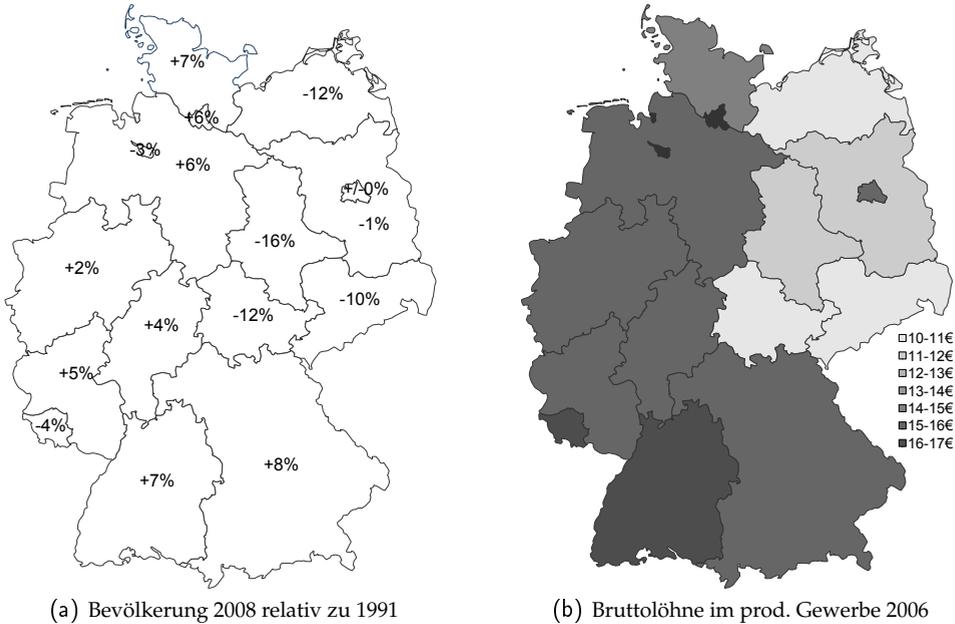


Abbildung 6.4: Bruttolohn und Wanderung (Daten siehe Anhang)

Abschließend sollen in Abbildung 6.4 die Bevölkerungsentwicklung 1991 - 2008 und der Lohnsatz 2006¹⁶¹ pro Bundesland dargestellt werden. Auch wenn dies keinen empirisch messbaren Wirkungszusammenhang darstellt, so ist doch deutlich sichtbar, dass Abwanderung eher dort stattfindet, wo die Löhne gering sind. Ein ähnliches Bild würde sich auch ergeben, wenn man die Wanderungsbewegungen und die Arbeitslosenquote gegenüberstellen würde. Darauf wurde hier

¹⁵⁹Im produzierenden Gewerbe und im Dienstleistungsbereich. Quelle: Statistisches Bundesamt.

¹⁶⁰Amtliche Daten zu regionalen Preisniveaus gibt es in Deutschland nicht. Dies ist ein Problem für viele Bereiche der Regionalökonomik. Es gibt zwar eine Reihe von Arbeiten, die auf Grundlage verfügbarer Daten versuchen, solche Indizes zu schätzen, so z.B. Roos (2006) oder Breuer und von der Lippe (im Erscheinen), diese verwenden aber meist Grundstückspreise oder Löhne und bieten keine solide Basis für eine Berechnung der Reallöhne.

¹⁶¹Neuere Werte sind auf Länderebene nicht verfügbar. Die Daten zur Abbildung finden sich in Tabelle A.3 im Anhang.

aber verzichtet, da Arbeitslosigkeit in den Modellen der NEG nicht enthalten ist. Dennoch kann daraus der Schluss gezogen werden, dass eine schlechte materielle Situation zu ökonomisch motivierter Abwanderung führt.

Zusammenfassend ist festzustellen, dass die eingangs aufgestellten Thesen empirisch gestützt wurden. Es ist eine Wanderung aus den neuen in die alten Länder zu beobachten. Das Bruttoinlandsprodukt pro Kopf ist im Westen deutlich höher als im Osten, was als Beleg für eine Agglomeration der Industrie, die mit steigenden Skalenerträgen produziert, im Westen gewertet werden kann. Außerdem besteht eine anhaltende Lohnsatzdifferenz zugunsten der alten Bundesländer. Im Folgenden sollen diese Phänomene durch ein erweitertes NEG-Modell erklärt werden.

6.4 Modell

Zur Erklärung der oben beschriebenen Entwicklung soll das modifizierte LW-Modell aus Abschnitt 3.4 verwendet werden. Dieses enthält zum einen beschränkte Mobilität der Industriearbeiter. Diese drückt sich in der Variablen φ_i aus, die als normal verteilt angenommen wird. Die Standardabweichung dieser Normalverteilung ist s . Über diesen Parameter wird nun das Maß der Mobilität ausgedrückt. Ein Wert von $s = 0$ würde bedeuten, dass die lokale Präferenz aller Arbeiter im Mittelwert der Verteilung bei 0,5 liegen würde. Dann liegen keine lokalen Präferenzen vor und alle Arbeitnehmer sind bereit, bei einem minimalen Lohnsatzunterschied zwischen den Regionen sofort zu wechseln. Ein Wert von $s > 0$ hingegen bedeutet, dass einige Arbeiter für den Wechsel entschädigt werden müssen, da sie eine Nutzeneinbuße erleiden. Sie wechseln daher nur, wenn der höhere Lohn diese kompensiert. Gerade bei der Anwendung des Modells auf den Fall der regionalen Wanderungen in Deutschland ist zu beachten, dass die Verwendung der auf das Intervall $[0, 1]$ standardisierten Normalverteilung mit Werten von $s < 0, 1$ bedeutet, dass keine Industriearbeiter vollkommen immobil sind und die Mobilität als schon recht hoch angenommen wird. Eine geringere Mobilität aber führt in diesem Modell nie zur Bildung von Agglomerationen. Im Ergebnis führt die Einbeziehung von eingeschränkter Mobilität im Vergleich zum CP-Modell, das von vollkommener Mobilität ausgeht, dazu, dass bei bestimmten Werten von T zwar eine Agglomeration in der einen Region und damit eine Entleerung der andern

entsteht, diese aber nicht vollständig ist.¹⁶² Das liegt daran, dass die Lohnsatzdifferenz nicht groß genug ist, um alle Arbeiter zur Wanderung zu bewegen. Ein Teil bleibt trotz des geringeren Lohns in der Region. In Abbildung 6.5 wäre dies z.B. bei $T = 1,45$ der Fall.

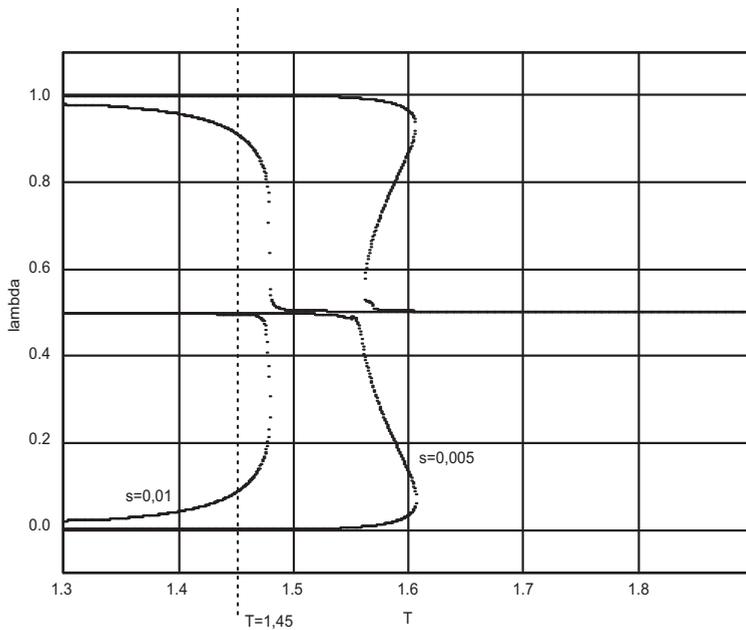


Abbildung 6.5: Bifurkationsdiagramm bei $s = 0,01$ und $s = 0,005$

Aus dieser Abbildung erkennt man aber noch etwas: je geringer die Mobilität der Arbeitskräfte ist, desto höher ist die Wahrscheinlichkeit, dass sich bei gegebenen Transportkosten keine Agglomeration bildet. So führen Transportkosten von 1,5 bei $s = 0,005$ zur Bildung einer Agglomeration, bei $s = 0,01$ aber zur Gleichverteilung der ökonomischen Aktivität. Ein Zusammenhang, auf den später noch genauer eingegangen werden soll.

Abschnitt 3.4 erweitert das Modell von Ludema und Wooton (1997) um die Annahme unterschiedlich großer Regionen. Dabei wird die Größe anhand der immobilien Arbeiter gemessen und in Region 1 ein Faktor ε eingeführt. Ist dieser 1, so sind beide Regionen gleich groß, ist er z.B. 1,05 so ist Region 1 um 5% größer als Region 2. Wenn eine Region größer im Sinne von mehr immobilien Arbeitern

¹⁶²Vgl. Abschnitt 3.2.

ist, so hat es dadurch eine höhere Kaufkraft nach Agrar- und Industriegütern. Nimmt man einmal extrem hohe Transportkosten an, so wäre auch die Verteilung der Industriearbeiter in diesem Fall ungleich. Die höhere Nachfrage nach Industriegütern müsste dann aus inländischer Produktion bedient werden.

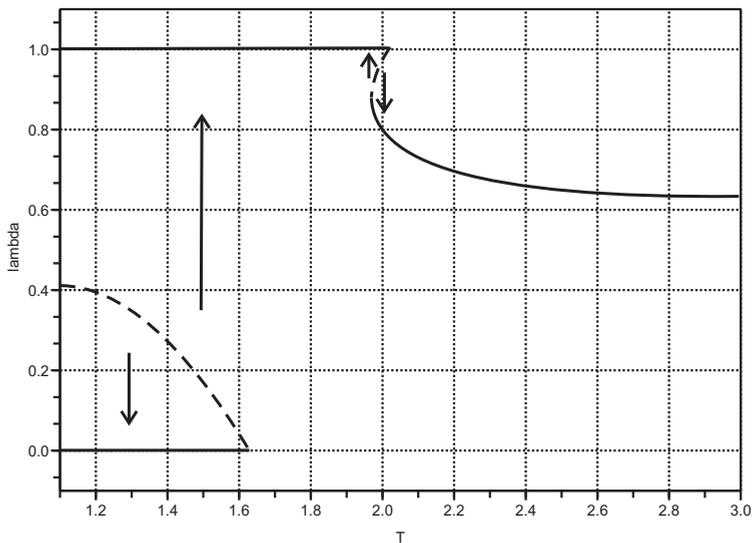
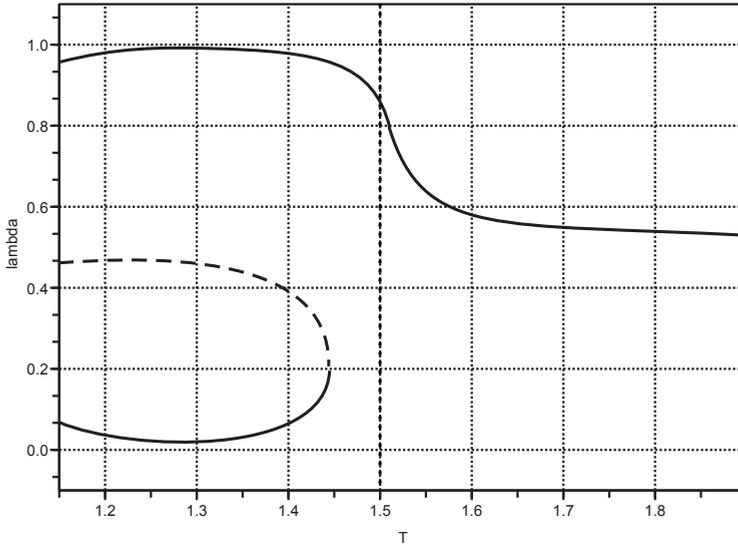


Abbildung 6.6: CP-Modell bei $\varepsilon = 1,5$

Diese Modellvariation führt, wie in Abschnitt 3.3 beschrieben, dazu, dass bei Transportkosten, die über den Break-Point hinaus sinken, sich eine Agglomeration im größeren Land bildet. Nimmt man einen deutlicheren Größenunterschied an, z.B. $\varepsilon = 1,5$, so wird unabhängig von der Ausgangssituation fast immer ein stabiles Gleichgewicht erreicht, bei dem eine Agglomeration in der größeren Region entsteht und sich die kleinere entleert. Diese, für Deutschland realistischere Situation, ist in Abbildung 6.6 dargestellt.

Bringt man nun beide Modellelemente zusammen, beschränkte Mobilität und Größenunterschiede der Regionen, so können sich stabile Gleichgewichte ergeben, die mit einer teilweisen Agglomeration in der größeren Region bei bestehenden Lohnsatzunterschieden verbunden ist. In Abbildung 6.7 ist das z.B. bei $T = 1,5$ der Fall.

Abbildung 6.7: LW-Modell bei $s = 0,01$ und $\varepsilon = 1,05$

Wir haben nun ein Modell, in dem über die Parameter T und s die Mobilität von Gütern und industriellen Arbeitskräften dargestellt werden kann. Bisher wurde in den Bifurkationsdiagrammen immer die Variation von T bei verschiedenen Werten von s untersucht. Im Zuge der deutschen Wiedervereinigung ist aber weit mehr als die Mobilität von Gütern die Mobilität der Menschen gestiegen. Es soll daher bei gegebenen Transportkosten, z.B. von $T = 1,5$, die Mobilität der Industriearbeiter verändert werden. In Abbildung 6.8 sind analog zum üblichen Bifurkationsdiagramm die stabilen und instabilen Gleichgewichte bei Variation von s dargestellt. Dabei werden stabile Gleichgewichte mit durchgezogenen Linien und instabile mit gestrichelten Linien gekennzeichnet.

Aus der Abbildung ist folgendes ersichtlich: Steigt die Mobilität von 0 (Situation mit geschlossener Grenze) kontinuierlich an¹⁶³, so verlagert sich das stabile Gleichgewicht von der Anfangsverteilung, die dem Größenverhältnis der Länder entspricht, in Richtung des größeren Landes. In diesem Bereich gibt es kein weiteres Gleichgewicht, so dass sich dieses unabhängig von der tatsächlichen Anfangsverteilung immer einstellt.

Welcher Punkt langfristig erreicht wird, hängt nun davon ab, bei welchem Wert

¹⁶³Sinkt also s ausgehend von ∞ .

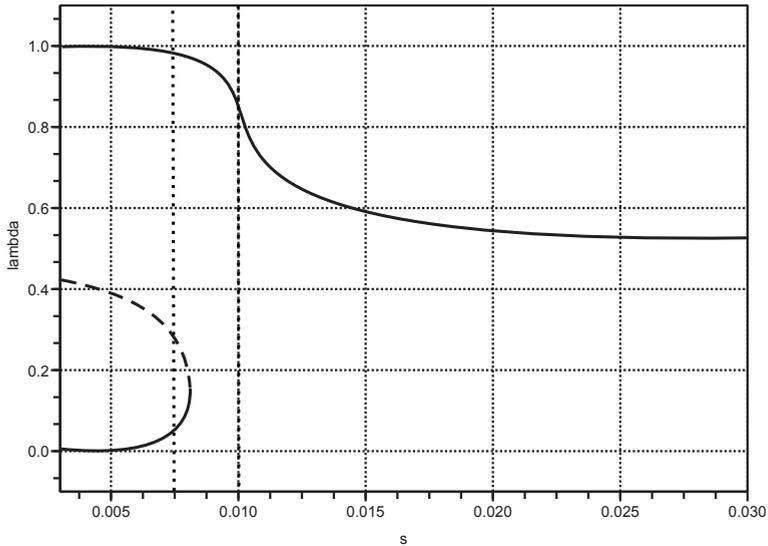


Abbildung 6.8: LW-Modell bei $T = 1,5$ und $\varepsilon = 1,05$

von s die Steigerung der Mobilität zum Halten kommt. Es wird dabei angenommen, dass die Arbeitskräfte nicht unendlich mobil sind. Nimmt die Standardabweichung z.B. einen Wert von $s = 0,01$ an, so wäre eine Verteilung, bei der ca. 85% der Industrie im Westen und 15% im Osten liegen langfristig stabil.¹⁶⁴ Das wäre genau die Aufteilung des BIP für 2007 aus Abschnitt 6.3. Betrachtet man zu dieser Situation einmal den Arbeitsmarkt (Abb. 6.9), so zeigt sich, dass hier eine Lohnsatzdifferenz zwischen Ost und West besteht, die dauerhaft ist.¹⁶⁵ Bei den gewählten Werten würde man z.B. zu einem Lohnsatzverhältnis von ca. 1,04 kommen. Obwohl das Modell nicht kalibriert ist und dies wegen der vielen schlecht schätzbaren Parameter auch kaum möglich sein dürfte, so erkennt man doch, dass die Prognose in der richtigen Größenordnung liegt.¹⁶⁶

Die beiden Ergänzungen, die als realistisch zu bezeichnen sind, führen nun dazu, dass neben der Agglomerationsbildung als solches nun auch erklärt werden kann, warum der Osten sich nicht vollständig entleert, warum wir andauernde Lohnsatzunterschiede beobachten und warum sich die Industrie vorwiegend im

¹⁶⁴Diese Situation ist mit einer roten Linie in Abbildung 6.8 markiert.

¹⁶⁵Arbeitsangebotskurve ist hier gestrichelt, die Arbeitsnachfragekurve durchgezogen dargestellt. Für eine weitere Erläuterung der Abbildung siehe auch Abschnitt 3.4.

¹⁶⁶Nimmt man die Zahlen für 2008, so ergibt sich ein Lohnsatzverhältnis von $\frac{3213}{2431} = 1,32$.

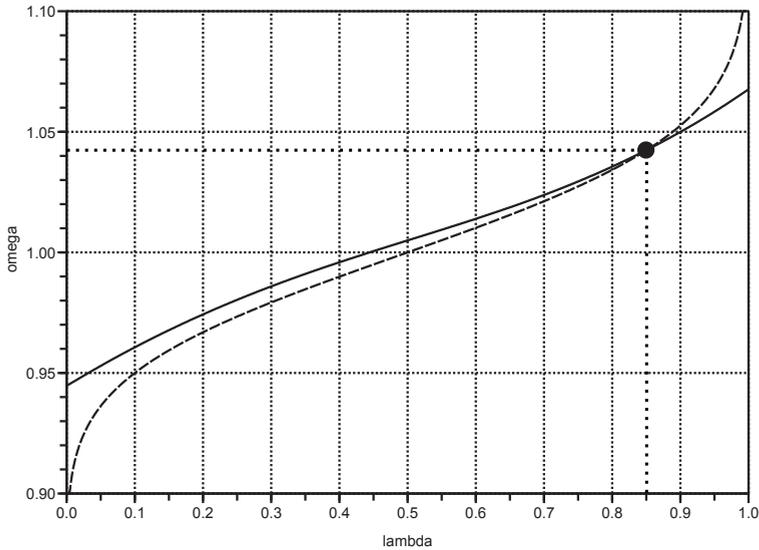


Abbildung 6.9: Dauerhafte Lohnsatzdifferenz bei $s = 0,01$, $\varepsilon = 1,05$ und $T = 1,5$

Westen ansiedelt und nicht im Osten. Trotz dieses überzeugenden Erklärungsgehalts des Modells, enthält es einige Annahmen, die kritisiert werden können. Zunächst enthält das Modell, wie alle mir bekannten Modelle der NEG, keine Arbeitslosigkeit.¹⁶⁷ Betrachtet man zunächst die Arbeitnehmer, ist dies nicht ganz so kritisch, denn Arbeitslose erhalten Arbeitslosengeld, das einem Einkommen entspricht. Dieses werden sie mit dem möglichen Arbeitseinkommen in einer anderen Region vergleichen und daraus ihre Entscheidung zu Wandern ableiten. Problematischer ist allerdings, dass Arbeitslose keine Güter produzieren und weder Sozialleistungen noch Steuern im Modell enthalten sind. Die Existenz von Arbeitslosigkeit widerspricht daher zwar nicht der Intuition hinter dem Modell, ist aber auch nicht darin enthalten. Dies ist deshalb ein Problem, weil es sich um ein Totalmodell handelt. Ein möglicher Ausweg könnte darin bestehen, die Höhe der Arbeitslosigkeit in den lokalen Präferenzen zu erfassen.¹⁶⁸ Zum einen ist dies

¹⁶⁷Laut Ottaviano (2002) ist dies ein Grundproblem der gesamten Neuen Ökonomischen Geographie.

¹⁶⁸Blien et al. (im Erscheinen) kommen in einer empirischen Studie, die auf einer Befragung von 18.954 Arbeitnehmern basiert, zu dem Ergebnis, dass der Reservationslohn sowohl durch die eigene Arbeitslosigkeit als auch durch die allgemeine Arbeitslosigkeit in der Region negativ beeinflusst wird.

aber nur ein unzureichender Ansatz, da er nur die psychologischen, nicht aber die wirtschaftlichen Auswirkungen von Arbeitslosigkeit erfasst. Zum anderen ist die Modellierung schwierig, da man sowohl die Standardabweichung als auch den Mittelwert verändern müsste, denn die Berücksichtigung von Arbeitslosigkeit ist nur sinnvoll, wenn diese in den beiden Regionen unterschiedlich hoch ist.

Der zweite Punkt, den man anführen kann, ist das Fehlen von Transportkosten innerhalb der Regionen. Transportkosten, und das gilt für die gesamte NEG, sind nicht an physische Entfernungen gebunden, sondern bestehen zwischen Regionen und nicht innerhalb. Das würde bedeuten, dass z.B. beim Transport eines Gutes von Kiel nach München keine Transportkosten anfallen, wird das gleiche Gut aber von Magdeburg nach Braunschweig transportiert, entstehen Transportkosten i.H.v. T . Das ist in der Tat nicht sonderlich realistisch.

Betrachtet man nun den Parameter s , so kann man im Prinzip die gleiche Kritik anbringen. Allerdings kann angenommen werden, dass der zu kompensierende Nutzenverlust bei einem Umzug der Arbeitskräfte nicht in erster Linie von der Entfernung abhängt, sondern eher davon, wie ähnlich die neue Region der alten ist. Dieser Überlegung folgend erscheint es dann nicht unrealistisch, dass z.B. ein Umzug von Kiel nach München mit einem geringeren Nutzenverlust verbunden ist als von Braunschweig nach Magdeburg. Dies gilt insbesondere, da zu vermuten ist, dass Einwohner aus dem Osten lieber in den Westen ziehen als umgekehrt. Diesen Zusammenhang, der aus Mangel an empirischen Daten jedoch nur vermutet werden kann, kann man mit einer Verschiebung des Mittelwertes der Verteilungsfunktion für die lokalen Präferenzen φ_i modelliert werden.

Dies ist in Abbildung 6.10 dargestellt, wobei hier der Mittelwert nur um 0,001 in Richtung Westen verschoben wurde. Die schwarzen Linien stellen die Situation mit dieser Modifikation dar, die grauen ohne. Bereits diese geringere Verschiebung verstärkt die Agglomerationstendenz im Westen deutlich.

6.5 Wirtschaftspolitische Konsequenzen

Welche Möglichkeiten hat nun der Staat, dieser Entwicklung entgegenzuwirken? Zum einen würden höhere Transportkosten tendenziell zu einer gleichmäßigeren

Dies würde für eine solche Modellierung sprechen.

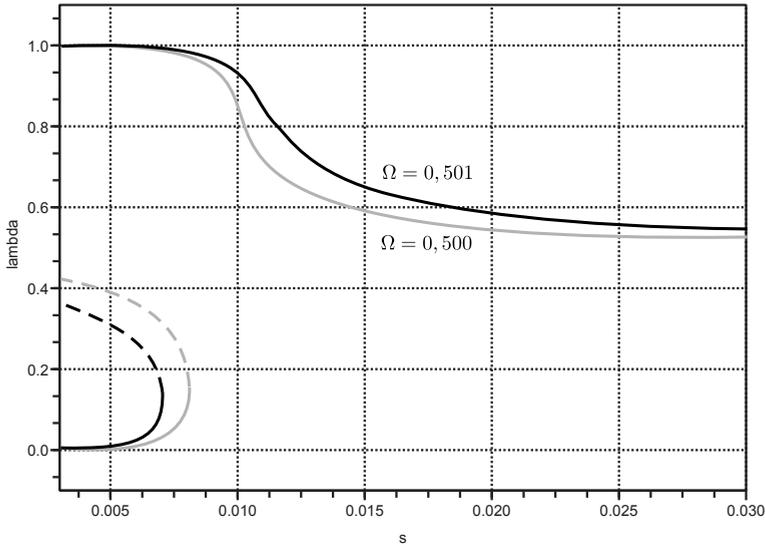


Abbildung 6.10: Unterschiede in der Attraktivität der Regionen ($\Omega = 0,501$, $T = 1,5$ und $\varepsilon = 1,05$)

Verteilung der Industrie führen (vgl. Abb. 6.7). In diesem Licht sind die erheblichen Investitionen in Infrastruktur nach der Wende auch eher negativ zu werten.¹⁶⁹ Die Instandsetzung der Infrastruktur war aber nach der Grenzöffnung ein wichtiges Projekt. Möchte man die Transportkosten wieder erhöhen, wäre dies weniger auf technische als vielmehr auf ökonomische Weise möglich. Hier zeigt sich eine Schwäche des Modells: Transportkosten fallen nur zwischen aber nicht innerhalb der Regionen an. Würde man z.B. die Abgaben auf Benzin oder die Kfz-Steuer erhöhen, so würde sich dies negativ auf die gesamte Produktion auswirken. Eine künstliche Verteuerung des Transports, der nur bei Überschreiten der Regionsgrenze zum Tragen kommt, ist nicht möglich, bzw. politisch nicht vertretbar. Bei solchen Maßnahmen wäre auch zu überlegen, ob dann nicht doch ein separater Transportsektor modelliert werden müsste, da es z.B. zu Beschäftigungseffekten bei Speditionen käme.

Eine weitere Möglichkeit, die Entleerung des Ostens zu verhindern, liegt in der Umverteilung durch Steuern und Subventionen. Da der Größenunterschied durch

¹⁶⁹Vgl. auch Schempp (2000), S. 319. Martin und Rogers (1995) weisen darauf hin, dass eine bessere Infrastruktur nicht nur Ansiedlungen in den peripheren Gebieten attraktiver macht, sondern auch dazu genutzt werden kann, die Peripherie aus der Kernregion zu versorgen.

eine unterschiedlich hohe Kaufkraft in Ost und West modelliert wird, sind Umverteilungen von Einkommen prinzipiell geeignet, diese Unterschiede auszugleichen. Ross (2001) untersucht verschiedene Arten von Subventionen, allerdings in einem etwas anderen Modell, das aber diesem recht ähnlich ist. Er kommt zu dem Ergebnis, dass Subventionen an die Unternehmen im Osten, die durch Besteuerung der Unternehmen im Westen finanziert werden, am effizientesten sind. Als politisch leichter durchsetzbar sieht Ross die Besteuerung von Haushalten im Westen und die Subventionierung der Produktion im Osten. Dieser, im Vergleich zur Subventionierung des Konsums im Osten immer noch recht effiziente Lösung, kann tatsächlich der Agglomerationsbildung im Westen entgegenwirken.¹⁷⁰

Ein weiterer Mechanismus, aus dem sich ebenfalls eine Option für die Politik ergibt, ist bereits in der Diskussion weiter oben angesprochen worden: Die lokalen Präferenzen der Haushalte mit ihren Parametern s (Standardabweichung) und Ω (Mittelwert). Wie aus Abbildung 6.8 ersichtlich ist, steigt der Grad der Konzentration wirtschaftlicher Aktivität im Westen mit steigender Mobilität.

Gelingt es bei den hier verwendeten Werten z.B. den Parameter s von den angenommenen 0,1 auf 0,15 zu bringen, ergibt sich nur noch eine Verteilung von 60% / 40% auf Westen und Osten. Die Standardabweichung ist ein Maß für die Streuung der Präferenzen. Ein hohes s weist auf eine Verbundenheit der Bürger mit der Region hin, in der sie leben. Dies ist unabhängig von der Region selbst und darf nicht mit der generellen Attraktivität einer Region verwechselt werden, die weiter unten thematisiert wird. Die Verbundenheit der Bürger mit ihrer Region könnte z.B. durch die Bereitstellung von preiswertem Bauland, einem besseren Angebot an Kultur oder der Förderung des sozialen Lebens, z.B. von Vereinen, gesteigert werden.

Der Parameter Ω gibt den Mittelwert der Verteilung der lokalen Präferenzen an. Liegt dieser Wert, wie hier angenommen, bei 0,5, so bedeutet das, dass zwar einige Bürger lieber im Westen und andere lieber im Osten leben, aber deren Anzahl immer gleich ist. Es gibt also keinen generellen Attraktivitätsvorteil für eine Region, in der die Einwohner im Mittel lieber wohnen würden. Gelänge es allerdings, Ω nur um 0,1% in Richtung Osten zu verschieben, so würde das den Größenunterschied fast ausgleichen.

Diese Maßnahme ist in Abbildung 6.11 dargestellt. Man erkennt, dass bei $\Omega =$

¹⁷⁰Vgl. Ross (2001).

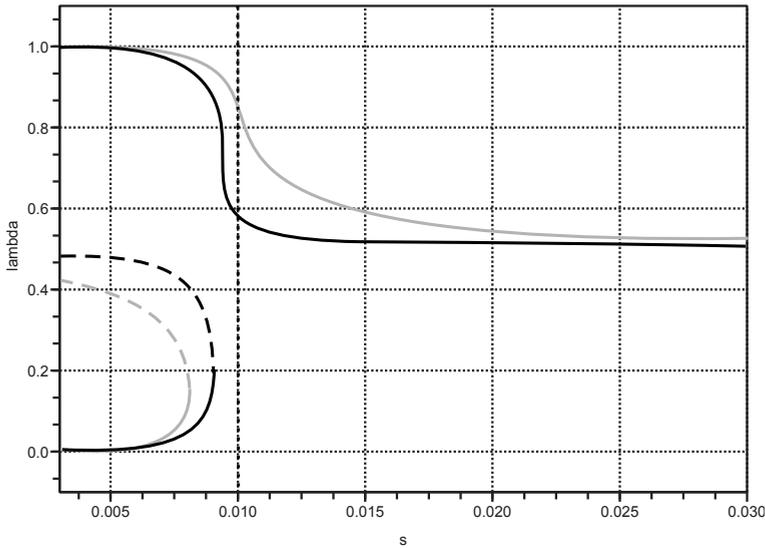


Abbildung 6.11: Unterschiede in der Attraktivität der Regionen ($\Omega = 0,499$, $T = 1,5$ und $\varepsilon = 1,05$)

0,499 (schwarze Kurve) das einzige stabile Gleichgewicht bei einer Aufteilung der industriellen Produktion von ca. 57% im Westen und 43% im Osten liegt. Um diese leichte Verschiebung der Präferenzen zu erreichen, müsste einseitig im Osten die Lebensqualität gesteigert und dies auch bekannt gemacht werden. Ob diese Maßnahme, die nur sehr langfristig wirkt, eine realistische Option ist, darf angezweifelt werden.¹⁷¹ Nach dem hier vorgestellten Modell ist sie allerdings recht wirksam. Denkbar ist auch eine Kombination der hier genannten Maßnahmen, z.B. eine leichte Verschiebung von Ω bei gleichzeitiger Erhöhung von s . Das gilt vor allem deshalb, weil die konkreten Maßnahmen vermutlich sowieso auf beide Parameter wirken.

Eine generelle Frage ist die, in welchem Zeitraum politische Maßnahmen wirken und in wie weit man deren Wirkung voraussagen kann. Dabei sind technische Instrumente, die z.B. die Transportkosten verändern, am schnellsten wirksam und auch am besten abschätzbar. Ökonomische Instrumente, wie beispielsweise Subventionen und Steuern, erfordern die Reaktion von Wirtschaftssubjekten und sind daher sowohl mit einer zeitlichen Verzögerung als auch mit einer gewissen Unsicherheit behaftet, die aber noch abschätzbar ist. Psychologische Mittel, z.B. die

¹⁷¹Dennoch sich solche Imagekampagnen zu beobachten, z.B. "Chemnitz - Stadt der Moderne".

Steigerung der Lebensqualität sowie Imagekampagnen für bestimmte Regionen sind da schon problematischer. Hier kann nur sehr schwer eingeschätzt werden, ob und wann die Wirtschaftssubjekte reagieren.

Die hier vorgestellten Möglichkeiten des Staates, Einfluss auf die Verteilung der industriellen Produktion zu nehmen, implizieren, dass der Staat dies will. Grundlage kann der Grundsatz gleicher Lebensverhältnisse sein, der oft zur Begründung von Regionalpolitik und regionaler Wirtschaftsförderung herangezogen wird.¹⁷² Dennoch ist anzumerken, dass Agglomerationen hier die Gesamteffizienz steigern, da Skaleneffekte ausgenutzt werden. Folglich haben Maßnahmen dagegen eine Effizienzeinbuße zur Folge. Es sei hier ausdrücklich darauf hingewiesen, dass dieses Ziel hier unterstellt wird ohne die Sinnhaftigkeit zu diskutieren. Gerade wenn psychologische Faktoren eine Rolle spielen und durch die Steigerung der Attraktivität einer Region versucht wird, die Menschen zum Bleiben zu bewegen, kann dies kritisch gesehen werden. Man kann hier auch argumentieren, dass, wenn eine Mehrheit aus psychologischen Gesichtspunkten heraus nicht in einer Region leben möchte, die Politik hier nicht gegensteuern sollte. Andererseits bedeutet die Abwanderung von Teilen der Bevölkerung, insbesondere von jüngeren und gebildeteren Mitgliedern, dass diejenigen, die bleiben dann schlechter dastehen. Dies auszugleichen kann sicher ein sinnvolles politisches Ziel sein.

Letztlich sind dies aber Werturteile, die nicht Gegenstand dieser Arbeit sind. Um überhaupt eine Basis für eine solche Diskussion zu haben, wäre zunächst eine Berechnung der Wohlfahrt notwendig. Baldwin et al. (2003) schlagen hier vor, die indirekten Nutzen der Haushalte aufzusummieren, was aber eine utilitaristische Soziale Wohlfahrtsfunktion impliziert.¹⁷³ Aber selbst wenn eine solche Wohlfahrtsbetrachtung zu dem Schluss käme, dass Agglomeration die Wohlfahrt steigert, lässt sich daraus ohne eine Wertentscheidung keine Handlungsempfehlung ableiten. Der Grund liegt darin, dass in jedem Fall der Nutzen der im Entleerungsgebiet verbleibenden Einwohner durch geringere Löhne und ein höheres Preisniveau sinkt, von den Folgen massiver Abwanderung auf das soziale System ganz abgesehen. Ob und wie ein Nutzensausgleich im Sinne Kaldors über die Regionen hinweg möglich ist, bleibt unklar.¹⁷⁴

¹⁷²Vgl. Abschnitt 6.1.

¹⁷³Vgl. Baldwin et al. (2003), S. 257.

¹⁷⁴Der Kaldor-Kompensations-Test prüft, ob die Summe, welche die nach der politischen Maßnahme besser gestellten Wirtschaftssubjekte bereit wären, für die Durchführung zu Zahlen, die Summe

Das hier vorgestellte Modell ist also nicht nur in der Lage, die anhand von empirischen Daten skizzierte Situation zu erklären, sie zeigt auch mögliche Auswege auf und bewertet diese hinsichtlich ihrer Wirksamkeit. Allerdings beinhaltet das Modell einige restriktive Annahmen, insbesondere die fehlende Arbeitslosigkeit. Außerdem wirken die gezeigten politischen Maßnahmen erst bei Erreichen bestimmter Schwellenwerte. Steigt in Abbildung 6.11 z.B. die Mobilität auf $s = 0,005$, so hat die Verschiebung des Mittelwertes nahezu keine Auswirkungen auf die Verteilung der Industrie.

übersteigt, welche die Geschädigten fordern würden, um der Maßnahme zuzustimmen. Vgl. Mas-Collel et al. (1995), S. 852. Ottaviano (2002) weist drauf hin, dass Wohlfahrtsüberlegungen in der Neuen Ökonomischen Geographie noch unzureichend implementiert sind und sich daraus auch ein Problem bei der Ableitung von Politikempfehlungen ergibt.

7 Fazit

In dieser Arbeit werden Asymmetrien in der Neuen Ökonomischen Geographie (NEG) aus modelltheoretischer und methodischer Sicht untersucht und deren wirtschaftspolitische Implikationen diskutiert.

Ein großer Teil der Modelle der NEG, und insbesondere die hier betrachteten grundlegenden Modelle, können nicht numerisch gelöst werden. Da es zwar Literatur zur Simulation aus mathematischer Sicht und Programmierhandbücher für die entsprechenden Softwarepakete gibt, jedoch keine Quelle zu Computersimulationen mit Bezug zur NEG gefunden werden konnte, werden in Kapitel 5 entsprechende Methoden entwickelt.

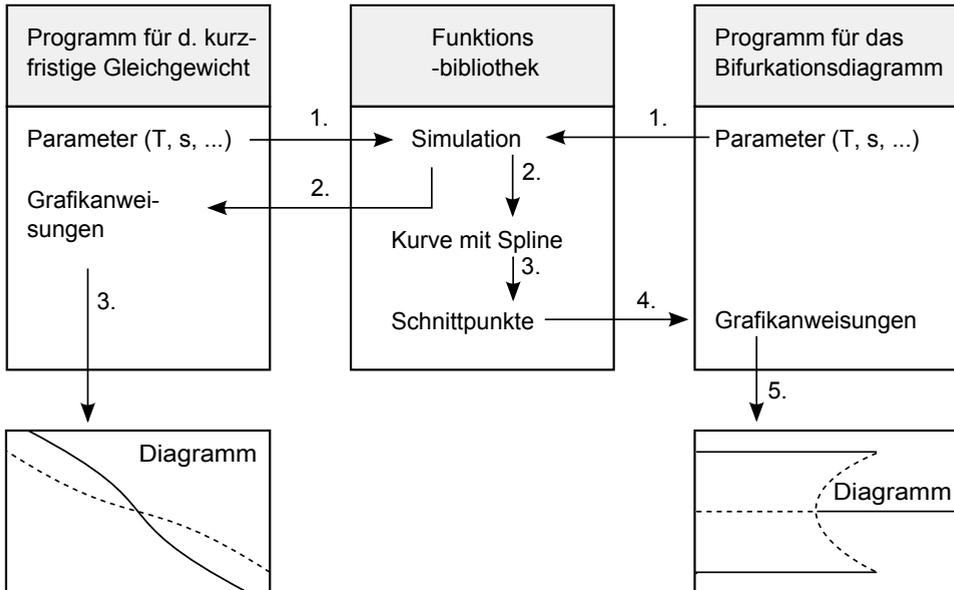


Abbildung 7.1: Schema der Simulation kurzfristiger Gleichgewichte

Ziel dieses Kapitels ist es, zum einen die Ergebnisse dieser Arbeit nachvollziehbar

und überprüfbar zu machen und zum anderen, ein Instrumentarium zu schaffen, mit dem auch andere Modelle der NEG gelöst werden können. Dabei werden sowohl mathematisch-methodische Aspekte diskutiert, als auch die konkrete Umsetzung im Programmcode vorgestellt.

Da sich die verschiedenen Modelle in der Formulierung, nicht jedoch im grundlegenden Aufbau unterscheiden, ist auch deren Lösungsweg im Prinzip gleich. Um die Programmierung zu vereinfachen, ist daher für jedes Modell eine Funktionsbibliothek erstellt worden, in der sich Funktionen für die Bestimmung des kurzfristigen Gleichgewichtes und einige Hilfsfunktionen befinden. Diese werden dann aus den eigentlichen Programmen heraus aufgerufen, die dadurch deutlich kürzer und übersichtlicher werden. Änderungen müssen dann nur noch an einer Stelle vorgenommen werden. Abbildung 7.1 zeigt dies schematisch.

Möchte man dieses Verfahren auf andere Modelle übertragen, müssen daher nur der Simulationscode für das kurzfristige Gleichgewicht und einige Funktionsaufrufe geändert werden. Besonders vorteilhaft ist dabei die Verwendung von kubischen Splines. Dadurch werden die Ergebnisse der Simulation auf eine einheitliche Skala transformiert. Dies führt dazu, dass die Weiterverarbeitung unabhängig von dem eigentlichen Simulationscode wird.

Aus theoretischer Perspektive besteht das Ziel der Arbeit darin, zu untersuchen, welche Ergebnisse sich einstellen, wenn die üblichen Symmetrieanahmen der Neuen Ökonomischen Geographie aufgegeben werden. In dieser Modellwelt gibt es inzwischen eine große Zahl verschiedener Modelle, sodass hier zunächst das Problem besteht, welche davon als Ausgangspunkt verwendet werden sollen. Da die Frage, wie Asymmetrien wirken, grundsätzlicher Natur ist, wurden möglichst grundlegende und allgemeine Modelle ausgewählt. Bei Modellen, die mobile Arbeitskräfte voraussetzen und sich daher auf Regionen beziehen, ist dies zunächst das Core-Periphery-Modell von Krugman (1991b) sowie die Erweiterung von Ludema und Wooton (1997). Letztere eignet sich aufgrund ihrer plausibleren Ergebnisse besonders für wirtschaftspolitische Überlegungen. Nimmt man statt Regionen, zwischen denen Arbeiter relativ frei wechseln, Länder an, sind die Modelle von Krugman und Venables (1995) und Krugman und Venables (1996) am bedeutendsten. Obwohl beide strukturell recht ähnlich sind, beschreibt das erste Agglomerationen und das zweite Konzentration von Branchen.

Zunächst gilt für alle betrachteten Modelle, dass eine Aufgabe der Symmetrie-

annahme nicht dazu führt, dass die Modelle nicht mehr lösbar sind, solange die Asymmetrien nicht zu stark ausfallen. Diese zunächst triviale Feststellung ist im Rahmen der Neuen Ökonomischen Geographie dennoch bemerkenswert, weil diese Modelle sehr komplex sind und verschiedene andere Veränderungen, beispielsweise der Produktionstechnologie, nicht (oder nur sehr begrenzt) zulassen.

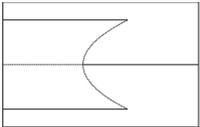
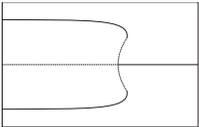
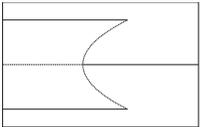
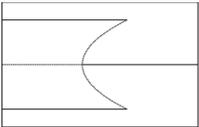
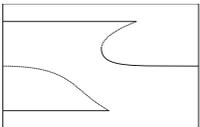
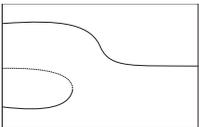
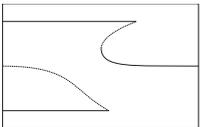
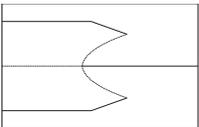
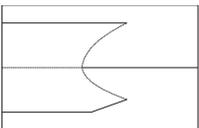
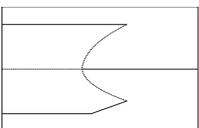
	CP-Modell	LW-Modell	IM-Modell	Cluster-Modell
Basis	 (Kap. 3.1)	 (Kap. 3.2)	 (Kap. 4.1)	 (Kap. 4.3)
Asym. Reg.	 (Kap. 3.3)	 (Kap. 3.4)	 (Kap. 4.2)	 (Kap. 4.5)
Asym. T				 (Kap. 4.4)
Asym. μ				 (Kap. 4.6)

Tabelle 7.1: Bifurkationsdiagramme im Vergleich

Zur besseren Übersicht sind in Tabelle 7.1 sind die Bifurkationsdiagramme aller betrachteten Modelle vereinfacht dargestellt. In allen Modellen hat die Berücksichtigung von Asymmetrien die Aussagen der Basismodelle nicht wiederlegt sondern präzisiert. Im Fall des CP-Modells ohne und mit lokale Präferenzen (LW-Modell) und des IM-Modells (Zwischenprodukte) führt die Annahme unterschiedlich großer Länder bei sinkenden Transportkosten immer zur Agglomerationsbildung im größeren Land. Abschnitt 2.5 hat gezeigt, dass sinkende Transportkosten in den vergangenen Jahrzehnten zu beobachten waren. Diese dürften die

heutige Wirtschaftsstruktur beeinflusst haben. Besonders interessant ist hier der Fall mit lokalen Präferenzen (LW-Modell). Zum einen verläuft die Entwicklung, anders als in den beiden anderen Modellen, ohne Sprünge ab. Zum anderen erlaubt das Modell unvollständige Agglomerationen bei bestehenden Lohndifferenzen. Aufgrund dieser Eigenschaften wird es auch in einer wirtschaftspolitischen Anwendung in Kapitel 6 verwendet.

Nimmt man unterschiedlich große Regionen an, kommen sowohl das CP-Modell als auch das Modell mit Zwischenprodukten qualitativ zum gleichen Ergebnis. Daraus ergibt sich, dass unabhängig von der Faktormobilität sowohl große Regionen als auch große Länder eine höhere Wahrscheinlichkeit haben, Agglomerationen anzuziehen. Der Grund liegt in der größeren Binnennachfrage und den damit verbundenen stärkeren Skaleneffekten auf dem Absatzmarkt, sowie an günstigeren Faktorpreisen auf dem Beschaffungsmarkt.

Das Clustermodell, auch wenn es dem Modell mit Zwischenprodukten formal sehr ähnelt, hat einen anderen Erklärungsgegenstand: Es beschreibt die Bildung von regionaler Spezialisierung. Agglomeration und Entleerung können mit diesem Modell nicht dargestellt werden, weil es keinen Agrarsektor gibt und folglich auch keine Entleerung von industrieller Aktivität möglich ist. Das Fehlen eines Sektors ohne Skaleneffekte bedingt auch die andersartigen Ergebnisse, was bereits aus der Form der Bifurkationsdiagramme in Abbildung 7.1 zu erkennen ist.

Zunächst fällt auf, dass keine Tendenz zu beobachten ist, dass das größere Land Vorteile hat. Dies gilt auch, wenn man ungleich große Sektoren, ausgedrückt in einem höheren Ausgabenanteil μ , unterstellt (Abschnitt 4.7). Kommt es allerdings zur räumlichen Konzentration der Branchen, so ist diese im kleineren Land stets vollständig. Im größeren Land dagegen ist auch eine unvollständige Spezialisierung möglich, bei der noch eine geringe Menge des Gutes produziert wird, dessen Produktion im kleineren Land konzentriert ist.

Solche unvollständigen Spezialisierungen (von Agglomeration kann man im Kontext des Cluster-Modells nicht sprechen) kommen bei ungleich großen Ländern in beiden Sektoren vor. Im Fall von sektorspezifischen Transportkosten oder ungleichen Ausgabenanteilen treten sie nur jeweils in einem Sektor auf, und zwar in demjenigen mit den geringeren Transportkosten oder dem höheren Ausgabenanteil. Insgesamt ist es im Cluster-Modell möglich, mehr Asymmetrien zu modellieren, als in den anderen Modellen, was daran liegt, dass es hier zwei Sektoren mit

unvollständiger Konkurrenz gibt.

Vergleicht man die Ergebnisse dieser Arbeit mit denen ähnlicher Modelle der Außenwirtschaft und der Stadtökonomik, wie sie in Abschnitt 1.2 kurz vorgestellt wurden, so lässt sich feststellen, dass sich diese nicht widersprechen. Einige kommen zu den gleichen Aussagen. Allerdings können die Ergebnisse nicht direkt gegenüber gestellt werden, da die Modelle zum Teil sehr unterschiedlich sind.

Um die hier beschriebenen Modelle für die Analyse von wirtschaftspolitischen Maßnahmen nutzen zu können, bedarf es einer Zielvorstellung. Daher wird in Abschnitt 6.1 zunächst diskutiert, welche ökonomischen und sozialen Argumente dafür sprechen könnten, Agglomerationsbildung zu begrenzen. Hier besteht ein Konflikt zwischen dem Ausgleichsziel und dem Effizienzziel, der nicht auf ökonomischer sondern auf politischer Ebene geklärt werden muss. Hilfreich wäre es allerdings, und hier besteht noch Forschungsbedarf, wenn Wohlfahrtsbetrachtungen in der NEG stärker entwickelt wären.

Allgemein ist zu erkennen, dass durch die veränderten Rahmenbedingungen Gleichgewichte möglich sind, in denen Regionen auf eine der Branchen spezialisiert sind, aber dennoch ein Teil der Produktion der anderen Branche dort angesiedelt ist. Sind die betrachteten Länder unterschiedlich groß, ist dieser Effekt eher im größeren Land zu erwarten. Dies Ergebnis ist deutlich befriedigender als die Lösung des symmetrischen Modells, bei dem sich entweder eine vollständige Spezialisierung einstellt oder gar keine.

Gerade bei Aussagen, die auf dem Cluster-Modell beruhen, muss allerdings beachtet werden, dass dieses nur zwei Sektoren beinhaltet, die zudem beide unter steigenden Skalenerträgen arbeiten. Außerdem werden nur zwei Länder angenommen. Diese Annahmen haben mehrere Implikationen: Zunächst ist eine Entleerung einer Region dadurch per Definition ausgeschlossen. Konzentriert sich eine Branche in einem Land, so ergibt sich daraus zwangsläufig auch die Konzentration der anderen Branche im anderen Land. Auch arbeiten beide Sektoren unter den gleichen steigenden Skalenerträgen. Dies ist im Modell, wenn überhaupt, nur schwer aufzuheben, weil damit die Standardisierungen nicht mehr in der Form möglich wären, diese aber zur Lösung notwendig sind. Aus der empirischen Forschung wissen wir aber, dass nur bei ca. 30% der Industrie überhaupt steigende Skalenerträge festzustellen sind.¹⁷⁵ Die Ergebnisse müssen also mit ei-

¹⁷⁵Vgl. Abschnitt 2.4.

ner gewissen Skepsis betrachtet werden. Das bedeutet allerdings nicht, dass sie für wirtschaftspolitische Empfehlungen ungeeignet sind, formalisieren sie doch nachvollziehbare Erklärungen für empirisch beobachtete Phänomene. Die Relevanz der NEG für solche Überlegungen, gerade auf internationaler Ebene, wird auch durch deren Verwendung im *World Development Report 2009* belegt. Speziell für Europa liegen eine Reihe von empirischen Studien vor, die sich auf die NEG beziehen¹⁷⁶. Schätzungen, die den Zusammenhang von Asymmetrien und dem Grad der Spezialisierung der Länder Europas belegen, würden den Rahmen dieser Arbeit sprengen, wären aber zweifellos interessant.

Im Rahmen von Agglomerationsbildung auf regionaler Ebene ist insbesondere das Modell von Ludema und Wooton (1997) von Interesse, da es auf dem CP-Modell aufbaut und durch die Annahme von lokalen Präferenzen unvollständige Agglomeration bei Lohnsatzunterschieden erlaubt. In Abschnitt 3.4 wurde dieses Modell durch ungleich große Regionen erweitert und in Abschnitt 6.4 statt der Transportkosten die Mobilität variiert. Zusammen haben diese Änderungen zu einem Modell geführt, das sehr schlüssig die Veränderungen der regionalen Wirtschaftsstruktur nach der Wiedervereinigung beschreibt.

In diesem Modellrahmen ist es außerdem möglich, verschiedene politische Maßnahmen im Hinblick auf ihre Wirkung auf die Agglomerationsbildung zu untersuchen. Zum einen sind dies Maßnahmen, die eine Erhöhung der Transportkosten zur Folge haben. Diese sind aber politisch schwer durchzusetzen. Zum anderen werden Instrumente betrachtet, welche die Bereitschaft der Arbeiter zum Wechsel der Region verringern. Dies mag verwundern, da eine hohe Mobilität oft als positiv gewertet wird. Allerdings muss man sich hier bewusst machen, dass in dieser Arbeit weder eine wohlfahrtstheoretische Analyse durchgeführt wurde, noch beinhaltet das Modell Arbeitslosigkeit. Die Aussage des Modells ist vielmehr die, dass eine geringere Bereitschaft umzuziehen, begründet durch Mittelwert und Varianz der Präferenzen für eine Region, die Agglomerationsbildung bremst. Ob dies ein sinnvolles Ziel ist, muss auf politischer Ebene entschieden werden. Tatsache ist aber, dass Kampagnen, die das Image einer Region aufwerten sollen, vielfach zu beobachten sind.

Kapitel 6 stellt ein Beispiel dar, welche regionalpolitischen Folgerungen aus den hier gewonnenen methodischen und empirischen Ergebnissen gezogen werden

¹⁷⁶Z.B. Klüver (2000), Midelfart-Knarvik et al. (2000) und Pflüger und Südekum (2005).

können. Es zeigt aber deutlich der Erklärungsgehalt der Neuen Ökonomischen Geographie, auch über das rein theoretische Interesse hinaus. Dieser Erklärungsgehalt lässt sich durch die Annahme der Asymmetrie noch deutlich erhöhen und es sind eine ganze Reihe weiterer Anwendungen denkbar. Hier ist noch Raum für zukünftige Forschung vorhanden.

Literaturverzeichnis

- Aghion, P. und Howitt, P. (2005). 'Groth with Quality-Improving Innovations: An Integrated Framework.' In P. Aghion und S. Durlauf (Hg.), *Handbook of Economic Groth*, Bd. 1a.
- Anderson, J.E. und van Wincoop, E. (2004). 'Trade Costs.' *Journal of Economic Literature*, Bd. XLII, S. 691–751.
- Baldwin, R., Forslid, R., Martin, P. et al. (2003). *Economic Geography and Public Policy*. Princeton University Press, Princeton und Oxford.
- Bickenbach, F. und Bode, E. (2006). 'Disproportionality Measures of Concentration, Specialization, and Polarization.' Working Paper 1276, Institut für Weltwirtschaft, Kiel.
- Blien, U., Messmann, S. und Trappmann, M. (im Erscheinen). 'Do Reservation Wages React to Regional Unemployment?' Discussion paper, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.
- Brakman, S. und Garretsen, H. (1993). 'The Relevance of Initial Conditions for the German Unification.' *Kyklos*, Bd. 46(2), S. 163–181.
- Brakman, S. und Garretsen, H. (2003). 'Rethinking the New Geographical Economics.' *Regional Studies*, Bd. 37, S. 637–648.
- Brakman, S., Garretsen, H., Gorter, J. et al. (2005). 'New Economic Geography, Empirics, and Regional Policy.' Working Paper 56, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.
- Brakman, S., Garretsen, H., Marrewijk, C. et al. (2000a). 'Empirical Research in Geographical Economics.' Paper presented at the conference The Monopolistic Competition Revolution after Twenty-Five Years.

- Brakman, S., Garretsen, H. und Schramm, M. (2000b). 'The Empirical Relevance of the New Economic Geography: Testing for a Spatial Wage Structure in Germany.' CESifo Working Paper 395, CESifo, München.
- Brakman, S., Garretsen, H. und van Marrewijk, C. (2001). *An Introduction to Geographical Economics: Trade, Location and Growth*. Cambridge University Press, Cambridge.
- Bröcker, J. (2004). 'Computable General Equilibrium Analysis in Transportation Economics.' In D. Hensher, K.J. Button, K.E. Haynes und P.R. Stopher (Hg.), *Handbook of Transport Geography and Spatial Systems*, S. 269–292. Elsevier.
- Breuer, C.C. und von der Lippe, P. (im Erscheinen). 'Datengewinnung im periodischen regionalen Preisvergleich - Die Problematik der Mieten und Immobilienpreise.' *Jahrbuch für Regionalwissenschaft*.
- Brühlhart, M. (2009). 'An Account of Global Intra-Industry Trade.' *The World Economy*, Bd. 32(3), S. 401–459.
- Cabrales, A. und Motta, M. (2001). 'Country Asymmetries, endogenous Product Choice and the Timing of Trade Liberalization.' *European Economic Review*, Bd. 45, S. 87–107.
- Chamberlin, E.H. (1933). *The Theory of Monopolistic Competition*. Harvard University Press, Cambridge, Mass.
- Davis, D.A. und Weinstein, D.E. (1998). 'Market Access, Economic Geography and Comparative Advantage: An Empirical Assessment.' Discussion Paper 6787, NBER, Cambridge.
- Dixit, A.K. und Stiglitz, J.E. (1977). 'Monopolistic Competition and Optimum Product Diversity.' *The American Economic Review*, Bd. 67, S. 297–308.
- Eckey, H.F. und Kosfeld, R. (2004). 'New Economic Geography.' *Volkswirtschaftliche Diskussionsbeiträge 65/04*, Universität Kassel, Kassel.
- Ethier, W.J. (1982). 'National and International Returns to Scale in the Modern Theory of International Trade.' *The American Economic Review*, Bd. 72(3), S. 389–405.
- Fackler, P.L. und Miranda, M.J. (2002). *Applied Computational Economics and Finance*. MIT Press.

- Forslid, R. und Ottaviano, G.I.P. (2003). 'An Analytically Solvable Core-Periphery Model.' *Journal of Economic Geography*, Bd. 3, S. 229–240.
- Forslid, R. und Wooton, I. (2003). 'Comparative Advantages and the Location of Production.' *Review of International Economics*, Bd. 11(4), S. 588–603.
- Frohwerk, S. (2008). 'Das Cluster-Modell bei sektoral unterschiedlichen Transportkosten.' In *Seminarbericht Nr. 51*. Gesellschaft für Regionalforschung, Heidelberg.
- Fujita, M. und Hamaguchi, N. (2001). 'Intermediate Goods and the Spatial Structure of an Economy.' *Regional Science and Urban Economics*, Bd. 31, S. 79–109.
- Fujita, M., Krugman, P. und Venables, A.J. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. The MIT Press, Cambridge, 2. Aufl.
- Fujita, M. und Thiesse, J.F. (2002). *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge University Press, Cambridge.
- Gallo, F. (2006). 'Increasing Returns, Input-Output Linkages, and Technological Leapfrogging.' *The B.E. Journal of Economic Analysis and Policy*, Bd. 6.
- Geppert, K. und Gornig, M. (2005). 'Regionale Konvergenz- und Polarisationsprozesse in der Europäischen Union.' *Vierteljahreshefte zur Wirtschaftsforschung*, Bd. 74, S. 8–25.
- Grossman, G.M. und Helpman, E. (1991). 'Quality Ladders in the Theory of Growth.' *Review of Economic Studies*, Bd. 58(1), S. 43–61.
- Grubel, H.G. und Lloyd, P. J. (1975). *Intra-Industry Trade: The Theory and Measurement of International Trade with Differentiated Products*. Macmillan, London.
- Haas, A. und Südekum, J. (2005). 'Spezialisierung und Branchenkonzentration in Deutschland.' Kurzbericht, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.
- Hanson, G.H. (2005). 'Market Potential, Increasing Returns and Geographic Concentration.' *Journal of International Economics*, Bd. 67(1), S. 1–24.
- Head, K. und Mayer, T. (2004). 'The Empirics of Agglomeration and Trade.' In V. Henderson und Thisse (Hg.), *Handbook of Regional and Urban Economics*, Bd. 4, S. 2609–2696. North Holland, Amsterdam.

- Hummels, D. (1999a). 'Have International Transportation Costs Declined?' Working Paper 1158, Center for Global Trade Analysis, Chicago.
- Hummels, D. (1999b). 'Toward a Geography of Trade Costs.' Working Paper 1196, Center for Global Trade Analysis, Chicago.
- Hummels, D. (2007). 'Transportation Costs and International Trade over Time.' *Journal of Economic Perspectives*, Bd. 21, S. 131–154.
- International Organization for Migration (2010). 'Global Estimates and Trends.' <http://www.iom.int>.
- Jehle, G.A. und Reny, P.J. (2001). *Advanced Microeconomic Theory*. Addison-Wesley, Boston, 2 Aufl.
- Judd, K.L. (1998). *Numerical Methods in Economics*. MIT Press, Cambridge.
- Klüver, A. (2000). *Die Auswirkung der Integration Europas auf die Agglomeration von Industriesektoren*. Dissertation, Universität Passau.
- Knaap, T. (2004). *Models of Economic Geography - Dynamics, Estimation and Policy Evaluation*. Labyrinth Publications, Ridderkerk (Netherlands).
- Krugman, P. (1980). 'Scale Economies, Product Differentiation, and the Pattern of Trade.' *American Economic Review*, Bd. 70(5), S. 950–959.
- Krugman, P. (1981). 'Intraindustry Specialization and the Gains from Trade.' *Journal of Political Economy*, Bd. 89(5), S. 959–973.
- Krugman, P. (1991a). *Geography and Trade*. MIT Press, Cambridge.
- Krugman, P. (1991b). 'Increasing Returns and Economic Geography.' *Journal of Political Economy*, Bd. 99(3), S. 483–499.
- Krugman, P. und Obstfeld, M. (2009). *Internationale Wirtschaft*. Pearson Education, 8 Aufl.
- Krugman, P. und Venables, A.J. (1995). 'Globalization and the Inequality of Nations.' *The Quarterly Journal of Economics*, Bd. 110(4), S. 857–880.
- Krugman, P. und Venables, A.J. (1996). 'Integration, Specialization and Adjustment.' *European Economic Review*, Bd. 40, S. 959–967.
- Ludema, R.D. und Wooton, I. (1997). 'Regional Integration, Trade, and Migration: Are Demand Linkages Relevant in Europe?' Discussion Paper 1656, Georgetown University.

- Maier, G. und Tödtling, F. (2002). *Regional- und Stadtökonomik 2*. Springer, 2 Aufl.
- Mare, D.C. (2005). 'Concentration, Specialisation and Agglomeration of firms in New Zealand.' Working Paper 05-12, Motu Economic and Public Policy Research.
- Martin, P. und Rogers, C.A. (1995). 'Industrial Location and Public Infrastructure.' *Journal of International Economics*, Bd. 39, S. 335–351.
- Martin, R. (1999). 'The new 'geographical turn' in Economics: Some Critical Reflections.' *Cambridge Journal of Economics*, Bd. 23, S. 65–91.
- Mas-Collel, A., Whinston, M.D. und Green, J.R. (1995). *Microeconomic Theory*. Oxford University Press, Oxford.
- Midelfart-Knarvik, K.H., Overman, H.G., Redding, S.J. et al. (2000). 'The Location of European Industry.' Economic Papers 142, European Commission, Brussels.
- Neary, P.N. (2001). 'Of Hype and Hyperbolas: Introducing the New Economic Geography.' *Journal of Economic Literature*, Bd. 39, S. 536–561.
- Obstfeld, M. und Taylor, A.M. (2003). 'Globalization and Capital Markets.' In M.D. Bordo, A.M. Taylor und J.G. Williamson (Hg.), *Globalization in Historical Perspective*. University of Chicago Press.
- Ottaviano, G. (2002). 'Regional Policy in the Global Economy: Insights from New Economic Geography.' Discussion Paper 211, HWWA, Hamburg.
- Ottaviano, G., Tabuchi, T. und Thisse, J.F. (2002). 'Agglomeration and Trade Revisited.' *International Economic Review*, Bd. 43, S. 409–435.
- Ottaviano, G. und Thiesse, J.F. (2001). 'On Economic Geography in Economic Theory: Increasing Returns and Pecuniary Externalities.' *Journal of Economic Geography*, Bd. 1, S. 153–179.
- Overman, H.G. (2004). 'Can we learn Anything from Economic Geograpfy Proper?' *Journal of Economic Geography*, Bd. 4, S. 501–516.
- Pflüger, M. (2008). 'Die Neue Ökonomische Geographie: Ein Überblick.' *WiSt*, Bd. 3/2008, S. 150–158.
- Pflüger, M. und Südekum, J. (2005). 'Die Neue Ökonomische Geographie und Effizienzgründe für Regionalpolitik.' *Vierteljahreshefte zur Wirtschaftsforschung*, Bd. 74, S. 26–46.

- Pflüger, M. und Südekum, J. (2008). 'A Synthesis of Footloose-Entrepreneur New Economic Geography Models: When is Agglomeration Smooth and Easily Reversible?' *Journal of Economic Geography*, Bd. 8, S. 39–54.
- Pincon, B. (k.A.). 'Eine Einführung in Scilab.' Handbuch, Université Henri Poincaré, Nancy.
- Puga, D. (1999). 'The Rise and Fall of Regional Inequalities.' *European Economic Review*, Bd. 43(2), S. 303–334.
- Puga, D. (2002). 'European regional Policies in Light of Recent Location Theories.' *Journal of Economic Geography*, Bd. 2, S. 373–406.
- Reichelt, R. (2008). *Deutschland und seine neue ökonomische Geographie*. Dissertation, Freie Universität Berlin.
- Romer, P. (1986). 'Increasing Returns in Long-Run Growth.' *Journal of Political Economy*, Bd. 94(5), S. 1002–1037.
- Roos, M. (2002). *Ökonomische Agglomerationstheorien - Die Neue Ökonomische Geographie im Kontext, Wirtschaftsgeographie und Wirtschaftsgeschichte*, Bd. 10. Eul Verlag, Lohmar.
- Roos, M. (2006). 'Regional price levels in Germany.' *Applied Economics*, Bd. 38(13), S. 1553–1566.
- Ross, M. (2001). 'Transfers, Agglomeration and German Unification.' Discussion Paper 144, HWWA, Hamburg.
- Rossi-Hansberg, E. (2005). 'A Spatial Theory of Trade.' *American Economic Review*, Bd. 95(5), S. 1464–1491.
- Samuelson, P.A. (1952). 'The Transfer Problem and Transport Costs: The Terms of Trade when Impediments are absent.' *Economic Journal*, Bd. 62, S. 278–304.
- Schempp, U. (2000). 'Über das lange Warten auf den Aufschwung Ost - Der Nordosten Deutschlands nach der Wende durch die theoretische Brille von Paul Krugman gesehen.' In *Wachstum, Strukturwandel, Wettbewerb*. Walter, H. and Hegner, S. and Schechler, J. M., Stuttgart.
- Schöler, K. (2005). *Raumwirtschaftstheorie*. Vahlen Verlag, München.
- Schöler, K. (2007). 'Transformationsprozesse und Neue Ökonomische Geographie - Erklärungsbeiträge der Neuen Ökonomischen Geographie zur Transformation

- der ostdeutschen Volkswirtschaft.' In S. Kauf (Hg.), *Regionalpolitik in der Transformationszeit - Ziele, Erfahrungen, Perspektiven, Festschrift zum 70. Geburtstag von Klaus Gloede*. Opole.
- Schoenberg, I.J. (1946). 'Contributions to the problem of approximation of equidistant data by analytic functions.' *Quarterly Applied Mathematics*, Bd. 4, S. 45–99 and 112–141.
- Südekum, J. (2006). 'Concentration and Specialization Trends in Germany since Re-unification.' *Regional Studies*, Bd. 40(8), S. 861–873.
- Shy, O. (2000). *Industrial Organization*. The MIT Press, Cambridge.
- Siebert, H. (1997). *Weltwirtschaft*. Lucius und Lucius, Stuttgart.
- Spence, M. (1976). 'Product Selection, Fixed Costs, and Monopolistic Competition.' *The Review of Economic Studies*, Bd. 43(2), S. 217–235.
- Sydsaeter, K., Stroem, A. und Berck, P. (1999). *Economists' Mathematical Manual*. Springer, Heidelberg, New York, 3 Aufl.
- Tabuchi, T. und Thisse, J.F. (2006). 'Regional Specialisation, Urban Hierarchy, and Commuting Costs.' *International Economic Review*, Bd. 47(4), S. 1295–1317.
- Takatsuka, H. und Zeng, D.Z. (2005). 'Regional Specialization via Differences in Transport Costs.' Paper presented at the ERSA Conference 2005.
- Tharakan, J. und Thisse, J.F. (2002). 'The Importance of Being Small. Or when Countries are Areas and not Points.' *Regional Science and Urban Economics*, Bd. 32, S. 381–408.
- Tibor, S. (1954). 'Two Concepts of External Economies.' *The Journal of Political Economy*, Bd. 62, S. 143–151.
- Tirole, J. (1994). *The Theory of Industrial Organization*. MIT Press, Cambridge, Mass.
- Varian, H. (2007). *Mikroökonomie*. Oldenbourg Verlag, München, 7 Aufl.
- Venables, A.J. (1995). 'International Location of Economic Activity: Economic Integration and the Location of Firms.' *The American Economic Review*, Bd. 85, S. 296–300.
- Venables, A.J. (1996). 'Equilibrium Location of vertically linked Industries.' *International Economic Review*, Bd. 37, S. 341–359.

Venables, A.J. (1999). 'The International Devision of Industries: Clustering and Competitive Advantage in a Multi-industry Model.' *Scandinavian Journal of Economics*, Bd. 101, S. 495–513.

Woll, A. (2003). *Allgemeine Volkswirtschaftslehre*. Vahlen, München, 14 Aufl.

World Bank (2009). *World Development Report 2009: Reshaping Economic Geography*. Washington.

Zogg, J.M. (2007). 'Arbeiten mit Scilab und Scicos.' Handbuch, HTW Chur.

Anhänge

A Simulationen

In diesem Abschnitt sind die für die Erstellung der Diagramme verwendeten Simulationen angegeben. Dabei können in der Regel durch geeignete Wahl der Werte für die Variablen mehrere Asymmetrien mit dem gleichen Programm simuliert werden. Daher ist hier immer die universelle Lösung angegeben. Formatierungsanweisungen für die grafische Ausgabe wurden im Sinne einer übersichtlicheren und kürzeren Darstellung größtenteils weggelassen.

A.1 CP-Modell und LW-Modell (Abschnitte 2.9, 3.1 bis 3.4 und 6.4)

Beide Modelle sind relativ ähnlich. Daher können teilweise die gleichen Simulationen verwendet werden und deshalb werden diese auch hier gemeinsam dargestellt. Die Arbeitsnachfrage in beiden Modellen wird gleich bestimmt.

Mathematische Formulierung der Arbeitsnachfrage:

$$Y_1 = \mu \lambda w_1 + \varepsilon \cdot \frac{1 - \mu}{2} \tag{A.1}$$

$$Y_2 = \mu (1 - \lambda) w_2 + \frac{1 - \mu}{2} \tag{A.2}$$

$$Q_1 = \left[\lambda w_1^{1-\sigma} + (1 - \lambda) (w_2 T)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{A.3}$$

$$Q_2 = \left[\lambda (w_1 T)^{1-\sigma} + (1 - \lambda) w_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{A.4}$$

$$w_1 = \left[Y_1 Q_1^{\sigma-1} + Y_2 T^{1-\sigma} Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \tag{A.5}$$

$$w_2 = \left[Y_1 Q_1^{\sigma-1} T^{1-\sigma} + Y_2 Q_2^{\sigma-1} \right]^{\frac{1}{\sigma}} \quad (\text{A.6})$$

$$\omega_1 = w_1 Q_1^{-\mu} \quad (\text{A.7})$$

$$\omega_2 = w_2 Q_2^{-\mu} \quad (\text{A.8})$$

Die Umsetzung im Programmcode dazu:

```

function [r1, r2] = mySimulateLWD(sigma, my, T, e)
L = -0.01;
w1 = 1;
w2 = 1;
for j = 1:101
    L = L+0.01;
    for i = 1:1000
        w11 = w1;
        w21 = w2;
        y1 = (my*L)*w1+(1+e)*(1-my)/2;
        y2 = (my*(1-L))*w2+(1-my)/2;
        Q1 = (L*(w1^(1-sigma))+(1-L)*((w2*T)^(1-sigma)))^(1/(1-sigma));
        Q2 = (L*((w1*T)^(1-sigma))+(1-L)*(w2^(1-sigma)))^(1/(1-sigma));
        w1 = (y1*(Q1^(sigma-1))+(y2*(Q2^(sigma-1)))*(T^(1-sigma)))^(1/
            sigma);
        w2 = ((y1*(Q1^(sigma-1)))*(T^(1-sigma))+y2*(Q2^(sigma-1)))^(1/
            sigma);
        if w1<=0 then
            break
        elseif w2<=0 then
            break
        end;
        if abs((w1-w11)/w11)<0.0001 then
            if abs((w2-w21)/w21)<0.0001 then
                break
            end;
        end;
        temp1 = w1*(Q1^(-my));
        temp2 = w2*(Q2^(-my));
    end;
    r1(j) = L;
    r2(j) = temp1/temp2;
end;
endfunction

```

Sind beide Regionen gleich groß, ist $e = 0$, sonst gibt e an, um wie viel Prozent

Region 1 größer ist als Region 2, z.B. $e = 0,05$ für einen Unterschied von 5%.

Da im CP-Modell das Arbeitsangebot vollkommen elastisch ist, wird keine spezielle Arbeitsangebotsfunktion benötigt. Um die Gleichgewichte zu finden, werden die Nullstellen der oben angegebenen Arbeitsnachfragefunktion bestimmt. Das LW-Modell hingegen beinhaltet eine Arbeitsangebotsfunktion. Die folgende Simulation stellt die verwendete *Truncated Normal Distribution Function* graphisch dar.

```

clear;
x0 = linspace(0,1,100)';
m = 0.5;
s = 0.3;
x1 = (x0-m)/s;
y0 = exp(-(x1.^2)/2)/sqrt(2*%pi);
[P1,Q0]=cdfnorf("PQ", (1-m)/s,m,s);
[P0,Q1]=cdfnorf("PQ", (0-m)/s,m,s);
P0
P1
y1 = ((1/s)*y0)/(P1-P0);
plot2d(x0,y1);
xlabel("Präferenz_für_Region_1");

```

Die Ausgabe der Variablen P0 und P1 erfolgt zur Kontrolle. Wenn Diese Variablen die Werte 1 und 0 annehmen, kann später für diese Standardabweichungen die Funktion `cdfnorm()` zur Berechnung der kumulierten Wahrscheinlichkeit verwendet werden, da die Verteilung dann die Grenzen 0 und 1 nicht überschreitet und diese daher vernachlässigt werden können. Dies ist für $s = 0,001$; $s = 0,005$ und $s = 0,05$ der Fall, nicht jedoch für $s = 0,1$. Daher wird in der Arbeitsangebotsfunktion des LW-Modells nur die Funktion `cdfnorm()` verwendet.

```

function [r1, r2] = mySimulateLWS(s)
m = 0.5;
r2 = linspace(0.9,1.1)';
for i = 1 : 100
    g = 1/(r2(i)+1);
    [P3, temp] = cdfnorf("PQ",g,m,s);
    r1(i) = temp;
end;
endfunction

```

Wie in Abschnitt 5.3.2 beschrieben, ist es für die Berechnung der Gleichgewichte,

aber auch für die Darstellung, oft sinnvoll, mittels Splines auf eine einheitliche Skala gebrachte Funktionen zu verwenden. Dies geschieht mit folgendem Code

```
// Arbeitsangebot im LW-Modell
function zz = myLWS(s)
[y1, y2] = mySimulateLWS(s);
zz = interp(L, y1, y2, splin(y1,y2,"fast"));
endfunction

//Arbeitsnachfrage im LW-Modell
function zz = myLWD(sigma, my, T, 0)
[k1, k2] = mySimulateLWD(sigma, my, T);
zz = interp(L, k1, k2, splin(k1,k2,"fast"));
endfunction
```

Die Abbildung der kurzfristigen Gleichgewichte im CP-Modell wird erstellt mit

```
clear;
chdir("D:\Aktuell\Dissertation\Simulation\LW");
getf("lib_lw.sci");
w = linspace(0.9,1.1)';
L = linspace(0,1,1001)';
T = 1.7;
dd1 = myLWD(5, 0.4, T);
ss = [0 1; 1 1];
plot2d(L,dd1, rect=[0,0.9,1,1.1]);
plot2d(ss(:,1), ss(:,2));
```

und für das LW-Modell

```
clear;
chdir("D:\Aktuell\Dissertation\Simulation\LW");
getf("lib_lw.sci");
w = linspace(0.9,1.1)';
L = linspace(0,1,1001)';
T = 1.5;
s = 0.01;
dd = myLWD(5, 0.4, T);
ss = myLWS(s);
plot2d(L, dd, rect=[0,0.9,1,1.1]);
xset('line_style',2);
plot2d(L, ss);
```

Das Bifurkationsdiagramm für das CP-Modell wird erstellt mit

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\LW");
getf("lib_lw.sci");
getf("../lib_ge.sci");
rts = [];
T0 = 1.3;
T1 = 2.0;
L = linspace(0,1,1001)';
T = linspace(T0,T1,500)';
for i = 1 : length(T)
    dd = myLWD(5, 0.4, T(i));
    xx = dd-1;
    r = myRoots(L,xx);
    for j = 1 : length(r) rts = [rts; T(i), r(j)]; end;
end;
plot2d(rts(:,1), rts(:,2), style=-0, rect=[T0,-0.1,T1,1.1]);

```

Dabei werden die Nullstellen berechnet mit

```

function [result]=myRoots(x,y)
result = [];
for i = 1 : length(y)-1
    if sign(y(i)) ~= sign(y(i+1)) then
        result = [result ((x(i)+x(i+1))/2)];
    end;
end;
endfunction

```

Diese Funktion wird auch für alle anderen Modelle verwendet.

Bifurkationsdiagramm für das LW-Modell:

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\LW");
getf("lib_lw.sci");
getf("../lib_ge.sci");
rts = [];
T0 = 1.3;
T1 = 1.9;
s = 0.01;
L = linspace(0,1,1001)';
T = linspace(T0,T1,500)';
for i = 1 : length(T)
    dd = myLWD(5, 0.4, T(i));
    ss = myLWS(s);

```

```

xx = dd-ss;
r = myRoots(L,xx);
for j = 1 : length(r) rts = [rts; T(i), r(j)]; end;
end;
plot2d(rts(:,1), rts(:,2), style=-0, rect=[T0,-0.1,T1,1.1]);

```

Code für das Bifurkationsdiagramm (Abb. 6.8) bei Variation von s aus Abschnitt 6.4:

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\LW");
getf("lib_lw.sci");
getf("../lib_ge.sci");
rts = [];
s0 = 0.003;
s1 = 0.03;
T = 1.5;
L = linspace(0,1,1001)';
s = linspace(s0,s1,500)';
for i = 1 : length(s)
    dd = myLWD(5, 0.4, T);
    ss = myLWS(s(i));
    xx = dd-ss;
    r = myRoots(L,xx);
    for j = 1 : length(r) rts = [rts; s(i), r(j)]; end;
end;
plot2d(rts(:,1), rts(:,2), style=-0, rect=[s0,-0.1,s1,1.1]);

```

A.2 Modell mit Zwischenprodukten (Abschnitte 4.1 und 4.2)

Mathematische Formulierung der Gleichgewichtsbedingung in einem Land (auch hier bereits mit der Möglichkeit, Größenunterschiede der Regionen zu berücksichtigen):

$$Q_h = \left[\varepsilon_h L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} + \varepsilon_f L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (\text{A.9})$$

$$Q_f = \left[\varepsilon_f L_f w_f^{1-\beta\sigma} Q_f^{-\alpha\sigma} + \varepsilon_h L_h w_h^{1-\beta\sigma} Q_h^{-\alpha\sigma} T^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (\text{A.10})$$

$$E_h = \mu [w_h \varepsilon_h L_h + \varepsilon_h (1 - L_h)] + \frac{\alpha w_h \varepsilon_h L_h}{\beta} \quad (\text{A.11})$$

$$E_f = \mu \left[w_f \varepsilon_f L_f + \varepsilon_f (1 - L_f) \right] + \frac{\alpha w_f \varepsilon_f L_f}{\beta} \quad (\text{A.12})$$

$$w_h = \left[\frac{\beta \left(E_h Q_h^\sigma Q_f + E_f Q_f^\sigma T^{1-\sigma} Q_h \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_h^{-\frac{\alpha}{\beta}} \quad (\text{A.13})$$

$$w_f = \left[\frac{\beta \left(E_f Q_f^\sigma Q_h + E_h Q_h^\sigma T^{1-\sigma} Q_f \right)}{Q_h Q_f} \right]^{\frac{1}{\sigma\beta}} \cdot Q_f^{-\frac{\alpha}{\beta}} \quad (\text{A.14})$$

Im symmetrischen Fall gilt $\varepsilon_h = \varepsilon_f = 1$.

Simulationscode für das kurzfristige Gleichgewicht:

```

function [r1, r2] = mySimulateIM2(T, e1, e2)
// Festsetzen der Konstanten
s = 5; // =sigma
m = 0.4; // =mu
a = 0.5; // =alpha
b = 1-a; // =beta
// Festsetzen der Startwerte
wh = 1; wf = 1;
Qh1 = 1; Qf1 = 1;
Lh = 0.9;
for l = 1:88
Lf = l/100;
for j = 1:100
// Die Preisindices (Länder i, a und Sektoren 1, 2)
for k = 1:50
for i = 1:80 // Q iterativ bestimmen
Qh0 = Qh1;
Qf0 = Qf1;
Qh1 = (e1*Lh*wh^(1-b*s)*Qh0^(-a*s)+e2*Lf*wf^(1-b*s)*Qf0^(-a*s)*
T^(1-s))^(1/(1-s));
Qf1 = (e2*Lf*wf^(1-b*s)*Qf0^(-a*s)+e1*Lh*wh^(1-b*s)*Qh0^(-a*s)*
T^(1-s))^(1/(1-s));
// Abbruchkriterium
if abs(Qh1-Qh0) < 0.001
if abs(Qf1-Qf0) < 0.001 then break
end;
end;
end; // for i
if i > max_i then max_i = i
end;

```

```

// E berechnen
Eh = m*(wh*e1*Lh+e1*(1-Lh))+((a*wh*e1*Lh)/b);
Ef = m*(wf*e2*Lf+e2*(1-Lf))+((a*wf*e2*Lf)/b);
// w berechnen
wh0 = wh;
wf0 = wf;
wh = ((b*(Qh1^s*Eh*Qf1+Qf1^s*Ef*T^(1-s)*Qh1))/(Qh1*Qf1))^(1/(b*s))
      )*Qh1^(-a/b);
wf = ((b*(Qf1^s*Ef*Qh1+Qh1^s*Eh*T^(1-s)*Qf1))/(Qh1*Qf1))^(1/(b*s))
      )*Qf1^(-a/b);
if abs(wh-wh0) < 0.01
    if abs(wf-wf0) < 0.01 then
        break
    end;
end;
end; // for k
if abs(wh-1) < 0.001 then break end;
if real(wh) > 1 then Lh = Lh + 0.001;
    else Lh = Lh - 0.001;
    end; //if
end; //for j
L(l) = Lh;
w(l) = wh;
Lx(l) = Lf;
end; // for l
r1 = Lx;
r2 = L;
endfunction

```

Kurzfristiges Gleichgewicht (hier ohne Splines, wegen des Problems der strengen Monotonie):

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\IM");
getf("lib_im.sci");
T = 2.15;
[k1, k2] = mySimulateIM2(T,1.05,1);
[v1, v2] = mySimulateIM2(T,1,1.05);
// Inlandskurve:
plot2d(k2, k1, rect=[0,0,1,1], style=5);
// Auslandskurven:
plot2d(v1, v2, rect=[0,0,1,1], style=5);

```

Bifurkationsdiagramm:

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\IM");
getf("lib_im.sci");
getf("../lib_ge.sci");
rts = [];
T0 = 1;
T1 = 3;
L = linspace(0,1,1001)';
T = linspace(T0,T1,300)';
for i = 1 : length(T)
    try
        hh = myIMI(T(i));
        ff = myIMA(T(i));
        xx = hh-ff;
        r = myRoots(L,xx);
        for j = 1 : length(r) rts = [rts; T(i), r(j)]; end;
        mprintf("T=%f_o.k._\n\r", T(i));
    catch
        mprintf("Fehler_bei_T=%f_\n\r", T(i))
    end
end;
plot2d(rts(:,1), rts(:,2), style=-0, rect=[T0,-0.1,T1,1.1]);

```

mit den Funktionen

```

function zz = myIMI(T)
[k1, k2] = mySimulateIM2(T,1.05,1);
zz = interp(L, k1, k2, splin(k1,k2,"fast"));
endfunction

function zz = myIMA(T)
[k1, k2] = mySimulateIM2(T,1,1.05);
m = size(k1);
s = m(1);
for i = 1 : s
    v1(i) = k1(s+1-i);
    v2(i) = k2(s+1-i);
end;
fprintfMat('.\matrix-v.txt',[v1, v2]);
zz = interp(L, v2, v1, splin(v2,v1,"fast"));
endfunction

```

Für eine Begründung dieser Funktionen sowie den Code der Funktion `myRoots()` siehe oben beim LW-Modell.

A.3 Cluster-Modell (Abschnitte 4.3 bis 4.8)

Mathematische Formulierung für das Cluster-Modell mit allen Erweiterungen:

$$Q_{h1} = \left[\varepsilon_h L_h w_{h1}^{1-\beta\sigma} Q_{h1}^{-\alpha\sigma} Q_{h2}^{-\nu\sigma} + \varepsilon_f L_f w_{f1}^{1-\beta\sigma} Q_{f1}^{-\alpha\sigma} Q_{f2}^{-\nu\sigma} T_1^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (\text{A.15})$$

$$Q_{h2} = \left[\varepsilon_h (1 - L_h) w_{h2}^{1-\beta\sigma} Q_{h2}^{-\alpha\sigma} Q_{h1}^{-\nu\sigma} + \varepsilon_f (1 - L_f) w_{f2}^{1-\beta\sigma} Q_{f2}^{-\alpha\sigma} Q_{f1}^{-\nu\sigma} T_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (\text{A.16})$$

$$Q_{f1} = \left[\varepsilon_f L_f w_{f1}^{1-\beta\sigma} Q_{f1}^{-\alpha\sigma} Q_{f2}^{-\nu\sigma} + \varepsilon_h L_h w_{h1}^{1-\beta\sigma} Q_{h1}^{-\alpha\sigma} Q_{h2}^{-\nu\sigma} T_1^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (\text{A.17})$$

$$Q_{f2} = \left[\varepsilon_f (1 - L_f) w_{f2}^{1-\beta\sigma} Q_{f2}^{-\alpha\sigma} Q_{f1}^{-\nu\sigma} + \varepsilon_h (1 - L_h) w_{h2}^{1-\beta\sigma} Q_{h2}^{-\alpha\sigma} Q_{h1}^{-\nu\sigma} T_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (\text{A.18})$$

$$E_{h1} = \mu \cdot (w_{h1} \varepsilon_h L_h + w_{h2} \varepsilon_h (1 - L_h)) + \frac{\alpha w_{h1} \varepsilon_h L_h + \nu w_{h2} \varepsilon_h (1 - L_h)}{\beta} \quad (\text{A.19})$$

$$E_{h2} = (1 - \mu) \cdot (w_{h1} \varepsilon_h L_h + w_{h2} \varepsilon_h (1 - L_h)) + \frac{\nu w_{h1} \varepsilon_h L_h + \alpha w_{h2} \varepsilon_h (1 - L_h)}{\beta} \quad (\text{A.20})$$

$$E_{f1} = \mu \cdot (w_{f1} \varepsilon_f L_f + w_{f2} \varepsilon_f (1 - L_f)) + \frac{\alpha w_{f1} \varepsilon_f L_f + \nu w_{f2} \varepsilon_f (1 - L_f)}{\beta} \quad (\text{A.21})$$

$$E_{f2} = (1 - \mu) \cdot (w_{f1} \varepsilon_f L_f + w_{f2} \varepsilon_f (1 - L_f)) + \frac{\nu w_{f1} \varepsilon_f L_f + \alpha w_{f2} \varepsilon_f (1 - L_f)}{\beta} \quad (\text{A.22})$$

$$w_{h1} = \left[\frac{\beta \left(Q_{h1}^\sigma \cdot E_{h1} \cdot Q_{f1} + Q_{f1}^\sigma \cdot E_{f1} \cdot T^{1-\sigma} \cdot Q_{h1} \right)}{Q_{h1} \cdot Q_{f1}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{h1}^{-\frac{\alpha}{\beta}} \cdot Q_{h2}^{-\frac{\nu}{\beta}} \quad (\text{A.23})$$

$$w_{h2} = \left[\frac{\beta \left(Q_{h2}^\sigma \cdot E_{h2} \cdot Q_{f2} + Q_{f2}^\sigma \cdot E_{f2} \cdot T^{1-\sigma} \cdot Q_{h2} \right)}{Q_{h2} \cdot Q_{f2}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{h2}^{-\frac{\alpha}{\beta}} \cdot Q_{h1}^{-\frac{\nu}{\beta}} \quad (\text{A.24})$$

$$w_{f1} = \left[\frac{\beta \left(Q_{f1}^\sigma \cdot E_{f1} \cdot Q_{h1} + Q_{h1}^\sigma \cdot E_{h1} \cdot T^{1-\sigma} \cdot Q_{f1} \right)}{Q_{f1} \cdot Q_{h1}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{f1}^{-\frac{\alpha}{\beta}} \cdot Q_{f2}^{-\frac{\nu}{\beta}} \quad (\text{A.25})$$

$$w_{f2} = \left[\frac{\beta \left(Q_{f2}^{\sigma} \cdot E_{f2} \cdot Q_{h2} + Q_{h2}^{\sigma} \cdot E_{h2} \cdot T^{1-\sigma} \cdot Q_{f2} \right)}{Q_{f2} \cdot Q_{h2}} \right]^{\frac{1}{\sigma\beta}} \cdot Q_{f2}^{-\frac{\alpha}{\beta}} \cdot Q_{f1}^{-\frac{\nu}{\beta}} \quad (\text{A.26})$$

Simulation zu dieser Formulierung:

```

function [r1, r2] = mySimulateCMI(T1, T2, e1, e2)
// Festsetzen der Konstanten
s = 3; // =sigma
a = 0.4; // =alpha
m = 0.52; // =mu
n = 0.1; // =nu
b = 1-a-n; // =beta
e1 = e1 + 1;
e2 = e2 + 1;
// Festsetzen der Startwerte
wh1 = 1; wh2 = 1; wf1 = 1; wf2 = 1;
Qf1 = 4; Qf2 = 4; Qh1 = 4; Qh2 = 4;
Li = 0.9;
for l = 1:101
    La = (l-1)/100;
    La1 = La*e2;
    La2 = (1-La)*e2;
    for j = 1:1000
        Lil = Li * e1;
        Li2 = (1-Li) * e1;
        // Die Preisindices
        for k = 1:100
            for i = 1:5 // Q iterativ bestimmen
                Qf1x = (e1*Lil*wh1^(1-b*s)*Qf1^(-a*s)*Qf2^(-n*s)+e2*La1*wf1
                    ^ (1-b*s)*Qh1^(-a*s)*Qh2^(-n*s)*T1^(1-s))^(1/(1-s));
                Qf2x = (e1*Li2*wh2^(1-b*s)*Qf2^(-a*s)*Qf1^(-n*s)+e2*La2*wf2
                    ^ (1-b*s)*Qh2^(-a*s)*Qh1^(-n*s)*T2^(1-s))^(1/(1-s));
                Qh1x = (e2*La1*wf1^(1-b*s)*Qh1^(-a*s)*Qh2^(-n*s)+e1*Lil*wh1
                    ^ (1-b*s)*Qf1^(-a*s)*Qf2^(-n*s)*T1^(1-s))^(1/(1-s));
                Qh2x = (e2*La2*wf2^(1-b*s)*Qh2^(-a*s)*Qh1^(-n*s)+e1*Li2*wh2
                    ^ (1-b*s)*Qf2^(-a*s)*Qf1^(-n*s)*T2^(1-s))^(1/(1-s));
                Qf1 = Qf1x;
                Qf2 = Qf2x;
                Qh1 = Qh1x;
                Qh2 = Qh2x;
            end; // for i (Q bestimmen)
        // E berechnen
        Eh1 = (m*(wh1*e1*Lil+wh2*e1*Li2))+((a*wh1*e1*Lil+n*wh2*e1*Li2)/
            b);
    end;

```

```

Eh2 = ((1-m)*(wh1*e1*Li1+wh2*e1*Li2))+((n*wh1*e1*Li1+a*wh2*e1*
    Li2)/b);
Ef1 = (m*(wf1*e2*La1+wf2*e2*La2))+((a*wf1*e2*La1+n*wf2*e2*La2)/
    b);
Ef2 = ((1-m)*(wf1*e2*La1+wf2*e2*La2))+((n*wf1*e2*La1+a*wf2*e2*
    La2)/b);
// w berechnen
wh1alt = wh1;
wh2alt = wh2;
wf1alt = wf1;
wf2alt = wf2;
wh1 = ((b*(Qf1^s*Eh1*Qh1+Qh1^s*Ef1*T1^(1-s)*Qf1))/(Qf1*Qh1))
    ^((1/(b*s))*Qf2^(-n/b)*Qf1^(-a/b));
wh2 = ((b*(Qf2^s*Eh2*Qh2+Qh2^s*Ef2*T2^(1-s)*Qf2))/(Qf2*Qh2))
    ^((1/(b*s))*Qf1^(-n/b)*Qf2^(-a/b));
wf1 = ((b*(Qh1^s*Ef1*Qf1+Qf1^s*Eh1*T1^(1-s)*Qh1))/(Qh1*Qf1))
    ^((1/(b*s))*Qh2^(-n/b)*Qh1^(-a/b));
wf2 = ((b*(Qh2^s*Ef2*Qf2+Qf2^s*Eh2*T2^(1-s)*Qh2))/(Qh2*Qf2))
    ^((1/(b*s))*Qh1^(-n/b)*Qh2^(-a/b));
if abs(wh1-wh1alt) < 0.001
    if abs(wh2-wh2alt) < 0.001 then
        break
    end;
end;
end; // for k
if abs(wh2-wh1) < 0.001 then
    break
end;
if real(wh1) > real(wh2) then Li = Li + 0.001;
else Li = Li - 0.001;
end; //if
end; //for j
r1(l) = La;
r2(l) = Li;
end; // for l
endfunction

```

Kurzfristiges Gleichgewicht:

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\CM");
getf("lib_cm-v2.sci");
L = linspace(0,1,1001)';
T1 = 2.15;
T2 = 2.15;

```

```

hh = myCMI2(T1,T2,0);
ff = myCMA2(T1,T2,0);
// Inlandskurven:
plot2d(L, hh, rect=[0,0,1,1], style=1);
// Auslandskurven:
plot2d(L, ff, rect=[0,0,1,1], style=1);

```

mit den Funktionen

```

//Inlandskurve im CM-Modell
function zz = myCMI2(T1, T2, e)
[k1, k2] = mySimulateCMI(T1, T2, e, 0);
m = size(k1);
s = m(1);
for i = 1 : s
    v1(i) = k1(s+1-i);
    v2(i) = k2(s+1-i);
end;
zz = interp(L, v2, v1, splin(v2,v1,"fast"));
endfunction

//Auslandskurve im CM-Modell
function zz = myCMA2(T1, T2, e)
[k1, k2] = mySimulateCMI(T1, T2, 0, e);
zz = interp(L, k1, k2, splin(k1,k2,"fast"));
endfunction

```

Bifurkationsdiagramm mit Ermittlung des "Knicks":

```

clear;
chdir("D:\Aktuell\Dissertation\Simulation\CM");
getf("lib_cm-v2.sci");
rts = [];
T0 = 2.15; // T1 von
T1 = 3.2; // T1 bis
T2 = 1.2; // T2 fest
e = 0.0; // Größenunterschied
L = linspace(0,1,101)';
T = linspace(T0,T1,500)';
for i = 1 : length(T)
    // Die folgende Zeile auskommentieren bei ungleichen
    // Transportkosten!
    T2 = T(i);
    // Inland:
    [k1, k2] = mySimulateCMI(T(i), T2, e, 0);

```

```

m = size(k1);
s = m(1);
for j = 1 : s
    v1(j) = k1(s+1-j);
    v2(j) = k2(s+1-j);
end;
dd = interp(L, v2, v1, splin(v2,v1,"fast"));
h0 = interp(0, k1, k2, splin(k1,k2,"fast"));
h1 = interp(1, k1, k2, splin(k1,k2,"fast"));
// Ausland:
[k1, k2] = mySimulateCMI(T(i), T2, 0, e);
ss = interp(L, k1, k2, splin(k1,k2,"fast"));
for j = 1 : s
    v1(j) = k1(s+1-j);
    v2(j) = k2(s+1-j);
end;
f0 = interp(0, v2, v1, splin(v2,v1,"fast"));
f1 = interp(1, v2, v1, splin(v2,v1,"fast"));
x0 = 1;
x1 = 0;
if (h0 > f0) then x0 = min(h0, 1); end;
if (h1 < f1) then x1 = max(h1, 0); end;
rts = [rts; T(i), x0];
rts = [rts; T(i), x1];
xx = dd-ss;
r = myRootsCM(L,dd, ss);
for j = 1 : length(r) rts = [rts; T(i), r(j)]; end;
end;
plot2d(rts(:,1), rts(:,2), style=-0, rect=[T0,-0.1,T1,1.1]);

```

Änderungen des Parameters μ , der die Ausgabenanteile angibt, erfolgt direkt in der Simulation oben. Bei Fehlern, aufgrund von Nichtstetigkeit der Vektoren in der Spline-Funktion sollten diese mit try-catch gekapselt werden.

B Daten zur Entwicklung in Deutschland

Bundesland	Lohnsatz 2006	BIP / Kopf 2007	BIP abs. 2007
Baden-Württemberg	16,72	33,26	357.575.250
Bayern	15,63	34,59	433.040.752
Berlin	15,66	24,81	84.757.401
Brandenburg	12,10	20,97	53.172.764
Bremen	18,05	40,36	26.765.178
Hamburg	17,95	48,61	86.062.266
Hessen	16,03	35,44	215.189.977
Mecklenburg-Vorpommern	11,49	20,71	34.781.865
Niedersachsen	16,46	26,00	207.273.702
Nordrhein-Westfalen	16,07	29,16	524.720.207
Rheinland-Pfalz	15,95	25,85	104.578.630
Saarland	16,89	29,23	30.297.202
Sachsen	11,21	21,98	92.746.789
Sachsen-Anhalt	11,65	21,40	51.616.690
Schleswig-Holstein	15,33	25,29	71.765.915
Thüringen	10,97	21,21	48.555.413

Tabelle A.2: Durchschnittlicher Brutto-Lohnsatz im produzierenden Gewerbe 2006 (in Euro), absolutes BIP und BIP pro Kopf 2007 (in Tausend Euro) auf Ebene der Bundesländer (Quelle: Statistisches Bundesamt)

	alte Länder	neue Länder
1991	1.987	924
1992	2.108	1.213
1993	2.188	1.425
1994	2.269	1.537
1995	2.358	1.652
1996	2.418	1.769
1997	2.460	1.819
1998	2.518	1.866
1999	2.589	1.917
2000	2.652	1.929
2001	2.718	1.982
2002	2.806	2.077
2003	2.889	2.141
2004	2.954	2.191
2005	3.009	2.239
2006	3.060	2.279
2007	3.134	2.344
2008	3.213	2.431

Table A.3: Durchschnittliche Bruttomonatsverdienste im produzierenden Gewerbe und im Dienstleistungsbereich ohne Sonderzahlungen (Quelle: Statistisches Bundesamt)

Bundesländer	1991	1992	1993	1994	1995	1996	1997	1998	1999
Baden-Württemberg	10.001.840	10.148.708	10.234.026	10.272.069	10.319.367	10.374.505	10.396.610	10.426.040	10.475.932
Bayern	11.595.970	11.770.257	11.863.313	11.921.944	11.993.484	12.043.869	12.066.375	12.086.548	12.154.967
Berlin	3.446.031	3.465.748	3.475.392	3.472.009	3.471.418	3.458.763	3.425.759	3.398.822	3.386.667
Brandenburg	2.542.723	2.542.651	2.537.661	2.536.747	2.542.042	2.554.441	2.573.291	2.590.375	2.601.207
Bremen	683.684	685.845	683.096	680.029	679.757	677.770	673.883	667.965	663.065
Hamburg	1.668.757	1.688.785	1.702.887	1.705.872	1.707.901	1.707.986	1.704.731	1.700.089	1.704.735
Hessen	5.837.330	5.922.639	5.967.305	5.980.693	6.009.913	6.027.284	6.031.705	6.035.137	6.051.966
Mecklenburg-Vorpommern	1.891.657	1.864.980	1.843.455	1.832.298	1.823.084	1.817.196	1.807.799	1.798.689	1.789.322
Niedersachsen	7.475.790	7.577.520	7.648.004	7.715.363	7.780.422	7.815.148	7.845.398	7.865.840	7.898.760
Nordrhein-Westfalen	17.509.866	17.679.166	17.759.300	17.816.079	17.893.045	17.947.715	17.974.487	17.975.516	17.999.800
Rheinland-Pfalz	3.821.235	3.880.965	3.925.863	3.951.573	3.977.919	4.000.567	4.017.828	4.024.969	4.030.773
Saarland	1.076.879	1.084.007	1.084.522	1.084.201	1.084.370	1.084.184	1.080.790	1.074.223	1.071.501
Sachsen	4.678.877	4.640.997	4.607.660	4.584.345	4.566.603	4.545.702	4.522.412	4.489.415	4.459.686
Sachsen-Anhalt	2.823.324	2.796.981	2.777.935	2.759.213	2.738.928	2.723.620	2.701.690	2.674.490	2.648.737
Schleswig-Holstein	2.648.532	2.679.575	2.694.875	2.708.392	2.725.461	2.742.293	2.756.473	2.766.057	2.777.275
Thüringen	2.572.069	2.545.808	2.532.799	2.517.776	2.503.785	2.491.119	2.478.148	2.462.836	2.449.082
alte Länder	62.319.883	63.117.467	63.563.191	63.836.215	64.171.639	64.421.321	64.548.280	64.622.384	64.828.774
neue Länder	14.508.650	14.391.417	14.299.510	14.230.379	14.174.442	14.132.078	14.083.340	14.015.805	13.948.034

Bundesländer	2000	2001	2002	2003	2004	2005	2006	2007	2008
Baden-Württemberg	10.524.415	10.600.906	10.661.320	10.692.556	10.717.419	10.735.701	10.738.753	10.749.755	10.749.506
Bayern	12.230.255	12.329.714	12.387.351	12.423.386	12.443.893	12.468.726	12.492.658	12.520.332	12.519.728
Berlin	3.382.169	3.388.434	3.392.425	3.388.477	3.387.828	3.395.189	3.404.037	3.416.255	3.431.675
Brandenburg	2.601.962	2.593.040	2.582.379	2.574.521	2.567.704	2.559.483	2.547.772	2.535.737	2.522.493
Bremen	660.225	659.651	662.098	663.129	663.213	663.467	663.979	663.082	661.866
Hamburg	1.715.392	1.726.363	1.728.806	1.734.083	1.734.830	1.743.627	1.754.182	1.770.629	1.772.100
Hessen	6.068.129	6.077.826	6.091.618	6.089.428	6.097.765	6.092.354	6.075.359	6.072.555	6.064.953
Mecklenburg-Vorpommern	1.775.703	1.759.877	1.744.624	1.732.226	1.719.653	1.707.266	1.693.754	1.679.682	1.664.356
Niedersachsen	7.926.193	7.956.416	7.980.472	7.993.415	8.000.909	7.993.946	7.982.685	7.971.684	7.947.244
Nordrhein-Westfalen	18.009.865	18.052.092	18.076.355	18.079.686	18.075.352	18.058.105	18.028.745	17.996.621	17.933.064
Rheinland-Pfalz	4.034.557	4.049.066	4.057.727	4.058.682	4.061.105	4.058.843	4.052.860	4.045.643	4.028.351
Saarland	1.068.703	1.066.470	1.064.988	1.061.376	1.056.417	1.050.293	1.043.167	1.036.598	1.030.324
Sachsen	4.425.581	4.384.192	4.349.059	4.321.437	4.296.284	4.273.754	4.249.774	4.220.200	4.192.801
Sachsen-Anhalt	2.615.375	2.580.626	2.548.911	2.522.941	2.494.437	2.469.716	2.441.787	2.412.472	2.381.872
Schleswig-Holstein	2.789.761	2.804.249	2.816.507	2.823.171	2.828.760	2.832.950	2.834.254	2.837.373	2.834.260
Thüringen	2.431.255	2.411.387	2.392.040	2.373.157	2.355.280	2.334.575	2.311.140	2.289.219	2.267.763
alte Länder	65.027.495	65.322.753	65.527.242	65.618.912	65.679.663	65.698.012	65.666.642	65.664.272	65.541.396
neue Länder	13.849.876	13.729.122	13.617.013	13.524.282	13.433.358	13.344.794	13.244.227	13.137.310	13.029.285

Tabelle A.4: Bevölkerungsentwicklung nach Bundesländern (Quelle: Statistisches Bundesamt)

Die Neue Ökonomische Geographie (NEG) erklärt Agglomerationen aus einem mikroökonomischen Totalmodell heraus. Zur Vereinfachung werden verschiedene Symmetrieanahmen getätigt. So wird davon ausgegangen, dass die betrachteten Regionen die gleiche Größe haben, die Ausgabenanteile für verschiedene Gütergruppen identisch sind und die Transportkosten für alle Industrieprodukte die selben sind.

Eine Folge dieser Annahmen ist es, dass zwar erklärt werden kann, unter welchen Bedingungen es zur Agglomerationsbildung kommt, nicht aber wo dies geschieht. In dieser Arbeit werden drei Standardmodelle der NEG um verschiedene Asymmetrien erweitert und die Veränderung der Ergebnisse im Vergleich zum jeweiligen Basismodell dargestellt. Dabei wird neben der Theorie auf die Methoden der Simulation eingegangen, die sich grundsätzlich auf andere Modelle übertragen lassen.

Darauf aufbauend wird eine asymmetrische Modellvariante auf die wirtschaftliche Entwicklung Deutschlands angewandt. So lässt sich das Ausbleiben eines flächendeckenden Aufschwungs in den neuen Ländern, die starken Wanderungsbewegungen in die alten Länder und das dauerhafte Lohnsatzgefälle in einem Totalmodell erklären.

