

Herwig Birg

Ein Prognosemodell zur Vorausschätzung der Angebots- und Nachfragekomponenten auf den regionalen Arbeitsmärkten in der Bundesrepublik Deutschland auf der Basis eines Optimierungsansatzes ^{*})

1. Ziel der Untersuchung

Ziel der Untersuchung ist es, für die 79 Regionen der Bundesverkehrswegeplanung die Zahl der angebotenen und nachgefragten Arbeitsplätze im Jahre 1990 zu prognostizieren. Die Angebots- und Nachfrageseite der regionalen Arbeitsmärkte wird jeweils durch eine große Zahl von Variablen (Zu- und Fortzüge aus anderen Regionen und dem Ausland, Pendler usw.) bestimmt. Es wird ein Modell dargestellt, das die Beziehungen zwischen den Variablen in der Vergangenheit und in der Prognoseperiode beschreibt. Hauptproblem der Prognose ist es, an Hand von zahlreichen Konsistenzbedingungen und bisher wenig beachteten Informationen zu überprüfen, ob und gegebenenfalls auf welche Weise die für die Vergangenheit ermittelten Strukturparameter in der Prognoseperiode modifiziert werden müssen. Dabei wird ein formalisiertes Verfahren angewandt, das es erlaubt, Informationen der unterschiedlichsten Art zu berücksichtigen.

2. Theoretische Vorüberlegungen

2.1 Die Struktur von Prognosemodellen

Unter einer Prognose wird im folgenden eine Aussage oder ein System von Aussagen verstanden, das Informationen über die künftigen Werte zeitabhängiger Variablen enthält. Je nachdem, ob die Gültigkeit einer prognostischen Aussage A_1 von der Gültigkeit einer zweiten Aussage A_2 abhängt oder nicht, unterscheidet man bedingte (konditionale) und unbedingte (nicht konditionale) Prognose-Aussagen. Die Qualität prognostischer Aussagen wird im allgemeinen nach dem Kriterium der Treffsicherheit bewertet. Wissenschaftliche Prognosen müssen darüber hinaus auch nachvollziehbar, d.h. intersubjektiv überprüfbar sein. Das Kriterium der Intersubjektivität wird von unbedingten Prognosen meist nicht erfüllt. Die folgenden Ausführungen beschränken sich daher auf den Bereich der konditionalen Prognosen.

In formalisierten und insbesondere in ökonometrischen Prognosemodellen entsprechen die abhängigen Variablen, die auch als "endogene" Va-

^{*}) Bericht über ein von der Deutschen Forschungsgemeinschaft gefördertes Projekt.

riablen bezeichnet werden, der Aussagenklasse A_1 , die unabhängigen (bedingenden) Variablen der Aussagenklasse A_2 . Die Wahrscheinlichkeit für das Eintreffen einer konditionalen Prognose-Aussage $P(A_1/A_2)$ hat als Obergrenze den Wert 1 (falls zwischen A_1 und A_2 ein stringenter Zusammenhang besteht) und als Untergrenze den Wert $P(A_1)$ (falls zwischen A_1 und A_2 keinerlei Zusammenhang besteht). Die Wahrscheinlichkeit dafür, daß die Prognoseaussage A_1 , gegeben A_2 , zutrifft, liegt also in dem Intervall

$$(1.1) \quad P(A_1) \leq P(A_1/A_2) \leq 1$$

Ausschlaggebend dafür, wie weit sich $P(A_1/A_2)$ dem angestrebten Wert 1 annähert, ist die Stringenz des Zusammenhanges zwischen den beiden Sachverhalten, auf die sich A_1 bzw. A_2 beziehen. Die Beziehung, die diesen Zusammenhang ausdrückt, wird als Prognosefunktion bezeichnet.

Der Vollständigkeit halber sei auch das Intervall für die Wahrscheinlichkeit bei unbedingten Prognosen angegeben:

$$(1.2) \quad P(A_1) \cdot P(A_2) \leq P(A_1A_2) \leq P(A_1)$$

Auch hier hängt die Wahrscheinlichkeit $P(A_1A_2)$ entscheidend von der Strenge des Zusammenhanges zwischen A_1 und A_2 ab. Ist der Zusammenhang stringent, so ist $P(A_1) = P(A_2)$. In diesem Fall kann die rechte Seite von (1.2) auch durch den Ausdruck $P(A_2)$ ersetzt werden.

Die Treffsicherheit einer konditionalen Prognose kann selbst dann sehr klein sein, wenn zwischen den abhängigen und den unabhängigen Variablen ein in der Zeit konstanter stringenter Zusammenhang besteht. In diesem Fall kann $P(A_1/A_2)$ den Wert 1 annehmen, ohne daß A_1 auch nur annähernd eintritt, nämlich dann, wenn A_2 falsch gesetzt wurde. Eine hohe Treffsicherheit erfordert deshalb sowohl eine möglichst stringente Beziehung zwischen A_1 und A_2 (Prognosefunktion) als auch einen hohen Wahrscheinlichkeitsgrad für die Gültigkeit der Bedingung A_2 . Damit die Bedingung A_2 realistisch gesetzt werden kann, ist wiederum eine Prognosefunktion für A_2 erforderlich. Wird in dieser Funktion A_2 von einer weiteren Größe abhängig gemacht usf., so gilt entsprechend (1.1)

$$(1.3) \quad P(A_1) \leq P(A_1/A_2 \wedge \dots \wedge A_n) \leq 1$$

Die Folge der Funktionen $A_1(A_2), A_2(A_3), \dots$ endet in der letzten Stufe bei einer Variablen A_n , die nur noch eine Funktion der Zeit ist: $A_n(t)$. Die Variable t als $n+1$ te Größe wirkt dabei als letzte Ursache aller Veränderungen kein Prognoseproblem auf (Zeit definiert als Kalenderzeit).

Das System der Funktionen $A_1(A_2), \dots, A_n(t)$ wird im folgenden mit dem Begriff Prognosemodell bezeichnet. Das System ist in Übersicht 1 dargestellt.

Übersicht 1
Systematische Darstellung eines mehrstufigen Prognosemodells

Modellstufe	abhängige Variable	Prognosefunktion	unabhängige Variable
n	A_n	$A_n(t)$	t
n-1	A_{n-1}	$A_{n-1}(A_n)$	A_n
⋮	⋮		⋮
1	A_1	$A_1(A_2)$	A_2

In ökonomischen Prognosemodellen werden die unabhängigen Variablen in die beiden Klassen der endogen-verzögerten und der exogenen Größen unterteilt. Die Prognose der endogen-verzögerten Variablen folgt dem Schema in Übersicht 1, die exogenen Größen werden dagegen meist nicht durch eine Abfolge von Prognosefunktionen vorausgeschätzt, sondern durch mehr oder weniger willkürliche Annahmen gesetzt. Enthält ein Prognosemodell keine endogen-verzögerten Variablen, sondern nur endogene (abhängige) und exogene Größen, so ist es durch die erste Stufe in Übersicht 1 darstellbar.

2.2 Das Problem der Instabilität von Prognosefunktionen und einige Lösungsvorschläge aus der Literatur

Eine hohe Treffsicherheit, d.h. eine möglichst gute Annäherung der Wahrscheinlichkeit $P(A_1/A_2 \cap \dots \cap A_n)$ an den Wert 1, setzt voraus, daß a) die Prognosefunktionen $A_1(A_2), \dots, A_n(t)$ auf möglichst stringenten Beziehungen basieren, und daß b) diese Beziehungen innerhalb des Prognosezeitraums unverändert gültig sind. In den Sozialwissenschaften sind in der Regel nicht nur die Werte der zu prognostizierenden Variablen, sondern auch ihre Beziehungen untereinander zeitlich variabel. Hierin liegt ein bedeutsamer Unterschied zu den Naturwissenschaften, die über ein großes Arsenal zeitlich invarianter Beziehungen ("Gesetze") verfügen. In der Übersicht 2 sind die Unterschiede zwischen den verschiedenen Disziplinen dargestellt.

Übersicht 2
Klassifikation des Prognoseproblems nach verschiedenen Disziplinen

		Meßwerte der Variablen	
		nicht konstant	konstant
Beziehungen zwischen den Variablen	nicht konstant	Fall 1: Sozialwissenschaftliche Prognosen	Fall 3: schwer interpretierbar
	konstant	Fall 2: Naturwissenschaftliche Prognosen	Fall 4: für Prognosen irrelevant

Gelten die Prognosefunktionen sowohl im Analysezeitraum, in dem sie getestet wurden, als auch im Prognosezeitraum, so besteht zwischen der Erklärung und der Prognose eines Phänomens kein prinzipieller Unterschied ("Symmetrie zwischen Erklärung und Prognose", Fall 2 in Übersicht 2). Sind dagegen wie in den Sozialwissenschaften die Beziehungen zwischen den Variablen in der Zeit veränderlich, so treten sowohl bei der Erklärung als auch bei der Prognose besondere Probleme auf:

- (I) Eine hinreichend sichere Überprüfung von Hypothesen setzt voraus, daß
 - (a) Messungen für eine möglichst große Zahl verschiedener Objekte (in der Regionalwissenschaft meist Regionen) vorliegen,
 - (b) daß diese Messungen sich alle auf den gleichen Zeitpunkt bzw. auf diejenige Zeitspanne beziehen, in der die zu testende Hypothese unverändert gilt, und
 - (c) daß die Objekte bzw. ihre Merkmalsausprägungen unabhängig voneinander sind.

Je mehr Objekte (Regionen) bei einer gegebenen Gesamtfläche durch Regionalisierung gebildet werden, desto größer ist die Sicherheit bezüglich des Testergebnisses (hohe Identifikation), aber desto weniger ist die Forderung (c) nach voneinander unabhängigen Merkmalsausprägungen erfüllt, weil die Regionen infolge von räumlichen Interaktionen nicht unabhängig voneinander sind (geringer Spezifikationsgrad). Der unlösbare "Antagonismus zwischen Spezifikation und Identifikation"¹⁾ tritt sowohl bei Zeitreihen- als auch bei Querschnittsanalysen auf. Auf Grund dieses Antagonismus ist es nicht ohne weiteres möglich, die Zahl der Beobachtungen in einer Querschnittsanalyse beispielsweise durch eine feinere Regionalisierung zu erhöhen.

- (II) Prognosen bei zeitabhängigen Beziehungen führen zum sogenannten Stabilitätsproblem. Selbst wenn für mehrere Teilperioden in der Vergangenheit eine Abfolge von jeweils nur in bestimmten Perioden gültigen Beziehungen hinreichend gut spezifiziert und identifiziert wäre: wie sollten die künftigen Beziehungen (Funktionsparameter) prognostiziert werden (Stabilitätsproblem in der Ökonometrie)?²⁾

Beide Probleme sind weitgehend ungelöst. Die folgenden Lösungsvorschläge aus der ökonometrischen Literatur kontrastieren mit der in der Ökonometrie sonst üblichen theoretischen Fundierung und methodologischen Klarheit.

-
- 1) Vgl. G. Menges: Die ökonomische Struktur und die Frage ihrer Konstanz. In: Strukturwandlungen in einer wachsenden Wirtschaft, 2.Bd., Schriften des Vereins für Socialpolitik, Berlin 1964, S. 995.
 - 2) Vgl. G. Menges und H. Diehl: Das Stabilitätsproblem in der Ökonometrie. In: Statistische Hefte, Jg. 1965, Vol. 6, S. 27 f.

(1) Klein und Theil haben vorgeschlagen, die in einem ersten Schritt mechanisch aus dem Prognosemodell abgeleiteten Vorausschätzungen, die auf Grund der Probleme I und II Mängel aufweisen müssen, in einem zweiten Schritt zu modifizieren, indem Informationen über bestimmte Eigenschaften der Residuen verwendet werden.¹⁾

(2) Hujer plädiert für eine Einbeziehung antizipierter Zielvorstellungen gesellschaftlicher Gruppen und eine stärkere Beachtung der vom Staat kontrollierten Instrumentvariablen.²⁾

(3) Jorgenson, Hunter und Nadiri haben am Beispiel von Investitionsfunktionen gezeigt, daß durch die Berücksichtigung antizipativer Variablen entscheidende Verbesserungen der Prognosegüte gelingen können: "No econometric model currently available can compete with the anticipated investment data in explanatory power."³⁾

Heute wird die Auffassung allgemein akzeptiert, daß bei der Anwendung von formalisierten Prognosemodellen auf die angeführten oder auf ähnliche heuristische Verfahren nicht verzichtet werden kann: "It is a fact of life that purely numerical methods cannot be used, but must be supplemented by special information and personal judgement. On the other hand, personal judgement alone is not suitable for prediction of a wide range of variables in a complicated industrial economy. The two must be used together. The objectively estimated model is needed as a framework in which to interpret special and subjective information."⁴⁾

2.3 Ein eigener Lösungsvorschlag

Die in ökonometrischen Modellen enthaltenen Informationen bestehen aus

- Fakten (Meßwerten),
- empirisch bestätigten Hypothesen und
- Annahmen (insbesondere hinsichtlich der statistischen Verteilungsgesetze der Residuen).

Informationen über Ziele werden in diesen Modellen lediglich implizit berücksichtigt, beispielsweise bei der Spezifikation von Verhaltensgleichungen (Konsumfunktion, Investitionsfunktion usw.). Da sozialwissenschaftliche Prozesse ohne Einbeziehung von Motiven, Werthaltungen, Normen und Präferenzen nicht verstanden werden können, führt ihre Vernachlässigung bei Prognosen auch mit großer Wahrscheinlichkeit zu Fehlurteilen. Es liegt deshalb nahe, die Klasse der Normen und alle übrigen

1) Vgl. L.R.Klein: A Textbook of Econometrics, 2. Auflage, New Jersey, 1974, S. 277.

2) Vgl. R. Hujer: Prognosegüte ökonometrischer Modelle - Ein Beitrag zum "Stabilitätsproblem in der Ökonometrie" -. In: Statistische Hefte, 16. Jg. 1975, Heft 1, S. 23.

3) D.W. Jorgenson, J. Hunter und M.I. Nadiri: A comparison of Alternative Econometric Models of Quarterly Investment Behavior, in: Econometrica, Vol. 38, März 1970, Nr. 2, S. 209.

4) C.R. Klein: A Textbook ..., a.a.O., S. 278.

Informationen, die infolge des formalen Aufbaus der meisten Prognosemodelle unberücksichtigt bleiben müssen, bei Prognosen zusätzlich heranzuziehen. Die heterogenen und für Prognosen unter Umständen sehr relevanten Informationen lassen sich nach folgenden Gesichtspunkten klassifizieren:

- nach der Referenzperiode: Expost-Informationen und antizipierte Informationen,
- nach der Herkunft: Empirische (induktive) und theoretische (deduktive) Informationen,
- nach dem Grad der Überprüfbarkeit: Objektive und subjektive Informationen,
- nach der Zuverlässigkeit: Spekulative und bestätigte Informationen,
- nach dem Grad der Vollständigkeit: Allgemeine (für alle Beobachtungsobjekte verfügbare) und punktuelle Informationen,
- nach der kategorialen Zugehörigkeit: Informationen über Fakten und Informationen über Normen.

Diese unterschiedlichen Teilinformationen lassen sich in einen formalen und intersubjektiv überprüfbaren Prognoseansatz einfügen, wenn das Prognoseproblem als ein Eingrenzungsproblem aufgefaßt wird:

Statt der üblichen Punkt- oder Intervallprognosen, bei denen ein Bereich angegeben wird, der vom Erwartungswert der zu prognostizierenden Variablen ausgehend mehr oder weniger stark nach außen erweitert werden kann, läßt sich der prognostizierte Bereich auch durch eine stufenweise Eingrenzung des zunächst allein auf Grund rein logischer Bedingungen möglichen und deshalb zunächst relativ großen Wertebereichs der zu prognostizierenden Variablen ermitteln. Dabei kann grundsätzlich jede zusätzlich berücksichtigte Information den bereits eingegrenzten Wertebereich weiter einengen.

Praktisch läßt sich dieses Verfahren dadurch durchführen, daß die ökonometrischen Gleichungen, die als Prognosefunktionen dienen, zu Ungleichungen umformuliert werden, die in ihrer Gesamtheit bei der Bestimmung der Variablenwerte eingehalten werden müssen. Gilt beispielsweise für die Region r in der Periode t die ökonometrische Beziehung

$$y^r(t) = f(x_1^r(t), \dots, x_n^r(t)) + u^r(t),$$

so wird diese Gleichung in die Ungleichungen

$$\left. \begin{array}{l} y^r(t) - f(x_1^r(t), \dots, x_n^r(t)) \leq u^r(t) \\ y^r(t) - f(x_1^r(t), \dots, x_n^r(t)) \geq 0 \end{array} \right\} \text{für } u^r(t) > 0$$

umgeformt.¹⁾ Dadurch wird gefordert, daß die zu prognostizierende Variable $y^r(t)$ in einem Intervall liegt, das durch die Differenz zwischen

¹⁾ Falls $u^r(t) < 0$, kehren sich in diesen Ungleichungen die Zeichen "<" und ">" um.

dem Prognosewert und dem Funktionswert ($u^f(t)$) gebildet wird. War das Störglied $u^f(t)$ in der Vergangenheit positiv (negativ) und kann davon ausgegangen werden, daß die Abweichung vom Funktionswert in erster Linie systematisch, kausal und weniger stochastisch bedingt ist, so kann für $u^f(t)$ auch in der Zukunft ein positiver (negativer) Wert angesetzt werden. Die für die Parameterschätzung getroffene Annahme normalverteilter Reste steht dazu nicht im Widerspruch.

Zur Ermittlung der Variablenwerte, die allen Ungleichungen genügen, bietet sich als Lösungsverfahren die Methode der linearen Optimierung an: Die aus den ökonomischen Gleichungen abgeleiteten Ungleichungen werden als Nebenbedingungen verwendet. Als Zielfunktion dient diejenige Norm, der unter den gegebenen Informationen in der Zukunft die höchste Priorität zukommt.¹⁾ In bezug auf die inhaltlichen Probleme interregionaler Prognosen bietet dieses Verfahren folgende Vorteile.

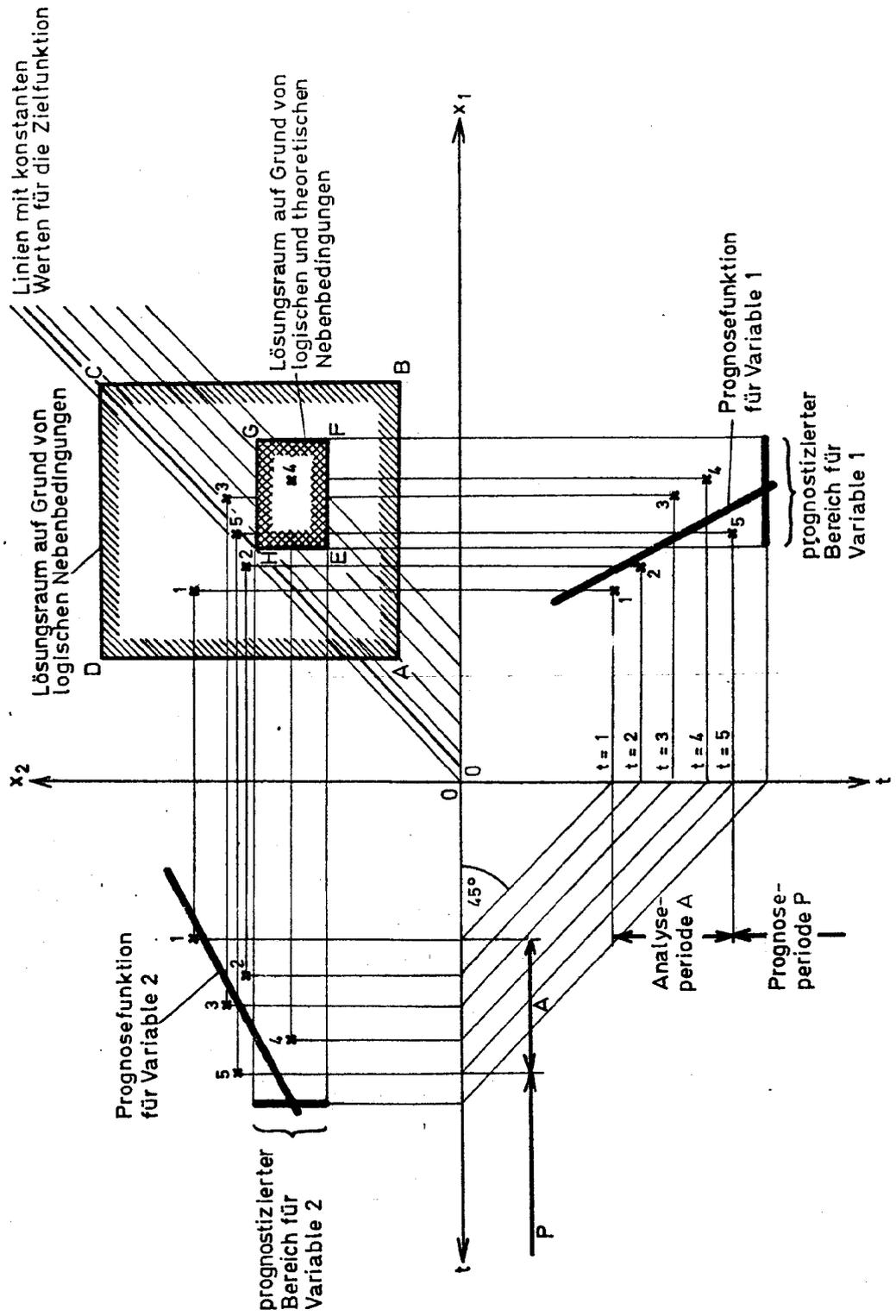
- (1) Es können durch Formulierung entsprechender Nebenbedingungen regionale und gesamtwirtschaftliche Prognosen miteinander konsistent gemacht werden.
- (2) Isolierte und punktuelle Informationen über die künftigen Werte einzelner Variablen in bestimmten Regionen können berücksichtigt werden.
- (3) Es ist möglich, Ziele und Normen explizit einzubeziehen.
- (4) Es können alternative Prognosefunktionen gleichzeitig nebeneinander verwendet werden.²⁾
- (5) Die logische Widerspruchsfreiheit aller in Form von Nebenbedingungen berücksichtigten Informationen und Annahmen ist stets gewährleistet.

Das Verfahren ist in schematischer Form für zwei Variablen in Schaubild 1 dargestellt. Der Punkt H in Schaubild 1 repräsentiert die prognostizierten Werte der beiden Variablen. Dieses Verfahren gestattet es, interdependente Beziehungen zwischen den Variablen zu berücksichtigen. Durch Einbeziehung einer Vielzahl von Informationen in Form von Nebenbedingungen dürfte es aber gegenüber den üblichen Modellen den Vorteil bieten, die Prognosewahrscheinlichkeit $P(A_1/A_2 \dots A_n)$ beträchtlich zu erhöhen. Diese Wahrscheinlichkeit hängt nicht zuletzt davon ab, ob es gelingt, die richtigen Prognoseparameter zu finden. Wie erste Versuche gezeigt haben, kann eine feasible solution oft nur dann gefunden werden, wenn bestimmte für die Vergangenheit ermittelte Funktionsparameter ge-

-
- 1) Genauer: der die höchste Priorität unter allen Normen zukommt, die sonst noch als Zielfunktion verwendet werden könnten. Denn Normen, die in der Form von Nebenbedingungen berücksichtigt werden, müssen stets voll erfüllt werden. Insofern haben sie stets höhere Priorität als die Zielfunktion, die immer nur mehr oder weniger gut erfüllt ist.
 - 2) Die gleichzeitige Verwendung alternativer Prognosefunktionen bietet sich dann an, wenn verschiedene Funktionstypen mit alternativen Erklärungsgrößen ein bestimmtes Phänomen gleich gut erklären. Nur ei-
ne Funktion auszuwählen würde in diesem Fall bedeuten, ohne Not auf Informationen zu verzichten.

Schaubild 1

Schematische Darstellung des Lösungsraumes für die Prognosewerte von zwei Variablen in einem Optimierungs-Ansatz



ändert werden. Dabei gibt das LP-Verfahren relativ eindeutige Hinweise, an welchen Stellen in einer großen Zahl von Nebenbedingungen nicht erfüllbare Beschränkungen existieren.

Das Verfahren bleibt zumindest in bezug auf die Art der Verwendung subjektiver Informationen und in bezug auf deren relative Bedeutung für die Ableitung des Prognoseergebnisses intersubjektiv überprüfbar, auch wenn Quelle, Relevanz und Zuverlässigkeit dieser Informationen ebenso wie bei den in der Ökonometrie vorgeschlagenen heuristischen Verfahren viele Probleme aufwerfen.

3. Das Modell

3.1 Vorbemerkung

Die kategoriale Trennung von normativen und positiven Aussagen hat dazu geführt, daß Entscheidungsmodelle unabhängig von Erklärungs- bzw. Prognosemodellen entwickelt und angewandt werden. Die Ausklammerung von Normen aus Prognosemodellen ist aber immer dann mit einem Informationsverlust verbunden, wenn die entsprechende Norm von den Entscheidungsträgern nicht nur propagiert, sondern auch realisiert wird. Ist die Annahme realistisch, daß die zur Erreichung von bestimmten Normen erforderlichen Aktivitäten tatsächlich ergriffen werden, dann läuft die Vernachlässigung der Normen auf eine Verringerung des Wahrscheinlichkeitsgrades des Prognoseergebnisses hinaus. In diesem Fall erscheint es gerechtfertigt, den formalen Apparat von Entscheidungsmodellen, in denen Normen explizit berücksichtigt werden können, für Prognosezwecke einzusetzen. Für dieses Vorgehen spricht auch die Überlegung, daß die beispielsweise vom Staat in der Vergangenheit angestrebten politischen Ziele vermutlich nicht ohne Einfluß auf die Parameterkonstellationen von ökonometrischen Modellen waren, auf deren Basis Prognosen durchgeführt werden. Deshalb sind auch die herkömmlichen ökonometrischen Prognosemodelle bzw. deren Aussagen konditional von der Relevanz von Normen (aus der Vergangenheit) abhängig, wenn auch nur implizit. Demgegenüber bieten Entscheidungsmodelle den Vorteil, Normen explizit zu berücksichtigen und den Einfluß alternativer Normen-Vorgaben auf das Prognoseergebnis zu quantifizieren.

3.2 Die Zielfunktion des Modells

Sowohl in gesamtwirtschaftlichen als auch in regionalisierten Entscheidungsmodellen wird oft das Wachstum des Bruttoinlandsprodukts als Zielfunktion gewählt. Ein nicht weniger plausibles und relevantes Ziel besteht in der Minimierung der Differenz zwischen der Zahl der nachgefragten und der Zahl der angebotenen Arbeitsplätze (Vollbeschäftigungs-

ziel).¹⁾ Dieses Ziel gewinnt zunehmend an Bedeutung.

Die Verwendung des Wachstumsziels würde voraussetzen, daß das Wachstum des Bruttoinlandsprodukts als Funktion der Modellvariablen numerisch beschrieben werden kann. Meist geschieht dies auf der Basis einer Produktionsfunktion, deren Elastizitäten in der Zielfunktion als Parameter verwendet werden²⁾.

Im folgenden wird das Vollbeschäftigungsziel zugrunde gelegt, weil Produktionselastizitäten aus statistischen Gründen bisher nur für relativ hoch aggregierte Produktionsfunktionen geschätzt werden können, deren Aussagekraft sehr beschränkt ist. Eine Zielfunktion, die als Differenz zwischen der Zahl der nachgefragten und der Zahl der angebotenen Arbeitsplätze definiert ist, hat demgegenüber den Vorteil, daß die Gewichte der entsprechenden Variablen in der Zielfunktion den Wert 1 besitzen, so daß hier Schätzprobleme, wie sie bei der empirischen Ermittlung von Produktionsfunktionen entstehen, entfallen. Für die Verwendung einer möglichst einfachen Zielfunktion spricht auch die Überlegung, daß die Art der Zielfunktion von vergleichsweise geringer Bedeutung ist, wenn bei der numerischen Lösung des Optimierungsansatzes das eigentliche Problem darin besteht, solche Variablenwerte zu finden, die sämtlichen Nebenbedingungen genügen.

(a) Regionales Angebot an Arbeitsplätzen im Prognosejahr

Das Angebot an Arbeitsplätzen im Prognosejahr 1990 wird aus einem sektorspezifischen Shift-Ansatz abgeleitet.³⁾ Für jede Region R gilt folgende Angebotsfunktion:

$$(2) \quad A(R, 90) = \sum_I [A(R, I, 70) \cdot VU(I) + VA(R, I)];$$
$$R = 1, \dots, 79$$

Hierin ist $A(R, I, 70)$ der Arbeitsplatzbestand im Sektor I der Region R im Basisjahr der Prognose (1970). $VU(I)$ ist der gesamtwirtschaftliche Wachstumsfaktor des Sektors I, der außerhalb des Modells bestimmt wird:

$$(2.1) \quad VU(I) = \sum_R A(R, I, 90) / \sum_R A(R, I, 70); \quad I = 1, \dots, 44$$

- 1) Dies ist eines der Ziele, die bei der Gemeinschaftsaufgabe "Verbesserung der regionalen Wirtschaftsstruktur" und im Rahmen des Stabilitätsgesetzes verfolgt werden.
- 2) Vgl. R. Thoss: Ein Vorschlag zur Koordinierung der Regionalpolitik in einer wachsenden Wirtschaft. In: Jahrbücher für Nationalökonomie und Statistik, Band 182, 1968/69, S. 490-529.
- 3) Angebotsfunktionen auf der Basis regionaler Invertitionsfunktionen wären theoretisch zwar befriedigender, doch gibt es für diese Funktionen bisher noch keine ausgereiften Parameterschätzungen.

Der Standorteinfluß $VA(R, I)$ im Prognosezeitraum wird definiert als

$$(2.2) \quad VA(R, I) = A(R, I, 90) - A(R, I, 70) \cdot VU(I); \quad R = 1, \dots, 79 \\ I = 1, \dots, 44$$

Gleichung 2 kann einfacher geschrieben werden, wenn man über alle Sektoren I aggregiert:

$$(2.3) \quad A(R, 90) = A(R, 70) \cdot VU(R) + VA(R); \quad R = 1, \dots, 79$$

In (2.3) ist $A(R, 70)$ der gesamte Arbeitsplatzbestand der Region im Jahre 1970. $VU(R)$ ist die Summe der mit den Anteilen der Sektoren am regionalen Arbeitsplatzbestand gewichteten sektoralen Wachstumsfaktoren:

$$(3) \quad VU(R) = \sum_I \left[A(R, I, 70) / A(R, 70) \right] \cdot VU(I)$$

und $VA(R)$ der über alle Sektoren summierte Standorteinfluß

$$(4) \quad VA(R) = \sum_I VA(R, I) \quad .$$

Die Standorteinflüsse $VA(R)$ in den Regionen werden durch den Optimierungsansatz bestimmt. Sie wurden darüber hinaus auch außerhalb des Modells nach einem im DIW entwickelten modifizierten Shift-Ansatz prognostiziert.¹⁾ Im vorliegenden Ansatz dienen die außerhalb des Modells gewonnenen Prognosen zur Bestimmung von Ober- und Untergrenzen für die im Optimierungsmodell abzuleitenden Werte der Standortfaktoren.

(b) Regionale Nachfrage nach Arbeitsplätzen im Prognosejahr

Die Nachfrage nach Arbeitsplätzen in der Region R im Prognosejahr ist eine Funktion zweier Komponenten, des Bevölkerungsbestandes in der Region R selbst und des Bevölkerungsbestandes in den angrenzenden Regionen. Die Nachfrage, die die in der Region lebende Bevölkerung ausübt, wird über eine Prognose der Erwerbsquoten bestimmt (Erwerbsquote = Quotient aus der Zahl der Erwerbspersonen mit Wohnsitz in der Region und der Zahl der Einwohner). Diese Nachfrage wird um denjenigen Teil, der durch Arbeitsplätze außerhalb der Region befriedigt wird, vermindert (Auspendler) und um diejenige Nachfrage erhöht, die die außerhalb der Region lebende Bevölkerung in der Region R ausübt (Einpendler).

Für die Zahl der in der Region R im Jahr 1990 nachgefragten Arbeitsplätze ergeben sich insgesamt folgende 7 Komponenten $N(R, 90, 1), \dots, N(R, 90, 7)$:

1) Vgl. H. Birg unter Mitarbeit von E. Jöhrens und W. Kirner: Prognose des regionalen Angebots an Arbeitsplätzen, Sonderheft Nr. 105 des DIW, Duncker und Humblot, Berlin, 1975.

1. Produkt des Bevölkerungsbestandes $BN(R,90)$ im Jahre 1990, der sich allein aus der natürlichen Bevölkerungsentwicklung des Anfangsbestandes ergibt (Fortschreibung des Anfangsbestandes auf Grund von Geburten und Sterbefällen), mit der Erwerbsquote $EN(R)$ dieses Teils des regionalen Bevölkerungsbestandes im Prognosejahr:

$$(5.1) \quad N(R,90,1) = EN(R) \cdot BN(R,90)$$

2. Kumulierte Zuzüge aus anderen Regionen ($ZB(R)$) bzw. aus dem Ausland ($ZA(R)$), multipliziert mit Faktoren, die das Wachstum sowie die Erwerbsbeteiligung der entsprechenden Bevölkerungsgruppen berücksichtigen. Nimmt man an, daß die über die einzelnen Jahre des Prognosezeitraums kumulierten Zuzüge $ZB(R)$ bzw. $ZA(R)$ sich jeweils aus gleich großen jährlichen Strömen zusammensetzen, deren alters- und geschlechtsspezifische Gliederung in jedem Jahr gleich bleibt, dann läßt sich für die kumulierten Zuzüge insgesamt ein einziger Faktor berechnen, mit dem der aus den Wanderungen resultierende Teil des Bevölkerungsbestandes im Jahre 1990 ermittelt werden kann. Dieser Faktor wird für die aus den anderen Regionen zuziehenden Personen mit $XZB(R)$, für die aus dem Ausland zuziehenden mit $XZA(R)$ bezeichnet. Die Erwerbsquoten der beiden Bevölkerungsgruppen werden mit $EZB(R)$ bzw. $EZA(R)$ symbolisiert. Mit diesen Größen kann die auf den Zuzügen beruhende Nachfrage wie folgt berechnet werden:

$$(5.2) \quad N(R,90,2) = EZB(R) \cdot XZB(R) \cdot ZB(R)$$

$$(5.3) \quad N(R,90,3) = EZA(R) \cdot XZA(R) \cdot ZA(R)$$

3. Durch Fortzüge in andere Regionen ($FB(R)$) und in das Ausland ($FA(R)$) verringert sich die Nachfrage (zur Definition der Multiplikatoren vgl. Punkt 2):

$$(5.4) \quad N(R,90,4) = - EFB(R) \cdot XFB(R) \cdot FB(R)$$

$$(5.5) \quad N(R,90,5) = - EFA(R) \cdot XFA(R) \cdot FA(R)$$

Die Koeffizienten in den Gleichungen (5.2) bis (5.5) werden auch in folgender Weise zusammengefaßt:

$$(5.5.1) \quad EXZB(R) = EZB(R) \cdot XZB(R)$$

$$EXZA(R) = EZA(R) \cdot XZA(R)$$

$$EXFB(R) = EFB(R) \cdot XFB(R)$$

$$EXFA(R) = EFA(R) \cdot XFA(R)$$

4. Berufspendler, die in die Region einpendeln ($PE(R,90)$) bzw. aus der Region auspendeln ($PA(R,90)$) erhöhen bzw. verringern die Nachfrage in der Region:

$$(5.6) \quad N(R, 90, 6) = PE(R, 90)$$

$$(5.7) \quad N(R, 90, 7) = - PA(R, 90)$$

Das Symbol $PS(R, 90) = PE(R, 90) - PA(R, 90)$ dient zur Bezeichnung des Pendlersaldos.

In jeder Region gilt damit für die Nachfrage nach Arbeitsplätzen insgesamt:

$$(6) \quad N(R, 90) = \sum_{J=1}^7 N(R, 90, J)$$

c) Regionaler Arbeitsmarktsaldo im Prognosejahr

Der Saldo $D(R, 90)$ zwischen der Zahl der nachgefragten und der Zahl der angebotenen Arbeitsplätze in jeder Region im Prognosejahr ergibt sich aus den Gleichungen (2.3) und (6) (unter Verwendung der Definitionen in 5.5.1):

$$(7) \quad \begin{aligned} D(R, 90) &= N(R, 90) - A(R, 90) \\ &= BN(R, 90) \cdot EN(R) \\ &\quad + EXZB(R) \cdot ZB(R) \\ &\quad + EXZA(R) \cdot ZA(R) \\ &\quad - EXFB(R) \cdot FB(R) \\ &\quad - EXFA(R) \cdot FA(R) \\ &\quad + PS(R, 90) \\ &\quad - [A(R, 70) \cdot VU(R) + VA(R)] \end{aligned}$$

d) Volkswirtschaftlicher Arbeitsmarktsaldo im Prognosejahr

Summiert man in Gleichung (7) über alle Regionen, so erhält man den gesamträumlichen Arbeitsmarktsaldo, der ein Minimum erreichen soll. Die Zielfunktion des Optimierungsansatzes lautet daher:

$$(8) \quad \begin{aligned} Z &= \sum_R D(R, 90) \longrightarrow \text{Min!} \\ &= \sum_R EN(R) \cdot BN(R, 90) \\ &\quad + \sum_R EXZB(R) \cdot ZB(R) \\ &\quad + \sum_R EXZA(R) \cdot ZA(R) \\ &\quad - \sum_R EXFB(R) \cdot FB(R) \end{aligned}$$

$$\begin{aligned}
 \text{(noch 8)} \quad & - \sum_R \text{EXFA}(R) \cdot \text{FA}(R) \\
 & + \sum_R \text{PS}(R, 90) \\
 & - \sum_R \left[A(R, 70) \cdot \text{VU}(R) + \text{VA}(R) \right]
 \end{aligned}$$

Würden sowohl der aus der natürlichen Bevölkerungsentwicklung stammende Bevölkerungsbestand $\text{BN}(R, 90)$ und die Zu- und Fortzüge als auch die mit diesen Größen multiplizierten Faktoren $\text{EN}(R)$, $\text{EXZB}(R)$, ..., $\text{EXFA}(R)$ als Variable behandelt, so wäre die Zielfunktion nichtlinear. Um die Lösung mittels eines linearen Optimierungs-Ansatzes ermitteln zu können, müssen wahlweise entweder der natürliche Bevölkerungsbestand und die Zu- und Fortzugs-Variablen oder die entsprechenden Faktoren als Parameter vorgegeben werden. Im folgenden wird davon ausgegangen, daß die Größen

Koeffizienten der Zielfunktion	Anzahl (bei 79 Regionen)	Nr. der Spalte in Übersicht 3
$\text{BN}(R, 90)$	79	1
$\text{EXZB}(R)$	79	2
$\text{EXZA}(R)$	79	3
$\text{EXFB}(R)$	79	4
$\text{EXFA}(R)$	79	5
$A(R, 70)$	79	7
insgesamt	474	

außerhalb des Modells bestimmt werden (vgl. Übersicht 3, Zeile 1).

Die Beschäftigtenbestände in den einzelnen Sektoren und Regionen des Basisjahres ($A(R, I, 70)$) und die aggregierten Größen $A(R, 70)$ sind bekannt, die gesamtträumlichen sektoralen Wachstumsfaktoren $\text{VU}(I)$ und die für die Sektoren aggregierten regionalen Faktoren $\text{VU}(R)$ werden außerhalb des Modells bestimmt. Daher ist auch die Summe

$$\sum_R \left[A(R, 70) \cdot \text{VU}(R) + \text{VA}(R) \right]$$

in der Zielfunktion gegeben, denn die Summe $\sum_R \text{VA}(R)$ ist definitionsgemäß Null. Die Zielfunktion (8) ließe sich durch Subtraktion dieser Summe vereinfachen. Auch die Variablen, mit denen die Ein- und Auspendler erfaßt werden, könnten aus der Zielfunktion eliminiert werden, denn die Summe der regionalen Pendlersalden ist gleich dem nationalen Pendlersaldo, der vorgegeben wird.

Die nicht vereinfachte Version der Zielfunktion ist aber vorzuziehen, weil sich dann die Schattenpreise in ökonomisch interpretierbaren Einheiten

berechnen lassen, nämlich als zusätzliches Überschußangebot bzw. zusätzliche Überschußnachfrage auf dem Arbeitsmarkt, die sich ergibt, wenn der in einer Nebenbedingung vorgegebene Schwellenwert für die jeweils betrachtete Variable um eine Einheit verändert wird.

Die folgende Aufstellung gibt eine Übersicht über die Zahl der zu prognostizierenden Größen (vgl. Übersicht 3, Zeile "0").

Variablen der Zielfunktionen	Anzahl (bei 79 Regionen)	Nr. der Spalte in Übersicht 3
EN(R)	79	1
ZB(R)	79	2
ZA(R)	79	3
FB(R)	79	4
FA(R)	79	5
PS(R)	79	6
VU(R)	79	7
VA(R)	79	8
Zwischensumme	632 Variablen	
zuzüglich	316 abgeleitete Größen	9 - 12
total	948 Unbekannte	

Eine sektorspezifische Differenzierung der Standorteinflüsse VA(R) beispielsweise nach den 44 Sektoren der Arbeitsstättenzählung würde die Zahl der endogenen Größen auf über 4000 ansteigen lassen. Da auch mit großen Rechenanlagen derartig umfangreiche Modelle mit tragbarem Aufwand nicht zu bewältigen sind, wird das Modell in aggregierter Form angewandt.

3.3 Nebenbedingungen

Nach der Art der in den Nebenbedingungen enthaltenen Informationen werden definatorische bzw. logische, empirische und normative Nebenbedingungen unterschieden. Die folgende Numerierung der Nebenbedingungen nimmt Bezug auf die Vorspalte der Übersicht 3.

Bedingung 2 enthält die Bilanzgleichung für die regionalen Bevölkerungsbestände im Prognosejahr.

Bedingung 3 sagt aus, daß die Summe der regionalen Bevölkerungsbestände auf Grund der natürlichen Bevölkerungsentwicklung dem entsprechenden nationalen Wert gleich sein muß.

Bedingung 4 und 5 definieren das Angebot von Arbeitsplätzen im Basisjahr und im Prognosejahr.

Bedingung 6 bis 8 stellen die Konsistenz zwischen den regionalen und nationalen Beständen an Arbeitsplätzen und Einwohnern her.

Bedingung 9 und 10 beruhen auf den logischen Forderungen, daß a) die Summe der regionalen Standorteinflüsse und b) die Summe der regionalen Binnenwanderungssalden Null ergibt.

Übersicht 3

LINEARES OPTIMIERUNGSMODELL ZUR PROGNOSE DER ANGEBOTS-

Zielfunktion:
Regionale Gliederung:

D = definitorische Nebenbedingung bzw.
logische Nebenbedingung
E = empirische Nebenbedingung
N = normative Nebenbedingung

	Zahl der Zeilen	Laufin- dex für R						
			(1)	(2)	(3)	(4)	(5)	(6)
(0) Variable der Zielfunktion	1	alle R	EN(R)	ZB(R)	ZA(R)	FB(R)	FA(R)	PS(R)
(1) Koeffizienten der Zielfunktion	1	alle R	BN(R,90)	EXZB(R)	EXZA(R)	-EXFB(R)	-EXFA(R)	1
(2) D für B(R,90)	79	R = 1 R = 79		XZB(1) XZB(79)	XZA(1) XZA(79)	-XFB(1) -XFB(79)	-XFA(1) -XFA(79)	
(3) D für BN(R,90)	79	R = 1 R = 79						
(4) D für A(R,70)	79	R = 1 R = 79						
(5) D für A(R,90)	79	R = 1 R = 79						
(6) D für B(90) im Gesamtraum	1	alle R						
(7) D für A(90) im Gesamtraum	1	alle R						
(8) D für VU(R) bzw. A(90)	1	alle R						
(9) D für VA(R)	1	alle R						
(10) D \sum_R Binnenwanderungssaldo = 0	1	alle R		1		-1		
(11) D Nachfrage nach Arbeitsplätzen im Gesamtraum	1	alle R	BN(R,90)	EXZB(R)	EXZA(R)	-EXFB(R)	-EXFA(R)	1
(12) D für Pendlersaldo über die Bundesgrenze	1	alle R						1
(13) E Schätzfunktion für ZB(R)	79	R = 1 R = 79		1 1		-0,948 -0,948		
(14) E Schätzfunktion für FB(R)	79	R = 1 R = 79		-0,403 -0,403		1 1		
(15) N Schätzfunktion für den Anteil der Ausländer an der Bevölkerung	79	R = 1 R = 79			1 1		-1 -1	
(16) E Schätzfunktion für FA(R)	79	R = 1 R = 79			-0,660 -0,660		1 1	
(17) E Strukturfaktoren VU(R)	79	R = 1 R = 79						
(18) E Standortfaktoren VA(R)	79	R = 1 R = 79						
(19) E Schätzfunktion für ZA(R)	79	R = 1 R = 79			1 1			
(20) E Erwerbsquoten EN(R)	79	R = 1 R = 79	1 1					
(21) E Pendlersaldo PS(R)	79	R = 1 R = 79						1 1
(22) E Gesamtwanderungssaldo	79	R = 1 R = 79		1 1	1 1	1 1	-1 -1	
(23) E Pendervolumen $\sum PS(R)$ für alle R mit PS(R) > 0	1	nur einige R						1
(24) E Binnenwanderungsvolumen	1	alle R		1				
(25) E Nord-Süd-Wanderung (für alle R im Süden der BRD)	1	nur einige R		1		-1		
(26) N Summe d. Außenwanderungszuzüge	1	alle R			1			
(27) N Summe d. Außenwanderungsfortzüge	1	alle R					1	
(28) N Regionaler Arbeitskräfteüberschuß	79	R = 1 R = 79	BN(1,90) BN(79,90)	EXZB(1) EXZB(79)	EXZA(1) EXZA(79)	-EXFB(1) -EXFB(79)	-EXFA(1) -EXFA(79)	1 1
(29) N Regionale Arbeitslosenquote kleiner als 3 vH	79	R = 1 R = 79	BN(1,90) BN(79,90)	EXZB(1) EXZB(79)	EXZA(1) EXZA(79)	-EXFB(1) -EXFB(79)	-EXFA(1) -EXFA(79)	1 1

Bedingung 11 stellt sicher, daß die Summe der in den Regionen nachgefragten Arbeitsplätze gleich der für die Bundesrepublik prognostizierten Nachfrage ist.

Bedingung 12 drückt aus, daß sich die regionalen Pendlersalden zu dem vorgegebenen Pendlersaldo über die Grenzen des Bundesgebiets addieren müssen.

Bedingung 13 beschreibt die kumulierten Zuzüge in eine Region aus anderen Regionen als Funktion der Veränderung der Zahl der Arbeitsplätze und der Fortzüge in andere Regionen. Diese Funktion wurde für die Periode 1961 bis 1970 auf der Basis einer multiplen linearen Querschnittsregression für die 79 Regionen geschätzt ($r^2 = 0,98$, t-Werte > 3).

Bedingung 14 beruht auf einer entsprechenden Schätzung für die Fortzüge ($r^2 = 0,97$, t-Werte > 3).

Bedingung 15 beruht auf der (gesetzlich verankerten) Forderung, daß der Anteil der Ausländer am Bevölkerungsbestand der Region 12 vH nicht überschreiten darf.

Bedingung 16 basiert auf der ökonometrischen Schätzfunktion für die Fortzüge ins Ausland ($r^2 = 0,99$, t-Werte > 3 , zum Schätzverfahren vgl. Bedingung 13).

Bedingung 17 bis 21 enthalten die außerhalb des Modells prognostizierten Werte für die regionalen Struktur- und Standortfaktoren, die Zuzüge aus dem Ausland, die Erwerbsquoten und die Pendlersalden. Für diese (geschätzten) Größen werden jeweils Ober- und Untergrenzen vorgegeben.

Bedingung 22 legt fest, in welchen Regionen die Summe aus Binnen- und Außenwanderungssalden positiv bzw. negativ sein muß. Die entsprechenden Vorgaben beruhen auf einer Analyse der Wanderungen für den Zeitraum 1961-70.

Bedingung 23 fordert, daß das Pendlervolumen (definiert als Summe aller regionalen Pendlersalden mit positivem Vorzeichen) einen bestimmten, durch Trendberechnungen ermittelten Wert nicht über- bzw. unterschreitet.¹⁾

Bedingung 24 und 25 beruhen ebenfalls auf durch Trendberechnungen vorgegebenen Werten.

Bedingung 26 und 27 entsprechen den politischen Vorgaben für die Zahl der Ausländer.

Bedingung 28 fordert, daß die Nachfrage nach Arbeitsplätzen in allen Regionen größer als das Angebot an Arbeitsplätzen ist, und zwar nachdem die interregionalen Ausgleichsprozesse durch Wanderungen erfolgt sind. Es wird also angenommen, daß ein Überangebot an Arbeitsplätzen durch die Arbeitskräftemobilität stets verhindert wird.

Bedingung 29 setzt als Norm für jede Region eine maximale Arbeitslosenquote von 3 vH. Diese Bedingung ist erforderlich, weil eine auf nationaler Ebene niedrige Arbeitslosenquote nicht ausschließt, daß auf den regionalen Arbeitsmärkten große Ungleichgewichte bestehen. (Obwohl der Begriff Arbeitslosenquote als vH-Satz in bezug auf die Zahl der Erwerbs-

1) Das Pendlervolumen ließe sich besser auf der Basis der Zahl der Ein- bzw. Auspendler definieren. Es waren jedoch nur Daten über die Pendlersalden verfügbar.

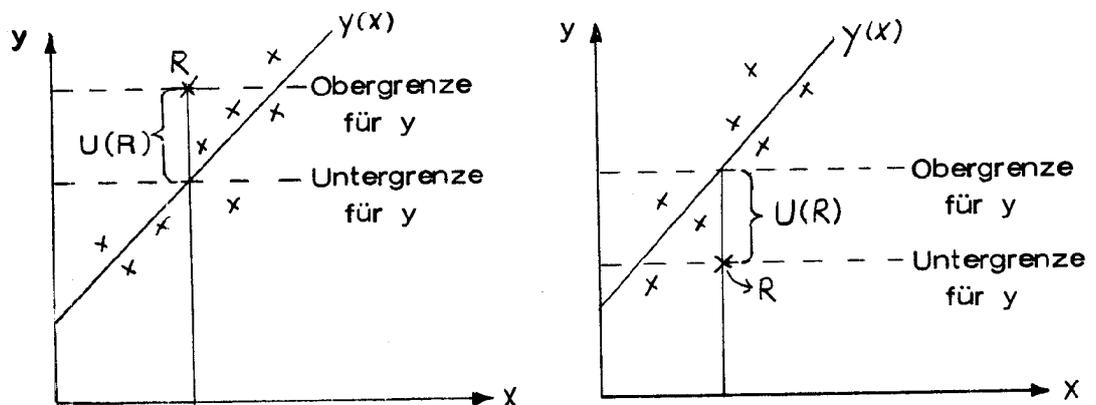
personen definiert ist, ist es möglich, die Zahl der Arbeitslosen auch auf die Zahl der Arbeitsplätze zu beziehen. Von dieser zweiten Art der Definition mußte hier aus programmtechnischen Gründen ausgegangen werden. Einer Arbeitslosenquote von 3 vH in der ersten Definition entspricht ein Prozentsatz von 3,09 vH in der zweiten Definition. Daher steht in Übersicht 3, Zeile 29, Spalte 12, nicht der Koeffizient 1,03, sondern der Koeffizient 1,0309.)

Bei der Festlegung von Ober- und Untergrenzen bzw. von "Ranges" für die in den Zeilen (13), (14) und (16) beschriebenen Nebenbedingungen wurde auf die Störglieder $U(R)$ der entsprechenden Regressionsfunktionen zurückgegriffen. Bei Regionen, für die in der Testperiode der Funktionswert unter dem effektiven Wert lag ($U(R) > 0$), wurde auch für die Prognoseperiode der Funktionswert als Untergrenze und der Funktionswert zuzüglich dem für die Testperiode errechneten Störglied $U(R)$ als Obergrenze festgelegt. Lag der Funktionswert über dem effektiven Wert, wurde der Funktionswert als Obergrenze und der Funktionswert abzüglich $U(R)$ als Untergrenze gewählt (vgl. Schaubild 2).

Schaubild 2

a) Ermittlung von Ranges für Regionen, deren Störglied in der Testperiode positiv war.

b) Ermittlung von Ranges für Regionen, deren Störglied in der Testperiode negativ war.



Dieses Verfahren impliziert eine inhaltliche Interpretation der Störglieder: Es wird angenommen, daß die Abweichungen von der Schätzfunktion auf kausalen Ursachen beruhen, die auch in der Prognoseperiode wirksam sind. Die schätztheoretische Annahme normalverteilter Reste steht dazu nicht im Widerspruch.

4. Stand der Arbeiten

Die für das Modell benötigten Daten wurden bereits aufbereitet, so daß mit den ersten Programm-Läufen begonnen werden konnte. Es wird das APEX-Programm von Control Data zur Linearen Optimierung verwendet (Rechenanlage der FU Berlin).

Erfahrungsgemäß ist es bei der hier vorliegenden relativ großen Zahl von Variablen (948) und Beschränkungen (1276) nicht sehr einfach, eine zulässige Lösung zu finden. Es zeigte sich bereits, daß die Annahme konstanter Parameter in den ökonomischen Gleichungen (insbesondere Zeile 16) die Auffindung einer Lösung sehr erschwert bzw. verhindert. Nach entsprechenden Parameteränderungen konnte die Zahl der nicht erfüllbaren Beschränkungen auf etwa 30 reduziert werden. Eine punktuelle Lockerung dieser verbleibenden Nebenbedingungen ist jederzeit möglich, doch soll versucht werden, eine Lösung ohne Lockerung der Beschränkungen zu finden, indem die besonders relevanten Parameter in der Koeffizientenmatrix gesucht und variiert werden.

Schaubild 3

