

† Münsteraner Beiträge
zur Cliometrie und
quantitativen
Wirtschaftsgeschichte

herausgegeben von

Richard Tilly

Institut für Wirtschafts- und Sozialgeschichte

Westfälische Wilhelms-Universität Münster

Band 1

Dr. Harald Frank

Regionale Entwicklungsdisparitäten im
deutschen Industrialisierungsprozeß 1849-1939

Eine empirisch-analytische Untersuchung

LIT

LIT

REGIONALE ENTWICKLUNGSDISPARITÄTEN IM
DEUTSCHEN INDUSTRIALISIERUNGSPROZESS

1849 – 1939

Eine empirisch-analytische Untersuchung

INAUGURAL – DISSERTATION

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft

durch den

Fachbereich Wirtschafts- und Sozialwissenschaften
der Westfälischen Wilhelms-Universität Münster

vorgelegt von
Harald Frank
aus Bochum

Juni 1993

Inhalt	
A. Zusammenfassende Betrachtungen zu Modernität und Lebensbedingungen.....	78
B. Zur Konvergenz/Divergenz – Diskussion über Einkommensentwicklung, Wirtschaftsstruktur und Lebensbedingungen.....	80
1. Darstellung verschiedener Ansätze.....	80
2. Eigene Berechnungen.....	84
C. Die Entwicklungsdynamik der Regierungsbezirke.....	92
Zweiter Teil	
Interregionaler Gütertausch als Determinante der Entwicklungstendenzen?	
I. Welfare, Außenhandel und Terms of Trade.....	100
A. Wohlfahrtseffekte des Außenhandels.....	100
B. Wirtschaftspolitische Relevanz des Austauschverhältnisses.....	108
C. Ausgewählte Terms of Trade-Konzepte.....	113
1. Commodity Terms of Trade.....	114
2. Income Terms of Trade.....	116
3. Single factorial Terms of Trade.....	118
II. Anwendung der Terms of Trade-Konzepte auf die regionalwirtschaftliche Fragestellung.....	121
A. Auswahl der Regionen.....	121
B. Vorstellung der Daten.....	122
1. Preisdaten.....	122
2. Mengendaten.....	126
a) Daten der Eisenbahnverkehrsstatistik.....	126
b) Daten der Binnenschifffahrtsstatistik.....	128
Erster Teil	
Entwicklungstendenzen in deutschen Regionen	
I. Einleitung.....	2
II. Forschungsstand – Ein Überblick.....	6
III. Vorstellung der Analysemethode.....	27
A. Zur Problematik der Raumabgrenzung.....	27
1. Systematisierungskonzept von Siebert.....	28
2. Auswahl der Aggregationsniveaus und Darlegung der Daten.....	30
3. Der Clusteralgorithmus.....	32
4. Die Wirtschaftsregionen.....	39
5. Exkurs: Darstellung der 7-Cluster-Lösung.....	45
B. Zur Problematik der Indikatorauswahl.....	50
1. Der Modernisierungsindex zur Messung der strukturellen Modernität der Regionen.....	50
2. Die Fäimdung nach einem Welfareindikator.....	56
a) Anmerkungen zu Lohndaten.....	58
b) Anmerkungen zu Körpergrößendaten.....	61
c) Zur Eignung von Mortalitätsdaten.....	66
3. Die regionalen Lebensbedingungen.....	73
4. Exkurs zur Mortalität.....	76
IV. Konsequenzen der angewandten Methodik.....	78

c) Daten der Seeschifffahrtsstatistik und Konsolidierung der Mengendaten.....	130	B. Kritik der North-Thesen.....	185
3. Produktivitätsdaten.....	140	C. Gültigkeit der Exportbasisthese für Deutschland.....	189
C. Empirische Terms of Trade-Ergebnisse.....	146	III. Zusammenfassung.....	193
1. Commodity Terms of Trade.....	148	IV. Ausblick.....	196
2. Single factorial Terms of Trade und Produktivitätsanstieg.....	151	Anhänge.....	I
3. Income Terms of Trade und Exportmengenentwicklung.....	152	Literaturverzeichnis.....	LXXVI
D. Interpretation der empirischen Terms of Trade-Ergebnisse unter Berücksichtigung des Umfangs des Außenhandelssektors.....	156	Quellenverzeichnis.....	LXXXVII
1. Die ländlichen Beispielregionen.....	157		
2. Die schwerindustriellen Beispielregionen.....	158		
3. Die regionalen Handelsbilanzalden und der Sonderfall Berlin.....	159		
E. Prüfen der Ergebnisse auf Zufälligkeit.....	162		
1. Die Terms of Trade auf reduzierter Güter- und Mengebasis im Raum östlich der Oder.....	164		
2. Die Terms of Trade auf reduzierter Güter- und Mengebasis in Oppeln.....	168		
3. Die Terms of Trade auf reduzierter Güter- und Mengebasis für das Ruhrrevier.....	170		
4. Erneuter Interregionenvergleich.....	172		
		Dritter Teil	
		Integration der Ergebnisse und Schluß	
I. Verknüpfungsversuch der Ergebnisse des ersten und zweiten Teils.....	176		
II. Diskussion der North-Thesen.....	182		
A. Darstellung der Argumentation von D.C. North.....	182		

III. Vorstellung der Analysemethode

A. Zur Problematik der Raumabgrenzung

Die Industrialisierung in Deutschland war kein Prozeß, der sich gleichförmig über das Land verbreitete. Die Entwicklungsschwerpunkte konzentrierten sich ab etwa der Mitte des 19. Jahrhunderts nicht nur auf bestimmte Branchen – Führungssektoren –, sondern auch auf Räume. Als Beispiel sei die Schwerindustrie des Ruhrgebiets genannt. In diesem Sinn ist die Industrialisierung als regionales Phänomen zu betrachten. Ein räumlich ungleichgewichtiger Entwicklungsprozeß vollzog sich, von dem manche Gebiete eine lange Zeit völlig unberührt blieben.

"... Eine Fülle von Modellen, wie sie etwa von Domar, Duesenberry, J. Robinson, Solow, Swan, Meade, Kaldor und Arrow konstruiert wurden, erheben die Expansion zum beherrschenden Erklärungsziel der ökonomischen Theorie. Diese Explikationsversuche richten ihr Augenmerk jedoch auf das Wachstum einer Punktwirtschaft – auf ein "Wunderland ohne räumliche Dimension" (*Isard*). Im Gegensatz zu einer solchen global orientierten Wachstumstheorie ist die Analyse der Expansion von Teilräumen einer Volkswirtschaft noch ein unterentwickeltes Gebiet der ökonomischen Theorie, stellen doch die Theorie der Wachstumspole in der französischen Literatur, die recht farblose Sektorthorie und die für sehr kleine Raumeinheiten anwendbare Theorie der Export-Basis die bisher einzigen Ansatzpunkte zur Erklärung regionalen Wirtschaftswachstums dar."⁹⁸

Für Siebert werden somit "die Faktoren Raum und Zeit ... meistens überhaupt nicht, oft nur versuchsweise und nie gleichzeitig berücksichtigt".⁹⁹ Insbesondere aus dieser Kritik entwickelte Siebert sein dynamisches regionales Wachstumsmodell. Ein derartiges Modell hat jedoch sehr hohe Datenanforderungen und damit realistischerweise keine Chance, in der Wirtschaftsgeschichte zum Einsatz zu kommen.

Der 2. Teil der vorliegenden Arbeit, der den interregionalen Gütertausch thematisiert, wird darum mit der komparativ-statischen Analyse anstelle einer dynamischen auskommen müssen. Der Raumbezug entsteht dadurch, daß Variablenwerte für bestimmte Regionen ermittelt und interregionale Prozesse im Rahmen eines Außenwirtschaftsmodells untersucht werden. Die Dimensionen Raum und Zeit sind hier also nur

⁹⁸H. Siebert, *Zur Theorie des regionalen Wirtschaftswachstums*, Tübingen 1967, Vorwort S. VII.

⁹⁹H. Siebert, a.a.O., S. 1.

fragmenthaft oder – um es im Siebertschen Terminus zu formulieren – "versuchsweise" integriert.

Methodische Probleme existieren jedoch bei der Abgrenzung von Räumen, und diese Abgrenzung von Wirtschaftsräumen ist für die vorliegende Untersuchung nötig und von besonderem Interesse. Eingehende Diskussionen dazu sind an anderer Stelle bereits umfassend thematisiert worden.¹⁰⁰

1. Systematisierungskonzept von Siebert

Um einen Überblick über denkbare Kriterien zur Einteilung eines Gesamttraumes in Regionen zu gewinnen, wird das theoretische Konzept von Siebert – insbesondere im Hinblick auf Operationalisierbarkeit und technische Verwertbarkeit für wirtschaftshistorische Analysen – dargestellt.¹⁰¹ Prinzipiell existieren danach drei Kriterien zur Bestimmung von Teilräumen, wobei sich die ersten beiden besonders an Ökonomen wenden und das dritte Kriterium auch von Historikern und Geographen mit unterschiedlichstem Sinngehalt zur Analyse herangezogen wird.

1. Interdependenz
2. Einheitlichkeit der Wirtschaftspolitik
3. Homogenität

zu 1. Interdependenzkriterien: Es geht um die Integration bzw. die Interdependenz aller Austauschbeziehungen, also um eine Verflechtung der Sektoren bzw. Branchen untereinander und mit dem Konsumsektor bei geringen Transaktionskosten. In diesem Zusammenhang wird Region von Leontieff als eine "Kombination von wirtschaftlichen Aktivitäten" verstanden.¹⁰² Diese Definiti-

¹⁰⁰Vgl. u.a. H. Siebert, *Zur Theorie des regionalen Wirtschaftswachstums*, Tübingen 1967, S. 6 ff; R. Fremdling, T. Pierenkemper, R. Tilly, *Regionale Differenzierung in Deutschland als Schwerpunkt wirtschaftshistorischer Forschung*, in: ders. (Hrsg.), *Industrialisierung und Raum*, Stuttgart 1979, S. 9 ff.; S. Pollard, *Einleitung*, in: ders. (Hrsg.), *Region und Industrialisierung*, Göttingen 1980, S. 11 ff; P. Steinbach, *Zur Diskussion über den Begriff "Region" - eine Grundsatzfrage der modernen Landesgeschichte*, in: Hessisches Jahrbuch für Landesgeschichte, 31 (1981), S. 208; H. Pohl, *Einführung*, in: ders. (Hrsg.), a.a.O., S. 8-11.

¹⁰¹H. Siebert, a.a.O., S. 15 ff.

¹⁰²W. Leontieff, *Interregionale Beziehungen wirtschaftlicher Aktivitäten*, in: *Probleme des räumlichen Gleichgewichts in der Wirtschaftswissenschaft*, hrsg. von W.G. Hoffmann, Schriften des Vereins für Socialpolitik, N.F. Bd. 14, Berlin 1959, S. 50.

on einer Austausch- oder Handelsregion¹⁰³ wird in der Raumwirtschaftstheorie oft bevorzugt, ist aber für historische Untersuchungen aufgrund der extrem hohen Datenanforderungen nicht praktikabel.

zu 2. Hier wird unter Region ein von der Wirtschaftspolitik festgelegter Teilraum verstanden.

Diese an Verwaltungsregionen anknüpfende Definition bringt bei Projektion auf das 19. Jahrhundert Probleme mit sich. Datentechnisch wäre dieser Weg zwar gangbar, aber es wurde im allgemeinen aufgrund liberaler Gesinnungen von einer aktiven Wirtschaftspolitik – insbesondere auf regionaler Ebene – abgesehen. Die nationale Wirtschaftspolitik sorgte für üblicherweise als günstig beurteilte Rahmenbedingungen, die sich natürlich regional unterschiedlich und verwaltungsgrenzenübergreifend auswirkten. Andererseits würden sozio-ökonomische Strukturen ausgeblendet, die bekanntlich gleichfalls an politischen Grenzen nicht halt machen.

zu 3. Darum wenden wir uns dem Homogenitätskriterium zu, von dem drei mögliche Ausprägungen vorgestellt werden.

Geographisch physikalische Homogenität ist hier unzweckmäßig, weil die daraus entstehenden natürlichen Regionen Landschaften sind, die ökonomische Strukturen und Prozesse nicht deutlich werden lassen.

"Auch die Einteilung nach der Einheitlichkeit des sozialen Systems (z.B. historische Gemeinsamkeiten, institutionell-administrative Einheitlichkeit) kann keine Anwendung finden,

¹⁰³Dieser Begriff von H.C. Binswanger umfaßt nach Siebert nicht nur die Konzepte von Christaller und Lösch, die durch zentrale Orte und das jeweils umgebende Hinterland oder die dazugehörigen Marktnetze gekennzeichnet sind, sondern auch die sogenannte "Polarisationsregion, die Boudeville als ein Gebiet formuliert, in dem die Intensität der internen Waren- und Dienstleistungsströme an jedem Punkt größer ist als die Intensität der externen Ströme."

Vgl. H. Siebert, a.a.O., S. 19;

H.C. Binswanger, *Das intra-regionale Gleichgewicht - zur Integration von Standorttheorie und Theorie des internationalen Handels*, in: Schweizerische Zeitschrift für Volkswirtschaft und Statistik, Bd. 97 (1961), S. 144;

W. Christaller, *Die zentralen Orte in Süddeutschland, Eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verteilung und Entwicklung der Siedlungen mit städtischen Funktionen*, Jena 1933;

A. Lösch, *Die räumliche Ordnung der Wirtschaft*, Jena 1944.

J.R. Boudeville, *L'économie régionale - Espace Opérationnel*, in: "Cahiers de l'Institut de Science Economique Appliquée", Serie L, No. 3, Paris 1958.

da die Frage, wie diesem Kriterium ökonomischer Sinngehalt zu geben ist, hier nicht gelöst werden kann."¹⁰⁴

Darum erscheint es in unserem Zusammenhang am sinnvollsten, auf die Homogenität der Produktionsstruktur abzustellen. Durch die Abgrenzung nach Ähnlichkeiten in der Produktion gelangt man zu landwirtschaftlichen Regionen, Industrieregionen oder zu ähnlichen sektorspezifischen oder durch ein bestimmtes 'Sektorenmix' charakterisierbaren Räumen. "Für den Zweck der Untersuchung des Wachstums scheint es am nützlichsten zu sein, eine Region durch die relative Homogenität ihrer Wirtschaftsstruktur zu definieren."¹⁰⁵

2. Auswahl der Aggregationsniveaus und Darlegung der Daten

Die o.g. Forderung nach wirtschaftsstruktureller Homogenität sei ... "jedoch kaum zu realisieren, da empirische Daten zumeist nur aggregiert vorliegen. Eine Näherungslösung dieses Problems besteht u. E. darin, möglichst kleine administrative Einheiten, für die Daten zu mobilisieren sind, als Region zu definieren, weil auf einer geringeren Aggregationsebene die Inhomogenität der Wirtschaftsstruktur vermutlich am wenigsten ausgeprägt ist. Diese kleinen Regionen lassen sich dann nachträglich zu größeren Einheiten zusammenfassen, ohne daß man dabei gezwungen wäre, den Grenzen der größeren administrativen Ebene zu folgen. Das bedeutet für die praktische Forschung, die Analyse bis auf *Kreis-* und *Gemeindeebene* zu disaggregieren."¹⁰⁶ Die hinter dem Homogenitätskonzept stehende statische Betrachtungsweise ist bei langen Untersuchungszeiträumen jedoch hinderlich, für das 19. Jh. bzw. unseren Untersuchungszeitraum aber kaum relevant, weil es – soviel sei vorweggenommen – keine so großen regional unterschiedlichen Verschiebungen der Produktionsstruktur gab.¹⁰⁷ Je nach Region mehr oder weniger schnell verlagerte sich die Erwerbsstruktur zum sekundären oder tertiären Sektor.

Das statistische Material der Berufs- und Gewerbezahlungen steht hierfür – sogar disaggregiert bis auf Kreisebene – zur Verfügung. Besondere Bedeutung muß dann – für die angestrebte Makroebene 'Deutschland' – der Zusammenfassung der wirtschaftlich homogenen 'Kleinräume' zu größeren Einheiten zu-

¹⁰⁴H. Siebert, a.a.O., S. 17.

¹⁰⁵R.H. Tilly/T. Pierenkemper/R. Fremdling, a.a.O., S. 17.

¹⁰⁶R.H. Tilly, T. Pierenkemper, R. Fremdling, a.a.O., S. 17 f.

¹⁰⁷Für die später einzuführende Analysemethode (Cluster) ist dieser Sachverhalt nicht problematisch, da Strukturänderungen berücksichtigt werden.

kommen. Auch wenn J. Schneider sicherlich nicht ohne Gründe davon spricht, daß "die Zuordnung einzelner Verwaltungsbezirke zu einem Wirtschaftsraum (...) wissenschaftlich nicht lösbar [ist] und es (...) immer Argumente für die eine oder andere Zuordnung geben" wird,¹⁰⁸ soll hier ein solcher Versuch unternommen werden.

Diese Studie bleibt durch die Disaggregation auf Regierungsbezirksebene eine Stufe früher stehen, als die theoretisch sicherlich sauberste Lösung einer Heranziehung der Kreis- oder Gemeindedaten.

Die Gründe dafür lagen neben dem leichten Zugriff auf Gehrmanns 'gleichnamig' gemachte Daten über die Beschäftigungsstrukturen in deutschen Regierungsbezirken zwischen 1849 und 1961 in einer nicht zu überblickenden Fülle von Gebietsänderungen im Lauf der Untersuchungsperiode. Bei der prinzipiell möglichen Disaggregation auf etwa 500 Kreise sind Gebietsveränderungen nur durch sehr großen rechnerischen Aufwand ausgleichbar, der in keinem Verhältnis zum Nutzen für die hiesige Fragestellung steht.

Der Datensatz (vgl. Anhang 2, Tabelle H) beinhaltet als Variable für die Zeitpunkte 1849, 1882, 1895, 1907 und 1939 die Beschäftigungsstrukturdaten der – incl. Berlin – 62 deutschen Regierungsbezirke. Jedoch stellte es sich als sinnvoll heraus, die Sektordaten von Landwirtschaft, Industrie und Handwerk, Handel und den Verkehrsdienstleistungen den Berufsdaten vorzuziehen und für die Analyse bereitzustellen. Aufgrund statistischer Zuordnungsprobleme sind nämlich die Berufe interregional nicht hinreichend genau abgegrenzt, so daß dann die Aggregatdaten der Sektoren aussagefähiger sind als disaggregierte Berufsdaten. Die bei den Berufs- und Gewerbezahlungen auch erhobenen Erwerbstätigen im öffentlichen Dienst bzw. im häuslichen Dienst stellten jeweils so heterogene Gruppen dar, daß eine sinnvolle Interpretation problematisch wäre. So umfaßt die Sektion 'öffentlicher Dienst' die Sparten Militär, Staats- und Gemeindedienst, Kirche, Erziehung und Unterricht, Gesundheitspflege, Privatgelehrte, Privatsekretäre, Schreiber usw., und Künstler aus Musik, Theater usw. Bei den häuslichen Diensten sind in der Regel Tagelöhner

¹⁰⁸J. Schneider, *Dynamik von Wirtschaftsräumen im Gefolge der europäischen Expansion und der Industrialisierung*, in: U. Bestmann u.a. (Hrsg.), *Hochfinanz, Wirtschaftsräume, Innovationen. Festschrift für W. v. Stromer*, Bd. II, Trier 1987, S. 868.

und andere Lohnarbeiten wechselnder Art mitverbucht.¹⁰⁹ Daß es zudem starke Überschneidungen zwischen der Berufsgruppe Landwirtschaft und der Gruppe der häuslichen Dienste und Tagelöhner gegeben hat, wird gesehen, bleibt jedoch unberücksichtigt.

Für die Clusteranalysen zum Finden von Wirtschaftsregionen wurde Berlin aus dem Datensatz herausgenommen, weil Ausreißer – also "Objekte, die im Vergleich zu den übrigen Objekten eine vollkommen anders gelagerte Kombination von Merkmalsausprägungen aufweisen ... – ... den Fusionierungsprozeß der übrigen Objekte stark beeinflussen und damit das Erkennen der Zusammenhänge zwischen den übrigen Objekten erschwert wird und Verzerrungen auftreten."¹¹⁰ Fehlende Werte und solche, die aufgrund von Gebietsänderungen unbrauchbar geworden sind und dabei nicht durch eine Neuzuteilung der beteiligten Kreise korrekt zu ermitteln waren, konnten durch Einfachregression plausibel geschätzt werden.¹¹¹ Ansonsten wurde die Gebietsstruktur von 1850 konstant gesetzt. Spätere Änderungen sind auf diese Struktur umgerechnet worden. Möglich war das, weil in den meisten Fällen einzelne Kreise unter den betroffenen Regierungsbezirken 'getauscht' oder 'abgegeben' wurden.¹¹²

3. Der Clusteralgorithmus

Bei der Zusammenfassung der Einzelregionen (also der Regierungsbezirke) zu homogenen größeren Einheiten, die unabhängig von der Verwaltungsstruktur sind, wurden Clusteranalysen durchgeführt.¹¹³ Diese Verfahren zur Datenstrukturierung können immer dann angewendet werden, wenn zu einer Reihe von Merkmalen (Variablen) Beobachtungen an verschiedenen Objekten vorliegen und die Objekte nach ihren Merkmalsausprägungen gruppiert werden sollen. Die Einteilung von Objekten in Gruppen (oder Cluster) erfolgt dabei so, daß die Ähnlichkeit zwischen den Objekten einer Gruppe möglichst groß, diejenige zwischen den Gruppen jedoch möglichst gering ist.¹¹⁴ Auch wenn

¹⁰⁹H. Kiesewetter, *Regionale Industrialisierung in Deutschland zur Zeit der Reichsgründung*, S. 50 f.

¹¹⁰K. Backhaus u.a., *Multivariate Analysemethoden*, Berlin u.a. 1990, S. 154 f.

¹¹¹Vgl. Anhang 2, IV.

¹¹²Vgl. Anhang 2, II mit den Tabellen F und G.

¹¹³Programmpaket 'Cluster', D. Steinhausen, Rechenzentrum Münster, vgl. D. Steinhausen, K. Langer, *Clusteranalyse*, Berlin, NY 1977.

¹¹⁴D. Hochstädter, U. Kaiser, *Clusteranalyse*, in: WISU 3/89, S. 157.

abweichende Erhebungssystematiken der Gewerbezahlungen nicht voll ausgleichbar waren, so ist das für Clusteranalysen relativ unerheblich, da die Distanzen der Regierungsbezirke untereinander in jedem Stichjahr verglichen werden, der Fehler dann für alle gilt und sich damit aufhebt. Aus dem gleichen Grund ist auch der Einfluß unterschiedlicher konjunktureller Situationen in den Analysejahren nicht so gravierend – außer bei ausgeprägten Regionalkonjunkturen.

Um allen Variablen das gleiche Gewicht bei der Clusteranalyse zu geben, war eine Z-Transformation der Daten nötig.¹¹⁵

Zur Erlangung einer größtmöglichen Gruppenhomogenität wurde der konservative WARD-Algorithmus auf euklidisch-quadratische Distanzen angewendet. Als 'konservativ' wird WARD bezeichnet, weil hier weder versucht wird, etwa gleichgroße Gruppen zu bilden (dilatierende Verfahren) noch ist dieser Algorithmus geneigt, zunächst möglichst wenige große Gruppen zu bilden, denen viele kleine gegenüberstehen (kontrahierende Verfahren).

Ohne im einzelnen darauf einzugehen, inwieweit sich WARD bei der Distanzbildung von anderen Verfahren unterscheidet, sollen doch Vorgehensweise und Ziele dieses Clusteralgorithmus bei der Fusion der Gruppen angeschnitten werden. Damit wird klar, warum WARD gerade für unsere Fragestellung interessant ist.

"Das Ward-Verfahren unterscheidet sich von den ... Linkage-Verfahren insbesondere dadurch, daß nicht diejenigen Gruppen zusammengefaßt werden, die die geringste Distanz aufweisen, sondern es werden die Objekte (Gruppen) vereinigt, die ein vorgegebenes Heterogenitätsmaß am wenigsten vergrößern. Das Ziel des Ward-Verfahrens besteht darin, jeweils diejenigen Objekte (Gruppen) zu vereinigen, die die Streuung (Varianz) in einer Gruppe möglichst wenig erhöhen. Dadurch werden möglichst homogene Cluster gebildet. Als Heterogenitätsmaß wird das Varianzkriterium verwendet."¹¹⁶ Die Fußnote 119 enthält die exakte Definition dieses Kriteriums.

Es zeigt sich, daß dieses Verfahren den gängigen Forderungen der Raumdefinition, die in sich homogene und gegen umliegende Regionen möglichst

¹¹⁵Transformation der Variablen auf Mittelwert 0 und Standardabweichung 1 mit $Z_k = (X_k - \bar{X}) / \sigma_x$

¹¹⁶Vgl. dazu K. Backhaus u.a., a.a.O. S. 141.

deutliche Abgrenzungen verlangen, entspricht.¹¹⁷ Eine Simulationsstudie konnte zeigen, daß WARD – im Vergleich zu anderen Algorithmen – "gleichzeitig sehr gute Partitionen findet und meistens die richtige Clusterzahl signalisiert".¹¹⁸

Als hierarchisch-agglomeratives Verfahren geht WARD von der feinsten Partition aus (jedes Objekt bildet eine Gruppe), fusioniert sukzessive Objekte bzw. Gruppen in jedem Verfahrensschritt und endet, wenn alle Objekte in einer Gruppe zusammengefaßt sind.

Fraglich ist nun, auf welcher Stufe die informativste Clusterstruktur gegeben ist. Dabei scheidet sowohl die Stufe 1 mit seiner Aussage 'jedes Objekt ist anders als die anderen' als auch die letzte Stufe, die signalisiert, 'alle untersuchten Objekte sind Regierungsbezirke' aus.

Ein Dendrogramm ist die graphische Darstellung des Fusionsprozesses im Zusammenhang mit einem Heterogenitätsmaß. Dort werden alle Clusterlösungen sichtbar und exakt nachvollziehbar.

Das in der Abbildung 1 vorgestellte Dendrogramm basiert auf der clusteranalytischen Verarbeitung des Datensatzes aus Tabelle H (vgl. Anhang 2). Es wurde durch das Programmpaket CLUSTER (Steinhausen und Langer, Rechenzentrum Universität Münster) mit der folgenden o.g. Konfiguration erstellt:

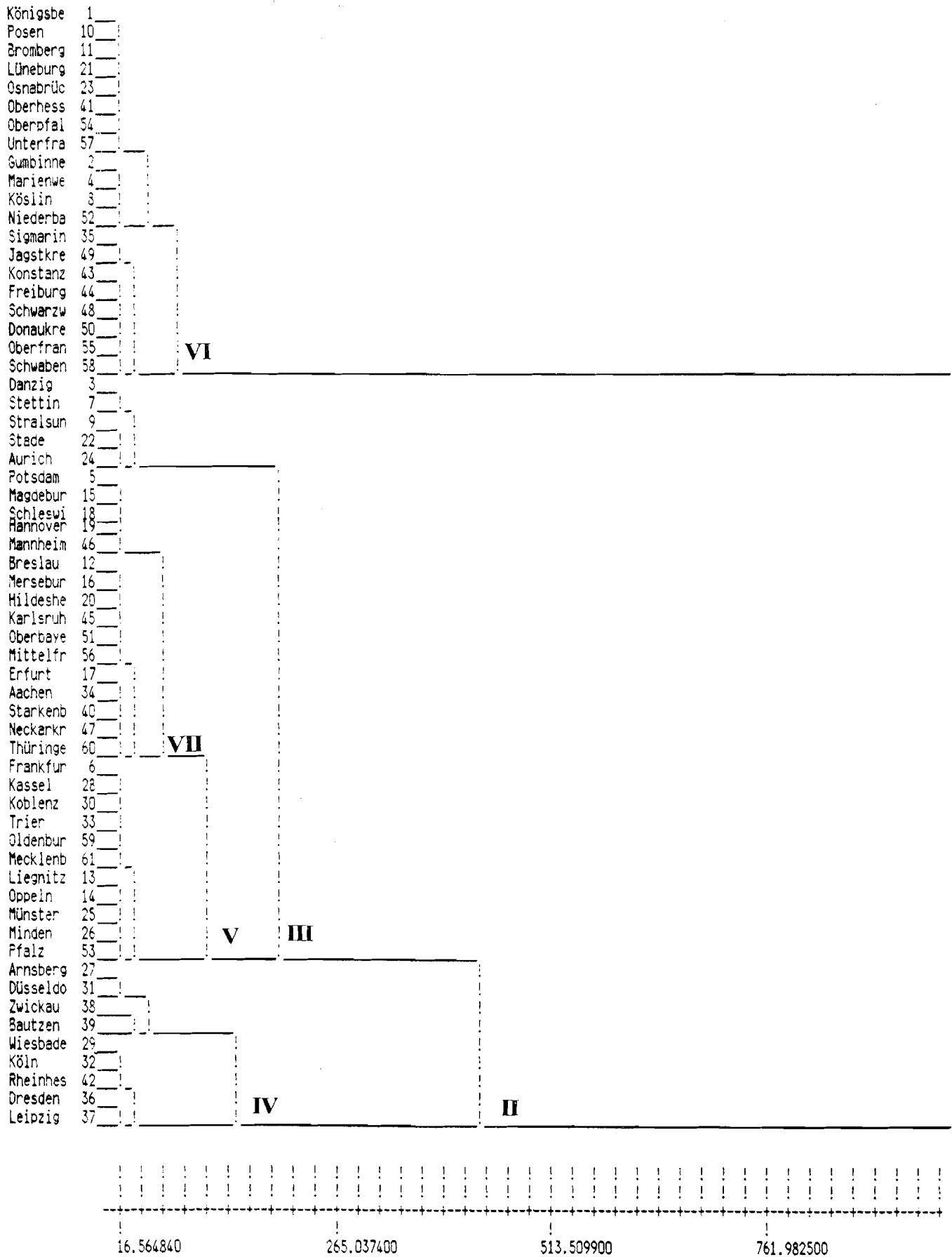
- WARD - Algorithmus,
- euklidisch-quadratische Distanzen und
- z-transformierte Variablen.

Um das Dendrogramm leicht interpretieren zu können, legt man einen senkrechten Schnitt bei der jeweils präferierten Clusteranzahl (römische Zahlen von rechts nach links fortlaufend) durch die Abbildung und verfolgt dann die 'Treppenstruktur' der Elemente (Regierungsbezirke) - von oben nach unten - bis zu den Schnittpunkten dieser Senkrechten mit einer waagerechten Linie. Bei den Elementen, die jeweils unter den 'Schnittpunktelementen' stehen, beginnt das neue Cluster.

¹¹⁷Vgl. H. Pohl, *Einführung*, in: ders. (Hrsg.), a.a.O., S. 9.

¹¹⁸S. Bergs, *Optimalität bei Cluster-Analysen*, Diss. Münster 1981, S. 96 f.

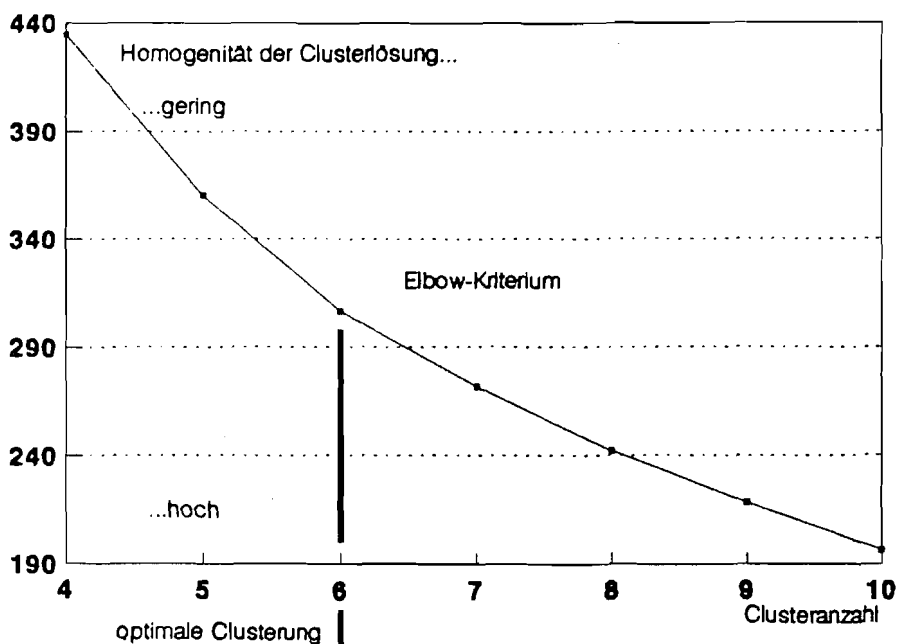
Abbildung 1: Dendrogramm zur Clusterung der Regierungsbezirke



Die Suche nach der optimalen Clusterzahl setzt eine Entscheidung über die sachlich relevante Gruppenanzahl voraus. Während eine Drei-Gruppenlösung ein für unser jetziges Untersuchungsziel lediglich triviales Ergebnis motiviert, indem ein ländliches, ein industrielles und ein durchschnittliches Cluster gebildet werden, treten bei sehr feinen Partitionen Interpretationsprobleme bezüglich der strukturellen Abgrenzung der einzelnen Cluster gegeneinander auf.

Mit Hilfe des 'Varianzkriteriums'¹¹⁹ werden die Homogenitäten aller relevanten Clusterlösungen ermittelt. Bei der graphischen Darstellung der Funktion des Varianzkriteriums in Abhängigkeit von der Clusteranzahl erscheint – allerdings nur bei genauer Betrachtung – eine sechser Gruppierung empfehlenswert, weil hier die vorhergehende grobere Teilung den relativ größten Heterogenitätszuwachs der Clusterstruktur zeigt und folgende feinere Teilungen nur relativ geringere Verbesserungen des Varianzkriteriums ergeben; der 'Ellbogen' bei Punkt 6 also den größten Knick bzw. kleinsten Winkel aufweist.

Abbildung 2: Der Wert des Varianzkriteriums



¹¹⁹Varianzkriterium für die gesamte Clusterstruktur. $Vk_g = \sum_{c=1}^C \sum_{k=1}^{n_c} \sum_{j=1}^J (X_{jkc} - \bar{X}_{jc})^2$

mit X_{jkc} : Wert der j-ten Variable für das k-te Objekt im Cluster C

und \bar{X}_{jc} : arithm. Mittel der j-ten Variable des Clusters C mit n_c Objekten.

Die Homogenität des einzelnen Clusters ist ablesbar an dem Varianzkriterium

$$Vk_c = \sum_{k=1}^{n_c} \sum_{j=1}^J (X_{jkc} - \bar{X}_{jc})^2$$

Aber auch, wenn die Bevorzugung der 6-Gruppen-Lösung nicht durch einen sehr starken Knick der Funktion des Varianzkriteriums gestützt wird, so ergeben andere Clusterlösungen bei hierarchisch-agglomerativen Verfahren wie WARD keine entscheidenden Änderungen, weil einmal zusammengefügte Gruppen hier in folgenden Fusionschritten nicht mehr auseinandergerissen werden.

In unserem Fall unterscheidet sich beispielsweise eine 5-Gruppen-Struktur von der hier präferierten 6-Cluster-Lösung durch die Verschmelzung des Clusters 3 (Potsdam u.a.) mit dem Cluster 4 (Frankfurt u.a.) der Tabelle 1.

Eine 7-Cluster-Struktur (vgl. 5. Exkurs: Darstellung der 7-Cluster-Lösung) würde c.g.s. die süddeutschen ländlichen Gebiete von den nord- und ostdeutschen Agrarbezirken trennen, also Cluster 1 der Tabelle 1 aufspalten.

Weil für uns ohnehin nur die räumlich zusammenhängenden Regierungsbezirke interessant sind, ist die Gruppen-Anzahl für unsere Analyse und unser Verfahren nicht überragend wichtig.¹²⁰

¹²⁰Würde hier mit partitionierenden Verfahren – z.B. dem Austauschverfahren – gearbeitet, so sind, gemessen am Varianzkriterium, in der Regel zwar etwas homogenere Lösungen möglich, die Ergebnisse aber kaum interpretierbarer, weil sie schon in Bezug auf kleine Veränderungen der Anzahl der Variablen, Objekte oder Gruppen nicht stabil waren.

Denn im Unterschied zu den hierarchischen Verfahren bleiben hier einmal konstruierte Cluster nicht zusammen, sondern es wird bei jeder vorgegebenen Clusterzahl und Startpartition für jedes Objekt untersucht, ob durch eine Verlagerung in eine andere Gruppe das Varianzkriterium vermindert werden kann. In unserem Beispiel mit $k = 61$ Objekten und $C = 6$ Gruppen existieren somit bei vollständiger Enumeration

$C^k = 6^{61} \cong 3 \cdot 10^{47}$ Einteilungsmöglichkeiten. Damit wird klar, daß sich "die partitionierenden Verfahren ... somit durch eine größere Variabilität [auszeichnen]. Sie haben jedoch bei praktischen Anwendungen nur wenig Verbreitung gefunden. Dieser Umstand ist vor allem durch folgende Punkte begründet:

- Die Ergebnisse der partitionierenden Verfahren werden verstärkt durch die der "Umordnung" der Objekte zugrunde liegenden Zielfunktion beeinflusst.
- Die Wahl der Startpartition ist häufig subjektiv begründet und kann ebenfalls die Ergebnisse des Clusterprozesses beeinflussen.
- Man gelangt bei partitionierenden Verfahren häufig nur zu lokalen und nicht zu globalen Optima, da selbst bei modernen EDV-Anlagen die Durchführung einer vollständigen Enumeration nicht wirtschaftlich möglich ist."

4. Die Wirtschaftsregionen

Die oben präferierte 6-Gruppen-Lösung mit jeweils hoher gruppeninternen wirtschaftsstrukturellen Ähnlichkeit zeigt Tabelle 1. Weil in der vorliegenden Studie Wirtschaftsräume thematisiert werden, sind in jeder Gruppe die enthaltenen Regionen ausgewiesen.

Insgesamt 13 Regierungsbezirke haben keine räumliche Zugehörigkeit. Sie sind – ausgenommen Lüneburg – Grenzregionen des Deutschen Reichs, so daß die Vermutung der beschäftigungsstrukturellen Ähnlichkeit mit den jeweils angrenzenden ausländischen Gebieten naheliegt.

Die Abbildung 3 zeigt die Räume der Tabelle 1, wobei aus pragmatischen Gründen Mecklenburg und Oldenburg in eine 'Küstenregion' (Cluster 2) integriert und Lüneburg und Osnabrück, die beide auch bei sehr feinen Clusterpartitionen immer in einer Gruppe erscheinen, aber durch ähnlich strukturierte Landschaften Oldenburgs und Hannovers getrennt sind, zusammengefügt werden.

Damit ist das erste Ergebnis dieser Studie beschrieben: es gibt Wirtschaftsregionen im Sinne beschäftigungsstruktureller Homogenität.¹²¹

Diese Regionen sind in der Regel unabhängig von Verwaltungsgrenzen. Sowohl Provinz- als auch Ländergrenzen konnten bezüglich der Erwerbsstruktur von der Clusteranalyse nicht erkannt werden. Dies kann zudem als Hinweis auf die Bedeutungslosigkeit der administrativen Wirtschaftspolitik auf Provinz- und Länderebene gedeutet werden. Aufgrund der im 19. Jahrhundert vorherrschenden liberalen Anschauungen wäre ein anderes Ergebnis allerdings als Überraschung zu betrachten.

Ob administrative Grenzen, so wie H. Theissen behauptet, "unter den deutschen Bedingungen des 19. Jahrhunderts für die Konstituierung regionaler Bindungen größere Bedeutung (hat) als beispielsweise in Frankreich oder England"¹²², mag nach den Clusteranalysen angezweifelt werden.

¹²¹Weil später nur der räumlich zusammenhängende Teil eines Clusters analysiert wird, bekommt man als Nebeneffekt zudem Regionen mit recht einheitlicher Gesamtstruktur. Gedacht sei hier insbesondere an Merkmale wie Klima, Kultur, Traditionen u.a.

¹²²H. Theissen, *Kommentar*, in: H. Kiesewetter u. R. Fremdling (Hrsg.), *Staat, Region und Industrialisierung*, Ostfildern 1984, S. 192.

