

I. Daten und Fakten

2. Methoden der empirischen Regionalanalyse

Joachim Möller

2.1 Einleitung

Die empirische Regionalanalyse gewinnt vor allem aus drei Gründen an Bedeutung. Erstens trägt hierzu die Wiederentdeckung der Raumdimension in der modernen Volkswirtschaftslehre bei, die etwa in der Neuen Ökonomischen Geographie (siehe *Kapitel 6*) zum Ausdruck kommt. Zweitens werden immer mehr Datensätze verfügbar, die die explizite Untersuchung der Raumdimension gestatten. Und schließlich eröffnet die Regionalisierung von empirischen Untersuchungen neue Möglichkeiten. Während beispielsweise zur Untersuchung der Beziehungen zwischen Konsum und verfügbarem Einkommen auf gesamtwirtschaftlicher Ebene nur jeweils eine Zeitreihe zur Verfügung steht, kann zur Analyse dieses Zusammenhangs bei Verwendung von Datensätzen etwa auf Bundesländerebene auf jeweils 16 Zeitreihen zurückgegriffen werden. Die zusätzliche Streuung der Beobachtungswerte in der räumlichen Dimension führt zu einer Zunahme des Informationsgehalts und damit letztlich zu einer besseren Absicherung der Schätzergebnisse.

Wissenschaftsgeschichtlich ist interessant, dass sich die empirische Regionalforschung einerseits der Verfahren bedient, die auch ohne Raumbezug eingesetzt werden (z. B. die Regressionsanalyse). Andererseits hat die räumliche Gliederung der Daten und die Notwendigkeit, den Einfluss von Lage und Distanz zu berücksichtigen, eigene methodische Entwicklungen innerhalb der Regionalökonomie hervorgebracht. Zu nennen sind hier beispielsweise das Gravitationsmodell, die räumlichen Konzentrations- und Spezialisierungsmaße sowie die in den letzten Jahren stark an Bedeutung gewinnende räumliche Ökonometrie.

Dass dieses Kapitel nicht eine umfassende Darstellung aller Methoden der empirischen Regionalforschung geben kann, versteht sich aus Platzgründen von selbst. Ausgewählt werden einige der bekanntesten Ansätze. Begonnen wird mit dem Gravitationsmodell und es wird dann die Shift-Share-Analyse vorgestellt. Ein weiterer Abschnitt ist den räumlichen Konzentrations- und Spezialisierungsmaßen gewidmet. Am Schluss des Kapitels wird noch auf die räumliche Ökonometrie eingegangen.

2.2 Das Gravitationsmodell

2.2.1 Grundlagen

Eines der erfolgreichsten regionalökonomischen Modelle ist das Gravitationsmodell. Für dieses Modell wurde eine Anleihe aus der (Astro-)Physik gemacht. Im Jahr 1686 erkannte Isaac Newton, dass die Planetenbahnen ebenso wie ein vom Baum fallender Apfel auf die gleiche Weise beschrieben werden können, nämlich durch das allgemeine Gravitationsgesetz. Dieses Gesetz besagt, dass sich zwei Körper i und j mit einer Kraft anziehen, die proportional ist zum Produkt ihrer Massen und m_j sowie umgekehrt proportional zum Quadrat ihres Abstandes. Es sei die Anziehungskraft mit a_{ij} bezeichnet, die Masse der beiden Körper mit m_i und m_j sowie der Abstand mit d_{ij} . Das Gesetz von Newton lautet also

$$a_{ij} = \gamma m_i m_j d_{ij}^{-2},$$

wobei der Proportionalitätsfaktor γ eine Naturkonstante ist.

Wie lässt sich dieses Modell auf regionalökonomische Fragestellungen übertragen? Die Analogie liegt auf der Hand. Auch in der Alltagssprache reden wir davon, dass z. B. eine Stadt, ein Einkaufszentrum oder ein touristisch interessanter Ort eine „Anziehungskraft“ entfaltet. Gemeint ist damit, dass diese Destinationen eine große Anzahl von Besuchern anlocken. Um den Bezug zum Gravitationsmodell herzustellen, ist eine genauere Betrachtung notwendig.

Zunächst lässt sich die Gesamtzahl der Besucher in Besucherströme zerlegen, die jeweils durch einen Ausgangs- und einen Zielort gekennzeichnet sind, sowie durch eine sich auf einen Zeitraum beziehende Maßzahl wie z. B. Personen pro Tag. Es bietet sich an, die Anziehungskraft mit Hilfe dieser Stromgrößen zu messen. Ein Besucherstrom, der im Ort i seinen Ausgang nimmt und im Ort j endet, sei mit t_{ij} bezeichnet. Wie in der Physik geht es nun darum, die durch t_{ij} erfasste Anziehungskraft zwischen i und j zu erklären. Beim Newtonschen Gravitationsgesetz werden dafür die Massen und die Distanz herangezogen. Letzteres erscheint zunächst einfach übertragen werden zu können, da sich die Distanz auch im regionalökonomischen Kontext leicht bestimmen lässt (z. B. als Entfernung in Kilometern). Die Luftlinie zwischen zwei Orten ist aber bei genauer Betrachtung vielfach kein ideales Maß. So können zwei Orte, die durch einen Fluss getrennt sind, zwar eine geringe Luftliniendistanz aufweisen, die tatsächliche Entfernung kann aber sehr hoch sein, wenn die nächstgelegene Brücke 10 Kilometer flussauf- oder -abwärts gelegen ist. Weiterhin kann die Überwindung einer bestimmten Entfernung auf einem gut ausgebauten Verkehrsweg mit geringeren Kosten (einschließlich der Zeitkosten und Unannehmlichkeiten der Reise) verbunden sein, als die Fahrt auf einer viel kürzeren, aber schlecht ausgebauten Strecke. Anstelle der Luftliniendistanz ist deshalb der *ökonomische Entfernungsbegriff* zu bevorzugen, der die angesprochenen Aspekte berücksichtigt. Damit wird versucht, die monetären und nicht-monetären Kosten der Entfernungsüberbrückung zu erfassen, etwa durch die Opportunitätskosten des Zeitverlustes oder die Unannehmlichkeiten der Fahrt.

Die nächste Frage ist, was dem physikalischen Massenbegriff in der Regionalökonomie entspricht. Die gesuchte Größe muss mit der „Bedeutung“ von Ursprungs- und

Zielort zu tun haben. Dabei lassen sich im Unterschied zur Physik auch unterschiedliche Maße von Ausgangspunkt und Destination wählen. Denkbar sind z. B. für den Ursprungsort Größen wie die Einwohnerzahl in Personen, die Einwohner einer bestimmten Alters- oder Zielgruppe, die Kaufkraft dieser Gruppe usw. Für den Zielort kommt z. B. die Zahl der Geschäfte oder deren Geschäftsfläche, die Zahl der Arbeitsplätze, der öffentlichen Einrichtungen, der Sehenswürdigkeiten usw. in Frage. Welcher Indikator jeweils der sinnvollste ist, hängt neben der Verfügbarkeit der Daten von der gewählten Fragestellung ab. Sollen etwa Pendlerströme, also Fahrten zwischen Wohn- und Arbeitsort, analysiert werden, so wäre die Bevölkerung im erwerbsfähigen Alter ein plausibler Indikator für die Bedeutung des Ursprungsortes, während die Bedeutung des Zielortes sinnvoll durch die Zahl Arbeitsplätze gemessen wird. Geht es um Warenströme, so wäre ein Indikator für die Wirtschaftskraft von Ursprungs- und Zielort wie das Bruttoinlandsprodukt geeigneter. Bei der Erklärung von Kaufkraftströmen beim Möbelerwerb könnte für den Ausgangsort das verfügbare Einkommen der Einwohner und für den Zielort die gesamte Geschäftsfläche der Möbelmärkte herangezogen werden.

Im Folgenden soll zugelassen werden, dass es nicht nur mehrere Ursprungsorte gibt, sondern auch alternative Zielorte. Um das Gravitationsmodell der Regionalökonomie formal zu beschreiben, seien die Maße zur Charakterisierung der Bedeutung von Ursprungs- bzw. Zielort mit u_i und z_j bezeichnet. In genauer Entsprechung zum Newtonschen Gravitationsgesetz lässt sich dann

$$t_{ij} = \kappa u_i z_j d_{ij}^{-2} \quad (1)$$

formulieren, wobei d_{ij} die (ökonomische) Distanz zwischen i und j bezeichnet und κ für einen Proportionalitätsfaktor steht.

Wird unterstellt, dass die gesamten Ströme zwischen m Ursprungs- und n Zielorten als Datenmaterial zur Verfügung stehen, dann ergibt sich eine weitere Möglichkeit: Als Maß für die Bedeutung des Ursprungsortes i kann das Gesamtvolumen der dort entstehenden Ströme eingesetzt werden, also die Summe aller n von i ausgehenden Ströme

$$u_i = \sum_{j=1}^n t_{ij}. \quad (2)$$

Entsprechend lässt sich für die Bedeutung des Zielortes die Summe der dort eintreffenden Ströme verwenden,

$$z_j = \sum_{i=1}^m t_{ij}. \quad (3)$$

Nehmen wir ein Beispiel. Sie wollen die Bahnreisen zwischen vier Großstädten Bayerns (Augsburg, München, Nürnberg, Regensburg) analysieren. Dazu soll ein Gravitationsmodell verwendet werden. Die Menge der Ursprungs- bzw. Zielorte ist $\{A, M, N, R\}$. Die Größe t_{ij} wäre demnach die Zahl der Bahnreisen pro Zeitperiode von i nach j . Beispielsweise entspricht t_{24} der Zahl der Reisenden von München nach Regensburg, t_{42} der der Reisenden in umgekehrter Richtung.

Dem einfachen Gravitationsmodell zufolge müsste die Zahl der Reisenden von Regensburg nach München t_{42} pro Zeitperiode näherungsweise proportional sein

- zur Zahl aller Reisenden, die pro Zeitperiode mit Zielort A , M oder N in R abfahren (u_4),
- zur Zahl aller Bahnreisenden aus A , N oder R , die pro Zeitperiode in M ankommen (z_2),
- zum Kehrwert der quadrierten Entfernung zwischen R und M , also zu $d_{42}^{-2} = d_{44}^{-2}$.

Die direkte Übertragung des Newtonschen Gravitationsmodells auf regionalökonomische Fragestellungen ist jedoch in der Literatur kritisiert worden. So besitzt der Erklärungsansatz von Gleichung (1) offenbar unrealistische Implikationen: Eine Verdoppelung aller u_i und z_j würde aufgrund der Modellstruktur zu einer Vervielfachung der t_{ij} führen, was unplausibel erscheint, da eher eine Verdoppelung erwartet werden sollte. Weiterhin ist in Zweifel zu ziehen, dass die Interaktionen zwischen zwei Orten gerade mit dem Quadrat der Entfernung abnehmen. Die Übernahme dieser astrophysikalischen Gesetzmäßigkeit auf regionalökonomische Zusammenhänge erscheint unbegründet. Insgesamt kann aus dieser Kritik der Schluss gezogen werden, dass das Newtonsche Gravitationsmodell für die Beschreibung räumlicher Interaktionen zu restriktiv ist. Um den berechtigten Einwänden Rechnung zu tragen, ohne den Grundgedanken der Gravitationstheorie aufzugeben, lässt sich das einfache Modell durch die folgende allgemeinere Formulierung ersetzen:

$$t_{ij} = \kappa u_i^\alpha z_j^\beta d_{ij}^\gamma, \quad (4)$$

wobei $\alpha, \beta > 0$ und $\gamma < 0$ Koeffizienten sind, die neben der Konstante κ mittels empirischer Verfahren aus den Daten bestimmt werden müssen. Noch genereller ist der Ansatz

$$t_{ij} = \kappa u_i^\alpha z_j^\beta f(d_{ij}),$$

in dem die Funktion $f(d_{ij})$ mit $f' < 0$ sicherstellt, dass ceteris paribus die Interaktionen zwischen zwei Orten im Raum mit zunehmender Entfernung abnehmen. Diese Funktion wird deshalb auch als *Abschreckungsfunktion* der Entfernung (distance deterrence function) bezeichnet. Als Alternative zur Festlegung $f(d_{ij}) = d_{ij}^\gamma$ wie in Gleichung (4) eignet sich beispielsweise die Exponentialfunktion. In diesem Fall lautet die Abschreckungsfunktion der Entfernung $f(d_{ij}) = e^{\gamma d_{ij}}$, mit $\gamma < 0$.

2.2.2 Transformation des Modells in einen Regressionsansatz

Um mit einem Ansatz wie in Gleichung (4) konkret arbeiten zu können – beispielsweise um Prognosen über zukünftige Verkehrsströme zu erstellen – ist es notwendig, die unbekannt Parameter α, β, γ und κ mit geeigneten empirischen Methoden zu bestimmen. Ein möglicher Weg dazu ist die Regressionsanalyse oder gewöhnliche *Kleinst-Quadrate-Methode* (method of ordinary least squares, OLS). Die OLS-Methode setzt eine in den Parametern lineare Schätzgleichung voraus. Das Gravitationsmodell nach Gleichung (4) erfüllt diese Voraussetzung zunächst nicht. Durch logarithmische Transformation lässt sich das Modell jedoch auf eine einfache Weise linearisieren,

$$\log t_{ij} = \log \kappa + \alpha \log u_i + \beta \log z_j + \gamma \log d_{ij}.$$

Nach Hinzufügen eines additiven stochastischen Störterms lautet die Schätzgleichung dann

$$\log t_{ij} = c + \alpha \log u_i + \beta \log z_j + \gamma \log d_{ij} + \varepsilon_{ij}, \quad (5)$$

mit $c = \log \kappa$. Für die empirische Analyse werden Daten für die Transportströme t_{ij} sowie für die u_i und z_j , also die Maße der Bedeutung von Ursprungs- und Zielorten, benötigt. Erforderlich sind ferner Informationen über die Entfernung oder die Fahrtzeiten zwischen den Orten (d_{ij}). Liegen diese Daten für eine oder besser für mehrere Zeitperioden vor, so lassen sich aus der mit Hilfe der OLS-Methode geschätzten Gleichung (5) Schätzwerte für die unbekannt Parameter c, α, β sowie γ gewinnen. Durch Einsetzen der Schätzwerte und Entlogarithmieren können dann die durch das Modell prognostizierten Transportströme leicht bestimmt werden als

$$\hat{t}_{ij} = \exp\left(\hat{c} + \hat{\alpha} \log u_i + \hat{\beta} \log z_j - \hat{\gamma} \log d_{ij}\right), \quad (6)$$

wobei die geschätzten Parameter durch ein Dach gekennzeichnet sind.

Eine Anwendung dieses Ansatzes könnte wie folgt aussehen: Angenommen, es veränderten sich durch den Bau einer neuen Straße oder einer neuen Schienentrasse die Distanzen bzw. die Fahrtzeiten zwischen den Orten im Raum. Mit Hilfe von Gleichung (6) wird es möglich, die sich daraus voraussichtlich ergebenden Auswirkungen auf die Transportströme zwischen den Orten zu berechnen.

Auf eine Komplikation bei Anwendung des skizzierten Verfahrens sollte allerdings hingewiesen werden: Die prognostizierten Schätzwerte aus Gleichung (6) verletzen im Allgemeinen die Bedingungen der Gleichung (2) und (3). Möchte man diese Inkonsistenz vermeiden, so ist ein aufwändigeres Schätzverfahren heranzuziehen, das die Einhaltung der Restriktionen erzwingt.

2.2.3 Dazwischen liegende Gelegenheiten (*intervening opportunities*)

Bei der Herleitung des Gravitationsmodells wurde bisher implizit unterstellt, dass zwischen den Orten A und B eine unbesiedelte Fläche liegt, die als Zielort außer Betracht bleibt, weil sie keinerlei ökonomische Möglichkeiten bietet. In der Realität ist diese Annahme natürlich in vielen Fällen verletzt. Die Vernachlässigung einer attraktiven Destination zwischen A und B kann aber zu gravierenden Verfälschungen der Analyse führen. Diese Erkenntnis ist in der Regionalökonomie spätestens seit der Arbeit von Samuel A. Stouffer (1940) bekannt. Stouffer hatte eine Art „Gesetz“ formuliert, das in seiner Anwendung auf einen Strom von Reisenden wie folgt dargelegt werden könnte: Die Zahl der Reisen zwischen einer Ausgangszone und einer Zielzone ist proportional zu der Anzahl der Gelegenheiten in der Zielzone und umgekehrt proportional zu der Anzahl der dazwischen liegenden Gelegenheiten.

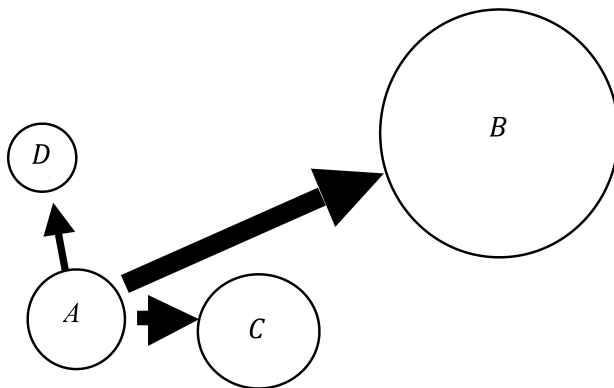
Wenn etwa Einkaufsfahrten betrachtet werden (z. B. zum Zweck eines Möbelkaufs), dann ist Stouffer zufolge die Zahl der Fahrten von A nach B proportional zu der Zahl oder der Verkaufsfläche der Möbelhäuser in B und umgekehrt proportional zu der Zahl oder der Verkaufsfläche der Möbelhäuser in A und der in allen Orten, die zwischen A und B liegen.

Stouffers Ansatz, der etwa in der berühmten *Chicago Area Transportation Study* (1960) Anwendung gefunden hat, ist im Kerngedanken zutreffend, bedarf jedoch aus heutiger Sicht einiger Modifikationen. Zunächst ist ebenso wie bei der originären Form des

Gravitationsmodells die postulierte funktionale Form des Zusammenhangs in Zweifel zu ziehen. Die von Stouffer behauptete Proportionalität ist ökonomisch unbegründet und sollte deshalb einer allgemeineren Formulierung weichen. Weiterhin erscheint Stouffers Ansatz auf einen eindimensionalen Raum zugeschnitten zu sein und bedarf bei der konkreten Anwendung in der Raumwirtschaft einer Verallgemeinerung. Um im Beispiel zu bleiben: Wird unterstellt, dass B genau nördlich von A liegt und die Entfernung 50 km beträgt, so wird auch ein Möbelhaus am Ort C, 20 km südlich von A gelegen, im verallgemeinerten Sinn eine „dazwischen liegende Gelegenheit“ sein. Dies würde auch auf alle Möbelhäuser zutreffen, die nicht in B gelegen sind, aber in einem Umkreis mit dem Radius von 50 km um A anzutreffen sind.

Betrachten wir als Beispiel eine Anwendung des Intervening-Opportunities-Modells auf den Pendlerverkehr. Als Pendlerverkehr bezeichnet man die arbeitstäglichen Fahrten zwischen Wohn- und Arbeitsort. Die Analyse der Pendlerströme ist z. B. für die Raum- und Verkehrsplanung von erheblicher Bedeutung. Eine wichtige Aufgabe der empirischen Regionalökonomie besteht darin, diese Ströme zu beschreiben, zu erklären und – bei Änderungen der Gegebenheiten – die sich daraus ergebenden Effekte zu prognostizieren. Die Situation stelle sich wie folgt dar:

Abbildung 2.1: Austauschbeziehungen zwischen den Orten A und B mit intervening opportunities in C und D



Die Modellregion besteht aus vier Städten, A bis D, wobei sich die Städte C und D in gleicher Entfernung zu A befinden sollen (Abbildung 2.1). Durch die unterschiedlichen Durchmesser soll die unterschiedliche Größe der Städte dargestellt werden. Bei der Untersuchung der Pendlerströme von Ort A nach Ort B wäre es nicht korrekt, die Orte C und D zu vernachlässigen, da sie ebenfalls mögliche Zielorte für die Auspendler aus A sind. Angenommen sei, dass die Größe der Städte die Zahl der Arbeitsmöglichkeiten widerspiegelt. Der Pendlerverkehr zwischen A und C wird dann den zwischen A und D übersteigen, da C bei gleicher Erreichbarkeit mehr Arbeitsmöglichkeiten für die Auspendler aus A bietet. Die Stadt B ist schlechter als C und D von A aus zu erreichen, ist aber wegen ihrer Größe attraktiv. Zweifellos wird die Existenz der dazwischen liegenden Gelegenheit in D und insbesondere in C den Pendlerverkehr zwischen A und B reduzieren.

Welche Konsequenzen besitzen Intervening Opportunities für die empirische Analyse? Eine Möglichkeit besteht darin, eine zunächst auf die Orte A und B beschränkte Untersuchung mit Hilfe des Gravitationsansatzes auf alle relevanten Orte des Untersuchungsraums auszudehnen. Falls die Zahl der Orte aber sehr groß ist, so könnte dies bei der praktischen Anwendung Probleme schaffen. Als Ausweg bietet sich an, die dazwischen liegenden Gelegenheiten zu gruppieren (z. B. nach solchen in einer Nah-, Mittel- und Fernzone) und sie dann in den Schätzansatz (Gleichung 5) mit aufzunehmen.

2.3 Die Shift-Share-Analyse

2.3.1 Grundlagen

Die *Shift-Share-Analyse* ist eine Methode, die in der Regionalökonomie eingesetzt wird, um die Entwicklung in einem betrachteten Wirtschaftsraum im Verhältnis zu einem Referenzraum zu charakterisieren. Typischerweise handelt es sich bei dem betrachteten Wirtschaftsraum um eine untergeordnete, bei dem Referenzraum hingegen um eine übergeordnete Gebietseinheit (z. B. Bundesland zu Bund, Regierungsbezirk zu Land oder Bund). Die Grundidee der Shift-Share-Analyse ist, dass sich die Einflüsse auf die ökonomische Entwicklung einer Region zum einen auf ihre Wirtschaftsstruktur, zum anderen auf die dort vorherrschenden Standortbedingungen zurückführen lassen. Beispielsweise könnte eine boomende Region sich deshalb im Vergleich zu anderen Wirtschaftsräumen gut entwickeln, weil besonders viele Unternehmen vorhanden sind, deren Geschäftsfelder in Wachstumsbranchen liegen. Möglich wäre aber auch, dass die Unternehmen vor Ort deshalb expandieren, weil sie auf eine besonders gute Infrastruktur zurückgreifen können bzw. aufgrund anderer günstiger Standortbedingungen Kostenvorteile gegenüber Wettbewerbern aus anderen Regionen besitzen. Kostenrelevante Faktoren können unter anderem geringe Lohnstückkosten und Bodenpreise, günstiger Zugang zu Ressourcen oder niedrige Steuern und Abgaben sein. Auch die so genannten „weichen“ Faktoren, die die Region für hoch spezialisierte Fachkräfte und innovative Unternehmer attraktiv machen, spielen hierbei möglicherweise eine wichtige Rolle. So ist es durchaus denkbar, dass sich aufgrund günstiger Standortfaktoren Wirtschaftszweige in einer Region positiv entwickeln, die anderswo zu den Krisenbranchen gerechnet werden.

Die Grundfragen, die die von Edgar S. Dunn (1960) und anderen entwickelte Shift-Share-Analyse zu beantworten sucht, sind:

- Wie lassen sich überhaupt die Bestimmungsfaktoren für die regionale Entwicklung in einen Struktur- und einen Standorteffect aufspalten?
- Sind Unterschiede im Wirtschaftswachstum von Regionen eher durch Struktur- oder Standorteffecte bestimmt?
- Können Informationen über Struktur- und Standorteffecte ausgenutzt werden, um regionale Entwicklungen vorherzusagen?

Den ersten beiden Fragen soll im Folgenden nachgegangen werden. Die Analyse der Daten für die Kreise und kreisfreien Städte Deutschlands zeigt zunächst, dass sich die Regionen in ihrer Wirtschaftsstruktur beträchtlich unterscheiden. Gemessen an den sozialversicherungspflichtig Beschäftigten im Jahr 2000 differiert der Anteil des Verar-

beitenden Gewerbes an der Gesamtwirtschaft zwischen 11,8% in Potsdam und 72,9% in Dingolfing/Landau, dem Standort eines großen Automobilwerkes. Spiegelbildlich dazu sind in Potsdam 87,8% der Beschäftigten im Dienstleistungssektor tätig, in Dingolfing/Landau hingegen nur 26,2%. Betrachtet man generell die Branchenstruktur auf kleinräumiger Ebene genauer, so lassen sich jeweils spezifische Schwerpunkte der regionalen Wirtschaft bestimmen (wie z. B. Maschinenbau, Chemische Industrie, Fahrzeugbau, Bankdienstleistungen, Medien usw.).

Das Prinzip der Shift-Share-Analyse besteht in der Komponentenzerlegung des *Regionalfaktors* in verschiedene Teileffekte. Der Regionalfaktor ist dabei nichts anderes als das beobachtete Wachstum einer Indikatorgröße (z. B. der Wertschöpfung oder der Beschäftigung) in einer Region. In der einfachsten Form lautet die Grundgleichung der Shift-Share-Analyse:

$$\text{Regionalfaktor} = \text{Strukturfaktor} + \text{Standortfaktor}.$$

Der *Strukturfaktor* gibt dabei die Antwort auf die fiktive Frage: „Wie wäre die Wirtschaft einer ausgewählten Region R in der betrachteten Zeitperiode gewachsen, wenn sich alle dort vorhandenen Sektoren bzw. Industriezweige genau so entwickelt hätten wie im übergeordneten Wirtschaftsraum A ?“. Der Strukturfaktor lässt sich leicht aus einem gewichteten arithmetischen Mittel der sektoralen Wachstumsraten im Referenzraum A berechnen. Als Gewichtungsfaktoren finden dabei die Sektor- bzw. Branchenanteile der Region R im Ausgangszeitpunkt Verwendung. Der *Standortfaktor* ergibt sich dann als Abweichung des Regionalfaktors vom Strukturfaktor. Der Standortfaktor wird somit als eine Restgröße bestimmt.

Ein positiver (negativer) Standortfaktor besagt, dass die Wirtschaftszweige in R im gewichteten Mittel eine bessere (ungünstigere) Wachstumsbilanz aufweisen als die entsprechenden Wirtschaftszweige in A . Ist der Strukturfaktor größer (kleiner) als die Wachstumsrate in A , dann sind in der Region R Wachstumsbranchen (Problembranchen) stärker vertreten als im Referenzraum A . Wenn die Wirtschaftsstruktur der Region ein genaues Abbild des Referenzraums ist, entspricht der Strukturfaktor genau dem Wachstum in A . Der Standortfaktor ist folglich Null, wenn sich die Wirtschaftszweige der Region genauso entwickeln wie in A .

Betrachten wir als Beispiel eine Shift-Share-Zerlegung der Erwerbstätigenentwicklung auf der Ebene der Bundesländer im Zehnjahreszeitraum 1994 bis 2003. Als Erwerbstätige zählen grundsätzlich alle Personen, die in einem Arbeitsverhältnis stehen, die selbstständig ein Gewerbe oder eine Landwirtschaft betreiben oder einen freien Beruf ausüben, einschließlich mithelfender Familienangehöriger, sowie Soldaten und Zivildienstleistende. Die verwendete Statistik des Statistischen Bundesamtes ist nach Wirtschaftsbereichen gegliedert und beruht auf dem Arbeitsortprinzip. Nach diesem Konzept werden Personen, die im Bundesland A leben und in Bundesland B arbeiten, also aus der Sicht von B Einpendler sind, in der Statistik von B mit erfasst. Analoges gilt für im Ausland lebende Personen, die im Inland beschäftigt sind.

Die *Tabelle 2.1* zeigt die Entwicklung der Erwerbstätigkeit in Deutschland differenziert nach den Sektoren der Wirtschaft. Der erste (oder Primäre) Sektor fasst die Bereiche Land- und Forstwirtschaft, Fischerei, Bergbau und Energiegewinnung zusammen. Dem zweiten (oder Sekundären) Sektor sind Erwerbstätige zugeordnet, die z. B. im Maschinen- oder Automobilbau, der Chemischen Industrie oder dem Hoch- oder

Tiefbau beschäftigt sind. Diese Wirtschaftszweige werden unter der Bezeichnung „Produzierendes Gewerbe“ in der Statistik geführt. Der dritte (oder Tertiäre) Sektor schließlich ist der Dienstleistungsbereich. Dem Tertiären Sektor werden Gebäudereiniger, Friseure, Verkäufer und Kellner genauso zugerechnet wie Personen, die in Ausbildungseinrichtungen, im Gesundheitswesen, in der öffentlichen Verwaltung, bei Banken und Versicherungen oder in Ingenieurbüros arbeiten. Die *Tabelle 2.1* zeigt nun sehr deutlich ein Phänomen, das in anderen hochentwickelten Ländern in ähnlicher Weise beobachtet wird: Der Primäre und Sekundäre Sektor verlieren im Zeitraum 1994–2003 mit –20,8% bzw. –15,9% in erheblichem Umfang Erwerbstätige, während der Dienstleistungssektor stark dazu gewinnt (+13,3%). Dieser Beschäftigungsaufbau im Dienstleistungsbereich ist offenbar in der Lage, die schrumpfenden Erwerbstätigenzahlen in den beiden ersten Bereichen aufzufangen. Denn insgesamt ergibt sich im Beobachtungszeitraum sogar ein Anstieg der Erwerbstätigen von 37,3 auf 38,2 Millionen, was einem relativen Zuwachs von 2,5% entspricht.

Tabelle 2.1: Erwerbstätige in Deutschland 1994 und 2003

Sektor	Erwerbstätige Personen				
	1994		2003		Veränd. Rate
	in Tsd.	Anteil	in Tsd.	Anteil	
1: Land- und Forstwirtschaft, Fischerei	1 172	3,1%	928	2,4%	–20,8%
2: Produzierendes Gewerbe	12 394	33,2%	10 418	27,2%	–15,9%
3: Dienstleistungsbereiche	23 738	63,6%	26 900	70,3%	13,3%
Summe	37 304		38 246		2,5%

Quelle: Statistisches Bundesamt.

Da Primärer und Sekundärer Sektor absolut und relativ rückläufige Erwerbstätigenzahlen aufweisen, liegt die Überlegung nahe, dass Bundesländer mit einem im Jahr 1994 hohen Anteil dieser Sektoren (und spiegelbildlich einem geringen Anteil des Dienstleistungssektors) eine im Hinblick auf die Wirtschaftsstruktur ungünstige Ausgangslage besaßen.

Die *Tabelle 2.2* enthält eine Shift-Share-Analyse anhand der Erwerbstätigen-Daten für Baden-Württemberg und Hamburg. Beide Bundesländer stehen für Extreme: Der Dienstleistungsanteil liegt im Stadtstaat Hamburg im Ausgangsjahr bei 78,4%, im Flächenland Baden-Württemberg hingegen nur bei 57,7%, was dem niedrigsten Niveau aller Bundesländer entspricht. Dafür weist das „Ländle“ mit fast 40% von allen Bundesländern den höchsten Anteil des Sekundären Sektors auf. Da sich auf aggregierter Ebene die Erwerbstätigkeit im Dienstleistungssektor positiv, die des Sekundären Sektors hingegen im Trend negativ entwickelt, lässt sich auf einen gravierenden Strukturnachteil Baden-Württembergs und eine günstige Ausgangslage Hamburgs schließen. Diese Erwartung drückt der Strukturfaktor in Spalte (8) aus. Wenn sich die Erwerbstätigkeit auf sektoraler Ebene in den beiden Bundesländern wie im Bund entwickelt hätte, wäre in Hamburg ein den Bundesdurchschnitt von 2,5% weit übertreffendes Wachstum von 7,0%, in Baden-Württemberg hingegen nur eines von 0,8% realisiert worden. Tatsächlich ist aber die Erwerbstätigkeit in Baden-Württemberg mit +6,5% so

stark angestiegen wie in keinem anderen deutschen Bundesland, während in Hamburg mit +0,8% eine mehr oder weniger als stagnierend zu bezeichnende Entwicklung der Erwerbstätigkeit zu beobachten war. Trotz des vermeintlichen Strukturvorteils verlief die Entwicklung in Hamburg somit deutlich schlechter als im Bundesdurchschnitt.

Der Shift-Share-Analyse zufolge ist dieses Ergebnis wie folgt zu interpretieren: Die guten Standortbedingungen in Baden-Württemberg haben den Nachteil der eigentlich ungünstigen Wirtschaftsstruktur überkompensiert. Die *Tabelle 2.2* zeigt in Spalte 9, dass dies sowohl im Sekundären als auch im Tertiären Sektor der Fall war. Die Schrumpfung des Sekundären (Industrie-) Sektors ist in diesem Bundesland somit vergleichsweise gering gewesen, während sich der Dienstleistungssektor überdurchschnittlich gut entwickelt hat. In Hamburg hingegen vollzog sich die Schrumpfung des bereits in der Ausgangslage unterproportional vertretenen Sekundären Sektors noch schneller als im Bund. Der in der Hansestadt stark vertretene Tertiäre Sektor konnte hingegen nicht in der Weise zulegen wie dies im übergeordneten Wirtschaftsraum, d. h. hier im Bund, der Fall war. Insbesondere die Schwäche im Dienstleistungsbereich ist Spalte 9 in der *Tabelle 2.2* zufolge für die unter dem Strich schwache Wachstumsbilanz des Bundeslands Hamburg verantwortlich.

Tabelle 2.2: Shift-Share-Analyse für die Erwerbstätigen am Beispiel der Bundesländer Baden-Württemberg und Hamburg

	Baden-Württemberg								
	Erwerbstätige				relative Veränderung		Shift-Share-Zerlegung		
	1994		2003		2003 gegenüber 1994		Regional-F.	Struktur-F.	Standort-F.
	in Tsd.	Anteil	in Tsd.	Anteil	i.d. Region	im Bund	(2)*(5)	(2)*(6)	(7)–(8)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sektor1	150	3,0%	113	2,1%	-24,7%	-20,8%	-0,7%	-0,6%	-0,1%
Sektor2	1 963	39,3%	1 827	34,3%	-6,9%	-15,9%	-2,7%	-6,3%	3,5%
Sektor3	2 884	57,7%	3 381	63,5%	17,2%	13,3%	9,9%	7,7%	2,3%
Summe	4 997		5 321		6,5%	2,5%	6,5%	0,8%	5,7%
	Hamburg								
	Erwerbstätige				relative Veränderung		Shift-Share-Zerlegung		
	1994		2003		2003 gegenüber 1994		Regional-F.	Struktur-F.	Standort-F.
	in Tsd.	Anteil	in Tsd.	Anteil	i.d. Region	im Bund	(2)*(5)	(2)*(6)	(7)–(8)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sektor1	6	0,6%	5	0,5%	-16,7%	-20,8%	-0,1%	-0,1%	0,0%
Sektor2	216	21,0%	167	16,2%	-22,7%	-15,9%	-4,8%	-3,3%	-1,4%
Sektor3	807	78,4%	859	83,3%	6,4%	13,3%	5,1%	10,4%	-5,4%
Summe	1 029		1 031		0,2%	2,5%	0,2%	7,0%	-6,8%

Eigene Berechnungen, Quelle der Ausgangsdaten: Statistisches Bundesamt.

Die *Abbildung 2.2* zeigt den Zusammenhang zwischen Struktureffekt und Regionaleffekt für die 16 Bundesländer. Eine Regressionsanalyse lässt auf keinen signifikanten Zusammenhang zwischen der Wirtschaftsstruktur und dem regionalen Wachstum der Erwerbstätigkeit schließen. Umgekehrt ergibt sich ein stark signifikanter Zusammenhang zwischen Standort- und Regionaleffekt, wie in der *Abbildung 2.3* deutlich wird. Dies unterstreicht nochmals, dass in unserer Analyse der Standortfaktor den Strukturfaktor klar dominiert. Entscheidend sind demnach die Bedingungen des Wirtschaftens (z. B. Infrastruktur, Forschungs- und Bildungseinrichtungen, spezialisierte Arbeitnehmerschaft), nicht so sehr der Schwerpunktbereich, in dem die Produktion stattfindet. Die relative Größe der drei Wirtschaftssektoren zueinander bestimmt damit nicht von vornherein die Erwerbstätigenbilanz der Bundesländer.

Es ist demnach zu fragen, ob die Sektorstruktur überhaupt eine Information ist, die für die Entwicklungsperspektive einer Region Bedeutung besitzt. Um eine Antwort auf diese Frage zu geben, wollen wir uns dem Sachverhalt nochmals von einer anderen Seite nähern. Dabei betrachten wir die sektorspezifischen Wachstumsraten für alle 16 Bundesländer im Zeitraum 1994 bis 2003. Untersucht werden soll, ob die im Jahr 1994 beobachtete Wirtschaftsstruktur überhaupt etwas mit der Wachstumsbilanz der Bundesländer im Zeitraum danach zu tun hat oder nicht. Wir klammern dabei den gesamtwirtschaftlich wenig bedeutsamen Primären Sektor aus und beschränken uns auf das Produzierende Gewerbe und den Dienstleistungsbereich. Die *Abbildungen 2.4* und *2.5* bilden den Zusammenhang zwischen dem Sektoranteil im Ausgangsjahr und der Veränderungsrate der Erwerbstätigkeit in diesem Sektor für den Zeitraum danach ab. Die Resultate einer Regressionsschätzung sind in der *Tabelle 2.3* enthalten. Dabei wurden die Unterschiede zwischen West- und Ostdeutschland mit Hilfe einer sogenannten Dummy-Variablen modelliert. Die Dummy-Variable ist wie folgt konstruiert: Sie nimmt den Wert Null (Eins) an, wenn sich der Beobachtungswert auf eines der alten (der neuen) Bundesländer bezieht. Ein Unterschied in den Mittelwerten der abhängigen Variable bei sonst gleichen Werten der übrigen erklärenden Variablen drückt sich dann im geschätzten Koeffizienten der Dummy-Variablen aus.

Die Ergebnisse sind wie folgt zu interpretieren: Der Koeffizient der Dummy-Variable für die Neuen Länder ist in beiden Fällen negativ und statistisch signifikant. Dies bedeutet, dass das Wachstum der Erwerbstätigkeit *ceteris paribus* in Ostdeutschland schwächer war als in Westdeutschland. Der vertikale Achsenabschnitt der Regressionsgerade ist für das Produzierende Gewerbe negativ, für den Dienstleistungsbereich hingegen positiv. Damit bildet die Schätzung den schon beschriebenen Sektorwandel zugunsten des Tertiären Sektors ab (immer mehr Erwerbstätige sind im Dienstleistungsbereich beschäftigt). Von besonderem Interesse ist aber der Koeffizient der Variable, die den Beschäftigtenanteil im Ausgangsjahr wiedergibt. Ein positiver Koeffizient würde anzeigen, dass die Entwicklung dort vergleichsweise positiv verlaufen ist, wo im Ausgangsjahr der Erwerbstätigenanteil in diesem Sektor hoch war. Umgekehrt impliziert ein negativer Koeffizient, dass ein bereits in der Ausgangslage hoher Anteil des Sektors ungünstig für das Wachstum ist.

Aus der *Tabelle 2.3* ist zu entnehmen, dass man für den Sekundären Sektor ein statistisch hoch signifikantes positives Vorzeichen erhält. Dies impliziert, dass sich die Er-

werbstätigkeit im Produzierenden Gewerbe in solchen Bundesländern vergleichsweise am günstigsten entwickelt hat, in denen im Jahr 1994 überproportional viele Erwerbstätige in der Industrie tätig waren. Umgekehrt ist der Industriesektor dort besonders stark geschrumpft, wo er bereits in der Ausgangslage schwach war. Die Ergebnisse sind demnach als eine räumliche Konzentrationstendenz im Sekundären Sektor zu deuten. Im Dienstleistungssektor ist der geschätzte Koeffizient des Ausgangsanteils hingegen negativ. Ein hoher Anteil an Dienstleistungsbeschäftigten im Jahr 1994 bedeutete demnach keine vorteilhafte Position. Im Gegenteil, ein hohes Wachstum der Beschäftigung im Tertiären Sektor lag eher dort vor, wo im Ausgangsjahr wenig vorhanden war. Die Ergebnisse sprechen somit für einen Aufholprozess, der dazu führen könnte, dass sich die regionalen Beschäftigtenanteile im Dienstleistungssektor tendenziell angleichen.

Abbildung 2.2: Der Zusammenhang zwischen dem berechneten Struktureffekt und dem Regionaleffekt

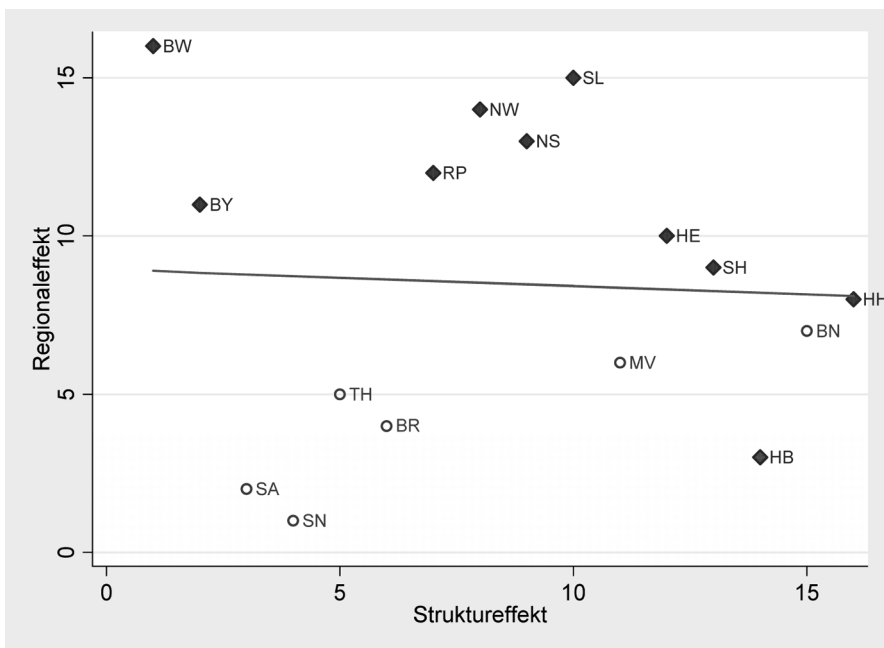


Abbildung 2.3: Der Zusammenhang zwischen dem berechneten Standorteffekt und dem Regionaleffekt

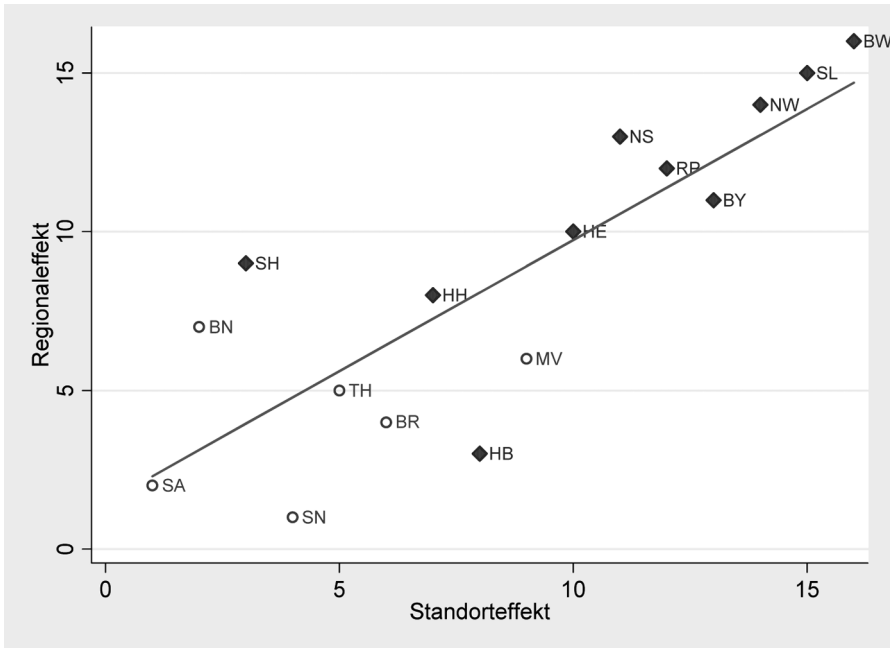


Abbildung 2.4: Der Zusammenhang zwischen der Veränderungsrate der Erwerbstätigkeit im Sekundären Sektor 1994–2003 und dem Anteil dieses Sektors im Ausgangsjahr

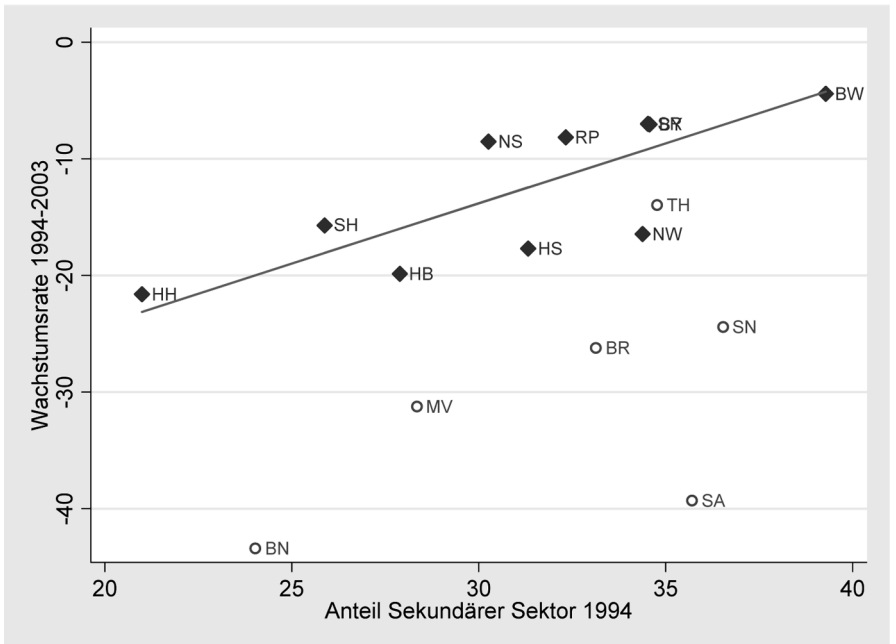


Abbildung 2.5: Der Zusammenhang zwischen der Veränderungsrate der Erwerbstätigkeit im Tertiären Sektor 1994–2003 und dem Anteil dieses Sektors im Ausgangsjahr

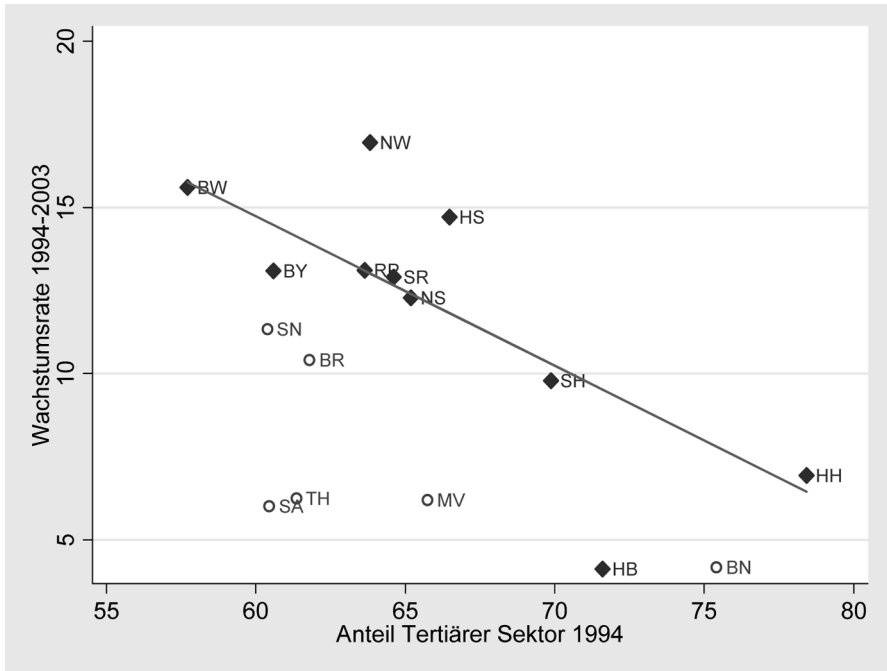


Tabelle 2.3: Ergebnisse einer Regressionsschätzung für den Zusammenhang zwischen der Veränderungsrate der Erwerbstätigkeit und dem Beschäftigungsanteil des jeweiligen Sektors

Abhängige Variable: Veränderungsrate der Erwerbstätigkeit im jeweiligen Sektor im Zeitraum 1994 bis 2003				
	Sekundärer Sektor		Tertiärer Sektor	
Erklärende Variablen:	Koeff.	Standard F.	Koeff.	Standard F.
Konstante	-44.79**	7.78	41.71**	6.08
Sektoranteil an der Erwerbstätigkeit 1994	1.03**	0.25	-0.45**	0.10
Dummy-Variable Neue Länder	-18.10**	3.77	-5.45**	1.21
Bestimmtheitsmaß (R ²)	0.74		0.68	

Anmerkungen: Eigene Berechnungen; Quelle für Ausgangsdaten: Statistisches Jahrbuch 2004, Tabelle 3.9: Erwerbstätige nach Ländern. Standard F.: Heteroskedastie-konsistente Standardfehler; ** bezeichnet Signifikanz auf dem 1%-Niveau.

2.4 Indizes der regionalen Spezialisierung und der geographischen Konzentration

2.4.1 Spezialisierung und Konzentration

Ein Großteil der regionalökonomischen Analyse hat damit zu tun, dass sich die ökonomische Aktivität nicht gleichförmig im Raum verteilt. Im Folgenden sei unterstellt, dass ein Gitternetz über ein Land gelegt wird und die dabei entstehenden Planquadrate als Wirtschaftsräume oder Regionen bezeichnet werden. Es soll untersucht werden, wie sich die wirtschaftlichen Aktivitäten in den Planquadraten unterscheiden. Drei Begriffe müssen dabei auseinander gehalten werden, obwohl sie miteinander verwandt sind:

- Eine *regionale Spezialisierung* liegt vor, wenn ein Wirtschaftsraum bei der Herstellung ganz bestimmter Güter und Dienste komparative Vorteile aufweist und deshalb einen vergleichsweise hohen Anteil seiner Kapazitäten für deren Produktion einsetzt.
- Von der *Konzentration einer spezifischen Produktionsaktivität im Raum* ist zu sprechen, wenn ein hoher Anteil der Gesamtproduktion eines bestimmten Gutes von nur wenigen Regionen erstellt wird.
- *Konzentration der gesamten Produktion im Raum* schließlich bedeutet, dass eine verhältnismäßig kleine Anzahl von Regionen einen hohen Anteil der gesamten Wirtschaftsleistung erbringt.

Naturgemäß ist die Konzentration einer speziellen Produktionsaktivität im Raum meist auch mit einer Spezialisierung von Regionen verbunden, während umgekehrt jede Spezialisierung auch mit einer gewissen räumlichen Konzentration dieser Produktionsaktivität einhergeht. Konzentration und Spezialisierung müssen aber nicht zusammenfallen. Wenn sich beispielsweise viele kleinere Regionen auf eine bestimmte Produktionsaktivität spezialisieren, so zieht dies keine nennenswerte Konzentration nach sich, da der Marktanteil der einzelnen Wirtschaftsräume gering bleibt.

In der traditionellen ebenso wie in der neuen regionalökonomischen Theorie spielen Zentrifugal- und Zentripetalkräfte gleichermaßen eine Rolle. Zentrifugalkräfte führen dazu, dass sich die wirtschaftliche Aktivität vom Zentrum in die Peripherie verlagert, während Zentripetalkräfte das Gegenteil bewirken. Welche Seite überwiegt, hängt in den theoretischen Modellen von Parametern wie der Höhe der Transportkostensätze, der Stärke der Skaleneffekte und der Mobilität der Produktionsfaktoren ab. Letztlich bleibt es also eine empirische Frage, in welcher Richtung die tatsächlichen Entwicklungsprozesse verlaufen. Neben der Konzentration der Produktionsaktivität im Raum spielt dabei auch die Spezialisierung von Regionen eine Rolle. Es bietet sich deshalb an, sich mit Indikatoren zur quantitativen Erfassung dieser Phänomene näher zu befassen.

Spezialisierungsindizes beziehen sich auf die Wirtschaftsstruktur von Regionen, Konzentrationsindizes auf die Verteilung von wirtschaftlicher Aktivität im Raum. Indikatoren, die sich zur Messung des Spezialisierungs- bzw. Konzentrationsgrades eignen, sind der nach Orris C. Herfindahl benannte Index und der räumliche Gini-Koeffizient (nach Corrado Gini). Daneben ist seit Paul Krugmans *Geography and Trade* (1991b) ein weiterer Index unter dem Namen *Krugman-Index* populär geworden, obwohl er auf

ältere Vorgänger wie den englischen Ökonomen und Soziologen Philip Sargant Florence (1890–1982) zurückgeht und schon bei Walter Isard (1960) als Standardansatz beschrieben wird.

Um die Konstruktion der Indikatoren zu erläutern, gehen wir von einfachen Beispielen mit $R = 4$ Regionen (R1 bis R4) und $I = 4$ Industrien (A, B, C und D) aus, in denen jeweils vier Unternehmen tätig sind. Jedes Unternehmen beschäftigt eine feste Zahl von Mitarbeitern, die wir auf eins festsetzen, um unnötige Komplikationen zu vermeiden. Ein Großbuchstabe bezeichne ein Unternehmen aus der entsprechenden Industrie.

Als Extremfälle ergeben sich die folgenden beiden Situationen.

Fall (i): Keinerlei Spezialisierung/vollkommene Dekonzentration, d. h. jede Region produziert alles.

R1	R2	R3	R4
ABCD	ABCD	ABCD	ABCD

Fall(ii): Komplette Spezialisierung/komplette Konzentration, d. h. jede Region weist nur Unternehmen aus einer einzigen Industrie auf und jede Industrie findet sich nur in einer einzigen Region.

R1	R2	R3	R4
AAAA	BBBB	CCCC	DDDD

Im Fall (i) sollten Spezialisierungs- und Konzentrationsindex somit ihren minimalen, im Fall (ii) hingegen ihren maximalen Wert annehmen.

In den Fällen (i) und (ii) fallen Spezialisierung und Konzentration zusammen. Das folgende Beispiel mit nur zwei Industrien A und B zeigt aber, dass völlige Spezialisierung nicht zwangsläufig eine komplette Konzentration bedingt.

Fall (iii): Beispiel einer kompletten Spezialisierung bei unvollkommener Konzentration

R1	R2	R3	R4
AAAA	BBBB	AAAA	BBBB

Alle Regionen weisen nur Unternehmen einer Industrie auf, sind also vollkommen spezialisiert. Die Produktion jeder Branche teilt sich aber hälftig auf zwei Regionen auf, ist also räumlich keineswegs vollständig konzentriert.

Fall (iv): Beispiel unvollkommener Spezialisierung und Konzentration – symmetrischer Fall

R1	R2	R3	R4
AABB	BBCC	CCDD	DDAA

Hier sind die Beschäftigten in allen Regionen zu gleichen Teilen in jeweils zwei Industriezweigen tätig. Dies bedeutet eine gewisse, aber keine vollständige Spezialisierung. Da jeweils zwei Regionen in einer Branche aktiv sind, liegt keine vollständige Konzentration vor.

Fall (v): Beispiel unvollkommener Spezialisierung und Konzentration – asymmetrischer Fall

R1	R2	R3	R4
AAAB	BBCD	ACCD	BCDD

In diesem Fall ist die Region R1 weitgehend, jedoch nicht vollständig auf Industrie A spezialisiert, während die Industriestruktur der Regionen R2 bis R4 stärker diversifiziert ist.

Die im Folgenden zu beschreibenden Indikatoren zielen darauf ab, die Industriestruktur der Regionen vergleichbar zu machen. Die Strukturen sind am ähnlichsten, wenn – wie in Fall (i) – die jeweiligen Anteile der Beschäftigung in den Regionen identisch sind. Stark voneinander abweichende Anteile weisen hingegen auf unterschiedliche Industriestrukturen hin. Genau diese Diskrepanzen messen die verschiedenen Indikatoren. Der Herfindahl-Index verwendet die aufsummierten quadrierten Anteile, der Krugman-Index die aufsummierten absoluten regionalen Abweichungen der Anteile von einem Normalwert auf gesamtwirtschaftlicher Ebene. Der räumliche Gini-Koeffizient lehnt sich an das aus der Statistik bekannte Maß zur Messung von Ungleichheit an. Die verschiedenen Indikatoren zielen darauf ab, denselben Sachverhalt zu messen, führen aber – wie zu zeigen sein wird – keineswegs zu identischen Resultaten.

2.4.2 Spezialisierungsindizes

Spezialisierungsmessung mit dem Krugman-Index

Wir beginnen mit der Darstellung des Krugman-Index zur Messung der regionalen Spezialisierung. Dazu ist es notwendig, eine Notation einzuführen. x sei ein Indikator, mit dem die wirtschaftliche Aktivität gemessen werden soll (z. B. die Produktion oder, wie im Folgenden angenommen, die Beschäftigung). Der Index i bezeichne die Industrie ($i = 1, \dots, I$), r die Region ($r = 1, \dots, R$). Die Größe x_{ir} steht also für die Zahl der Beschäftigten in Industrie i in der Region r . x_i ist die Gesamtbeschäftigung der Industrie i im Gesamttraum, x_r die Summe der Beschäftigten aller Industrien in Region r und x die Summe der Beschäftigten aller Industrien. Es gilt also

$$x_r := \sum_{i=1}^I x_{ir}, \quad x_i := \sum_{r=1}^R x_{ir}, \quad x := \sum_{i=1}^I \sum_{r=1}^R x_{ir}.$$

Weiterhin seien die folgenden Anteilsgrößen definiert:

$$a_{ir} := \frac{x_{ir}}{x_r}$$

ist der Anteil der Industrie i in der Region r an der Gesamtaktivität dieser Region.

$$A_{ir} := \frac{x_{ir}}{x_i}$$

ist der Anteil der in der Region Beschäftigten in Industrie i an der Beschäftigung dieser Industrie auf aggregierter, gesamtäumlicher Ebene. Wenn beispielsweise die erste Region das Bundesland Baden-Württemberg ist und der Maschinenbau den Index $i = 2$ besitzt, dann bezeichnet a_{21} den Anteil der im Maschinenbau Beschäftigten an der Gesamtbeschäftigung in Baden-Württemberg. Entsprechend gibt die Größe A_{21} den Anteil der Beschäftigten im Maschinenbau Baden-Württembergs an der Gesamtzahl aller in dieser Branche Beschäftigten in Deutschland an.

Weiterhin spielt eine Rolle, welchen Anteil die über alle Branchen summierte Beschäftigung einer Region an der Beschäftigung im Gesamtraum besitzt und welcher Anteil der Beschäftigung im Gesamtraum auf die einzelnen Industrien entfällt,

$$a_r := \frac{x_r}{x} \text{ und } A_i := \frac{x_i}{x}.$$

Der *Krugman-Index* misst regionale Spezialisierung durch den Vergleich der Industriestruktur der Region mit der Industriestruktur im Gesamtraum. Präziser formuliert: Für jede Industrie der Region ist die Differenz zwischen dem tatsächlichen Beschäftigtenanteil in der Region und dem Beschäftigtenanteil, den diese Industrie im Gesamtraum aufweist, zu berechnen. Der Absolutbetrag dieser Differenzen wird dann über alle Industrien aufsummiert. Formal ergibt sich

$$SPEZ_r^{K(\text{rugman})} := \sum_i |a_{ir} - A_i|.$$

Betrachten wir den oben dargestellten Fall (i). Der Anteil jeder Industrie in jeder Region beträgt genau 25%, ebenso auf aggregierter Ebene. Es ergibt sich ein Spezialisierungsindex von Null für jede der vier Regionen. Damit wird der Minimalwert des Index erreicht, wie es bei völliger Gleichverteilung der Industrien über die Regionen auch der Fall sein sollte. Im Fall (ii) ist der Anteil jeder Industrie auf aggregierter Ebene (im Gesamtraum) wiederum 25%. In Region 1 besitzt die erste Industrie A einen Anteil von 100%, alle anderen sind nicht vorhanden. Dies bedeutet für den Index:

$$SPEZ_{r=1}^K = \sum_{i=1}^4 |a_{i1} - A_i| = |1 - 0,25| + |0 - 0,25| + |0 - 0,25| + |0 - 0,25| = 1,5.$$

Ein identischer Wert errechnet sich für die anderen Regionen. Zwar kann der Krugman-Index im allgemeinen Fall (bei einer großen Anzahl von Industrien) maximal den Wert Zwei annehmen, im hier betrachteten Fall ist der größtmögliche Wert des Indikators jedoch nur 1,5.

Im Fall (iv) berechnet sich der Indikator für Region 1 wie folgt

$$SPEZ_{r=1}^K = \sum_{i=1}^4 |a_{i1} - A_i| = |0,5 - 0,25| + |0,5 - 0,25| + |0 - 0,25| + |0 - 0,25| = 1.$$

Der Indikatorwert signalisiert also eine Spezialisierung, deren Ausmaß zwischen den Extremfällen (i) und (ii) liegt. Da die Industriestruktur in allen Regionen hier symmetrisch modelliert wurde, ist der Indikatorwert für alle Regionen identisch.

Interessant erscheint der asymmetrische Fall (v). Für die vergleichsweise stark spezialisierte Region 1 ergibt sich

$$SPEZ_{r=1}^K = \sum_{i=1}^4 |a_{i1} - A_i| = |0,75 - 0,25| + |0,25 - 0,25| + |0 - 0,25| + |0 - 0,25| \\ = 1,25$$

und für die stärker diversifizierte Region 2

$$SPEZ_{r=2}^K = \sum_{i=1}^4 |a_{i2} - A_i| = |0 - 0,25| + |0,5 - 0,25| + |0,25 - 0,25| + |0,25 - 0,25| \\ = 0,5,$$

für die Regionen 3 und 4 ergeben sich analoge Werte.

Spezialisierungsmessung mit dem Herfindahl-Index

Die Gleichung für den *Herfindahl-Index* zur Messung der regionalen Spezialisierung lautet

$$SPEZ_r^{H(\text{erfindahl})} := \sum_i a_{ir}^2.$$

Die *Tabelle 2.4* enthält den Herfindahl-Index für die fünf im Beispiel konstruierten Fälle im Vergleich mit dem Krugman-Index.

Tabelle 2.4: Spezialisierungsindikatoren nach Herfindahl und Krugman für die Beispiele (i) bis (v)

Beispiel	Herfindahl-Index für Region				Krugman-Index für Region			
	R1	R2	R3	R4	R1	R2	R3	R4
(i)	0.25	0.25	0.25	0.25	0	0	0	0
(ii)	1.0	1.0	1.0	1.0	1.5	1.5	1.5	1.5
(iii)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
(iv)	0.5	0.5	0.5	0.5	1.0	1.0	1.0	1.0
(v)	0.625	0.375	0.375	0.375	1.25	0.5	0.5	0.5

Die Ergebnisse zeigen, dass im Fall vollkommener Gleichverteilung der Herfindahl-Index nicht auf den Wert Null zurückgeht, sondern auf 0,25. Man kann zeigen, dass der niedrigste Wert für den Herfindahl-Index bei N Industrien $1/N$ beträgt. Wir sehen auch, dass der Index im Fall (v) für Region 1 eine erheblich höhere Spezialisierung anzeigt als für die übrigen Regionen.

Auch wenn die Werte zwischen den beiden Indizes aufgrund der unterschiedlichen Konstruktion nicht übereinstimmen, so ergibt sich doch eine qualitative Ähnlichkeit der Resultate für die gewählten Beispiele.

Spezialisierungsmessung mit dem Gini-Koeffizienten

Der aus der Statistik bekannte *Gini-Koeffizient* zur Messung der Ungleichheit einer Verteilung kann auch zur Messung der Spezialisierung einer Region eingesetzt werden. Dabei geht man wie folgt vor. Ausgangspunkt der Berechnung sind die Anteile der einzelnen Industrien an der Gesamtaktivität einer Region, a_{ir} . Durch Division mit dem entsprechenden Anteil der Industrie im Gesamttraum ergibt sich der *Standortquotient* oder *Balassa-Index*

$$B_{ir} := a_{ir} / A_i.$$

Zu beachten ist, dass sich der Balassa-Index alternativ auch berechnen lässt, indem man den Anteil, den die Industrie in einer Region an dieser Industrie auf gesamt-räumlicher Ebene hält, durch den Anteil der regionalen Gesamtaktivität an der Gesamtwirtschaft dividiert. Die Äquivalenz lässt sich wie folgt zeigen:

$$B_{ir} = \frac{A_{ir}}{a_r} = \frac{x_{ir}/x_i}{x_r/x} = \frac{x_{ir}/x_r}{x_i/x} = \frac{a_{ir}}{A_i}.$$

Der Balassa-Index wird nun für eine ausgewählte Region r für alle Industrien $i = 1, \dots, I$ berechnet und in aufsteigender Reihenfolge geordnet, d. h. die Industrie mit dem niedrigsten Balassa-Index in der Region erhält die Ordnungsnummer 1, die mit dem zweitniedrigsten die Ordnungsnummer 2 usw. Im Folgenden seien alle in dieser Reihenfolge angeordneten Größen mit einer „Tilde“ ($\tilde{}$) gekennzeichnet. Benötigt werden neben den \tilde{a}_{ir} die entsprechend angeordneten Beschäftigungsanteile der einzelnen Industrien auf gesamt-räumlicher Ebene, \tilde{A}_i .

Die geordneten Anteile werden kumuliert:

$$\tilde{s}_{jr} := \sum_{i=1}^j \tilde{a}_{ir} \quad \text{und} \quad \tilde{S}_{jr} := \sum_{i=1}^j \tilde{A}_i \quad \text{für } j = 1, \dots, I.$$

In einem Diagramm können nun (für $j = 1, \dots, I$) die kumulierten Anteile \tilde{S}_{jr} auf der Abzisse und \tilde{s}_{jr} auf der Ordinate gegeneinander abgetragen werden. Werden die Punkte miteinander verbunden, so entsteht eine Darstellung, die sich an die bekannte *Lorenz-Kurve* aus der Statistik anlehnt. Die doppelte Fläche zwischen dieser modifizierten Lorenz-Kurve und der Winkelhalbierenden ergibt ein Spezialisierungsmaß analog zum Gini-Koeffizienten. Die Gleichung dafür lautet

$$SPEZ_r^{G(im)} := 1 - \sum_{i=1}^I (\tilde{s}_{i-1,r} + \tilde{s}_{i,r}) \tilde{A}_i \quad \text{mit } \tilde{s}_{0,r} = 0.$$

Wie der Gini-Koeffizient liegt auch dieses Maß zwischen den Werten Null und Eins. Ein Wert von Null bedeutet, dass die betreffende Region keine Spezialisierung aufweist, während ein Wert in der Nähe von Eins eine hohe Spezialisierung anzeigt.

Erwartungsgemäß errechnet sich für den Fall (i) z. B. für Region 1 $SPEZ_{r=1}^G = 1 - (0 + 0,25)0,25 - (0,25 + 0,50)0,25 - (0,50 + 0,75)0,25 - (0,75 + 1,00)0,25 = 0$, da hier keine Spezialisierung vorliegt. Hingegen ergibt sich im Fall (v) für die vergleichsweise stark spezialisierte Region 1 ein relativ hoher Wert

$$SPEZ_{r=1}^G = 1 - (0 + 0)0,25 - (0 + 0)0,25 - (0 + 0,25)0,25 - (0,25 + 1,00)0,25 = 0,625.$$

Das Spezialisierungsmaß ist in diesem Fall für die stärker diversifizierte Region 2 den Erwartungen entsprechend geringer:

$$SPEZ_{r=2}^G = 1 - (0 + 0)0,25 - (0 + 0,25)0,25 - (0,25 + 0,50)0,25 - (0,50 + 1,00)0,25 = 0,375.$$

2.4.3 Konzentrationsindizes

In Analogie zu den Spezialisierungsindizes kann die geografische Konzentration einer Branche durch die Abweichung ihrer geografischen Verteilung von der geografischen Verteilung der Gesamtwirtschaft gemessen werden. Die unter *Abschnitt 2.4.2* genannten Indizes eignen sich – nach entsprechender Anpassung – auch zur Messung räumlicher Konzentration industrieller Aktivität.

Überhaupt keine Konzentration liegt vor, wenn der Anteil, den eine Region an der nationalen Gesamtaktivität eines bestimmten Sektors oder einer Industrie besitzt, in allen Regionen gleich ist. Die höchste Konzentration ist natürlich dann gegeben, wenn eine einzige Region die Aktivität dieses Sektors in der Gesamtwirtschaft auf sich vereinigt.

Konzentrationsmessung mit dem Krugman-Index

Die für die Konzentrationsmessung einer Industrie i modifizierte Gleichung des *Krugman-Index* lautet

$$KONZ_i^K := \sum_r |A_{ir} - a_r|.$$

Konzentrationsmessung mit dem Herfindahl-Index

Demgegenüber ist der entsprechende *Herfindahl-Index* wir folgt zu berechnen:

$$KONZ_i^H := \sum_r A_{ir}^2.$$

Für die einfachen Fälle (i) bis (v) ergeben sich nur im Fall (iii) Abweichungen in den numerischen Werten im Vergleich von Spezialisierungs- und Konzentrationsindizes. Im Fall (iii) bestimmt sich der Indexwert nach Krugman für die der Reihenfolge nach erste Industrie

$$AalsKONZ_{i=1}^K = \sum_{r=1}^4 |A_{1r} - a_r| = |0,5 - 0,25| + |0 - 0,25| + |0,5 - 0,25| + |0 - 0,25| = 1,$$

analog für Industrie B . Der Herfindahl-Konzentrationsindex errechnet sich für den Fall (iii) als

$$KONZ_{i=1}^H = \sum_{r=1}^4 A_{1r}^2 = 0,5^2 + 0^2 + 0,5^2 + 0^2 = 0,5.$$

Konzentrationsmessung mit dem Gini-Koeffizienten

Bei der Verwendung des *Gini-Koeffizienten* zur Messung der räumlichen Konzentration einer Industrie geht man von den Anteilen an der gesamtäumlichen Beschäftigung dieser Industrie aus, die auf die einzelnen Regionen entfallen. Für eine ausgewählte Industrie berechnet man diese Anteile A_{ir} für alle Regionen. Analog zum Vorgehen in *Abschnitt 2.4.3* werden nun die Regionen nach dem Balassa-Index in aufsteigender Reihenfolge angeordnet und die entsprechenden Werte der Variablen mit einer Tilde (\sim) gekennzeichnet. Neben den \tilde{A}_{ir} sind die in die entsprechende Reihenfolge gebrachten Anteile der regionalen Beschäftigung an der Beschäftigung im Gesamtraum, \tilde{a}_r , erforderlich. Die kumulierten Werte seien definiert als

$$\tilde{Z}_{ij} := \sum_{r=1}^j \tilde{A}_{ir} \quad \text{und} \quad \tilde{z}_{ij} := \sum_{r=1}^j \tilde{a}_r \quad \text{für } j = 1, \dots, R.$$

Die grafische Darstellung geschieht in der schon bekannten Weise, indem die \tilde{z}_{ij} und die \tilde{Z}_{ij} gegeneinander aufgetragen werden. Die Gleichung für das Konzentrationsmaß auf Basis des Gini-Koeffizienten lautet

$$KONZ_i^G := 1 - \sum_{r=1}^R (\tilde{Z}_{i,r-1} + \tilde{Z}_{i,r}) \tilde{a}_i \quad \text{mit} \quad \tilde{Z}_{i,0} = 0.$$

Wie der Gini-Koeffizient im Standardfall liegt auch der räumliche Gini-Koeffizient zwischen den Werten Null und Eins. Ein Wert von Null bedeutet, dass die räumliche Verteilung der betreffenden Industrie und die räumliche Verteilung der Gesamtbeschäftigung identisch sind, während ein Wert von Eins eine vollständige räumliche Konzentration der Industrie anzeigt.

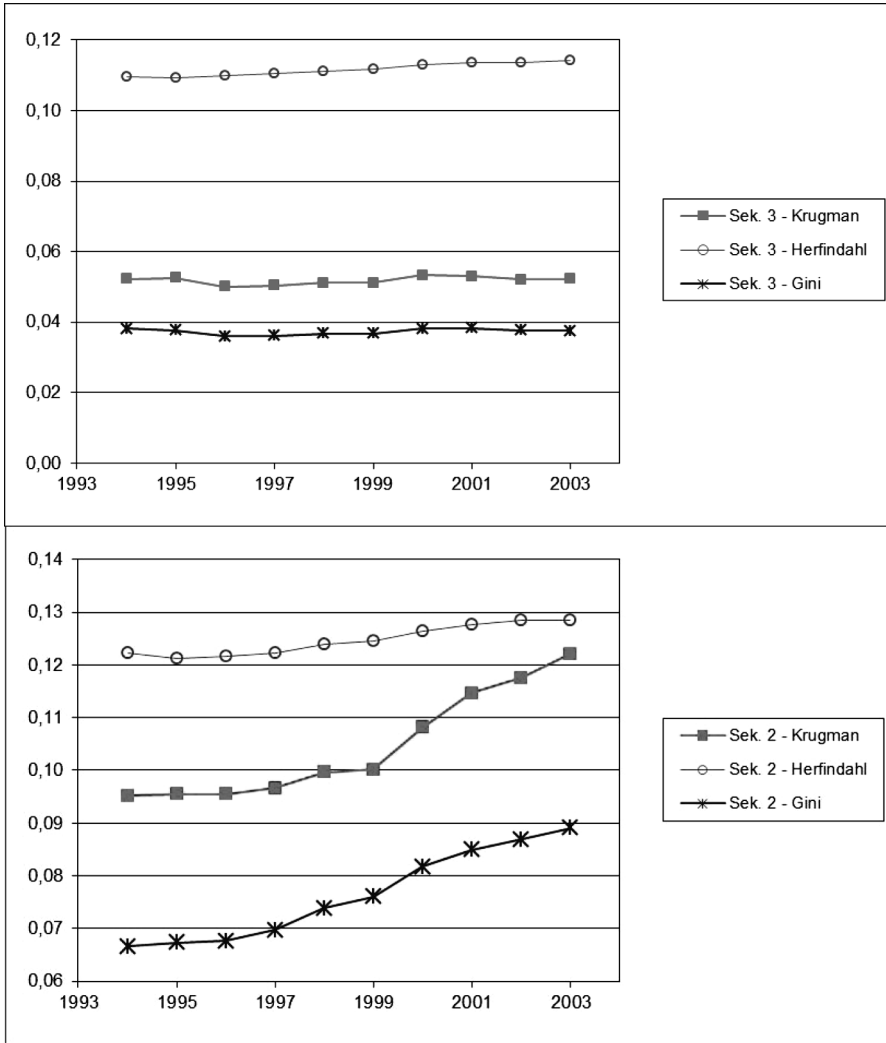
2.4.4 Praktische Anwendung auf die Sektorstrukturdaten

Die *Abbildung 2.6* zeigt den Verlauf der räumlichen Konzentrationsindizes nach Krugman, Herfindahl und Gini für den Sekundären und Tertiären Sektor. Im Tertiären Sektor verlaufen die Indizes demnach relativ flach, nur der Herfindahl-Index weist eine leichte Aufwärtstendenz auf. Für die Entwicklung der räumlichen Konzentration im Sekundären Sektor ist den Berechnungen zufolge hingegen ein positiver Zeittrend nachweisbar. Besonders der Krugman-Index und der Gini-Koeffizient weisen ab dem Jahr 1999 deutlich nach oben.

Wie ausgeführt wurde, misst der Krugman-Index den Unterschied zwischen der regionalen und der aggregierten Produktionsstruktur. Die Diskrepanz zwischen der tatsächlichen Erwerbstätigkeit im Sekundären Sektor und der Erwerbstätigkeit, die in den jeweiligen Bundesländern bei gleicher Wirtschaftsstruktur wie im Bund zu erwarten gewesen wäre, nimmt im Beobachtungszeitraum deutlich zu. Gleichzeitig wächst dem Gini-Koeffizienten zufolge die Abweichung von der räumlichen Gleichverteilung. Zusammengefasst ist dies eine starke Evidenz dafür, dass die räumliche Konzentration der Wirtschaftsaktivität im Produzierenden Gewerbe seit Ende der 1990er Jahre tendenziell angestiegen ist.

Abbildung 2.6: Indizes der räumlichen Konzentration für den Sekundären (unten) und den Tertiären Sektor (oben) nach Krugman, Herfindahl und Gini

Erwerbstätige in den deutschen Bundesländern 1994–2003



Die *Abbildungen 2.7* und *2.8* zeigen die berechneten Indizes der sektoralen Spezialisierung nach Krugman bzw. Herfindahl für die 16 deutschen Bundesländer. Zunächst ergibt sich nach beiden Indizes eine besonders hohe Spezialisierung für die Stadtstaaten Hamburg und Berlin. Hier weicht demnach die Wirtschaftsstruktur am stärksten von der im Aggregat (in Deutschland) ab. Niedrige Werte ergeben sich hingegen für Nordrhein-Westfalen und Rheinland-Pfalz. Interessant erscheint die zeitliche Entwicklung. Den Berechnungen zufolge ist eine generelle Tendenz für den Krugman-Index nicht zu erkennen. In sieben von 16 Bundesländern ist die sektorale Spezialisie-

rung gesunken, in neun gestiegen. Hingegen weist der Herfindahl-Index der sektoralen Spezialisierung eindeutig einen steigenden Trend der sektoralen Spezialisierung aus. Für alle Bundesländer liegt der Spezialisierungswert von 2003 nach diesem Messkonzept über dem von 1994.

Abbildung 2.7: Krugman-Indizes der sektoralen Spezialisierung in den deutschen Bundesländern (1994–2003)

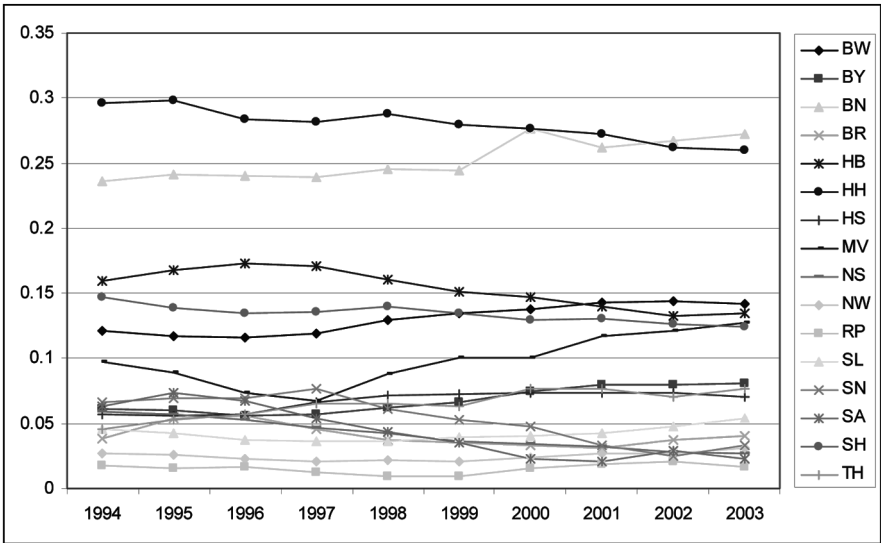
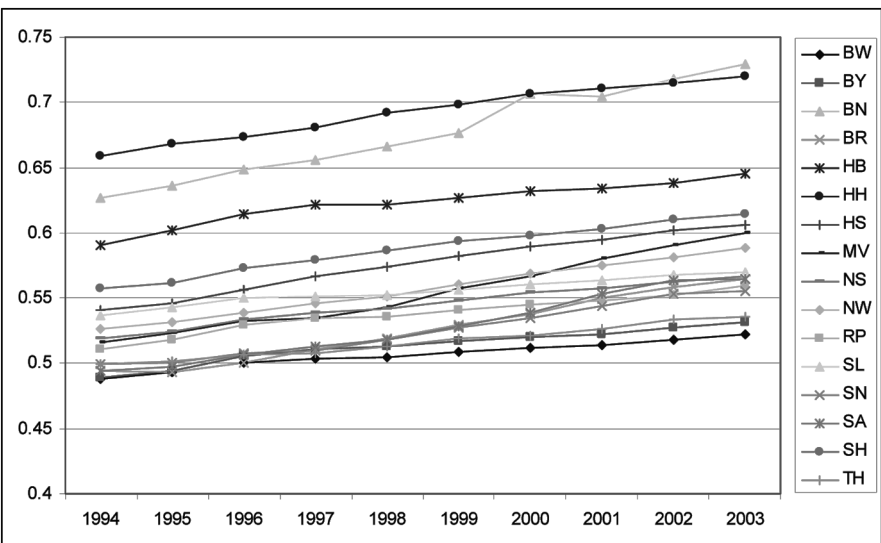


Abbildung 2.8: Herfindahl-Indizes der sektoralen Spezialisierung in den deutschen Bundesländern (1994–2003)



Beim Vergleich der Ergebnisse nach Krugman und Herfindahl sollte man sich die unterschiedliche Konstruktion dieser Maßzahlen in Erinnerung rufen. Der Krugman-Index vergleicht die Wirtschaftsstruktur der Länder mit dem Bund. Ein hoher Indexwert bedeutet eine hohe Diskrepanz der Struktur zwischen dem jeweiligen Land und dem Bund. Der Herfindahl-Index misst die Unterschiede in den Sektoranteilen jedes Bundeslandes, ohne einen Vergleich mit dem Bund vorzunehmen. Ein Anwachsen des stärksten Sektors (des Dienstleistungsbereichs) bewirkt zwangsläufig einen Anstieg des Herfindahl-Index. Hinter dem zunehmendem Trend in *Abbildung 2.8* steckt deshalb vermutlich zu einem beträchtlichen Teil der allgemeine Sektorwandel.

2.5 Räumliche Ökonometrie (Spatial Econometrics)

2.5.1 Was ist die räumliche Ökonometrie?

Regionen sind keine isolierten Gebilde. Sie stehen in vielfältigen Wechselbeziehungen zu ihren räumlichen Nachbarn. Empirische Studien auf der Basis von Querschnitts- oder kombinierten Längs-/Querschnittsbeobachtungen vernachlässigen häufig diese Zusammenhänge und werten die verfügbaren Informationen damit nur suboptimal aus. Die Konsequenzen daraus sind mit denjenigen beim Vorliegen von Autokorrelation oder Heteroskedastie in der Regressionsanalyse vergleichbar, d. h. es ist mit ineffizienten Schätzungen der Parameter sowie mit verzerrten und inkonsistenten Schätzungen der Standardfehler zu rechnen. Ineffizienz bedeutet, dass bei voller Ausnutzung der vorhandenen Informationen Parameterschätzungen mit geringerer Streuung erreichbar wären. Verzerrte und inkonsistente Standardfehler führen zu falschen Schlüssen hinsichtlich der Signifikanz von Parameterschätzungen, und auch mit wachsender Stichprobe verschwindet dieses Problem nicht.

Die räumliche Ökonometrie ist ein Spezialgebiet der Ökonometrie, das sich mit der Behandlung räumlicher Wechselbeziehungen (räumlicher Korrelation) und Raumstrukturen (räumlicher Heterogenität) befasst. Die räumliche Ökonometrie hat sich in den letzten 30 Jahren rasant entwickelt und wird aller Voraussicht nach noch an Bedeutung zunehmen.

2.5.2 Grundlagen

Methoden der räumlichen Ökonometrie werden z. B. relevant, wenn die Beobachtungen für eine Gebietseinheit systematisch mit den Beobachtungen für solche Gebiets-einheiten korrelieren, die durch Nähe, Nachbarschaft oder Intensität des Austausches mit der betrachteten eng verbunden sind. Man spricht in diesem Zusammenhang von räumlicher Abhängigkeit (*spatial dependence*). Dabei werden grundsätzlich zwei Formen unterschieden. Die erste bezieht sich auf die räumlichen Interdependenzen in den untersuchten (abhängigen) Variablen. Beispielsweise könnten bei der Untersuchung regionalen Innovationsverhaltens Ausstrahlungseffekte von Interesse sein. So lässt sich etwa die Hypothese formulieren, dass Nachbarregionen einer hoch innovativen Region tendenziell ebenfalls überdurchschnittlich innovativ sind. Die zweite

Form der räumlichen Abhängigkeit bezieht sich auf die Korrelation in den Fehlern eines Schätzmodells. Dies kann durch räumliche Zusammenhänge in nicht erfassten oder nicht beobachtbaren Variablen verursacht sein. Ein Beispiel bietet die Untersuchung von Determinanten der Miethöhe mit Hilfe eines ökonomischen Modells. Dabei werden üblicherweise Variablen wie Gebäudealter, Wohnungsgröße und Ausstattungsmerkmale herangezogen. In solchen Ansätzen können räumliche Korrelationen in den Störtermen in Form von Nachbarschaftseffekten auftreten, die durch Ausstrahlungseffekte guter oder schlechter Wohnviertel begründet sind.

Eine ganz andere Begründung, Methoden der räumliche Ökonometrie zum Einsatz zu bringen, besteht in der räumlichen Heterogenität. Diese liegt dann vor, wenn in unterschiedlichen Teilräumen die Schätzkoeffizienten oder die Residualstreuung eines ökonomischen Modells systematisch variieren. Beispielsweise können sich verschiedene Regionstypen unterschiedlich verhalten.

2.5.3 Raumgewichte und Raumverzögerungen

Räumliche Abhängigkeiten bedeuten, dass die Kovarianzen von Variablen Informationen enthalten, die sich durch raumstrukturelle Gegebenheiten (Distanz, Nachbarschaft usw.) erklären lassen. Nur bei Vorliegen von kombinierten Längs-/Querschnittsdaten sind genügend Beobachtungen vorhanden, um die Varianz-/Kovarianzmatrix empirisch zu schätzen. In diesem Fall könnten auch mit Hilfe nicht-parametrischer Verfahren Raumstrukturen aufgezeigt werden. Typischerweise wird jedoch a priori eine bestimmte Struktur der räumlichen Abhängigkeiten vorgegeben. Dies geschieht mit Hilfe von Interaktionsgewichten w_{ij} , die die Intensität der Beziehung zwischen zwei Gebietseinheiten i und j beschreiben. Diese Interaktionsgewichte können auch auf einfachen Indikatoren der Nachbarschaft oder reinen Distanzmaßen beruhen. Dabei werden Interaktionen einer Gebietseinheit mit sich selbst ausgeschlossen ($w_{ii} = 0$ für alle i) und eine Normierung in der Regel so vorgenommen, dass die Summe der Interaktionsgewichte für jede Gebietseinheit i den Wert Eins annimmt ($\sum_{j=1}^N w_{ij} = 1$ für alle i).

Die Interaktionsgewichte werden in einer Matrix \mathbf{W} angeordnet. Aufgrund ihrer Konstruktion weist diese *räumliche Gewichtungsmatrix* Nulleinträge auf der Diagonalen und Zeilensummen von Eins auf.

Es sei \mathbf{y} ein $(N \times 1)$ -Vektor, der Werte für eine für N Gebietseinheiten im räumlichen Querschnitt erhobene Variable zusammenfasst, beispielsweise das verfügbare Einkommen der Bundesländer. Es existiere ferner eine $(N \times N)$ -Matrix \mathbf{W} , die die Interaktionsintensität zwischen den Bundesländern beschreibt. Das Produkt $\mathbf{W} \cdot \mathbf{y}$ enthält dann für alle Bundesländer den nach dem Umfang der Interaktionen gewichteten Durchschnitt des verfügbaren Einkommens der Partnerländer. Wenn sich beispielsweise das i -te Element $[\mathbf{W} \cdot \mathbf{y}]_i = \sum_{j=1}^{16} w_{ij} y_j$ auf Bayern bezieht, so enthält dieses einen gewichteten Durchschnitt der verfügbaren Einkommen in den außerbayerischen Bundesländern. Typischerweise wäre das Gewicht des Nachbarlandes Baden-Württemberg dabei natürlich deutlich höher als das von Mecklenburg-Vorpommern.

Für die konkrete Wahl von \mathbf{W} sind zwei Dinge entscheidend: Auf der einen Seite soll die Matrix die Interaktionen zwischen den Gebietseinheiten plausibel abbilden, auf der anderen Seite ist jedoch als methodisches Erfordernis zu beachten, dass \mathbf{W} strikt

exogen in Bezug auf das betrachtete Modell sein muss. Strikt exogen sind beispielsweise Entfernungsmaße oder Nachbarschaftsindikatoren. Viele raumwirtschaftliche Studien verwenden deshalb solche Maße.

2.5.4 Modellklassen bei räumlichen Abhängigkeiten

Einige Begriffe der räumlichen Ökonometrie sind in Anlehnung an die Zeitreihenanalyse gebildet worden. Allerdings besteht ein wesentlicher Unterschied zum Zeitreihenkontext darin, dass eine lineare Aneinanderreihung räumlicher Beobachtungen in der Regel nicht möglich ist, ein Vor- oder Zurückgehen wie entlang eines Zeitstrahls damit nicht in Frage kommt. Die Analogie trägt deshalb nur sehr begrenzt.

Unterschieden wird zwischen der räumlichen Entsprechung eines autoregressiven Modells (*spatial autoregressive, SAR*) und dem gleitenden Durchschnittsprozess (*spatial moving average, SMA*). Das SAR-Modell weist folgende Struktur auf:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad (7)$$

Hierbei ist \mathbf{y} ein $(N \times 1)$ -Vektor der abhängigen Variable, \mathbf{X} eine $(N \times k)$ -Matrix von präterminierten Variablen und $\boldsymbol{\beta}$ ein passender Koeffizientenvektor. Der auf den Wertebereich zwischen -1 und $+1$ beschränkte Parameter ρ wird als der *räumliche Verzögerungsparameter* bezeichnet. Für den Vektor der Zufallstörungen $\boldsymbol{\varepsilon}$ wird angenommen

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0 \text{ für } i \neq j \quad \text{und} \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2).$$

Im Falle $\rho \neq 0$ führt eine traditionelle Kleinst-Quadrate-Schätzung, die den Term $\rho \mathbf{W}\mathbf{y}$ vernachlässigt, zu ineffizienten Schätzern von $\boldsymbol{\beta}$ sowie verzerrten und inkonsistenten Schätzern der Kovarianzmatrix des geschätzten $\boldsymbol{\beta}$.

Wird die Gleichung (7) nach der abhängigen Variable aufgelöst, so ergibt sich

$$\mathbf{y} = \tilde{\mathbf{X}}\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}},$$

wobei $\tilde{\mathbf{X}}$ und $\tilde{\boldsymbol{\varepsilon}}$, die so genannten *räumlich gefilterten* Variablen, definiert sind als

$$\tilde{\mathbf{X}} := (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \quad \text{und} \quad \tilde{\boldsymbol{\varepsilon}} := (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}.$$

Das SMA-Modell als Alternative zum SAR-Ansatz lässt sich wie folgt angeben:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{mit} \quad \boldsymbol{\varepsilon} = \lambda \mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^*, \quad (8)$$

wobei $\boldsymbol{\varepsilon}^* = (\mathbf{I} - \lambda \mathbf{W})\boldsymbol{\varepsilon}$ unabhängig und identisch verteilt ist mit Varianz $\sigma_{\varepsilon^*}^2$. Es seien räumlich gefilterte Variablen definiert als

$$\mathbf{y}^* = (\mathbf{I} - \lambda \mathbf{W})\mathbf{y} \quad \text{und} \quad \mathbf{X}^* = (\mathbf{I} - \lambda \mathbf{W})\mathbf{X}.$$

Für das SMA-Modell (8) ergibt sich dann die Formulierung

$$\mathbf{y}^* = \mathbf{X}^* \boldsymbol{\beta} + \boldsymbol{\varepsilon}^*.$$

Die Modelle lassen sich mit einem Maximum-Likelihood-Ansatz schätzen, bei dem auch die Parameter ρ bzw. λ bestimmt werden.

Hinzuweisen ist auf einen gravierenden Strukturunterschied zwischen SMA- und SAR-Ansatz. Im ersten Fall wird bei der räumlichen Filterung die Inverse einer mit

Hilfe der Gewichtungsfaktoren gebildeten Matrix $(\mathbf{I} - \rho\mathbf{W}\mathbf{y})^{-1}$ verwendet. Dies bedeutet, dass neben den in \mathbf{W} erfassten direkten Interaktionen auch indirekte räumliche Abhängigkeiten ins Spiel kommen. Diese Verkettung der Effekte ist mit einem Multiplikatorprozess vergleichbar, in dem ein Ausgangsimpuls letztlich das ganze System der betrachteten Regionen tangiert. Im Gegensatz dazu beschränkt sich die Reichweite der räumlichen Interaktion im SAR-Ansatz auf die a priori in \mathbf{W} vorgegebenen direkten Abhängigkeiten.

2.5.5 Tests

Ein genereller Test auf das Vorliegen räumlicher Abhängigkeiten wurde bereits von Patrick A.P. Moran (1950) konzipiert. Der so genannte *Moran-I-Test* kann als Analogon zum Durbin-Watson-Test im Zeitreihenkontext betrachtet werden. Im Fall der üblichen Normierung der räumlichen Gewichtungsmatrix berechnet sich die Testgröße als

$$I = \hat{\mathbf{e}}' \mathbf{W} \hat{\mathbf{e}} / \hat{\mathbf{e}}' \hat{\mathbf{e}},$$

wobei $\hat{\mathbf{e}}$ den Vektor von Residuen der Kleinst-Quadrate-Schätzung bezeichnet. Weiterhin lassen sich mit Hilfe von Likelihood-Ratio-Tests die Hypothesen $\rho = 0$ und $\lambda = 0$ gegen die Alternative einer existierenden räumlichen Abhängigkeitsbeziehung ($\rho \neq 0$ bzw. $\lambda \neq 0$) auf einfache Weise testen.

2.6 Zusammenfassung

In der Regionalökonomik steht man häufig vor der Aufgabe, Austauschbeziehungen zwischen Regionen wie z. B. Pendel- oder Handelsströme in einem Modell abzubilden. Dafür hat sich das Gravitationsmodell sehr gut bewährt. Es führt das Niveau der Ströme auf Massevariablen an der Quelle und am Ziel sowie auf Distanzen zwischen Quelle und Ziel zurück und lässt sich einfach schätzen. Das Modell kann erweitert werden, indem Intervening Opportunities berücksichtigt werden.

Ein anderer gängiger Untersuchungsgegenstand sind Daten zur Wirtschaftsaktivität, differenziert nach Branchen und Regionen. Die Entwicklung im Zeitablauf lässt sich in Standort- und Struktureinflüsse zerlegen. Neben der traditionellen Shift-Share-Analyse kommt dafür eine Zerlegung mit Hilfe der Regressionsanalyse zum Einsatz. Man interessiert sich auch für die Frage, in welchem Maße sich Regionen spezialisieren und in welchem Maße sich Branchen auf bestimmte Regionen konzentrieren. Zur Messung stehen diverse Indikatoren, darunter die Indizes nach Krugman bzw. Florence, Herfindahl und Gini zur Verfügung.

Schließlich kommt in der Regionalökonomik der Regressionsanalyse von räumlichen Querschnittsdaten überragende Bedeutung zu, um Theorien zu testen sowie Zusammenhänge zu quantifizieren. Dabei muss man die Besonderheiten beachten, die durch die räumlichen Wechselwirkungen begründet sind. Dafür stehen Standardmodelle wie das SAR-Modell und das SMA-Modell zur Verfügung.

Die neuere Literatur hat sich darüber hinaus intensiv mit so genannten Paneldaten befasst, in denen die Beobachtungen sowohl eine Raum- als auch eine Zeitdimension

haben. Hier ergeben sich schwierige ökonometrische Fragen, weil die Korrelationsstrukturen in Zeit und Raum komplex sein können. Dazu muss auf die Literatur verwiesen werden. Verzichten mussten wir auch auf Modelle mit diskreten abhängigen Variablen wie z. B. Haushaltsentscheidungen über Wohnort- oder Verkehrsmittelwahl, obgleich sie in der Regionalwissenschaft von großer Bedeutung sind.

Literaturhinweise zu Kapitel 2

Ein Klassiker der regionalökonomischen Methodenlehre ist *Isard* (1960). *Anselin* (1988) sowie *LeSage, Pace* (2009) bieten einen Überblick über Methoden der räumlichen Ökonometrie. Über die Analyse von Panel-Daten, darunter auch solchen mit räumlicher Dimension, informiert *Baltagi* (2005). Populäre Computer-Programme zur Auswertung räumlicher Daten sind *GeoDa* (*Anselin*, 2005), die MATLAB-Toolbox von *LeSage* und die Spatial-Econometrics-Funktionen in der Computer-Sprache „R“ (*Bivand*, 2002). Im Statistik-Programm STATA werden von Nutzern diverse Routinen mit ähnlicher Funktionalität angeboten. *Combes, Overman* (2004) geben einen kritischen Überblick über Verfahren zur Spezialisierungs- und Konzentrationsmessung. *Aiginger, Leitner* (2002), *Amiti* (1999), *Brühlhart* (1998) und *Kim* (1995, 1997) liefern empirische Ergebnisse für die Vereinigten Staaten und Europa. Der regressionsanalytische Ansatz zur Shift-Share-Analyse findet sich bei *Möller, Tassinopoulos* (2000).

3. Grundzüge der Raumstruktur

Johannes Bröcker und Hayo Herrmann

Die Regionalökonomik befasst sich mit dem Wirtschaftsgeschehen eines regional differenzierten Gesamttraumes, insbesondere mit der Verteilung wirtschaftlicher Aktivitäten zwischen den Regionen und mit ihren räumlichen Wohlstands- und Entwicklungsunterschieden. Die Bedeutung der Regionalökonomik ergibt sich daraus, dass Bevölkerung, Wirtschaftstätigkeit und Einkommen räumlich äußerst ungleich verteilt sind. Diese Beobachtung gilt vor allem im weltweiten Maßstab, aber auch für Europa und Deutschland.

Über viele Jahrhunderte hat sich eine differenzierte Wirtschaftslandschaft herausgebildet. Die wichtigsten Merkmale, mit denen sie sich charakterisieren lässt, sind die Bevölkerung, die Betriebe mit ihrer wirtschaftlichen Aktivität und den dort tätigen Arbeitskräften, es sind Institutionen und Infrastruktureinrichtungen und die im Zuge der wirtschaftlichen Tätigkeit entstehenden Einkommen. Aus der *Lokalisation*, also der räumlichen Verankerung dieser Merkmale, entstehen Verflechtungen und Mobilitäten zwischen den Akteuren.

Die regionalen Unterschiede, die man bei der Beschreibung der Wirtschaftslandschaft entdeckt, sind häufig über längere Zeiträume recht stabil und weisen eine Reihe auffälliger Muster und Regelmäßigkeiten auf.

Selbst in einem im Weltmaßstab dicht besiedelten und von hoher Wirtschaftskraft geprägten Land wie Deutschland offenbaren sich erhebliche regionale Unterschiede, wenn man das Land z. B. in seine über 400 Kreise und kreisfreien Städte unterteilt. So schwankt die Bevölkerungsdichte (Einwohner pro Quadratkilometer) zwischen 40 in Mecklenburg-Vorpommern (Kreis Müritz) und über 4.000 (Stadt München), die jährliche Wirtschaftsleistung (Bruttoinlandsprodukt) je Arbeitsplatz lag im Jahr 2007 zwischen 40 Tsd. Euro in Vorpommern und 115 Tsd. Euro im Landkreis München, und das verfügbare pro Kopf-Einkommen der Haushalte bewegte sich im Jahr 2007 zwischen 13 Tsd. Euro in Vorpommern (Kreis Uecker-Randow) und knapp 30 Tsd. Euro im Umland von München (Kreis Starnberg).

Beim Bruttoinlandsprodukt pro Kopf bewegten sich die Werte in Deutschland 2007 (nach Kreisen und kreisfreien Städten) zwischen 14 Tsd. und 84 Tsd. Euro, d. h. der höchste Wert war mehr als sechsmal so hoch wie der kleinste. Im Gebiet der aktuellen Europäischen Union (27 Staaten) waren die Unterschiede 2007 allerdings zwischen den über 1300 (NUTS 3-) Regionen weitaus stärker: Der Durchschnittswert (in *Kaufkraftstandards* KKS) der zehn reichsten Regionen war 14 mal so hoch wie der Durchschnitt der zehn ärmsten Regionen (allesamt aus Rumänien und Bulgarien). Kaufkraftstandards (KKS) sind fiktive Währungseinheiten, die die Kaufkraftunterschiede aufgrund verschiedener Preisniveaus zwischen den europäischen Regionen eliminie-